Statistical Analysis Plan Amendment 2

Study ID: 207495

Official Title of Study: A Phase III, Open-Label, Randomized Study to Evaluate the Efficacy and Safety of Single Agent Belantamab Mafodotin Compared to Pomalidomide Plus Lowdose Dexamethasone (Pom/Dex) in Participants With Relapsed/Refractory Multiple Myeloma (RRMM) (DREAMM 3)

NCTID: NCT04162210

Date of Document: 02 September 2022

Information Type:	Statistical Analysis Plan (SAP)

TITLE PAGE

Protocol Title: A Phase III, Open-Label, Randomized Study to Evaluate the Efficacy and Safety of Single Agent Belantamab Mafodotin Compared to Pomalidomide plus Low-dose Dexamethasone (pom/dex) in Participants with Relapsed/Refractory Multiple Myeloma (RRMM) (DREAMM 3)

Study Number: 207495

Compound Number: GSK2857916

Abbreviated Title: Phase III Study of Single Agent Belantamab Mafodotin versus Pomalidomide plus Low-dose Dexamethasone in Participants with Relapsed/Refractory Multiple Myeloma

Acronym: DREAMM3

Sponsor Name: GlaxoSmithKline Research & Development Limited

Regulatory Agency Identifier Number(s)

Registry	**ID**
Clinicaltrials.gov	NCT04162210
EudraCT	2018-004252-38
IND	119333

TABLE OF CONTENTS

PAGE

Version history

This Statistical Analysis Plan (SAP) amendment 2 for study 207495 is based on the protocol amendment 4 (Version: GSK Document Number TMF-14577505), Dated 20-Apr -2022.

SAP Version	Approval Date	Protocol Version (Date) on which SAP is Based	Change	Rationale
SAP	11-DEC-2019	2017N336101_00 (30-SEP-2019)	Not Applicable	Original version
SAP amendment 1		TMF-14577505 (20-Apr-2022)	1) Migrated the SAP to new SAP template released on March 2021. Sections/Section number were updated accordingly. Added Section 6 supporting documentation to the SAP. 2) Section 1.2 updated to incorporate changes made in protocol amendment 2 and protocol amendment 3 including End of Study definition 3) Section 2.1 updated to incorporate changes to the information fraction for the interim PFS analysis (and now based on first 120 participants randomized) and introduction of an additional interim OS analysis made in protocol amendment 2 4) Section 4.7 updated to incorporate changes to the event trigger for the interim PFS analysis and an additional interim OS analysis made in protocol amendment 2 5) Section 4.7.1 align text with changes made in protocol amendment 2 6) Section 5.2 updated to incorporate changes to the event trigger for the final OS analysis, estimated timing for the final OS analysis and removal of enrolment cap for North East Asia Countries made in protocol amendment 2 and protocol amendment 3 7) Section 4.2, 4.3.1, 4.3.2: Remove piecewise constant Hazard ratio analysis according to protocol amendment 2. Non-proportional hazards will be handled by Restricted Mean Survival Time (RMST) method; specified analyses based on mITT population.	Updated based on protocol Amendment 2, protocol Amendment 3 and protocol Amendment 4 Further clarification

SAP Version	Approval Date	Protocol Version (Date) on which SAP is Based	Change	Rationale
			8) Section 4.2.1: move the planed analyses to Section 4.2.2 9) Section 4.2.2: remove the sensitivity analysis of supplementary estimand 1 and 2 on investigator response 10) Section 4.2.2: Minor updates to the footnote of PFS censoring rule for more clarification; 11) Section 4.3.1.3 Added sensitivity analysis for OS censoring any subjects who started a subsequent anti-cancer treatmen 12) Section 4.3.2: Clarify the timing of analyses for secondary efficacy endpoints per updates in the protocol amendment 2; update analysis methods for TTP. 13) Section 4.3.2.1: Add response confirmation algorithm. 14) Section 4.3.2.1 Removed supplementary analyses of supportive secondary efficacy endpoints based on the mITT CCI [redacted] 16) Section 4.3.2.4 remove summary of change of fatigure domain in QLQ-C30 and IL52 by responder status CCI [redacted] 18) Section 4.5.2: Add safety analyses for COVID-19. 19) Section 4.5.3.1: Update analyses for ocular exam findings based on protocol amendment 1. 20) Section 4.5.3.2: Add KVA analyses based on protocol amendment 2. 21) Section 4.3.2.4: Update Secondary PRO analyses based on protocol amendment 2. Added more clarifications to the analyses. CCI [redacted]	

SAP Version	Approval Date	Protocol Version (Date) on which SAP is Based	Change	Rationale
			CCI expression levels in the main SAP. 23) Section 4.5.4.3: Remove summary analyses for ECG since protocol amendment 2 removed routine ECG monitoring; CCI CCI CCI 27) Section 4.5.1: Update dose intensity and duration of exposure analyses for belantamab mafodotin arm 28) Section 4.5.3.1: remove listings of intraocular pressure and dilated fundoscopic examination results 29) Section 4.6.1: Update language for Pop-PK analyses CCI 31) Section 6.2.3 updated to align treatment emergent and on treatment definition with collection of AEs per CSP CCI CCI 34) Section 4.2.3 add PFS rate at 6 months 35) Section 4.3.1 add OS rate at 6, 12 and 18 months 36) Section 5.2 update sample size determination for 90% power 37) Section 4.7 update Table 7 summary of PFS and OS analyses; remove interim analysis for PFS; update OS analyses 38) Section 1.2 study design; table for Analysis CCI	
SAP amendment 2		TMF-14577505 (20-Apr-2022)	1) PFS primary estimand to be based on investigator-assessed response. PFS based on derived	Change of primary analyses to

SAP Version	Approval Date	Protocol Version (Date) on which SAP is Based	Change	Rationale
			response removed. Similar changes for ORR, CBR, MRD negativity, DoR, TTR, TTP and [redacted] 2) Section 1.1.2 amend primary estimands and add supplementary estimands for DoR 3) Section 4.2.3 add unstratified cox model 4) 4.2.4.2 add sensitivity aanalysis on independent review committee (IRC) - assessed response [redacted]	using investigator-assessed response per FDA comments; Amend definition of DoR per FDA comments; Add additional statistical analyses;

1. INTRODUCTION

The purpose of this SAP (Version 3) is to describe the planned analyses to be included in the Clinical Study Report for Study 207495.

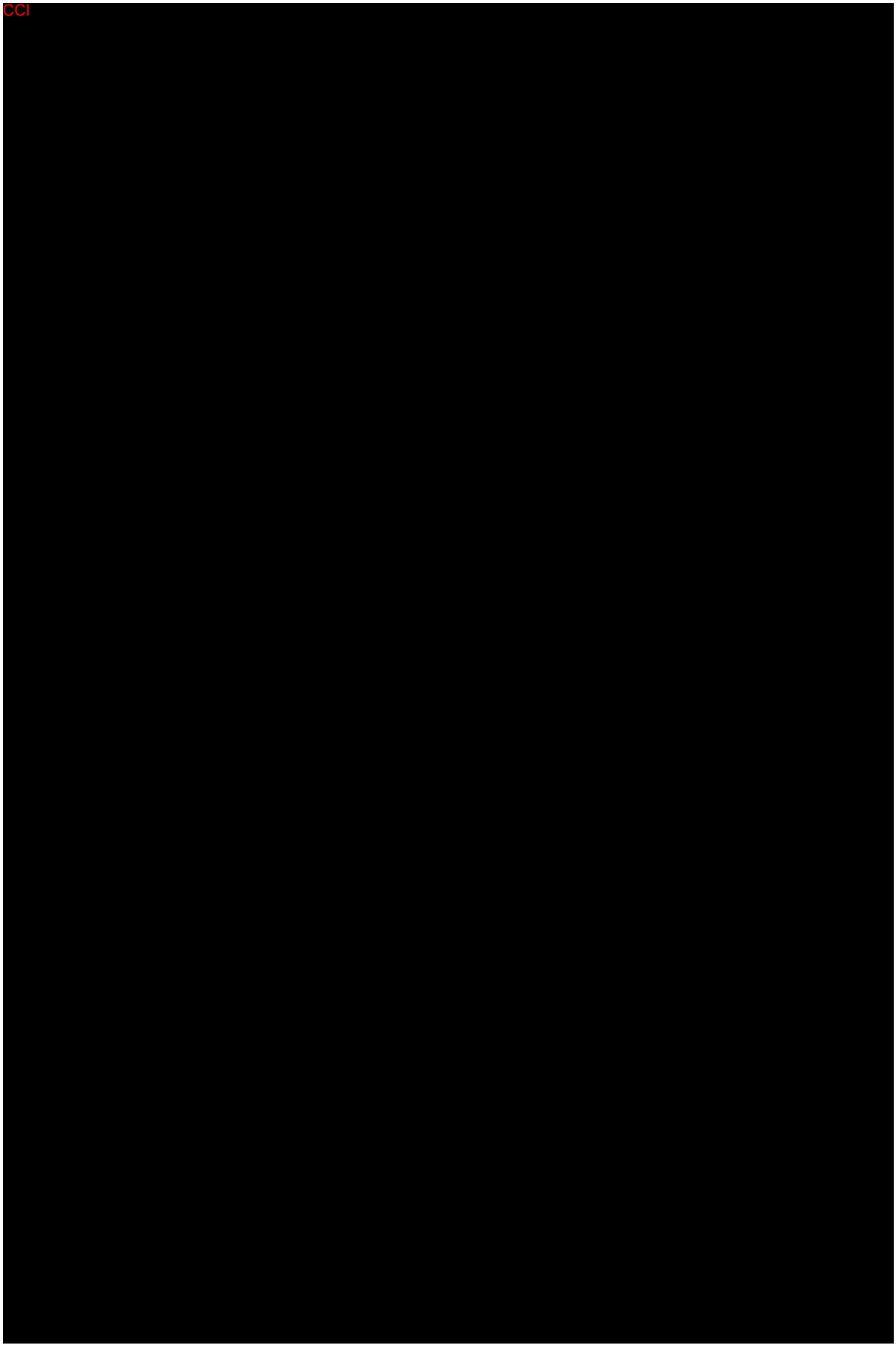
Additional detail with regards to data handling conventions and the specification of data displays will be provided in the Output and Programming Specification (OPS) document.

1.1. Objectives, Estimands and Endpoints

1.1.1. Objectives and Endpoints

Objectives	Endpoints
Primary	
• To compare the efficacy with belantamabmafodotin vs pomalidomide plus low dose dexamethasone (pom/dex) in participants with relapsed/refractory multiple myeloma (RRMM)	• PFS, defined as the time from the date of randomization until the earliest date of documented disease progression (according to IMWG Response Criteria) or death due to any cause
Key Secondary	
• To compare the overall survival with belantamab mafodotin vs Pom/Dex in participants with RRMM	• OS, defined as the time from randomization until death due to any cause
Secondary	
• To compare other markers of efficacy of belantamab mafodotin vs pom/dex in participants with RRMM	• ORR, defined as the percentage of participants with a confirmed PR or better per IMWG • Clinical benefit rate (CBR), defined as the percentage of participants with a confirmed minimal response (MR) or better per IMWG • DoR, defined as the time from first documented evidence of PR or better until PD per IMWG or death due to any cause among participants who achieve confirmed PR or better • TTR, defined as the time between the date of randomization and the first documented evidence of response (PR or better) among participants who achieve confirmed PR or better. • TTP, defined as the time from the date of randomization until the earliest date of documented PD (per IMWG Response Criteria) or death due to PD
• To evaluate the safety and tolerability of belantamab mafodotin vs pom/dex in participants with RRMM	• Incidence of adverse events (AEs) and changes in laboratory parameters • Ocular findings on ophthalmic exam
• To evaluate the pharmacokinetic profile of belantamab mafodotin	• Plasma concentrations of belantamab mafodotin, total mAb, and cys-mcMMAF
• To assess anti-drug antibodies (ADAs) against belantamab mafodotin	• Incidence and titers of ADAs against belantamab mafodotin
• To evaluate the tolerability of belantamab mafodotin vs pom/dex based on self-reported symptomatic adverse effects	• Symptomatic adverse effects as measured by the PRO-CTCAE and OSDI
• To evaluate and compare changes in symptoms and health-related quality of life (HRQOL) of belantamab mafodotin to pom/dex.	• Change from baseline of Health-related QOL as measured by EORTC QLQ-C30, EORTC IL52* and EORTC QLQ-MY20*
• To assess Minimal Residual Disease (MRD) in participants who achieve ≥VGPR or better for belantamab mafodotin vs pom/dex	• MRD negativity rate, defined as; the percentage of participants who are MRD negative by NGS method
Exploratory	

CCI



CCI [redacted]

ADA= anti-drug antibodies; AE = adverse event; CCI [redacted]; CCI [redacted]; CCI [redacted] DOR = duration of response; CCI [redacted]; EORTC-IL52 = Disease Symptoms domain of EORTC QLQMY20; EORTC QLQC30 = European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire 30 item Core module; EORTC QLQMY20 = European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire 20 item Multiple Myeloma module; CCI [redacted]; HRQoL = health related quality of life; IMWG = International Myeloma Working Group; KVA = Keratopathy Visual Acuity; MRD = minimal residual disease; NGS = Next Generation Sequencing; ORR = overall response rate; OS = overall survival; OSDI = Ocular Surface Disease Index; PD = progressive disease; PFS = progression-free survival; CCI [redacted]; pom/dex = pomalidomide/dexamethasone; PR = partial response; PROCTCAE = Patient Reported Outcomes version of the Common Terminology Criteria for Adverse Events; QOL = quality of life; RRMM = relapsed/refractory MM; SAE = serious adverse event; CCI [redacted]; TTP = time to progression; TTR -= time to response; VGPR= very good partial response

* = EORTC IL52 applies to participants enrolled under the original protocol; EORTC QLQ-MY20 applies to participant enrolled under protocol amendment 1;

CCI [redacted]

1.1.2. Estimands

Table 1 Estimands

Objective	Estimand Category	Estimand			
		Variable/ Endpoint	Analysis Set	Intercurrent Event Strategy	Population Level Summary Measure
Primary Objective: To demonstrate the superiority of belantamab mafodotin compared to pomalidomide plus low dose dexamethasone (pom/dex) in PFS in participants with relapsed/refractory multiple myeloma (RRMM) [1]	Primary	PFS	ITT	• Disease assessments between scheduled visits: treatment policy • New anti-cancer therapy: while on treatment • Extended loss to follow-up: while on treatment • Treatment discontinuation: treatment policy • Death: composite	Hazard ratio for belantamab mafodotin vs pom/dex
	Supplementary 1	PFS	ITT	• Disease assessments between scheduled visits: hypothetical • New anti-cancer therapy: while on treatment • Extended loss to follow-up: while on treatment • Treatment discontinuation: treatment policy • Death: composite	Hazard ratio for belantamab mafodotin vs pom/dex
	Supplementary 2	PFS	ITT	• Disease assessments between scheduled visits: treatment policy • New anti-cancer therapy: composite • Extended loss to follow-up: treatment policy • Treatment discontinuation: composite • Death: composite	Hazard ratio for belantamab mafodotin vs pom/dex
Key Secondary Objective: To demonstrate the superiority of belantamab mafodotin compared to pomalidomide plus low dose dexamethasone (pom/dex) in OS in participants with relapsed/refractory multiple myeloma (RRMM) [1]	Primary	OS	ITT	Survival benefit regardless of subsequent anticancer therapies or treatment compliance • New anti-cancer therapy: treatment policy • Treatment discontinuation: treatment policy	Hazard ratio for belantamab mafodotin vs pom/dex

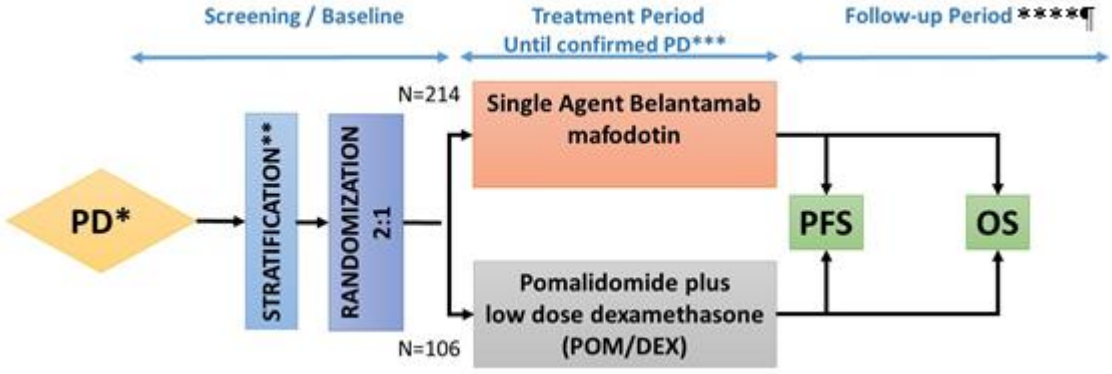
Objective	Estimand Category	Estimand			
		Variable/ Endpoint	Analysis Set	Intercurrent Event Strategy	Population Level Summary Measure
Secondary Objectives (Efficacy): To demonstrate the superiority of belantamab mafodotin compared to pomalidomide plus low dose dexamethasone (pom/dex) in ORR/CBR/TTR/TTP/DoR/MRD in participants with relapsed/refractory multiple myeloma (RRMM) [1]	Primary	ORR	ITT	• New anti-cancer therapy: while on treatment • Treatment discontinuation: treatment policy	>=PR percentage by treatment arm
		CBR	ITT	• New anti-cancer therapy: while on treatment • Treatment discontinuation: treatment policy	>=MR percentage by treatment arm
		DoR	ITT	• Disease assessments between scheduled visits: treatment policy • New anti-cancer therapy: while on treatment • Extended loss to follow-up: while on treatment • Treatment discontinuation: treatment policy • Death due to non-PD: composite • Death due to PD: composite	Summarized using the Kaplan-Meier method by treatment arm
		TTR	ITT	• New anti-cancer therapy: while on treatment • Treatment discontinuation: treatment policy	Descriptive summary of TTR by treatment arm
		TTP	ITT	• Disease assessments between scheduled visits: treatment policy • New anti-cancer therapy: while on treatment • Extended loss to follow-up: while on treatment • Treatment discontinuation: treatment policy • Death due to non-PD: while on treatment • Death due to PD: composite	Hazard ratio for belantamab mafodotin vs pom/dex
		MRD	ITT	• New anti-cancer therapy: while on treatment • Treatment discontinuation: treatment policy	MRD negativity percentage by treatment arm
	Supplementary	DoR	ITT	• Disease assessments between scheduled visits: treatment policy • New anti-cancer therapy: while on treatment • Extended loss to follow-up: while on treatment • Treatment discontinuation: treatment policy • Death due to non-PD: while on treatment • Death due to PD: composite	Summarized using the Kaplan-Meier method by treatment arm

Objective	Estimand Category	Estimand			
		Variable/ Endpoint	Analysis Set	Intercurrent Event Strategy	Population Level Summary Measure

[1]. have received at least 2 prior lines of anti-myeloma treatments, including at least 2 consecutive cycles of both lenalidomide and a proteasome inhibitor (given separately or in combination), and must have documented disease progression as defined by IMWG (a) on, or within 60 days of completion of the last therapy or (b) must be non-responsive while on last treatment, where non-responsive is defined as not achieving at least Minimal Response (MR) after 2 complete treatment cycles. In such cases lack of achieving of MR must be determined no earlier than at least 4 weeks after the last treatment

1.2. Study Design

Overview of Study Design and Key Features



* Progression (a) on, or within 60 days of completion of the last therapy or (b) must be non-responsive while on last treatment, where non-responsive is defined as not achieving at least Minimal Response (MR) after 2 complete treatment cycles

** Stratification based on International Staging System (ISS), number of prior lines of therapy and prior usage of anti-CD38 antibody treatment

*** Until progressive disease (PD), death, unacceptable toxicity, withdrawal of consent, lost to follow-up or end of study, whichever comes first

CCI [redacted]

