Integrated Analysis Plan

Clinical Study Protocol Identification No.	MS201430-0001
Title	Phase I, First-in-Human, Open-Label, Multiple Ascending Dose Study to Investigate the Safety, Tolerability, Pharmacokinetics, Pharmacodynamics and Clinical Activity of M6223 (an Inhibitor of TIGIT) as Single Agent and in Combination with Bintrafusp alfa (anti-PDL1/TGFß Trap in Participants with Metastatic or Locally Advanced Solid Unresectable Tumors
Study Phase	I
Investigational Medicinal Product(s)	M6223
Clinical Study Protocol Version	1 December 2021/Version 3.0
Integrated Analysis Plan Author	**Coordinating Author** Biostatistics, Merck PPD **Function** — **Author(s) / Data Analyst(s)** Quantitative Pharmacology, Merck — PPD
Integrated Analysis Plan Date and Version	19 July 2023 / Version 4.0
Integrated Analysis Plan Reviewers	**Function** — **Name** PPD PPD Biostatistics, Merck — PPD QP, Merck — PPD Programming, Merck — PPD Medical lead, Merck — PPD Global Patient Safety, Merck — PPD Clinical Biomarker representative, Merck — PPD

Approval Page

Integrated Analysis Plan: MS201430-0001

Phase I, First-in-Human, Open-Label, Multiple Ascending Dose Study to Investigate the Safety, Tolerability, Pharmacokinetics, Pharmacodynamics and Clinical Activity of M6223 (an Inhibitor of TIGIT) as Single Agent and in Combination with Bintrafusp alfa (anti-PD-L1/TGFß Trap in Participants with Metastatic or Locally Advanced Solid Unresectable Tumors

Approval of the IAP by all Merck Data Analysis Responsibles has to be documented within Veeva Vault via eSignature. With the approval, the Merck responsible for each of the analysis also takes responsibility that all reviewers' comments are addressed adequately.

By using eSignature, all signatures will appear at the end of the document.

1 Table of Contents

Approval Page 2
1 Table of Contents.....3
2 List of Abbreviations and Definition of Terms 6
3 Modification History 8
4 Purpose of the Integrated Analysis Plan.....9
5 Objectives and Endpoints 9
6 Overview of Planned Analyses.....11
6.1 Analyses for SMC meetings 12
6.2 Interim Analyses 12
6.3 Final analysis 12
7 Changes to the Planned Analyses in the Clinical Study Protocol 12
8 Analysis Sets and subgroups 13
8.1 Definition of Analysis Sets.....13
8.2 Subgroup definition and parameterization.....15
9 General Specifications for Data Analyses 15
9.1 Definition of baseline and change from baseline 18
9.2 Study day / study treatment day.....19
9.3 Definition of duration and 'time since' variables 19
9.4 Conversion factors 19
9.5 Definition of on-treatment period.....19
9.6 Imputation of missing data 19
10 Study Participants 22
10.1 Disposition of Participants and Discontinuations.....22
10.2 Protocol Deviations / Exclusion from Analysis Sets.....23
10.2.1 Important Protocol Deviations.....23
10.2.2 Reasons Leading to the Exclusion from an Analysis Set 23
11 Demographics and Other Baseline Characteristics.....23
11.1 Demographics 23
11.2 Medical History 24

11.3 Other Baseline Characteristics....25
11.4 Prior Anti-cancer Therapy....25
12 Previous or Concomitant Medications/Procedures....26
13 Study treatment: Compliance and Exposure....27
14 Efficacy Analyses....28
14.1 Overall Response....29
14.2 Disease Control....30
14.3 Figures....30
14.4 Progression Free Survival....31
14.5 Overall survival....32
14.6 Tumor growth dynamics....33
15 Safety Analyses....35
15.1 DLT analysis....35
15.2 Adverse Events....37
15.2.1 All Adverse Events....38
15.2.2 Adverse Events Leading to Discontinuation of Study Treatment....40
15.3 Deaths, Other Serious Adverse Events, and Other Significant Adverse Events....40
15.3.1 Deaths....40
15.3.2 Serious Adverse Events....41
15.3.3 Other Significant Adverse Events....41
15.4 Clinical Laboratory Evaluation....42
15.5 Vital Signs....47
15.6 Other Safety or Tolerability Evaluations....48
16 Analyses of Other Endpoints....50
16.1 Pharmacokinetics....50
16.1.1 Criteria for Handling Concentrations Below the Limit of Quantification in Pharmacokinetic Analysis....52
16.1.2 Criteria for calculation of the apparent terminal elimination rate constant and half-life....52
16.1.3 Calculation of Area Under the Concentration-time Curve....53

16.1.4 Anomalous Values....54
16.1.5 Dose Proportionality Assessment....54
16.1.6 Presentation of Pharmacokinetic Data....55
16.1.6.1 Descriptive Statistics of PK Concentration Data....55
16.1.6.2 Descriptive Statistics of PK Parameter Data....56
16.1.6.4 Graphical Summaries and Individual Plots (PK Analysis Set)....58
16.2 Pharmacodynamics....59
16.2.1 TIGIT target occupancy profile:....59
16.2.2 Immunophenotyping in blood....60
16.2.3 Immunophenotyping in tumor tissue....62
16.3 Immunogenicity (ADA)....63
CCI
17 References....66
18 Appendices....67
18.1 Appendix 1 – IAP for SMC....67

2 List of Abbreviations and Definition of Terms

ADA	Anti-drug antibody
ADaM	Analysis Data Model
AE	Adverse Event
AESI	Adverse Events of Special Interest
ALT	Alanine aminotransferase
aPTT	Activated partial thromboplastin time
AST	Aspartate aminotransferase
BOR	Best overall response
CDISC	Clinical Data Interchange Standards Consortium
CI	Confidence Interval
CR	Complete Response
(e)CRF	(electronic) Case Report Form
cSCC	Cutaneous squamous cell carcinomas
CSR	Clinical Study Report
CT	Computed tomography
DC	Disease control
DLT	Dose Limiting Toxicity
DR	Duration of response
ECG	Electrocardiogram
ECOG	Eastern Cooperative of Oncology Group
EoT	End of Treatment

GBS	Global Biostatistics
IAP	Integrated Analysis Plan
IB	Investigator's Brochure
ICH	International Council for Harmonisation
INR	International normalized ratio
irAE	Immune-related adverse event
IRR	Infusion-related reactions
IV	Intravenous
MedDRA	Medical Dictionary for Regulatory Activities
MTD	Maximum Tolerated Dose
NA	Not Applicable
NCI-CTCAE	National Cancer Institute – Common Terminology Criteria for Adverse Events
ORR	Objective Response Rate
Pd	Pharmacodynamics
PD	Progressive Disease
PD-1	Programmed death 1
PD-L1	Programmed death ligand 1
PFS	Progression Free Survival
PK	Pharmacokinetics
PP	Per Protocol
PR	Partial Response
PT	Preferred Term
PT	Prothrombin time

Q2W	Every 2 weeks
Q3W	Every 3 weeks
QTcF	QT interval corrected by Fridericia formula
RBC	Red blood cell
RDE	Recommended Dose for Expansion
SAE	Serious Adverse Event
SCR	Screening analysis population
SD	Stable Disease or Standard Deviation
SDTM	Study Data Tabulation Model
SMC	Safety Monitoring Committee
SOC	System Organ Class
TEAE	Treatment-Emergent Adverse Event
TLF	Tables, Listings, and Figures
ULN	Upper Limit of Normal
WHO-DD	World Health Organization Drug Dictionary

3 Modification History

Unique Identifier for Version	Date of IAP Version	Author	Changes from the Previous Version
1.0	25 March 2021		Initial version
2.0	27 September 2022	PPD	A lot of adaptations to new developments in study as well as clarifications to existing text.

Unique Identifier for Version	Date of IAP Version	Author	Changes from the Previous Version
3.0	05 May 2023	PPD	Specification of analyses for immunophenotyping data from blood and tumor tissue added, correction of dose intensity definition, decrease of TLFs for interim end of mono dose escalation analysis (part 1A) as time between cut-off for interim of end of part 1A and end of study will be short and interim outputs not urgently needed, clarification of several items
4.0	19 July 2023	PPD	CCI [redacted]

4 Purpose of the Integrated Analysis Plan

The purpose of this Integrated Analysis Plan (IAP) is to document technical and detailed specifications for the final analysis and other analyses of protocol MS201430-0001.

The current version specifies the analyses that will be performed at the end of monotherapy dose escalation period (part 1A) and at the end of combination dose escalation period (part 1B) and final analysis after end of study.

Results of the analyses described in this IAP will be included in the Clinical Study Report (CSR). Additionally, the planned analyses identified in this IAP may be included in regulatory submissions or future manuscripts. Any post-hoc, or unplanned analyses performed to provide results for inclusion in the CSR but not identified in this prospective IAP will be clearly identified in the CSR.

The IAP is based upon section 9 (Statistical considerations) of the Protocol version 3.0 dated 1 December 2021 and is prepared in compliance with ICH E9. It describes analyses planned in the protocol. Details of the Safety Monitoring Committee (SMC) analyses for regular review of the participants' safety are provided in appendix.

