Statistical Analysis Plan

NCT Number: NCT03215030

Title: A Phase 1/2 Open-label Study to Investigate the Safety and Tolerability, Efficacy, Pharmacokinetics, and Immunogenicity of Modakafusp Alfa (TAK-573) as a Single Agent in Patients With Relapsed Refractory Multiple Myeloma

Study Number: TAK-573-1501

Document Version and Date: Version 4.0, 15 March 2024

Certain information within this document has been redacted (ie, specific content is masked irreversibly from view) to protect either personally identifiable information or company confidential information.

STATISTICAL ANALYSIS PLAN
STUDY NUMBER: TAK-573-1501
A Phase 1/2a Open-label Study to Investigate the Safety and Tolerability, Efficacy, Pharmacokinetics, and Immunogenicity of Modakafusp Alfa (TAK-573) as a Single Agent in Patients With Relapsed Refractory Multiple Myeloma
PHASE 1/2a

Version**:** Final 4.0

Date**:** 15th March 2024

Prepared by:

[redacted]

Statistics and Quantitative Sciences

Based on:

Protocol Version: Amendment 9

Protocol Date: 02 August 2023

Amendment History:

Date	Version Number	Amendment Type
15 March 2024	Version 4.0	Substantial
30 March 2022	Version 3.0	Substantial
04 December 2020	Version 2.0	Substantial
20 February 2020	Initial SAP	Not applicable

1.1 Approval Signatures

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

Approvers:

[redacted], Ph.D.

[redacted] (Statistical and Quantitative Sciences), Data Science Institute

1.2 Summary of Changes

Summary of Changes from Version 3 to Version 4 of the SAP		
Location	***Description***	***Rationale***
Section 4.3	Part 3 extension primary objective was clarified. Part 3 extension secondary objective was added regarding the PK profile in Chinese patients. Part 3 extension exploratory objective was clarified regarding EQ-5D-5L.	To align with the protocol.
Section 4.4	Text added stating that the sponsor may decide to discontinuation 1 arm if the interim analysis reveals an unfavorable risk-benefit profile.	To align with the protocol.
Section 5.1	Part 1 and Part 2 secondary endpoint was clarified to include NAb.	To align with the protocol.
Section 5.3	Part 3 secondary endpoint added regarding PK parameters in Chinese patients. Part 3 exploratory endpoint regarding EQ-5D-5L was clarified.	To align with the protocol.
Section 5.3.4	Supplementary Estimand removed.	To algin with the removal of sensitivity analyses.
Section 6.0	Added clarification regarding the China continuation cohort.	To align with the protocol.
Section 7.2	Pharmacokinetic analysis set added for Part 3. Per-Protocol Analysis Set removed for Part 3.	To clarify the analysis set used for pharmacokinetic analysis. To algin with the removal of sensitivity analyses.
Section 7.4	Updated refractory definition.	To align with the protocol.
Section 7.7	Study Drug Exposure and Compliance section updated	Algorithm details will be documented in analysis dataset specifications.
Section 7.8.2.1	Clarified programmatic algorithm will be reported at primary analysis. Sensitivity analyses removed.	To align with the protocol.
Section 7.8.2.2	Time to progression endpoint updated to remove death due to PD as an event.	To clarify endpoint definition.
Section 7.8.2.4	Subgroup table moved here (previously under section 7.8).	To provide clarification of subgroups.

	Added two subgroups.	
Section 7.9	Added clarification regarding the PRO endpoints EORTC QLQ-MY20 and EQ-5D-5L.	To align with the protocol and provide additional details for reporting.
Section 7.10	Moved HCRU to a separate section. Clarified the analysis population.	To clarify which analysis set will be used.
Section 7.11.2	Pharmacodynamic analysis was revised to align with protocol.	To align with the protocol.
Section 7.11.3	Clarified NAb will be collected and reported for Parts 1 and 2.	To align with the protocol.
Section 7.12.1	Clarified that IRR signs and symptoms will be reported separately by preferred terms under the AESI tables only. Added summary tabulations grade 3 or higher treatment-emergent AESI and haemorrhage SMQ AEs. Removed the summary tabulation for: - Treatment-emergent AESIs Grade 3 or Grade 4. - Treatment-emergent AESIs Grade 5.	To provide clarification for reporting.
Section 7.14	Moved changes in the previous SAP versions to this section.	Clarification.
Appendix A	Streamlined listings reported.	To provide details for reporting.
Appendix F	Added appendix to detail the visit windowing for PRO data.	To provide additional details for reporting.

2.0 TABLE OF CONTENTS

1.0 TITLE PAGE 1
1.1 Approval Signatures 3
1.2 Summary of Changes 4
2.0 TABLE OF CONTENTS 6
3.0 LIST OF ABBREVIATIONS 8
4.0 OBJECTIVES 10
4.1 Part 1 Dose Escalation Objectives 10
4.1.1 Part 1 Primary Objective 10
4.1.2 Part 1 Secondary Objectives 10
4.1.3 Part 1 Exploratory Objectives 10
4.2 Part 2 Dose Expansion Objectives 10
4.2.1 Part 2 Primary Objectives 10
4.2.2 Part 2 Secondary Objectives 10
4.2.3 Part 2 Exploratory Objective 10
4.3 Part 3 Dose Extension Objectives 11
4.3.1 Part 3 Primary Objective 11
4.3.2 Part 3 Secondary Objectives 11
4.3.3 Part 3 Exploratory Objectives 11
4.4 Study Design 11
5.0 ANALYSIS ENDPOINTS 17
5.1 Part 1 Dose Escalation Endpoints 17
5.1.1 Part 1 Primary Endpoints 17
5.1.2 Part 1 Secondary Endpoints 17
5.1.3 Part 1 Exploratory Endpoints 18
5.2 Part 2 Expansion Endpoints 18
5.2.1 Part 2 Primary Endpoint 18
5.2.2 Part 2 Secondary Endpoints 18
5.2.3 Part 2 Exploratory Endpoints 19
5.3 Part 3 Extension Endpoints 19
5.3.1 Part 3 Primary Endpoint 19
5.3.2 Part 3 Secondary Endpoints 19
5.3.3 Part 3 Exploratory Endpoints 20
5.3.4 Estimands 20
6.0 DETERMINATION OF SAMPLE SIZE 21
7.0 METHODS OF ANALYSIS AND PRESENTATION 22

7.1 General Principles 22
7.1.1 Data Presentations 23
7.1.2 Definition of Study Days 24
7.1.3 Methods for Handling Missing Data 24
7.1.4 Extended Loss-to-Follow-up 24
7.2 Analysis Sets 25
7.3 Disposition of Subjects 26
7.4 Demographic and Other Baseline Characteristics 26
7.5 Medical History and Concurrent Medical Conditions 29
7.6 Medication History and Concomitant Medications 30
7.7 Study Drug Exposure and Compliance 30
7.8 Efficacy Analysis 30
7.8.1 Part 1 and Part 2 Efficacy Analysis 30
7.8.2 Part 3 Efficacy Analysis 34
7.9 Patient-Reported Outcomes (Part 3 only) 43
7.10 Healthcare Resource Utilization Analysis (Part 3 only) 46
7.11 Pharmacokinetic/Pharmacodynamic Analysis 46
7.11.1 Pharmacokinetic Analysis 46
7.11.2 Pharmacodynamic Analysis 48
7.11.3 Immunogenicity Analysis 48
7.12 Safety Analysis 48
7.12.1 Adverse Events 49
7.12.2 Clinical Laboratory Evaluations 51
7.12.3 Vital Signs 52
7.12.4 12-Lead ECGs 52
7.12.5 Other Observations Related to Safety 53
7.13 Interim Analyses and Criteria for Early Termination 53
7.14 Changes in the Statistical Analysis Plan 55
8.0 REFERENCES 57

LIST OF APPENDICES

Appendix A By-Subject Listings: 59
Appendix B Date Imputation Rules 60
Appendix C Algorithms of DLT-like Events 62
Appendix D Operating Characteristics of the Futility Stopping Rules 65
Appendix E Statistical Guidance on Unacceptable Toxicity and Treatment Related Death 67
Appendix F Preparation of Patient-Reported Outcomes Data for Analyses (Part 3 only) 70

3.0 LIST OF ABBREVIATIONS

Abbreviation	Term
ADA	antidrug antibody
AE	adverse event
AESI	adverse event of special interest
ALT	alanine aminotransferase
ANC	absolute neutrophil count
AST	aspartate aminotransferase
AUC	area under the serum concentration-time curve
AUC_{τ}	area under the concentration-time curve during a dosing interval
AUC_{∞}	area under the serum concentration-time curve from time 0 to infinity
AUC_{last}	area under the serum concentration-time curve from time 0 to time of the last quantifiable concentration
BLQ	below the limit of quantitation
CBR	clinical benefit rate
CI	confidence interval
CL	total clearance after administration
C_{max}	maximum observed serum concentration
CR	complete response
DCR	disease control rate
DLT	dose-limiting toxicity
DOR	duration of response
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
EOT	end of treatment
FLC	free light chain
HRQOL	health-related quality of life
ICF	informed consent form
ICU	intensive care unit
IDMC	independent data monitoring committee
Ig	immunoglobulin
IMWG	International Myeloma Working Group
IRC	Independent review committee
IRR	Infusion-related reaction
IV	intravenous(ly)
κD	dissociation constant
LLOQ	lower limit of quantification
LS	least squares
MedDRA	Medical Dictionary for Regulatory Activities
MM	multiple myeloma
MTD	maximum tolerated dose
NAb	neutralizing antibody
NCI CTCAE	National Cancer Institute Common Terminology Criteria for Adverse Events
OBD	optimal biological dose

Abbreviation	Term
ORR	objective response rate
PD	progressive disease; disease progression
PFS	progression-free survival
PoS	probability of success
PI	Proteasome inhibitor
PK	pharmacokinetic(s)
PO	orally
PPOS	posterior probability of success
PR	partial response
QD	quaque die (once daily)
QTcF	QT interval with Fridericia correction method
RBC	red blood cells
MID	minimally important difference
MMRM	mixed model for repeated measures
RNA	ribonucleic acid
SAE	serious adverse event
SAP	statistical analysis plan
sCR	stringent complete response
SD	stable disease
SMQ	standardised MedDRA query
SOC	system organ class
SPEP	serum protein electrophoresis
$t_{1/2}$	terminal disposition phase half-life
TEAE	treatment-emergent adverse event
t_{max}	time of first occurrence of maximum observed serum modakafusp alfa concentration
TTP	Time to progression
TTR	Time to response
ULN	upper limit of normal
UPEP	urine protein electrophoresis
VGRP	very good partial response
V_{ss}	volume of distribution at steady state
λ_z	terminal disposition phase rate constant

4.0 OBJECTIVES

4.1 Part 1 Dose Escalation Objectives

4.1.1 Part 1 Primary Objective

To determine the safety and tolerability of single agent modakafusp alfa in patients with relapsed/refractory MM.

4.1.2 Part 1 Secondary Objectives

- To determine the maximum tolerated dose (MTD)/optimal biological dose (OBD) of modakafusp alfa with 1 or more schedules of administration.
- To characterize the pharmacokinetics (PK) profile of modakafusp alfa.
- To evaluate the immunogenicity of modakafusp alfa.
- To provide a preliminary evaluation of the clinical activity of modakafusp alfa.

4.1.3 Part 1 Exploratory Objectives

To explore potential biomarkers to test their correlation with clinical efficacy and safety parameters.

4.2 Part 2 Dose Expansion Objectives

4.2.1 Part 2 Primary Objectives

To provide a preliminary evaluation of the clinical activity of modakafusp alfa as a single agent and in combination with dexamethasone in patients with relapsed/refractory multiple myeloma (RRMM).

4.2.2 Part 2 Secondary Objectives

- To further evaluate efficacy and safety and to determine the suitability of the modakafusp alfa MTD/OBD as a single agent and in combination with dexamethasone for further trials as the recommended dose and schedule.
- To further characterize the PK profile of modakafusp alfa as a single agent and in combination with dexamethasone.
- To further characterize the immunogenicity of modakafusp alfa as a single agent and in combination with dexamethasone.

4.2.3 Part 2 Exploratory Objective

To explore potential biomarkers to test their correlation with clinical efficacy and safety parameters.

4.3 Part 3 Dose Extension Objectives

4.3.1 Part 3 Primary Objective

To determine the objective response rate (ORR) as assessed by an independent review committee (IRC) according to International Myeloma Working Group (IMWG) criteria (Protocol Amendment 9 Appendix F) of modakafusp alfa in patients who have MM defined by the IMWG criteria with evidence of disease progression and are in need of additional myeloma therapy as determined by the investigator, have previously received at least 3 lines of myeloma therapy, and are refractory to at least 1 IMiD (ie, lenalidomide or pomalidomide [thalidomide excluded]), at least 1 PI (ie, bortezomib, ixazomib, or carfilzomib), and refractory to at least 1 anti-CD38 antibody and who have demonstrated disease progression during or after the last therapy.

4.3.2 Part 3 Secondary Objectives

- To determine ORR by investigator assessment, duration of response (DOR), clinical benefit rate (CBR), duration of clinical benefit, disease control rate (DCR), duration of disease control, progression-free survival (PFS), time to progression (TTP), and overall survival (OS).
- To assess minimal residual disease (MRD) negativity in patients achieving complete response (CR).
- To further characterize safety and tolerability of modakafusp therapy.
- To collect PK data to evaluate population PK and exposure-response (safety/efficacy) analysis.
- To characterize the PK profile of modakafusp alfa in Chinese patients.
- To further characterize the immunogenicity of modakafusp alfa.
- To assess healthcare resource utilization.
- To evaluate patient-reported disease symptoms (includes bone aches or pain, back pain, hip pain, arm or shoulder pain, chest pain, and pain increasing with activity).

4.3.3 Part 3 Exploratory Objectives

- To explore predictive biomarkers of response and resistance.
- To characterize MRD negativity at a higher sensitivity in patients achieving CR.
- To evaluate generic health-related quality of life (HRQOL)/health status, measured by the EuroQoL-5-Dimensions-5 Levels (EQ-5D-5L) instrument.

4.4 Study Design

This is a multicenter, open-label, phase 1/2 study designed to determine the safety and tolerability of modakafusp alfa as a single agent and in combination with dexamethasone in patients with relapsed/refractory MM. The study will be conducted in 3 parts.

Part 1 Dose Escalation: Single-agent modakafusp alfa with 1 or more schedules of administration.

Part 2 Dose Expansion: At least 1 cohort of modakafusp alfa combined with dexamethasone using the dose and schedule determined in Part 1.

Treatment cycle duration is 28 days (21 days for Schedule C). For Parts 1 and 2, modakafusp alfa with or without dexamethasone will be administered for up to 12 cycles or until disease progression, unacceptable toxicity, or any other discontinuation criterion is met. Patients with demonstrated clinical benefit may continue treatment beyond 12 cycles with the agreement of the sponsor.

Part 3 Dose Extension: Single-agent modakafusp alfa with 1:1 randomization to receive modakafusp alfa 120 mg or 240 mg. Patients will be stratified by their cytogenetics risk (high risk [(del17, t(4;14) and/or t(14;16)] vs standard risk) and myeloma type (IgA vs other). Cytogenetic results from samples taken within 5 weeks prior to first dose are acceptable for stratification. All patients should also have a sample from screening sent for central analysis. If a previous result is not available and the patient is known to have high-risk disease [i.e., del17, t(4;14) and/or t(14;16)] from prior cytogenetic testing, regardless of timing, they should be stratified as high risk for the purpose of enrollment.

Response assessments in Part 3 will be made on the basis of central laboratory data.