Design Features

- This study is a Phase III, open-label, randomized, multicenter study evaluating the efficacy and safety of single agent belantamab mafodotin compared to pom/dex in participants with RRMM.
- The study will include a screening period, study treatment period, and follow-up.
  - During Screening, participants will be evaluated for study eligibility per protocol as defined in the Inclusion/Exclusion criteria (see Section 6 of Protocol Amendment 3). Eligible participants must have been previously treated with at least two prior lines of therapy, including at least 2 consecutive cycles of lenalidomide and a proteasome inhibitor (PI), (given separately or in combination) and must have documented progression (a) on, or within 60 days of completion of the last therapy or (b) must be non-responsive while on last treatment, where non-responsive is defined as not achieving at least Minimal Response (MR) after 2 complete treatment cycles. In such cases lack of achieving of MR must be determined no earlier than at least 4 weeks after the last treatment.
  - Following Screening, approximately 320 participants will be centrally randomized in a 2:1 ratio to either Arm 1 (single agent belantamab mafodotin) or Arm 2 (pom/dex), as described in Section 1.1 of Protocol Amendment 3. Participants will be stratified based on the following: previous treatment with anti-CD38 (Y/N), with a 40% global enrolment cap for participants with prior anti-CD38 treatment, stage (International Staging System [ISS]) (I/II or III), with a 55% global enrolment cap for participants with ≤3 prior lines. No cross-over will be allowed until final OS analysis.
  - During the Study Treatment Period, safety and disease assessments will be performed regularly according to the schedule of activities (Section 2 of Protocol Amendment 3) for each arm. Participants in both arms will be treated until PD, death, unacceptable toxicity, withdrawal of consent, loss to

Overview of Study Design and Key Features	
	follow-up or end of study, whichever comes first. End of study is defined in Section 5.4 of Protocol Amendment 3. Dose interruptions or reductions may be required following potential drug-associated toxicities. CCI [redacted] ○ For participants who discontinue study treatment for reasons other than PD, disease evaluations will continue to be performed every 3 weeks (±3 days) until confirmed (documented) PD, death, start of a new anticancer treatment, withdrawal of consent, or end of the study, whichever occurs first. In case of PD, participants will be followed to ascertain survival status CCI [redacted] every 12 weeks (±14 days) until withdrawal of consent, lost to follow-up, death or the end of the study. • The final PFS analysis (primary analysis) will be conducted at the time of observing approximately 151 PFS events. • A participant is considered to have completed the study if they are followed until death or end of study. • End of study is defined as when the planned 250 deaths for final OS analysis have occurred, or when all participants have died or are lost to follow up or have withdrawn consent, whichever occurs first. At the first OS interim analysis (~40% OS information fraction), the study will not be stopped regardless of the results. GSK will continue to collect the OS data to conduct the 2nd OS interim analysis to test for efficacy at ~70% OS information fraction. If OS is significant at 70% IF, OS data will continue to be collected until all participants have died, are lost to follow up, or withdrawn consent, or for 2 years after the final OS analysis at 70% IF, whichever occurs first.
Study intervention	• **Arm 1: Single agent Belantamab mafodotin.** Belantamab mafodotin will be administered intravenously at 2.5 mg/kg on Day 1 (D1) of a Q3W schedule (e.g., 21-day cycle). • **Arm 2: Pomalidomide and low-dose dexamethasone (Pom/Dex).** Pomalidomide will be administered orally at the approved starting dose of 4 mg daily on Days 1 to 21 of each 28-day cycle, with dexamethasone administered orally at a dose of 40 mg once weekly (Days 1, 8, 15, and 22). For participants over 75 years old, dexamethasone should be administered at the lower dose of 20 mg once weekly (Days 1, 8, 15, and 22).
Study intervention Assignment	• Approximately 320 participants will be randomized in a 2:1 ratio, in favor of Arm 1. Arm 1 (single agent belantamab mafodotin) will enroll approximately 214 participants and Arm 2 (pom/dex) will enroll approximately 106 participants. Participants will be stratified based on the following: previous treatment with anti-CD38 (Y/N), with a 40% global enrolment cap for participants with prior anti-CD38 treatment, stage (International Staging System [ISS]) (I/II or III), and number of prior lines of therapy (≤3 vs >3), with a 55%

<table>
<tr><th colspan="2">Overview of Study Design and Key Features</th></tr>
<tr><td></td><td>global enrolment cap for participants with ≤3 prior lines. No cross-over will be allowed until final OS analysis.</td></tr>
<tr><td>Analysis</td><td>
<table>
<tr><th>Analyses Timing from Randomization</th><th>Planned PFS analyses</th><th>Planned OS analyses</th></tr>
<tr><td>~25 months</td><td>PFS final
Approximately 151 PFS events and the first 320 randomized subjects have been followed for a minimum of 4 months[c]</td><td>OS IA for efficacy[a]
~100 OS events
~40% OS info</td></tr>
<tr><td>48 months</td><td>N/A</td><td>OS IA for efficacy[b]
~175 OS events
~70% OS info</td></tr>
<tr><td>>60 months</td><td>N/A</td><td>OS final[b]
~250 OS events</td></tr>
</table>
a. Provided that final PFS analysis is significant.
b. If PFS is significant and null hypothesis is not rejected at the 2nd OS IA for efficacy.
c. First dose of 320th subject happens around 21 months
</td></tr>
</table>

2. STATISTICAL HYPOTHESES

Primary endpoint PFS

The primary efficacy analysis will be the comparison of the distribution of PFS between the two treatment groups. Assuming proportional hazards for PFS, the following statistical hypothesis will be tested to address the primary efficacy objective at one-sided alpha level of 2.5%:

$$H_{01}: \Theta_1 \geq 1 \quad \text{VS.} \quad H_{A1}: \Theta_1 < 1$$

where, θ_1 is the PFS HR (belantamab mafodotin arm vs. pom/dex arm).

Key secondary endpoint OS

Assuming proportional hazards for OS, the following statistical hypotheses will be tested at one-sided alpha level of 2.5% if PFS is statistically significant:

$$H_{02}: \Theta_2 \geq 1 \quad \text{VS.} \quad H_{A2}: \Theta_2 < 1$$

where, θ_2 is the OS HR (belantamab mafodotin arm vs. pom/dex arm).

2.1. Multiplicity Adjustment

The family-wise type I error for this study is strongly controlled at 2.5% (one-sided). A hierarchical testing procedure is adopted and the hypothesis testing of key secondary endpoint OS will only be performed if the primary endpoint PFS is statistically significant at PFS final analysis [Bretz, 2009; Li, 2017].

The hypothesis testing for key secondary endpoint OS will be conducted provided that the primary endpoint PFS is statistically significant at PFS final analysis. The type I error for OS is strongly controlled at 2.5% (one-sided) by a group sequential design with a Lan-DeMets (O'Brien-Fleming) alpha spending function [Lan, 1983] using information fractions of (0.4, 0.7, 1) with total one-sided type I error of 2.5%. More details about the hypothesis testing and overall alpha control are provided in Section 4.7.

3. ANALYSIS SETS

Analysis Set	Definition / Criteria	Analyses Evaluated
All Screened	• The All Screened Population will consist of all participants who sign the ICF to participate in the clinical trial. Participants in this population will be used for screen failure summary.	• Study Population
Enrolled	• All participants who entered the study. • Note screening failures (who never passed screening even if rescreened) and participants screened but never enrolled into the study (Reserve, Not Used) are excluded from the Enrolled analysis set as they did not enter the study.	• Study Population
Intent-to-Treat (ITT)	• ITT Population will consist of all randomized participants whether or not randomized treatment was administered. This population will be based on the treatment and strata to which the participant was randomized and will be the primary population for the analysis of efficacy data. Any participant who receives a treatment randomization number will be considered to have been randomized.	• Study Population • Efficacy
mITT (modified ITT)	• All participants who have received at least 2 lines of prior therapies; randomized and received at least one dose of study treatment (participant randomized to the belantamab mafodotin arm but received pom/dex will be excluded and vice versa); with measurable disease at baseline[1].	• Efficacy
Safety	• All randomized participants who take at least 1 dose of study intervention (any drug component). Participants will be analyzed according to the intervention • they actually received.	• Safety
Pharmacokinetic (PK)	• All participants in the Safety analysis set who had at least 1 non-missing PK assessment (Non-quantifiable [NQ] values will be considered as non-missing values). • Data should be reported according to the actual treatment	• PK

CCI [redacted]

1. measurable disease at baseline is defined as: a patient has at least one of the following measurements: a. Serum M-protein ≥0.5 g/dL (≥5 g/L) or b. Urine M-protein ≥200 mg/24h or c. Serum FLC assay: Involved FLC level ≥10 mg/dL (≥100 mg/L) and an abnormal serum free light chain ratio (<0.26 or >1.65)

4. STATISTICAL ANALYSES

4.1. General Considerations

4.1.1. General Methodology

The Intent-to-Treat (ITT) analysis set will be used for all study population analyses and efficacy analyses, unless otherwise specified and Safety analysis set will be used for all safety analyses.

The stratified log-rank test and stratified Cox proportional hazards models will include the randomization stratification factors as “strata”. Unless otherwise specified, the stratification factors entered for randomization will be used in the primary analysis. If there is any mis-stratification, a supplementary analysis will be performed using the stratification data based on the clinical database.

Confidence intervals will use 95% confidence levels unless otherwise specified.

Unless otherwise specified, continuous data will be summarized using descriptive statistics: n, mean, standard deviation (std), median, minimum and maximum. Categorical data will be summarized as the number and percentage of participants in each category.

For efficacy analyses, only central lab values will be used.

4.1.2. Baseline Definition

For all endpoints, unless otherwise specified, the baseline value will be the latest pre-dose assessment with a non-missing value, including those from unscheduled visits. If time is not collected, Day 1 assessments are assumed to be taken prior to first dose and used as baseline. For participants who did not receive study treatment during the study, baseline will be defined as the latest, non-missing collected value.

For laboratory data, baseline will be the latest non-missing pre-dose value from central lab. If no central lab value is available, the latest non-missing pre-dose value from local lab will be used. For efficacy lab tests, only central lab values will be used.

For ECG analyses, if the latest, non-missing pre-dose values is from triplicate, the subject level baseline is defined as the mean of triplicate baseline assessments.

Unless otherwise stated, if baseline data is missing no derivation will be performed and baseline will be set to missing.

4.2. Primary Endpoint Analyses

4.2.1. Definition of endpoint

Progression-Free-Survival (PFS)

PFS is the primary endpoint of this study; it is defined as the time from randomization until the earliest date of PD, or death due to any cause. The analyses of PFS will be based on the ITT Analysis Set, unless otherwise specified.

4.2.2. Planned analysis of PFS

Overview of the planned analyses

- Primary analysis of PFS will be based on investigator-assessed response (per IMWG Kumar, 2016)
- Handling of intercurrent events for primary estimand (Table 1) are specified as the primary event and censoring rules in Table 2; Handling of intercurrent events for supplementary estimands 1&2 (Table 1) are specified as the alternative event and censoring rules 1&2 in Table 2, respectively;

The following sets of analyses will be conducted:

[1]. Primary analysis of primary estimand (investigator-assessed response + primary censoring rules);

[2]. Primary analysis of supplementary estimand 1 (investigator-assessed response + alternative censoring rules 1);

[3]. Primary analysis of supplementary estimand 2 (investigator-assessed response + alternative censoring rules 2);

Table 2 Assignments for Primary and Alternative Progression and Censoring Dates for PFS Analysis

Situation	Date of Event (Progression/Death) or Censoring	Outcome Event (Progression/Death) Or Censored
No (or inadequate) baseline tumor assessments [1] and the subject has not died (if the subject has died follow the rules for death indicated at the bottom of the table)	Randomization	Censored
No post-baseline assessments and the subject has not died (if the subject has died follow the rules for death indicated at the bottom of the table)	Randomization	Censored
Progression documented at scheduled visits and Progression documented without extended loss-to-follow-up time [4]	Date of assessment of progression	Event
Progression documented between scheduled visits and Progression documented without extended loss-to-follow-up time [4]	Date of assessment of progression	Event
	Date of next scheduled response assessment [SA1]	*Event [SA1]*
No progression (or death)	Date of last 'adequate' assessment of response [2]	Censored
New anticancer treatment started (prior to documented disease progression or death) [3].	Date of last 'adequate' assessment of response [2] *(on or prior to starting anti-cancer therapy)*	Censored
	Date of starting new anticancer therapy [SA2]	*Event [SA2]*
Death without extended loss-to-follow-up time [4]	Date of death	Event
Death or progression after an extended loss-to-follow-up time [4]	Date of last 'adequate' assessment of response [2] prior to PD/death (prior to missed assessments)	Censored
	Date of death or progression [SA2]	*Event [SA2]*
Death or progression after an extended loss-to-follow-up time [4] from randomization	Date of randomization	Censored
Treatment discontinuation due to clinical PD before PD or death [SA2]	*Date of treatment discontinuation [SA2]*	*Event [SA2]*

[SA1]. Alternative rule 1 for handling of intercurrent events for supplementary estimand of PFS

[SA2]. Alternative rule 2 for handling of intercurrent events for supplementary estimand of PFS

[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 (≥5 g/L) or b. Urine M-protein ≥200 mg/24h or c. Serum FLC assay: Involved FLC level ≥10 mg/dL (≥100 mg/L) and an abnormal serum free light chain ratio (<0.26 or >1.65)

[2]. An adequate assessment is defined as an assessment where the response is sCR, CR, VGPR, PR, MR, or SD.

[3]. If PD or death and New anti-cancer therapy occur on the same day assume the progression or death was documented first (e.g., outcome is progression or death and the date is the date of the assessment of progression or death). If anti-cancer therapy is started prior to any adequate assessments, censoring date should be the date of randomization.

[4]. Extended loss-to-follow-up time = 6 weeks + 7 day window = 49 day window; Without extended loss-to-follow-up time is defined as: <= 49 days; after an extended loss-to-follow-up time is defined as: >49 days. More details about extended loss-to-follow-up time are provided in Section 6.2.9.

4.2.3. Main Analytical Approach

Progression-Free-Survival (PFS)

The non-parametric Kaplan-Meier method will be used to estimate the survival curves for PFS. Kaplan-Meier plots of PFS will be presented by treatment arm. Kaplan-Meier estimates for the median PFS, the first and third quartiles, and 6-month PFS rate will be presented, along with 95% CIs. CIs for quartiles will be estimated using Brookmeyer-Crowley method [Brookmeyer, 1982]. The treatment difference in PFS will be compared by the stratified log-rank test at one-sided alpha level of 0.025. The stratified log-rank test (stratified by randomization factors) will only be performed for the primary analysis of primary estimand of PFS (i.e. based on investigator-assessed response and primary event and censoring rules) based on ITT Analysis Set.

Hazard ratio (HR) and its corresponding 95% CI will be estimated from Cox proportional hazard model stratified by randomization factors with treatment arm as the sole explanatory variable. The Cox models will be fitted using SAS PROC PHREG with the Efron method to control for ties.

Hazard ratio (HR) and its corresponding 95% CI will also be estimated from unstratified Cox proportional hazard model with treatment arm as the sole explanatory variable. The Cox models will be fitted using SAS PROC PHREG with the Efron method to control for ties.

Stratification factors entered for randomization will be used in the primary analysis. If there is any mis-stratification, supplementary analyses will be performed using the stratification data based on the clinical database.

Statistical Methodology Specification

Endpoint / Variables
• PFS
Model Specification
• PFS will be analyzed across treatment arms using Kaplan-Meier analysis (PROC LIFETEST). • 95% Confidence intervals will be estimated using the Brookmeyer-Crowley method [Brookmeyer, 1982]. • The treatment difference in PFS will be tested by the stratified log-rank test. • A stratified Cox proportional hazard model with Efron's method of tie handling and treatment arm as the sole explanatory variable will be used to assess the magnitude of the treatment difference (i.e., the hazard ratio) in PFS between the treatment arms.
Model Checking & Diagnostics

• The proportional hazards assumption will be assessed using the following methods:
  ○ Kaplan-Meier plot by treatment arm
  ○ Plot of log(time) against log(-log(survival)) by treatment arm
  ○ Plot of Schoenfeld residuals for treatment
  ○ Evaluation of time-dependency of treatment effect by adding an interaction term of treatment and time in the Cox model. If the interaction term is significant (p< [0.10]), it is considered that the proportional hazards assumption is violated.
• If one or more of the procedures above demonstrates clear violation of the proportional hazards assumption in PFS, it is considered the proportional hazards assumption does not hold. Hazard ratio and corresponding 95% CI estimated from the Cox model will still be reported.
• More details for handling possible non-proportional hazards effect are provided in Section 4.2.4.1.

Model Results Presentation
• Kaplan-Meier estimates for the median PFS and the first and third quartiles will be presented, along with 95% CIs. • The p-value from the stratified log-rank test will be reported. • Hazard ratio and corresponding 95% confidence interval from the Cox model will be reported.

4.2.4. Sensitivity analyses

4.2.4.1. Non-Proportional Hazards Effect

If there is possible non-proportional hazards effect in PFS, the Restricted Mean Survival Time (RMST) method [Uno, 2015] may be conducted as appropriate, while the hazard ratio (HR) and its corresponding 95% CI based on Cox proportional hazard model will still be reported.

Restricted Mean Survival Time (RMST)

RMST method may be conducted to account for the possible non-proportional hazards effect. The RMST is the expected survival time restricted to a specific time horizon t^*. The cutoff t^* for determining the RMST will be the smallest value among the largest observed time across study interventions.

Statistical Methodology Specification

Endpoint / Variables
• PFS
Model Specification
• Additional analysis based on RMST will be conducted if the proportional hazard assumption does not hold. • RMST at *t** will be estimated from the Kaplan-Meier curve for each treatment arm: $$\mu_{t^*} = \int_0^{t^*} S(t)\,dt$$ • RMST difference at *t** ($\hat{\Delta}_{t^*}$) between treatment arms will be estimated as: $$\hat{\Delta}_{t^*} = \int_0^{t^*} \left[\hat{S}_T(t) - \hat{S}_C(t)\right] dt$$ • 95% CI for RMST difference and the p-value will be estimated using the following formula under normal approximation (Klein, 2005): $$Var\left(\hat{\Delta}_{t^*}\right) = V[\hat{\mu}_{t^*}(T)] + V[\hat{\mu}_{t^*}(C)]$$ $$V[\hat{\mu}_{t^*}] = \sum_{i=1}^{D} \left[\int_{t_i}^{t^*} \hat{S}(t)dt\right]^2 \frac{d_i}{Y_i(Y_i - d_i)}$$ where d_i is the number of events and Y_i is number of participants at risk at t_i.
SAS Procedure
• SAS/STAT 15.1 will be used for the statistical analysis. • Proc LIFETEST will be used with RMST option in order to obtain the RMST in both the treatment groups. • Proc RMSTREG will be used to obtain the RMST difference between the groups and corresponding 95% CI. The option link=linear will be specified. "Mean Plot" with "CLBAND" option will be used to generate the RMST plot with confidence bands.
Model Results Presentation
• If the proportional hazard assumption does not hold, the p-value based on the RMST test will also be reported in addition to the stratified log-rank test p-value. • RMST difference at *t**, and the corresponding 95% confidence interval will be presented. • A plot of RMST up to *t** and the corresponding 95% simultaneous confidence bands will be generated.

4.2.4.2. Sensitivity Analysis of Primary Estimand of PFS on Independent Review Committee (IRC) assessed response

If a review by IRC is conducted on all the disease assessments, a sensitivity analysis of the primary estimand of PFS (i.e. handling of intercurrent events based on primary event and censoring rules) will be conducted based on the IRC-assessed response. The sensitivity analyses will only use the Kaplan-Meier estimates, Cox proportional hazards model stratified by randomization factors described in Section 4.2.3. This analysis won't be available at the time of Study Analysis Complete (SAC) but will only be conducted

after the full IRC review is finished.If a review by IRC is conducted only on a randomly selected subjects, statistical analyses specified in the IRC audit plan on the IRC- assessed response will be conducted. This analysis won't be available at the time of SAC but will only be conducted after the IRC review on the randomly selected subjects is finished.

4.2.5. Supplementary Analyses

4.2.5.1. Analysis of Supplementary Estimand of PFS

Additional analyses of the supplementary estimands 1&2 (i.e. handling of intercurrent events based on alternative event and censoring rules 1&2) of PFS will be conducted based on investigator-assessed response.