5 Objectives and Endpoints

Objectives	Endpoints (Outcome Measures)	IAP section
Primary		
• To determine safety and tolerability and (if observed) the maximum tolerated dose (MTD) of M6223 as a single agent (Part 1A) for both Q2w and Q3W regimen and of M6223 combined with bintrafusp alfa (Part 1B) for the Q2W regimen in participants with advanced solid tumors who are refractory to or have progressed under standard treatment and have no other treatment options known to confer clinical benefit.	• Occurrence of DLTs during the first 4 weeks (Day 1 to Day 28) of study intervention. • Occurrence of treatment-emergent adverse events (TEAEs) and treatment-related adverse events. • Incidence, nature, and severity of TEAEs and deaths, including cause of death, from Screening up to the 30 days Safety Follow-up Visit. • Changes (from start of treatment until 30 days after end of treatment) in clinical laboratory measures from baseline, safety ECGs measures, vital signs, ECOG performance status.	Section 15 Safety Analyses
• To determine the recommended doses of M6223 monotherapy (Part 1A) for both the Q2W and Q3W regimens and of M6223 combined with bintrafusp alfa (Part 1B) for the Q2W regimen for further exploratory clinical development. Primary endpoints (see above) and available data for secondary and exploratory endpoints (see below).	Primary endpoints (see above) and available data for secondary and exploratory endpoints (see below)	Section 15 Safety Analyses Section 16.1 Pharmacokinetics Section 15.6 Other Safety or Tolerability Evaluations Section 16.2 Pharmacodynamics
Secondary		
• To characterize the PK profile of M6223 when given alone (Part 1A) or when combined with bintrafusp alfa (Part 1B).	• PK parameters of M6223 in terms of AUC0-t, AUC0-∞, AUCτ, Cmax, Ctrough, tmax, $t_{1/2}$ and λ_z (Part 1A and Part 1B) on Day 1 of Treatment of Cycle 1, 2 and 4 from time zero to 14 (Q2W) or 21 (Q3W) days postdose.	Section 16.1 Pharmacokinetics
• To characterize the PK of bintrafusp alfa in combination with M6223 (Part 1B) by sparse sampling.	• PK parameters of bintrafusp alfa in terms of Cmax and Ctrough on Day 1 of Treatment of Cycle 1 and 4 from time zero to 14 days postdose.	Section 16.1 Pharmacokinetics

Objectives	Endpoints (Outcome Measures)	IAP section
• To characterize immunogenicity of M6223 when given alone and immunogenicity of M6223 and bintrafusp alfa when given in combination and its relationship to drug exposure.	• Immunogenicity of M6223 (Part 1A and Part 1B) and bintrafusp alfa (Part 1B), as measured by anti-drug-antibody (ADA) assays, from pre-dose of Cycle 1 Day 1 through 30 days safety follow-up Visit.	Section 16.3 Immunogenicity
• To characterize the relationship between exposure and QT interval of M6223 alone and in combination with bintrafusp alfa.	• QT interval from Cycle 1 to Cycle 7.	Section 15.6 Other Safety or Tolerability Evaluations
• To characterize preliminary clinical activity parameters of M6223 in monotherapy (Part 1A) and in combination with bintrafusp alfa (Part 1B) using RECIST 1.1.	• OR, DR, time to response, disease control (DC), progression free survival (PFS) using RECIST 1.1, per Investigator, and overall survival (OS).	Section 14 Efficacy Analyses
Tertiary/Exploratory		
• To assess the peripheral TIGIT target occupancy with M6223 alone (Part 1A) and combined with bintrafusp alfa (Part 1B) and exposure/ target occupancy	• Target occupancy at Cycle 1, 2 and 4.	Section 16.2 Pharmacodynamics

CCI

[redacted]

6 Overview of Planned Analyses

The following analyses are planned for this trial:

- SMC analyses (See IAP for SMC in Appendix 18.1)
- End of monotherapy dose escalation analysis (part 1A)
- End of combination dose escalation analysis (part 1B)
- Final analysis

Statistical analyses (except for SMCs) will be performed on the basis of CDISC SDTM data. These SDTM data contain as clean as possible eCRF data as well as external data for e.g. central laboratory and DLT decisions from SMC.

A data review meeting will be held prior to any database lock. In addition, no database can be locked until this IAP has been approved (except for emergency).

6.1 Analyses for SMC meetings

Details of analyses for SMC meetings are specified in a separate IAP, provided in Appendix 18.1.

6.2 Interim Analyses

The cut-off for an exploratory interim analysis of the safety, available PK, available Pd and preliminary antitumor activity data from each study part (end of monotherapy dose escalation part 1A, end of combination dose escalation part 1B) will be triggered when all participants enrolled in the respective study part either:

- Have reached the first on-study intervention tumor assessment (including confirmatory scan), or
- Have died, or
- Have been withdrawn from the study intervention or from the study for any reason.

  For the interim analysis 1B, no PK outputs will be produced as only preliminary PK data would be available and additionally to avoid delays in analysis.

6.3 Final analysis

The final analysis is performed after the study is completed.

7 Changes to the Planned Analyses in the Clinical Study Protocol

Definition of TEAE differs from what is mentioned in protocol. The protocol mentions: “from Screening up to the 30 days Safety Follow-up Visit”. The definition used is “from start of treatment until (including) 30 days after the end of treatment.”

Definition of DLT analysis set adapted to reflect that final decision on DLT analysis set is with the SMC.

Definition of Pk, PD and ECG analysis set is changed to include all participants with baseline and at least one post-dose measurement, i.e.no exclusion due to clinically important protocol deviations or other important effects. If participants had important events that may affect the PK, PD or ECG sensitivity analyses may be done additionally excluding such participants.

As analyses for duration of response and time to response can only be performed in patients with response and there are no responders in this trial, no analyses for duration of response and time to response will be performed.

As cut-off for end of monotherapy dose escalation occurred close to the end of the study, only a limited set of TLFs will be produced for end of mono dose escalation (1A) interim.

8 Analysis Sets and subgroups

8.1 Definition of Analysis Sets

The analysis sets are specified below. The final decision to exclude participants from any analysis set will be made during a data review meeting prior to database lock except for the DLT analysis set. For the DLT analysis sets, decisions for DLTs are taken by the SMC.

Screening Analysis Set (SCR)

The Screening analysis set includes all participants who signed the informed consent.

Safety Analysis Set (SAF)

The Safety analysis set will include all participants who received at least one dose of study intervention. Analyses will consider participants as treated meaning that participants will be classified according to the study intervention actually received (regarding dose level, see section 9 General Specifications for Data Analyses including explanation on how to deal with participants that did receive a dose different to the planned dose)

Dose Limiting Toxicity Analysis Set (DLT)

The Dose Limiting Toxicity analysis set (DLT) will include all participants who received at least one dose of study intervention and meet at least one of the following criteria:

- Experienced at least one DLT during the DLT period, regardless of the administered number of doses of study treatment/completion of the DLT period.
- Received at least 80% of the planned cumulative dose of each treatment during the DLT period and completed the DLT period.
- Additionally, participants that did not receive 80% of the planned total dose of study treatment during the DLT period, but at least 80% dosing of a different (lower) dose cohort during the completed DLT period are eligible for the DLT analysis set to be analyzed in the highest dose cohort for which they received 80% of dosing.

In case the received cumulative dose is higher than any dose level that was to be administered per SMC decision (and all other criteria for inclusion in DLT analysis set are fulfilled) the participant will be analyzed in the highest dose level that was recommended to be administered by the SMC.

The above items are only guidance. The final decision on DLT analysis set is with SMC. Any participant who has been declared to have had a DLT or to be DLT evaluable by SMC is part of the DLT analysis set

Pharmacokinetics Analysis Set (PK)

Pharmacokinetics Analysis Set will include all participants, who receive at least one dose of study intervention and provide at least one measurable post-dose concentration. Participants will be analyzed per the actual study intervention they received.

All PK analyses will be based on this analysis set.

Pharmacodynamic Analysis Set (Pd)

Pharmacodynamic analysis set will include all participants, who receive at least one dose of study intervention and provide at least one measurable Pd endpoint post-dose in addition to a pre-dose measurement.

Participants will be analyzed per the actual study intervention they received. All Pd analyses will be based on this analysis set.

Immunogenicity Analysis Set (IM)

Immunogenicity analysis set will include all participants who receive at least one dose of study intervention and have at least one valid antidrug antibody (ADA) result. All ADA analyses will be based on this analysis set.

Electrocardiogram Analysis Set (ECG)

Electrocardiogram analysis set will include all participants, who receive at least one dose of study intervention and provide at least one measurable post-dose ECG endpoint.

The following table summarizes the use of the analysis sets in the different analyses.

