Janssen Research & Development

Statistical Analysis Plan

A Phase 3, Randomized, Controlled, Open-label Study of VELCADE (Bortezomib) Melphalan-Prednisone (VMP) Compared to Daratumumab in Combination with VMP (D-VMP), in Subjects with Previously Untreated Multiple Myeloma who are Ineligible for High-dose Therapy

Protocol 54767414MMY3007; Phase 3

JNJ-54767414 (daratumumab)

Authors: [redacted]

Status: Approved
Date: 19 July 2017
Prepared by: Janssen Research & Development, LLC;
Document No.: EDMS-ERI-142546005, 3.0

Compliance: The study described in this report was performed according to the principles of Good Clinical Practice (GCP).

TABLE OF CONTENTS

ABBREVIATIONS

AE	adverse event
ALT	alanine aminotransferase
ASCT	Autologous Stem Cell Transplant
AST	aspartate aminotransferase
BSA	body surface area
CI	confidence interval
CR	complete response
CrCL	creatinine clearance
CRF	case report form
CSR	clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
DOR	duration of response
DPS	data presentation specification
D-VMP	Daratumumab, bortezomib, melphalan, prednisone
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
FISH	fluorescence in situ hybridization
FLC	free light chain
IA	interim analysis
ICF	informed consent form
IDMC	Independent Data Monitoring Committee
IMWG	International Myeloma Working Group
IRR	infusion-related reaction
ISS	International Staging System
ITT	intent-to-treat
IWRS	interactive web response system
MedDRA	Medical Dictionary for Regulatory Activities
M-protein	monoclonal protein, monoclonal paraprotein
MRD	minimal residual disease
NCI-CTCAE	National Cancer Institute Common Terminology Criteria for Adverse
NGS	Next generation sequencing
ORR	overall response rate
OS	overall survival
PD	progressive disease
PBMC	peripheral blood mononuclear cell
PFS	progression-free survival
PP	per-protocol
PR	partial response
SAE	serious adverse event
SAP	statistical analysis plan
SC	subcutaneous
sCR	stringent complete response
SD	stable disease
SD	standard deviation
SOC	system organ class
SPEP	serum protein electrophoresis
SMQ	Standardized MedDRA Queries
TEAEs	treatment-emergent adverse events
TTR	time to response
TTP	time to disease progression
UPEP	urine protein electrophoresis
VGPR	very good partial response
VMP	Bortezomib, melphalan, prednisone
WBC	white blood cells

1. INTRODUCTION

This statistical analysis plan (SAP) contains definitions of analysis sets, derived variables, and statistical methods for the planned analyses as specified in the protocol JNJ-54767414MMY3007, Amendment 4.

1.1. Overview of Trial Design

This is a Phase 3, randomized, open-label, active-controlled, parallel-group, multicenter study comparing daratumumab, VELCADE (bortezomib), melphalan and prednisone (D-VMP) with VELCADE, melphalan and prednisone (VMP) in subjects with previously untreated multiple myeloma who are ineligible for high dose chemotherapy and autologous stem cell transplant (ASCT). The primary objective is to evaluate if daratumumab in combination with VMP prolongs progression-free survival (PFS) compared with VMP alone. The secondary objectives are to compare the 2 treatment groups with respect to time to disease progression (TTP), overall response rate (ORR), very good partial response (VGPR) or better rate, complete response (CR) or better rate, MRD negativity rate, progression-free survival on next line of therapy (PFS2), and overall survival (OS) and as well as to assess the safety and tolerability of daratumumab when administered in combination with VMP.

Approximately 700 subjects (350/arm) will be randomized in a 1:1 ratio to receive either D-VMP or VMP. Randomization will be stratified by International Staging System (ISS) (I, II, or III) at screening, region (Europe vs. Other), and age (<75 vs ≥ 75). Within each stratum, subjects will be randomized using an equal allocation ratio of 1:1.

Subject participation will include a Screening Phase, a Treatment Phase, and a Follow-up Phase. The Screening Phase will be up to 21 days before Cycle 1, Day 1. All subjects will receive up to 9 cycles of the VMP regimen (1 cycle = 6 weeks) with or without daratumumab.

The Treatment Phase will extend from Day 1 of Cycle 1 to discontinuation of all study treatments of a given treatment arm. Subjects in both treatment arms will receive 1.3 mg/m^2 VELCADE by subcutaneous injection (SC) twice weekly (Weeks 1, 2, 4, and 5) in Cycle 1 followed by once weekly (Weeks 1, 2, 4, and 5) in Cycles 2 to 9. Melphalan orally at 9 mg/m^2 and prednisone orally at 60 mg/m^2 will be self-administered on Day 1-4 of each VELCADE cycle. For subjects randomized to the D-VMP regimen, 20 mg of dexamethasone will substitute for the planned dose of prednisone on Day 1 of each cycle. In this setting, dexamethasone will be utilized as the treatment dose of steroid for that particular day, as well as the required pre-medication prior to daratumumab infusion. Daratumumab 16 mg/kg will be administered to subjects in D-VMP regimen by IV infusion once every week for 6 weeks (Cycle 1; 1 VELCADE cycle); then once every 3 weeks for 16 additional doses (Cycles 2-9). Measures to prevent infusion-related

reactions (IRR) will include pre-infusion medication with dexamethasone, paracetamol, and antihistamine before each daratumumab infusion.

After completion of the VMP cycles, subjects in the VMP arm will enter the Follow-up Phase. Subjects in D-VMP regimen will continue to receive daratumumab every 4 weeks until documented progression, unacceptable toxicity, or the study ends (see below for definition). Subjects who need to discontinue treatment with any one component of study treatment (VELCADE, melphalan, prednisone, or daratumumab) may continue to receive treatment with the other components of study treatment, as assigned. Upon discontinuation of daratumumab, subjects in the D-VMP arm will also enter the Follow-up Phase.

In the Follow-up Phase, subjects who discontinued for reasons other than disease progression will continue to have disease evaluations per the Time and Events Schedule, which will continue until confirmed progressive disease (PD), death, lost to follow up, subsequent antimyeloma therapy, withdrawal of consent, or the study ends, whichever occurs first.

Two interim analyses are planned for this study. The first interim analysis, with a purpose to evaluate safety, will be performed after a total of approximately 100 subjects have been treated for at least 2 cycles or discontinued the study treatment. The second interim analysis will be performed when 216 PFS events (60% of the total events) have been accumulated. The purpose of this interim analysis is to evaluate cumulative interim safety and efficacy data. The significance level at this interim analysis to establish the superiority of D-VMP over VMP with regard to PFS will be determined based on the observed number of PFS events at the interim analysis, using the O'Brien-Fleming boundaries as implemented by the Lan-DeMets alpha spending method. If the experimental arm (D-VMP) is numerically worse than the control arm in terms of PFS (observed hazard ratio >1 favoring the control arm), then the study may be terminated for futility.

The primary PFS analysis will occur when approximately 360 PFS events have been observed if the second interim analysis does not result in an early stop due to efficacy or futility. The date established for the primary PFS analysis will serve as the clinical cut-off date, after which data collection in the study will be reduced. Investigators will be informed when each interim analysis is to occur. All available data prior to that time will be included in each of the respective analyses.

The end of the study will occur when 330 subjects have died, or 5 years after the last subject is randomized, whichever comes first. The sponsor will ensure that subjects benefiting from treatment with daratumumab will be able to continue treatment after the end of the study.

An Independent Data Monitoring Committee (IDMC) will be commissioned for this study to review efficacy and safety results at the planned interim analyses. After the interim reviews, they will make recommendations regarding the continuation of the study. In addition, the IDMC may also review cumulative safety data every 6 months in addition to the planned 2 interim analyses.

Assessment of tumor response and disease progression will be conducted in accordance with the International Myeloma Working Group (IMWG) response criteria. An assessment of MRD will be conducted on bone marrow samples. Safety evaluations will include adverse event monitoring, physical examinations, electrocardiogram (ECG) monitoring, clinical laboratory parameters (hematology and chemistry), vital sign measurements, and Eastern Cooperative Oncology Group (ECOG) performance status. Blood samples will be drawn for assessment of pharmacokinetic parameters and immunogenicity.

1.2. Statistical Hypotheses for Trial Objectives

The primary efficacy endpoint of this study is PFS. The null hypothesis is that there is no difference in PFS between daratumumab in combination with VMP and VMP alone in subjects with newly diagnosed multiple myeloma who are ineligible for high dose chemotherapy and ASCT.

The secondary endpoints such as TTP, ORR, VGPR or better rate, CR or better rate, MRD negativity rate, PFS2, time to response, duration of response and OS will be evaluated as well.

1.3. Sample Size Determination

The sample size calculation is performed based on the assumption that the median PFS for the VMP group in this study is estimated to be approximately 21 months.