These additional analyses will only use the Kaplan-Meier estimates and Cox proportional hazards model stratified by randomization factors described in Section 4.2.3.

4.2.5.2. Analysis of PFS based on mITT Analysis Set

The following sets of analyses will also be conducted based on the mITT analysis set using investigator-assessed response:

[1]. Primary analysis of primary estimand (investigator-assessed response + primary censoring rules);

[2]. Primary analysis of supplementary estimand 1 (investigator-assessed response + alternative censoring rules 1);

[3]. Primary analysis of supplementary estimand 2 (investigator-assessed response + alternative censoring rules 2);

The analytical approach for each analysis above is: Kaplan-Meier estimates and Cox proportional hazards model stratified by randomization factors as described in Section 4.2.3.

4.2.5.3. Analysis of PFS based on Stratification Data from the Clinical Database

If there is any mis-stratification for stratification factors entered for randomization, the following supplementary analysis will be performed using the stratification data based on the clinical database.

[1]. Primary analysis of primary estimand (investigator-assessed response + primary censoring rules);

The analytical approach is: Cox proportional hazards model stratified by randomization factors (based on data from the clinical database).

4.3. Secondary Endpoint(s) Analyses

4.3.1. Key secondary endpoint

Key secondary endpoints are those secondary endpoints for which a label claim is pursued as part of the confirmatory hypotheses for which the type 1 error is controlled (multiplicity adjustment). The analyses of OS will be based on the ITT Analysis Set, unless otherwise specified. In addition, pending on maturity of data, the survival probability at 6, 12 and 18 months with 95% CI will be estimated using Kaplan-Meier method.

4.3.1.1. Definition of endpoint

The key secondary endpoint in this study is Overall Survival (OS), defined as the interval of time from randomization to the date of death due to any cause. Participants who are alive will be censored at the last known date of last contact. Participants without documented death will be censored at last known alive date. The last date will be determined by the maximum collection/assessment date from among selected data domains within the clinical database; details will be provided in a separate Output and Programming Specification (OPS) document. When calculating overall survival, all deaths following subsequent anti-cancer therapy will be included. This is the primary estimand of OS, and there is no supplementary estimand of OS.

4.3.1.2. Main analytical approach

Refer to Section 4.2.3 (i.e. Kaplan-Meier estimates, stratified log-rank test, Cox proportional hazards model stratified by randomization factors, and examination of non-proportional hazards effect).

4.3.1.3. Sensitivity analyses

The analyses in this section may only be conducted at the 2nd interim and final analysis for OS when OS data is more mature.

If there is possible non-proportional hazards effect, refer to Section 4.2.4.1 for RMST method.

To obtain a less biased estimate of treatment effect, an IPCW method will be performed for OS adjusting for the post study therapy [Watkins, 2013]. To compensate the impact of switching to another therapy, patients with similar characteristics to those patients that switched treatment are obtained and assigned a higher weight to the patient that stayed on treatment and a lower weight to the patients receiving post study therapy. To calculate these weights, the likelihood of remaining uncensored will be estimated by logistic regression. Specifically, two logistic regression models, one using only baseline covariates and other using both baseline and time dependent covariates will be performed. The coefficient between these two estimated probabilities of switching is going to give the assigned weights. Subsequently, patients who switched will have a lower weight than patients that did not.

4.3.1.4. Supplementary Analyses

Analysis of OS based on mITT Analysis Set

Analyses of OS will also be conducted based on the mITT analysis set. The analytical approach for each analysis above is Kaplan-Meier estimates and Cox proportional hazards model stratified by randomization factors as described in Section 4.2.3.

Analysis of OS based on Stratification Data from the Clinical Database

If there is any mis-stratification for stratification factors entered for randomization, the following supplementary analysis will be performed using the stratification data based on the clinical database. The analytical approach is Cox proportional hazards model stratified by randomization factors (based on data from the clinical database).

4.3.2. Supportive secondary endpoint(s)

4.3.2.1. Supportive Secondary Efficacy Endpoint(s)

Primary analysis of supportive secondary efficacy endpoints will be based on investigator-assessed response. Analyses of supportive secondary efficacy endpoints will be based on the ITT Analysis Set, unless otherwise specified.

Definition of Endpoint(s)

At final PFS/OS analysis, the analyses for the following supportive secondary efficacy endpoints will be conducted:

- **Overall response rate (ORR)**, is defined as the percentage of participants with a confirmed PR or better (i.e., PR, VGPR, CR, and sCR).
- **Clinical benefit rate (CBR)**, is defined as the percentage of participants with a confirmed minimal response (MR) or better.
  ORR and CBR will be analyzed based on the confirmed responses, which will be derived based on the algorithm specified in Table 3. Only the assessments from the start of treatment up to the earlier of confirmed disease progression or the start of new anti-cancer therapy will be considered. Only new 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). Participants 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.
- **Duration of response (DoR)** is defined as the time from first documented evidence of PR or better until the earliest date of disease progression (PD), or death due to any cause among participants who achieve a response (i.e., confirmed PR or better). Responders without disease progression will be censored at the censoring time point for PFS.
- **Time to response (TTR)** is defined as the time between the date of randomization and the first documented evidence of response (PR or better), among participants who achieve a response (i.e., confirmed PR or better).

- **Time to progression (TTP)** is defined as the time from randomization until the earliest date of PD, or death due to PD. Determination of dates of TTP event and dates for censoring are described in the Table 4.
- **Minimal Residual Disease (MRD) Negativity Rate**, is defined as: the percentage of participants who are MRD negative by Next Generation Sequencing (NGS).

Table 3 Response confirmation algorithm based on visit-level response

#	Response at the First Time Point	Response at Subsequent Disease Assessment[1]	Confirmed Response at the First Time Point
1	sCR	sCR	sCR
2	sCR	CR	CR
3	CR	sCR/CR	
4	sCR/CR	VGPR	VGPR
5	VGPR	sCR/CR/VGPR	
6	sCR/CR/VGPR	PR	PR
7	PR	sCR/CR/VGPR/PR	
8	sCR/CR/VGPR/PR	MR	MR
9	MR	sCR/CR/VGPR/PR/MR	
10	sCR/CR/VGPR/PR/MR	SD	SD
11	sCR/CR/VGPR/PR/MR	PD (any reason) OR No subsequent disease assessment: **subject died or discontinued study or started new anti-cancer therapy** before further adequate disease assessment	SD
12	PD (due to reasons other than imaging, i.e., plasmacytoma or bone lesion)	PD (any reason) including PD after initiation of new anti-cancer therapy OR No subsequent disease assessment: subject **died due to PD** before further adequate disease assessment and within 49 days of PD at First Time Point (including death due to PD after initiation of new anti-cancer therapy)	PD
13	PD (due to reasons other than imaging, i.e., plasmacytoma or bone lesion)	sCR/CR/VGPR/PR/MR/SD OR No subsequent disease assessment: subject died **due to reasons other than**	NE

#	Response at the First Time Point	Response at Subsequent Disease Assessment[1]	Confirmed Response at the First Time Point
		PD before further adequate disease assessment OR No subsequent disease assessment: subject **discontinued study** before further adequate disease assessment	
14	sCR/CR/VGPR/PR/MR/PD (due to reasons other than imaging, i.e., plasmacytoma or bone lesion)	No subsequent disease assessment: subject has not died, discontinued from study or (except for PD) started new anti-cancer therapy; but as yet has no further adequate disease assessments	Unconfirmed sCR/CR/VGPR/PR/MR/PD. Will be categorized as NE for final ORR analysis.
15	SD	Any OR No subsequent disease assessment	SD
16	PD due to imaging (plasmacytoma or bone lesion)	Any OR No subsequent disease assessment	PD
17	NE or missing	Any OR No subsequent disease assessment	NE

1. Subsequent disease assessment is defined as the next adequate (not missing or NE) disease assessment following the first timepoint before (or on the same date of) start of new anti-cancer therapy except for confirmation of PD, for which PD or death due to PD after new anti-cancer therapy are considered for confirmation of PD. No minimal time interval is required for the subsequent disease assessment, but a different sample is required for confirmation.
2. SD does not need to be confirmed.
3. PD due to imaging (i.e., plasmacytoma or bone lesion) does not need to be confirmed.
4. Where criteria are not mutually exclusive, take the first that applies.

Table 4 Assignments for Progression and Censoring Dates for TTP Analysis

Situation	Date of Event (Progression/Death) or Censoring	Outcome Event (progression/Death) Or Censored
No *(or inadequate)* baseline tumor assessments [1] and the subject has not died (if the subject has died follow the rules for death indicated at the bottom of the table)	Randomization	Censored
No post-baseline assessments and the subject has not died (if the subject has died follow the rules for death indicated at the bottom of the table)	Randomization	Censored
Progression documented at or between scheduled visits, without extended loss-to-follow-up time [4]	Date of assessment of progression	Event
No progression *(or death)*	Date of last 'adequate' assessment of response [2]	Censored
New anticancer treatment started *(prior to documented disease progression or death)* [3].	Date of last 'adequate' assessment of response [2] *(on or prior to starting anti-cancer therapy)*	Censored
Death due to progression without extended loss-to-follow-up time [4]	Date of death	Event
Death from causes other than progression without extended loss-to-follow-up time [4]	Date of death	Censored
Death or progression after an extended loss-to-follow-up time [4]	Date of last 'adequate' assessment of response [2] prior to PD/death *(prior to missed assessments)*	Censored
Death or progression after an extended loss-to-follow-up time [4] from randomization	Date of randomization	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 (≥5 g/L) or b. Urine M-protein ≥200 mg/24h or c. Serum FLC assay: Involved FLC level ≥10 mg/dL (≥100 mg/L) and an abnormal serum free light chain ratio (<0.26 or >1.65)

Situation	Date of Event (Progression/Death) or Censoring	Outcome Event (progression/Death) Or Censored

[2]. An adequate assessment is defined as an assessment where the response is sCR, CR, VGPR, PR, MR, or SD.

[3]. If PD or death and New anti-cancer therapy occur on the same day assume the progression or death was documented first (e.g., outcome is progression or death and the date is the date of the assessment of progression or death). If anti-cancer therapy is started prior to any adequate assessments, censoring date should be the date of randomization.

[4]. Extended loss-to-follow-up time = 6 weeks + 7 day window = 49 day window; Without extended loss-to-follow-up time is defined as: <= 49 days; after an extended loss-to-follow-up time is defined as: >49 days. More details about extended loss-to-follow-up time are provided in Section 6.2.9.

Main Analytical Approach

- **ORR**: The number and percentage of participants with the best confirmed response in the following response categories at will be summarized by treatment arm: sCR, CR, VGPR, PR, overall response (sCR+CR+VGPR+PR), minimal response (MR), stable disease (SD), progressive disease (PD), and not evaluable (NE). The corresponding exact 95% CI for ORR will also be provided. Participants with unknown or missing responses will be treated as non-responders, i.e., these participants will be included in the denominator when calculating percentages of response. The exact 95% CI for the difference will be calculated.
- **CBR**: summaries of CBR by treatment arms will be provided in the same way as ORR.
- **DoR**: Distribution of DoR will be summarized using the Kaplan-Meier method by treatment arm. The median, 25th and 75th percentiles of DoR will be estimated and corresponding 95% confidence intervals will be estimated using the Brookmeyer-Crowley method [Brookmeyer, 1982].
- **TTR**: TTR will be summarized descriptively by treatment arm using medians and quartiles in the subset of participants with a confirmed response of PR or better as the Best Overall Response (BOR).
- **TTP:** Distribution of TTP will be summarized using the Kaplan-Meier method by treatment arm. The median, 25th and 75th percentiles of TTP will be estimated and corresponding 95% confidence intervals will be estimated using the Brookmeyer-Crowley method [Brookmeyer, 1982]. TTP analysis will also be conducted using Cox proportional hazards model stratified by randomization factors. Refer to Section 4.2.3 for details of the analytical approaches.
- **MRD Negativity Rate**: For MRD negativity rate based on bone marrow testing using Next Generation Sequencing (molecular negativity), the number and percentage of participants who have achieved MRD negativity will be summarized by treatment arm. For analysis purposes, participants in the ITT population without MRD assessment will be considered as having positive MRD. The corresponding exact 95%

CI for MRD negativity rate will also be provided. Number and percentage of participants who have sustained MRD negativity (≥6 months, ≥ 12months) will be provided. Information of MRD will be included in the listing of response. If data are available, imaging-based assessment of MRD (i.e. PET-CT) will also be included in the listing and related to NGS testing.

4.3.2.2. Supplementary analyses

Additional analyses of DoR will be conducted for the supplementary estimand of DoR.

4.3.3. Pharmacokinetic Analyses

Plasma concentrations of belantamab mafodotin, total mAb, and cys-mcMMAF will be summarized using descriptive statistics, graphically presented (where appropriate) and listed based on the Pharmacokinetic (PK) Analysis Set. Details of the planned displays are provided in the Output and Programming Specification (OPS) document.

4.3.4. Immunogenicity (Anti-Drug Antibody) Analyses

For each subject, the anti-belantamab mafodotin antibody results, titers, and neutralizing antibody assay results, and total antibody concentration will be listed for each assessment time point. The frequency and percentage of participants with positive and negative anti-drug antibody and neutralizing antibody assay results will be summarized for each assessment time and overall for each subject by dose cohort. The conclusive results will be based on the total antibody concentration. The immunogenicity analyses will be based on the Safety Analysis Set.

4.3.5. Secondary Patient Reported Outcome Analyses

The EORTC QLQ-C30 (version 3.0), EORTC QLQ-IL52 (disease symptom domain of EORTC QLQ-MY20), EORTC QLQ-MY20 and the PRO-CTCAE are three oncology-specific Health-Related Quality-of-Life (HRQoL) assessments that will be analysed in this study as supportive secondary endpoints. EORTC QLQ-IL52 will be included in the EORTC QLQ-MY20 analyses.

In addition, the impact of potential corneal event on function and health-related quality-of-life will be assessed with the Ocular Surface Disease Index (OSDI) visual function questionnaire as a supportive secondary endpoint.

The analysis of EORTC QLQ-C30 and EORTC QLQ-MY20 (including EORTC QLQ-IL52) will be based on the ITT Analysis Set; while the analysis of PRO-CTCAE, and OSDI will be based on the Safety Analysis Set.

Patient Reported Outcome Version of the Common Term Criteria for Adverse Events (PRO-CTCAE)

The Patient-Reported Outcomes Version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE) is a patient-reported outcome measure developed to

evaluate symptomatic toxicity in patients on cancer clinical trials [Basch, 2014]. The PRO-CTCAE was designed to be used as a companion to the Common Terminology Criteria for Adverse Events (CTCAE), the standard lexicon for adverse event reporting in cancer trials. The PRO-CTCAE includes an item library of 124 items representing 78 symptomatic toxicities drawn from the CTCAE. PRO-CTCAE provides a systematic yet flexible tool for descriptive reporting of symptomatic treatment side effects in cancer clinical trials. In the present study, a subset of items selected from the PRO-CTCAE Version 1.0 Item library will be administered.

The levels and related code values for PRO-CTCAE are shown below:

	Levels and related code values				
Response scale	0	1	2	3	4
Frequency	Never	Rarely	Occasionally	Frequently	Almost Constantly
Severity	None	Mild	Moderate	Severe	Very severe
Interference	Not at all	A little bit	Somewhat	Quite a bit	Very much
Present/Absence	No	Yes			

For each selected item from the library: proportion of PRO-CTCAE scores for attributes (frequency, severity and/or interference) will be presented-with stacked bar charts by visit. Maximum PRO-CTCAE score at post-baseline for each item attribute will be summarized by counts and proportions. Proportion of patients with a maximum score of 3 or 4 for each item attribute (severe or very severe, frequently or almost constantly, quite a bit or very much) will also be reported. Proportions will be based on the number of patients with available data and subject with missing response will be excluded from analysis. A listing of the PRO-CTCAE score will be provided for each attribute (frequency, severity, interference, presence).

Ocular Surface Disease Index

The Ocular Surface Disease Index (OSDI) is a 12-item questionnaire designed to assesses both the frequency of dry eye symptoms and their impact on vision-related functioning [Schiffman, 2000; Dougherty, 2011]. The OSDI has demonstrated good reliability, validity, sensitivity, and specificity, and can be used as a complement to other clinical and subjective measures of dry eye disease by providing a quantifiable assessment of dry eye symptom frequency and the impact of these symptoms on vision-related functioning.

For the OSDI, the total score will be calculated as well as scores for the three subscales (Ocular Symptoms: item 1-3; Vision-related Function: item 4-9; and Environmental Triggers: item 10-12).

The total OSDI score = ([sum of scores for all questions answered × 100]/[total number of questions answered ×4]). Subscale scores are computed similarly with only the questions from each subscale used to generate its own score. A score of 100 corresponds to complete disability (a response of “all of the time” to all questions answered), while a score of 0 corresponds to no disability (a response of “none of the time” to all questions answered). Therefore, decrease in score from baseline means improvement.

For total score and each of the three sub-scales, the descriptive summary of the actual value and change from baseline at selected time points will be provided. Plots of mean change from baseline (including baseline) and 95% confidence interval over time by visit, and at end of study treatment, last follow-up, and worst case post-baseline for individual domains will also be provided.

European Organization for Research and Treatment of Cancer Quality of Life Questionnaire 30-item Core Module (EORTC QLQ-C30)

The EORTC QLQ-C30 is a 30-item questionnaire containing both single- and multi-item measures [Aaronson, 1993]. These include five functional scales (Physical, Role, Cognitive, Emotional, and Social Functioning), three symptom scales (Fatigue, Pain, and Nausea/Vomiting), a Global Health Status/QoL scale, and six single items (Constipation, Diarrhea, Insomnia, Dyspnea, Appetite Loss, and Financial Difficulties). Scores for each scale and single-item measure are averaged and transformed linearly to a score ranging from 0–100. Details of deriving domain scores (9 scales and 6 single items) and summary score can be found in Section 6.2.8.1 and more details will be provided in OPS document.

- A high score for functional scales and for Global Health Status/QoL and summary score represent better functioning ability or Health-Related Qualify of Life (HRQoL), (higher score indicates improvement)
- whereas a high score for symptom scales and single items represents significant symptomatology [Proskorovsky, 2014]. (lower score indicates improvement)

Descriptive summaries (mean, SD, median, min and max) of the actual value and change from baseline at selected time points will be provided for EORTC QLQ-C30 summary score and each domain score. The number and percentage of participants with post-baseline score improved by ≥ 10, and ≥ 5 points, respectively from baseline score will be summarized at selected time points. The number and percentage will be provided for summary score and each domain score. Should new thresholds be available at the time of the analysis (i.e. from ongoing EORTC group work) these modified thresholds will be used and specified in OPS.

Plots of mean change from baseline (including baseline) and 95% confidence interval over time by visit, and at end of study treatment, last follow-up, and worst case post-baseline for individual domain score and summary scores will also be provided.

Longitudinal changes from baseline by treatment group for EORTC QLQ-C30 domain Fatigue will be explored using a restricted maximum likelihood-based mixed model for repeated measures (MMRM) to compare between-treatment difference adjusting for correlations across multiple time points within a patient and controlling for the baseline value. Adjusted mean difference and 95% CIs will be presented to illustrate the effect of treatment.

The MMRM model will include patient, treatment, analysis visit, and treatment-by-visit interaction as explanatory variables, the baseline value as a covariate along with the baseline-by-visit interaction. Treatment, visit, and treatment-by-visit interactions will be fixed effects in the model; patient will be treated as a random effect. An unstructured

covariance matrix will be used to model the within subject variance and the Kenward-Roger approximation will be used to estimate the degrees of freedom. Restricted maximum likelihood (REML) estimation will be used. If the fit of the unstructured covariance structure fails to converge, the following covariance structures will be used in order until convergence is reached: toeplitz with heterogeneity (TOEPH), autoregressive with heterogeneity (ARH(1)), Toeplitz (TOEP), and autoregressive (AR(1)). If there are still issues with the fit of the model or estimation of the treatment effects, SUBJECT will be treated as a fixed effect.