	Analysis Set					
Analyses	**SAF**	**DLT**	**PK**	**Pd**	**IM**	**ECG**
DLT		✓				
Baseline Characteristics	✓					
Previous and Concomitant Therapies	✓					
Compliance and Exposure	✓					
Efficacy	✓					
Safety and Tolerability	✓					
Pharmacokinetics			✓			
Pharmacodynamic biomarkers				✓		
Immunogenicity					✓	
Electrocardiogram						✓

Note:

For details regarding the dose level a subject is analyzed in (in conjunction with the analysis) see section 9 General Specifications for Data Analyses including.

8.2 Subgroup definition and parameterization

As is the nature of dose escalation studies, analyses will be performed by dose (and regimen (e.g. Q2W or Q3W) if different regimens are tested as well as summarized by study part (monotherapy/combination).

9 General Specifications for Data Analyses

This section describes any general specifications not included in subsequent sections.

Study intervention is defined and labelled as "M6223" in Part 1A and "M6223 + bintrafusp alfa", in Part 1B.

All analyses will be separate for each part (dose escalation in monotherapy (1A) and dose escalation in combination (1B)), and so will decisions by the SMC. Tables for final analysis will be separate for each part, unless otherwise specified.

Analyses will be displayed separately by regimen (if different regimens are administered, e.g. Q2W and Q3W) and dose.

The "start date" for this study is the date of first study drug administration (in Part 1A M6223 and in Part 1B M6223 + bintrafusp alfa).

Dose level

General rule for actual dose level to consider participant in, if actual dose differs from planned dose. This is determined during the DLT period:

<table>
<tr><td colspan="2">Subject had a DLT
Use planned dose as actual dose, unless actual cumulative dose received ≥ 80% of any higher prespecified dose level, then use this higher prespecified dose level as actual**</td></tr>
<tr><td colspan="2">Subject has been classified as not evaluable for DLT by SMC:

DLT analysis: Not evaluable
Safety analysis: Use lowest single dose received actual dose
Efficacy analysis: Use planned dose</td></tr>
<tr><td colspan="2"></td></tr>
<tr><td colspan="2"></td></tr>
<tr><td>Actual cumulative dose received ≥ 80% of cumulative planned dose and has no DLT during the DLT period and has not been classified as not evaluable for DLT by SMC</td><td>Actual cumulative dose received < 80% of cumulative planned dose and has no DLT during the DLT period and has not been classified as not evaluable for DLT by SMC</td></tr>
</table>

Actual cumulative dose received ≤ Cumulative Planned dose : Use planned dose as actual dose	**Actual cumulative dose received ≥ 80% of a lower dose level (cumulative)** DLT/safety analysis, PK: Use this lower dose as actual Pd and Efficacy analysis: Use planned dose
Actual cumulative dose received ≥ cumulative planned dose: Use planned dose level unless actual dose received ≥ 80% of higher planned dose level, then use this higher planned dose level as actual**	**Actual cumulative actual dose received ≤ 80% of any tested dose (cumulative)** DLT analysis: Not evaluable Safety analysis, PK: Use lowest single dose received as actual dose Pd and Efficacy analysis: Use planned dose

* subjects with DLT may be flagged in outputs showing individual data with footnotes showing their actual cumulative dose or end of treatment day.
**In case a subject received >120% cumulative dose of the dose level he/she is assigned to as actual dose level the data of this subject will not be considered in Pd and efficacy summary outputs. The data will be displayed in listings and line plots but with a footnote stating the actual cumulative dose.

Disposition, compliance and exposure will be described by planned dose and actual dose if there are participants that are assigned differently. For more details see in the corresponding sections.

Significance level:

There is no formal significance level for this study and all analyses are considered descriptive.

There will be no statistical tests performed. If confidence or credibility intervals are to be calculated, these will be two-sided with a confidence probability of 95%, unless otherwise specified in this IAP.

Presentation of continuous and qualitative variables:

Continuous variables will be summarized using descriptive statistics, i.e.

- number of participants, number of participants with non-missing values
- mean, standard deviation

- median, 25th Percentile - 75th Percentile (Q1-Q3)
- minimum, maximum

If there are fewer than 5 observations summarized, only the number of subjects (N), number of subjects with non-missing values, the mean, and the min and max will be given.

Qualitative variables will be summarized by counts and percentages.

Unless otherwise stated the calculation of proportions will be based on the number of subjects in the analysis set of interest. Therefore, counts of missing observations will be included in the denominator and presented as a separate category.

Deviations from this definition might apply to the PK analysis (see section 16.1 for definitions)

Kaplan-Meier estimates (product-limit estimates) will be presented for the analysis of OR, BOR, DR, PFS and OS, together with a summary of associated statistics (median survival time, 6-, 12-month survival rate estimates and estimates for every 6 months thereafter if applicable) including the corresponding two-sided 95% Cis if sample size is sufficient.

Descriptive statistics will be computed by treatment day (C1D1, C1D8, C2D1, …, EOT). A windowing approach will be used. Unless differently stated below the windows will be the visit windows allowed per protocol. Unscheduled visits will be included in the derivation of baseline (if occurred within the allowed screening period) or worst on-treatment values and in any analysis using treatment day (e.g. plots of lab data).

Statistical software:

All analyses will be performed using SAS® Software version 9.2 or higher, and otherwise as specified (e.g. for Bayes modeling)

9.1 Definition of baseline and change from baseline

Definition of baseline:

In general, the last non-missing measurement prior to the first study treatment administration will be used as the baseline measurement.

If an assessment that is planned to be performed before treatment per protocol is performed on the same day as the start of treatment, respectively, but the assessment time is not available, it will be assumed that it was performed prior and will be considered as baseline.

If a scheduled pre-dose measurement actually occurred post-dose, then the corresponding measurement will be analyzed similar to an unscheduled post-dose measurement.

Definition of change from baseline

Absolute change from baseline = visit value – baseline value

Percent Change from Baseline = 100 * (visit value – baseline value) / baseline value

9.2 Study day / study treatment day

Day 1 is the day of start of study treatment administration (in part 1A M6223 and in part 1B M6223 or bintrafusp-alpha) , the day before is Day -1 (no Day 0 is defined). Study treatment day is defined relative to Day 1.

9.3 Definition of duration and 'time since' variables

If not otherwise specified, duration will be calculated by the difference of start and stop date + 1 (e.g. survival time (days) = date of death – date of first study drug administration + 1).

The time since an event (e.g. time since first diagnosis) will be calculated as reference date minus date of event.

9.4 Conversion factors

The following conversion factors will be used to convert days into months or years: 1 month = 30.4375 days, 1 year = 365.25 days.

9.5 Definition of on-treatment period

The on-treatment period is defined as the time from the first study intervention to the last study intervention date + 30 days. For participants with treatment ongoing at cut-off date, all data from the first study intervention up to the cut-off date will be considered under the on-treatment period.

9.6 Imputation of missing data

Unless otherwise specified all data will be evaluated as observed, and no imputation method for missing values will be used.

In all participant data listings, imputed values will be presented and imputed information will be flagged.

Missing statistics, e.g. when they cannot be calculated, should be presented as "nd". For example, if n=1, the measure of variability (SD) cannot be computed and should be presented as "nd".

Where tables are presented over different time points, the total of missing and non-missing observations at each time-point should reflect the population still in the trial at that time. This does not apply when imputations are made beyond trial withdrawal. For example, if a subject is still in the trial at the time-point but with missing data, they should be counted in the number of missing observations.

The following table for imputation rules will be considered:

Adverse events	Incomplete AE-related dates will be imputed as follows: • If the AE onset date is missing completely, then the onset date will be replaced by the start of study treatment. • If only the day part of the AE onset date is missing, but the month and year are equal to the start of study treatment, then the AE onset date will be replaced by the start of study treatment. For example, if the AE onset date is --/JAN/2015, and study treatment start date is 15/JAN/2015, then the imputed AE onset date will be 15/JAN/2015. If the end date or resolution date indicates that the AE has stopped before start of treatment, this date will be used for imputation instead of start of treatment date. • If both the day and month of the AE onset date are missing but the onset year is equal to the start of study treatment, then the onset date will be replaced by the start of study treatment. For example, if AE onset date is --/---/2014, and study treatment start date is 19/NOV/2014, then the imputed AE onset date will be 19/NOV/2014. • In all other cases the missing onset day or missing onset month will be replaced by 1. • Incomplete stop date will be replaced by the last day of the month (if day is missing only), if not resulting in a date later than the date of participant's death. In the latter case the date of death will be used to impute the incomplete stop date. • In all other cases the incomplete stop date will not be imputed. If stop date of AE is after date of cut-off this date will be kept.
Previous and concomitant medication	For identification of previous or concomitant medications/procedures, no formal imputation will be performed on missing or incomplete dates. Rules presented in Table 1 will be used to define if a medication/procedure is considered as a previous, concomitant or both previous and concomitant medication/procedure.