Assuming the addition of daratumumab can reduce the risk of the disease progression or death by 27.6%, i.e., assuming the hazard ratio (D-VMP vs. VMP) of 0.724, which translates to a median PFS of 29 months for the D-VMP arm, a total of 360 PFS events is needed to achieve a power of 85% to detect this hazard ratio with a log-rank test (two-sided alpha is 0.05). With a 20-month accrual period and an additional 21-month follow-up, the total sample size needed for the study is approximately 700 (350/treatment group) subjects. The sample size calculation has taken into consideration an annual dropout rate of 5%.

Long-term survival follow-up will continue until 330 deaths have been observed or 5 years after the last subject is randomized, whichever is first. This study will achieve approximately 80% power to detect a 27% reduction in the risk of death (hazard ratio = 0.73) with a log-rank test (two-sided alpha = 0.05) if 330 death events are observed at the study end.

1.4. Interim Analyses

Two interim analyses are planned. The first interim analysis has a purpose to evaluate safety after a total of approximately 100 subjects treated for at least 2 cycles or discontinued the study treatment. The second interim analysis is to evaluate the cumulative interim safety and efficacy of daratumumab in combination with VMP when approximately 216 PFS events have been accumulated. More details about the conduct of these interim analyses can be found in a stand-alone IDMC charter.

2. GENERAL ANALYSIS DEFINITIONS

2.1. Visit Windows

For analyses of data by cycle, if data are collected by date (e.g., AE onset), the corresponding study evaluations will be assigned to actual sequential cycles, which are derived from the study treatment administration data. The start date of a particular cycle is defined as the date of the first scheduled dose of any component of the study treatment, and the end date of a cycle is the start date of the next cycle minus 1. For the last cycle, the end date is defined as the end of treatment visit date or the minimum of last study treatment date plus 30 days and subsequent antimyeloma therapy minus 1 day, if the end of treatment visit date is not available.

In general, if data (e.g., laboratory and vital sign etc.) are collected by cycle, the nominal cycle will be used to summarize data. However, due to possible cycle delays, assessment performed in the same cycle may not be well aligned in time scale for different subjects. To address this, by-week windowing rules may be applied in the overtime data summaries by study week.

2.2. Pooling Algorithm for Analysis Centers

All participating centers in the study will be pooled together for analyses.

2.3. Study Treatment and Study Drug

Study treatment refers to bortezomib, melphalan, prednisone, and/or daratumumab. Study drug refers to daratumumab.

2.4. Study Treatment Dosing Date

Study treatment dosing date is the date on which a subject actually received study treatment (partial or complete) and will be recorded in the study treatment administration dataset.

For subjects who receive D-VMP treatment, the first study treatment date is defined as the earliest date of non-zero dose of the following administration: bortezomib, melphalan, prednisone or daratumumab. The last study treatment date is defined as the latest date of non-zero dose of the following administration: bortezomib, melphalan, prednisone or daratumumab.

For subjects who receive VMP treatment, the first study treatment date is defined as the earliest date of non-zero dose of the following administration: bortezomib, melphalan or prednisone. The last study treatment date is defined as the latest date of non-zero dose of the following administration: bortezomib, melphalan or prednisone.

2.5. Baseline Measurement

Baseline measurement is defined as the closest non-missing measurement taken on or prior to the first study treatment administration (including time if time is available, with exception of parameters associated with disease-related efficacy assessment such as SPEP, UPEP, kappa, lambda, kappa/lambda ratio, serum calcium, and albumin).

2.6. Unique Lab Value

In general, in instances when there are multiple records at a given visit date for lab parameters associated with disease assessment, the following rules will be applied to select the unique lab value for analysis: a.) multiple records from both central and local lab, central lab value always takes precedence over local lab value; b) multiple records from central lab, select the latest value by time, visit number, or sequence number as the unique lab value; c.) multiple records from local lab, select the latest lab value by time, visit number, or sequence number as the unique lab value.

2.7. Imputation of Partial Dates

Unless specified otherwise, no data imputation will be applied for missing safety and efficacy evaluations. For analysis and reporting purpose, partial dates in adverse event (AE onset date; AE end date), concomitant therapies (start date; end date), MM diagnosis date, and start date of subsequent antimyeloma therapy will be imputed.

2.7.1. Missing/Partial Adverse Event Onset Date

If the onset date of an adverse event is missing completely or partially, the following imputation rules will be used.

- When month and year are present and the day is missing,
  - If the onset month and year are the same as the month and year of first study treatment, the day of first study treatment or the day-component of the AE end date (possibly imputed) is imputed, whichever is earlier
  - If the onset month and year are not the same as the month and year of first study treatment, then 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 first study treatment. If AE end date is available and is prior to first study treatment, the day and month of AE end date are imputed. Otherwise, the day and month of first study treatment are imputed

  - If the onset year is different from the year of first study treatment, the 1st of January is imputed
- If the onset date is completely missing, the earlier one of the date of first study treatment and the AE end date is imputed as the onset date.

No imputation will be done for partial or missing AE onset time.

2.7.2. Missing/Partial Adverse Event End Date

If the end date of an adverse event is missing completely or partially, the following imputation rules will be used.

- If month and year are present and the day of the month is missing, the last day of the month is imputed.
- If only a year is present, the 31st of December is used.
- If the imputed date is later than the date of death (if available), the date of death will be used as the imputed date instead.
- If the year of end date is missing, no imputation will be applied.

No imputation will be done for partial or missing AE end time.

2.7.3. Partial Concomitant Medication Start/End Date

In case of partially missing concomitant medication start/end dates, the following imputation rules will be applied. If the date is completely missing, no imputation will be performed.

- If only the day is missing, the 15th day of the month will be used
- If both the day and month are missing, the 30th of June will be used. If the medication was taken prior to study start, and the imputed start date is after first treatment date, further adjust of the imputed start date as the day prior to first dosing date; If the medication was taken after study start, and the imputed start date is prior to first dosing date, further adjust the imputed start date as first dosing date. Also, adjust the imputed medication end date so that it is on or after first dosing date.

2.7.4. Partial Multiple Myeloma Diagnosis Date

For partial date of original multiple myeloma diagnosis, the following imputation rules will apply:

- If only the day is missing, set day=15 and pick minimum of imputed date, date of collection and date of randomization.
- If both the day and month are missing, set to January 1 and pick minimum of imputed date, date of collection and date of randomization.
- If year is missing, no imputation will be applied.

If the imputed date of original diagnosis is after the randomization date, further adjust the imputed start date as the day before the randomization date.

2.7.5. Partial Subsequent Antimyeloma Therapy Start Date

If year or month of subsequent antimyeloma therapy start date is missing or no components of the start date are present, no imputation will be performed.

If only the day is missing, the following steps apply:

- If the month and year of the start date are the same as the month and year of last dosing date, the day of last dosing date or the day-component of the stop date of subsequent antimyeloma therapy is imputed, whichever is earlier.
- If the start month and year are not the same as the month and year of last dosing date, the first day of the month is imputed.

No imputation will be applied for missing or partial subsequent antimyeloma therapy end date.

2.8. General Analysis Method

In general, continuous variables will be summarized using descriptive statistics such as mean, standard deviation (SD), median and range. Categorical variables will be summarized using frequency and percentage. For time-to-event variables, which is defined as from the date of randomization to the date of the event, the Kaplan-Meier method will be used for descriptive summaries. For the calculation of time-to-event and duration-of-event variables, the difference between the start date and the end date plus 1 day will be used.

2.9. Analysis Sets

The following analysis sets are defined.

- Intent-to-treat (ITT): defined as subjects who have been randomly assigned to the D-VMP or VMP group. Analyses of demographics, baseline characteristics, and efficacy endpoints will be primarily analyzed based on this population.
- Safety population: defined as subjects who have received at least 1 administration of any study treatment (partial or complete). This population will be used for all safety analyses. The safety analyses grouping will be according to treatment actually received.
- Response-evaluable: defined as subjects who have a confirmed diagnosis of multiple myeloma and measurable disease at baseline or screening visit. In addition, subjects must have received at least one administration of study treatment and have adequate post-baseline disease assessments. Measurable disease is defined as follows:
  - IgG myeloma: Serum monoclonal paraprotein (M-protein) level ≥1.0 g/dL or urine M-protein level ≥200 mg/24 hours; or

  - IgA, IgD, IgE or IgM multiple myeloma: serum M-protein level ≥0.5 g/dL or urine M-protein level ≥200 mg/24 hours; or
  - Light chain multiple myeloma without measurable disease in serum or urine: Serum immunoglobulin free light chain (FLC) level ≥10 mg/dL and abnormal serum immunoglobulin kappa lambda free light chain ratio.