EORTC QLQ-MY20 and EORTC IL52

The EORTC Quality of Life Questionnaire 20-item Multiple Myeloma module (QLQ-MY20) is a supplement to the QLQ-C30 instrument used in patients with multiple myeloma [Aaronson, 1993; Cocks, 2007]. The module comprises 20 questions that address four myeloma-specific HRQoL domains: Disease Symptoms (EORTC IL52), Side Effects of Treatment, Future Perspective, and Body Image. Three of the four QLQ-MY20 domains are multi-item scales: Disease Symptoms (includes bone aches or pain, back pain, hip pain, arm or shoulder pain, chest pain, and pain increasing with activity); Side Effects of Treatment (includes drowsiness, thirst, feeling ill, dry mouth, hair loss, upset by hair loss, tingling hands or feet, restlessness/agitation, acid indigestion/heartburn, and burning or sore eyes); and Future Perspective (includes worry about death and health in the future, and thinking about illness). The Body Image scale is a single-item scale that addresses physical attractiveness. As with the QLQ-C30, QLQ-MY20 domain scores are averaged and transformed linearly to a score ranging from 0–100. A high score for Disease Symptoms and Side Effects of Treatment represents a high level of symptomatology or problems [Proskorovsky, 2014] (lower score indicates improvement), whereas a high score for Future Perspective and Body Image represents better outcomes (higher score indicates improvement). Details of deriving domain score can be found in Section 6.2.8.2. and more details will be provided in OPS document.

Descriptive summaries (mean, SD, median, min and max) of the actual value and change from baseline at selected time points will be provided for each domain score. The number and percentage of participants with post-baseline score improved by ≥ 10, and ≥ 5 points, respectively from baseline score will be summarized at selected time points. Should new thresholds be available at the time of the analysis (ie. from ongoing EORTC group work) these modified thresholds will be used and specified in OPS.

Plots of mean change from baseline (including baseline) and 95% confidence interval over time by visit, and at end of study treatment, last follow-up, and worst case post-baseline for each domain will also be provided. In addition, the plot will also be provided in the subgroup for participants achieving a confirmed partial response (PR) or better based on the investigator-assessed response.

Longitudinal changes from baseline by treatment group for EORTC QLQ-IL52 domain score will be explored using a restricted maximum likelihood-based mixed model for repeated measures (MMRM), using the same approach described in Section for EORTC QLQ-C30 analysis.

Compliance of PRO-CTCAE, OSDI, EORTC QLQ-C30, EORTC QLQ-MY20 and EORTC QLQ-IL52

For each of the PROs PRO-CTCAE, OSDI, EORTC QLQ-C30, EORTC QLQ-MY20 and EORTC IL52, overall compliance and compliance by visit will be summarized, based on the following definitions.

- Number of patients expected to complete PRO form: Date of study discontinuation and/or date of death will be used to determine the last visit at which a patient is still expected under PRO follow-up.
- Evaluable forms:
    - PRO-CTCAE: with at least one non-missing item score
    - OSDI: with at least one non-missing total score or subscale score
    - EORTC QLQ-C30, EORTC QLQ-MY20: with at least one non-missing scale/domain score
    - EORTC IL52: with non-missing EORTC IL52 scale/domain score

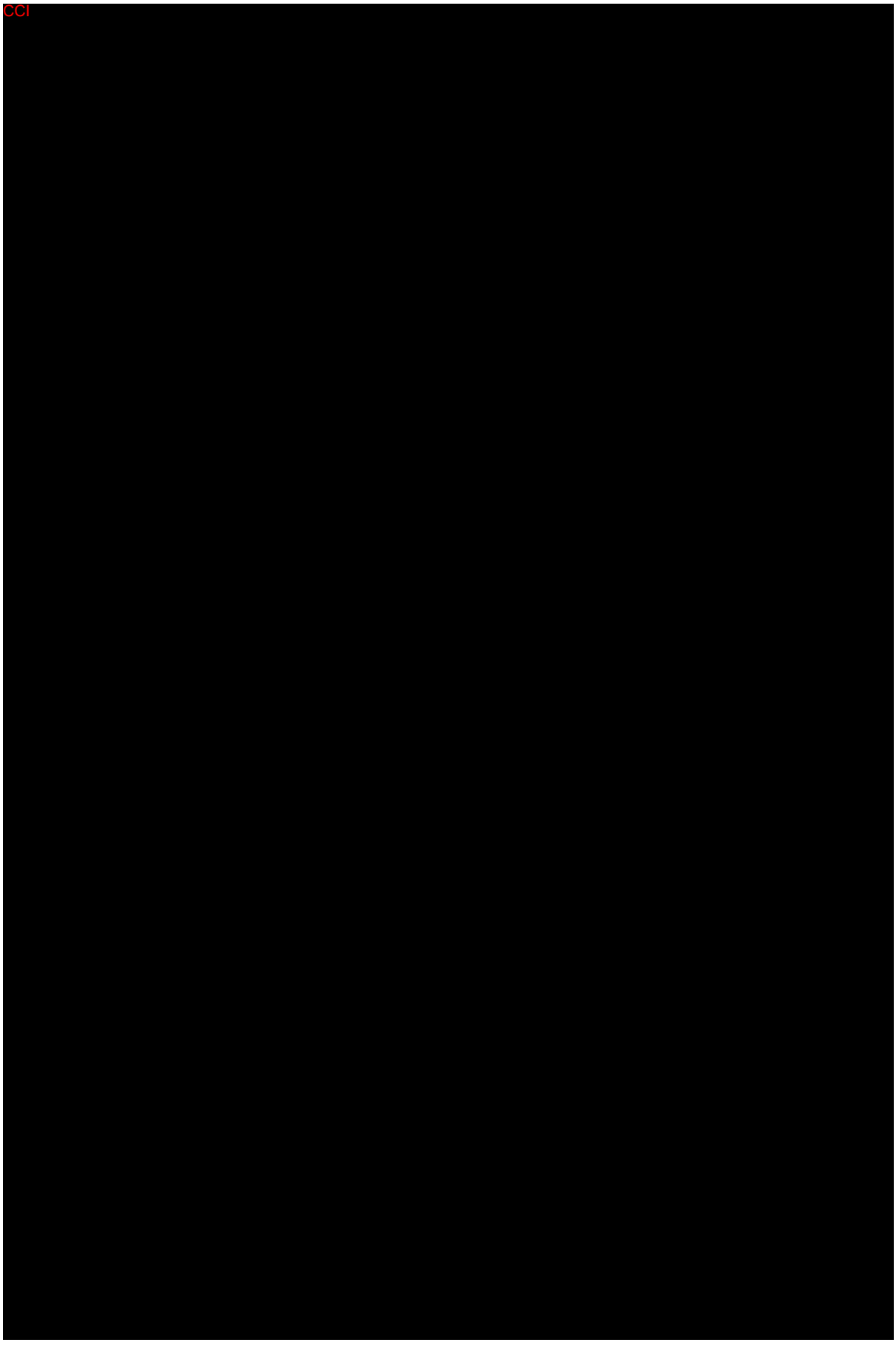
The overall compliance rate is defined as the number of patients with an evaluable baseline form and at least one evaluable post-baseline form, divided by the number of patients expected to complete the baseline form.

Compliance by visit will be calculated as the number of patients with an evaluable form at that visit, divided by the number of patients expected to complete the form at that visit.

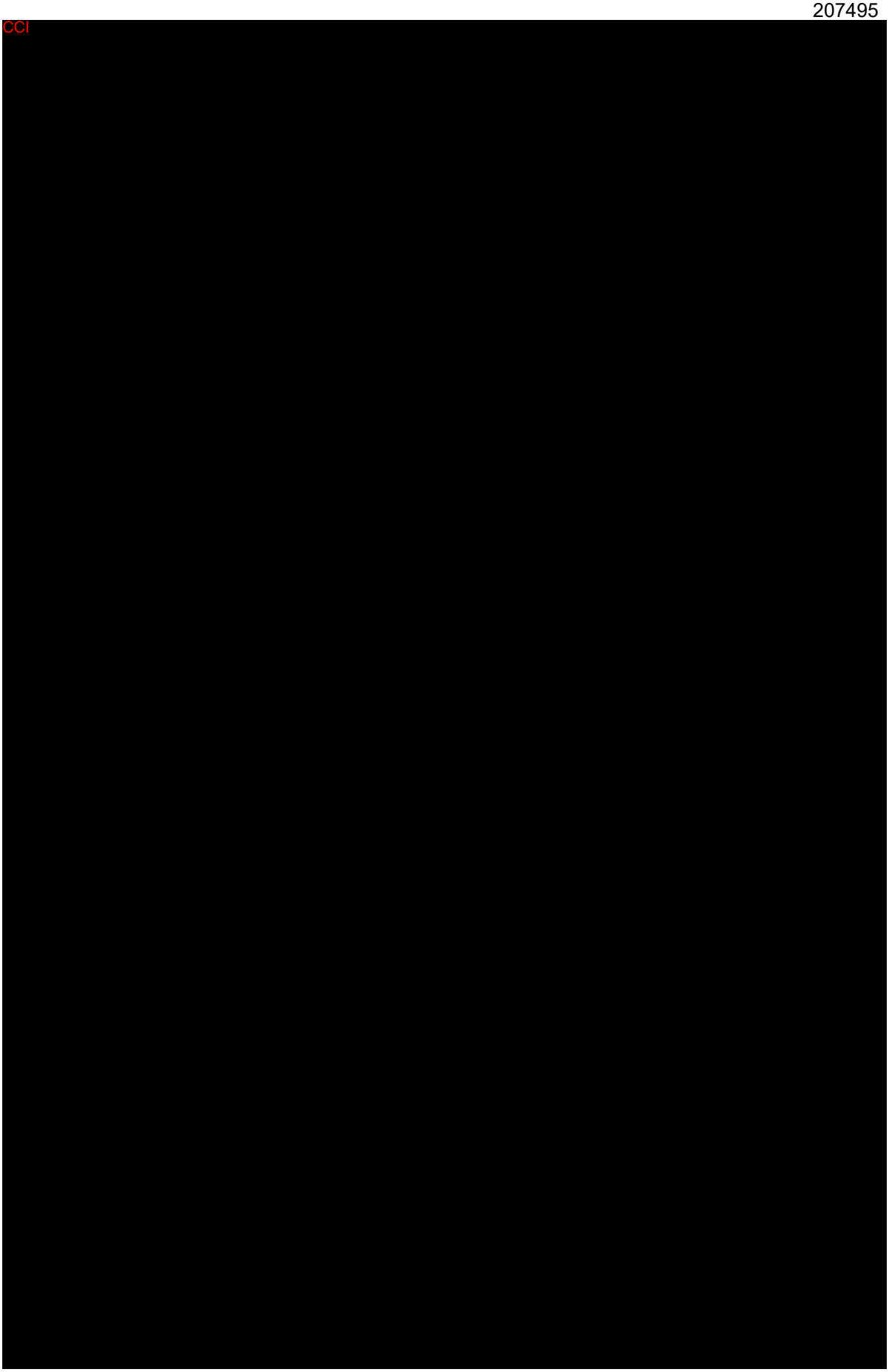
4.4. Exploratory Endpoint(s) Analyses

CCI

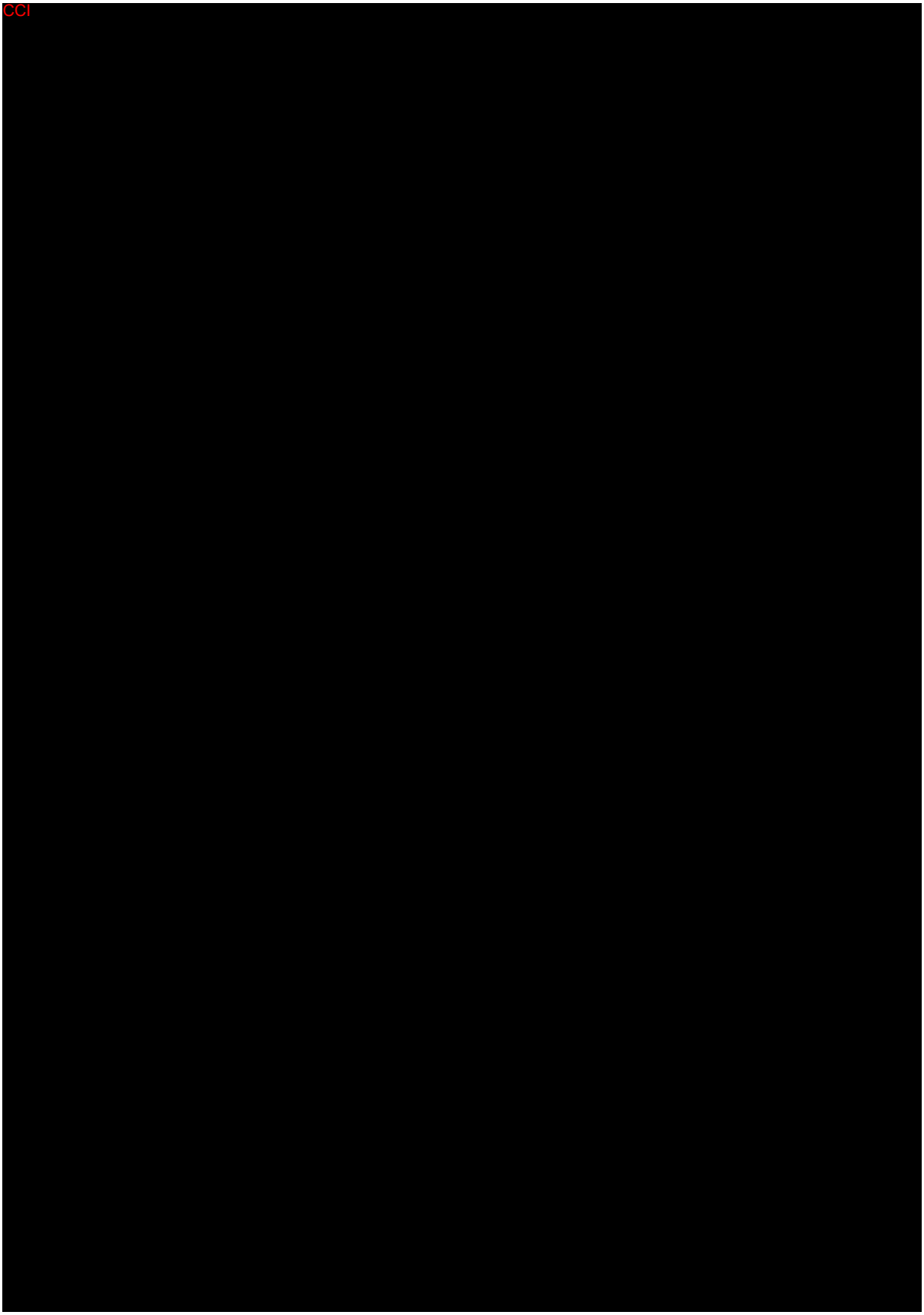
CCI



CCI



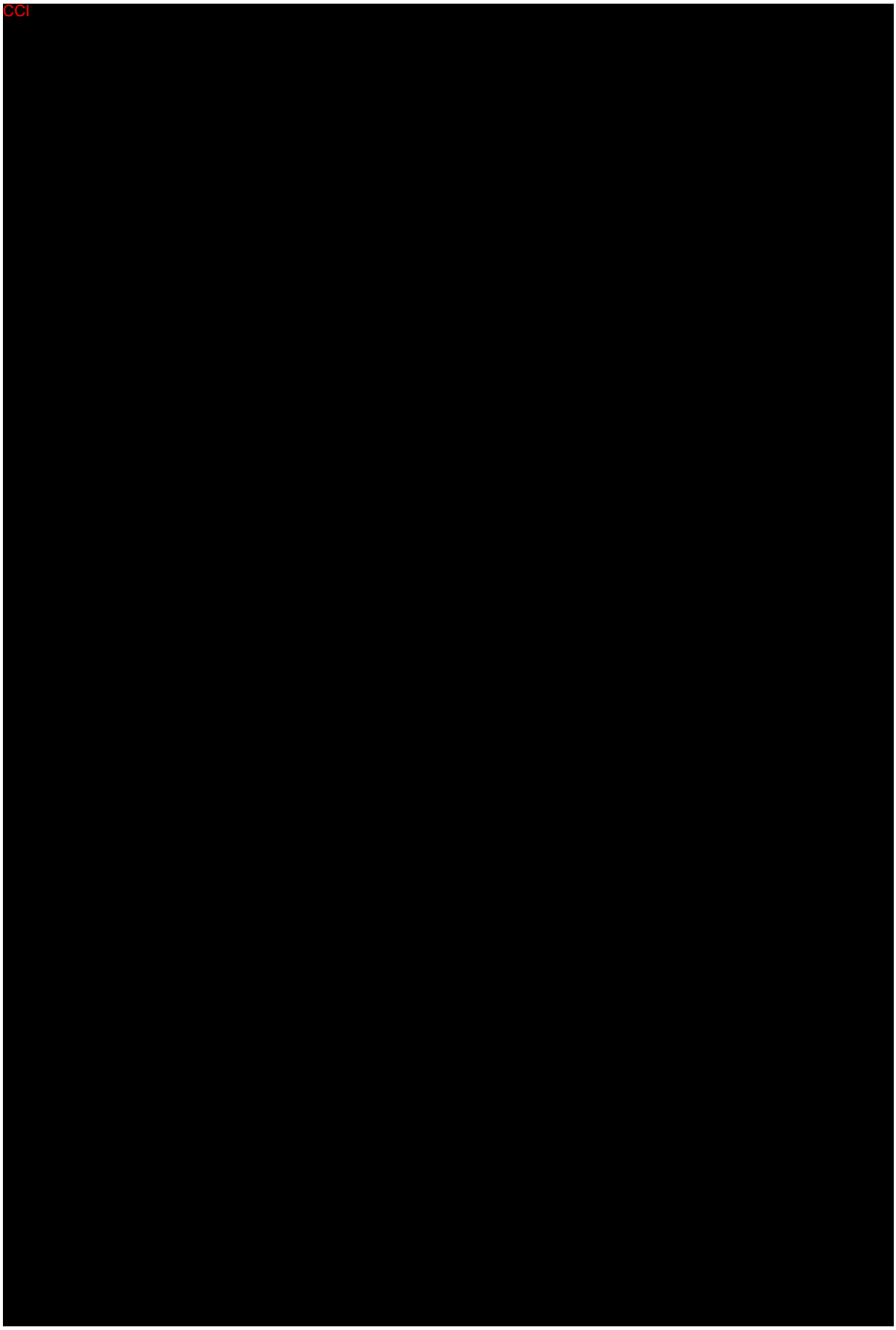
CCI



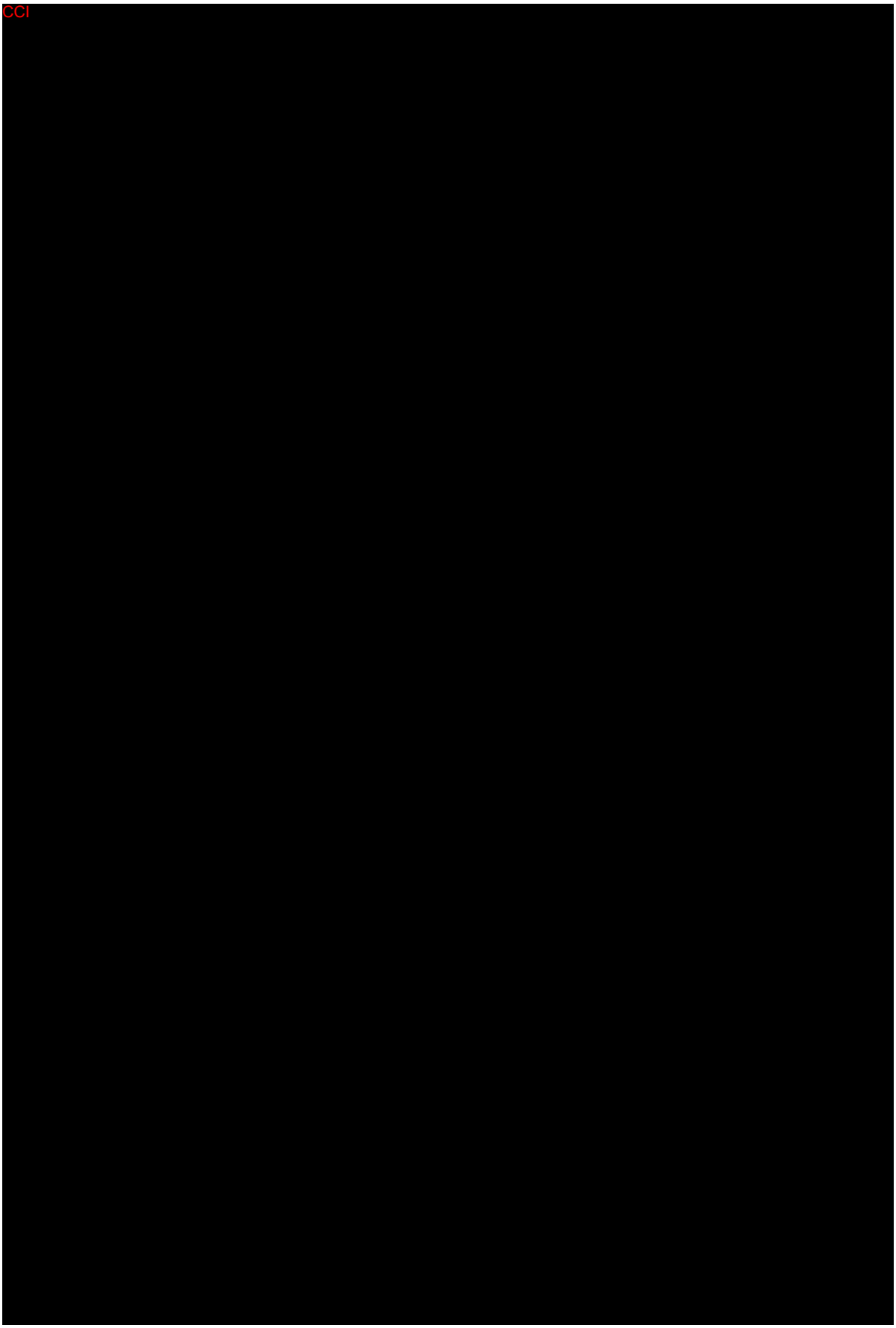
CCI

[redacted]