Part 1 Dose Escalation
The primary objective of Part 1 of the study is to determine the safety and tolerability of one or more schedules of single-agent modakafusp alfa in patients with RRMM.

Part 1 will follow a 3 + 3 dose escalation design to evaluate up to 4 different schedules of administration of modakafusp alfa with the objective of determining, with one or more schedules, either an MTD based on the observation of dose-limiting toxicities (DLTs), or an OBD.

An OBD is a dose level that is below or coincides with the MTD for which there is evidence of anti-myeloma activity plus data supporting significant pharmacodynamic effects in one or more biomarkers. An OBD for one or more schedules can be selected before or after identifying MTD. This MTD/OBD will be further refined based on the safety, PK and pharmacodynamic findings in Part 2 to produce a recommended dose and schedule that may be taken into future efficacy studies. The dose-escalation design is described in the Protocol Section **Error! Reference**

source not found.. The safety and tolerability of modakafusp alfa will be assessed by recording treatment-emergent adverse events (TEAEs), dose modifications, treatment discontinuations, vital signs, physical examination, chemistry and hematology, urinalysis, electrocardiograms (ECGs), and concomitant medications according to National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE; Version 5.0). DLTs are defined in Protocol Section **Error! Reference source not found.**. The most favorable dose levels for Schedules B (once every 2 weeks [Q2W]), C (once every 3 weeks [Q3W]), and D (Q4W) based on safety, activity and other parameters will be selected for further evaluation in Part 2 as a single agent and in combination with dexamethasone.

Up to 4 single agent schedules were to be evaluated during Part 1:

- Schedule A (initial): Modakafusp alfa infusion is given at weekly intervals in a 28-day cycle for 2 cycles followed by 1 infusion every 2 weeks (twice per cycle) in Cycles 3 to 6, and once every 4 weeks from Cycle 7 until treatment discontinuation.
  Schedule A is the most intensive in terms of administration frequency. Three additional less-intensive schedules are added to introduce longer recovery times. The aim is to identify the optimal schedule in terms of tolerability and dose maximization. The investigators and sponsor will decide which schedule will be tested based on the available safety, efficacy, PK and pharmacodynamic information. The decision will be documented in writing.
- Schedule B: Modakafusp alfa infusions will be administered on a 28-day (4-week) cycle once every 2 weeks until treatment discontinuation criterion is met.
- Schedule C: Modakafusp alfa infusions will be administered on a 21-day (3-week) cycle once every 3 weeks until treatment discontinuation criterion is met.
- Schedule D: Modakafusp alfa infusions will be administered on a 28-day (4-week) cycle once every 4 weeks until treatment discontinuation criterion is met.

During the dose escalation phase and after the first cycle, patients who have not experienced a DLT, have not shown signs of progressive disease (PD), and in the opinion of the investigator would continue to benefit from additional modakafusp alfa, may receive additional cycles of treatment. A window of ±2 days will be permitted for the weekly and every-2-week doses, and a window of ±4 days will be permitted for the every 3-week and every 4-week doses.

Part 2 Dose Expansion

The primary objective of Part 2 of the study is to provide a preliminary evaluation of the clinical activity of one or more schedules of modakafusp alfa given as a single agent in patients with RRMM. Additional cohort(s) of modakafusp alfa (using the same dose used in a single-agent cohort) in combination with dexamethasone will be implemented, with the selected schedule(s)

to be determined once the modakafusp alfa MTD or OBD is established. Treatment allocation will be assigned based on patient availability and eligibility. Dexamethasone will be given as a once weekly oral dose of 40 mg. Patients over 75 years of age will receive a reduced dose of dexamethasone (20 mg, same schedule). The aim of this approach is to obtain preliminary information on the effect of standard doses of dexamethasone on modakafusp alfa safety, efficacy, and pharmacodynamic endpoints.

Preliminary activity of modakafusp alfa as a single agent and in combination with dexamethasone will be evaluated by measuring the confirmed ORR (partial response [PR] or better) according to IMWG Uniform Response Criteria (Protocol Appendix F). The efficacy of modakafusp alfa will further assessed by assessing DOR, PFS, OS, and time to response.

Part 3 Dose Extension

Part 3 will determine the confirmed ORR per IRC assessment, evaluate other efficacy endpoints, assess MRD negativity in patients achieving CR, further characterize the safety, tolerability, and immunogenicity of modakafusp alfa therapy, collect PK data, and assess healthcare resource utilization.

The modakafusp alfa doses of 1.5 mg/kg and 3 mg/kg Q4W will be translated into 2 fixed doses of 120 mg and 240 mg based on the median bodyweight of approximately 80 kg (from Parts 1 and 2 of Study TAK-573-1501) for the Part 3 evaluation.

Two interim analyses for futility are planned for the study when approximately 15 and 48 patients of the planned 118 patients per arm are enrolled and treated for at least 3 cycles or have discontinued treatment prematurely (Figure 1).

The study will only proceed with enrolling additional patients into a treatment arm where the futility boundary is not crossed; more specifically:

Scenario 1: If the futility boundary is not crossed in one arm (e.g., the 120mg arm) but is crossed in the second treatment arm (e.g., the 240mg arm), patient enrollment will continue only in the arm that did not cross the futility boundary (e.g., the 120mg arm).

Scenario 2: If the futility boundary is not crossed in the 120mg treatment arm and is not crossed in the 240mg treatment arm, patient enrollment will continue for both arms to identify the optimal dose, defined as having the more favorable risk-benefit profile based on the totality of data from both arms and Parts 1 and 2.

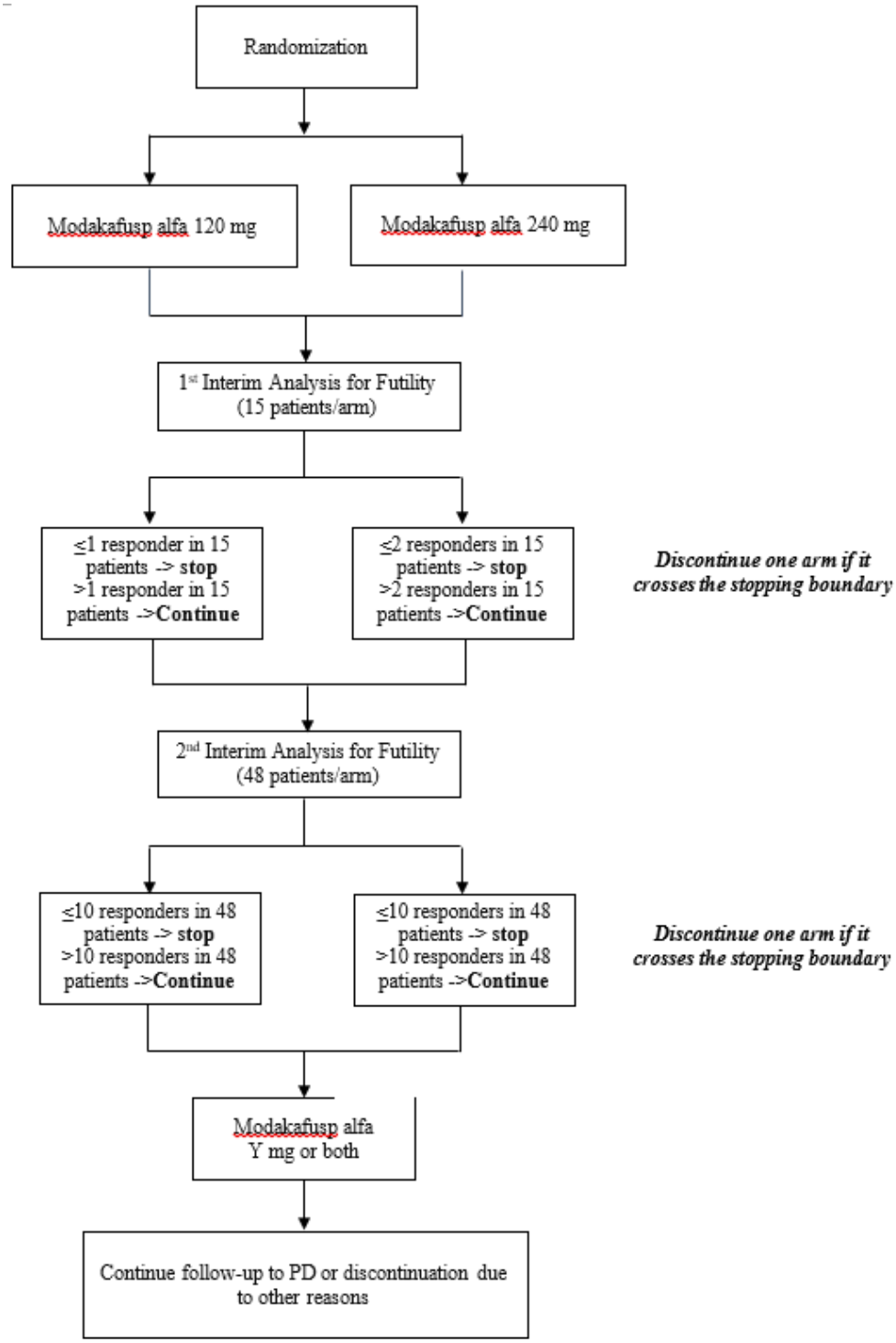
Scenario 3: If the futility boundary is crossed in the 120mg treatment arm and is crossed in the 240mg treatment arm, the study will be terminated.

In addition, the sponsor may decide to discontinue 1 arm if there is early detection of an unfavorable risk-benefit profile based on the interim data analysis data.

Part 3 China Continuation Cohort

After global patient enrollment is completed in Part 3, enrollment will continue for the China continuation cohort until about 15% of the total sample size is reached at the selected dose(s) (eg, 18 patients based on 118 patients treated at the selected dose[s]).

Figure 1 Part 3 Study Schematic



PD: progressive disease.
Dosing is once every 4 weeks in both cohorts.

5.0 ANALYSIS ENDPOINTS

5.1 Part 1 Dose Escalation Endpoints

5.1.1 Part 1 Primary Endpoints

- The number of patients with TEAEs overall and per dose level
- Patients with DLTs at each dose level
- Patients with Grade ≥3 TEAEs
- Patients with serious adverse events (SAEs)
- Patients who discontinue because of TEAEs
- Patients with dose modifications (delays, interruptions, dose reductions)
- Clinically significant laboratory values
- Clinically significant vital sign measurements.

5.1.2 Part 1 Secondary Endpoints

- DLT-like (TEAEs meeting DLT definition that occur after phase 1 Cycle 1) frequencies and other TEAEs occurring over the course of extended treatment with modakafusp alfa, including information about dose modification, treatment discontinuation, and clinically significant laboratory values and vital signs
- Summary statistics by dose level and cycle day for the following PK parameters:
  - Single-dose maximum observed concentration (C_{max}).
  - Time of first occurrence of C_{max} (t_{max}).
  - Area under the serum concentration-time curve from time 0 to infinity (AUC_{∞}).
  - Area under the serum concentration-time curve from time 0 to time of the last quantifiable concentration (AUC_{last}).
  - Apparent serum modakafusp alfa terminal disposition rate constant (λ_z).
  - Apparent serum modakafusp alfa terminal elimination phase half-life ($t_{1/2z}$).
  - Total clearance after administration (CL).
  - Volume of distribution at steady state (V_{ss}).
- Anti–modakafusp alfa antibody incidence and characteristics (eg, titer and specificity), neutralizing antibody (NAb).
- Preliminary evaluation of antitumor activity of modakafusp alfa:
  - For patients with measurable disease only: ORR, defined as the proportion of patients who achieved a PR or better during the study: stringent complete response (sCR), complete response (CR), very good partial response (VGPR), and PR as defined by IMWG Uniform Response Criteria; CBR (includes patients with a response of sCR, CR, VGPR, PR, or minimal response [MR]); DCR (includes patients with a response of sCR, CR, VGPR, PR, MR, or stable disease [SD]).

- For patients with measurable disease only: median DOR, with DOR defined as the time from the date of first documentation of response PR or better to the time of disease progression or death, whichever occurs first.
- For patients with measurable disease only: time to response, defined as the time from first dose to the date of first documentation of response (PR or better).
- PFS, defined as the time from the date of first dose until the sooner of the date of PD, defined by IMWG criteria, or the date of death due to any cause.

5.1.3 Part 1 Exploratory Endpoints

- CD38 expression on MM cells and other immune cells in bone marrow aspirate (BMA) and its correlation with clinical outcome.
- Pharmacodynamic biomarkers including, but not limited to, neopterin, complement, cytokines/chemokines, and gene expression and their correlation with clinical outcome.
- Pharmacodynamic analysis of the presence and changes of immune cells from whole blood and bone marrow and their correlation with clinical outcome.
- Effects of modakafusp alfa administration on the electrocardiographic QT/QT interval with Fridericia correction method (QTcF).

5.2 Part 2 Expansion Endpoints

5.2.1 Part 2 Primary Endpoint

- ORR, defined as the proportion of patients who achieved a PR or better during study; sCR, CR, VGPR, and PR as defined by IMWG Uniform Response Criteria.

5.2.2 Part 2 Secondary Endpoints

- DOR
- CBR
- DCR
- PFS, defined as the time from the date of first dose until the sooner of the date of PD, defined by IMWG criteria, or the date of death due to any cause
- OS, defined as the date from the first dose to the date of death due to any cause
- Time to response, defined as the time from the first dose to the date of first documentation of response (PR or better)
- DLT-like events (TEAEs meeting DLT definition that occur after Part 1 Cycle 1) frequencies and other TEAEs occurring over the course of extended treatment with modakafusp alfa, including information about dose modification, treatment discontinuation, and clinically significant laboratory values and vital signs.

- Summary statistics by dose level and cycle day for the following PK parameters: C_{max}, AUC_{∞}, AUC_{last}, λ_z, t_{max}, CL, V_{ss}, and $t_{1/2z}$.
- Anti–modakafusp alfa antibody incidence and characteristics (eg, titer and specificity), NAb.

5.2.3 Part 2 Exploratory Endpoints

- CD38 expression on MM cells and other immune cells in BMA and its correlation with clinical outcome.
- Pharmacodynamic biomarkers including, but not limited to, neopterin, cytokines/chemokines, and gene expression and their correlation with clinical outcome.
- Pharmacodynamic analysis of the presence and changes of immune cells from whole blood and bone marrow and their correlation with clinical outcome
- Effects of modakafusp alfa administration on the electrocardiographic QT/QT interval with Fridericia correction method (QTcF)

5.3 Part 3 Extension Endpoints

5.3.1 Part 3 Primary Endpoint

- ORR (PR or better) assessed by IRC according to modified IMWG criteria

5.3.2 Part 3 Secondary Endpoints

- ORR by investigator assessment
- DOR by IRC and investigator assessment
- CBR by IRC and investigator assessment.
- Duration of clinical benefit
- DCR by IRC and investigator assessment
- Duration of disease control
- PFS (time from the date of first dose until the sooner of the date of PD, defined by IMWG criteria or the date of death due to any cause) by IRC and investigator assessment
- TTP by IRC and investigator assessment
- OS (date from first dose to the date of death due to any cause)
- Rate of MRD negative status at a sensitivity of 10^{-5} in patients achieving CR
- Duration of MRD negativity at a sensitivity of 10^{-5} in patients achieving CR
- AEs, SAEs, laboratory assessments, supportive care use
- Eastern Cooperative Oncology Group (ECOG) status
- ADA incidence and characteristics (eg, titer and specificity) and NAb
- Length of hospital stay, types of hospital stay, and other healthcare utilization data

- No worsening of disease symptoms (includes bone aches or pain, back pain, hip pain, arm or shoulder pain, chest pain, and pain increasing with activity) from baseline at 12 weeks across doses of modakafusp alfa as measured by the patient-reported outcome (PRO) instrument European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Multiple Myeloma Module (EORTC QLQ-MY20).
- Summary statistics by dose level and cycle day for the following PK parameters in Chinese patients with intensive PK schedule: C_{max}, AUC_{∞}, AUC_{last}, λ_z, t_{max}, CL, V_{ss}, and $t_{1/2z}$.