  This population will be used as sensitivity analyses for selective response-related secondary endpoints such as CR or better rate.
- Per-protocol (PP) population: defined as subjects who are randomized and meet all eligibility criteria.
- Pharmacokinetics-evaluable: defined as subjects assigned to D-VMP group who received at least 1 administration of daratumumab and have at least 1 pharmacokinetic sample concentration value after the first infusion. All pharmacokinetics analyses are based on the pharmacokinetic evaluable population.
- Immune response-evaluable: defined as subjects assigned to D-VMP group who have at least 1 immunogenicity sample obtained after their first daratumumab administration.
- Molecular genetic evaluable: defined as subjects who meet one of the following biomarker criteria for risk assessment based on the next generation sequencing (NGS) data:
  - Standard risk: subjects that are negative (=molecular aberration absent) for all del17p, t(14;16), t(4;14).
  - High risk: subjects that are positive (=molecular aberration present) for any of del17p, t(14;16), t(4;14).
- Cytogenetic evaluable: defined as subjects who meet one of the following cytogenetic risk categories:
  - Standard risk: subjects that are negative for del17p, t(14;16), t(4;14) by FISH/Karyotype.
  - High risk: subjects that are positive for any of del17p, t(14;16), t(4;14) by FISH/Karyotype.

2.10. Definition of Subgroups

The following pre-specified subgroup analyses are to be performed at the 2nd IA for the efficacy and safety endpoints. Additional subgroup analyses may be performed, if requested and deemed necessary by IDMC for their analyses and decision making.

Table 1: Subgroup Analyses of Efficacy and Safety Endpoints

Subgroup	Definition	Analysis Type
Sex	Male, Female	E, S
Age	E: <75 years, ≥75 years S: <65 years, 65 to <75, ≥75	E, S
Race	White, Other	E, S
Baseline renal function (CrCl):	E: >60 mL/min; ≤60 mL/min; S: <30, 30 to <60, 60 to <90, ≥90 mL/min	E, S
Baseline hepatic function	Normal, Impaired[a]	E, S
Region	Europe, Other	E, S
International Staging System (ISS)	I, II, and III	E
Type of MM	IgG, Non-IgG	E
Cytogenetic risk	High risk, Standard risk	E
ECOG performance score	0, ≥1	E

E: efficacy (PFS, CR or better rate); S: TEAE

[a] Includes mild, moderate and severe. mild: (total bilirubin ≤ ULN and AST > ULN) or ULN < total bilirubin ≤ 1.5×ULN; moderate: 1.5×ULN < total bilirubin ≤ 3×ULN; severe: total bilirubin > 3×ULN

3. SUBJECT INFORMATION

3.1. Demographics and Baseline Characteristics

Unless specified otherwise, all demographic and baseline characteristics variables will be summarized for the ITT population. No statistical comparison between the 2 treatment groups is planned.

The distribution of subject enrollment will be presented for each treatment group according to region and country. A list of subjects who did not meet study inclusion/exclusion criteria will be provided. This listing will include subject ID, treatment group, category of study selection criteria not met and specific criteria not met.

Subject demographic and baseline characteristic variables: age (<65 years, 65 to <75 years, and ≥75 years), sex, ethnicity, race, weight (kg), height (cm), body surface area (BSA) (m^2), and ECOG performance status will be summarized by treatment group and overall. A listing of subject demographic and baseline characteristics will be provided as well.

Baseline disease characteristics including type of multiple myeloma (IgG, IgA, IgM, IgD, IgE, light chain only, biclonal, or negative immunofixation), type of measurable disease (IgG, IgA, Other (IgD, IgM, IgE and biclonal), Serum and urine, Urine only, or Serum FLC), ISS staging at screening by central laboratory assessment (I, II, III), time since initial MM diagnosis (months), number of lytic bone lesions (None, 1-3, 4-10, more than 10), presence of diffuse myeloma-related osteopenia (Yes, No), number of extramedullary plasmacytomas (Yes, No), bone marrow % plasma cells (<10, 10 –

30, >30), standard-risk and high-risk cytogenetic abnormalities (del17p, t(4;14), t(14;16)), will be summarized and tabulated by treatment group and overall.

A descriptive summary of selected hematology and chemistry laboratory analytes at baseline will be provided for each treatment group and overall. In addition, baseline toxicity grade of each selected laboratory analyte in hematology and chemistry panel will be summarized by treatment group using frequency.

Medical history collected at baseline or screening visit will be summarized by system-organ class and preferred term for each treatment group and overall.

A summary of stratification factors (ISS staging, region, and age) used in the randomization based on IWRS will be provided to evaluate whether or not randomization process was appropriately executed in the study. The stratified log-rank test and stratified cox model use these stratification factors in randomization.

3.2. Disposition Information

An overview of subject disposition in the study will be provided. The overview includes a summary of total number of subjects who are randomized to each treatment group, the number and percentage of subjects who are randomized but not treated in each treatment group, and total number of subjects who are treated in each treatment group. For all treated subjects (defined as subjects who have received at least 1 administration of any study treatment), the number and percentage of subjects who discontinued treatment including reason for discontinuation as indicated by the investigators will be summarized. Similar summaries will be presented for all randomized subjects who discontinued from study.

A list of subjects who discontinued study treatment will be provided for safety population. This listing will include subject ID, treatment group, date of treatment discontinuation, study day of last dose, reason for discontinuation as well as the specific adverse events (MedDRA preferred term/verbatim term) if discontinuation due to AEs and primary cause of death if discontinuation due to death.

A similar list of subjects who discontinued study will be provided for ITT population.

3.3. Extent of Exposure

Extent of exposure to study treatments will be summarized and presented based on the safety population.

The number and percentage of subjects treated within each cycle will be summarized by treatment group. The maximum number of treatment cycles received for each subject will be summarized by frequency and descriptive statistics.

Duration of study treatment, defined as the number of days from the date of the first administration of study treatment to the date of the last administration of study treatment, will be summarized.

The number of daratumumab administrations (continuous and categorical variables) will be summarized for subjects treated with D-VMP. The total dose administered for daratumumab (mg/kg), bortezomib (mg/m^2), melphalan (mg/m^2) and prednisone (mg/m^2) will be summarized overall, by cycle, by high-intensity (Cycle 1) or low-intensity (Cycles 2-9) for bortezomib, and by Cycle 1, Cycles 2-9 and Cycles 10+for daratumumab.

The dose intensity, which is defined as the sum of total dose administered in all cycles divided by the number of treatment cycles, will be calculated for each study treatment and summarized accordingly. Additionally, the daratumumab dose intensity will be summarized for cycles with high-intensity (Cycle 1) or low-intensity (Cycles 2-9 and Cycles 10+).

The relative dose intensity (%) defined as the ratio of total dose actually received and total planned dose (planned dose level times the number of administered infusions/medications) will be calculated for each study treatment and summarized by treatment group using descriptive statistics.

The number of subjects with treatment cycle delay, dose delays/skipping for each study treatment, dose reduction for bortezomib, melphalan, and prednisone will be summarized for each treatment group. The reasons (AE or other) for treatment cycle delay, dose delays or skipping for each study treatment as well as dose reduction for bortezomib, melphalan, and prednisone will also be reported. In addition, a summary of study treatment dose modifications by cycle will be provided.

3.4. Protocol Deviations

Major protocol deviations will be summarized for the ITT population by the following types of deviation for each treatment group:

- Entered but did not satisfy inclusion/exclusion criteria
- Developed withdrawal criteria but not withdrawn
- Received wrong treatment or incorrect dose
- Received an excluded concomitant treatment
- Efficacy assessment deviation
- Safety assessment deviation
- Other – protocol non-compliance

All major protocol deviations will be summarized. A listing of subjects with major protocol deviations including subject ID, type of deviation, and reasons for deviation will also be provided.

3.5. Prior, Concomitant and Subsequent Therapies

With the study population of newly diagnosed subjects with multiple myeloma who are not candidates for high dose chemotherapy and ASCT, prior systemic use of corticosteroids is limited to a short course of emergency use to treat multiple myeloma symptoms. If any, a listing of all prior systemic use of corticosteroids will be provided.

Concomitant medications collected in the CRF page during the study will be summarized by therapeutic class, pharmacologic class, and drug name for each treatment group. A similar summary will be provided for subjects who received growth factor support, pre-infusion medication and post-infusion medication, respectively. In addition, systemic steroids as concomitant medication use during the study will be summarized. Additionally, prophylactic antiviral medication use will be tabulated.

The total number of subjects who received subsequent antimyeloma therapy will be reported for safety population in each treatment group. A summary of subsequent antimyeloma therapy will be presented by therapeutic class, pharmacologic class and drug name. In addition, for subjects who received subsequent antimyeloma therapy, their best response to the first subsequent antimyeloma therapy will be summarized.

4. EFFICACY

A validated computerized algorithm, which is based on the IMWG response criteria (Durie 2006, Rajkumar 2011)[1,3] and has been used and validated by an independent review committee (IRC) in Study MMY2002, also used in MMY3003 and MMY3004, will be used to determine response and disease progression for each subject. Unless specified otherwise, relapse from CR by positive immunofixation or trace amount (defined as less than 0.5 g/dL) of M-protein is not considered to be progressive disease in the IMWG response criteria. As a sensitivity analysis, investigator assessment of response and disease progression using the IMWG response criteria will also be performed.