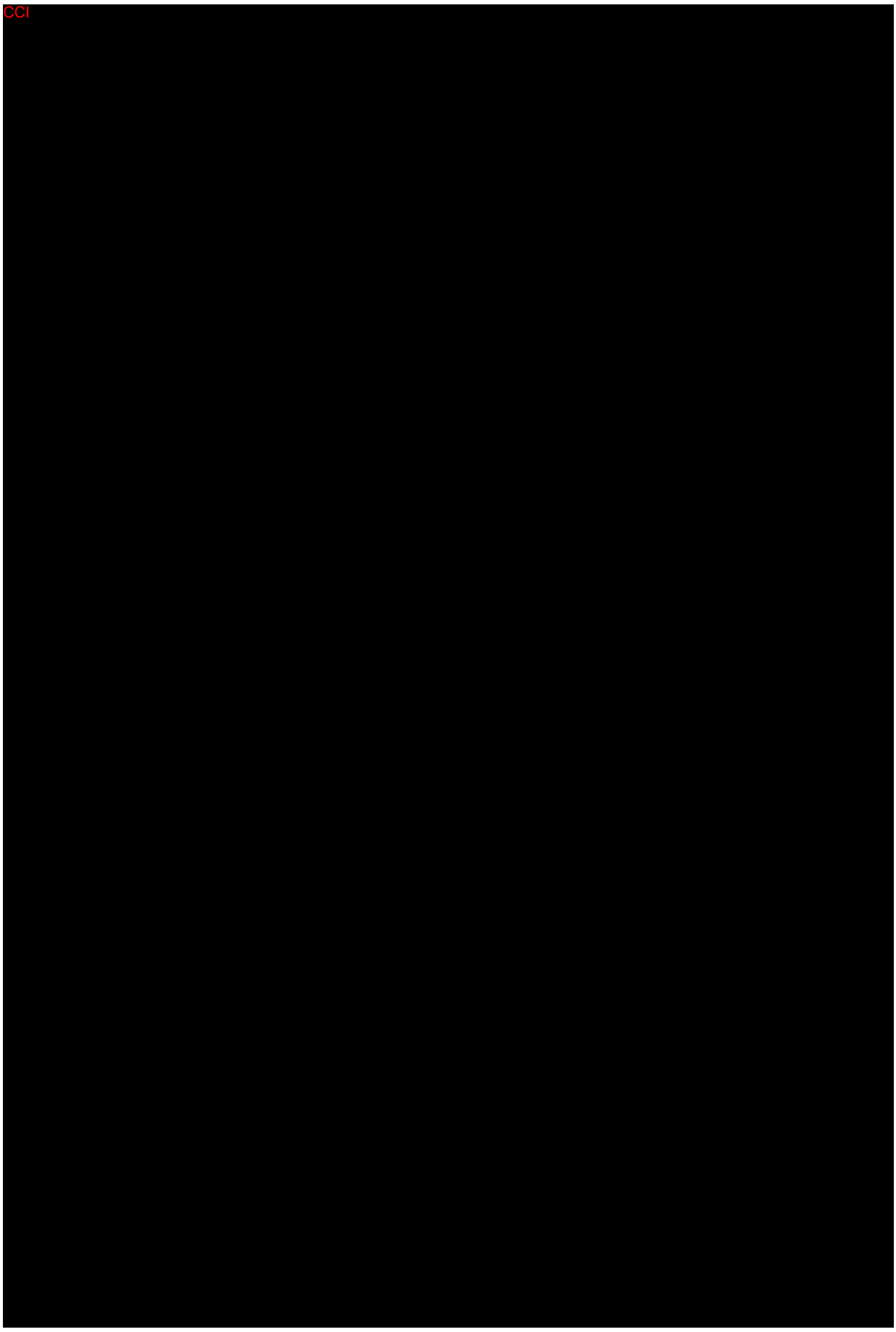
CCI



CCI



CCI



4.5. Safety Analyses

The safety analyses will be based on the Safety Analysis Set, unless otherwise specified.

4.5.1. Extent of Exposure

Belantamab mafodotin Arm:

Extent of exposure to belantamab mafodotin will be summarized.

The number of cycles administered by study treatment will be summarized with mean, median, standard deviation, minimum, and maximum.

The dose intensity (mg/kg/3 weeks), which is calculated as the cumulative actual dose (mg/kg) divided by expected duration of exposure in 3 weeks [(last infusion date – first infusion date+21)/21], will also be summarized. A by subject summary listing of data on exposure to all study treatments will be produced.

The dose intensity (mg/kg/3 weeks) up to a cycle or infusion except the last (cycle or infusion), which is calculated as cumulative actual dose (mg/kg) up to the current cycle, divided by expected duration of exposure up to the current cycle in 3 weeks [(next

infusion date-1 – first infusion date+1)/21], will also be summarized. The dose intensity up to the last cycle or infusion will be calculated in the same way as the dose intensity earlier.

The duration of exposure to study treatment (from first day to last day of treatment) will be calculated and summarized using mean, median, standard deviation, minimum, and maximum. A horizontal bar graph of duration of treatment will be produced that displays duration of treatment in months for each participant.

Dose reductions will be summarized by number of reductions and reasons for reductions. Dose delays will be summarized by number of delays, reasons for the delays, and delay duration (days). The number and percentage of the delays for intervals of 1-21, 22-42 and >42 days will be computed. Primary reasons for dose reductions and dose delays will also be summarized by cycle.

Duration of delays is defined as period from the expected start date of dose to actual start date of current dose. Calculation: (actual start date of current dose - expected start date of dose). Expected start date of dose = actual start date of previous dose + 21.

The summaries of dose modifications will be provided. All the dose reductions, dose escalations, infusion interruptions, incomplete infusions and dose delays will be listed. A plot showing the number and percentage of participants treated at different dose levels over time will be provided.

A patient profile plot will also be produced of all responders (PR or better based on investigator-assessed response) with any dose delays >63 days.

Pomalidomide / Dexamethasone Arm:

The start date of the overall study treatment is defined as the first dose date of pomalidomide or dexamethasone, whichever is earlier (i.e. the first study drug start date). For pom/dex arm, for each cycle, the cycle start date is defined as the pomalidomide start date; the day before the start date of the current cycle (after the 1st cycle) is defined as the end day of the previous cycle. The end date of the last cycle will be calculated as the earliest of: start date of last treatment cycle + 27 days, treatment discontinuation date , which is the date of the last dose, or the death date if the subject discontinues study or dies before the expected end of the last cycle (start date of last treatment cycle + 27 days).

The overall study treatment end date is defined as the end date of the last cycle. The overall treatment duration (days) is defined as:

Overall treatment duration (days)=the overall study treatment end date – the first study drug start date + 1

Separately for pomalidomide and dexamethasone:

- Duration of treatment = the end date of the last cycle of the study drug – the first dose date of the study drug+1
- Cumulative dose = sum of all actual doses taken across the treatment.

- Dose exposure = total number of days on the study drug during the treatment phase, periods of dose break per protocol or dose interruptions will be excluded
- Average daily dose = cumulative dose / dose exposure (mg/day).
- Dose intensity = cumulative dose / duration of treatment (mg/day).
- Relative dose intensity = dose intensity / plan dose intensity:
    - planned dose intensity for pomalidomide = 4mg*21/28 days;
    - planned dose intensity for dexamethasone (for participants <=75 years old) = 40mg*4/28days;
    - planned dose intensity for dexamethasone (for participants with >75 years old) = 20mg*4/28days;

Descriptive statistics of cumulative dose, dose exposure, average daily doses, dose intensity and relative dose intensity will be summarized by cycle and by pomalidomide and dexamethasone, separately. The overall summary across cycles will also be provided. Duration of treatment by drug and overall treatment duration will be summarized by descriptive statistics.

The dose modifications (dose reductions, dose interruptions) will be summarized by study drug and listed.

4.5.2. Adverse Events

Adverse events analyses including the analysis of adverse events (AEs), Serious AEs (SAEs) and other significant AEs will be based on GSK Core Data Standards.

An overview summary of AEs, including counts and percentages of participants with any AE, AEs related to study intervention, Grade 3 and 4 AEs, Grade 3 and 4 AEs related to study intervention, AEs leading to permanent discontinuation of study intervention, study intervention related AEs leading to permanent discontinuation of study intervention, AE leading to dose reductions, AEs leading to dose delays, SAEs, SAEs related to study intervention, fatal SAEs, and fatal SAEs related to study intervention will be produced. Unless otherwise specified, AEs will be summarized by treatment arms. AEs related to pomalidomide and dexamethasone separately, OR pomalidomide and dexamethasone in combo will all be considered as AEs related to the pom/dex regimen and summarized under the pom/dex treatment arm. Adverse events will be coded using the standard Medical Dictionary for Regulatory Affairs (MedDRA dictionary) and graded by the investigator according to the NCI-CTCAE, (version 5.0).

A summary of number and percentage of participants with any adverse events by maximum grade will be produced. AEs will be sorted by Preferred term (PT) in descending order. The summary will use the following algorithms for counting the participant:

- **Preferred term row**: Participants experiencing the same AE preferred term several times with different grades will only be counted once with the maximum grade.

- **Any event row**: Each participant with at least one adverse event will be counted only once at the maximum grade no matter how many events they have.

The frequency and percentage of AEs (all grades) will be summarized and displayed in two ways: 1) in descending order by PT only and 2) in descending order by SOC and PT. In the SOC row, the number of participants with multiple events under the same SOC will be counted once.

A summary of number and percentage of participants with treatment-emergent AEs by maximum grade will be also produced.

A separate summary will be provided for study intervention-related AEs. A study intervention-related AE is defined as an AE for which the investigator classifies the possible relationship to study intervention as “Yes”. A worst-case scenario approach will be taken to handle missing relatedness data, i.e. the summary table will include events with the relationship to study intervention as ‘Yes’ or missing. The summary table will be displayed in in two ways: 1) by maximum grade sorted by PT in descending order and 2) in descending order by SOC and PT.

A summary of cumulative incidence of AE by number of cycles (<=1, <=2, <=4, <=6, <=8, <=10, Any) received at first occurrence will be provided.

In addition, AEs of maximum grade of 3 or higher will be summarized separately by PT.

All SAEs will be tabulated based on the number and percentage of participants who experienced the event. Separate summaries will also be provided for study intervention-related SAEs. The summary tables will be displayed by PT. The summary of all SAEs will also be created by SOC and PT. In addition, a summary of cumulative incidence of SAE by number of cycles (<=1, <=2, <=4, <=6, <=8, <=10, Any) received at first occurrence will be provided.

A study intervention-related SAE is defined as an SAE for which the investigator classifies the relationship to study intervention as “Yes”. A worst-case scenario approach will be taken to handle missing data, i.e. the summary table will include events with the relationship to study intervention as ‘Yes’ or missing.

A summary of non-serious AEs that occurred in 5% of the participants or above will be provided (no rounding for the percentage will be used in terms of 5% threshold, e.g., event with 4.9% incidence rate should not be included in this table). The summary will be displayed by SOC and PT.

All AEs will be listed. Additionally, a listing of subject IDs for each individual AE will be produced. Separate supportive listings with participant-level details will be generated for fatal and non-fatal SAEs, respectively. The relationship between MedDRA SOC, PT, and Verbatim Text will be listed.

4.5.2.1. Adverse Events of Special Interest

The following will be considered adverse events of special interest (AESI) for the purpose of analyses: corneal events (CTCAE), thrombocytopenia and infusion-related reactions.

For corneal events (CTCAE), thrombocytopenia and infusion-related reactions (IRR), in addition to events identified and collected in eCRF, a comprehensive list of MedDRA terms based on clinical review will be used to identify each type of event. Specifically for IRR, events identified by the comprehensive list of MedDRA terms based on clinical review would only be considered IRR if the event was reported on an infusion day after the start of infusion or within 24 hours following end of infusion, AND led to a temporary interruption or prolongation of infusion time or treatment withdrawal. Changes to the MedDRA dictionary could occur between the start of the study and the time of reporting and/or emerging data from on-going studies may highlight additional adverse events of interest, therefore the list of terms to be used for each event of interest and the specific events of interest will be based on the SRT agreements in place at the time of reporting.

Summaries of the number and percentage of participants with these events will be provided for each type of events separately by preferred term and maximum grade. The time to onset and duration of first occurrence for each type of events will be summarized using summary statistics mean, standard deviation, median, minimum value, and maximum. The number and percentage of participants who have time to onset of first occurrence in categories of (1-21, 22-42, 43-63, >63 days) will be reported for corneal events (CTCAE) and thrombocytopenia. The number and percentage of participants who have time to onset of first occurrence in categories of (0-6, >6-12, >12-18, >18-24, >24 hours) will be reported for infusion-related reactions. The number and percentage of participants who have duration of first occurrence in categories of (1-21, 22-42, 43-63, 64-84, 85-105, >105 days) will be reported for corneal events (CTCAE) and thrombocytopenia. The number and percentage of participants who have duration of first occurrence in categories of (0-12, >12-24, >24 hours) will be reported for infusion-related reactions. For an AESI which is based on a single adverse event term, the onset and duration will be calculated based on the start and end dates of the single term. For an AESI which is based on multiple adverse event terms, the onset and duration will be calculated by looking across all terms for the AESIs. The derived start date is identified as the onset of any term defined as the AESI. The derived end date is identified as last end date for any terms once all concurrent terms for the AESI have resolved, i.e., the first time a subject is free of any adverse event term defined as the AESI.

The summary of event characteristics will be provided for each AESI respectively, including number of participants with any event, number of events, number of participants with any event that is serious, number of participants with any event that is related to study intervention, number of occurrences (One, Two, Three or more), maximum grade, maximum grade for events related to study intervention, outcomes and the action taken for the event. The percentage will be calculated in two ways, one with number of participants with event as the denominator and the other with total number of participants as the denominator. The worst-case approach will be applied at participant

level for the maximum grade, i.e. a participant will only be counted once as the worst case from all the events experienced by the participant. For action taken to an event, a participant will be counted once under each action, e.g. if a participant has an event leading to both study intervention discontinuation and dose reduction, the participants will be counted once under both actions.

An overview summary of corneal events (CTCAE), including counts and percentages of participants with any AE, AEs related to study intervention, Grade 3 and 4 AEs, Grade 3 and 4 AEs related to study intervention, AEs leading to permanent discontinuation of study intervention, study intervention related AEs leading to permanent discontinuation of study intervention, AE leading to dose reductions, AEs leading to dose delays, SAEs, SAEs related to study intervention will be produced. Unless otherwise specified, AEs will be summarized by treatment arms. AEs related to pomalidomide and dexamethasone separately, OR pomalidomide and dexamethasone in combo will all be considered as AEs related to the pom/dex regimen and summarized under the pom/dex treatment arm.

For each of these AESI, a summary of cumulative incidence by number of cycles (<=1, <=2, <=4, <=6, <=8, <=10, Any) received at first occurrence will be provided.

For thrombocytopenia, number and percentage of participants with grade 3 or 4 platelet count decreased (based on lab data) and concomitant grade 2 or above bleeding event will be summarized. A bleeding event will be considered as concomitant only if the start date is within ± 3 days of the lab event.

4.5.2.2. Deaths

All deaths will be summarized based on the number and percentage of participants. This summary will classify participants by time of death relative to the last dose of medication (>30 days or ≤30 days) and analyze the primary cause of death in the order listed in the CRF. A supportive listing will be generated to provide participant-specific details on participants who died.

4.5.2.3. Adverse Events Leading to Discontinuation of Study Treatment and Other Significant Adverse Events

The following categories of AEs will be summarized separately in descending order of total incidence by PT only and separate supportive listings will be generated with participant level details for those participants:

- AEs leading to permanent discontinuation of study treatment
- AEs leading to dose delays
- AEs leading to dose interruptions
- AEs leading to dose reductions

Only listings will be provided for the following:

- AEs leading to infusion stopped early and not completed
- AEs leading to infusion interrupted but completed

4.5.2.4. COVID-19 Assessment and COVID-19 AEs

A standardized MedDRA Query (SMQ) will be used to identify all COVID-19 AEs.

The overall incidence of AEs and SAEs of COVID-19, COVID-19 AEs leading to study intervention discontinuation, COVID-19 AEs leading to study withdrawal, and Grade 3 and 4 COVID-19 AEs will be summarized. The incidence of these events at individual PT level can be obtained from the standard AE/SAE summaries.

4.5.2.5. Impact of COVID-19 Pandemic on Safety Results

The study began in most countries after the pandemic began. Hence, the impact of the COVID-19 pandemic on the safety results will not be assessed, e.g. tables showing incidence rates for events occurring before and after the start of the COVID-19 pandemic won't be produced.

4.5.3. Ocular Findings on Ophthalmic Exam

4.5.3.1. Ocular Exam and Visual Acuity

As outlined in study protocol (GlaxoSmithKline Document Number:2017N336101_00) Schedule of Activities (Section 2 of the protocol), ocular exams are scheduled at screening, while on treatment, and at end of treatment for participants in both arms. Ocular exams in follow-up period (if needed) will only be conducted for Arm 1(belantamab mafodotin). The ocular findings and visual acuity from ocular exams will be summarized descriptively:

- **From baseline to last follow-up, the following analyses will be performed**

i. Visual acuity

- The best corrected visual acuity (BCVA) summary will be based on the Logarithm of the Minimum Angle of Resolution (logMAR score), where:

$$\text{logMAR score} = -\log_{10}(\text{Snellen Acuity Score})$$

- The following categories of logMAR score changes from baseline are defined: No change/improved vision is defined as a change from baseline <0.12; a possible worsened vision is defined as a change from baseline >=0.12 to <0.3; a definite worsened vision is defined as a change from baseline >=0.3 logMAR score.
- A summary of characteristics of worsened vision (logMAR Score change from baseline >=0.12) by subject will be provided, including time to onset of first occurrence: summary statistics and frequency/percentage in categories (1-21, 22-42, 43-63, 64-105, >105 days); outcome of first occurrence, duration of first occurrence: summary statistics and frequency/percentage in categories (1-21, 22-42, 43-63, 64-105, >105 days); number of occurrences based on participants with worsened vision; outcome post treatment exposure; time to resolution post treatment exposure, outcome of last event. The duration is defined as time from onset of worsened vision to the first time the subject is free of worsened vision (i.e. free of >=0.12 logMAR Score change from

baseline). It requires at least one day gap between the resolution of all worsened vision from first occurrence to the onset of second occurrence.