5.3.3 Part 3 Exploratory Endpoints

- CD38 expression on MM cells and other immune cells within the bone marrow and its correlation with clinical outcome.
- Evaluation of somatic mutations and polymorphisms and their associations with response and/or acquired resistance.
- Evaluation of anti-interferon antibodies at baseline and correlation with response.
- Rate of MRD negative status at a sensitivity of 10^{-6} in patients achieving CR.
- Duration of MRD negativity (10^{-6}) in patients who achieve MRD negative status at CR.
- Evaluation of generic HRQOL/health status, measured by the EQ-5D-5L instrument.

5.3.4 Estimands

Estimand: ORR in Part 3	**Primary Estimand**
Definition	The treatment effect of modakafusp alfa in the full analysis set
Treatment	Modakafusp alfa 120mg and 240mg
Population	Full analysis set: including all patients who receive at least 1 dose of the study drug even an incomplete dose, of modakafusp alfa, in the Part 3 extension cohorts.
Variable (or Endpoint)	ORR, defined as the proportion of patients who achieved a confirmed sCR, CR, VGPR, or PR during the study as defined by IMWG Uniform Response Criteria
Strategy for Addressing Intercurrent Event	Hypothetical strategy for subsequent anti-cancer therapy (intercurrent event): The disease assessments after subsequent therapy will be censored. Treatment policy strategy for treatment discontinuation (intercurrent event): The disease assessment after treatment discontinuation will be included for evaluation
Population-Level Summary	Observed ORR with 95% CI

6.0 DETERMINATION OF SAMPLE SIZE

Part 1 of the study will follow a standard 3+3 dose escalation design. Under this design, the probability of dose escalation according to the true DLT rate is as follows:

True DLT rate	5%	10%	20%	30%	40%	50%	60%	70%
Probability of escalating	97%	91%	71%	49%	31%	17%	8%	3%

DLT: dose-limiting toxicity.

The above probability is calculated by summing up the probability of the two below events that leads to dose-escalation:

a) no DLT in the first cohort of 3 patients at a dose level, or
b) One DLT in the first cohort of 3 patients and no DLT in the second cohort of 3 patients at a dose level.

The probability of not observing a DLT in a sample size of 6 patients according to the true DLT rate is as follows:

True DLT rate	5%	10%	20%	30%	40%	50%	60%	70%
Probability of no toxicity	74%	53%	26%	12%	4.7%	1.6%	0.4%	<0.1%

DLT: dose-limiting toxicity.

Patients will be considered evaluable in phase 1 for MTD/OBD assessment provided that they have not missed any of their infusions of modakafusp alfa in Cycle 1 or have had a DLT. Patients who are not considered evaluable as defined above will be replaced.

Part 2 Expansion

The number of patients in the dose expansion phase was determined based on a binomial exact test design. A sample size of 25 patients will allow a cohort to have over 80% power to rule out an uninteresting response rate of 10% if the true response rate is 30% with a one-sided alpha of 0.05. A Bayesian predictive probability design is used to allow multiple interim analyses to stop early for futility. The study will enroll up to 25 patients in each cohort to further characterize safety of the OBD/MTD and obtain preliminary clinical efficacy data. With a vague beta prior with shape parameters 0.428 and 1 for ORR (i.e., mean of 30%), 11 responders out of 25 gives a greater than 90% posterior probability of true ORR being greater than 30%. An OBD/MTD dose will be considered clinically relevant if 11 out of 25 response evaluable patients have a confirmed PR or better. The design has reasonable operating characteristics. With a true ORR of 30%, the probability of declaring success is 7%. In contrast, with a true ORR of 50%, the probability of declaring success is 70%.

See Section 7.13 for details of futility analyses after 10 and 15 evaluable patients are enrolled in the expansion phase.

Part 3 Extension

The primary objective of Part 3 of the study is to determine the confirmed ORR, including PR, VGPR, CR, and sCR assessed according to the IMWG criteria to modakafusp alfa in patients with MM previously treated with ≥3 lines and with evidence of disease progression who are refractory to, or intolerant of, a PI, an IMiD, and an anti-CD38 mAb (triple refractory).The sample size was determined to claim that the true ORR is greater than the threshold response rate of 20% within each treatment arm. A total sample size of approximately 236 enrolled patients (118 patients per arm) will allow the study to have over 90% power to rule out an uninteresting response rate of 20% if the true rate is 35% with a 1-sided alpha of 0.025.

The purpose of the China continuation cohort is to allow continued evaluation of efficacy and any emerging safety signals in Chinese patients enrolled in China. After global patient enrollment is completed in Part 3, enrollment will continue for the China continuation cohort until about 15% of the total sample size is reached at the selected dose(s) (eg, 18 patients based on 118 patients treated at the selected dose[s]).

7.0 METHODS OF ANALYSIS AND PRESENTATION

For Part 3, two interim analyses for futility are planned when approximately 15 and 48 patients of the planned 118 patients per arm are enrolled and treated for at least 3 cycles or discontinued treatment prematurely. Details of the interim analyses can be referred to Section 7.13.

The planned primary analysis will be performed once the study reaches 6 months after the last patient is randomized and treated. An updated analysis for PFS, DOR, and OS will be performed at the end of study.

7.1 General Principles

All statistical analyses will be conducted using SAS® Version 9.4 or later.

For categorical variables, the count and proportions of each possible category value will be tabulated. The denominator for the proportion will be based on the number of subjects in the analysis sets. For continuous variables, the mean, median, standard deviation, minimum, and maximum values will be presented for all patients in the relevant analysis set.

Means and medians will be presented to 1 more decimal place than the recorded data. The standard deviations will be presented to 2 more decimal places than the recorded data.

Confidence intervals about a parameter estimate will be presented using the same number of decimal places as the parameter estimate.

All confidence intervals, statistical tests, and resulting P-values will be reported as 2-sided. P-values should be presented to 3 decimal places, with values less than 0.001 presented as <0.001.

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

7.1.1 Data Presentations

Part 1 and Part 2

1. **Baseline evaluations/disposition/medical history/concurrent medical condition/concomitant medication/medication history/prior therapy:**
   Data will be presented by schedule within dose escalation, dose expansion, and in overall study population.

Dose escalation

Schedule A	Schedule B	Schedule C	Schedule D	Total

Dose expansion

Schedule A MTD/OBD	Schedule B MTD/OBD	Schedule C MTD/OBD	Schedule D MTD/OBD	Schedule X + Dex MTD/OBD	Total

Study overall

Dose Escalation Total	Dose Expansion Total	Total

2. **Safety/Efficacy/Pharmacokinetic/Pharmacodynamic**
   Data will be presented by dose level within dose escalation and dose expansion, and in overall study population.

Dose escalation: by dose level within each schedule
Schedule A (repeat for schedule B, C, and D):

Dose Level 1	Dose Level 2	Dose Level 3	Continue for each dose level	Total

Dose expansion:
Present for each schedule MTD/OBD dose level explored in the expansion.

Schedule: TBD Dose Level mg/kg (+ Dexamethasone)

Schedule A MTD/OBD	Schedule B MTD/OBD	Schedule C MTD/OBD	Schedule D MTD/OBD	Schedule X MTD/OBD + Dex

Study overall

Dose Escalation Total	Dose Expansion Total	Total

Part 3

1. Baseline evaluations/disposition/medical history/concurrent medical condition/concomitant medication/medication history/prior therapy/safety:

Data will be presented by dose level and the overall study population.

120mg	240mg	Total

2. Efficacy/Pharmacokinetic/Pharmacodynamic

Data will be presented by dose level.

120mg	240mg

7.1.2 Definition of Study Days

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

7.1.3 Methods for Handling Missing Data

For efficacy and safety data, no imputation of values for missing data will be performed. Imputation rules for incomplete dates are described in Appendix B.

7.1.4 Extended Loss-to-Follow-up

For subjects, if two or more scheduled disease assessments are missed and are then followed by an assessment of PD or death, PFS, DOR, and TTP will be censored at the last adequate assessment prior to PD or death. If the scheduled disease assessment is every 4 weeks (Schedule A, B, and D), a window of 64 days will be used to determine whether there was an extended time without adequate assessment. That is, if the time difference between PD/death and last adequate assessment is more than 64 days, then PFS, DOR and TTP will be censored at the last adequate assessment prior to PD/death. If the scheduled disease assessment is every 3 weeks (Schedule C), a window of 50 days will be used to determine whether there was an extended time without adequate assessment.

7.2 Analysis Sets

Part 1 and Part 2

- Safety Analysis Set: The safety analysis set will include all enrolled patients who receive at least 1 dose of the study drug even an incomplete dose, of modakafusp alfa or Dexamethasone. The safety population will be used for safety related analyses such as AE, concomitant medications, laboratory tests, and vital signs. In addition, it will also be used for the efficacy analyses.
- Response-Evaluable Analysis Set: The subset of the safety analysis set of patients treated at the MTD/OBD from the Part 2 dose expansion cohorts with measurable disease at baseline and at least 1 post-treatment efficacy evaluation. The response-evaluable analysis set will be used in the futility analysis of the Part 2 expansion.
- DLT Evaluable Analysis Set: Patients who receive all Cycle 1 doses of modakafusp alfa or experience a DLT in Cycle 1 in the Part 1 dose escalation cohorts.
- Pharmacokinetic (PK) Analysis Set: The PK analysis set includes patients from the safety analysis set who have sufficient data to report at least 1 PK concentration for modakafusp alfa. The PK analysis set will be used for PK analysis.
- Immunogenicity-evaluable Analysis Set: The immunogenicity-evaluable analysis set will include patients from the safety analysis set with a baseline assessment and at least 1 post-baseline immunogenicity assessment. The immunogenicity-evaluable analysis set will be used for immunogenicity analysis.

Part 3

- Full Analysis Set (Safety Analysis Set): The full analysis set will include all enrolled patients who receive at least 1 dose of the study drug even an incomplete dose, of modakafusp alfa. The full analysis set will be used for baseline assessment, efficacy analysis and PRO compliance. The safety analysis set will be used for safety analysis.
- Intent-to-treat (ITT) Analysis Set: The ITT analysis set will include all randomized patients regardless of whether they receive study drug or adhere to the assigned dose. The ITT analysis set will be used for listings.
- Pharmacokinetic (PK) Analysis Set: The PK analysis set includes patients from the safety analysis set who have sufficient data to report at least 1 PK concentration for modakafusp alfa. The PK analysis set will be used for PK analysis.
- Immunogenicity-evaluable Analysis Set: The immunogenicity-evaluable analysis set will include patients with a baseline assessment and at least 1 post-baseline immunogenicity assessment. The immunogenicity-evaluable analysis set will be used for immunogenicity analysis.

- PRO Analysis Set: The PRO analysis set includes all patients with a baseline and at least one post-baseline measurement of any PRO. The analyses of PROs will be based on the PRO analysis set.

7.3 Disposition of Subjects

Study information, including the date first subject signed the informed consent form (ICF), date of last subject's last visit/contact, date of last subject's last procedure for collection of data for primary endpoint, MedDRA Version, WHO Drug Version, and SAS Version will be presented. The reasons for screen failures will be generated in a summary table.

The disposition includes the number and percentage of patients in the following categories: patients in each of the study populations, primary reason off treatment, ongoing (if applicable at the time of database lock/data cut-off), participating in PFS follow-up, and primary reason discontinued from the study. This information will be presented by safety analysis set in Part 1 escalation cohorts and Part 2 expansion cohorts and by full analysis set in Part 3 extension cohorts.

7.4 Demographic and Other Baseline Characteristics

All demographics and baseline characteristics will be summarized by safety analysis set in Part 1 escalation cohorts and Part 2 expansion cohorts and by full analysis set in Part 3 extension cohorts.

Demographics

Demographic data to be evaluated will include age, sex, race, ethnicity, height, and weight.

Stratification factors (Part 3 only)

- Type of Myeloma
  - IgA
  - Other
- Cytogenetic
  - High risk [(del17, t(4;14) and/or t(14;16)]
  - Standard risk

Disease characteristics:

- Years since initial diagnosis [(date of first dose – date of diagnosis) / 365.25)].
- Type of Myeloma
  - IgG, IgA, IgD, IgM, Biclonal (if known, otherwise summarize with Light chain below).

  - Light Chain Kappa, Light Chain Lambda, Light chain Biclonal, Light Chain unknown.
- Measurable disease
  - Measurable by both SPEP (baseline serum M-protein ≥0.5 g/dL) and UPEP (baseline urine M-protein ≥200 mg/24 hr)
  - Measurable by SPEP only
  - Measurable by UPEP only
  - Measurable by FLC only (met all three criteria below):
    1. Involved either Kappa or Lambda free light chain level at baseline ≥10 mg/dL;
    2. and FLC ratio (Kappa FLC/Lambda FLC) is abnormal (serum FLC ratio >1.65 or <0.26)
    3. and not measurable by SPEP or UPEP
  - Non-measurable (not measurable by SPEP, UPEP, or FLC)
- Durie-Salmon Stage.
- International Staging System (ISS) at initial diagnosis.
- International Staging System (ISS) at study entry.
- Evidence of Lytic Bone Disease at initial diagnosis.
- Evidence of Extramedullary Disease at initial diagnosis.
- Eastern Cooperative Oncology Group (ECOG) performance status.

Baseline disease characteristics– laboratory values

- β_2-microglobulin category (<2.5, 2.5-5.5, >5.5 mg/L).
- Serum albumin category (<3.5, ≥3.5 mg/dL).
- Serum creatinine category (≤2, >2 mg/dL).
- Calculated creatinine clearance category (<30, 30 - <60, 60 - <90, ≥90 mL/min).
  Creatinine clearance will be calculated using the Cockcroft-Gault formulas as follows:

  For male patients:

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

  For female patients multiply by 0.85.
- Hemoglobin (g/L).
- Platelet count (10^9/L).
- Absolute neutrophil count (10^9/L).
- Lactate dehydrogenase (LDH) > Upper limit of normal (ULN).

Baseline Bone Marrow Evaluation

- Baseline bone marrow aspirate/biopsy performed and adequate for interpretation.
- Percent plasma cells in bone marrow (if both biopsy and aspirate were performed use larger value; unable to detect is considered as zero).

Cytogenetics

- Cytogenetic testing performed.

The number and percentage of patients with each type of chromosomal aberration or abnormality will be presented including but not limited to the following:

- high risk cytogenetics [del 17, t(4;14) or t(14;16)].
- expanded high risk cytogenetics [del 17, t(4;14), (14;16), or 1q21 amplification/gain].
- del 13.
- del 17.
- t(4:14).
- t(11:14).
- t(14:16).
- t(14:20).
- 1q21 amplification/gain.
- hyperdiploidy.
- hypodiploidy.

(percentages are based on the number of patients with cytogenetic testing performed)

Extramedullary Disease Assessment (imaging)

- Imaging lytic lesions (yes, no, indeterminate).
- Imaging extramedullary plasmacytomas (yes, no indeterminate).
- Location of plasmacytomas (Liver, Visceral; Lung, Visceral; Node; Soft Tissue; Other).
- Number of plasmacytomas (1, 2, ≥3).
- Skeletal survey results (normal, abnormal not clinically significant, abnormal clinically significant, and not done).
- Skeletal survey lytic lesions (yes, no, indeterminate).