4.1. Analysis Specifications

4.1.1. Level of Significance

All statistical hypothesis tests and 95% confidence intervals presented will be 2-sided.

The primary hypothesis is to be tested at the 0.05 significance level (overall). The exact significance level at the second interim analysis is to be determined by the observed number of events per the O'Brien-Fleming alpha spending function. Assuming 216 PFS events are observed at the second interim analysis, the alpha to be spent will

be 0.0076 (2-sided) for the interim analysis and 0.0476 (2-sided) for the primary PFS analysis (360 PFS events occur). If the observed two-sided p-value is smaller than this significance level as specified above, the superiority of D-VMP versus VMP with respect to PFS will be established.

If the primary endpoint of PFS is statistically significant, the following secondary endpoints ordered below will be sequentially tested, each with an overall two-sided alpha of 0.05, by utilizing a hierarchical testing approach as proposed by Tang and Geller 1999[4] that strongly controls Type I error rate:

1) ORR
2) VGPR or better rate
3) CR or better rate
4) MRD negativity rate
5) OS.

These secondary endpoints will be tested at the second IA and the primary PFS analysis. The primary PFS analysis will be skipped if PFS is positive at the second IA. If this is the case, the protocol will be amended to add an analysis for secondary endpoints at a similar timing as the planned primary PFS analysis (note that OS will also be tested at its protocol-specified final analysis). The significance level at the second interim and the primary PFS analyses will be determined by the alpha-spending function specific to endpoints:

- For ORR, VGPR or better rate, CR or better rate, and MRD negativity rate, the information fraction is expected to be 80% at the second IA. The O'Brien-Fleming alpha-spending function as implemented by the Lan-DeMets method will be used for alpha spending: 0.0244 (two-sided) at the second IA and 0.0428 (two-sided) at the primary PFS analysis.

- For OS, a modified linear alpha spending function will be used to determine the alpha level at three looks (second IA, primary PFS analysis, and final OS analysis). The alpha level will be 0.0001 for the first OS look, and linear spending function will be used to determine the alpha level for the second and last looks. For example, if 59% targeted OS events are observed at the second look, the corresponding alpha level will be 0.0295.

If the null hypothesis for any of these endpoint fails to be rejected at the second interim analysis, then any subsequent endpoint(s) listed above will not be tested until the next analysis time point (e.g., primary PFS analysis), if applicable. If the null hypothesis

for an endpoint is rejected at the second interim analysis, it will remain being rejected and will not be re-tested at any subsequent time points, if any.

4.1.2. Data Handling Rules

There is no imputation planned for missing efficacy endpoint values.

4.2. Primary Efficacy Endpoint

The primary efficacy endpoint is progression-free survival (PFS) based on the computerized algorithm.

4.2.1. Definition

PFS is defined as the duration from the date of randomization to either progressive disease, according to the IMWG response criteria, or death, whichever occurs first. Subjects who start subsequent antimyeloma therapies for multiple myeloma without disease progression will be censored at the last disease assessment before the start of subsequent therapies. Subjects who withdrew consent from the study before disease progression will be censored at the last disease assessment. Subjects who are lost to follow-up will be censored at the last disease assessment before subjects are lost to follow-up. Subjects who have not progressed and are still alive at the cutoff date for analysis will be censored at the last disease assessment. Subjects without any post-baseline disease assessment will be censored at the randomization.

Determination of dates of PFS event and dates for censoring is summarized in Table 2 as follows.

Table 2: PFS Event and Censoring Method

Situation	Date of Progression or Censoring	Outcome
Disease progression prior to start of subsequent antimyeloma therapy	Earliest date that indicates disease progression	PFS event
Death prior to start of subsequent antimyeloma therapy	Date of death	PFS event
No post-baseline disease assessment	Randomization	Censored
Other (e.g., withdrawal of consent to study participation, lost to follow-up, start of subsequent antimyeloma therapy etc.)	Date of last disease assessment prior to withdrawal of consent to study participation, lost to follow-up, or subsequent antimyeloma treatment	Censored

4.2.2. Analysis Methods

Analysis of PFS will be based on the ITT population. The Kaplan-Meier method will be used to estimate the distribution of overall PFS for each treatment group. The median PFS with 95% CI will be provided. In addition, the number and percentage of subjects who had a PFS event or were censored will be reported. The Kaplan-Meier PFS curve will also be plotted by treatment group.

The primary treatment comparison of the distribution of overall PFS will be based on a stratified log-rank test. The p-value from a stratified log-rank test will be reported. Hazard ratio and its 95% confidence interval will be estimated based on a stratified Cox's regression model with treatment as the sole explanatory variable. Stratification factors used in the analyses include ISS staging (I, II, III), region (Europe vs other), and age (<75 years vs ≥75 years).

In addition, 12, 18 and 24-months of PFS rate with 95% CI will be estimated by Kaplan-Meier method and reported for each treatment group.

Additionally, reasons for PFS censoring will be summarized for ITT population.

4.2.3. Sensitivity Analysis of PFS

The following sensitivity analysis is planned to evaluate the robustness of the primary endpoint analysis of PFS.

4.2.3.1. Progressive Disease Based on Investigator Assessment

A sensitivity analysis of PFS, in which progressive disease is based on investigator assessment according to the IMWG response criteria, will be performed in a similar manner as described in the Section 4.2.2.

The PFS definition used in the sensitivity analysis is similar to that defined in the Section 4.2.1, except for date of progressive disease and date of censoring. The date of progressive disease is the date of initial disease progression recorded in the Disease Progression CRF page or earliest date of confirmed progressive disease recorded in the Evaluation of Response CRF page, based on investigator assessment. Similarly, the censoring date is the latest date of disease response recorded in the Evaluation of Response CRF page, based on investigator assessment.

In addition, reasons for PFS and censoring based on investigator assessment will be summarized for ITT population.

4.2.3.2. Not Censored for Start of Subsequent Antimyeloma Therapies

A sensitivity analysis of PFS derived from the algorithm by not censoring data due to start of subsequent antimyeloma therapies for subjects, who have not developed a

confirmed progressive disease, will be performed in a similar manner as described in Section 4.2.2.

The PFS definition used in the sensitivity analysis is similar to that defined in Section 4.2.1, except for censoring date due to start of subsequent antimyeloma therapies. Subjects who start subsequent antimyeloma therapies for multiple myeloma without disease progression will NOT be censored at the last disease assessment before the start of subsequent therapies. If there is no confirmed progressive disease, the subjects will be censored at the last disease assessment before subjects are lost to follow-up or withdrawal of consent to study.

Determination of dates of PFS event and dates for censoring is summarized in Table 3 as follows.

Table 3: PFS Event and Censoring Method

Situation	Date of Progression or Censoring	Outcome
Disease progression	Earliest date that indicates disease	PFS event
Death	Date of death	PFS event
No post-baseline disease assessment	Randomization	Censored
Other (e.g., withdrawal of consent to study participation, lost to follow-up, start of subsequent antimyeloma therapy etc.)	Date of last disease assessment prior to withdrawal of consent to study participation, lost to follow-up	Censored

4.2.3.3. Censored for Death/PD after Missing More Than One Disease Evaluation

A sensitivity analysis of PFS derived from the algorithm by censoring for death or progression after missing two or more consecutive evaluations will be performed in a similar manner as described in Section 4.2.2.

The PFS definition used in the sensitivity analysis is similar to that defined in Section 4.2.1, except for death or progression after missing more than one disease evaluation. For any PFS (death or progression) event identified by the computer algorithm, if the event date and the latest date of scheduled disease evaluation (includes serum M-protein, urine M-protein, serum FLC and corrected calcium only) immediately preceding the event differs more than 1.25 cycles for year 1 (Cycle 1 to Cycle 9), 2.5 cycles for the second year, and 5 cycles thereafter, which indicates that subject missed at least one scheduled disease evaluation, then this event will not be considered as a PFS event in the sensitivity analysis. Instead, the subject will be censored at the date of last disease evaluation (includes serum M-protein, urine

M-protein, serum FLC and corrected calcium only) prior to the PFS event originally identified.

4.2.3.4. Per-protocol Analysis of PFS

A sensitivity analysis of PFS derived from the algorithm based on Per-protocol population will be performed in a similar manner as described in Section 4.2.2 .

4.2.3.5. Unstratified Analysis of PFS

A sensitivity analysis of PFS by using unstratified log-rank test and unstratified Cox's regression model will be performed in a similar manner as described in Section 4.2.2.

4.3. Secondary Endpoints

Secondary efficacy endpoints include overall response rate (ORR), VGPR or better rate, CR or better rate, MRD negativity rate, OS, PFS2, sCR rate, TTP, time to and duration of response, time to subsequent antimyeloma treatment, and proportion of subjects with best M-protein/dFLC responses (maximal reduction of 100%, ≥90% or ≥50% from baseline).