- In addition, a summary of worst change from baseline (based on the eye with worst change) in BCVA Score (logMAR Score) will be provided for categories "increase >=0.12 to <0.3", "increase >=0.3 to <0.6", "increase >=0.6".
- Shift table for visual acuity from baseline to worst case post-baseline by eye (R/L) will be provided

ii. Corneal Exam

- Shift table from baseline to worst case post-baseline by eye (R/L) for corneal epithelium findings:
  - Corneal epithelium (Normal to Abnormal),
  - Microcyst-like deposits (No to Yes).
  - Subepithelial haze (No to Yes)
  - Stromal opacity (No to Yes)
  - Epithelial edema (No to Yes)
  - Corneal epithelia defect (No to Yes)
  - Corneal erosion (No to Yes)
  - Corneal ulcer (No to Yes)
  - Corneal neovascularization (No to Yes)
  - Superficial punctate keratopathy severity (No to yes)
  - Stippled peripheral corneal staining ±vortex/whorl staining pattern (No to Yes), summarize only for patients enrolled under protocol amendment 1.

iii. Lens

- Shift table from baseline to worst case post-baseline by eye (R/L) for lens findings:
  - Cataract (No to Yes)

Supportive listings may be provided, e.g.

- Listing of participants which fall into each of the two categories of change from baseline in BCVA: "possible worsened vision", "definite worsened vision" with sub-categories "increase >=0.3 to <0.6", and "increase >=0.6".
- Listing of participants who fall into each of the two categories of decline in BCVA to 'light perception' (LP) or 'no light perception' (NLP) anytime post-baseline.
- Listings of participants with cataracts at baseline, pseudophakia at Baseline, who developed cataracts post-baseline, and who underwent cataract surgery post-baseline will provided.
- Listing of impact on driving and reading if data are available
- Listing of corneal exam results

Details of the displays are to be provided in Output and Programming Specification document.

4.5.3.2. Corneal Events Based on Keratopathy and Visual Acuity Scale (KVA Scale)

For ocular exam visits based on the ocular worksheet under the original protocol, KVA grade is not expected to be collected. For ocular exam visits based on the ocular worksheet under the protocol amendment 1, KVA grade is expected to be collected. Analyses in this section will be based on collected / investigator reported KVA scale only.

Unless otherwise specified, for the following analyses, corneal events (KVA scale) will be summarized by treatment arms based at subject level (subject level KVA grade at each visit is defined as worse-eye grade at each visit).

An overview summary of corneal events (KVA scale) will be provided, including numbers and percentages of participants with any corneal events (KVA scale), and corneal events Grade 2 or above and Grade 3 or above (KVA scale).

In addition to the overview summary described above, an additional overview summary of corneal events (KVA scale) will be provided for only those participants enrolled under protocol amendment 1, including numbers and percentages of participants with any corneal events (KVA scale) leading to permanent discontinuation of study intervention, corneal events (KVA scale) leading to dose reductions, and corneal events (KVA scale) leading to dose delays.

A summary of characteristics of corneal events (KVA scale) of grade 2 or above by subject will be provided, including time to onset of first occurrence: summary statistics and frequency/percentage in categories (1-21, 22-42, 43-63, 64-105, >105 days); outcome of first occurrence, duration of first occurrence: summary statistics and frequency/percentage in categories (1-21, 22-42, 43-63, 64-105, >105 days); number of occurrences based on participants with corneal events (KVA scale) of grade 2 or above; outcome post treatment exposure; time to resolution post treatment exposure; outcome of last event. The duration is defined as time from onset of any corneal events (KVA scale) of grade 2 or above to the first time the subject is free of overall KVA grade >=2. It requires at least one day gap between the resolution of all events from first occurrence to the onset of second occurrence.

A summary of characteristics of corneal exam findings (KVA scale) of grade 2 or above by subject will be provided, including time to onset of first occurrence: summary statistics and frequency/percentage in categories (1-21, 22-42, 43-63, 64-105, >105 days); outcome of first occurrence, duration of first occurrence: summary statistics and frequency/percentage in categories (1-21, 22-42, 43-63, 64-105, >105 days); number of occurrences based on participants with corneal exam findings (KVA scale) of grade 2 or above; outcome post treatment exposure; time to resolution post treatment exposure; outcome of last event. The duration is defined as time from onset of any corneal exam findings (KVA scale) of grade 2 or above to the first time the subject is free of any corneal exam findings with KVA scale grade >=2. It requires at least one day gap between the resolution of all events from first occurrence to the onset of second occurrence.

A summary of characteristics of visual acuity (KVA scale) of grade 2 or above by subject will be provided, including time to onset of first occurrence: summary statistics and

frequency/percentage in categories (1-21, 22-42, 43-63, 64-105, >105 days); outcome of first occurrence, duration of first occurrence: summary statistics and frequency/percentage in categories (1-21, 22-42, 43-63, 64-105, >105 days); number of occurrences based on participants with visual acuity (KVA scale) of grade 2 or above; outcome post treatment exposure; time to resolution post treatment exposure; outcome of last event. The duration is defined as time from onset of any visual acuity (KVA scale) of grade 2 or above to the first time the subject is free of visual acuity with KVA scale grade>=2. It requires at least one day gap between the resolution of all events from first occurrence to the onset of second occurrence.

For the summary of characteristics tables above, the end of treatment exposure is defined as last infusion date +20 days for belantamab Mafodotin arm; the end of treatment exposure is defined as overall study treatment end date (i.e. the end date of last cycle) for Pom/Dex arm as defined in Section 4.5.1.

A summary of corneal events (KVA scale) by grade and visit will be provided, worst-case post-baseline overall KVA grade will also be provided.

A summary of first dose delay due to corneal events (KVA Scale) will be provided for patients enrolled under protocol amendment 1 and belantamab mafodotin arm only, since under the original protocol, KVA scale is not used as guidance for dose modification. The table includes number of participants with dose delay due to corneal events (KVA Scale), grade of corneal events (KVA Scale) at onset of the dose delay; maximum grade of corneal events (KVA Scale) during the first dose delay; grade of corneal events (KVA Scale) at restart following the first dose delay; duration of first dose delay; frequency and percentage of whether dose is reduced at restart. The onset of dose delay is defined as previous dose date + 21 days (cycle length) + 3 days (visit window).

Details of the displays are to be provided in Output and Programming Specification document.

4.5.4. Additional Safety Assessments (if applicable)

4.5.4.1. Laboratory Data

Laboratory evaluations including the analyses of Chemistry laboratory tests, Haematology laboratory tests, Urinalysis, and liver function tests will be based on GSK Core Data Standards.

Summary of change from baseline by scheduled visits using mean, median, standard deviation, minimum and maximum will be provided.

Summaries of worst case grade increase from baseline grade will be provided for all the lab tests that are gradable by CTCAE v5.0. These summaries will display the number and percentage of participants with a maximum post-baseline grade increasing from their baseline grade. Any increase in grade from baseline will be summarized along with any increase to a maximum grade of 3 and any increase to a maximum grade of 4. Missing baseline grade will be assumed as grade 0. For laboratory tests that are graded for both low and high values, summaries will be done separately and labelled by direction, e.g., sodium will be summarized as hyponatremia and hypernatremia separately.

For lab tests that are not gradable by CTCAE v5.0, summaries of worst case changes from baseline with respect to normal range will be generated. Decreases to low, changes to normal or no changes from baseline, and increases to high will be summarized for the

worst case post-baseline. If a subject has a decrease to low and an increase to high during the same time interval, then the subject is counted in both the "Decrease to Low" categories and the "Increase to High" categories.

Separate summary tables for haematology, and chemistry laboratory tests will be produced. Liver function laboratory tests will be included with chemistry lab tests.

For spot urine albumin/creatinine ratio (mg/g), a shift table from baseline to worst post-baseline will be provided.

A supporting listing of laboratory data for participants with abnormalities of potential clinical concern will be provided. A separate listing of laboratory data with character values will also be provided.

Unless otherwise specified, the denominator in percentage calculation at each scheduled visit will be based on the number of participants with non-missing value at each particular visit.

Summaries of hepatobiliary laboratory events including possible Hy's law cases will be provided in addition to what has been described above. Possible Hy's law cases are defined as any elevated alanine aminotransferase (ALT)>3×upper limit of normal (ULN), total bilirubin≥2×ULN and alkaline phosphatase (ALP)<3×ULN/missing. Total bilirubin≥2×ULN can be within 28 days following the ALT elevation and if direct bilirubin is available on the same day, it must be ≥ 35% of total bilirubin. ALP<3×ULN/missing means it is satisfied unless the ALP is ≥3xULN at the time of bilirubin elevation. The summary will be produced for worst case post baseline only.

An e-DISH plot of maximum post baseline total bilirubin versus maximum post baseline ALT will be created.

The following plots will also be provided:

- maximum post baseline ALT versus baseline ALT
- maximum post baseline AST versus maximum post baseline LDH
- maximum post baseline AST versus maximum post baseline Creatinine Kinase
- maximum post baseline LDH versus maximum post baseline Creatinine Kinase

maximum post baseline urine albumin versus maximum post baseline urine CreatinineA summary of Liver Monitoring/Stopping Event Reporting will be provided. The medical conditions data for participants with liver stopping events will be listed. The substance use data for participants with liver stopping events will be listed.

4.5.4.2. Vital Signs

Values of vital signs (temperature, systolic and diastolic blood pressure, pulse rate, respiratory rate, and oxygen saturation (measured by pulse oximetry) as well as the change from baseline will be summarized by scheduled visit using mean, median, standard deviation, minimum and maximum.

A summary of changes in heart rate comparing the baseline value to the worst-case post baseline value will be provided. Heart rate will be categorized into 'Decrease to <60', 'Change to Normal or No Change' and 'Increase to >100'. The determination of the worst-case post baseline considers both scheduled and unscheduled assessments. If a

participant has a decrease to low and an increase to high, then the participant is counted in both the “Decrease to <60” categories and the “Increase to >100” categories.

In addition, summaries of grade increase in systolic blood pressure (SBP) and diastolic blood pressure (DBP) will be provided separately. These summaries will display the number and percentage of participants with any grade increase, increase to Grade 2 and increase to Grade 3 for worst-case post-baseline only. The grade definition for SBP (mmHg) is: Grade 0 (<120), Grade 1 (120-139), Grade 2 (140-159), Grade 3 (>=160). The grade definition for DBP is: Grade 0 (<80), Grade 1 (80-89), Grade 2 (90-99), Grade 3 (>=100). The summaries will be produced for worst-case post baseline only.

4.5.4.3. ECG

A listing of QTc values of potential clinical importance will be provided using the collected values based on Fridericia formula.

4.5.4.4. Performance Status

ECOG performance status will be summarized at baseline and each post-baseline scheduled visit. Summaries will use frequency and percentage of participants at each planned assessment time. A summary of change from baseline by scheduled visits will be performed, as well as the worst case post-baseline and the best case post-baseline changes during the study (improved, no change, deteriorated).

A supporting listing will also be provided.

4.5.4.5. Pregnancies

While pregnancy itself is not considered to be an AE or SAE, any pregnancy complication or elective termination of a pregnancy for medical reasons will be recorded as an AE or SAE as described in the protocol. If participants become pregnant while on the study, the information will be included in the narratives and no separate table or listing will be produced.

CCI

CCI

4.6. Other Analyses

4.6.1. Population Pharmacokinetic Analysis

Plasma belantamab mafodotin, total mAb, and/or cys-mcMMAF concentration-time data may be combined with data from other studies and analysed using a population pharmacokinetic approach. The initial analysis will use the population pharmacokinetic model developed for Study BMA117159 and Study 205678 to generate post hoc pharmacokinetic parameter estimates for the individual participants in Arm 1of Study 207495. Based on the individual post hoc parameter values, dosing information, and sample collection times, belantamab mafodotin plasma concentrations at the time of sample collection will be predicted for each participant. Model evaluation will consist of comparison of model-predicted and observed concentrations. If necessary, model estimation or refinement will be performed.

Details of the population PK analyses will be reported under a separate RAP, and the results of this analysis will be provided in a separate report.

4.6.2. Subgroup analyses

4.6.2.1. Subgroup analyses of PFS

Subgroup analyses of PFS will be based on the ITT analysis set.

The following subgroup analyses will be performed to compare the primary estimand of PFS between treatments, based on investigator-assessed response, respectively, if data permit.

Subgroup	Categories [1]
Previous treatment with anti-CD38 therapy [2]	Yes, No
ISS staging [2]	I/II; III
Number of prior lines of therapy [2]	≤3; >3
Previous treatment with anti-CD38 therapy and number of prior lines of therapy	Yes and >3; Other
Previous treatment with melphalan	Yes, No
Previous stem cell transplant	Yes, No
Age at randomization	<65, ≥65 to <75, ≥75 years age
Gender	Male, Female
Ethnicity	Hispanic, Non-Hispanic
Race	Black, White, Other
Region of Enrollment	North America, Europe, North East Asia, Rest of the world
Baseline renal impairment status per eGFR (ml/min/1.73 m^2)	Normal (≥ 90), Mild (≥ 60, < 90), Moderate (≥ 30, < 60), Severe (≥ 15, < 30)
Type of myeloma	IgG, Non-IgG

Subgroup	Categories [1]
Baseline extramedullary disease	Yes, No
Cytogenetics Risk [3]	High, Other (non-high risk - all others)
Refractory to prior anti-cancer therapy	Any Proteasome Inhibitor (PI) Bortezomib Carfilzomib Ixazomib Any Immunomodulator (IMiD) Thalidomide Lenalidomide Pomalidomide Any Monoclonal Antibody Elotuzumab Isatuximab Daratumumab Daratumumab alone [4] Daratumumab in combination [5] PI+IMiD Anti-CD38 antibody+PI+IMiD

[1]. If the percentage of participants is small within a particular subgroup, then the subgroup categories may be refined prior to unblinding the trial.
[2]. Randomization factors
[3]. A subject is considered as high risk if the subject has any of the following cytogenetics: t(4;14), t(14;16), and del(17p13).
[4]. Defined as prior CTX regimen with Daratumumab as the only drug in the regimen.c
[5]. Defined as prior CTX regimen with Daratumumab and other drugs.

Additional subgroups CCI [redacted] may be assessed if there is clinical justification or an imbalance is observed between the treatment arms. The purpose of the subgroup analyses is to assess the consistency of treatment effect across expected prognostic factors.

The ITT analysis set will be used in the subgroup analyses. The subgroup analyses will be based on values recorded on the eCRF (or vendor data if collected outside of eCRF).

CCI [redacted]

For each subgroup, the hazard ratio (HR) and 95% CI will be calculated from an unstratified Cox proportional hazards model with treatment as the only covariate. HR and 95% CI for each subgroup level will be provided separately; refer to Section 4.2.2 for the analysis method and the subgroup analysis will only be performed for PFS based on primary event/censoring rule. The HRs and associated two-sided 95% CIs will be summarized and presented on a forest plot, along with the results of the overall analysis (i.e. analysis based on the whole ITT analysis set).

If there are too few events available for a meaningful analysis of a particular subgroup (it is not considered appropriate to present analyses where there are less

than 20 events in a subgroup), the subgroup will not be formally analyzed by unstratified Cox proportional hazards model. In this case, only descriptive summaries will be provided.

Subgroup PFS analysis will be conducted at final PFS analysis.

4.6.2.2. Subgroup analyses of OS

Subgroup analyses of OS will be based on the ITT analysis set.

Subgroup OS analyses will be performed based on subgroups as specified in Section 4.6.2.1. The HRs and associated two-sided 95% CIs will be summarized and presented on a forest plot, along with the results of the overall analysis (i.e. analysis based on the whole ITT analysis set). Refer to Section 4.6.2.1. for subgroups and analysis details.

Subgroup OS analysis will be conducted at final OS analysis

4.7. Planned Analyses

Table 8 summarizes the planned analyses for PFS and OS.

Table 8 Summary of PFS and OS analyses

Analyses Timing from Randomization	Planned PFS analyses	Planned OS analyses
~25 months	PFS final ~151 PFS events	OS IA for efficacya ~100 OS events ~40% OS info
48 months	N/A	OS IA for efficacyb ~175 OS events ~70% OS info
>60 months	N/A	OS finalb ~250 OS events

a. Provided that final PFS analysis is significant.
b. If PFS is significant and null hypothesis is not rejected at the 2nd OS IA for efficacy.
c. First dose of 320th subject happens around 21 months

Planned Analyses for Key secondary endpoint OS

The key secondary efficacy endpoint, OS, will be compared between the 2 treatment arms (belantamab mafodotin arm vs pomalidomide/ dexamethasone arm) using a group sequential log-rank test corresponding to three analyses: one interim analysis at the same time of PFS final (~40% information), one interim analysis at ~70% information and one final at 100% information. The boundary for declaring superiority of belantamab mafodotin arm over pomalidomide/ dexamethasone arm is based on a Lan-DeMets (O'Brien-Fleming) alpha spending function [Lan, 1983] with overall alpha = 0.025, one-

tailed, and will be adjusted based on actual observed number of deaths (actual information fraction).

At the first OS interim analysis for efficacy (~40% OS information fraction), the study will not be stopped regardless of the results. The OS data will continue to be collected to conduct the 2nd OS interim analysis to test for efficacy at ~70% OS information fraction.

At the OS interim analysis for efficacy (~70% OS information fraction) (Table 9), assuming 175 events are observed, if the calculated p<0.0073 (corresponding to HR<0.676), it would be considered that the efficacy boundary is crossed and the OS analysis would be significant.

A hierarchical testing procedure will be adopted and the OS interim/final analysis for efficacy will only be performed if the primary efficacy endpoint PFS is statistically significant at PFS final analysis [Bretz, 2009; Li, 2017]. The testing procedures are detailed as follows:

Step 1: Test PFS at the final PFS analysis. If significant, go to Step 2 and overall one-sided alpha of 0.025 will be carried forward to test for OS; if not significant, stop testing;

Step 2: Test OS at the first OS interim analysis (at the same time of PFS final, ~40% OS information fraction) for efficacy. (cumulative one-sided alpha spent for OS = 0.0004)

Step 3: Test OS at the second OS interim analysis (~70% OS information fraction) for efficacy. If significant, stop testing; if not significant, go to Step 4. (cumulative one-sided alpha spent for OS = 0.0074)

Step 4: Test OS at the time of the final OS analysis (cumulative one-sided alpha spent for OS = 0.025).

Table 9 Stopping boundaries for OS

Information fraction	N of events	Cum. alpha Spent	Efficacy Boundary	Efficacy Boundary	Boundaries crossing probabilities (incremental)	
			(p-value)	(HR)	Under H0	Under H1
0.4	100	0.0004	0.0004*	0.491*	0.04%	5.8%
0.7	175	0.0074	0.0073	0.676	0.7%	41.2%
1	250	0.025	0.0227	0.765	1.8%	33.1%

*: The threshold will be is used for the 1st OS analysis, however, the study will not be stopped for efficacy regardless of the results. GSK will continue to collect the OS data to conduct the 2nd OS interim analysis to test for efficacy at ~70% OS information fraction.

If OS is significant at 70% IF, OS data will continue to be collected until end of study: defined as all participants have died, are lost to follow up, or withdrawn consent, or for 2 years after the OS analysis at ~70% IF, whichever occurs first. For the two scenarios above, analysis of OS will be performed at the end of study using only Kaplan-Meier method and stratified Cox proportional hazards model.

4.7.1. Independent Data Monitoring Committee

An Independent Data Monitoring Committee (IDMC) consisting of at least 2 physicians and one statistician as defined in the IDMC Charter will review data at defined time points. Additional details, including the list of outputs supporting decision making at the interim analysis, will be provided in the IDMC charter.

The first IDMC safety review meeting is planned when approximately 60 participants have been on study for at least 8 weeks. Subsequent IDMC safety review meeting is planned approximately every 6 months hereafter.

4.8. Changes to Protocol Defined Analyses

There were no changes or deviations to the originally planned statistical analysis specified in the protocol amendment 3 (Version: GSK Document Number TMF-13954921, Dated: [20-SEP-2021]).

