Janssen Research & Development

Statistical Analysis Plan
Independent Data Monitoring Committee (IDMC) Review

A Phase 3 Randomized Study Comparing JNJ-68284528, a Chimeric Antigen Receptor T cell (CAR-T) Therapy Directed Against BCMA, versus Pomalidomide, Bortezomib and Dexamethasone (PVd) or Daratumumab, Pomalidomide and Dexamethasone (DPd) in Subjects with Relapsed and Lenalidomide-Refractory Multiple Myeloma

Protocol 68284528MMY3002; Phase 3
Amendment 1

JNJ-68284528 (ciltacabtagene autoleucel)

Status:	Approved
Date:	08 November 2022
Prepared by:	Janssen Research & Development, LLC
Document No.:	EDMS-RIM-305627, 2.0

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

TABLE OF CONTENTS

TABLE OF CONTENTS **2**

LIST OF IN-TEXT TABLES **3**

AMENDMENT HISTORY **4**

ABBREVIATIONS **6**

1. INTRODUCTION **8**
1.1. Trial Objectives 8
1.2. Trial Design 8
1.3. Statistical Hypotheses 10
1.4. Objective of IDMC Reviews 10
1.5. Timing of IDMC Reviews 10
1.6. Safety Evaluation at IDMC Reviews 10
1.7. Efficacy Evaluation at the IDMC Review 11
1.8. Level of Blinding 11

2. GENERAL ANALYSIS DEFINITIONS **11**
2.1. Visit Windows 11
2.2. Pooling Algorithm for Analysis Centers 11
2.3. Analysis Sets 12
2.3.1. All Consented Analysis Set 12
2.3.2. Efficacy Analysis Set(s) 12
2.3.3. Safety Analysis Set 12
2.4. Study Day and Relative Day 12
2.5. Baseline 12
2.6. Measurable Disease of Multiple Myeloma at Baseline and Measurable Type 12

3. SUBJECT INFORMATION **13**
3.1. Demographics and Baseline Characteristics 13
3.2. Disposition Information 14
3.3. Extent of Exposure 14
3.4. Prior and Concomitant Medications 15
3.4.1. Prior Exposure to Multiple Myeloma Therapies 15
3.4.2. Refractory Disease 16
3.4.3. Concomitant Medications 16
3.5. Subsequent Anti-myeloma Therapy 16

4. EFFICACY **16**
4.1. Analysis Specifications 16
4.1.1. Level of Significance 16
4.1.2. Computerized Algorithm 17
4.1.3. Data Handling Rules 18
4.2. Primary Endpoint(s) 18
4.2.1. Definition 18
4.2.2. Analysis Methods 18
4.2.3. Sensitivity Analysis 19
4.2.3.1. Progressive Disease Based on Investigator Assessment 19
4.2.3.2. Unstratified Weighted Analysis of PFS 19
4.2.3.3. Standard “Unweighted” Stratified Analysis of PFS 19
4.3. Key Secondary Endpoints 19
4.3.1. ORR 19
4.3.1.1. Definition 19
4.3.1.2. Analysis Methods 19
4.3.2. Response rate of CR/sCR 20

4.3.2.1. Definition 20
4.3.2.2. Analysis Methods 20
4.3.3. Overall MRD Negativity Rate 20
4.3.3.1. Definition 20
4.3.3.2. Analysis Methods 20
4.3.4. Overall Survival 20
4.3.4.1. Definition 20
4.3.4.2. Analysis Methods 21

5. SAFETY 21
5.1. Adverse Events 21
5.2. Deaths 22
5.3. Adverse Events of Special Interest 22
5.3.1. Cytokine Release Syndrome 22
5.3.2. Neurologic Adverse Events 23
5.3.3. Second Primary Malignancies 23
5.4. Clinical Laboratory Tests 23

REFERENCE 25

LIST OF IN-TEXT TABLES

Table 1 Demographic Variables 13
Table 2 Baseline Characteristics Variables 14

AMENDMENT HISTORY

Initial Release: 31 October 2020

Amendment 1: 08 November 2022

Section	Change	Rationale
Abbreviations	Updated the abbreviations.	Revised for consistency.
1.1, 4.1, 4.3	Updated the list of major secondary endpoints and the hierarchical testing order: ▪ Removed two MRD endpoints from the list: Rate of MRD negativity and CR/sCR at 12 months±3 months, and Rate of sustained MRD negative status ▪ Added ORR in the list, it will be tested after the rate of CR/sCR and before Overall MRD negativity rate.	Health authority feedback.
1.2, 1.5, 4.1	Updated the information fraction from 66% (165 PFS events) to 75% (188 PFS events) at the PFS interim analysis.	Health authority feedback.
2.6	Added the measurable disease section	Aligned with protocol amendment 2 which was released on 02 Jul 2021.
3.1	Add gain/amp (1q) to the high-risk cytogenetic abnormalities.	Aligned with protocol amendment 2 which was released on 02 Jul 2021.
3.1	The replacement of "extramedullary plasmacytomas" with "soft tissue plasmacytomas".	Revised for clarity.
4.1.1	Added OS futility analysis by using hazard ratio >1 as a guidance.	Health authority feedback.
4.2.2	Added following clarification for primary analysis. "Within each stratification factor, if one of the strata has very small sample size, i.e., less than 20% of total sample size, this stratum may be pooled with other strata in the stratified analysis."	To provide flexibility for stratified analysis when there's small stratum in a stratification factor.
4.2.2	▪ Updated the primary analysis of PFS to constant piecewise weighted (CPW) log-rank test. ▪ Moved the original primary analysis of PFS (standard stratified log-rank test) to the section 4.2.3.3 as a sensitivity analysis	To reflect the delayed treatment effect according to health authority feedback.