4.3.1. Overall Response Rate (ORR)

4.3.1.1. Definition

ORR is defined as the proportion of subjects who achieve a partial response or better (i.e., PR, VGPR, CR, or sCR) based on the computerized algorithm, according to IMWG response criteria, during or after the study treatment but before the start of subsequent antimyeloma therapy.

4.3.1.2. Analysis Methods

ORR will be calculated for each treatment group based on the ITT and response-evaluable population. The number and percentage of subjects in the following response categories will be presented by treatment group: stringent complete response (sCR), complete response (CR), sCR+CR, very good partial response (VGPR), VGPR or better (sCR+CR+VGPR), partial response (PR), overall response (sCR+CR+VGPR+PR), stable disease (SD), progressive disease (PD), and not evaluable (NE). The corresponding 95% exact CI will be provided.

Stratified CMH test will be used to test treatment difference in ORR, VGPR or better rate, CR or better rate. The CMH estimate of odds ratio and its 95% confidence interval and p-value for testing treatment difference will be reported. Stratification factors used in the analysis include ISS staging (I, II, III), region (Europe vs. Other), and age (<75 years vs. ≥ 75 years).

A sensitivity analysis, which disease response is based on investigator assessment according to the IMWG response criteria, will be performed in a similar manner as described above.

4.3.2. VGPR or Better Rate

4.3.2.1. Definition

VGPR or better rate is defined as the proportion of subjects with a response of VGPR or better (i.e., VGPR, CR or sCR) based on the computerized algorithm, according to IMWG response criteria, during or after the study treatment but before the start of subsequent antimyeloma therapy.

4.3.2.2. Analysis Methods

Similar statistical methods will be applied as described in Section 4.3.1.2 for ORR analysis.

4.3.3. CR or Better Rate

4.3.3.1. Definition

CR or better rate is defined as the proportion of subjects with a response of CR or better (i.e., CR or sCR) based on the computerized algorithm, according to IMWG response criteria, during or after the study treatment but before the start of subsequent antimyeloma therapy.

4.3.3.2. Analysis Methods

Similar statistical methods will be applied as described in Section 4.3.1.2 for ORR analysis.

4.3.4. MRD Negativity Rate

For this study, three threshold values, 10^{-4}, 10^{-5} and 10^{-6}, will be used to evaluate MRD negativity status and its predictive value for PFS.

4.3.4.1. Definition

MRD negativity rate is defined as the proportion of subjects who have negative MRD by bone marrow aspirate at any time point after the randomization and before disease progression or starting subsequent therapy. MRD positive subjects include subjects of which all tested samples were found to be MRD positive or ambiguous. Subjects with missing or unevaluable MRD status will be considered as MRD positive.

4.3.4.2. Analysis Methods

The MRD negativity rate will be calculated for each treatment group based on the ITT population. The corresponding 95% exact CI will be provided. Reasons for missing or unevaluable MRD status will be tabulated by treatment group.

For each threshold value, the stratified CMH estimate of odds ratio and its 95% confidence interval and p-value from Fisher's exact test will be used to test if the MRD negativity rate is the same between the two treatment groups in the previously untreated

myeloma setting. Stratification factors used in the analysis include ISS staging (I, II, III), region (Europe vs. Other), and age (<75 years vs. ≥75 years). For the purpose of hierarchical testing, the threshold value of 10^{-5} will be employed.

In addition, time point analysis of MRD negativity rate will be performed at 12, 18, 24 and 30 months after the first dose of study treatment. Time to MRD negativity will be descriptively summarized by treatment groups, as well as the proportion of subjects with durable MRD negativity (i.e., lasted at least 12 months after the start of any MRD negativity) if sufficient data becomes available.

Exploratory analyses correlating MRD negativity with long-term clinical endpoints (e.g. PFS) are described in Section 7.1.2.

4.3.5. Overall Survival

4.3.5.1. Definition

Overall survival (OS) is measured from the date of randomization to the date of death due to any cause. Subjects who withdraw consent from the study or are lost to follow-up will be censored at the time of withdrawal or lost to follow-up. Subjects who are still alive at the cutoff date for the analysis will be censored at the last known alive. The date of last known alive will be determined by the maximum collection/assessment date from among selected data domains within the clinical database.

4.3.5.2. Analysis Methods

OS, including 12, 24, and 36-months survival rate, will be analyzed for the ITT population. The Kaplan-Meier method will be used to estimate the distribution of OS for each treatment group. Median OS with 95% CI will be provided. In addition, the number and percentage of subjects who had died or were censored will be reported. Additionally, the survival rate with 95% CI at 12, 24 and 36 -months will be estimated using Kaplan-Meier method for each treatment group.

Due to the expected small number of death events at the second interim analysis, the distribution of OS for the 2 treatment groups will be compared based on an un-stratified log-rank test. A p-value from an un-stratified log-rank test will be reported. Hazard ratio and its 95% confidence interval will be estimated based on an un-stratified Cox's regression model with treatment as the sole explanatory variable.

If the null hypothesis of OS is not rejected at the second interim analysis, for OS at the primary PFS analysis when 360 PFS events observed, the alpha to be spent will be determined by a linear alpha spending function based on the observed number of deaths at that time, i.e., the cumulative alpha to be spent will be the total alpha (0.05) multiplied by the proportion of the observed number of deaths out of the total planned number of deaths (330, expected at time of the final OS). See Section 4.1.1 for more details on alpha spending and type I error control.

If a large number of subjects crossed over to receive daratumumab treatment, at the final OS analysis, exploratory analysis may be performed that adjust for the effect of crossover, such as an analysis based on the inverse-probability-of-censoring weighting (IPCW) by Robins (2000)[5] or the accelerated failure time model with iterative parameter estimation method described by Branson and Whitehead (2002).[6] If a large number of patients are censored, the rank preserving structural failure time model described by Robins and Tsiatis (1991)[7] may also be considered in the overall assessment of the survival endpoint.

4.3.6. Progression-free Survival on Next Line of Therapy (PFS2)

4.3.6.1. Definition

PFS2 is defined as the duration from randomization to progression on the next line of subsequent antimyeloma therapy or death due to any cause, whichever comes first. Disease progression on next line of treatment will be based on investigator judgment.

For those subjects who progress on study treatment, but are still alive and not yet progressed on the next line of therapy, they will be censored on the date of last follow-up. Subjects who withdraw consent or lost to follow-up prior to any subsequent antimyeloma therapy will be censored at the date of last disease assessment during the course of study. Subjects without any post-baseline follow-up will be censored at the randomization.

Determination of dates of PFS2 event and dates for censoring is summarized in Table 4 as follows:

Table 4: PFS2 Event and Censoring Method

Situation	Date of Progression or Censoring	Outcome
No post-baseline disease assessment	Randomization	Censored
No death on study treatment and withdrew or lost to follow-up when next line of subsequent antimyeloma therapy not yet started	Date of last disease assessment on study treatment	Censored
Death when next line of subsequent antimyeloma therapy not yet started	Date of death	PFS2 event
No disease progression on study treatment, and no disease progression and no death after the start of next line of subsequent antimyeloma therapy	Date of last disease assessment on study treatment	Censored
Disease progression on study treatment, and no disease progression and no death after the start next line of subsequent antimyeloma therapy	The start date of 2nd line of subsequent antimyeloma therapy if any (the one after the next line of subsequent therapy) minus 1 day, or date of last follow-up, whichever occurs first	Censored
Disease progression (or death) after the start of next line of subsequent antimyeloma therapy	Earliest date that indicates progression on the next line of subsequent antimyeloma therapy (or date of death). If both occurred, pick the minimum between the two dates.	PFS2 event

4.3.6.2. Analysis Methods

Similar statistical methods will be applied as described in Section 4.2.2 for PFS analysis.

4.3.7. sCR Rate

4.3.7.1. Definition

sCR rate is defined as the proportion of subjects with a response of sCR based on the computerized algorithm, according to IMWG response criteria, during or after the study treatment but before the start of subsequent antimyeloma therapy.

4.3.7.2. Analysis Methods

sCR rate will be calculated for each treatment group based on the ITT and response-evaluable population. The corresponding 95% exact CI will be provided.

4.3.8. Time to Disease Progression (TTP)

4.3.8.1. Definition

TTP is defined as the time between the date of randomization and the date of first documented evidence of confirmed PD, as defined in the IMWG response criteria, or death due to PD, whichever occurs first. Subjects who start subsequent antimyeloma therapies for multiple myeloma without disease progression will be censored at the last disease assessment before the start of subsequent therapies. Subjects who withdraw consent to study or are lost to follow-up or die without disease progression will be censored at the last disease assessment. Subjects who have not progressed at the cutoff date for analysis will be censored at the last disease assessment. Subjects without any post-baseline disease assessment will be censored at the randomization.

Determination of dates of TTP event and dates for censoring is summarized in Table 5 as follows.