Table 1 Imputation rules for medication/procedure end dates

End date of medication/procedure			imputation rule
Day	**Month**	**Year**	
UNK	UNK	UNK	After treatment start (ongoing)
UNK	UNK	< Treatment start (year)	Before treatment start
UNK	UNK	>= Treatment start (year)	After treatment start
UNK	< Treatment start (month and year)		Before treatment start
UNK	>= Treatment start (month and year)		After treatment start
< Treatment start (complete date)			Before treatment start
>= Treatment start (complete date)			After treatment start

UNK = Unknown

Table 2 Rules to define previous and/or concomitant medications

Start date of medication/procedure			imputation rule (see Table 1)	Medication/procedure
Day	Month	Year		
UNK	UNK	UNK	Before treatment start	Previous
UNK	UNK	UNK	After treatment start	Previous and concomitant
UNK	UNK	<= Treatment start (year)	Before treatment start	Previous
UNK	UNK	<= Treatment start (year)	After treatment start	Previous and concomitant
UNK	UNK	> Treatment start (year) and <= Treatment end + 30 days (year)	After treatment start	Concomitant
UNK	<= Treatment start (month and year)		Before treatment start	Previous
UNK	<= Treatment start (month and year)		After treatment start	Previous and concomitant
UNK	> Treatment start (month and year) and <= Treatment end + 30 days (month and year)		After treatment start	Concomitant
<= Treatment start (date)			Before treatment start	Previous
<= Treatment start (date)			After treatment start	Previous and concomitant
> Treatment start (date) and <= Treatment end + 30 days (date)			After treatment start	Concomitant

UNK = Unknown

Death date	For the purpose of survival analyses partially missing death dates will be imputed as follows:

	If only the day is missing, the death date will be imputed to the maximum of the (non-imputed) day after the date of last contact and the 15th day of the month. Otherwise it will not be imputed.

10 Study Participants

The subsections in this section include specifications for reporting participant disposition and study treatment/study discontinuations. Additionally, procedures for reporting protocol deviations are provided.

10.1 Disposition of Participants and Discontinuations

A table of disposition should display the following (per study part, regimen (if a change occurred, e.g. Q2W and Q3W), dose level and in total). This table is the only one that will come with a total "total" column and not be separated by part:

- Number of participants in each analysis population (SCR, SAF, DLT, PK, Pd, IM, ECG)
- Number of participants who discontinued from the study prior to treatment overall and grouped by the main reason (participants did not meet eligibility criteria, withdrew consent, other)
- Number and percentage of participants still in trial
- Number and percentage of participants withdrawn from trial
- Number of participants by reason for withdrawal

This table will be produced up to four times if there are participants with differing planned dose and actual dose, one by planned dose (including SCR, SAF, Pd, PK, IM, ECG) and one by actual dose (one for DLT, one for SAF and PK, and one for Pd, IM, ECG). This table is to be provided for all analyses.

Additionally a listing displaying the actual dose, disposition status and reason for end of trial will be provided in all analysis (as part of section 9.2).

Additionally a listing will display subjects excluded from analysis sets and reason for exclusion (also as part of section 9.2) for end of dose escalation combination 1B and final analysis.

For final analysis the number of participants in analysis sets will additionally be provided by site.

10.2 Protocol Deviations / Exclusion from Analysis Sets

10.2.1 Important Protocol Deviations

Important protocol deviations (IPDs) are protocol deviations that might significantly affect the completeness, accuracy, and/or reliability of the study data or that might significantly affect a participant's rights, safety, or well-being.

The IPDs will be identified for all participants by either site monitoring, medical review processes or programming and confirmed prior to or at the Data Review Meeting at the latest.

All important protocol deviations are documented in SDTM datasets whether identified through site monitoring, medical review or programming.

A full list of potential protocol deviations including definition and categorization is maintained by Covance in the "Study Specific Protocol Deviation List" attached to the Protocol Deviation Management Plan.

Important protocol deviations will be listed by participant for the end of dose escalation part 1B analysis and final analysis. For the final analysis additionally a frequency table for important, separated for such pre-/post inclusion deviations will be provided based on the Safety analysis set.

Referring to FDA, a listing of all participants affected by the COVID-19 related study disruption will be produced by unique subject number identifier and by investigational site, and a description of how the individual's participation was altered for the end of dose escalation part 1B analysis and final analysis.

10.2.2 Reasons Leading to the Exclusion from an Analysis Set

The reason for exclusion from the SAF/DLT/PK/Pd/IM/ECG analysis set will be included in the listing of subjects excluded from analysis sets, as specified in section 10.1.

11 Demographics and Other Baseline Characteristics

If not stated otherwise, the following analyses will be performed based on the SAF, per study part, regimen (if a change occurred), dose level and in total.

11.1 Demographics

Demographic characteristics and physical measurements will be summarized descriptively using the following information from the Screening eCRF pages.

The following demographic characteristics will be included:

- Sex: male, female
- Race: White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Not collected at this site, Other
- Ethnic origin 1: Hispanic or Latino/Not Hispanic or Latino,
- Ethnic origin 2: Japanese/Not Japanese
- Age (years)
- Age categories :< 65 years, ≥ 65 years
- Pooled Region: North America, Europe, Asia, Rest of the World
- Geographic Region: North America, Latin America, Western Europe, Eastern Europe, Middle East, Australia, Asia
- Weight (kg)
- BMI (kg/m^2) at Baseline
- Eastern Cooperative Oncology Group (ECOG) Performance status (0,1,2,3,4)

Specifications for computation:

- Age [years]: (date of given informed consent – 01JANYYYY + 1) / 365.25, where YYYY is the year of birth as collected in eCRF. The integer part of the calculated age will be used for reporting purposes.
- BMI [kg/m²] = $\frac{weight\ [kg]}{height[cm]^2} \times 10000$
- Site codes will be used for the determination of the participant's geographic region.

There will be a listing in section 9.2 displaying the demographics for each participant.

Demographics tables and listing will be done for end of combination dose escalation 1B and final analysis.

11.2 Medical History

TLFs for medical history are not produced for end of dose escalation analyses, only for final analysis. The medical history will be summarized from the "Medical History" eCRF page, using the most recent MedDRA version at time of database lock, preferred term as event category and system organ class (SOC) body term as Body System category. Each participant will be counted only once within each PT or SOC.

Medical history will be displayed in terms of frequency tables: ordered by primary SOC and PT in alphabetical order.

11.3 Other Baseline Characteristics

Information on other baseline characteristics collected at baseline will be summarized. Summary statistics will be presented for (only at final analysis, except for site and disease stage that will also be included in a table in end of combination escalation 1B):

- Childbearing Potential (not applicable for end of dose escalation analyses)
- Site of primary tumor
- Time since initial diagnosis
- Disease stage at initial diagnosis including TNM status (CRF) (summarized in T0, T1, T2, T3, T4, Tx or Tis, N0, N1, N2 and N3, NX, M0 and M1, MX)
- Disease stage at study entry including TNM status (CRF)
- TGF-Beta mediated skin lesions (not applicable for end of dose escalation analyses)

Baseline characteristics with respect to vital signs, physical examinations, and hematology/biochemistry will be part of Section 15 (Safety Evaluation).

11.4 Prior Anti-cancer Therapy

The prior anti-cancer therapies are collected under the "Prior Anti-Cancer Drug Therapies Details" and "Prior Anti-Cancer Radiotherapy Details" eCRF pages. No TLFs in this category are to be produced for end of dose escalation analyses, only for final analysis.

The number and percentage of participants in each of the following anti-cancer therapy categories will be tabulated:

- Participants with at least one prior anti-cancer therapy
- Participants with at least one prior anti-cancer drug therapy
- Participants with at least one prior anti-cancer radiotherapy

Prior anti-cancer drug therapy will be summarized as follows based on the number and percentage of participants with the following:

- Any prior anti-cancer radiotherapy regimens: missing / 1 -2 / ≥3
- Any prior anti-cancer drug therapy regimens: missing / 1 -3/ 4-6 / ≥7

The summary table will also contain summary stats of number of prior lines of therapy

The prior anti-cancer drugs will also be summarized based on the number and percentage of participants by the drug class and preferred term. A participant will be counted only once within a given drug class and within a given drug name, even if he/she received the same medication at different times. The summary will be sorted on decreasing frequency of drug class and decreasing frequency of drug name in a given drug class. In case of equal frequency regarding drug class (respectively drug name), alphabetical order will be used.

The listings of prior anti-cancer treatments and procedures will also be provided as follows. These will include the participant identification number, and all the relevant collected data-fields on the corresponding eCRF pages.

- Listing of prior anti-cancer drug therapies, additionally containing: Intent of Therapy: Neo-Adjuvant / Adjuvant/ Metastatic or Locally advanced and Best response
- Listing of prior anti-cancer radiotherapy

12 Previous or Concomitant Medications/Procedures

The following analyses will be performed based on the safety analysis set by regimen (if different regimens are administered) and by dose. No TLFs in this category are to be produced for end of dose escalation analyses, only for final analysis

Concomitant treatments are medications, other than study treatment, which are taken by participants any time during the on-treatment period, see section 9.5.

Previous medications are medications, other than study treatment and pre-medications for study treatments, which started before first administration of study treatments.

A medication may be classified as both concomitant and previous. The respective flags will be derived based on start and end date.