Prior therapy

- Prior radiation (yes).
- Prior surgery (yes).
- Prior bone marrow transplant or stem cell transplant.
  - Type of stem cell transplant (autologous, allogeneic, syngeneic, unknown).

- Prior systemic therapy (yes).
- Lines of prior therapy (descriptive statistics and categorical summary: 1, 2, 3, 4, 5, 6, >6).
- Lines of prior therapy initiated within 12 months of first dose of modakafusp alfa (categorical summary).
- Type of prior therapy may include but not limited to:
  - Anti-CD38 monoclonal antibody (Daratumumab, Isatuximab).
  - CAR-T.
  - Anti-BCMA.
  - IMiD.
  - Proteasome inhibitor (PI).
  - Elotuzumab.
- Type of last line of prior therapy.
- Best response to any prior anti-CD38.
- Best response to last line of prior therapy
- Refractory to prior therapy may include but not limited to:
  - Refractory to anti-CD38.
  - Refractory to IMiD.
  - Refractory to PI.
  - Refractory to IMiD and PI.
  - Refractory to IMiD, PI, and anti-CD38.
  - Refractory to last line of prior therapy.

Refractory to therapy is defined as: <25% reduction in M-protein or progression of disease during treatment or within 60 days after cessation of treatment.

7.5 Medical History and Concurrent Medical Conditions

Medical history and concurrent medical conditions will be coded using the latest MedDRA dictionary. Concurrent medical conditions are the ones ongoing or started on or after the day informed consent was signed. Medical history and concurrent medical conditions will be summarized separately by safety analysis set for Part 1 escalation cohorts and Part 2 expansion cohorts and by full analysis set for Part 3 extension cohorts.

If both start date and stop date are missing, medical condition will be assumed to start before informed consent and continue after treatment discontinuation. If only start date is missing, then medical condition will be assumed to start before informed consent. If stop date is missing, then medical condition will be assumed to continue after treatment discontinuation.

7.6 Medication History and Concomitant Medications

Medication history and concomitant medications will be coded using the latest World Health Organization (WHO) Drug Dictionary. Medication history is defined as the medication stopped before the first dosing date of the study drug. Concomitant medication is defined as the medication ongoing or started on or after the first dosing date of the study drug. Medication history and concurrent medications will be summarized separately by safety analysis set for Part 1 escalation cohorts and Part 2 expansion cohorts and by full analysis set for Part 3 extension cohorts. The number and percentage of patients taking concomitant medications will be tabulated by WHO standardized medication name.

If both start date and stop date are missing, medication will be assumed to start before the first dosing date of the study drug and continue after treatment discontinuation. If only start date is missing, then medication will be assumed to start before the first dosing date of the study drug. If only stop date is missing, then medication will be assumed to continue after treatment discontinuation.

7.7 Study Drug Exposure and Compliance

Summaries and descriptive statistics of duration of treatment with modakafusp alfa (in weeks for Parts 1 and 2, and months for Part 3), number of treated cycles (1, 2, 3, 4, 5, 6, 7-12, >12 and summary statistics), cumulative dose, planned cumulative dose (Part 3 only), actual and planned dose per cycle, and relative dose intensity will be summarized by safety analysis set in Part 1 escalation cohorts and Part 2 expansion cohorts and by full analysis set for Part 3 extension cohorts. A treated cycle is defined as a cycle in which any amount of modakafusp alfa for at least one of the dosing days in the cycle. Dose unit will be mg/kg for Part 1/Part 2 and mg for Part 3.

The algorithm details will be documented in the analysis dataset specifications.

Action on Study Drug

The reason for dose modification (e.g. dose increased, dose reduced, interrupted, withdrawn, delayed, rate increased, rate reduced, drug infusion interrupted) of modakafusp alfa will be summarized by cycle (Cycles 1-8), greater than 8 cycles and overall.

7.8 Efficacy Analysis

7.8.1 Part 1 and Part 2 Efficacy Analysis

Response evaluations after the start of subsequent anti-cancer therapy will not be considered in the calculation of efficacy endpoints, while the response evaluation after the treatment discontinuation will be considered in the calculation of efficacy endpoints. Please refer to Section 7.9.2 Part 3 Efficacy Analysis for details. When appropriate, data from patients in the

expansion cohorts will be summarized together with data from patients in the dose escalation phase.

Objective Response Rate (ORR)

The ORR is defined as the proportion of patients who achieved a confirmed PR or better during the study per investigator assessment as defined by IMWG Uniform Response Criteria. The ORR will be summarized by frequencies, percentages, and the exact two-sided 95% CIs based on the safety analysis set.

Clinical Benefit Rate (CBR)

The CBR includes patients with a confirmed response of sCR, CR, VGPR, PR, or MR during the study per investigator assessment as defined by IMWG Uniform Response Criteria. The CBR will be summarized by frequencies, percentages, and the exact two-sided 95% CIs based on the safety analysis set.

Disease Control Rate (DCR)

The DCR includes patients with a confirmed response sCR, CR, VGPR, PR, MR, or SD during the study per investigator assessment as defined by IMWG Uniform Response Criteria. The DCR will be summarized by frequencies, percentages, and the exact two-sided 95% CIs based on the safety analysis set.

Duration of Response (DOR)

The DOR will be calculated for those patients with a confirmed PR or better in the safety analysis set. The DOR is defined as the time from the first documentation of a confirmed response until progressive disease or death, whichever comes first. If there is no progressive disease or death observed, it will be censored at the last adequate response assessment. Please refer to Table 1 for the detailed censoring rules.

DOR (months) = (date of progression/death or censor – date of confirmed response + 1)/30.4375. The Kaplan-Meier method will be used to estimate the distribution of DOR. The Kaplan-Meier survival curves, 25th, 50th (median), and 75th percentiles, along with the associated 2-sided 95% confidence intervals (CIs) based on Brookmeyer and Crowley method, and Kaplan-Meier probability estimates with 95% CIs at 6 and 12 months (or later time points if data permits) will be presented. The number of patients with events and the number of patients censored will be summarized.

Progression Free Survival (PFS)

PFS is defined as the time from the date of first dose to the date of the first documentation of a confirmed PD per IMWG criteria, or the date of death due to any cause, whichever occurs first. Patients without documentation of confirmed PD or death will be censored at the date of the last adequate response assessment prior to the date of initiation of subsequent anti-cancer therapy. Patients with no post baseline response assessment will be censored on day 1, unless patient died without extended loss-to-follow-up time. Please refer to Table 1 for the detailed censoring rules.

PFS (months) = (earliest date of progression or death or censor – date of first dose + 1)/30.4375. The analysis of PFS will be based on safety analysis set. The Kaplan-Meier method will be used to estimate the distribution of PFS. The Kaplan-Meier survival curves, the 25th, 50th (median), and 75th percentiles, along with associated 2-sided 95% confidence intervals (CIs) based on Brookmeyer and Crowley method, and Kaplan-Meier PFS probability estimates with 95% CIs at 3 and 6 months (or later time points if data permits) will be presented. The number of patients with events along with the type of events (death or progressive disease) and the number of patients censored will be summarized.

Time to Response (TTR)

Time to response is defined as the time from first dose to the date of first documentation of confirmed response (PR or better). Patients without confirmed response will not be included in the analysis of TTR.
TTR (months)= (date of confirmed response – date of first dose + 1)/30.4375).
Descriptive statistics including mean, median, standard deviation, minimum, maximum, Q1 and Q3 will be summarized in responders.
In addition, time to the best response (the time from the first dose to the date of the first documentation of the best confirmed response [PR or better]) will be analyzed by using the same approach.

Overall Survival (OS)

Overall survival is defined as the time from the date of first dose to the date of death from any cause. Patients without documentation of death at the time of analysis will be censored at the date last known to be alive.
OS (months) = (date of death or censor – date of first dose +1)/30.4375
The analysis of OS will be based on safety analysis set. The Kaplan-Meier method will be used to analyze the distribution of OS. Kaplan-Meier survival curves, the 25th, 50th (median), and 75th percentiles, along with the associated 2-sided 95% CIs based on Brookmeyer and Crowley method, and Kaplan-Meier estimates with 95% CIs at 6 and 12 months (or later time points if

data permits) will be presented. The number of patients with events and the number of patients censored will be summarized.

Best Overall Response

Best overall response is defined as the best response recorded after the first dose of study drug until subsequent anti-cancer therapy.

Confirmation of Response: All response categories except SD (sCR, CR, VGPR, PR, PD) require 2 consecutive adequate assessments made at any time before the initiation of subsequent anti-cancer therapy. There is no requirement on the time interval between the two visits. Consecutive visits are needed for response confirmation, except if there is one or two "not evaluable" (NE) or "not done" (ND) assessments in between.
Either one or two consecutive occurrences of "ND" or "NE" between responses can be skipped when confirming a response such as:

PR -> NE -> PR: PR is considered to be confirmed in this case.
CR -> ND -> CR: CR is considered to be confirmed in this case.

Confirmation of PD:

1. If only one PD is recorded as the last available assessment prior to the subsequent anti-cancer therapy, it can be counted as a confirmed PD. It will be considered as an event for DOR, TTP, PFS analyses.

BL → SD → PD → Subsequent anti-cancer therapy/End – It is a confirmed PD

2. If only one PD value is recorded in the middle of responses better than PD as below, it will not be counted as a confirmed PD. It will not be considered as an event for DOR, TTP, PFS analyses.

BL → SD → PD → SD → SD → Subsequent anti-cancer therapy/End – It is not a confirmed PD.

3. If the PD is due to imaging (plasmacytomas or bone lesion), the PD can be considered as a confirmed PD.
4. If only one PD is recorded as the last available assessment before the patient discontinues the treatment due to PD, it can be counted as a confirmed PD.

Best M-Protein Response

The best m-protein response is defined as the percent change from baseline to the best (lowest) value post-baseline. For subjects with both measurable serum and urine m-protein at baseline, the best m-protein will be the worse (i.e. lower percent reduction) of the best serum m-protein and the best urine m-protein. For subjects with only measurable serum m-protein at baseline, the

best m-protein response is based on serum m-protein. For subjects with non-measurable serum m-protein at baseline, but measurable urine m-protein at baseline, the best m-protein response is based on the urine m-protein. If not measurable by serum or urine m-protein the best m-protein response is based on the difference between involved and uninvolved FLC. A waterfall plot of the best m-protein response will be generated. The plot will have the following information:

1. Response.
2. Dose.
3. Schedule.
4. Measurable categories: Both Serum and Urine M-protein/Serum M-protein/Urine M-protein/FLC.

7.8.2 Part 3 Efficacy Analysis

7.8.2.1 Primary Efficacy Analysis

ORR, defined as the percentage of patients with a confirmed PR or better according to the IMWG Response Criteria. ORR based on IRC assessment will be analyzed as the primary endpoint. The primary analysis of ORR will be based on the full analysis set.

At the interim futility analyses, response status will be determined by a computer algorithm based on the IMWG Response Criteria. The algorithm will be specified in a separate document prior to the database lock. The algorithm based response will also be displayed at the time of primary analysis. A summary table of IRC-assessed, investigator-assessed, and algorithm-based best response with confirmation will be provided to assess the concordance rate at the time of primary analysis.

ORR at primary analysis will be analyzed based on the confirmed responses. Only the assessments from the start of treatment up to the earlier of confirmed disease progression or the start of subsequent anti-cancer therapy will be considered. Only subsequent systemic anti-cancer drugs taken are considered as anti-cancer therapy (radiotherapy and surgeries are not considered as systemic anti-cancer therapy for the purpose of this analysis). However, assessment after the treatment discontinuation will be considered. Subjects with only assessments of Not Evaluable or missing response will be treated as non-responders, i.e. they will be included in the denominator when calculating the percentage.

To confirm a response or progression, please refer to the details in Section 7.9.1 Best Overall Response.

The number and percentage of patients with BOR in the following response categories will be summarized by dose level: sCR, CR, VGPR, PR, MR, SD, PD, and NE/ND. The number and

percentage of patients with ORR (sCR+CR+VGPR+PR) will be presented with the corresponding 2-sided 95% exact CI. A list of IRC-assessed response at each visit will be provided.

A waterfall plot showing the maximum percent reduction from baseline in serum M-protein, or urine M-protein, or difference between two types of serum FLC [Kappa light chain (Kappa LC) and Lambda light chain (Lambda LC)] for each subject will be generated by dose level. The plot will be color-coded for M-protein types and serum FLC.
Only the assessments from the start of treatment up to the start of a subsequent anti-cancer therapy will be considered.
The maximum percent reduction will be plotted in the following hierarchical order:
[1] Plot the worse (higher) value of the maximum percent reduction from baseline in serum and urine M-protein if the patient has both serum and urine M-protein measurable;
[2] Plot Serum M-protein maximum percent reduction from baseline if data is available;
[3] If [2] is not feasible, plot Urine M-protein maximum percent reduction from baseline if data is available;
[4] If both [2] and [3] are not feasible, plot maximum percent reduction from baseline for difference between two types of Serum FLC if data is available.

Difference between two types of serum free light chain (FLC)
The percent change from baseline for difference between two types of serum FLC is defined as:

(post-baseline difference-baseline difference) / baseline difference *100%

To calculate the difference, the “involved” and “non-involved” light chains must be determined at first based on the ratio of non-missing values for Serum Kappa LC protein and Serum Lambda LC protein at baseline.
The detailed algorithm is provided as below:

- If the baseline ratio of (Kappa LC/Lambda LC)>1.65, then Kappa LC is defined as involved FLC, and Lambda LC is defined as non-involved FLC. Then:
  Difference between involved and uninvolved = Kappa LC-Lambda LC
- If the baseline ratio of (Kappa LC/Lambda LC) <0.26, then Lambda LC is defined as involved FLC, and Kappa LC is defined as non-involved FLC:
  Difference between involved and uninvolved = Lambda LC-Kappa LC
- If the baseline ratio of (Kappa LC/Lambda LC) ≤1.65 and ≥ 0.26, then “involved” and “non-involved” FLC cannot be determined (ratio is normal), and maximum percent reduction from baseline for difference between two types of Serum FLC won’t be available.

7.8.2.2 Secondary Efficacy Analysis

ORR based on Investigator Assessment

ORR based on investigator-assessed response will be analyzed as one of the secondary endpoints. It will be analyzed based on the confirmed responses by using the same analysis approach as for IRC-assessed ORR.

DOR

DOR is defined as the time from first documented evidence of confirmed PR or better until the earliest date of a confirmed PD per IMWG, or death among patients who achieve a response (i.e., confirmed PR or better). Responders without confirmed PD or death will be censored at the censoring time point for PFS as specified in Table 1.

DOR will be analyzed at the time of primary ORR analysis, also at final analysis (study close out), based on the responses assessed by both IRC and investigator.

DOR will be summarized using the Kaplan-Meier method by dose level for patients with a confirmed PR or better. Median DOR, first and third quartiles, 3-month, 6-month, 9-month, 12-month DOR rate (or later time points if data permits), and 95% CIs will be estimated using the Brookmeyer-Crowley method. A Kaplan-Meier plot of DOR time will also be provided.

CBR

CBR is defined as the percentage of patients with a confirmed MR or better, according to the IMWG Response Criteria. CBR will be summarized in the same way as ORR. At the primary analysis, CBR based on both IRC and investigator assessment will be summarized. At interim analysis, investigator-assessed CBR will be summarized.

DCR

DCR is defined as the percentage of patients with a SD or confirmed MR or better, according to the IMWG Response Criteria. DCR will be summarized in the same way as ORR. At the primary analysis, DCR based on both IRC and investigator assessment will be summarized. At interim analysis, investigator-assessed DCR will be summarized.