5. SAMPLE SIZE DETERMINATION

5.1. Statistical Hypotheses

Primary endpoint PFS

The primary efficacy analysis will be the comparison of the distribution of PFS between the two treatment groups. Assuming proportional hazards for PFS, the following statistical hypothesis will be tested to address the primary efficacy objective at one-sided alpha level of 2.5%:

$$H_{01}: \Theta_1 \geq 1 \quad \text{VS.} \quad H_{A1}: \Theta_1 < 1$$

where, θ_1 is the PFS HR (belantamab mafodotin arm vs. pom/dex arm).

Key secondary endpoint OS

Assuming proportional hazards for OS, the following statistical hypotheses will be tested at one-sided alpha level of 2.5% if PFS is statistically significant:

$$H_{02}: \Theta_2 \geq 1 \quad \text{VS.} \quad H_{A2}: \Theta_2 < 1$$

where, θ_2 is the OS HR (belantamab mafodotin arm vs. pom/dex arm).

5.2. Sample Size Determination

Primary Endpoint PFS

Based on available data from literature, the median PFS in the pom/dex arm is expected to be around 4 months [San Miguel, 2013]. It is expected that treatment with belantamab mafodotin will result in a 43% reduction in the hazard rate for PFS, i.e. an expected HR of 0.57 (corresponding to an increase in median PFS from 4 months to 7 months under the exponential model assumption).

The final PFS analysis will be conducted at the time of observing approximately 151 events and the first 320 randomized subjects have been followed for a minimum of 4 months. With 151 events, the study has a power of 90% to detect a hazard ratio of 0.57 at 1-sided alpha of 0.025 (corresponding to a critical value of 0.713 for the hazard ratio). This calculation assumes participants randomized to the two treatment arms in a 2:1 ratio. Assuming that enrolment will continue for approximately 20 months at a uniform rate of 16 participants per month, a total of 320 participants will be randomized in a 2:1 ratio to receive single agent belantamab mafodotin or pom/dex. It is estimated that the targeted 151 PFS events will be observed approximately 23 months after the first participant is randomized based on a lognormal cure rate model (Chen, 2016). These calculations were made using the software package East 6.5 and a proprietary SAS macro.

Power for Analysis of Key Secondary Endpoint OS

OS, as the key secondary endpoint, will be formally statistically tested, provided that the primary endpoint PFS is statistically significant. Based on available data from literature, the median OS in the pom/dex arm is expected to be around 13 months [San Miguel, 2013]. It is hypothesized that treatment with belantamab mafodotin will result in a 32% reduction in the hazard rate for OS, i.e., an expected HR of 0.68 (which corresponds to an increase in median OS to 19 months under the exponential model assumption). In order to ensure 80% power to test the null hypothesis: OS HR = 1, versus the specific alternative hypothesis: OS HR = 0.68, a total of 250 deaths need to be observed. This calculation assumes analysis by a one-sided log-rank test at the overall 2.5% level of significance, participants randomized to the two treatment arms in a 2:1 allocation ratio, and a group sequential design with a Lan-DeMets (O'Brien-Fleming) alpha spending function [Lan, 1983] using information fractions of (0.4, 0.7, 1). Based on the same number of participants that are planned to be enrolled in this study to provide sufficient power for the primary endpoint (i.e., 320 participants), it is estimated that these 250 deaths will be observed approximately 55 months after the randomization date of the first participant under H_{A2} (assuming a similar loss to follow-up, i.e. 5% per year). Therefore, the cut-off date for the final analysis of OS will be approximately 35 months after the cut-off date for the final analysis of PFS. These calculations were made using the software package East 6.5.

If the number of participants required by local regulatory agencies are not recruited within the planned recruitment target, enrolment may continue in separate cohorts until the country enrolment requirements, as required by local regulatory bodies, have been reached. Additional participants that are enrolled in separate cohorts will not be included in the analysis portion of the study planned for the marketing application, which is based on approximately 151 events. However, these additional participants will be included in country-specific supplemental analyses, as detailed in country specific Statistical Analysis Plan (SAP), requested by the applicable regulatory authorities concerned.

6. SUPPORTING DOCUMENTATION

6.1. Appendix 1 Study Population Analyses

Unless otherwise specified, the study population analyses will be based on the ITT Analysis Set. Study population analyses including analyses of subject's disposition, protocol deviations, demographic and baseline characteristics, prior and concomitant medications, and subsequent anti-cancer therapy will be based on GSK Core Data Standards.

6.1.1. Participant Disposition

A summary of the number of participants in each of the analysis set described will be provided (for primary analysis, Evaluable population will not be included). In addition, the number of participants enrolled by centre will be summarized by dose level using the "Enrolled" population. A summary of subject status and reason for study withdrawal will be provided. This display will show the number and percentage of participants who withdrew from the study, including primary reasons for study withdrawal. Reasons for study withdrawal will be presented in the order they are displayed in the eCRF.

A summary of study treatment status will be provided. This display will show the number and percentage of participants who are ongoing or discontinued study treatment and a summary of the primary reasons for discontinuation of study treatment. Reasons for study treatment discontinuation will be presented in the order they are displayed in the eCRF. A listing of study treatment discontinuation will be generated. The listing will include last dose date, and reasons for study treatment discontinuation.The 'Summary of Subject Status and Subject Disposition for the Study Conclusion Record' and the 'Summary of Treatment Status and Reasons for Discontinuation of Study Treatment' will be repeated, with the reason for withdrawal/discontinuation categorised as due to the COVID-19 pandemic, or non-due to the COVID-19 pandemic based on information collected on the COVID-19 Pandemic Study Impact form.

In this multicenter global study, enrolment will be presented by country and site.

Data from all participating centers will be integrated and no controlling for center-effect will be considered in the statistical analyses. It is anticipated that patient accrual will be spread thinly across centers and summaries of data by center is unlikely to be informative and will not be provided.

6.1.2. Demographic and Baseline Characteristics

The demographic characteristics (e.g., age, race, ethnicity, sex, baseline height, and baseline body weight and baseline BMI) will be summarized and listed. Age, height, weight and BMI will be summarized using the mean, standard deviation, minimum, median, and maximum. The count and percentage will be computed for sex and ethnicity.

Race and racial combinations will be summarized and listed.

Disease history and characteristics (e.g. time since initial diagnosis in years, stage at initial diagnosis, date of initial diagnosis) at initial diagnosis and screening will be listed. Separate summaries of disease characteristics at initial diagnosis and screening will be provided. Disease characteristics at screening, including stage, type of multiple myeloma, number of prior lines, and types of therapy, myeloma light chain and myeloma immunoglobulin, extramedullary disease, lytic bone lesion, and genetic characteristics (including high cytogenetic risk) will be summarized and listed.

Medical conditions collected at screening will be listed and will be summarized by past and current and by cancer-related and non-cancer related categories.

Substance use, including smoking history and alcohol use will be summarized.

Prior anti-cancer therapy will be coded using GSK Drug coding dictionary, then summarized by type of therapy and listed. A listing of prior anti-cancer therapy will show the relationship between ATC Level 1, Ingredient, and verbatim text. A summary of the best response to the most recent prior anti-cancer therapy will be provided. A summary of the number of prior anti-cancer therapy regimens will also be produced.

Prior anti-cancer therapy for multiple myeloma participants will also be summarized by type of therapy, and drug class. A summary of multiple myeloma participants' refractory to prior anti-cancer therapy by drug class will be provided.

Anti-cancer radiotherapy will be listed. Prior cancer and non-cancer related surgeries will be summarized. Prior and on treatment cancer and non-cancer related surgeries will be listed.

6.1.3. Protocol Deviations

Important protocol deviations will be summarized.

Protocol deviations will be tracked by the study team throughout the conduct of the study. These protocol deviations will be reviewed to identify those considered as important as follows:

- Data will be reviewed prior to freezing the database to ensure all important deviations (where possible without knowing the study intervention details) are captured and categorised in the protocol deviations dataset.
- This dataset will be the basis for the summaries of important protocol deviations.

In addition to the overall summary of important protocol deviations, separate summaries will be produced for important protocol deviations related to COVID-19, and important protocol deviations not related to COVID-19 respectively.

6.1.4. Prior and Concomitant Medications

Concomitant medications will be coded using both the GSK Drug and WHO Drug dictionaries. However, the summary will be based on GSK Drug dictionary only. The summary of concomitant medications will show the number and percentage of participants taking concomitant medications by Ingredient. Multi-ingredient products will

be summarized by their separate ingredients rather than as a combination of ingredients. Anatomical Therapeutic Chemical (ATC) classification Level 1 (Body System) information will be included in the dataset created but will not appear on the listing or summary.

In the summary of concomitant medications, each subject is counted once within each unique ingredient. For example, if a subject takes Amoxycillin on two separate occasions, the subject is counted only once under the ingredient "Amoxycillin". In the summary of concomitant medications, the ingredients will be summarized by the base only, using CMBASECD and CMBASE. Note: In order to be considered a concomitant medication, the concomitant medication must have been taken at some point during the on-treatment window.

Prophylactic medication for infusion-related reactions and prophylactic topical eye medications will be summarized by drug class and drug name and listed separately.

Blood products or blood supportive care products with onset date within the on-treatment window will be included in the summary tables. The frequency and percentage of participants using blood products and blood supportive care products after the start of study medication will be provided. Supporting listings will also be provided.

6.1.5. Subsequent Anti-Cancer Therapies

The number and percentage of participants that received chemotherapy, immunotherapy, hormonal therapy, biologic therapy, radioactive therapy, small molecule targeted therapy, as subsequent anti-cancer therapy will be summarized. Time from study treatment discontinuation to the first post study treatment anti-cancer therapy will also be included in this summary table, if available.

Follow-up anti-cancer therapy will be coded using GSK Drug coding dictionary, then summarized by ingredient. A listing of the type of follow-up anti-cancer therapy received (chemotherapy, immunotherapy, hormonal therapy, biologic therapy, radioactive therapy, and small molecule targeted therapy) for each subject will be provided.

6.1.6. Additional Analyses Due to the COVID-19 Pandemic

A participant is defined as having a suspected, probable or confirmed COVID-19 infection during the study if the answer is "Confirmed", "Probable" or "Suspected" to the case diagnosis question from the COVID-19 coronavirus infection assessment eCRF.

Summaries and listings of the numbers of participants with a suspected, probable or confirmed COVID-19 infection, and of COVID-19 test results will be based on GSK Core Data Standards.

6.2. Appendix 2 Data Derivations Rule

6.2.1. Criteria for Potential Clinical Importance

6.2.1.1. Laboratory Values

Reference ranges for all laboratory parameters collected throughout the study are provided by the laboratory. A laboratory value that is outside the reference range is considered either high abnormal (value above the upper limit of the reference range) or low abnormal (value below the lower limit of the reference range). Note: a high abnormal or low abnormal laboratory value is not necessarily of clinical concern. The laboratory reference ranges will be provided on the listings of laboratory data. Clinical laboratory test results outside of the reference range will be flagged in the listings.

To identify laboratory values of potential clinical importance, National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE v5.0) will be used to assign grades to the relevant laboratory parameters. NCI-CTCAE v5.0 can be found at http://ctep.cancer.gov/reporting/ctc.html.

For laboratory data which are not listed in the NCI CTCAE v5.0, a summary of values outside the normal range will be provided.

6.2.1.2. ECG Parameters and vital signs

For ECG and vital signs, outputs per the most updated IDSL standard up to the SAP effective date will be provided.

Unless otherwise specified, ECG displays will be based on central reading.

6.2.2. Study Period

Assessments and events will be classified according to the date/time of occurrence relative to date/time of first dose of study treatment.

Study Phase	Definition
Pre-Treatment	Date/time ≤ Study Treatment Start Date/time
On-Treatment	Study Treatment Start Date/time < Date/time ≤ Last Dose Date + 70 days
Post-Treatment	Date/time > Study Treatment Stop Date + 70 days

For assessment or event on the first dosing day, whether it is Pre-Treatment or On-Treatment should be based on time if available. If time is not available, the first dosing day (Day 1) is considered Pre-Treatment for ECOG, ECG, vital signs, liver events, lab tests, cardiac scan, and other safety domains, and On-Treatment for adverse events and concomitant medications.

Concomitant Medication and Blood and Blood Supportive Care Product start and end dates will be assigned to study time periods in relation to first dose of study treatment as defined below. The start date references time flag variables and end date reference time flag variables will be added to the concomitant medications and blood and blood supportive products datasets, respectively.

- Start relative to treatment: Assign to 'BEFORE' if start date is prior to study treatment start date or if subject has not taken any study treatment or (start date is missing and end date is before study treatment start date). Else assign to 'DURING' if the start date falls into the on-treatment period as defined above or if subject is ongoing (not all study treatment discontinuation records completed) or start date is missing. Else assign to 'AFTER' if start date is after the on-treatment period.
- End relative to treatment: Assign to 'BEFORE' if end date is prior to study treatment start date or if subject has not taken any study treatment. Else assign to 'DURING' if start date falls into the on-treatment period or if subject is ongoing (not all study treatment discontinuation records completed) or (end date is missing and start relative to treatment not 'AFTER'). Else assign to 'AFTER' if start date is after the on-treatment period or (end date is missing and start relative to treatment='AFTER').

Only on-treatment blood and blood supportive care products that start after the start of study treatment are included in the Blood Products and Blood Supportive Care Product summaries. Therefore, for summary tables, include blood and blood supportive care product records where start relative to treatment in ('DURING') and end relative to treatment in ('DURING','AFTER'). All data will be reported in listings.

Concomitant medication starts relative to treatment and end relative to treatment flags are used to select data to include in the Concomitant Medication summaries as follows:

- Summary of Concomitant Medications: This summary will contain medications including those with start date prior to study treatment start date and continue (missing end date or end date after study treatment start date) on therapy. Note that any medications with start date and end date prior to study treatment start date will be excluded. In addition, any medication that was started during post-therapy will be excluded. Include concomitant medication records where start relative to treatment in ('BEFORE','DURING') and end relative to treatment in ('DURING','AFTER').
- Summary of Concomitant Medications with On-Therapy Onset: This summary will contain medications with start date after study treatment start date. In addition, any medication that was started during post-therapy (see above for definition of posttherapy) will be excluded. Include concomitant medication records where start relative to treatment in ('DURING') and end relative to treatment in ('DURING','AFTER').

6.2.3. Treatment Emergent Flag for Adverse Events

Flag	Definition
Treatment Emergent	• Study Treatment Start Date ≤ AE Start Date ≤ min(Study Treatment Stop Date + 70 day, Start of anti-cancer therapy) • AE Start Date is missing

All AE displays will use the "treatment emergent" definition. One additional AE display by preferred term may be generated for the AEs between the start of anti-cancer therapy and Study Treatment Stop Date + 70 day if a subject starts an anti-cancer therapy before the Study Treatment Stop Date + 70 day.

6.2.4. Study Day and Reference Dates

Calculated as the number of days from First Dose Date:

- Ref Date = Missing → Study Day = Missing
- Ref Date < First Dose Date → Study Day = Ref Date – First Dose Date
- Ref Data ≥ First Dose Date → Study Day = Ref Date – (First Dose Date) + 1

6.2.5. Assessment Window

For data summaries by visit, scheduled visits with nominal visit description as well as the worst-case post baseline will be displayed. Unscheduled visits will not be displayed or slotted into a visit window but will be included in the derivation of worst-case post baseline assessment. All un-scheduled visits will be displayed in the listing.

6.2.6. Multiple measurements at One Analysis Time Point

When triplicate ECG assessments are taken, mean of the measurement will be calculated first and summary statistics will be based on the calculated mean. This will apply to both baseline and post baseline assessments.

For lab tests on a study day, if more than one assessment is taken on the same day, the test from a central lab will be taken over the test from a local lab. If multiple assessments are taken from the same type of lab, the worst case will be used.

6.2.7. Handling of Partial Dates

<table>
<tr><th>Element</th><th>Reporting Detail</th></tr>
<tr><td>General</td><td>• Partial dates will be displayed as captured in participant listing displays.
• However, where necessary, display macros may impute dates as temporary variables for sorting data in listings only. In addition, partial dates may be imputed for 'slotting' data to study phases or for specific analysis purposes as outlined below.
• Imputed partial dates will not be used to derive study day, time to onset or duration (e.g., time to onset or duration of adverse events), or elapsed time variables (e.g., time since diagnosis). In addition, imputed dates are not used for deriving the last contact date in overall survival analysis dataset.</td></tr>
<tr><td>Adverse Events</td><td>• Partial dates for AE recorded in the CRF will be imputed using the following conventions:
<table><tr><td>Missing start day</td><td>First of the month will be used unless this is before the start date of study treatment; in this case the study</td></tr></table></td></tr>
</table>

<table>
<tr><th>Element</th><th>Reporting Detail</th></tr>
<tr><td></td><td>
<table>
<tr><td></td><td>treatment start date will be used and hence the event is considered on-treatment</td></tr>
<tr><td>Missing start day and month</td><td>No Imputation</td></tr>
<tr><td>Missing end day</td><td>Last day of the month will be used.</td></tr>
<tr><td>Missing end day and month</td><td>No Imputation</td></tr>
<tr><td>Completely missing start/end date</td><td>No imputation</td></tr>
</table>
</td></tr>
<tr><td>Concomitant Medications/ Medical History</td><td>
Partial dates for any concomitant medications recorded in the CRF will be imputed using the following convention:
<table>
<tr><td>Missing start day</td><td>A '01' will be used for the day</td></tr>
<tr><td>Missing start day and month</td><td>A '01' will be used for the day and 'Jan' will be used for the month</td></tr>
<tr><td>Missing end day</td><td>A '28/29/30/31' will be used for the day (dependent on the month and year).</td></tr>
<tr><td>Missing end day and month</td><td>A '31' will be used for the day and 'Dec' will be used for the month.</td></tr>
<tr><td>Completely missing start/end date</td><td>No imputation</td></tr>
</table>
</td></tr>
<tr><td>New Anti-Cancer Therapy/Radiotherapy/Surgical Procedures for Efficacy Evaluation (e.g., response rate, time to event)</td><td>Start dates for follow-up anti-cancer therapy, radiotherapy (where applicable), and surgical procedures (where applicable) will be temporarily imputed in order to define event and censoring rules for progression-free survival, response rate, time to progression, duration of response or time to response (i.e. start date for new anticancer therapy). The imputed dates will be stored on the anti-cancer therapy, radiotherapy, or surgical procedure datasets.

If missing start day, month, and year, then no imputation for completely missing dates
If missing start day and month, then no imputation should be done
If missing start day, then do the following:

If partial date falls in the same month as the last dose of study treatment, then assign to earlier of (date of last dose of study treatment+1, last day of month).

</td></tr>
</table>

Element	Reporting Detail
	○ If partial date falls in the same month as the subject's last assessment and the subject's last assessment is PD, then assign to earlier of (date of PD+1, last day of month). ○ If both rules above apply, then assign to latest of the 2 dates ○ Otherwise, impute missing day to the first of the month. • If missing end date, then no imputation should be done.
Covariates for efficacy analysis (Date of initial diagnosis/Last recurrence/Last progression)	• If both month and day are missing, first of January will be used If only day is missing, first of the month will be used
Death date	If there are partial death dates, then partial death dates will be imputed in order to define event dates for overall survival. The imputed dates will be stored on the time to event dataset. • If missing start day and month, then do the following: ○ If partial date corresponds to the same year as the last known date of last contact, then assign to last known date of last contact. ○ Otherwise, impute missing day and month to the first of January. • If missing start day, then do the following: ○ If partial date falls in the same month as the last known date of last contact, then assign to last known date of last contact. ○ Otherwise, impute missing day to the first of the month.

6.2.8. Patient Reported outcome Analyses

6.2.8.1. EORTC QLQ-C30

The EORTC QLQ-C30 is a 30-item questionnaire containing both single- and multi-item measures [Aaronson, 1993]. These include five functional scales (Physical, Role, Cognitive, Emotional, and Social Functioning), three symptom scales (Fatigue, Pain, and Nausea/Vomiting), a Global Health Status/QoL scale, and six single items (Constipation, Diarrhea, Insomnia, Dyspnea, Appetite Loss, and Financial Difficulties). The below image shows the details.