Concomitant and previous treatment each will be summarized by number and percentage of participants from the “Previous and Concomitant medication and/or Therapies” eCRF. ATC-2nd level and preferred term will be tabulated as given from the WHO-DD dictionary most current version. If any previous or concomitant medication is classified into multiple ATC classes, the medication will be summarized separately under each of these ATC classes. The summary tables will be sorted by decreasing frequency of drug class and decreasing frequency of drug name in a given drug class. In case of equal frequency regarding drug class (respectively drug name), alphabetical order will be used. In case any specific medication does not have an ATC classification level 2 coded term, it will be summarized under “Unavailable ATC classification” category. Each participant will only be counted once, even if he/she received the same medication at different times.

All **concomitant procedures**, which were undertaken during the on-treatment period will be listed according to the CRF page "Concomitant Procedures".

13 Study treatment: Compliance and Exposure

The following analyses will be performed based on the safety analysis set, per study part, regimen (if a change occurred), dose level and in total. The TLFs in this category are to be produced twice (by actual dose and by planned dose) if there are subjects for which planned and actual dose differs.

All dosing calculations and summaries will be based on "M6223 Administration – Infusion" and "Bintrafusp alfa Administration – Infusion" CRFs pages in particular for first study drug administration.

The date of last study drug administration will be defined as the last date in the "M6223 Administration – Infusion" and "Bintrafusp alfa Administration – Infusion" CRFs pages, with dose>0.

No imputation of missing start dates of study treatments will be done.

In case the last date of study drug is incomplete the date of last study drug administration will be taken from the End of Treatment page.

Interruption of treatment will be identified in the "M6223 Administration – Infusion" and "Bintrafusp alfa Administration – Infusion" CRFs pages, with no dose.

Each study drug (M6223 and bintrafusp alfa) will be analyzed separately.

A dose is regarded to be administered, if the actual dose received ("Dose per administration") is > 0

As this is a study with cyclic dosing,

end date of study drug = start date of last cycle with non-zero dose of study drug + intended cycle duration in days – 1.

Intended cycle duration in our study is 14 days in Q2W schedule and 21 in Q3W schedule.

For the end of combination dose escalation (1B) analyses and final analysis, the summary of treatment exposure will include the following information:

- Exposure duration (in weeks) as defined by (end date of study drug–date of first dose of study drug +1)/7

- Number of dosing days (in days) as defined by (end date of study drug–date of first dose of study drug – sum of days with dose interruption +1)
- Total number of cycles received with at least 80% of planned dose as defined sum of cycles with received dose >= 80% of planned dose
- Total number of cycles received
- Dose intensity (definition below)
- Number and percentage of participants with at least one dose reduction and number of participants with at least one dose delay ≥7 days or missed cycle will be presented.

  For the end of dose escalation analyses a listing will show the planned dose, actual dose for SAF, actual dose for DLT analysis and the doses received (including the study day) for M6223 and Bintrafusp-alpha (if applicable).

  For final analysis, this listing will provide all administrations (in section 9).

Additionally a line plot will show the dose received versus the study day (color code by planned dose) for end of combination dose escalation (1B) only. As swimmer plot is updated to include breaks, this plot is no longer needed.

The cumulative dose (mg) per participant is the sum of all doses that the participant received in a study part.

The dose intensity (mg/cycle) is defined as

Cumulative dose / (exposure duration in weeks (including the +2 for the last cycle)/2)

14 Efficacy Analyses

The following analyses will be performed based on the safety analysis set, per study part, regimen (if a change occurred) and by dose level. See section 9 for description on how to handle subjects that did not receive the planned dose.

14.1 Overall Response

Best Overall response **(BOR)** will be assessed based on reported overall responses at different evaluation time points from the study treatment start date until documented disease progression in accordance to RECIST v1.1 by investigator, taking requirements for confirmation into account as detailed below.

Only tumor assessments performed before the start of any subsequent anti-cancer therapies will be considered in the assessment of BOR. If a tumor assessment was performed on the same day as start of new anti-cancer therapy, it will be assumed that the tumor assessment was performed prior to the start of the new anti-cancer therapy, therefore the tumor assessment will be included in the assessment of BOR.

Clinical deterioration will not be considered as documented disease progression. The date of the overall response assessment is the earliest date for imaging of target, non-target and new lesions of images taken at that response assessment.

Confirmed Objective Response (OR) is defined as a confirmed BOR of complete response (CR) or partial response (PR) according to RECIST v1.1.

Participants who do not have an on-treatment radiographic tumor assessment due to early progression, who receive anti-tumor treatments other than the study treatments prior to reaching confirmed CR or PR, or who die, progress, or drop out for any reason prior to reaching confirmed CR or PR will be counted as non-responders in the assessment of OR. Each participant will have an objective response status (0: ‘no OR’; 1: ‘OR’).

OR rate (ORR) is the proportion of participants with OR in the analysis set.

The ORR per study part, regimen (if a change occurred) and by dose level will be calculated along with the two-sided 95% CI using the Clopper-Pearson method (exact CI for a binomial proportion as computed by default by the SAS FREQ procedure using the EXACT option).

In addition, the frequency (number and percentage) of participants with BOR of CR, PR, SD, PD, non-CR/non-PD (applicable only to participants with non-measurable disease at baseline), and NE will be tabulated.

Additionally the proportion of confirmed responses will be shown (CR or PR with a confirmation no sooner than 6 weeks)

- *CR = at least two determinations of CR at least 6 weeks apart (with no PD in between)*
- *PR = at least two determinations of PR or better (PR followed by PR or PR followed by CR) at least 6 weeks apart (and not qualifying for a CR), with no PD in between*

Listings of efficacy results data for section 9.2 will be presenting overall response and sum of longest diameters

All these analyses will only be done for end of combination dose escalation 1b and final analyses.

14.2 Disease Control

Disease control **(DC)** for at least 14 weeks is defined as having an assessment of SD, PR or CR at least 14 weeks (-1 day is allowed as in visit schedule) after start of treatment and not having had a PD assessment before that. DC for at least 28 weeks is defined as having an assessment of SD, PR or CR at least 28 weeks (-1 day) after start of treatment and not having had a PD assessment before that.

DC rate (DCR) is the proportion of participants with DC. DCR for at least 14 weeks and DCR for at least 28 weeks will be calculated along with the two-sided 95% CI using the Clopper-Pearson method. This will be done for combination 1B end of dose escalation and final analyses.

14.3 Figures

The following figures are to be provided at each analysis (end of escalation and final). Figures will be provided separately for mono and combination. Color code for dose levels should be the same in each figure:

- Swimmer plot. The swimmer plot will show horizontal bars for time on treatment of each participant (x axis: time in days). Symbols will denote tumor assessments, using different symbols for CR, PR, SD or PD. All lines will start at baseline (day 0) at 0. At the end of treatment a symbol will denote the reason for end of treatment. In the figure for combination cohorts, a symbol will denote the time of last administration of BA.

- Spider plot. The spider plots will show the percent change in sum of longest diameters $\frac{(post\ baseline\ sum\ of\ longest\ diameters - baseline\ sum\ of\ longest\ diameters)}{baseline\ sum\ of\ longest\ diameters}$ on y-axis and the time (in days) on treatment on x-axis. The plot will start at 0/0 for each patient. Reference lines are to be drawn at 0%, -30% and +20%.

- Waterfall plot. The waterfall plot will show the best percent change in sum of longest diameters as a bar for each participant. Different dose levels will be depicted by different fill colors for the bars (e.g. bars for 100 mg patients all filled in orange, for 900 mg all in red etc). Additionally if PD-L1 data are available, bars will have a different structure depending on PD-L1 expression. PD-L1 negative bars will be just the dotted, PD-L1 between 1% and 49% will have alternating horizontal lines (looks like hyphens) within the bars and PD-L1 >= 50% will be with diagonal stripes. Unknown PD-L1 status will be depicted by just having the colored bars without structure. Below there is a mock for the waterfall plot

14.4 Progression Free Survival

Progression Free Survival (PFS) time is defined as the time from start date of treatment to the date of the first documentation of objective progression of disease (PD) or death due to any cause, whichever occurs first. The tumor response will be determined according to RECIST 1.1 and assessed by the investigator.

PFS time (in months) = (Date of PD or death – start date + 1)/ 30.4375 (months)

Progression Free Survival data will be censored on the date of the last adequate tumor assessment for participants who do not have an event (PD or death), for participants who start a new anticancer therapy prior to an event or for participants with an event after 2 or more missing tumor assessments. Participants who do not have a baseline tumor assessment or who do not have any post-baseline tumor assessments will be censored on the treatment start date unless death occurred on or before the time of the second planned tumor assessment in which case the death will be considered an event.

The last adequate tumor assessment is defined as the last tumor assessment result that is not "NE" or "NA".