	▪ Added section 4.2.3.2 unstratified weighted analysis of PFS by using CPW log-rank test	
5.1	Re-phrased the definition of TEAE.	Revised for clarity.
5.3	Removed the tumor lysis syndrome from the AEs of special interest	Aligned with protocol amendment 2 which was released on 02 Jul 2021.
References	Added reference of CPW log-rank test and Polack et al.	Add appropriate references.
Globally	The replacement of "MRD negative rate" with "MRD negativity rate".	Revised for consistency.
Globally	Minor grammatical, formatting, spelling changes	Minor errors were noted

ABBREVIATIONS

AE	adverse event
ALT/SGPT	alanine aminotransferase
AST/SGOT	aspartate aminotransferase
ASTCT	American Society for Transplantation and Cellular Therapy
BCMA	B-cell maturation antigen
CAR-T	Chimeric Antigen Receptor T cell
CR	complete response
CRS	cytokine release syndrome
CRF	case report form
CMH	Cochran Mantel Haenszel
CPW	Constant piecewise weighted
CSR	Clinical Study Report
CTCAE	Common Terminology Criteria for Adverse Events
DOR	duration of response
DPd	Daratumumab, Pomalidomide and Dexamethasone
DPS	Data Presentation Specifications
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic case report form
eGFR	estimated glomerular filtration rate
FLC	Free-light chain
HLGT	high level group term
HLT	high level term
IA	interim analysis
ICANS	immune effector cell-associated neurotoxicity syndrome
ICE	immune-effector cell-associated encephalopathy
IDMC	Independent Data Monitoring Committee
IMiD	Immunomodulatory drug
IMWG	International Myeloma Working Group
ISS	International Staging System
ITT	intent-to-treat
IV	intravenous
MedDRA	Medical Dictionary for Regulatory Activities
MM	Multiple Myeloma
MRD	minimal residual disease
NGS	next generation sequencing
ORR	overall response rate
OS	overall survival
PD	progressive disease
PFS	Progression-free Survival
PFS2	PFS on next-line of therapy
PI	proteasome inhibitor
PO	orally
PR	partial response
PT	preferred term
PVd	Pomalidomide, Bortezomib and Dexamethasone
SAE	serious adverse event
SAP	Statistical Analysis Plan
SC	subcutaneously
sCR	stringent complete response
SD	standard deviation
SOC	system organ class
SPM	Secondary primary malignancies
SSG	Statistical Support Group
TEAE	treatment-emergent adverse event
VGPR	very good partial response

WHO-DD	World Health Organization Drug Dictionary

1. INTRODUCTION

This statistical analysis plan (SAP) contains definitions of the analysis sets, derived variables and statistical methods for the planned independent data monitoring committee (IDMC) review of the safety and efficacy data at the time of interim analysis for the primary efficacy endpoint and periodic safety data review as specified in the protocol 68284528MMY3002.

1.1. Trial Objectives

The primary objective of the trial is to compare the efficacy of cilta-cel with standard therapy (either PVd or DPd) by progression-free survival (PFS) in subjects with relapsed and lenalidomide-refractory MM.

Key secondary objectives include the following: comparing the efficacy of cilta-cel with standard therapy by rate of complete response (CR)/stringent complete response (sCR), overall response rate (ORR), overall minimal residual disease (MRD) negativity rate, overall survival (OS); assessing the safety profile of cilta-cel by incidence and severity of adverse events.

1.2. Trial Design

This is a Phase 3, randomized, open-label, multicenter study to determine whether treatment with cilta-cel will provide efficacy benefit compared to standard therapy (PVd or DPd) in subjects with relapsed and lenalidomide-refractory multiple myeloma (MM).

The study will be conducted in 3 phases: Screening, Treatment and Follow-Up. During the screening phase, all subjects will provide written consent for study participation and will be screened for study eligibility with 28 days prior to randomization.

Approximately 400 subjects will be randomized 1:1 to receive either standard therapy with PVd or DPd (Arm A) or to receive cilta-cel (Arm B).

Randomization will be stratified by investigator's choice of PVd or DPd, International Staging System (ISS) (I vs. II vs. III) and number of prior lines of therapy (1 vs. 2 or 3).

Arm A

Prior to screening the investigator will determine if the subject will be treated with PVd or DPd, as standard therapy based on the subject's prior exposure to anti-myeloma therapies. After meeting eligibility criteria defined in Section 5 of the protocol, subjects randomized to Arm A will receive either PVd or DPd. Subjects should start treatment with PVd or DPd within 7 days after randomization. Details of interventions administration are in protocol section 6.

Subjects randomized to Arm A will continue to receive PVd or DPd until confirmed progressive disease (PD), death, intolerable toxicity, withdrawal of consent, or end of the study. All subjects in Arm A will have an end-of-treatment visit. Subjects who discontinue PVd or DPd for any reason, other than PD or withdrawal of consent, will continue to be followed for response assessment until confirmed PD or the start of a subsequent anti-myeloma therapy. After confirmed PD, subjects will then be followed for survival status, subsequent anti-myeloma therapies,

response to subsequent anti-myeloma therapies including the date of subsequent progression (PFS2), and the occurrence of second primary malignancies (SPMs) every 16 weeks until the end of the study (defined when approximately 250 deaths have occurred). The occurrence of SPMs must be reported throughout study treatment and must continue to be reported until the end of the study.