For CBR and DCR, the number and percentage of patients with the CBR and DCR will be summarized by dose level. The corresponding exact 95% CI for CBR and DCR will also be provided. Patients with unknown or missing responses will be included in the denominator when calculating percentages of response.

Duration of Clinical Benefit

Duration of clinical benefit is defined as the time from first documented evidence of confirmed MR or better until the earliest date of a confirmed PD per IMWG, or death among patients who achieve a confirmed MR or better. Patients without confirmed PD or death will be censored at the censoring time point for PFS as specified in Table 1.

Duration of clinical benefit will be analyzed at the time of primary ORR analysis, also at final analysis (study close out), based on responses assessed by both IRC and investigator.

Duration of clinical benefit will be summarized using the Kaplan-Meier method by dose level for patients with a confirmed MR or better. Median duration of clinical benefit, first and third quartiles and 95% CI, will be estimated using the Brookmeyer-Crowley method.

Duration of Disease Control

Duration of disease control is defined as the time from first documented evidence of SD or better until the earliest date of a confirmed PD per IMWG, or death among patients who achieve a SD or better. Patients without confirmed PD or death will be censored at the censoring time point for PFS as specified in Table 1.

Duration of disease control will be analyzed at the time of primary ORR analysis, also at final analysis (study close out), based on responses assessed by both IRC and investigator.

Duration of disease control will be summarized using the Kaplan-Meier method by dose level for patients with a SD or confirmed MR or better. Median duration of disease control, first and third quartiles and 95% CI, will be estimated using the Brookmeyer-Crowley method.

PFS

PFS is defined as the time from the date of first dose until the earliest date of confirmed PD per IMWG, or death due to any cause. PFS will be analyzed at the time of primary analysis of ORR, also at final analysis (study close out), based on responses assessed by both IRC and investigator.

A summary of the assignments for progression and censoring dates for PFS are specified in Table 1 below.

Table 1 Assignments for Progression and Censoring Date for PFS Analysis

Scenarios	Event date/censoring date	Event status
No (or inadequate) baseline tumor assessments[1]	Date of the first dose	Censored
No post-baseline assessments	Date of the first dose	Censored
Confirmed progression documented without extended loss-to-follow-up time (two or more missed cycles)[4]	Date of assessment of the earliest date of the confirmed progression	Event
No confirmed progression or death	Date of last adequate assessment of response[2]	Censored
Subsequent anti-cancer therapy started (before documented confirmed progression or death)[3]	Date of last adequate assessment of response[2] (prior to the initiation of the subsequent anti-cancer therapy)	Censored
Death without extended loss-to-follow-up time	Death date	Event
Death or progression after an extended loss-to-follow-up time (two or more missed cycles)[4]	Date of last adequate assessment of response[2] (prior to missed assessments)	Censored

1. Adequate baseline assessment is defined as at baseline, a patient has at least one of the following measurements:
    a. Serum M-protein ≥ 0.5g/dL (≥500 mg/dL) or
    b. Urine M-protein ≥200 mg/24 hours or
    c. Serum FLC assay: Involved FLC level ≥10 mg/dL (≥100 mg/L) and an abnormal serum FLC ratio (<0.26 or >1.65)
2. An adequate assessment is defined as an assessment where the IRC/investigator determined response is sCR, CR, VGPR, PR, MR, or SD prior to a subsequent anti-cancer therapy. An unconfirmed PD assessment is also considered as an adequate assessment.
3. If PD or death and the subsequent anti-cancer therapy occur on the same day, it is assumed that the progression or death was documented first (e.g., event status is an event, and the date of event is the date of progression or death). If the subsequent anti-cancer therapy is initiated prior to any adequate assessment, censoring date should be the date of the first dose.
4. Refer to Section 7.1.4 “Extended Loss-to-follow-up”

PFS will be summarized using the Kaplan-Meier method by dose level. Median PFS, first and third quartiles, 3-month, 6-month, 9-month, 12-month PFS rate (or later time points if data

permits), and 95% CIs will be estimated using the Brookmeyer-Crowley method. A Kaplan-Meier plot of PFS time will also be provided.

TTP

TTP is defined as the time from the date of the first dose until the earliest date of confirmed PD per IMWG.

TTP will be analyzed at the time of primary analysis of ORR, also at final analysis (study close out), based on responses assessed by both IRC and investigator.

A summary of the assignments for progression and censoring dates for TTP are specified in Table 2.

Table 2 Assignments for Progression and Censoring Date for TTP Analysis

Scenarios	Event date/censoring date	Event status
No (or inadequate) baseline tumor assessments[1]	Date of the first dose	Censored
No post-baseline assessments	Date of the first dose	Censored
Confirmed progression documented without extended loss-to-follow-up time (two or more missed cycles)[4]	Date of assessment of the earliest date of the confirmed progression	Event
No confirmed progression or death	Date of last adequate assessment of response[2]	Censored
Subsequent anti-cancer therapy started (before documented confirmed progression)[3]	Date of last adequate assessment of response[2] (on or prior to the initiation of the subsequent anti-cancer therapy)	Censored
Death	Date of last adequate assessment of response[2] prior to death	Censored
Progression after an extended loss-to-follow-up time (two or more missed cycles)[4]	Date of last adequate assessment of response[2] (prior to missed assessments)	Censored

1. Adequate baseline assessment is defined as at baseline, a patient has at least one of the following measurements:
    a. Serum M-protein ≥ 0.5 g/dL (500 mg/dL) or
    b. Urine M-protein ≥200 mg/24 hours or

c. Serum FLC assay: Involved FLC level ≥10 mg/dL (≥100 mg/L) and an abnormal serum FLC ratio (<0.26 or >1.65)

2. An adequate assessment is defined as an assessment where the IRC/investigator determined response is sCR, CR, VGPR, PR, MR, or SD prior to a subsequent therapy. An unconfirmed PD assessment is also considered as an adequate assessment.

3. If PD and the subsequent anti-cancer therapy occur on the same day, it is assumed that the progression was documented first (e.g., event status is an event, and the date of event is the date of progression). If the subsequent anti-cancer therapy is initiated prior to any adequate assessment, censoring date should be the date of the first dose.

4. Refer to Section 7.1.4 "Extended Loss-to-follow-up"

TTP will be summarized using the Kaplan-Meier method by dose level. Median TTP, first and third quartiles, 3-month, 6-month, 9-month, 12-month TTP rate (or later time points if data permits), and 95% CI will be estimated using the Brookmeyer-Crowley method. A Kaplan-Meier plot of TTP time will also be provided. Analysis of TTP will be based on both IRC and investigator-assessed responses.

OS

OS is defined as the time from the date of the first dose until death due to any cause. Patients who do not have a death record at the clinical cut-off date for the analysis will be censored at the last known alive date. The last known alive date will be determined by the latest collection/assessment date from among all data domains within the clinical database.

An OS analysis will be performed at the time of primary analysis of ORR, also at final analysis (study close out).

OS will be summarized using the Kaplan-Meier method by dose level. Median OS, first and third quartiles, and 95% CIs will be estimated using the Brookmeyer-Crowley method. A Kaplan-Meier plot and listing of OS time will also be provided. In addition, pending on maturity of data, the survival probability at 6, 12 and 18 months with 95% CI will be estimated using the Kaplan-Meier method.

MRD negativity at a sensitivity of 10^{-5}

MRD negativity at a sensitivity of 10^{-5} as a secondary endpoint is defined as patients who are MRD negative at a sensitivity of 10^{-5} in patients achieving suspected CR. Duration of MRD negativity (10^{-5}) is defined as the time from the first MRD negative status (10^{-5}) to the earliest date of the MRD positive status (10^{-5}), confirmed PD per IMWG or death.

For MRD negativity (10^{-5}) rate based on bone marrow testing, the number and percentage of patients who have achieved MRD negativity (10^{-5}) will be summarized by dose level. The corresponding exact 95% CI for MRD negativity (10^{-5}) rate will also be provided. The rate of maintaining MRD negativity (10^{-5}) will be reported in patients who achieve CR and MRD negativity (10^{-5}). Duration of MRD negativity (10^{-5}) will be analyzed using the Kaplan-Meier method in patients who have achieved MRD negativity (10^{-5}) when data permit; otherwise, a listing will be generated.

7.8.2.3 Exploratory Efficacy Analysis

MRD negativity at a sensitivity of 10^{-6}

MRD negativity at a sensitivity of 10^{-6} as an exploratory endpoint is defined as patients who are MRD negative at a sensitivity of 10^{-6} in patients achieving suspected CR. Duration of MRD negativity (10^{-6}) is defined as the time from the first MRD negative status (10^{-6}) to the earliest date of the MRD positive status (10^{-6}), confirmed PD per IMWG or death.

For MRD negativity (10^{-6}) rate based on bone marrow testing, the number and percentage of patients who have achieved MRD negativity (10^{-6}) will be summarized by dose level. The corresponding exact 95% CI for MRD negativity (10^{-6}) rate will also be provided. The rate of maintaining MRD negativity (10^{-6}) will be reported in patients who achieve CR and MRD negativity (10^{-6}). Duration of MRD negativity (10^{-6}) will be analyzed using the Kaplan-Meier method in patients who have achieved MRD negativity (10^{-6}) when data permit; otherwise, a listing will be generated.

Time to Response (TTR)

Time to response is defined as the time from first dose to the date of first documentation of confirmed response (PR or better). Patients without confirmed response will not be included in the analysis of TTR.

Descriptive statistics including mean, median, standard deviation, minimum, maximum, Q1 and Q3 will be summarized in responders.

In addition, time to the best response will be analyzed by using the same approach in responders.

7.8.2.4 Subgroup Efficacy Analysis

For the primary endpoint ORR by IRC assessment, subgroup analysis will be performed in the following selected subgroups.

Subgroups	**Categories**
Age group (at screening)	e.g. 18 to <70; ≥=70
Gender	Male; female

Race	White; Black; Asian; Other
Type of myeloma	IgA; IgG; light chain only
Cytogenetics risk	High risk [defined as 17p-, t(4;14) and/or t(14;16)]; complementary standard risk (non-high risk) Expanded high risk [defined as High risk and/or 1q21 amplification/gain]; complementary standard risk (non-expanded high risk)
Prior treatment	Anti-CD38 in last line of therapy; IMiD in last line of therapy; Recently approved agents (Selinexor, Anti-BCMA CAR-T, belantamab); Bispecific T-cell engagers; Anti-BCMA exposed; Anti-BCMA exposed (CAR-T); Anti-BCMA exposed (ADC); Anti-BCMA exposed (Bispecific T-cell engagers); Penta exposed.
Number of prior lines of therapy	≤4, >4
Refractory to prior anti-cancer therapy	Anti-CD38 monoclonal antibody; CAR-T; IMiD (Lenalidomide, Pomalidomide, Thalidomide); Proteasome Inhibitor; Triple refractory (IMiD, PI, and anti-CD38); Penta refractory (Bortezomib, Carfilzomib, Lenalidomide, Pomalidomide and anti-CD38); Anti-BCMA refractory; Triple refractory and bispecific T-cell engager exposed.

In the subgroup analysis, ORR will be presented using forest plot within each treatment arm for the following categories: age, gender, race, type of myeloma, cytogenetic risk group, prior treatment, number of prior lines of therapy, and refractory to prior anti-cancer therapy. Other subgroups may be explored if data permits, but not required by this SAP. If the percentage of subjects is small within a particular subgroup, then the subgroup categories may be refined prior to the primary analysis of Part 3 extension.

7.9 Patient-Reported Outcomes (Part 3 only)

Patient-reported outcome measures will be analyzed using the PRO analysis set. The primary PRO score of interest will be the EORTC QLQ-MY20 disease symptoms subscale at approximately 12 weeks (C4D8). Data preparation steps will include those described in Appendix F. PRO analyses will only include observations that fall within "windows" specified in Appendix F.

EORTC QLQ-MY20

The actual value and nominal change from baseline of the EORTC QLQ-MY20 subscale scores will be summarized using descriptive statistics overall and by dose arm over time, with 95% CIs, and presented both in tables and figures.

Considering sufficient sample sizes and follow-up data for model convergence, mixed models for repeated measures (MMRM) may be conducted to examine the change from baseline in EORTC QLQ-MY20 single or multi-item subscales. The models will include the following covariates: treatment, visit, baseline score, and any variables that were used as randomization factors (cytogenetic risk, myeloma type). The models for all EORTC QLQ-MY20 subscales will be first conducted using unstructured covariance structure; if any subscale model does not converge, then compound symmetry or autoregressive (1) [AR(1)] covariance structures will be considered. The repeated measures analysis will use assessments collected per the schedule of events in the protocol (see Appendix F). The estimated least squares (LS) mean change in score from baseline with 95% CIs will be reported at each treatment cycle visit (time point) for each treatment arm. However, no comparison will be made between the treatment arms. LS mean changes from baseline with 95% CIs will be graphically presented over time.

In addition, there will be empirical cumulative distribution functions of change in scores from baseline at each visit per treatment group, to provide information to help support the interpretation of the results. Calculate the $P(X \leq x)$ for each EORTC QLQ-MY20 subscales, and plot as cumulative distribution functions (CDFs), characterized by an S-shape rising from 0% to 100% on the y-axis (i.e., lower to higher, from left to right). For EORTC QLQ-MY20 functional subscales, the arrow pointing to the left of zero ($X<0$) is "Worsening," and the arrow pointing to the right of zero ($X>0$) is "Improvement"; for the EORTC QLQ-MY20 symptom subscales, the arrow pointing to the left of zero ($X<0$) is "Improvement," and the arrow pointing to the right of zero ($X>0$) is "Worsening". These CDF curves depict the cumulative probability on the y-axis that a patient will report a change from baseline on the EORTC QLQ-MY20 subscale less than or equal to the observed value "x" on the x-axis. The CDF curves will be provided for the visits if the number of patients per treatment arm is over 20.

The subscales of EORTC QLQ-MY20 are defined as shown in Table 3. Details of scoring and initial handling of missing data are included in the EORTC QLQ-MY20 scoring guidelines, which specifies that multi-item scales can be scored whenever at least half of the items in the subscale are completed: e.g., for the Disease Symptom subscale, at least 3 of items 1-6 will need to be completed. For EORTC QLQ-MY20, no imputations of missing PRO assessments will be performed.

Table 3 Subscales of EORTC QLQ-MY20

Subscale scores	Included items
Future perspective	18-20
Body image	17
Disease symptoms	1-6
Side effects of treatment	7-16

For each subscale score of EORTC QLQ-MY20, the number and percentage of patients with either a stable score, an improvement in score or a worsening in score from study entry based on a minimally important difference (MID) of 10 will be summarized by treatment arm from baseline to week 12 (C4D8). MID thresholds for functioning subscales: Improved = Change from baseline in subscale score ≥10; Stable = Change from baseline in subscale score ≤9 and ≥-9; Worsening = Change from baseline in subscale score ≤-10. MID thresholds for symptom subscales: Improved = Change from baseline in subscale score ≤-10; Stable = Change from baseline in subscale score ≥-9 and ≤9; Worsening = Change from baseline in subscale score ≥10.

Other MIDs and responder metrics may also be considered: e.g., Sully et al. (2019) defines the smallest difference in mean score between groups which could be considered clinically meaningful, as the minimally important difference (MID), while the threshold of within-patient change deemed meaningful, used to define a patient as a responder if their change in score exceeds this threshold, is referred to as the responder definition (RD); where EORTC QLQ-MY20 MID: Disease Symptoms (DS 10 points), Side Effects of Treatment (SE 10 points), Body Image (BI 13 points), and Future Perspective (FP 9 points); and, Responder definitions (RD): DS (16 improvement; 11 worsening), SE (6 improvement; 9 worsening), BI (33 improvement; 33 worsening), and FP (11 improvement; 11 worsening).