Table 4 PFS Censoring Reasons and Hierarchy

Hierarchy	Condition	Censoring Reason
1	No event and none of the conditions in the below hierarchy are met	Administrative censoring: Ongoing in the study without an event

2	No adequate baseline assessment	No baseline assessment
3	No event and no adequate post-baseline tumor assessment	No post-baseline tumor assessment
4	Start of subsequent anti-cancer therapy before event.	Start of subsequent anti-cancer therapy
5	Event more than 2*8 weeks (8 weeks is the scheduled time between tumor assessments) after last adequate post-baseline tumor assessment//date of first study treatment administration	Event after 2 or more missing assessments
6	No event and [withdrawal of consent date ≥ date of randomization] OR [End of study (EOS) = Withdrew consent]	Withdrawal of consent

PFS will be displayed graphically and analyzed using Kaplan-Meier methodology. As the number of participants in this trial is expected to be small and the number of participants with PFS event might be even smaller, the Kaplan-Meier method may not provide reliable estimates. In this case, only descriptive statistics or listings will be provided.

Kaplan-Meier estimates (product-limit estimates) will be presented by study part together with a summary of associated statistics including the median PFS time with two-sided 95% CIs. In particular, the PFS rate at 6 and 12 months and estimates for every 6 months thereafter will be estimated with corresponding two-sided 95% CIs. The CIs for the median will be calculated according to Brookmeyer and Crowley and the CIs for the survival function estimates at the time points defined above will be derived using the log-log transformation according to Kalbfleisch and Prentice (conftype=loglog default option in SAS Proc LIFETEST) with back transformation to a CI on the untransformed scale. The estimate of the standard error will be computed using Greenwood's formula.

The PFS time or censoring time and the reasons for censoring will also be presented in a participant listing.

These analyses will only be performed at the final analysis.

14.5 Overall survival

Overall survival (OS) is defined as the time from start date of treatment to the date of death due to any cause. OS for participants without death prior to cut-off will be censored at date of last contact. Overall survival in months is calculated as follows:

OS (months) = [date of death or censoring– start date +1]/30.4375

The Kaplan-Meier based analysis described in Section 14.5 for the derivation of progression-free related estimates will be repeated for the survival related estimates.

Frequency (number and percentage) of participants with an event (death) and censoring reasons will be presented by treatment group. Censoring reasons are as follows:

- Alive
- Withdrawal of consent
- Lost to follow-up

Lost to follow-up will include participants that the investigator states were lost to follow-up prior to the analysis cut-off as well as participants with a last contact date > 8 weeks prior to the analysis cut-off date.

The OS time or censoring time and the reasons for censoring will also be presented in a participant listing.

These analyses are only applicable for final analyses

CCI

CCI

15 Safety Analyses

The subsections in this section include specifications for summarizing safety endpoints that are common across clinical studies such as adverse events, laboratory tests and vital signs.

The primary endpoint during dose escalation is the occurrence of DLTs, AEs and other safety related measurements.

The analysis population for the DLT analysis will be the DLT analysis set, for all other safety analyses it will be the SAF analysis set.

If there is reference to dose then this means the actual dose for DLT for DLT analysis (section 15.1) and actual dose for safety in all other safety sections.

15.1 DLT analysis

Of Note, DLT as per SMC decision will be available in the table "Review of data" in the SMC minutes or recommendation form.

In case any additional information leads to change in judgement of the SMC regarding DLTs, it will be added to the minutes.

For end of combination dose escalation (part 1A and part 1B) and final analysis, a table will display the following (per regimen and dose level and in total) for the DLT analysis set:

- Number and percentage of participants with a DLT per investigator's judgement as recorded in eCRF (using DLT flag from the AE CRF page)

- Number and percentage of participants with a DLT per SMC decision

A listing of DLTs (per investigator and/or per SMC) will be provided by dose level and participant, including whether it was judged to be a DLT by investigator and/or by SMC) only for end of combination dose escalation (part 1B) and final analysis.

This listing will additionally contain planned dose, age, sex, site of primary tumor, end of treatment date,
SOC, PT, investigator term, start date (+treatment day), stop date (+treatment day), relatedness to M6223 (yes/no), to bintrafusp alpha (yes/no) (if applicable, in part 1B only), grade, action taken with M6223, outcome, SAE (yes/no).

A listing of participants excluded from DLT analysis set along with their reason for exclusion will also be provided as part of listings 9.2 (as already defined in section 10.2.2) only for end of combination (1B) and final analysis.

In addition, the following will be presented in a table for all analyses:

- Mean and Quantiles (2.5%, 25%, 50%, 75% and 97.5%) for the posterior probability of a participant experiencing a DLT at each of the dose levels used in the study according to the same Bayesian Logistic Regression Model (BLRM) as described in the SMC IAP (see Appendix 18.1) or SMC charter (if change occurred). For this analysis either SAS or R (using the package bcrm or crmPack) will be used. Note that 3 models need to be run separately (but all using the same prior: one for mono Q2W, one for mono Q3W and one for combination)

This table will also contain in a separate row the **MTD suggestion from the Bayesian logistic regression model** (per part and regimen) called "fitted dose with 30% DLT probability"

CCI

The dose with fitted median probability of 30% will be (item2)

For above analysis either SAS or R (using the package bcrm or crmPack) will be used.

Additionally, there will be a table showing the **frequentist MTD suggestion** (per part and regimen) only for end of combination escalation 1B and final analysis.

As second approach, the MTD as suggested by the frequentist estimation will also be derived:

1) All DLT data (in DLT analysis set) from the dose escalation will be included in a two-parameter logistic regression model (intercept and slope over log of the scaled dose, no prior; if the model does not converge, it will be reported as not calculable; the program code for this analysis can be found in the TLF shells).

2) The dose with targeted toxicity probability of 30% will be identified.

3) The next lower tested dose will be selected.

The two-sided confidence interval for the DLT probability will be provided as well as the dose with fitted probability 30% displayed.

Both these analyses (Bayesian suggestion of MTD and frequentist suggestion) will be contained in end of dose escalation analyses and repeated for final analysis (if changes occurred).

15.2 Adverse Events

AEs will be analyzed among the Safety analysis set and presented by regimen (if different regimens are tested), and dose level and summarized by study part ("Part 1A", and "Part 1B").

Treatment-emergent adverse events (TEAE) are those events with onset dates occurring during the on-treatment period or if the worsening of an event is during the on-treatment period as defined in section 9.5. If time of onset of AE and time of infusion start for M6223 and bintrafusp-alfa (if applicable) is captured, Aes with time of onset before the start of the first study drug infusion will not be TEAEs. Events that started before the start of treatment but had a decrease in grade will not be considered as TEAEs.

- **Related Adverse Events**: Aes with relationship missing, unknown or yes. Relationship is judged separately for M6223, and bintrafusp alfa.
- **SAEs**: Serious adverse events (as recorded on the AE eCRF page, Serious Adverse Event = 'Yes').
- **Aes Leading to Treatment Discontinuation**: Aes leading to permanent discontinuation of study treatment (as recorded on the AE eCRF page, Action taken with study treatment = 'Drug withdrawn'). Treatment discontinuation is recorded separately for M6223, and bintrafusp alfa.
- **Aes Leading to Death**: Aes leading to death (as recorded on the AE eCRF page, Outcome = 'Fatal').
- **Adverse Events of Special Interest**: Aes for which the investigator ticked "YES" on: "Is this an adverse event of special interest"

All analyses described in Section 15.2 will be based on TEAEs if not otherwise specified.

TEAEs will be summarized by number and percentage of participants with at least one TEAE in the category of interest, by primary MedDRA and PT (both sorted alphabetically), unless otherwise stated. Each participant will be counted only once within each SOC or PT. If a participant experience more than one AE within a SOC or PT for the same summary period, only the AE with the strongest relationship or the worst severity, as appropriate, will be included in the summaries of relationship and severity.

The severity of adverse events will be graded using the NCI-CTCAE version 5.0, except where CTCAE grades are missing. No imputation of missing grades will be performed.

Adverse events will be coded according to the latest MedDRA version at the time of the data cut-off.

Incomplete AE-related dates will be handled as defined in section 9.6.

15.2.1 All Adverse Events

The analyses in this section will be performed both at end of dose escalation (part 1A and part 1B) and final analyses unless otherwise specified.