Arm B

Eligible subjects randomized to Arm B will receive cilta-cel and will undergo the following steps:

- Apheresis to collect peripheral blood mononuclear cells (PMBCs). Apheresis should be performed 3 to 6 days after randomization.
- Bridging therapy should be started after apheresis but not more than 7 days after randomization. One cycle of bridging therapy with either PVd or DPd (determined by the investigator based on the subject's prior anti-myeloma therapy). Additional cycles of PVd or DPd may be considered based on subject's clinical status and timing of availability of cilta-cel. Investigator must contact the sponsor for approval.
- A wash-out period must occur from the last dose of bridging therapy until prior to initiating the conditioning regimen, the length of which will depend on whether PVd or DPd was given as a bridging therapy (See Protocol section 6.1.5.4).
- Conditioning regimen consists of IV cyclophosphamide 300 mg/m^2 and fludarabine 30 mg/m^2 daily for 3 days. The dose of fludarabine should be reduced to 24mg/m^2 for subjects with an estimated glomerular filtration rate (eGFR) of 30 to 70 mL/min/1.73 m^2.
- Cilta-cel will be administered with the target dose of 0.75 x 10^6 CAR-positive viable T cells/kg (range: 0.5-1.0 x 10^6 CAR-positive viable T cells/kg) 5 to 7 days after the start of the conditioning regimen.
- Post-infusion Follow Up: Intensive safety, pharmacokinetics (PK), biomarkers, and efficacy monitoring during the first 112 days following cilta-cel administration.
- Post-treatment Follow Up: Starts once the post-infusion follow-up is complete (on Day 112) and lasts until end of study. In the post treatment follow-up phase, subjects will continue to be monitored for efficacy until confirmed PD, death, or withdrawal of consent. After confirmed PD, subjects will be followed for survival, subsequent anti-myeloma therapies, response to subsequent anti-myeloma therapies, including the date of subsequent progression (PFS2), and SPMs every 16 weeks until the end of the study (defined when approximately 250 deaths have occurred).

For both Arm A and Arm B, disease status will be evaluated according to the International Myeloma Working Group (IMWG) consensus recommendations for MM treatment response criteria (see Protocol appendix 7). The primary endpoint and response-related secondary endpoints will be determined using a validated computerized algorithm.

One interim analysis for efficacy is planned for PFS. The interim analysis will be performed when approximately 188 PFS events are observed (75% information fraction). The final analysis of PFS will be performed after approximately 250 PFS events are observed. The end of the study is

defined as the time when approximately 250 deaths have occurred or the sponsor terminates the study, whichever occurs first.

All study evaluations will be conducted according to the Schedule of Activates as per protocol Section 1.3.

Safety data will be periodically reviewed by an Independent Data Monitoring Committee (IDMC). Efficacy data will also be reviewed at the planned interim analysis for the primary efficacy endpoint by IDMC.

1.3. Statistical Hypotheses

The primary efficacy endpoint of the study is PFS. The null hypothesis is that there is no difference in PFS between cilta-cel and standard therapy (PVd or DPd) in subjects who have previously received 1 to 3 prior lines of therapy, that included a proteasome inhibitor (PI) or immunomodulatory drug (IMid), and who are refractory to lenalidomide.

1.4. Objective of IDMC Reviews

The objective of the planned periodic safety review is to have a comprehensive evaluation of the safety of the study treatment in Arm A and Arm B.

The objective of the interim analysis is to provide aggregated summary of efficacy and safety data to the IDMC so the IDMC can make recommendations on the study conduct based on the benefit/risk ratio observed.

1.5. Timing of IDMC Reviews

The IDMC will review unblinded reports for safety after the first approximately 20 subjects with at least 3 months follow-up after randomization and thereafter every 6 months until the study end, inform the Study Team and make recommendations regarding continuation of the study. More frequent review of the safety data may be necessary upon agreement between the IDMC and sponsor.

The IDMC will also review unblinded reports for safety and efficacy data at the first interim analysis for PFS. The interim analysis will be performed after all subjects have enrolled and when approximately 188 PFS events, which is approximately 75% of the total planned PFS events, have been accumulated. The significance level for PFS at this interim analysis to establish the superiority of cilta-cel over standard therapy will be determined based on the observed number of PFS events at the interim analysis.

1.6. Safety Evaluation at IDMC Reviews

Safety evaluation will focus on assessment of study treatment discontinuation, adverse events, serious adverse events, Grade ≥3 adverse events, adverse events leading to study treatment discontinuation, adverse events of special interest, deaths due to adverse event, and clinical laboratory parameters. The detailed analyses of safety are described in Section 5.

1.7. Efficacy Evaluation at the IDMC Review

IDMC will review the efficacy data at the first interim analysis for PFS. Efficacy evaluation will focus on assessment of the primary endpoint PFS and the key secondary efficacy endpoints. The detailed analyses of efficacy are described in Section 4.