PRO completion over time and in overall will be calculated at the questionnaire level for the EORTC QLQ-MY20 in the full analysis set. Proportion of PRO score compliance will be

calculated as follows: (number of patients with the PRO score completed on Cycle X Day Y) / (number of patients for whom a PRO score is expected on Cycle X Day Y). A subject will be considered to have a PRO score "completed" on Cycle X Day Y if they have enough item scores to calculate at least one subscale score on the EORTC QLQ-MY20 questionnaire. A subject is "expected" to have a PRO score on Cycle X Day Y if treatment was received on Day 1 of the same cycle.

PRO 'available data rate' will also be calculated over time, and in overall at the questionnaire level for EORTC QLQ-MY20 in the full analysis set. Proportion of PRO score compliance will be calculated as follows: (number of patients with the PRO score completed on Cycle X Day Y) / the number of patients in the full analysis set i.e. number of patients consented and eligible to participate in the PRO data collection.

EQ-5D-5L

The EQ-5D-5L patient-reported questionnaire consists of 5 dimensions (5 items: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) each rated on 5 levels. The EQ-5D-5L also contains the Visual Analogue Scale (VAS); EQ VAS ranges from 0 (worst imaginable health state) to 100 (best imaginable health state).

EQ-5D-5L data will be analyzed using the PRO analysis set. Data preparation steps will include those described in Appendix F. EQ-5D-5L summaries and analyses will only include observations that fall within "windows" specified in Appendix F.

EQ-5D-5L (i.e., 5 individual item scores) will be analyzed by summarizing the raw item scores descriptively over time.

For EQ VAS, ranging from 0 to 100, the actual mean values and mean change from baseline, with 95% CIs, will be summarized using descriptive statistics in overall and by treatment group over time and presented both in tables and figures. As with EORTC QLQ-MY20, MMRM analyses may be considered, if sample sizes and follow-up data are sufficient.

PRO compliance over time and overall will be calculated at the questionnaire level for the EQ-5D-5L in the full analysis set. Proportion of PRO score completion and PRO 'available data rate' will be calculated following the approach for EORTC QLQ-MY20 above.

For EQ-5D-5L, no imputations of missing PRO assessments will be performed.

7.10 Healthcare Resource Utilization Analysis (Part 3 only)

Healthcare resource utilization data will be summarized in descriptive statistics of medical encounters in the full analysis set. Specifically, the number of hospitalizations, number of intensive care unit (ICU) stays, number of emergency room stays, number of outpatient visits, length of hospital stay, length of stay in ICU, length of stay in hospital floor bed. Visit reasons for each of these and primary diagnosis at discharge will be summarized over time.

7.11 Pharmacokinetic/Pharmacodynamic Analysis

7.11.1 Pharmacokinetic Analysis

7.11.1.1 Serum Modakafusp alfa Concentrations

Blood samples will be collected at prespecified time points as described in the study protocol for the measurement of serum modakafusp alfa concentrations. Individual serum concentration data will be listed by patient, dosing schedule (A, B, C, and D), dose level, study day, and sampling time. Both nominal (scheduled) and actual sampling times will be presented in the listings.

Serum modakafusp alfa concentrations will be summarized by nominal time post dose, grouped by dosing schedule, dose level, nominal infusion duration, and study day by phases (escalation and expansion). Summary statistics will be reported at nominal sampling times; means will be reported if the number of observations above the lower limit of quantitation (NALQ) is ≥50% of the number of patients. The summary statistics will consist of: N, NALQ, arithmetic mean, standard deviation (SD), coefficient of variation (CV), geometric mean, geometric CV, median, min, and max. The SD and CV will be reported on at least 3 non-missing values. Modakafusp alfa concentrations that are below the limit of quantitation (BLQ) will be set to zero for calculation of summary statistics, except for geometric means, where BLQ values will be considered missing.

Concentration data that are considered anomalous may be excluded from the concentration summaries and plots. Evidence or explanations will be provided in the clinical study report to justify the exclusion of concentration data.

Mean and individual modakafusp alfa serum concentration data will be plotted over nominal sampling time, grouped by dosing schedule, dose level, nominal infusion duration, and study day on both linear and semi-logarithmic scales. Mean and individual serum concentrations will be plotted over time (labeled with cycle number), grouped by dosing schedule and dose level on a linear scale by phases (escalation and expansion). Visual inspection of the serum concentration-time plots will be used to make inferences regarding the attainment of PK steady-state by cycle.

BLQ values will be plotted as zero on a linear scale and treated as missing on a semi-logarithmic scale.

7.11.1.2 Serum PK Parameters

The serum PK concentration-time course data will be used to calculate standard PK parameters using noncompartmental methods with Phoenix WinNonlin.

Actual sampling times will be used for the calculation of PK parameters. In the event that actual collection times are either unreliable or missing, nominal collection times will be used. For the calculation of PK parameters, serum concentrations of modakafusp alfa that are BLQ will be treated as zero prior to t_{max}, missing between t_{max} and the time of the last measurable concentration, and the concentration-time curve will be considered to have terminated at the time of the last measurable concentration. If measurable concentrations are near the lower limit of quantification (LLOQ) or imbedded between BLQ concentrations, these values may be excluded at the discretion of the Clinical Pharmacologist. Concentration data that are considered anomalous may not be used in the calculation of PK parameters; evidence or explanations will be provided in the clinical study report to justify the exclusion of data. In addition, patients with several missing samples around the expected T_{max} and patients for whom AUC estimation was considered not be reliable due to several missing samples (50% or more of the data are missing) may be excluded from the PK parameter (Cmax and/or AUC) calculation and PK parameter summary.

The following PK parameters will be determined in Parts 1 and 2, and in Chinese patients with the intensive PK schedule in Part 3,, as permitted by data:

- C_{max}.
- t_{max}.
- AUC_{∞}.
- AUC_{last}.
- λ_z.
- $t_{1/2z}$.
- CL.
- V_{ss}.
- Accumulation ratio based on AUC_{last}.

Individual PK parameters will be presented in listings, PK parameters for Cycle 1 Day 1, Cycle 1 Day 15, Cycle 2 Day 1, and Cycle 2 Day 15, as appropriate, will be summarized and grouped by dosing schedule, dose level, and study day by phases (escalation and expansion). The SD and CV

will be reported on at least 3 non-missing values. Except for t_{max}, the summary statistics will consist of: N, mean, SD, CV, geometric mean, geometric CV, median, min, and max. The summary statistics for t_{max} will consist of: N, median, min, and max.

The PK data collected in this study are intended to contribute to future population PK analyses of modakafusp alfa. These population PK analyses may include data collected in other modakafusp alfa clinical studies. The analysis plan for the population PK analysis will be separately defined, and the results of these analyses will not be reported in the clinical study report.

7.11.2 Pharmacodynamic Analysis

During the clinical development of modakafusp alfa, a range of biomarkers will be assessed to test for their correlation with safety and efficacy. Markers that will be studied are markers linked to the drug itself, to the treated disease, or to patient factors.

A pharmacodynamic analysis will be included in CSR only if there is a trend.

7.11.3 Immunogenicity Analysis

The immunogenicity of modakafusp alfa will be assessed by determining anti–modakafusp alfa antibody (ADA) incidence and characteristics (eg, titer and specificity) and by determining neutralizing antibody (NAb) incidence. All ADA and NAb data will be listed. Percent of patients with positive ADA or NAb, percent of patients with positive ADA by domain specificity, and summary statistics of ADA titer at each timepoint will be summarized by dose schedules and dose levels by study phases. Percent of patients with positive ADA and NAb at baseline and positive ADA (final result and by domain specificity) and NAb at any post-baseline time point during the study will be summarized. Additional summary of the ADA kinetics including transient- vs. persistent-ADA, preexisting ADA, treatment-induced, or treatment-boosted ADA will be summarized, as applicable. ADA titer distribution over time will be presented through a box-plot on a logarithmic scale. The impact of ADA on the PK profile, drug efficacy, pharmacodynamic profile, and clinical safety will be evaluated and summarized, if possible. These analyses will be descriptive and exploratory in nature and may be included in CSR if there is a trend.

7.12 Safety Analysis

For Part 3 extension cohorts, analysis and review will assess the safety data across the treatment arms as well as in the patients with low body weight. Data to be reviewed will be summarized by treatment arm as well as by body weight (e.g., ≤ 80 kg and > 80 kg). With additional cumulative data, the weight subgroups may be further subdivided. The data might be included in the CSR if there is a safety trend identified in a specific low body weight group.

7.12.1 Adverse Events

Adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) Version 24.0 or later (based on version at time of database lock). Treatment-emergent is defined as any AE that occurs after administration of the first dose of any study treatment through 30 days after the last dose of any study treatment.

Treatment-emergent AEs will be summarized by MedDRA system organ class (SOC), and preferred term. For summary tabulations the following hematologic abnormalities coded to MedDRA preferred terms in the Investigations SOC will be pooled with the appropriate clinical terms in the Blood and lymphatic system disorders SOC:

MedDRA Preferred Term (Investigation SOC)	**Mapped to (Blood and lymphatic system disorders SOC)**
Neutrophil count decreased	Neutropenia
Platelet count decreased	Thrombocytopenia
Hemoglobin decreased	Anemia
While blood cell count decreased	Leukopenia

Treatment-emergent AEs will also be summarized by the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) Version 5.0 or later.

Summary tabulations will include the following categories:

- Treatment-emergent AEs.
- Drug-related treatment-emergent AEs.
- Grade 3 and Grade 4 treatment-emergent AEs (based on maximum severity).
- The most commonly reported treatment-emergent AEs (i.e., those events reported by >= 10% of all patients).
- Serious treatment-emergent adverse events (SAEs).
- Drug-related serious treatment-emergent adverse events.
- Grade 3 and Grade 4 treatment-emergent SAEs (based on maximum severity).
- Grade 3 and Grade 4 drug-related treatment-emergent SAEs (based on maximum severity).
- Grade 5 treatment-emergent AEs.
- Grade 5 drug-related treatment-emergent AEs.
- Treatment-emergent AEs resulting in discontinuation of study drug.
- Treatment-emergent AEs resulting in study drug modifications (including discontinuation, delay, reduction, and interruption).
- Treatment-emergent AEs resulting in dose delays.

- Treatment-emergent AEs resulting in dose reductions.
- Treatment-emergent AEs resulting in drug interruptions.
- Treatment-emergent AEs resulting in drug infusion interruptions.
- Treatment-emergent AESI – all Grades.
- Grade 3 or higher treatment-emergent AESI.
- Serious treatment-emergent AESIs
- Treatment-emergent haemorrhage standardised MedDRA query (SMQ) AEs
- Drug-related treatment-emergent haemorrhage SMQ AEs.
- Grade 3 and Grade 4 treatment-emergent haemorrhage SMQ AEs.
- Serious treatment-emergent haemorrhage SMQ AEs
- Grade 5 treatment-emergent haemorrhage SMQ AEs.

All summaries will be by dose schedule and dose levels by phases (escalation, expansion, and extension cohorts) for patients in the safety analysis set.

Patients with the same AE more than once will have that event counted only once within each SOC, and once within each preferred term. Patients with the same AE more than once will have only the maximum intensity of that event counted once within each SOC, and once within each preferred term.

Most commonly reported (at least 10% of all patients) treatment-emergent AEs will be presented by preferred term only. Patients with multiple occurrences of the same AE will have that event counted only once within each preferred term.

AE of special interest from a scientific and medical perspective (AESI) will be reported by preferred terms. In this study, infusion-related reactions (IRR) are considered as AESIs. The IRR signs and symptoms will be reported separately by preferred terms under the AESI tables only.

An overall summary treatment-emergent AE table will include numbers and percentages of patients who had any treatment-emergent AE, drug-related treatment-emergent AE, Grade 3 or higher treatment-emergent AE, grade 3 or higher drug-related treatment-emergent AE, serious treatment-emergent AE, drug-related serious treatment-emergent AE, serious treatment-emergent AE leading to study drug discontinuation, treatment-emergent AE resulting in study drug discontinuation, and on-study deaths. On-study death is defined as deaths that occur between the first dose of study drug and up to 30 days after the last dose of study drug and deaths that occur after 30 days after the last dose of study drug but are assessed as related to study drug.

A by-patient listing of DLTs as identified by the investigator that occur during Cycle 1 of treatment will be presented by dose schedule and dose levels for all patients enrolled during the dose escalation portion of this study. Patients will be grouped by the dose level to which they were originally assigned, including those who receive subsequent treatment at a lower or higher dose level. DLT-like events identified programmatically after Cycle 1 in Part 1 escalation cohorts and Part 2 expansion cohorts will be summarized and also be presented in a by-patient listing. Please refer to Appendix C for details.

A by-patient listing of all deaths regardless of whether on study or not will also be presented. In this listing, on-study deaths will be flagged.

7.12.2 Clinical Laboratory Evaluations

For the purposes of summarization, all laboratory values will be converted to standardized units. If a lab value is reported using a non-numeric qualifier (eg, less than (<) a certain value, or greater than (>) a certain value), the given numeric value will be used in the summary statistics, ignoring the non-numeric qualifier.

If a patient has repeated laboratory values for a given time point, the value from the last evaluation will be used. Laboratory test results will be summarized according to the scheduled sample collection time point. Change from baseline will be presented. Scheduled laboratory along with unscheduled lab test results will be listed.

Lab parameters to be analyzed are as follows:

- Hematology: Hemoglobin, Lymphocyte count, Neutrophil count, Platelet count, White blood cell count.
- Chemistry: albumin, alanine aminotransferase (ALT), alkaline phosphatase (ALP), aspartate aminotransferase (AST), bilirubin (total), corrected calcium, creatinine, glucose, potassium, sodium.

Whenever available, laboratory values will be assigned toxicity grades using the NCI CTCAE Version 5.0. Shift tables will be constructed for laboratory parameters either using CTCAE grade or based on low/medium/high compared to normal ranges on patients who have both baseline and at least one post-baseline assessment. Individual platelet and ANC profiles will be generated for each schedule/dose level. Mean values over time for platelets and ANC will be produced for the expansion cohort only.

7.12.3 Vital Signs

For dose escalation cohorts, descriptive statistics for vital sign results (diastolic and systolic blood pressure, respiratory rate, pulse rate, oxygen saturation and body weight) will be summarized by schedules and dose levels as follows:

- Baseline value (C1D1 or screening if C1D1 is not available).
- Minimum post-baseline value.
- Change to Minimum post-baseline value.
- Maximum post-baseline value.
- Change to Maximum post-baseline value.

Changes to the minimum and maximum post-baseline values will be calculated relative to the baseline value.

For blood pressure values collected during the infusion, the number and percentage of patients with changes from pre-infusion blood pressure values will be summarized in the following categories:

- Increase of diastolic blood pressure of at least 10 mm Hg during any infusion.
- Decrease of diastolic blood pressure of at least 10 mm Hg during any infusion.
- Increase of systolic blood pressure of at least 20 mm Hg during any infusion.
- Decrease of systolic blood pressure of at least 20 mm Hg during any infusion.

For expansion and extension cohorts, descriptive statistics for vital sign results will be summarized by visit up to end of treatment (EOT) visits (including visits where blood pressure was measured every 30 minutes during the first 4 infusions, and after the end of the infusion). In addition, a summary of diastolic blood pressure and systolic blood pressure during the infusion will be summarized as described above for escalation cohorts.