The overall summary of Aes table will include the frequency (number and percentage) of participants with each of the following, tabulated by regimen, dose level and overall:

- Any TEAE
- Any NCI-CTCAE Grade ≥ 3 TEAE
- Any NCI-CTCAE Grade ≥ 4 TEAE
- Any M6223 related TEAE
- Any bintrafusp alfa related TEAE (for part 1B only)
- Any to both M6223 and bintrafusp alfa related TEAE (for part 1B only)
- Any M6223 related TEAE with NCI-CTCAE Grade ≥ 3
- Any bintrafusp alfa related TEAE with NCI-CTCAE Grade ≥ 3 (for part 1B only)
- Any to both M6223 and bintrafusp alfa related TEAE with NCI-CTCAE Grade ≥ 3 (for part 1B only)
- Any M6223 related TEAE with NCI-CTCAE Grade ≥ 4
- Any bintrafusp alfa related TEAE with NCI-CTCAE Grade ≥ 4 (for part 1B only)
- Any to both M6223 and bintrafusp alfa related TEAE with NCI-CTCAE Grade ≥ 4 (for part 1B only)

- Any TEAE leading to permanent discontinuation of M6223
- Any TEAE leading to permanent discontinuation of bintrafusp alfa (for part 1B only)
- Any M6223 related TEAE leading to permanent discontinuation of M6223
- Any bintrafusp alfa related TEAE leading to permanent discontinuation of bintrafusp alfa (for part 1B only)
- Any Serious TEAE
- Any M6223 related Serious TEAE
- Any bintrafusp alfa related Serious TEAE (for part 1B only)
- Any to both M6223 and bintrafusp alfa related Serious TEAE (for part 1B only)
- Any TEAE leading to death (Aes with Grade 5 or outcome "fatal" if grade 5 not applicable)
- Any M6223 related TEAE leading to death
- Any bintrafusp alfa related TEAE leading to death (for part 1B only)
- Any to both M6223 and bintrafusp alfa related TEAE leading to death (for part 1B only)
- Any AE and related AE of special interest

Also, TEAEs will be summarized in a table by worst toxicity (according to NCI-CTCAE version 5.0) per participant, using preferred term as event category and primary system organ class (SOC) (sorted alphabetically) only for end of combination dose escalation 1B and final analysis. This table will be provided by regimen, dose level and overall. If an adverse event is reported for a given participant more than once during study intervention, the worst toxicity and the worst relationship to study intervention will be tabulated. In case a participant had events with missing and non-missing grades, the maximum of the non-missing grades will be displayed.

Such a SOC and PT table will also be produced only containing TEAEs grade >=3 at final analysis. SOC and PT tables for all TEAEs and only grade >=3 TEAES will additionally be produced containing only TEAEs judged to be related by investigator.

Additionally a table will display the frequency of each PT (per dose level and overall per study part) for PTs with an incidence of >=10% in the respective part 1A mono or 1B combination. Sorting should be by descending absolute frequency for final analysis.

A listing of all AEs will be provided (in section 9) only for end of combination dose escalation 1B and final analysis. This listing will be sorted by study part, regimen, actual dose level and participant. This listing will additionally contain planned dose, age, sex, site of primary tumor, SOC, PT, grade, SAE (yes/no), DLT per investigator (yes/no), DLT per SMC, relatedness to

M6223 (yes/no), to bintrafusp alfa (yes/no) (if applicable, in part 1B only), start date (+treatment day), stop date (+treatment day) (and if applicable change date), duration (in days), action taken with M6223, if treatment was stopped before start of AE: days since stop of treatment, and outcome.

Similarly, listing of all adverse events grade >=3 (incl non TEAEs) will be provided with a flag for onset outside of the on-treatment period (for all analyses).

15.2.2 Adverse Events Leading to Discontinuation of Study Treatment

A listing of TEAEs leading to treatment discontinuation, interruption, or dose reduction of each study drug will be provided (one for M6223, additionally in part 1B: one for bintrafusp alfa). This listing, sorted by regimen, actual dose level and participant ID, will include: dose level (actual for safety and planned), participant ID, age, sex, site of primary tumor, AE investigator term, SOC, PT, relationship to M6223 and bintrafusp alfa (if applicable, in part 1B only), seriousness, action taken with M6223 or bintrafusp alfa (if applicable, in part 1B only) (displayed as: Disc/IR/DR), start date (+treatment day), stop date (+treatment day), duration (in days), if treatment was stopped before start of AE: days since stop of treatment, and outcome.

15.3 Deaths, Other Serious Adverse Events, and Other Significant Adverse Events

15.3.1 Deaths

All deaths, deaths within 30 days after last dose of study treatment, death within 60 days after first dose of study treatment as well as primary reason for death will be tabulated based on information from the "Death" eCRF page.

- Number of Deaths and by primary reason
- Number of Deaths within 30 days after last dose of study treatment and by primary reason
- Number of deaths within 60 days after the first dose of study treatment

Primary Reason for Death

- progressive disease and/or disease related condition
- Event unrelated to study treatment
- Event related to study treatment

- Unknown

This table is only to be produced for final analysis

A listing of Deaths will be provided in all analyses. This will be sorted by study part, regimen, actual dose level and participant. This listing will additionally contain planned dose, age, sex, site of primary tumor, SOC, PT, grade, SAE (yes/no), DLT per investigator (yes/no), relatedness to M6223 (yes/no), to bintrafusp alfa (yes/no) (if applicable, in part 1B only), start date (+treatment day), stop date (+treatment day), duration (in days), action taken with M6223, if treatment was stopped before start of AE: days since stop of treatment, and outcome.

15.3.2 Serious Adverse Events

A listing of SAEs) with a flag for SAEs with onset outside of the on-treatment period will be provided.

In final analysis the number of participants with serious AEs (SAEs, treatment emergent only)) will be described by PT (descending order of frequency per study part):

- SAEs
- Related SAEs

15.3.3 Other Significant Adverse Events

A listing of AEs of special interest (AESI) by investigator judgement will be provided at end of combination dose escalation (1B) and final analysis. This will be sorted by actual dose level and participant. This listing will additionally contain planned dose, age, sex, site of primary tumor, SOC, PT, grade, SAE (yes/no), DLT per investigator/SMC (yes/no), relatedness to M6223 (yes/no), to bintrafusp alfa (yes/no) (if applicable, in part 1B only), start date (+treatment day), stop date (+treatment day), duration (in days), action taken with M6223, if treatment was stopped before start of AE: days since stop of treatment, and outcome. Listing will only include those AESIs that investigator ticked as AESIs.

The following Aes may be AESIs:

- Infusion-Related Reactions including Immediate Hypersensitivity (IRRs) (SMQ Hypersensitivity)
- Immune-Related Adverse Events (irAEs) (SMQ Immune-mediated/autoimmune disorder)
- Thromboembolic events (SMQ Embolic and thrombotic events)

- Skin lesions possibly due to TGFβ inhibition (tick mark "Is this event a TGF-beta mediated skin event" ticked on AE page
- anemia
- And additionally (not in eCRF definition of AESI, to be included additionally): bleeding events. (SMQ haemorrhage terms (excl laboratory terms))

The listing should also include any other event that the investigator ticked as AESI.

The listing will only be produced for end of combination dose escalation 1B and final analysis.

A frequency table of AESIs by PT will be produced for final analysis. Sort order is by decreasing frequency per study part.

15.4 Clinical Laboratory Evaluation

No lab value TLFs will be produced for end of mono dose escalation part 1A.

Laboratory values (including corresponding normal ranges) from the local laboratories will be used for analyses of part 1A and part 1B for summary statistics and shift tables.

Laboratory results will be classified according to the NCI-CTCAE criteria version 5.0. Some of the toxicity gradings are based on laboratory measurements in conjunction with clinical findings. Non-numerical qualifiers (with the exception of fasting flags) will not be taken into consideration in the derivation of CTCAE criteria (e.g., hypokalemia Grade 1 and Grade 2 are only distinguished by a non-numerical qualifier and therefore Grade 2 will not be derived).

Additional laboratory results that are not part of NCI-CTCAE will be presented according to the categories: below normal limit, within normal limits and above normal limit (according to the laboratory normal ranges).

Parameters with NCI-CTC grades available:

Hematology:

Hemoglobin (HB), Leukocytes (white blood cell decreased), Lymphocytes (lymphocyte count increased/decreased), Neutrophils / Absolute Neutrophils Count (ANC) (neutrophil count decreased), Platelet Count (PLT) (platelet count decreased),

Chemistry:

Albumin (hypoalbuminemia), Alkaline Phosphatase (alkaline phosphatase increased), Alanine Aminotransferase (ALT) (ALT increased), Amylase (serum amylase increased), Aspartate Aminotransferase (AST) (AST increased), Total Bilirubin (blood bilirubin increased, Cholesterol (cholesterol high), Creatinine (creatinine increased), Creatine Kinase (CPK increased), Potassium (hypokalemia/ hyperkalemia), Sodium (hyponatremia/ hypernatremia), Magnesium (hypomagnesemia/hypermagnesemia), Calcium

(hypocalcemia/ hypercalcemia), Glucose (hypoglycemia/hyperglycemia), Gamma Glutamyl Transferase (GGT) (GGT increased), Lipase (lipase increased), Phosphates (hypophosphatemia)

Hemostaseology: activated partial thromboplastin time (aPTT), prothrombin international normalized ratio (INR).

Parameters with NCI-CTC grades not available:

Hematology:

Hematocrit, Red Blood Cell (RBC), Reticulocytes, Reticulocyte Volume Distribution Width Coefficient of Variation, Mean Corpuscular Hemoglobin (MCH), Mean Corpuscular Volume (MCV), Mean Corpuscular Hemoglobin Concentration (MCHC).

Chemistry:

C-Reactive Protein, Lactate Dehydrogenase (LDH), Total Protein, Total Urea, Uric Acid, Chloride, Iron, Ferritin, Iron Binding Capacity Saturation Serum Folate, B12, Mycobacterium tuberculosis Nucleic Acid.

Hemostaseology: Prothrombin time (PT)

Urinalysis and microscopic urinalysis: all urinalysis parameters.