1.8. Level of Blinding

Central randomization will be implemented in this study. Subjects will be assigned randomly in a 1:1 ratio to either treatment Arm A (standard therapy, PVd or DPd) or treatment Arm B (cilta-cel) based on a computer-generated randomization schedule prepared before the study by or under the supervision of the sponsor. The randomization will be balanced by using randomly permuted blocks and will be stratified according to 3 baseline factors: investigator's choice of PVd vs. DPd, ISS at screening (I vs. II vs. III), and number of prior lines of therapy (1 vs. 2 or 3). Investigators will screen and if eligible randomize each subject using an interactive web response system (IWRS). Each subject will be assigned a unique subject number.

Prior to the scheduled IDMC review meetings, a Statistical Support Group (SSG) independent of the Sponsor will prepare the data package and send the data package to the IDMC members. The data available for IDMC review will be unblinded with respect to both treatment groups and individual subject. However, all participating investigators and the Sponsor's study team will be blinded to the aggregated data until analysis of the primary endpoint of PFS, which will occur after 250 PFS events have occurred unless the sponsor decides to unblind the study results at the interim PFS analysis.

2. GENERAL ANALYSIS DEFINITIONS

Study treatment is defined as following:

- Arm A: Study treatment refers to the standard therapy (investigator's choice of PVd or DPd)
- Arm B: Study treatment refers to a sequence of apheresis, bridging therapy (investigator's choice of PVd or DPd), conditioning regimen (cyclophosphamide and fludarabine), and cilta-cel infusion.

Study drug refers to cilta-cel.

Continuous data will be summarized by the mean, standard deviation (SD), median, minimum and maximum. Categorical data will be presented by absolute and relative frequency (n and percentages).

2.1. Visit Windows

Unless otherwise specified, data to be analyzed or listed over time will be presented by day and time point (as appropriate) that are recorded in the electronic case report form (eCRF).

2.2. Pooling Algorithm for Analysis Centers

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

2.3. Analysis Sets

2.3.1. All Consented Analysis Set

All consented analysis set includes all subjects who signed the informed consent for the study.

2.3.2. Efficacy Analysis Set(s)

The intent-to-treat (ITT) analysis set includes all subjects who were randomized in the study.

The summary of study populations, analyses of disposition, demographics, baseline characteristics, and efficacy endpoints will be primarily analyzed based on this analysis set. All subjects in the ITT analysis set will be analyzed according to their randomized treatment group, regardless of the actual treatment received.

2.3.3. Safety Analysis Set

The safety analysis set includes all subjects who received any part of study treatment.

This analysis set will be used for safety analyses. Subjects will be analyzed according to the study treatment they actual receive.

2.4. Study Day and Relative Day

Study Day 1 is defined as following:

- For safety assessments, Study Day 1 refers to the initial date of study treatment
- For efficacy assessments, Study Day 1 refers to the date of randomization.

Study day or relative day for a visit is defined as:

- Visit date - (date of Study Day 1) +1, if visit date is ≥ date of Study Day 1
- Visit date - date of Study Day 1, if visit date < date of Study Day 1

There is no 'Day 0'.

CAR-T Day 1 refers to the start of initial cilta-cel infusion in Arm B.

2.5. Baseline

Baseline measurement is defined as the last non-missing measurement prior to the initiation of study treatment (including time if time is available).

For assessments specific to cilta-cel infusion in Arm B (such as ICE neurologic test, CAR-T chemistry), the last non-missing measurement prior to the start of cilta-cel infusion will be used).

2.6. Measurable Disease of Multiple Myeloma at Baseline and Measurable Type

Measurable disease at baseline based on the central laboratory assessments is defined by any of the following:

- 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 the serum or the urine: Serum free light chain (FLC) ≥10 mg/dL and abnormal serum FLC ratio.

If a subject meets the criteria for serum M-protein, the measurable disease type is serum; otherwise, if a subject meets the criteria for urine M-protein, the measurable disease type is urine; otherwise, if a subject meets the criteria for FLC, the measurable disease type is FLC. If a subject meets both of the criteria for serum M-protein and urine M-protein, then the measurable disease type is "serum and urine".

3. SUBJECT INFORMATION

Analysis of subject disposition, demographic and baseline disease characteristics will be conducted on the ITT analysis set. No statistical comparisons between the two treatment groups will be performed.

3.1. Demographics and Baseline Characteristics

Table 1 presents a list of the demographic variables that will be summarized by treatment group and overall, for the ITT analysis set.

Table 1 Demographic Variables

Continuous Variables:	**Summary Type**
Age (years)	Descriptive statistics (N, mean, standard deviation [SD], median and range [minimum and maximum]
Weight (kg)	
Height (cm)	
Body Surface Area (BSA) (m^2)	
Categorical Variables	
Age (<65 years, 65-75 years, >75 years)	Frequency distribution with the number and percentage of subjects in each category.
Sex (male, female, undifferentiated, unknown)	
Race[a] (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or other Pacific Islander, White, Other, Not Reported)	
Ethnicity (Hispanic or Latino, not Hispanic or Latino)	
Region (European union [EU], North America, Other)	
Baseline ECOG performance status (0, 1)	
Childbearing potential (of childbearing potential, permanently sterilized, postmenopausal, not applicable)	

[a]If multiple race categories are indicated, the Race is recorded as 'Multiple'

Table 2 presents a list of the baseline characteristics variables that will be summarized by treatment group and overall, for the ITT analysis sets.