Table 5: TTP Event and Censoring Method

Situation	Date of Progression or Censoring	Outcome
Disease progression prior to start of subsequent antimyeloma therapy	Earliest date that indicates disease progression	TTP event
Death due to disease progression prior to start of subsequent antimyeloma therapy	Date of death	TTP event
No post-baseline disease assessment	Randomization	Censored
Other (e.g., withdrawal of consent to study participation, lost to follow-up, start of subsequent antimyeloma therapy etc.)	Date of last disease assessment prior to subsequent antimyeloma treatment	Censored

4.3.8.2. Analysis Methods

Similar statistical methods will be applied as described in Section 4.2.2 for PFS analysis.

4.3.9. Time to Response

4.3.9.1. Definition

Time to response (i.e., time to first response) is defined as the time between the date of randomization and the first efficacy evaluation that the subject has met all criteria for PR or better based on the computerized algorithm for patients who had PR or better as their best response.

4.3.9.2. Analysis Methods

For subjects who achieve a confirmed response, descriptive statistics (n, mean, SD, median, and range) will also be provided to summarize time to response, time to VGPR or better response and time to CR or better response.

4.3.10. Duration of Response

4.3.10.1. Definition

Duration of response (DOR) is defined for subjects with a confirmed response (PR or better) as the time between first documentation of response and disease progression based on the computerized algorithm, according to IMWG response criteria, or death due to PD, whichever occurs first. Responders without disease progression will be censored at the censoring time point for TTP.

4.3.10.2. Analysis Methods

No formal statistical comparison of DOR between the 2 treatment groups is planned. Analysis of DOR will be based on subjects who achieved a confirmed response of PR or better. Median DOR with 95% CI will be estimated based on the Kaplan-Meier method for each treatment group. The Kaplan-Meier duration of response curve will be plotted by treatment group.

4.3.11. Time to Subsequent Antimyeloma Treatment

4.3.11.1. Definition

Time to subsequent antimyeloma treatment is defined as the time from randomization to the start of subsequent antimyeloma treatment. Death due to PD without start of subsequent therapy will be considered as event. Subjects who withdrew consent to study or are lost to follow, or die due to causes other than disease progression will be censored at the date of death or the last date known to be alive.

4.3.11.2. Analysis Methods

The Kaplan-Meier method will be used to estimate the distribution of time to subsequent antimyeloma treatment for the ITT population. Median time to subsequent antimyeloma treatment with 95% CI will be tabulated for each treatment group. In addition, a Kaplan-Meier curve for time to subsequent antimyeloma treatment will be plotted. The hazards ratio and its 95% CI will be obtained through a stratified Cox's regression model with treatment as the sole explanatory variable. Treatment comparison will be made via a stratified log-rank test.

4.3.12. Best M-protein/dFLC Response

4.3.12.1. Definition

Best M-protein response is defined as the maximal percent reduction or the lowest percent increase from baseline in serum M-protein for subjects with measurable heavy chain at baseline or urine M-protein for subjects without measurable heavy chain, but with measurable light chain disease at baseline.

For subjects without measurable heavy chain and light chain disease at baseline, best response in serum FLC is defined as the maximal percent reduction or the lowest percent increase from baseline in the difference between involved and uninvolved serum FLC level.

4.3.12.2. Analysis Methods

The number and percentage of subjects in each response category, along with presentation of ≥90% reduction and ≥50% reduction will be tabulated for each treatment group.

4.4. Subgroup Analysis of Efficacy Endpoints

For assessment of internal consistency and investigation of homogeneity of the treatment effect across subgroups, subgroup analyses of the primary and selected secondary efficacy endpoints will be conducted based on pre-specified subgroups defined in Section 2.10.

A forest plot of subgroup analysis on PFS and selected secondary endpoints (e.g. CR or better rate) will be generated, respectively.

4.5. Functional Status and Well-being

4.5.1. Definition

Functional status and well-being will be assessed using two PRO measures, the EORTC- QLQ-C30 and the EQ-5D-5L. The EORTC QLQ-C30 and EQ-5D-5L will be scored based on the instrument developer guidelines.

The EORTC QLQ-C30 includes 30 items resulting in 5 functional scales (physical functioning, role functioning, emotional functioning, cognitive functioning, and social functioning), 1 Global Health Status scale, 3 symptom scales (fatigue, nausea and vomiting, and pain), and 6 single items (dyspnea, insomnia, appetite loss, constipation, diarrhea, and financial difficulties). The recall period is 1 week (the past week).

The instrument contains 28 items using a Likert scale with 4 response options: “Not at All,” “A Little,” “Quite a Bit,” and “Very Much” (scored 1 to 4). Two additional items use response options (1 to 7): 1 = Very Poor, to 7 = Excellent. All scale and item scores will be linearly transformed to be in the range from 0 to 100 according to the algorithm in

EORTC QLQ-C30 scoring manual, version 3.0 (Fayers 2001).[2] A higher score represents a higher ("better") level of functioning, or a higher ("worse") level of symptoms.

The EQ-5D-5L is a generic measure of health status. For purposes of this study, the EQ-5D-5L will be used to generate utility scores for use in cost effective analyses. The EQ-5D-5L is a 5 item questionnaire that assesses 5 domains including mobility, self-care, usual activities, pain/discomfort and anxiety/depression plus a visual analog scale rating "health today" with anchors ranging from 0 (worst imaginable health state) to 100 (best imaginable health state). The scores for the 5 separate questions are categorical and cannot be analyzed as cardinal numbers. However, the scores for the 5 dimensions are used to compute a single utility score ranging from zero (0.0) to 1 (1.0) representing the general health status of the individual.

4.5.2. Analysis Methods

Compliance rates for completion of EORTC QLQ-C30 and EQ-5D-5L at each time point will be generated based on number of expected.

Key PRO endpoints

- EORTC-QLQ-C30 global health status/quality of life subscale
- EQ-5D-5L VAS and utility score.

The change from baseline at each time point will be summarized descriptively by treatment group. Line plot of mean with standard error over time will be displayed by treatment group. A mixed effects model with repeated measures analysis will be conducted estimating change from baseline at each time point between two treatments. ITT subjects who have a baseline value and at least one post-baseline value are included in the analysis. Change from baseline will be fitted to a mixed effects model including subjects as a random effect, and baseline value, treatment group, time in month, treatment-by-time interaction, and stratification factors as fixed effects.

A distribution based method will be used to define improvement/worsening in scores, i.e., half SD away from the mean score at baseline combining both treatment groups.

Time to improvement will be summarized by using descriptive statistics such as mean, standard deviation (SD), median and range.

Time to worsening will be estimated using Kaplan-Meier methods. The hazard ratio for D-VMP relative to VMP and its associated 95% confidence interval (CI) will be calculated based on the stratified Cox proportional hazards model by the stratification factor at randomization. Death due to disease progression will be considered as worsening. Subjects who have not met the definition of worsening will be censored at the last PRO assessment. Subjects without baseline assessment or post-baseline assessment will be censored at date of randomization.

Secondary PRO endpoints

These may include other QLQ-C30 scales:

- functional scales: physical, role, cognitive, emotional, and social
- symptom scales: fatigue, pain, and nausea and vomiting
- single-item score: dyspnea, loss of appetite, insomnia, constipation, diarrhea, and financial difficulties.

The change from baseline at each time point may be summarized descriptively by treatment group. Line plot of mean with standard error over time may be displayed by treatment group.

Time to improvement/worsening and the mixed effect model analysis, as described for the key PRO endpoints, may be performed as appropriate.

5. SAFETY

Safety assessment will be evaluated through AEs and clinical hematology and chemistry laboratory tests. Safety analyses will be based on the safety population and presented by the actual treatment received.

5.1. Adverse Events

All adverse events whether serious or non-serious, will be reported from the time a signed and dated informed consent form (ICF) is obtained until 30 days after the last dose of study treatment, until the subject withdraws consent for study participation, or until the subject starts subsequent antimyeloma therapy, whichever occurs first. AEs will be recorded in standard medical terminology and graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE), Version 4.03. For AE reporting, the verbatim term used in the CRF by investigators to identify adverse events will be coded using the latest version of Medical Dictionary for Regulatory Activities (MedDRA) coding dictionary.

Unless otherwise specified, at each level (e.g., system organ class and/or preferred term) of subject summarization in reporting the incidence of the AE, a subject is counted once if one or more events were recorded. For summarizing new onset events, all event records of the same preferred term from the same subject are to be linked by the onset date and the end date. If an event is followed by another event of the same preferred term with an onset date (or date/time) the same as or 1 day (or 1 minute if applicable) after the end date (or date/time) of the previous record and any features of the adverse event (i.e.: toxicity grades/seriousness/action taken) are different between these two records, these 2 records should be linked together and considered as one event. A Grade 5 event will be linked to previous event of the same preferred term if the onset date of Grade 5 record is the same or one day after the end date of previous record.

All summaries of AEs will be based on treatment-emergent adverse events (TEAEs), which are defined as any AE that occurs after start of the first study treatment through 30 days after the last study treatment; or the day prior to start of subsequent antimyeloma therapy, whichever is earlier; or any AE that is considered drug-related (very likely, probably, or possibly related) regardless of the start date of the event; or any AE that is present at baseline but worsens in toxicity grade or is subsequently considered drug-related by the investigator.