Table 1: Scoring the QLQ-C30 version 3.0

	Scale	Number of items	Item range*	Version 3.0 Item numbers	Function scales
Global health status / QoL					
Global health status/QoL (revised)†	QL2	2	6	29, 30	
Functional scales					
Physical functioning (revised)†	PF2	5	3	1 to 5	F
Role functioning (revised)†	RF2	2	3	6, 7	F
Emotional functioning	EF	4	3	21 to 24	F
Cognitive functioning	CF	2	3	20, 25	F
Social functioning	SF	2	3	26, 27	F
Symptom scales / items					
Fatigue	FA	3	3	10, 12, 18	
Nausea and vomiting	NV	2	3	14, 15	
Pain	PA	2	3	9, 19	
Dyspnoea	DY	1	3	8	
Insomnia	SL	1	3	11	
Appetite loss	AP	1	3	13	
Constipation	CO	1	3	16	
Diarrhoea	DI	1	3	17	
Financial difficulties	FI	1	3	28	

* *Item range* is the difference between the possible maximum and the minimum response to individual items; most items take values from 1 to 4, giving *range* = 3.

† (revised) scales are those that have been changed since version 1.0, and their short names are indicated in this manual by a suffix "2" – for example, PF2.

Scores for each scale and single-item measure are averaged and transformed linearly to a score ranging from 0–100. (see below image for details). A high score for functional scales and for Global Health Status/QoL represent better functioning ability or HRQoL, whereas a high score for symptom scales and single items represents significant symptomatology [Proskorovsky, 2014].

Technical Summary

In practical terms, if items I_1, I_2, ... I_n are included in a scale, the procedure is as follows:

Raw score
Calculate the raw score
$RawScore = RS = (I_1 + I_2 + ... + I_n)/n$

Linear transformation
Apply the linear transformation to 0-100 to obtain the score S,

Functional scales: $S = \left\{1 - \frac{(RS-1)}{range}\right\} \times 100$

Symptom scales / items: $S = \{(RS-1)/range\} \times 100$

Global health status / QoL: $S = \{(RS-1)/range\} \times 100$

Range is the difference between the maximum possible value of *RS* and the minimum possible value. The QLQ-C30 has been designed so that all items in any scale take the same range of values. Therefore, the range of *RS* equals the range of the item values. Most items are scored 1 to 4, giving *range* = 3. The exceptions are the items contributing to the global health status / QoL, which are 7-point questions with *range* = 6, and the initial yes/no items on the earlier versions of the QLQ-C30 which have *range* = 1.

Scoring of the QLQ-C30 Summary Score

The EORTC QLQ-C30 Summary Score is calculated from the mean of 13 of the 15 QLQ-C30 scales (the Global Quality of Life scale and the Financial Impact scale are not included). Prior to calculating the mean, the symptom scales need to be reversed to obtain a uniform direction of all scales. The summary score should only be calculated if all of the required 13 scale scores are available.

QLQ-C30 Summary Score = [Physical Functioning+ Role Functioning+ Social Functioning+ Emotional Functioning+ Cognitive Functioning+ (100- Fatigue)+ (100-Pain)+ (100-Nausea_Vomiting)+ (100-Dyspnoea)+ (100-Sleeping Disturbances)+ (100-Appetite Loss)+ (100-Constipation)+ (100-Diarrhoea)]/13.

Handling of missing items

Single-item measures: if the item is missing, the score S will be set to missing.

Scales requiring multiple items: if at least half of the items from the scale are available, the score S will be calculated based on available items. If more than half of the items from the scale are missing, the score S will be set to missing (Fayers, 2001).

Minimal Important Difference (MID): In a sample of patients who received chemotherapy for either breast cancer or small-cell lung cancer (n=246, n=80 respectively), the mean change in EORTC QLQ-C30 score between baseline and follow-

up was about 5 to 10 points on a 0-100 scale for patients who indicated “a little” change on the Subjective Significance Questionnaire (SSQ), either for better or for worse (Osoba, 1998).

6.2.8.2. EORTC QLQ-MY20

EORTC QLQ-IL52 is the disease symptom domain of EORTC QLQ-MY20. The EORTC QLQ-MY20 is a supplement to the QLQ-C30 instrument used in patients with multiple myeloma [Aaronson, 1993; Cocks, 2007]. The module comprises 20 questions that address four myeloma-specific HRQoL domains: Disease Symptoms, Side Effects of Treatment, Future Perspective, and Body Image. Three of the four QLQ-MY20 domains are multi-item scales: Disease Symptoms (includes bone aches or pain, back pain, hip pain, arm or shoulder pain, chest pain, and pain increasing with activity); Side Effects of Treatment (includes drowsiness, thirst, feeling ill, dry mouth, hair loss, upset by hair loss, tingling hands or feet, restlessness/agitation, acid indigestion/heartburn, and burning or sore eyes); and Future Perspective (includes worry about death and health in the future, and thinking about illness). The Body Image scale is a single-item scale that addresses physical attractiveness. (see below image for details).

Scoring

	Scale name	Number of items	Item range	QLQ-MY20 item numbers
Functional scales / items				
Future perspective	MYFP	3	3	18 – 20
Body image	MYBI	1	3	17
Symptom scales				
Disease symptoms	MYDS	6	3	1 – 6
Side effects of treatment	MYSE	10	3	7 – 16

Remarks

- Question 12 is considered scored “not at all” if question 11 is scored “not al all”.

As with the QLQ-C30, QLQ-MY20 domain scores are also averaged and transformed linearly to a score ranging from 0–100 (see below for details).

1) Raw score

For each multi-item scale, calculate the average of the corresponding items.

$Raw\ Score = RS = \{(I_1 + I_2 + \ldots + I_n)/n\}$

For the single-item measure, the score of the concerning item corresponds to the raw score.

2) Linear Transformation

To obtain the Score S, standardize the raw score to a 0 – 100 range following the appropriate transformation:

Symptom scales: $S=\{(RS-1)/range\} \times 100$

Functional scales: $S= \{1-(RS-1)/range\} \times 100$

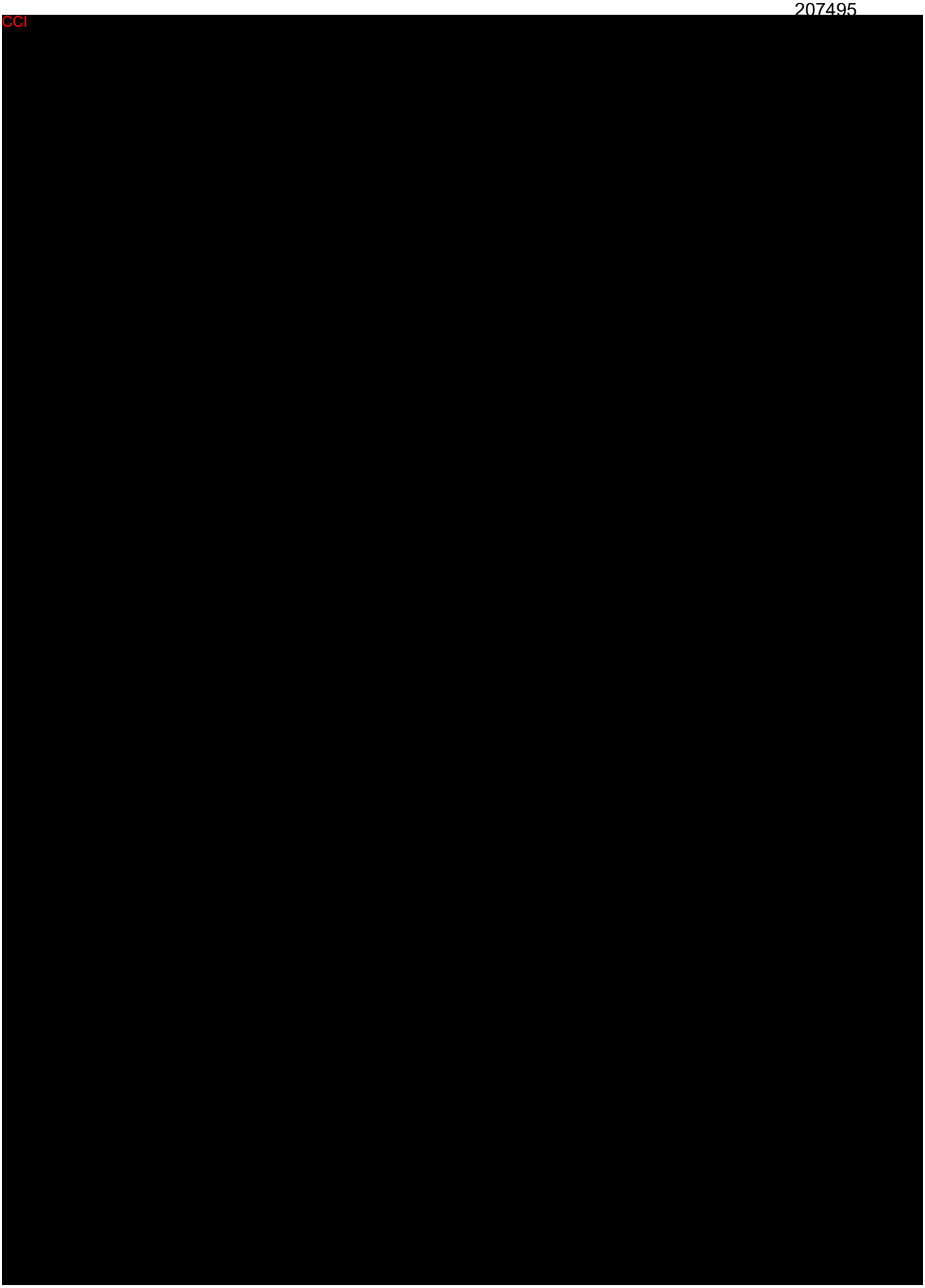
A high score for Disease Symptoms and Side Effects of Treatment represents a high level of symptomatology or problems, whereas a high score for Future Perspective and Body Image represents better outcomes [Proskorovsky, 2014].

Missing items can be handled similarly to EORTC QLQ-C30 as described in Section 6.2.8.1.

6.2.9. Extended Loss to Follow-up or Extended Time without an Adequate Assessment

For participants, if two or more scheduled disease assessments are missed and are then followed by an assessment of PD or death, PFS will be censored at the last adequate assessment prior to PD or death. When the scheduled disease assessment is every 3 weeks, a window of 49 days (6 weeks + 7 day window) 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 49 days, then PFS will be censored at the last adequate assessment prior to PD/death. In case there is no adequate assessment between PD/death and randomization date, and the time difference between PD/death and randomization date is more than 49 days, then PFS will be censored at the randomization date.

CCI



CCI

[redacted]

6.2.11. List of Abbreviations

Abbreviation	Description
ADA	Anti-drug antibody
ADC	Antibody drug conjugate
AE	Adverse event
ALT	Alanine transaminase
CCI [redacted]	
CCI [redacted]	
CCI BNP	B-type natriuretic peptide
BP	Blood pressure
CBR	Clinical benefit rate
CI	Confidence interval
CL	Clearance
CCI [redacted]	
CR	Complete response
CTCAE	Common Terminology Criteria for Adverse Events
CT	Computed tomography
CCI [redacted]	
DNA	Deoxyribonucleic acid
DOR	Duration of response
ECG	Electrocardiogram
ECOG (PS)	Eastern Cooperative Oncology Group (Performance Status)
eCRF	Electronic case report form
eGFR	Estimated glomerular filtration rate
CCI [redacted]	
EOI	End of infusion
EORTC IL52	European Organisation for Research and Treatment of Cancer - Disease Symptoms domain of EORTC-QLQ-MY20
EORTC QLQ-C30	European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire 30-item Core Module
EORTC QLQ-MY20	European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire 20-item Multiple Myeloma Module
EOT	End of treatment
FLC	Free light chain
GSK	GlaxoSmithKline
HR	Hazard ratio
HRQoL	Health-related quality of life
IA	Interim Analysis
ICF	Informed consent form
IMiD	Immunomodulatory imide drug
IMWG	International Myeloma Working Group

Abbreviation	Description
IRC	Independent Review Committee
ITT	Intent-To-Treat
kg	Kilogram
KVA	Keratopathy Visual Acuity
λz	Terminal phase elimination rate constant
L	Liter
LVEF	Left ventricular ejection fraction
mAb	Monoclonal antibody
MedDRA	Medical Dictionary for Regulatory Activities
mg	Milligram
µg	Microgram
min	Minute
mm	Millimeter
MM	Multiple myeloma
MMAF	Microtubular inhibitor monomethyl auristatin-F
MMRM	Mixed Model Repeated Measures
MR	Minimal response
MRD	Minimal residual disease
NCI	National Cancer Institute
NCI-CTCAE	National Cancer Institute- Common Toxicity Criteria for Adverse Events
NGS	Next generation sequencing
ORR	Overall response rate
OPS	Output and Programming Specification
OS	Overall survival
PD	Progressive disease
PET	Positron emission tomography
PFS	Progression-free survival
CCI	[redacted]
PK	Pharmacokinetic(s)
pom/dex	Pomalidomide plus Low-Dose Dexamethasone
PopPK	Population PK
PRO-CTCAE	Patient-Reported Outcome Version of the Common Toxicity Criteria for Adverse Events
PDMP	Protocol Deviation Management Plan
PR	Partial response
Q3W	Every three weeks
QTc	Corrected QT interval
QTcF	Frederica's QT Interval Corrected for Heart Rate
RRMM	Relapsed/refractory multiple myeloma
SAE	Serious adverse event
SAC	Statistical Analysis Complete
SAP	Statistical Analysis Plan
SAS	Statistical Analysis System
CCI	[redacted]

Abbreviation	Description
sCR	Stringent complete response
SD	Stable disease
$t_{1/2}$	Terminal phase half-life
CCI [redacted]	
CCI [redacted]	
TTR	Time to response
CCI [redacted]	
ULN	Upper limit of normal
V	Volume of distribution
VGPR	Very good partial response

Trademarks

Trademarks of the GlaxoSmithKline Group of Companies
belantamab mafodotin-blmf
BLENREP

Trademarks not owned by the GlaxoSmithKline Group of Companies
NONMEM
SAS
WinNonlin

7. REFERENCES

Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst. 1993 Mar 3;85(5):365-376.

Basch E, Reeve BB, Mitchell SA, Clauser SB, Minasian LM, Dueck AC, et al. Development of the National Cancer Institute patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTCAE). J Natl Cancer Inst. 2014; 106:1-11.

Bretz F, Maurer W, Brannath W, Posch M. A graphical approach to sequentially rejective multiple test procedures. Statistics in medicine. 2009;28(4):586-604.

Brookmeyer R, Crowley J. A confidence interval for the median survival time. Biometrics. 1982 Mar 1:29-41.

Chen, TT. Predicting analysis times in randomized clinical trials with cancer immunotherapy. BMC Med Res Methodol 16, 12 (2016).

Cocks K, Cohen D, Wisløff F, Sezer O, Lee S, Hippe E, et al. An international field study of the reliability and validity of a disease-specific questionnaire module (the QLQ-MY20) in assessing the quality of life of patients with multiple myeloma. European Journal of Cancer. 2007;43(11):1670-8.

Dougherty BE, Nichols JJ, Nichols KK. Rasch analysis of the Ocular Surface Disease Index (OSDI). Invest Ophthalmol Vis Sci. 2011 ;52(12): 8630-8635.

Eliason L, Loubert A, Gorsh B , Meunier J , Kleinman D , Sapra S , Regnault A. Methodology for Evaluating Meaningful Change Thresholds in Patient-Reported Outcome (PRO) Measures for Corneal Events in Patients with Relapsed/Refractory Multiple Myeloma (RRMM) Receiving Single-Agent Belantamab Mafodotin (BELAMAF). International Society for Pharmacoeconomics and Outcomes Research. European Congress. Nov. 16-19, 2020.

Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A on behalf of the EORTC Quality of Life Group. EORTC QLQ-C30 Scoring Manual (3rd edition). Brussels: EORTC, 2001.

GlaxoSmithKline Document Number 2017N336101_00 Study ID 207495. DREAMM-3: A Phase III, Open-Label, Randomized Study to Evaluate the Efficacy and Safety of Single Agent Belantamab Mafodotin Compared to Pomalidomide plus Low-dose Dexamethasone (pom/dex) in Participants with Relapsed/Refractory Multiple Myeloma (RRMM) (DREAMM 3). Report Date 30-SEP-2019.

GlaxoSmithKline Document Number TMF-13954921 Study ID 207495. A Phase III, Open-Label, Randomized Study to Evaluate the Efficacy and Safety of Single Agent Belantamab Mafodotin Compared to Pomalidomide plus Low-dose Dexamethasone

(pom/dex) in Participants with Relapsed/Refractory Multiple Myeloma (RRMM) (DREAMM 3). Report Date 23-OCT-2020.

Guy W (ed). ECDEU Assessment Manual for Psychopharmacology. Rockville, MD: US Department of Health, Education, and Welfare Public Health Service Alcohol, Drug Abuse, and Mental Health Administration, 1976

Hwang IK, Shih WJ and De Cani JS. Group sequential designs using a family of type 1 error probability spending functions. Statistics in medicine. 1990; 9(12): 1439-1445.

Kenward MG, Roger JH. An improved approximation to the precision of fixed effects from restricted maximum likelihood. Computational Statistics & Data Analysis. 2009 May 15;53(7):2583-95.

Klein J, Moeschberger M. Survival analysis: techniques for censored and truncated data. Springer Science and Business Media. 2005

Kumar S, Paiva B, Anderson KC, Durie B, Landgren O, Moreau P, et al. International Myeloma Working Group consensus criteria for response and minimal residual disease assessment in multiple myeloma. The Lancet Oncology. 2016;17: e328-46

Lan KKG, DeMets DL. Discrete sequential boundaries for clinical trials. Biometrika. 1983; 70: 649-53.

Li X, Wulfsohn MS, Koch GG. Considerations on testing secondary endpoints in group sequential design. Statistics in Biopharmaceutical Research. 2017 Oct 2;9(4):333-7.

National Cancer Institute (NCI) CTCAE Version 5. Available at: https://ctep.cancer.gov/protocolDevelopment/electronic_applications/docs/CTCAE_v5_Quick_Reference_5x7.pdf.

NCCN (National Comprehensive Cancer Network) Multiple myeloma guidelines. 2018. https://www.nccn.org/professionals/physician_gls/pdf/myeloma.pdf (Accessed 3 Sept 2018).

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 Jan;16(1):139-44

Proskorovsky I, Lewis P, Williams CD, et al. Mapping EORTC QLQ-C30 and QLQ-MY20 to EQ-5D in patients with multiple myeloma. Health Qual Life Outcomes. 2014 Mar 11;12:35.

Rajkumar SV, Dimopoulos, M.A., Palumbo, A., Blade, J., Merlini, G., Mateos, M.V., et al. International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. Lancet. 2014; 15:e538-48.

Richardson PG, Barlogie B, Berenson J, Singhal S, Jagannath S, Irwin D, et al. A phase 2 study of bortezomib in relapsed, refractory myeloma. New England Journal of Medicine. 2003;348(26):2609-17.

San Miguel J, Weisel K, Moreau P, Lacy M, Song K, Delforge M, et al. Pomalidomide plus low-dose dexamethasone versus high-dose dexamethasone alone for participants tients with relapsed and refractory multiple myeloma (MM-003): a randomised, open-label, phase 3 trial. The lancet oncology. 2013;14(11):1055-66.

Schiffman RM, Christianson MD, Jacobsen G, Hirsch JD, Reis BL. Reliability and validity of the Ocular Surface Disease Index. Arch Ophthalmol. 2000; 118:615-21.

Siegel RL, Miller KD, Jemal A. "Cancer statistics, 2019." CA Cancer J Clin 2019; 69:7-34.

Uno H, Tian L, Claggett B, Wei LJ. A versatile test for equality of two survival functions based on weighted differences of Kaplan–Meier curves. Statistics in medicine. 2015; 34(28):3680-95.

Watkins C, Huang X, Latimer N, Tang Y and Wright EJ. Adjusting overall survival for treatment switches: commonly used methods and practical application. Pharmaceutical Statistics, s2013; 12(6): 348-57