Table 2 Baseline Characteristics Variables

Continuous Variables	**Summary Type**
Time since initial multiple myeloma diagnosis (years)	Descriptive statistics (N, mean, standard deviation [SD], median and range [minimum and maximum]).
Categorical Variables	
Baseline disease characteristics	Frequency distribution with the number and percentage of subjects in each category.
Type of multiple myeloma (IgG, IgA, IgM, IgD, IgE, free light chain only, biclonal, or negative immunofixation)	
Type of measurable disease (Serum only, Serum and urine, Urine only, or Serum FLC)	
Number of lytic bone lesions (None, 1-3, 4-10, more than 10)	
Number of soft tissue plasmacytomas [Yes, No]	
Presence of evaluable bone marrow assessment (yes, no)	
Bone marrow biopsy % plasma cells (<10, 10-30, >30-<60, ≥60)	
Bone marrow cellularity (hypercellular, normocellular, hypocellular, indeterminate) by biopsy	
Standard-risk and high-risk cytogenetic abnormalities (del17p, t(4;14), t(14;16), gain/amp(1q))	
Baseline safety related variables	
Medical history collected at screening visit (by system organ class, preferred term)	
ISS staging (I, II, III)	
Number of prior therapies (1, 2 or 3)	
Intention to be treated with – PVd, DPd	

3.2. Disposition Information

The number of subjects who are randomized, treated, ongoing, and discontinued treatment with reasons of discontinuation reported on eCRF will be summarized. The number of subjects who discontinued from study with the reported reasons will also be presented.

Listings of subjects who discontinued study treatment and reasons will be provided.

3.3. Extent of Exposure

Extent of exposure to study treatment will be summarized and presented based on the safety analysis set.

Arm A

The number and percentage of subjects within each cycle will be summarized.

The maximum number of treatment cycles received for each subject will be summarized by frequency and descriptive statistics. The maximum number of treatment cycles for each subject is the largest cycle number in which a subject receives any non-zero dose of PVd or DPd.

Descriptive statistics for summary of total number of treatment cycles (N, mean, SD, median, and range (minimum, maximum)) will be presented by study treatment for the safety analysis set.

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 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 treatment group and summarized accordingly.

Arm B

The number and percentage of subjects within each cycle of bridging therapies will be summarized.

Descriptive statistics (N, mean, SD, median, and range) for cilta-cel cells infusion duration (minutes), total volume (mL) of the cilta-cel cells infusion bag, and time from apheresis (days) to cilta-cel cells infusion will be presented.

The total CAR-positive viable T cells infused ($x10^6$ cells), weight-adjusted CAR-positive viable T cells infused ($x10^6$ cells/kg) will be summarized (based on weight at apheresis and cilta-cel infusion date).

Similarly, the total dose of the cyclophosphamide and fludarabine infusion (mg/m^2) will be respectively summarized.

3.4. Prior and Concomitant Medications

Prior and Concomitant medications will be coded using the World Health Organization (WHO) Drug Dictionary (WHO-DD).

3.4.1. Prior Exposure to Multiple Myeloma Therapies

A summary of prior exposure to multiple myeloma therapies (systemic therapy, stem cell transplant, radiotherapy, or cancer-related surgery/procedure) will be provided by treatment group and overall. Specifically, the number of prior lines of therapy will be calculated and summarized by the following categories: 1, 2-3, >3, 2, 3 through frequency and descriptive statistics. Additionally, the summary of prior systemic therapies will be presented by therapeutic class, pharmacologic class and drug name.

The therapy classes include proteasome inhibitors (PI), immunomodulatory drugs (IMiD), anti-CD38 antibody, steroids, alkylating agents and anthracyclines. Therapies included in the PI class are: bortezomib, carfilzomib, oprozomib, marizomib, and ixazomib; IMiD class: lenalidomide, pomalidomide, and thalidomide; anti-CD38 antibody class: daratumumab; monoclonal antibody class: elotuzumab and isatuximab; and steroids class: dexamethasone and prednisone, among others.

The number of subjects who had prior exposure to multiple therapy classes (e.g., PI + IMiD) or multiple therapies (e.g., bortezomib + lenalidomide) may be provided, if the number of subjects who exposed to those therapy classes or therapies is sufficient.

3.4.2. Refractory Disease

Refractory is defined as being nonresponsive (either failure to achieve minimal response or development of progressive disease) while on a therapy or progressed within 60 days of the end of the therapy.

The number and percentage of subjects' refractory status to PI, IMiD, or anti-CD38 antibody therapy class will be summarized by the following categories: none, PI and IMiD, anti-CD38 and IMiD, PI and anti-CD38 and IMiD. Refractory to specific prior multiple myeloma therapy, such as bortezomib, carfilzomib, ixazomib, lenalidomide, pomalidomide, thalidomide, daratumumab or isatuximab, and the relevant combinations of the aforementioned therapies will be provided separately.

The incidence of subjects who are refractory to their last line of therapy will be reported.

3.4.3. Concomitant Medications

Concomitant medications collected on the eCRF page during the study will be summarized by therapy class and drug name. Concomitant medications will be coded using the latest version of WHO Drug Dictionary (WHO-DD).

3.5. Subsequent Anti-myeloma Therapy

The total number of subjects who received subsequent anti-myeloma therapy will be reported for each treatment group. A summary of subsequent anti-myeloma therapy will be presented by preferred ATC classification and drug name, coded using the latest WHO-DD version.

In addition, for subjects who received subsequent anti-myeloma therapy, their best response to the first subsequent anti-myeloma therapy will be summarized.