The incidence of TEAEs will be summarized overall, by MedDRA system organ class (SOC) and preferred term, by toxicity grade, and by relationship to study treatment administration. Specifically, the following AE summaries will be presented by treatment group:

5.1.1. Overview of TEAEs

An overview of TEAEs reported through the study will be provided for each treatment group. The overview will include summaries of subjects with TEAEs, with TEAEs related to study treatment, with TEAEs of maximum toxicity Grade of 1 to 5, SAEs, TEAEs leading to discontinuation of any study treatment, and deaths due to TEAE.

5.1.2. All TEAEs

- Incidence of TEAEs by MedDRA SOC and preferred term
- Most commonly reported (>10%) TEAE by MedDRA SOC and preferred term

5.1.3. Toxicity Grade 3 or 4 TEAEs

- Incidence of toxicity Grade 3 or 4 TEAEs, by MedDRA SOC and preferred term
- List of subjects with any toxicity Grade 3 or 4 TEAEs.
- Most commonly reported (>5%) Grade 3 or 4 TEAE by MedDRA SOC and preferred term

5.1.4. Study Treatment-Related TEAEs

- Incidence of TEAEs considered by the investigator to be related to study treatment, by MedDRA SOC, preferred term and relationship to study treatment
- Incidence of TEAEs with toxicity Grade 3 or 4 considered by the investigator to be related to study treatment, by MedDRA SOC and preferred term and relationship to study treatment.

5.1.5. Serious Adverse Events (SAEs)

- Incidence of treatment-emergent SAEs, by MedDRA SOC and preferred term
- Incidence of treatment-emergent SAEs considered by the investigator to be related to study treatment, by MedDRA SOC, preferred term and relationship to treatment

- Most commonly reported (>2%) SAEs by MedDRA SOC and preferred term
- List of subjects with any treatment-emergent SAEs

5.1.6. TEAEs Leading to Cycle Delays or Dose Modifications

Incidence of TEAEs leading to treatment cycle delays or dose modifications will be summarized by MedDRA SOC and preferred term. The summaries will be presented by all grades and Grade 3 or 4 for each treatment. This table will include TEAEs leading to cycle delays or dose modification of at least one of the study treatments, the dose modifications include dose delays, dose skipping, or dose reduction (applicable to bortezomib, melphalan and steroids).

5.1.7. TEAEs Leading to Discontinuation of Any Study Treatment

A summary of number of subjects who discontinued any study treatment because of 1 or more TEAEs by MedDRA system-organ class and preferred term will be provided. The summaries will be presented by all grades and Grade 3 or 4 for each treatment group. The AEs leading to discontinuation of any study treatment are based on AEs recorded in the AE CRF page with an action taken of drug withdrawal for any study treatment.

5.1.8. TEAEs Leading to Discontinuation of All Study Treatments

A summary of number of subjects who discontinued all study treatment because of 1 or more TEAEs by MedDRA system-organ class and preferred term will be provided. The summaries will be presented by all grades and Grade 3 or 4 for each treatment group. This table includes AEs leading to discontinuation of all study treatment for those subjects indicated as having discontinued study treatment due to an adverse event on the end of treatment CRF page.

A listing of subjects who discontinued all study treatment because of TEAEs will be provided.

5.2. Deaths

5.2.1. Death Due to TEAEs

The number of subjects who died due to treatment-emergent adverse events will be summarized by preferred term and relationship to study treatment for each treatment group. The TEAEs included in this table are AEs with outcome death or toxicity Grade of 5 recorded in the AE CRF page.

A listing of subjects who died due to treatment-emergent adverse events will be provided.

5.2.2. All Deaths

A summary of all death and cause of death will be tabulated overall and by treatment group. Specifically, the number of subjects who died during the study will be

summarized for the ITT population. The primary cause of death collected on the death information CRF page will be reported. If the primary cause of death is an AE, the number of subjects who have a related AE and unrelated AE will be further reported. The similar summaries will be presented for subjects who died within 30 days of last study treatment dose and within 60 days of first study treatment dose, respectively.

5.3. Adverse Events of Clinical Interest

5.3.1. Infusion-Related Reactions (IRR)

Subjects with any IRR associated with daratumumab administration will be summarized by MedDRA system-organ class and preferred term. The summaries will be presented by all grades, Grade 3, 4, and 5. In addition, the total number of subjects with an IRR in more than 1 infusion will be reported. Additionally, the timing of an IRR associated with daratumumab administration will be evaluated through a summary of an IRR by event onset time.

A listing of subjects with Grade 3 or higher treatment-emergent IRR associate with daratumumab administration will be provided. In addition, subjects with treatment-emergent IRR results in discontinuation of daratumumab will be listed.

5.3.2. Infections and infestations

Infections and infestations refer to adverse events with SOC of infections and infestations. A summary of number of subjects with 1 or more toxicity Grade 3 or 4 treatment-emergent infections and infestations by MedDRA preferred term and relationship to treatment will be provided. Additional summary analyses may include by onset time (i.e., ≤6 months vs. 6- ≤12 months vs. >12 months).

5.3.3. Peripheral Neuropathies

Peripheral neuropathies (PNs) refer to adverse events with high level term (HLT) of peripheral neuropathies NEC. Incidences of PNs will be summarized by MedDRA high level term and preferred term. The summaries will be presented by all grades and Grade 3 or 4 for each treatment group.

5.3.4. Hemorrhage Events

Hemorrhage events refer to the adverse events defined by Standardized MedDRA Queries (SMQ) with the first subcategory SMQ of hemorrhage terms (exclude laboratory terms). Incidences will be summarized by MedDRA system-organ class and preferred term. The summaries will be presented by all grades and maximum toxicity grade for each treatment group.

5.3.5. Tumor Lysis Syndrome

Tumor lysis syndrome (TLS) events refer to the adverse events defined by narrow Standardized MedDRA Queries (SMQ) of tumor lysis syndrome (haemorrhagic tumour

necrosis, tumour lysis syndrome, or tumour necrosis). A listing of subjects who reported any treatment-emergent TLSs during the study will be provided.

5.3.6. Second Primary Malignancies

A listing of subjects who reported second primary malignancies during the study will be provided. This listing will include diagnosis, study day of diagnosis, recurrence of a prior existing malignancy (yes, no) and pathology diagnosis (biopsy, aspirate, etc.) information whenever a second primary malignancy is observed. In addition, cumulative study treatment exposure, the treatment for second primary malignancy and the outcome information will also be presented in the listing. Second primary malignancies will be clinically reviewed and categorized as cutaneous/non-invasive, non-cutaneous/invasive or hematologic malignancies, which will be summarized accordingly.

5.3.7. Adverse Events by Subgroups

The following subgroup analysis of adverse events will be performed based on subgroups specified in Section 2.10:

- Overview of TEAEs
- All TEAEs
- Toxicity Grade 3 or 4 TEAEs
- SAEs.

5.4. Clinical Laboratory Tests

The evaluation of clinical laboratory tests will focus on the following selected laboratory analytes:

Hematology panel:

- hemoglobin
- platelet count
- white blood cell (WBC) count with absolute neutrophils and lymphocytes

Blood chemistry panel:

- alkaline phosphatase
- creatinine
- uric acid
- glucose
- bilirubin
- aspartate aminotransferase (AST)

- alanine aminotransferase (ALT)
- sodium
- potassium
- corrected calcium
- lactate dehydrogenase (LDH)
- creatinine clearance.

Blood samples for serum hematology are taken at the screening visit, on Days 1, 8, 15, and 22 of Cycles 1, Days 1 and 22 of Cycles 2-9, Day 1 every 4 weeks of Cycles 10+, and at the End-of-Treatment visit. Blood sample for biochemistry are taken at the screening visit, on Days 1, and 22 of Cycles 1-9, on Day 1 every 4 weeks of Cycle 10+ and at the End-of-Treatment visit.

Descriptive statistics (mean, SD, median, range) will be used to summarize observed laboratory values and change from baseline in observed value at each scheduled visit for each treatment group. Line plot of mean with standard error for each laboratory analyte over time will be displayed by treatment group for hemoglobin, neutrophils, lymphocytes, platelets, WBC, AST, ALT, creatinine, and creatinine clearance.

The worst toxicity grade in hematology and chemistry during the treatment will be summarized by treatment group and toxicity grade. Shift tables from baseline to worst toxicity grade during the treatment will be provided for each laboratory analyte listed above. These tables will summarize the number of subjects with each baseline CTC grade and changes to the maximum CTC grade.

5.5. Vital Signs and Physical Examination Findings

Baseline weight and vital signs (systolic and diastolic blood pressure, heart rate, and temperature) values will be summarized.

Post baseline physical examination findings were collected as AEs, and therefore will not be summarized.