7.12.4 12-Lead ECGs

ECG data (ventricular rate, RR interval, PR interval, QT interval, and QTcF interval) will be summarized by schedules and dose levels as follows:

- Baseline value (C1D1 or screening if C1D1 is not available).
- Minimum post-baseline value.
- Change to Minimum post-baseline value.
- Maximum post-baseline value.
- Change to Maximum post-baseline value.

Changes to the minimum and maximum post-baseline values will be calculated relative to the baseline value.

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

7.12.5 Other Observations Related to Safety

Shifts from baseline to the worst post-baseline ECOG score will be tabulated.

7.13 Interim Analyses and Criteria for Early Termination

Part 2

A futility analysis will be conducted when 10 response-evaluable patients in a phase 2 arm have been followed for at least 3 cycles or have already responded or withdrawn. An arm with the predictive probability <15% (i.e., 3 or fewer responders in 10 patients) will stop enrollment for futility. Overall, avoidance of futility requires 11 or more responders out of the 25 patients to achieve 90% of the overall posterior probability of success (PPOS) at the final analysis. Unless all 25 patients have already been enrolled, another futility analysis with the same decision rule based on PPOS will be conducted when 15 response-evaluable patients in a phase 2 arm have been followed for at least 3 cycles or have already responded or withdrawn. If there are 5 or fewer responders in 15 patients, enrollment will stop for futility. However, the study team reserves the rights to continue the study if the patients still benefit from the study drug.

Part 3

In the Part 3 extension cohorts, an independent data monitoring committee (DMC) will review accumulating safety, tolerability, and efficacy data. The first DMC meeting will occur after approximately 20 patients (approximately 10 patients from each arm) have been enrolled and treated. The second and third DMC meeting will occur at the first and second interim futility analysis when approximately 30 and 96 patients (approximately 15 and 48 patients from each arm) have been enrolled and have completed 3 cycles of treatment or discontinue the treatment prematurely. The subsequent DMC meetings will be approximately every 6 months thereafter.

Futility Stopping Rules

Two interim analyses for futility are planned for the Part 3 extension cohorts when approximately 15 and 48 patients of the planned 118 patients per arm are enrolled and treated for at least 3 cycles or have discontinued treatment prematurely. The futility stopping boundary will be determined based on the predictive probability of success at the primary analysis, using a Bayesian efficacy monitoring approach. The treatment arm will be dropped for futility if the predictive probability of success at the primary analysis is found to be <10% based on the data

from the interim analysis. Various prior distributions for the ORR are used for each treatment arm due to different prior efficacy information at different dose levels collected in Parts 1 and 2 of the study. Based on the data cutoff of October 22, 2021, the observed ORR was 40% for the overall 30 patients enrolled in 1.5 mg/kg cohorts including both escalation and expansion cohorts. To reflect the prior information for the 120 mg arm collected in the Part 1 and Part 2, a weakly informative prior of Beta (1.2, 1.8) is used for the 120 mg arm. This will translate to a Beta prior with a response rate of 40% based on the information of 3 patients (10% of the sample size in Part 1 and 2). However, because of comparatively limited efficacy data available at the 3 mg/kg (240 mg) dose, a noninformative prior of Beta (0.5, 0.5) is used. Based on these prior distributions, Table 3 shows the stopping rules used for the treatment arms of 240 mg and 120 mg at each interim futility analysis based on the prespecified number of patients. The actual futility boundaries will be recalculated based on the observed number of treated patients at each interim futility analysis using the same posterior probability stopping rule if it is different from the planned number of patients. The details of the operating characteristics of the futility stopping rules are described in Appendix D.
There is no plan to stop for efficacy based on the interim analysis data.

Table 3 Stopping Boundaries for 240mg and 120mg Arms at Each Interim Futility Analysis

	Number of patients enrolled and followed up for 3 months	Recommend terminating the corresponding treatment arm due to lack of efficacy when number of responders is:	
		240mg Arm	120mg Arm
1st Interim Futility Analysis	15	≤2	≤1
2nd Interim Futility Analysis	48	≤10	≤10

Continuous Safety Monitoring Plan and Stopping Rules

Grade ≥4 nonhematologic treatment-related TEAEs, as well as treatment-related deaths, will be continuously monitored at the following timepoints using a Bayesian toxicity monitoring approach based on posterior probability (Table 4).

Table 4 Safety Monitoring Time Points Based on Bayesian Stopping Rules

	Timepoint
1st DMC safety review	20 patients have been treated (10 patients in each arm)
2nd DMC safety review (first interim futility analysis)	15 patients in each arm treated for at least 3 cycles or discontinued treatment prematurely

3rd DMC safety review (second interim futility analysis)	48 patients in each arm treated for at least 3 cycles or discontinued treatment prematurely

The stopping rules will be determined based on the posterior probability of the unacceptable toxicity event rate >25%, and the unacceptable treatment-related death rate >5%. If there is at least an 80% probability that the true toxicity event (Grade ≥4 nonhematologic treatment-related TEAEs) rate is greater than 25% given the observed toxicity data, the enrollment will be paused until the IDMC reviews the data to avoid putting more patients at risk. Similarly, if there is at least an 80% probability that the true treatment-related death rate is greater than 5% given the observed death data, enrollment will be paused until the IDMC reviews the data. Details of the Bayesian toxicity monitoring plan and the operating characteristics can be found in Appendix E.

7.14 Changes in the Statistical Analysis Plan

Summary of Changes from Version 2 to Version 3 of the SAP		
Location	***Description***	***Rationale***
Throughout protocol	TAK-573 was changed to modakafusp alfa.	To align with the protocol.
Title	"Relapsed" was added.	To align with the protocol.
Section 4.3 Section 4.4 Section 5.3 Section 6.0 Section 7.9.2 Section 7.10 Section 7.13	Part 3 extension study design, objectives, endpoints, sample size and analysis approaches were added to this study	To align with the protocol
Section 4.1 Section 4.4 Section 5.1	Part 1 escalation exploratory objectives and endpoints were clarified	To align with the protocol
Section 4.2 Section 4.4 Section 5.2	Part 2 expansion secondary objectives and endpoints were clarified Part 2 expansion exploratory objectives and endpoints were clarified	To align with the protocol
Section 5 Section 7.9	Clarified the definition of duration of response	To align with the protocol
Section 7.14	Extended loss-to-follow-up rule for PFS and TTP analysis was added	To provide details in the derivation algorithm

Section 7.4	The definition of refractory was updated	To align with the definition in the protocol
Section 7.9	Best unconfirmed response analysis was removed; The algorithm of response confirmation was updated	To provide details in the derivation algorithm
Section 7.11.3	Additional ADA analyses were added under immunogenicity analysis	To include additional ADA analyses for assessment
Section 7.9.1	The analysis approach for time to response was changed from time-to-event analysis to descriptive analysis	To evaluate the endpoint in an appropriate approach
Appendix A	The listing of DLT-like events was added	To include for evaluation of DLT-like events
Appendix C	The algorithms of DLT-like events were added	To provide details in the derivation algorithm

Summary of Changes from Version 1 to Version 2 of the SAP		
Location	***Description***	***Rationale***
Section 6	Changed planned sample size for phase 2	To align sample size with change in study endpoints.
Section 7.2	Added response-evaluable analysis set	To align statistical content with change in study endpoints.
Section 7.4	Added the description of the measurable disease at baseline	To clarify the algorithm to define patients with measurable disease at baseline
Section 7.8	Added clinical benefit rate and disease control rate as a secondary endpoint for phase 1 and as a primary endpoint for phase 2	To adjust endpoints to be more relevant in this early-phase study.
Section 7.8	Removed overall survival as a secondary endpoint	To align with the protocol and reflect determination that the endpoint of overall survival is not informative without a comparator group.
Section 7.8	Added the definition of confirmed response and the scenarios to claim a confirmed response	To clarify the algorithm of confirmed responses.

Section 7.11	Added description of analysis of AESI	To align with the protocol and monitor the AESI.
Section 7.11	Added oxygen saturation as a vital sign	To align with the protocol.
Section 7.12	Added criteria for early termination	To clarify conditions for early termination and clarify that no interim analysis is planned.
Appendix B	Added imputation rules for partial dates	To clarify the imputation rules for different types of partial dates.

8.0 REFERENCES

1. Simon R (1989), "Optimal two-stage designs for phase II clinical trials", Controlled Clinical Trials,10(1):1-10.
2. New NCCN Clinical Practice Guidelines in Oncology - Multiple Myeloma. Version 2.2019 National Comprehensive Cancer Network. November 16, 2018.
3. Rajkumar SV, Dimopoulos MA, Palumbo A, Blade J, Merlini G, Mateos MV, et al. (2014), "International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma" Lancet Oncology(2014),15(12):e538-48.
4. Brookmeyer R, Crowley J (1982), "A confidence interval for median survival time", Biometrics, Vol 38(1):29-41
5. Lee, J. J., & Liu, D. D. (2008), "A predictive probability design for phase II cancer clinical trials", Clinical trials, 5(2), 93-106.
6. Kumar, S., Paiva, B., Anderson, K. C., Durie, B., Landgren, O., Moreau, P., ... & Avet-Loiseau, H. (2016). International Myeloma Working Group consensus criteria for response and minimal residual disease assessment in multiple myeloma. The lancet oncology, 17(8), e328-e346.
7. Rajkumar, S. V., Harousseau, J. L., Durie, B., Anderson, K. C., Dimopoulos, M., Kyle, R., ... & International Myeloma Workshop Consensus Panel 1. (2011). Consensus recommendations for the uniform reporting of clinical trials: report of the International Myeloma Workshop Consensus Panel 1. Blood, The Journal of the American Society of Hematology, 117(18), 4691-4695.
8. Durie, B. G., Harousseau, J. L., Miguel, J. S., Blade, J., Barlogie, B., Anderson, K., ... & Rajkumar, S. V. (2006). International uniform response criteria for multiple myeloma. Leukemia, 20(9), 1467-1473.
9. Van Hout, B., Janssen, M.F., Feng, Y., Kohlmann, T., Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value Health. 2012;15(5):708-15.

10. Dolan P. Modeling valuations for EuroQol health states. Med Care. 1997;35(11):1095-108.
11. National Institute for Health and Care Excellence. Position statement on use of the EQ-5D-5L value set for England (updated October 2019). Available at: https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/technology-appraisal-guidance/eq-5d-5l. 2019.
12. EORTC QLQ-C30 Scoring Manual. Third edition, 2001. Copyright © 1995, 1999, 2001 EORTC, Brussels. D/2001/6136/001. ISBN 2-9300 64-22-6.
13. Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. J Clin Oncol 1998;16:139–44.
14. Cella D, Pickard AS, Duh MS, et al. Health-related quality of life in patients with advanced renal cell carcinoma receiving pazopanib or placebo in a randomized phase III trial. Eur J Cancer. 2012 Feb;48(3):311-23.
15. Kvam AK, Fayers PM, Wisloff F. Responsiveness and minimal important score differences in quality-of-life questionnaires: a comparison of the EORTC QLQ-C30 cancer-specific questionnaire to the generic utility questionnaires EQ-5D and 15D in patients with multiple myeloma. Eur J Haematol. 2011;87(4):330-337.
16. Stewart AK, Dimopoulos MA, Masszi T, et al. Health-Related Quality-of-Life Results From the Open-Label, Randomized, Phase III ASPIRE Trial Evaluating Carfilzomib, Lenalidomide, and Dexamethasone Versus Lenalidomide and Dexamethasone in Patients With Relapsed Multiple Myeloma. J Clin Oncol. 2016 Nov 10;34(32):3921-3930.
17. Sully K, Trigg A, Bonner N, et al. Estimation of minimally important differences and responder definitions for EORTC QLQ-MY20 scores in multiple myeloma patients. Eur J Haematol. 2019;103(5):500-509. doi:10.1111/ejh.13316
18. Coens C, Pe M, Dueck AC, Sloan J, Basch E, Calvert M, Campbell A, Cleeland C, Cocks K, Collette L, Devlin N, Dorme L, Flechtner HH, Gotay C, Griebsch I, Groenvold M, King M, Kluetz PG, Koller M, Malone DC, Martinelli F, Mitchell SA, Musoro JZ, O'Connor D, Oliver K, Piault-Louis E, Piccart M, Quinten C, Reijneveld JC, Schürmann C, Smith AW, Soltys KM, Taphoorn MJB, Velikova G, Bottomley A; Setting International Standards in Analyzing Patient-Reported Outcomes and Quality of Life Endpoints Data Consortium. International standards for the analysis of quality-of-life and patient-reported outcome endpoints in cancer randomised controlled trials: recommendations of the SISAQOL Consortium. Lancet Oncol. 2020 Feb;21(2):e83-e96. doi: 10.1016/S1470-2045(19)30790-9. PMID: 32007209

Appendix A By-Subject Listings:

In addition to the analysis outputs outlined above in the main text, separate by-patient listings will also be generated to include the following information. This includes but is not limited to:

- Disposition (date of first dose, date of last dose, number of cycles, reason for discontinuation of study treatment).
- Populations.
- Significant protocol deviations.
- Demographics.
- Study drug exposure.
- TEAEs.
- TEAEs leading to study drug discontinuation.
- Serious AEs.
- All deaths.
- DLTs during Cycle 1 in dose escalation.
- Pharmacokinetic concentrations.
- ADA.
- Efficacy (response assessments, All time to events).

Appendix B Date Imputation Rules

Incomplete Dates in the Screening Period

1. If only the day-component is missing, the first day of the month will be used if the year and the month are the same as those for the first dose of study drug. Otherwise, the fifteenth will be used.
2. If only the year is present, and it is the same as the year of the first dose of study drug, the fifteenth of January will be used unless it is later than the first dose, in which case the date of the first of January will be used.
3. If only the year is present, and it is not the same as the year of the first dose of study drug, the fifteenth of June will be used.

Incomplete Adverse Event Onset Date

Assumption: For on-study Adverse Events.

If *year* is missing (or completely missing): set to the date of first dose.

If (*year* is present and *month* and *day* are missing) or (*year* and *day* are present and *month* is missing):

If *year* = year of first dose: set the date to the first dose date.

If *year* < year of first dose: set *month* and *day* to December 31st.

If *year* > year of first dose: set *month* and *day* to January 1st.

If *month* and *year* are present and *day* is missing:

If *year* = year of first dose, and:

If *month* = month of first dose: set *day* to day of first dose.

If *month* < month of first dose: set *day* to last day of *month*.

If *month* > month of first dose: set *day* to 1st day of *month*.

If *year* < year of first dose: set *day* to last day of month.

If *year* > year of first dose: set *day* to 1st day of month.

For all other cases: set to date of first dose.

Incomplete Adverse Event Resolution Date

Assumption: For on-study Adverse Events.

If *day* is missing but *month* and *year* are non-missing, impute as the earliest of:

- Last day of the *month*
- Data cutoff date
- Death date

If *day* and *month* are missing, impute as the earliest of:

- December 31st
- Data cutoff date
- Death date

If date is completely missing (i.e. AE is ongoing), impute as earliest of:

- Data cutoff date

- Treatment discontinuation date (i.e. last dose date) + 30 days
- Death date

Incomplete Concomitant Medication Start Date/Prior Therapy Start Date/Progression Date during Prior Therapy

If *year* is missing (or completely missing): do not impute.

If (*year* is present and *month* and *day* are missing) or (*year* and *day* are present and *month* is missing):

Set *month* and day to January 1st.

If *year* and *month* are present and *day* is missing:

Set *day* to 1st day of month.

Incomplete Concomitant Medication End Date

If *year* is missing (or completely missing): do not impute.

If (*year* is present and *month* and *day* are missing) or (*year* and *day* are present and *month* is missing):

Set *month* and day to December 31st.

If *year* and *month* are present and *day* is missing:

Set *day* to last day of the month.