4. EFFICACY

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 for superiority at the PFS interim analysis is to be determined by the observed number of events using the O'Brien-Fleming boundaries as implemented by the Lan-DeMets alpha spending method. Assuming 188 PFS events are observed at the interim analysis, the two-sided alpha to be spent in the interim analysis is 0.0193 and it will be 0.0442 for the final analysis. If the observed two-sided p-value is smaller than 0.0193 or 0.0442 at the corresponding interim or

final analysis respectively, the superiority of cilta-cel over PVd or DPd with respect to PFS will be established.

If the primary endpoint PFS is statistically significant the following major secondary endpoints ordered below will be sequentially tested for superiority by utilizing a hierarchical procedure as proposed by Tang and Geller (1999)[4] to control familywise Type I error rate at a 2-sided significance level of 0.05 (overall). The major secondary endpoints will be sequentially tested in the following order at the interim analysis and the final analysis for PFS when PFS is statistically significant:

1. Rate of CR/sCR
2. Overall response rate
3. Overall MRD negativity rate
4. Overall survival

The significance level at the interim analysis will be determined by the alpha-spending function specific to that endpoint.

- For the rate of CR/sCR: if at the interim analysis for PFS the information level is 80% then alpha to be spent for each of these endpoints is 0.0244 (2-sided).
- For the overall response rate: if at the interim analysis for PFS the information level is 90% then alpha to be spent is 0.0363 (2-sided)
- For the overall MRD negativity rate: if at the interim analysis for PFS the information level is 75% then alpha to be spend for each of these endpoints is 0.0193 (2-sided).
- For OS, the first and second interim analyses will be performed at the same time as the interim and final analysis for PFS, respectively. The alpha spending function for OS will be power (Kim-DeMets) spending function with parameter=2. For example, if the observed number of OS events is 114 at the time of the first interim analysis for PFS (i.e., approximately 188 PFS events), the alpha to be spent for OS is 0.0104 (2-sided). IDMC will evaluate futility or potential for harm using OS hazard ratio >1 as a guidance for futility decision, taking into consideration the totality of data.

4.1.2. Computerized Algorithm

Efficacy assessment of response and disease progression will also be performed using a validated computerized algorithm, following the IMWG criteria[2,3] (Attachment 10.7 of Protocol). As a sensitivity analysis, investigator assessment of response and disease progression using the IMWG response criteria[2,3] will also be summarized for the primary endpoint and the key secondary endpoints.

4.1.3. Data Handling Rules

There is no imputation planned for missing efficacy endpoint values.

4.2. Primary Endpoint(s)

4.2.1. Definition

PFS is defined as the duration from the date of randomization to either progressive disease (PD) or death, whichever comes first. Disease progression will be determined according to the International Myeloma Working Group (IMWG) criteria[2,3], based on computerized algorithm. For Subjects who have not progressed and are alive, data will be censored at the last disease assessment before the start of any subsequent anti-myeloma therapy. Subjects without any post-baseline disease assessment will be censored at the date of randomization.

4.2.2. Analysis Methods

Analysis of PFS will be based on the ITT analysis set. 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.

Based on the study design, both treatment arms are expected to receive the same bridging therapy (approximately 2 cycles) immediately after randomization. Therefore, it is expected that there would be no separation of the Kaplan-Meier curves between the two arms initially. Such "delayed treatment effect" is similar to those observed in vaccine trials, where the primary analyses only count new infections after certain time period. For example, the Phase 3 Covid-19 vaccine study analyzed only new SARS-COV-2 infections 1 week after the second vaccine dose (Polack 2020)[6].

To account for the delayed effect due to the same bridging therapies, only PFS events that occurred more than 8 weeks post-randomization will be included in the treatment comparison of overall PFS distributions between the two arms. This is essentially a weighted log-rank test with a threshold lag of length t^* (Zucker 1990)[5] where treatment has no detectable effect during the period [0, t^*] and is fully effective after t^* with t^*=End of Week 8, which can be equivalently expressed as the constant piecewise weighted (CPW) log-rank test, where the weight is 0 for the first 8 weeks post-randomization, and 1 afterwards. The p-value from a stratified CPW log-rank test will be reported. Hazard ratio (Arm B vs. Arm A) and its 95% confidence interval will be estimated based on a stratified Cox's regression model, similarly, including only PFS events that occurred more than 8 weeks post-randomization, with treatment as the sole explanatory variable.

Stratification factors used in the analyses include investigator's choice of PVd or DPd, ISS staging (I, II, III), and number of prior lines (1 vs. 2 or 3). Within each stratification factor, if one of the strata has very small sample size, i.e., less than 20% of total sample size, this stratum may be pooled with other strata in the stratified analysis. Noninformative censoring is assumed for all types of censoring and distinct baseline hazard for each stratum and common proportional hazard ratio across strata.

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

4.2.3. Sensitivity Analysis

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[2,3], 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.

4.2.3.2. Unstratified Weighted Analysis of PFS

A sensitivity analysis of PFS will be performed using the same analysis procedures described in Section 4.2.2 except that stratification will not be used.

4.2.3.3. Standard "Unweighted" Stratified Analysis of PFS

A sensitivity analysis of PFS by using standard "unweighted" stratified log-rank test and stratified Cox's regression model will be performed using the same stratification factors and stratum size consideration described in Section 4.2.2.