5.6. Electrocardiogram (ECG)

Electrocardiograms (ECG) will be performed at Screening, Day 1 of Cycle 1, Cycle 3, and Cycle 6 immediately after end of daratumumab infusion, and End-of-Treatment visit.

The number and percentage of subjects with normal or abnormal 12-lead ECG results will be summarized.

5.7. ECOG Performance Score

ECOG performance status, which evaluates the effect of the disease status on the activities of daily living, will be assessed at Screening, every 3 months in the first year

from Cycle 1 Day 1, thereafter every 6 months until PD, 8 weeks and 16 weeks post-PD. Descriptive statistics will be used to summarize ECOG performance status at baseline, and post-baseline visits (including change from baseline), worst score during treatment (including change from baseline) for each treatment group. Shift table from baseline to worst score during the treatment may be provided.

6. PHARMACOKINETICS/IMMUNOGENICITY/PHARMACODYNAMICS

Unless specified otherwise, descriptive statistics (e.g., number of observations, mean, SD, median, and range) will be used to summarize pharmacokinetics and pharmacodynamics data. In addition, coefficient variation and geometric mean will be provided in the pharmacokinetic concentration summary.

6.1. Pharmacokinetics

6.1.1. Sampling Timepoints

For subjects assigned to D-VMP, blood samples to assess serum concentration (pharmacokinetics) of daratumumab will be obtained at Day 1 of Cycle 1, 3 and 6, End-of-Treatment and Post-Treatment Week 8. On a daratumumab dosing day, blood samples need to be collected before (up to 2 hours but not after the start of infusion) and immediately after (up to 2 hours but not before the end of infusion) daratumumab administration.

6.1.2. Pharmacokinetic Parameters

The pharmacokinetic parameters are defined as:

- Minimum observed concentration (C_{min}) - the concentration observed immediately before infusion.
- Maximum observed concentration (C_{max}) – the concentration observed after the end of infusion.

For daratumumab, the pharmacokinetic evaluations include Cmin and Cmax.

6.1.3. Analysis Methods

Pharmacokinetic analyses will be performed on the pharmacokinetic-evaluable population. All serum concentrations below the lowest quantifiable concentration or missing data will be labeled as such in the concentration data presentation. Concentrations below the lowest quantifiable concentration will be treated as zero in the summary statistics. All subjects and samples excluded from the analysis will be clearly documented in the study report.

Descriptive statistics will be used to summarize daratumumab serum concentrations at each sampling time point. Plot of mean (±SD) daratumumab serum peak and trough concentrations over time will be provided.

If sufficient data are available, population pharmacokinetic analysis of serum concentration-time data of daratumumab may be performed using nonlinear mixed-effects modeling. If population pharmacokinetic analysis is conducted, it may include data from other clinical studies; details will be provided in a population pharmacokinetic analysis plan and results will be presented in a separate report.

6.2. Immunogenicity

6.2.1. Sampling Timepoints

Samples to assess the generation of antibodies to daratumumab (immunogenicity) will be obtained from all subjects in the D-VMP group at Cycle 1 Day 1 predose, End-of Treatment, and Post-Treatment Week 8. In addition, any time an infusion-related reaction is observed during the study, an unscheduled blood sample should be drawn as soon as possible after the reaction for potential immune response analysis.

6.2.2. Analysis Methods

The incidence of antibodies to daratumumab (immunogenicity) will be summarized for all subjects who receive a dose of daratumumab and have appropriate samples for detection of antibodies to daratumumab. In addition, subjects who are positive for antibodies to daratumumab will also be listed. A listing of daratumumab concentrations at the time of each immunogenicity sample will also be provided.

6.3. Pharmacokinetic/Pharmacodynamic Analyses

If sufficient data are available, other pharmacokinetic/pharmacodynamic modeling may be performed, including exploring the relationship between serum concentrations of daratumumab and endpoints of clinical efficacy. If analysis is conducted, details and results of the analysis will be presented in a separate report.

7. BIOMARKER

Biomarker studies are designed to identify markers predictive of response (or resistance) to daratumumab. Planned analyses are based on the availability of clinically valid assays and may be deferred if emerging study data show no likelihood of providing useful scientific information. Results of biomarker analyses may be presented in a separate report.

Minimal residual disease (MRD) will be assessed for all subjects who achieve a CR/sCR. Molecular subtyping will be done through next generation sequencing (NGS) to evaluate daratumumab response rates in risk-stratified multiple myeloma subpopulations.

7.1. Minimal Residual Disease (MRD)

Bone marrow aspirates will be collected at baseline from all patients, as well as on treatment in those subjects who attain or suspect to have a CR/sCR to monitor MRD. MRD will be monitored using ClonoSEQ Assay 2.0 on bone marrow aspirate.

7.1.1. Sampling Timepoints

For all subjects, a fresh bone marrow aspirate will be obtained at baseline, as well as on treatment for subjects who were suspected to have a CR/sCR, at the time of the suspected CR/sCR, and 4 landmark time points: 12, 18, 24 and 30 months (+2 months) after first dose of study treatment.

7.1.2. Analysis Methods

Details on MRD negativity rate analyses are described in Section 4.3.4.

In addition, exploratory landmark analyses will be conducted to correlate MRD negativity results (as either binary or continuous values) with long-term clinical endpoints such as PFS. Subgroup analysis of PFS by MRD negativity status may also be performed. Similar analysis may be performed for subjects who achieved CR or better.

7.2. Molecular Subtyping

7.2.1. Sample collection and characterization

Viable, frozen CD138+ bone marrow aspirate samples were collected at baseline for DNA/RNA sequencing analysis. Peripheral blood mononuclear cell (PBMC) samples were collected at baseline to use as controls.

7.2.2. Molecular Subtypes

A portion of the bone marrow aspirate samples collected as specified in the Section 7.2.1 will be utilized for translocation/mutation/genomic analysis using whole exome-sequencing and RNA-sequencing to assess whether specific molecular subtypes having chromosomal aberrations such as del17p, t(4;14), t(14;16), are responsive to daratumumab treatment. PBMC samples will be sent for whole exome sequencing and used as controls for mutation and copy number analysis. Examples of molecular subtypes are:

- DNA sequence
  - del17p
  - amp1q
- RNA sequence
  - t(4;14)
  - t(14;16).

7.2.3. Molecular Risk Subgroup Analysis

To determine if daratumumab combined with VMP (D-VMP) will lead to improved clinical responses in standard as well as high-risk molecular subgroups, over VMP the following exploratory analysis will be conducted by using the similar analysis methods specified in Sections 4.3.1.2 and 4.2.2.

- To evaluate ORR for subjects in high-risk molecular subgroup and subjects with specific molecular subtypes such as del17p, t(14;16), t(4;14).
- To evaluate MRD negativity rate for subjects in standard-risk and high-risk molecular subgroups.
- To evaluate PFS and OS for subjects in high-risk molecular subgroup and subjects with specific molecular subtyping such as del17p, t(14;16), t(4;14).

Subgroup exploratory analysis of ORR and PFS by molecular risk within treatment group, as well as between molecular subgroups across treatment group will be conducted.

8. MEDICAL RESOURCE UTILIZATION

Medical resource utilization (excluding study infusion administration) will be descriptively summarized by treatment group. Frequencies of hospitalization, outpatient visits, type of hospitalization or outpatient visit, reasons for hospitalization or outpatient visit, durations of hospitalization or outpatient visit will be calculated and tabulated.

REFERENCES

1. Durie BG, Harousseau JL, Miguel JS, et al. International uniform response criteria for multiple myeloma. Leukemia 2006;20:1467–1473. Corrigenda/Erratum in: Leukemia. 2007; 21:1134-1135.
2. Fayers P, Aaronson N, Bjordal K, Curran D, Groenvold M, on behalf of the EORTC Quality Of Life Study Group. EORTC QLQ-C30 Scoring Manual: Third Edition. 2001. ISBN 2-9300 64-22-6.
3. Rajkumar SV, Harousseau J-L, Durie B, et al. Consensus recommendations for the uniform reporting of clinical trials: report of the International Myeloma Workshop Consensus Panel 1. Blood. 2011; 4691-4695.
4. Tang and Geller, Closed Testing Procedures for Group Sequential Clinical Trials with Multiple Endpoints, Biometrics, Dec;55(4):1188-92.
5. Robins JM, Finkelstein DM. Correcting for non-compliance and dependent censoring in an AIDS clinical trial with Inverse Probability of Censoring Weighted (IPCW) log-rank tests. Biometrics 2000; 56:779-788.
6. Branson M and Whitehead J. Estimating a treatment effect in survival studies in which patients switch treatment. Stat Med 2002; 21:2449-2463.
7. Robins JM and Tsiatis AA. Correcting for non-compliance in randomized trials using Rank Preserving Structural Failure Time models. Commun Statist Theory Meth 1991; 20: 2609-2631.

ATTACHMENT 1: ADDITIONAL EXPLORATORY ANALYSIS TO SUPPORT HEMAR