Incomplete Subsequent Anti-Cancer Therapy Start Date

If *year* is missing (or completely missing): set to date of last dose of study treatment + 1

If (*year* is present and *month* and *day* are missing) or (*year* and *day* are present and *month* is missing):

If *year* > year of the last dose: Set *month* and *day* to January 1st.

If *year* = year of the last dose: Set *month* and *day* to date of last dose of study treatment +1.

If *year* and *month* are present and *day* is missing:

Set *day* to 1st day of month if the resulting imputed date is greater than date of last dose.

Otherwise set the imputed date to date of last dose + 1.

Appendix C Algorithms of DLT-like Events

Wording from Protocol Section 8.2	Programmatic algorithm
1) Nonhematologic TEAEs of NCI CTCAE Grade ≥3 clearly unrelated to the underlying disease and occurring during the first cycle will be considered DLTs (see Section 10.2 for relatedness guidance), **with the following exceptions:**	For DLT-like events from AE data, check for Grade 3 or higher drug-related nonhematologic TEAEs, i.e. AEs with AETOXGR in ("GRADE 3" "GRADE 4" "GRADE 5") and AEREL="RELATED", then remove the exceptions from 1b to 1e For nonhematologic DLT-like events from laboratory data see instructions in 1a.
a. Asymptomatic laboratory changes (**other than renal and hepatic laboratory values, and Grade 4 lipase/amylase**) that can be successfully supplemented (reversion of Grade 4 events to Grade ≤2, reversion of Grade 3 events to Grade ≤1 or baseline) within 72 hours.	1) Pull out any Grade 3 or higher renal and hepatic chemistry lab values i.e. where LBTOXGR in ("GRADE 3" "GRADE 4" "GRADE 5") and LBTESTCD in ("CREAT" "BUN" "BILI" "ALT" "AST"). All Grade 3 or higher results are considered DLT-like events. 2) Pull out any Grade 3 or higher chemistry lab values not included in point 1, i.e. where LBTOXGR in ("GRADE 3" "GRADE 4" "GRADE 5") and LBCAT="CHEMISTRY" and LBTESTCD not in ("CREAT" "BUN" "BILI" "ALT" "AST"), and keep the records that don't return to baseline (or go from Grade 3 to 1, or from grade 4 to 2) within 3 days. These results are considered DLT-like events. Note: • Exclude pre-existing abnormalities, e.g. if a patient has Grade 3 ALT at baseline and post-baseline, this is not considered a DLT-like event since the abnormality was present at baseline. • Renal includes: creatinine, blood urea nitrogen • Hepatic includes: bilirubin, ALT, AST • Lipase and amylase are not collected
b. Grade 3 nausea/vomiting that can be managed subsequently with anti-emetics (Grade 3 nausea or vomiting that persists	AEDECOD in ("Nausea" "Vomiting") and AETOXGR in ("GRADE 3") and AEDUR <= 2

beyond 48 hours with or without appropriate medical intervention will be considered a DLT).	
c. Grade 3 fatigue lasting less than 72 hours.	AEDECOD in ("Fatigue") and AETOXGR in ("GRADE 3") and AEDUR <= 3
d. Grade 3 elevation of ALT or AST that resolves to Grade ≤1 or baseline within 7 days.	AEDECOD in ("Alanine aminotransferase increased" "Aspartate aminotransferase increased") and AETOXGR in ("GRADE 3") and AEDUR <=7 and AEOUT="RECOVERED/RESOLVED"
e. Grade 3 IRR that responds to symptomatic treatment, without recurrence of Grade 3 symptoms.	IRRFL="Y" and AETOXGR in ("GRADE 3") Note: IRR events and signs and symptoms are considered for this.
2) The following hematologic TEAEs of Grade ≥3 clearly unrelated to the underlying disease and occurring during the first cycle will be considered DLTs:	For DLT-like events from AE data, check for drug-related hematologic TEAEs then search for the following event in 2a. For hematologic DLT-like events from laboratory data see instructions in 2e.
a. Grade ≥3 hemolysis.	AEDECOD in ("Hemolysis") and AETOXGR in ("GRADE 3" "GRADE 4" "GRADE 5")
b. Grade 4 neutropenia lasting more than 7 consecutive days	N/A – see 2e.
c. Grade 4 thrombocytopenia lasting more than 14 consecutive days.	N/A – see 2e.
d. Grade 3 thrombocytopenia with clinically significant bleeding.	N/A – see 2e.
e. Any other Grade ≥4 hematologic toxicity with the following exception: a. Grade 4 lymphopenia	Check for Grade 4 or higher hematology lab values, except lymphocytes per the exclusion note i.e. LBTOXGR in ("GRADE 4" "GRADE 5") and LBCAT="HEMATOLOGY" and LBTESTCD not in ("LYM"). Apply the following to the remaining parameters: 1) For all parameters except platelets and neutrophils i.e. LBTESTCD not in ("NEUT" "PLAT"), any occurrence of a Grade 4 or higher result will be considered a DLT-like event. 2) For platelets and neutrophils, i.e. where LB.LBTESTCD in ("NEUT" "PLAT"),

	keep the record if the patient's result doesn't return to the following prior to or including the next scheduled dose date (use exposure dates for this): Platelets >= 50 10^9/L Neutrophils >= 1 10^9/L **Scheduled dose date** = date the next dose should have occurred, e.g. schedule D would be 28 days after the previous dose.
3) An incomplete recovery from treatment-related toxicity causing a >2-week delay in the next scheduled infusion before the initiation of Cycle 2 (or next cycle if after cycle 1) will be considered a DLT.	Any TEAEs with AEREL="RELATED" and AEACN="DOSE DELAYED" and the next dose was delayed for >14 days

Assumptions:

1) Hematologic AEs include the following PTs:

Hematology collection from protocol	Preferred term
Hematocrit	"Anaemia" "Hemolysis"
Hemoglobin	"Anaemia" "Hemolysis"
Leukocytes with differential	"Leukopenia" "White blood cell count decreased"
Neutrophils (ANC)	"Neutropenia" " Neutrophil count decreased" "Febrile neutropenia"
Platelet (count)	"Thrombocytopenia" "Platelet count decreased"
Lymphocytes	"Lymphopenia" "Lymphocyte count decreased"

Note: PTs may change with coding updates.

2) Only include TEAEs which are related to study drug.
3) Duration is calculated in days, not hours e.g. 48 hours = 2 days (This would check the day of AE and the next day).
4) For 'Clinically significant bleeding', this is included in the non-hematologic AE part. Considering 'clinically significant' as Grade 3 or higher. No need to check for the thrombocytopenia, that is checked via the laboratory data.

Appendix D Operating Characteristics of the Futility Stopping Rules

Based on the proposed two interim futility analyses criteria (Section 7.13), the design operating characteristics were calculated for each of the treatment arms under different true ORR scenarios.

Appendix D Table 1 displays the corresponding operating characteristics for the treatment arm of 240 mg, including cumulative probability of terminating the arm early at each interim analysis, average number of patients enrolled and followed up for 3 months, average number of responders observed, and probability of success (PoS) at primary analysis. As there is no plan for efficacy claim based on any of the interim analyses, the only time to claim a positive efficacy conclusion (success for each arm) will be at the primary analysis. The success criterion will be a minimal of 33 responders among 118 treated patients, achieving the goal of lower limit of 95% CI above 20% uninteresting rate.

Appendix D Table 1 Operating Characteristics for the Treatment Arm of 240 mg

True ORR	Cumulative probability of terminating the arm early at		PoS at primary analysis	Average number of patients enrolled and followed up for 3 months	Average number of observed responders
	1st IA for Futility	2nd IA for futility			
10%	82%	99%	0%	21	2
20%	40%	70%	2%	56	11
30%	13%	19%	64%	101	30
35%	6%	8%	89%	111	39
40%	3%	3%	97%	115	46
50%	<1%	<1%	100%	118	59

IA: interim analysis; ORR: objective response rate; PoS: probability of success.

As shown in Appendix D Table 1, when the true ORR is 20%, there is about 40% chance to terminate the 240 mg arm at the first interim futility analysis; about 70% chance to terminate the arm at either the first or second interim futility analysis; probability of success at primary analysis is around only 2%. When the true response rate is 35%, the cumulative probabilities of terminating the arm are reduced to 6% and 8% at the first and second interim futility analyses, the probability of success at primary analysis is 89%.

A similar operating characteristics table was created for the treatment arm of 120 mg (Appendix D Table 2**Appendix D**).

Appendix D Table 2 Operating Characteristics for the Treatment Arm of 120 mg

True ORR	Cumulative probability of terminating the arm early at		PoS at primary analysis	Average number of patients enrolled and followed up for 3 months	Average number of observed responders
	1st IA for Futility	2nd IA for futility			
10%	55%	99%	0%	30	3
20%	17%	66%	2%	67	13
30%	4%	13%	68%	108	32
35%	1%	4%	93%	115	40
40%	1%	1%	99%	117	47
50%	<1%	<1%	100%	118	59

IA: interim analysis; ORR: objective response rate; PoS: probability of success

Similarly, as shown in Appendix D Table 2, when the true ORR is 20%, there is about 17% chance to terminate the 120 mg arm at the first interim futility analysis; about 66% chance to terminate the arm at either the first or second interim futility analysis; PoS at primary analysis is around only 2%. When the true response rate is 35%, the cumulative probabilities of terminating the arm are reduced to 1% and 4% at the first and second interim futility analyses, the PoS at primary analysis is about 93%.

Appendix E Statistical Guidance on Unacceptable Toxicity and Treatment Related Death

The Bayesian toxicity monitoring approach based on posterior probability will be implemented to each DMC data review up to the second interim futility analysis to continuously monitor the unacceptable toxicity (particularly for Grade ≥4 nonhematologic treatment-related TEAEs) and treatment-related death within each treatment arm separately at various time points.

These toxicity monitoring rules can be translated to toxicity boundaries in terms of the number of patients with Grade ≥4 nonhematologic treatment-related TEAEs and the number of treatment related death among all patients treated for each arm at each safety review. Appendix I Table 1 illustrates the toxicity boundaries with the expected numbers of treated patients for each IDMC safety review based on our most recent enrollment projection. The actual toxicity boundaries will be recalculated based on the observed number of treated patients at each safety review using the same posterior probability stopping rule.

Appendix E Table 1 Bayesian Toxicity Monitoring Boundaries for Grade ≥4 Nonhematologic Treatment-Related TEAEs and Treatment Related- Deaths for Each Treatment Arm

Number of treated patients per arm	Unacceptable toxicity if the number of patients with Grade ≥4 nonhematologic treatment-related TEAEs	Unacceptable toxicity if the number of patients with treatment-related death
10	≥4	≥1
35 [a]	≥11	≥3
65 [a]	≥20	≥5

[a] The expected numbers of patients treated at IA1 and IA2 based on most recent enrollment projection.

Based on the toxicity boundaries presented in Appendix E Table 1, the corresponding operating characteristics for Grade ≥4 nonhematologic treatment-related TEAE monitoring for each treatment arm and the cumulative probability of early detection of unacceptable toxicity under different true toxicity rates with the expected number of patient treated at each safety review are displayed in Appendix E Table 2.

Appendix E Table 2 Operating Characteristics for Monitoring Grade≥4 Nonhematologic Treatment-Related TEAEs within Each Treatment Arm

True Toxicity rate	Cumulative probability of early detection with number of treated patients		
	10 patients	35 patients	65 patients
15%	5%	6%	6%
20%	12%	16%	17%
25%	22%	35%	39%
30%	35%	58%	67%
35%	49%	78%	88%

As shown in Appendix E Table 2, when the true toxicity rate is as low as 15%, there are low cumulative probabilities of 5% to 6% for early calls of unacceptable toxicity at each safety review; when the true toxicity rate is as high as 35%, there are about 49% to 88% cumulative probabilities of early call for unacceptable toxicity at each safety review. When the true DLT rate is 25% (the maximum toxicity allowed), there is about 39% cumulative probability to have an early call for unacceptable toxicity.

Similar operating characteristics table was created for monitoring of treatment-related death for each treatment arm, shown in Appendix E Table 3.

Appendix E Table 3 Operating Characteristics for Monitoring Treatment-Related Death within Each Treatment Arm

True Toxicity rate	Cumulative probability of early detection with number of treated patients		
	10 patients	35 patients	65 patients
1%	10%	10%	10%
3%	26%	29%	30%
5%	40%	48%	52%
7%	52%	64%	71%
9%	61%	76%	85%

As shown in Appendix E Table 3, when the true treatment-related death rate is as low as 1%, there is about 10% cumulative probability of early call of unacceptable toxicity at each safety review; when the true treatment related death rate is as high as 9%, there are about 61% to 85% cumulative probabilities of early toxicity call at each safety review. When the true treatment-related death rate is 5% (the maximum toxicity allowed), there is about 52% cumulative probability to alarm the unacceptable toxicity early.

Appendix F Preparation of Patient-Reported Outcomes Data for Analyses (Part 3 only)

During the (extension) Part 3 of the study, the EORTC QLQ-MY20 and the EQ-5D-5L will be administered electronically per the SOE (Section **Error! Reference source not found.** in the protocol) to all enrolled and treated patients. The PRO assessments will be collected regularly during treatment and at the EOT visit: PROs are administered on D8 of each cycle, except for C1 when the EORTC QLQ-MY20 is administered on D1 and D8 (i.e., C1D1 and C1D8) and the EQ-5D-5L on D1 (i.e., C1D1). On C1D1, PROs may be completed within a +1-day window; on EOT visit, PROs may be completed within a +10-day window; and on the rest of days, PROs may be completed within a ±2-day window.

Given PROs are administered at every cycle during treatment, each PRO assessment is then associated with a specific Cycle and Day number (e.g., Cycle X, Day Y).

Per SOE in protocol, each PRO assessment should be assigned a cycle label (Cycle X, Day Y), using 'dates of dosing at each cycle' and 'dates of ePRO assessments throughout the study period', based on when the respective ePRO assessment date falls within the following windows during each cycle:

PRO	Cycle and Day label	Window	Algorithm
EORTC-QLQ-MY20	Cycle 1 Day 1	+ 1 day	PRO between the C1D1 dose date and the C1D1 dose date+1
	Cycle 1 Day 8	± 2 days	PRO between the C1D1 dose date+5 and the C1D1 dose date+9
	Cycle 2 Day 8– Cycle 48* Day 8	± 2 days	PRO between the CxD1 dose date+5 and the CxD1 dose date+9
	EOT	+ 10 days	PRO between the EOT visit date and EOT visit date+10
EQ-5D-5L	Cycle 1 Day 1	+ 1 day	PRO between the C1D1 dose date and the C1D1 dose date+1
	Cycle 2 Day 8 – Cycle 48* Day 8	± 2 days	PRO between the CxD1 dose date+5 and the CxD1 dose date+9
	EOT	+ 10 days	PRO between the EOT visit date and EOT visit date+10

*Or, max number of cycles observed

Only one PRO assessment should be observed for each measure per SOE during each cycle (e.g., one for MY20, and one for EQ-5D); however, if more than one assessment for either the MY20 or EQ-5D appears in a single window, use the ePRO assessment that appears the closest to SOE

assessment date: i.e., Cycle 1 Day 1 or Cycle 1 Day 8, and Cycle 2+ Day 8, for the visit ID remapping.

Descriptive summaries and statistical analyses of PRO data will only include observations that fall within the specified "windows," while the data that falls outside the specified windows will be only reported in the listings.

Signature Page for TAK-573-1501 SAP v4.0_15Mar2024
Title:

Approval	[redacted] Statistics 15-Mar-2024 13:28:58 GMT+0000

Document Number: TDN-000248278 v1.0