4.3. Key Secondary Endpoints

4.3.1. ORR

4.3.1.1. Definition

Overall response rate (ORR) is defined as the proportion of subjects who achieve PR or better (i.e PR, VGPR, CR or sCR) based on the computerized algorithm, according to the IMWG response criteria [2,3] after the date of randomization but before the start of subsequent anti-myeloma therapy.

4.3.1.2. Analysis Methods

ORR will be calculated for each treatment group based on the ITT analysis set. Subjects with no post-baseline data will be considered as non-responders.

The corresponding 95% Clopper-Pearsons exact CI will be provided. 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 ORR is the same between the two treatment groups. Stratification factors used in the analysis include investigator's choice of PVd or DPd, ISS (I vs. II vs. III) and number of prior lines of therapy (1 vs. 2 or 3).

4.3.2. Response rate of CR/sCR

4.3.2.1. Definition

Response rate of CR or better is defined as the proportion of subjects with a response of a CR or better (CR or sCR) according to IMWG response criteria[2,3], during or after the study treatment but prior to the start of subsequent anti-myeloma therapy.

4.3.2.2. Analysis Methods

The response rate of CR or sCR will be calculated for each treatment group on ITT analysis set. The corresponding 95% Clopper-Pearson exact CI will be provided.

The stratified Cochran Mantel Haenszel (CMH) estimate of odds ratio and its 95% confidence interval and p-value will be used to test if the CR/sCR rate is the same between the two treatment groups. Stratification factors used in the analysis include investigator's choice of PVd or DPd, ISS (I vs. II vs. III) and number of prior lines of therapy (1 vs. 2 or 3).

4.3.3. Overall MRD Negativity Rate

4.3.3.1. Definition

Overall MRD negativity rate, defined as the proportion of subjects who have MRD negativity status (at 10^{-5}) by bone marrow aspirate after the date of randomization and prior to the start of subsequent anti-myeloma therapy. MRD positive subjects include subjects of which all tested samples were found to be MRD positive or ambiguous. The reasons for the missing or unevaluable MRD status will be summarized accordingly.

4.3.3.2. Analysis Methods

For this study, threshold value of 10^{-5} will be used for the primary MRD negativity analysis.

The overall MRD negativity rate will be calculated for each treatment group based on the ITT analysis set. The corresponding 95% Clopper-Pearson exact CI will be provided.

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 overall MRD negativity rate is the same between the two treatment groups. Stratification factors used in the analysis include investigator's choice of PVd or DPd, ISS (I vs. II vs. III) and number of prior lines of therapy (1 vs. 2 or 3).

4.3.4. Overall Survival

4.3.4.1. Definition

Overall survival (OS) is defined as the time from the date of randomization to the date of the subject's death due to any cause. Subjects who are lost to follow-up will be censored at the time of lost to follow-up. Subjects who died after consent withdrawal will be considered as having an OS event. If the subject is alive at the cutoff date for the analysis or the survival status is unknown, then the subject's data will be censored at the date the subject was last known to be 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.

4.3.4.2. Analysis Methods

The analysis will consist of a stratified log-rank test for the comparison of the OS distribution between the 2 treatment groups, stratified by the 3 factors: investigator's choice of PVd vs. PDd, ISS at screening (I vs. II vs. III), and number of prior lines of therapy (1 vs. 2 or 3). The Kaplan-Meier methods will be used to estimate the distribution of overall OS for each treatment. The treatment effect (HR) and its 2-sided 95% CI are to be estimated using a stratified Cox regression model with treatment as the sole explanatory variable. If the number of events is low at the first or second IA for OS then an unstratified analysis may be performed.

5. SAFETY

Safety assessment will be evaluated at the each IDMC review through AEs and clinical laboratory tests, and results of ancillary testing (radiographic studies, cardiac studies, etc.)

All safety analyses will be based on the safety analysis set and presented by the actual treatment received, unless otherwise specified. In addition, selected summaries will be provided for CAR-T treatment-emergent AEs separately.

For all continuous safety variables, descriptive statistics will include the N, mean, standard deviation, median, minimum, and maximum. Categorical variables will be summarized using frequency counts and percentages.

5.1. Adverse Events

Adverse Events will be recorded in standard medical terminology and graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE), Version 5.0, with the exception of (1) Cytokine Release Syndrome (CRS) and Immune Effector Cell-associated Neurotoxicity Syndrome [ICANS], which will be evaluated according to the American Society for Transplantation and Cellular Therapy (ASTCT[1]) consensus grading system, and (2) AE associated with changes in handwriting, which will be graded according to the protocol criteria (Attachment 27 of Protocol Amendment 1). The verbatim terms used in the eCRF by investigators to identify AEs will be coded using the latest Medical Dictionary for Regulatory Activities (MedDRA).

Unless otherwise specified, at each level (e.g., system organ class [SOC] and/or preferred term [PT]) of subject summarization in reporting the incidence of the AE, a subject is counted once if one or more events were recorded.

TEAEs are defined as any AE that occurs at or after the start of study treatment until 30 days after the last treatment for Arm B subjects who discontinued the treatment prior to cilta-cel infusion, or the start of subsequent therapy whichever comes first, and Arm A subjects until 30 days of the last dose of PVd or DPd, or until the start of subsequent therapy, whichever comes first, or until day 112 post-infusion of cilta-cel for subjects in Arm B who received cilta-cel infusion, or the day

prior to start of subsequent anti-myeloma therapy, whichever is earlier; or any AE that is considered related to study drug 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. If the event occurs on the day of the initiation of study treatment and either event time or time of the initiation of study treatment are missing, then the event will be assumed to be treatment emergent. If the event date is recorded as partial or completely missing, then the event will be considered as treatment-emergent unless it is known to be prior to the initiation of study treatment based on partial onset date or resolution date.

Summary tables will be provided for treatment-emergent adverse events:

- An overview of the TEAE, including subjects with TEAE, treatment-emergent serious AE (SAE), TEAE related to study drug, TEAE of maximum grade 1 to 5, TEAE with outcome of death
- TEAE by SOC, PT and toxicity grade 3/4
- Treatment-emergent SAE by SOC and PT
- Toxicity grade 3 or 4 TEAE by SOC, PT, and relationship to study drug
- TEAE with outcome death by SOC, PT and relationship to study drug
- Treatment-emergent infusion related reactions

In addition to the summary tables, listings will be provided for subjects who:

- Had Serious TEAEs
- Had TEAEs leading to discontinuation of study treatment

5.2. Deaths

Number of subjects who died during the study and the primary cause of death will be summarized for the safety analysis set. In addition, all deaths within 30 days after the last dose of PVd or DPd in Arm A and within 112 days after the initial dose of cilta-cel infusion in Arm B will be summarized respectively.

A listing will be generated for safety subjects who died during the study.

5.3. Adverse Events of Special Interest

The Cytokine Release Syndrome and neurologic adverse events will be summarized for Arm B only.

5.3.1. Cytokine Release Syndrome

Subjects with any CRS will be summarized by the maximum toxicity grades (according to ASTCT[1] consensus grading system). In addition, the time from initial cilta-cel infusion to first onset of CRS, the duration of CRS in days, the outcome of CRS, and the treatment for CRS will be summarized as well.

5.3.2. Neurologic Adverse Events

Neurologic adverse events refer to CAR-T cell neurotoxicity (ICANS, other neurotoxicity) and other neurologic adverse events. CAR-T cell neurotoxicity events (ICANS, other neurotoxicity) are adverse events of special interest. Other neurologic adverse events, while not a protocol-defined adverse event of special interest, will be included to allow a comprehensive summary on the neurologic adverse events.

Subjects with any neurologic adverse events reported after cilta-cel infusion will be summarized by MedDRA SOC, HLGT, HLT, PT, and grade 3/4.

Subjects with any ICANS will be summarized by the maximum toxicity grades (according to ASTCT[1] consensus grading system). In addition, the time from initial cilta-cel infusion to the first onset of ICANS, the duration of ICANS in days, the outcome of ICANS, the treatment of ICANS, and concurrent/non-concurrent CRS will be summarized as well. One shift table from baseline to worst toxicity grade during the post infusion period will be provided for ICE scores.

5.3.3. 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, stage of disease, recurrence of a prior existing malignancy (yes, no) and pathology diagnosis (biopsy, aspirate, etc.) information whenever a second primary malignancy is observed. In addition, 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.

5.4. Clinical Laboratory Tests

Selected laboratory analytes are as follows:

- Hematology parameters include the following:
  - Hemoglobin
  - Platelets
  - Absolute lymphocyte counts
  - White blood cell count
  - Absolute neutrophil count
- Biochemistry parameters include the following ([a]Arm B only):
  - AST
  - ALT
  - Alkaline phosphatase
  - Creatinine
  - Total bilirubin
  - Calcium

- Gamma-glutamyltransferase (GGT)[a]
- Ferritin[a]
- C-reactive protein[a]
- Magnesium[a]
- Creatine Phosphokinase[a]
- eGFR

- Coagulation parameters include the following:
  - Prothrombin time/INR
  - Activated partial thromboplastin time
  - Fibrinogen
  - D-dimer

Applicable laboratory results will be graded according to NCI-CTCAE version 5.0. Shift tables from baseline to worst toxicity grade during the post infusion period will be provided for selected laboratory analytes. These tables will summarize the number of subjects with each baseline CTC grade and changes to the maximum CTC grade.

The time to onset and time to recovery of neutropenia and thrombocytopenia will be analyzed based on the post-baseline absolute neutrophil counts and platelets. In addition, the time to recovery for the subjects who were not recovered by the first month will be summarized.

REFERENCE

1. Lee DW, Santomasso BD, Locke Fl, et al. ASTCT consensus grading for cytokine release syndrome and neurological toxicity associated with immune effector cells. Biol Blood Marrow Transplant. 2019; 25 (4):625-638
2. Kumar S, Paiva B, Anderson KC, et al. International Myeloma Working Group consensus criteria for response and minimal residual disease assessment in multiple myeloma. Lancet Oncol. 2016;17(8): e328-e346.
3. Rajkumar SV, Harousseau JL, Durie B, et al. Consensus recommendations for the uniform reporting of clinical trials: report of the International Myeloma Workshop Consensus Panel 1. Blood. 2011;117(18):4691-4695.
4. Tang D and Geller NL. Closed Testing Procedures for Group Sequential Clinical Trials with Multiple Endpoints. Biometrics 55, 1188-1192. December 1999
5. Zucker D.M., Lakatos E. Weighted log rank type statistics for comparing survival curves when there is a time lag in the effectiveness of treatment. Biometrika. 1990;77(4):853–864.
6. Polack, FP, Thomas SJ, Kitchin N, et al, 2020, Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine, NEJM, 383:2603-2615.