Pregnancy test.

Serology.

The last measurement before study treatment (including unscheduled measurements) will serve as the baseline measurement.

Values below the detection limit will be imputed by half of the detection limit.

In case just a text value with an "> x" is reported it will be analyzed as +1 significant digit, e.g. "> 7.2 mmol" will be analyzed as 7.3.

For **WBC differential counts** (total neutrophil, lymphocyte, monocyte, eosinophil, and basophil counts), the absolute value will be used when reported. When only percentages are available (this is mainly important for neutrophils and lymphocytes, because the CTCAE grading is based on the absolute counts), the absolute value is derived as follows:

Derived differential absolute count = (WBC count) * (Differential %value / 100)

If the range for the differential absolute count is not available (only range for value in % is available) then Grade 1 will be attributed to as follows:

- Lymphocyte count decreased:
  - derived absolute count does not meet Grade 2-4 criteria, and
  - % value < % LLN value, and
  - derived absolute count ≥ 800/mm3
- Neutrophil count decreased
  - derived absolute count does not meet Grade 2-4 criteria, and
  - % value < % LLN value, and
  - derived absolute count ≥ 1500/mm3

For **calcium**, CTCAE grading is based on Corrected Calcium and Ionized Calcium (CALCIO), if available. Corrected Calcium is calculated from Albumin and Calcium as follows

Corrected calcium (mmol/L) = measured total Calcium (mmol/L) + 0.02 (40 - serum albumin [g/L])

The time windows given in Table 4 will be applied to allocate measurements to treatment day. In case of multiple blood samples are collected in the same window, the value closest to the target day will be used for analysis. If there are two values with the same time before and after the target day, the earlier value will be used for analysis. This windowing approach will be used for summary tables (if by visit data are displayed). In line plots the actual date (study day) will be used.

Liver function tests: Alanine aminotransferase (ALT), aspartate aminotransferase (AST), and total bilirubin are used to assess possible drug induced liver toxicity. The ratios of test result over upper limit of normal (ULN) will be calculated and classified for these three parameters during the on-treatment period.

Table 5 Time windows for safety laboratory analysis part 1A and part 1B

Treatment day	Window Definition
C1D1/baseline	As per baseline definition
C1D8	As per schedule of assessments
C2D1	
C3D1	
C4D1	

Treatment day	Window Definition
C5D1 C6DI C7D1	
CXD1 (after cycle 7)	
End of Study Intervention (within 7 days after last intervention/intake)	As per schedule of assessments
Follow-up 1 (30 days after last dose)	As per schedule of assessments
Follow-up 2 (90 days after last dose)	As per schedule of assessments

Since the number of participants will decrease over time, time windows will not be applied for the complete treatment period (stopped after C7D1). For summaries by treatment day, the last available laboratory measurement will be presented in addition to the treatment day defined above.

Qualitative data based on reference ranges will be described according to the categories (i.e. Low, Normal, and High). Abnormalities classified according to NCI-CTCAE toxicity grading version will be described using the worst on-treatment grade. For those parameters which are graded with two directions of toxicities such as potassium (hypokalemia/hyperkalemia), the toxicities will be summarized separately. Low direction toxicity (e.g., hypokalemia) grades at baseline and post-baseline will be set to 0 when the variables are derived for summarizing high direction toxicity (e.g., hyperkalemia), and vice versa.

TLFs:

For end of combination dose escalation and final analyses quantitative data will be summarized for actual values and absolute changes from baseline to each visit (until C7 visit). The End of treatment and safety-follow-up (30d) visit will be summarized separately.

Parameters with NCI-CTC grades available:

A shift table will summarize baseline CTCAE grade versus the worst on-treatment CTCAE grade. The highest CTCAE grade during the on-treatment period is considered as the worst grade for the summary (for end of combination 1B dose escalation analysis and final analysis).

Parameters with NCI-CTC grades not available:

Hematology and chemistry evaluations which cannot be graded per CTCAE criteria will be summarized for end of combination 1B dose escalation analysis and final analysis as frequency (number and row percentage) of participants with:

- shifts from baseline normal to at least one result above normal during on-treatment period
- shifts from baseline normal to at least one result below normal during on-treatment period

The Hemostaseology parameters, urinalysis parameters, microscopic urinalysis parameters and pregnancy test will be listed in dedicated listings presenting all corresponding collected information on the eCRF (only for final analysis).

The listings of laboratory results will be provided for all laboratory parameters (grouped in hematology, chemistry, coagulation, hemostaseology, urinanalysis and pregnancy test), only for final analysis. The listings will be sorted by parameters and assessment dates for each participant. Laboratory values that are outside the normal range will be flagged in the data listings, along with corresponding normal ranges and CTCAE grades.

Clinically significant abnormalities (identified as laboratory values having CTCAE grades >= 3) will be also presented in a separate listing final analysis only.

Line plots for Hematology

Graphical display (line plots) of neutrophils, platelets, lymphocytes and hemoglobin will be provided by actual time in days for all analyses. (with x-axis time, y-axis lab value), using different colors per regimen, dose level and different line types to identify participants. Below the plot the time on treatment will be displayed for end of combination 1B and final analysis.

If available (and feasible due to potential differences between labs) reference lines for ULN and LLN should be added to the plots.

Hepatotoxicity assessment

A plot of peak ALT versus peak total bilirubin, both relative to the upper limit of normal (ULN) will be provided for end of combination 1B and final analysis. This eDISH plot (evaluation of drug-induced serious hepatotoxicity) will have
reference lines at 3×ULN for ALT and at 2×ULN for total bilirubin. Subjects outside the lower left box will be identified by subject-ID.

A listing of all total bilirubin, ALT, AST and ALP values for participants with a post-baseline total bilirubin ≥ 2×ULN, ALT≥ 3×ULN or AST ≥ 3×ULN will be provided for final analysis.

.

15.5 Vital Signs

Vital sign summaries will include all vital sign assessments from the on-treatment period. All vital sign assessments will be listed, and those collected outside the on-treatment period will be flagged in the listing. All vital signs analyses are only applicable for final analysis.

All vital sign parameters will be summarized using descriptive statistics (mean, SD, median, Q1, Q3, minimum, and maximum) of actual values and changes from baseline for each visit over time. End of treatment visit will be summarized separately. The changes computed will be the differences from baseline. This is only applicable for final analysis.

The maximum changes of vital sign measurements screening/baseline to maximum changes after start of 1st study treatment will be grouped as follows:

Body temperature increase	< 1°C , 1-<2°C , 2-<3°C, ≥ 3 °C
Heart rate increase from baseline <100 bpm ; ≥ 100 bpm	≤20 bpm, >20 – 40 bpm, >40 bpm
Heart rate decrease from baseline <100 bpm ; ≥ 100 bpm	≤20 bpm, >20 – 40 bpm, >40 bpm
SBP increase from baseline <140 mmHg; ≥ 140 mmHg	≤20 mmHg, >20 – 40 mmHg, >40 mmHg
SBP decrease from baseline <140 mmHg; ≥ 140 mmHg,	≤20 mmHg, >20 – 40 mmHg, >40 mmHg
DBP increase from baseline <90 mmHg; ≥ 90 mmHg	≤20 mmHg, >20 – 40 mmHg, >40 mmHg
DBP decrease from baseline <90 mmHg; ≥ 90 mmHg,	≤20 mmHg, >20 – 40 mmHg, >40 mmHg
Respiration rate increase from baseline <20 bpm ; ≥ 20 bpm	≤5 bpm, >5 – 10 bpm, >10 bpm
Respiration rate decrease from baseline <20 bpm ; ≥ 20 bpm	≤5 bpm, >5 – 10 bpm, >10 bpm

For each participant the worst on-treatment value will be calculated. For the definition of baseline values see section 9.6. Missing values will define a separate category.

The following summaries will be prepared for vital sign parameters as grouped above:

- Maximal Shifts (changes in categories)

An additional participant data listing will present all changes from baseline reported in the highest categories.

15.6 Other Safety or Tolerability Evaluations

ECOG

Analysis of ECOG will be performed. A frequency shift table will be provided with the number and percentage of participants' shifts to worst on treatment ECOG status category and, by dose level and overall. Corresponding listing will be provided. All ECOG analyses only for final analysis.

ECG Endpoints by Triplicate

Triplicate 12-lead ECG in digital format will be obtained using an ECG machine that automatically calculates the heart rate and measures PR, QRS, QT, and QTc intervals until Cycle 7. Subsequent ECGs will be obtained as single recordings (normal nontriplicate safety ECGs). For immediate safety assessments, ECG will be locally analyzed at each time point.

This analysis will be done for electrocardiogram analysis set.

For each of the ECG parameters (HR, and QT, QTcF, QRS, PR intervals), descriptive statistics for the average at baseline, at each post-baseline time point and changes from baseline at each post-baseline time point will be presented at end of combination 1B dose escalation analysis (if data available) and final analysis. This table will be split by part, regimen and actual dose and overall.

Qualitative ECG abnormalities according to the categories specified in Table 6 will be listed for each participant and time point and the corresponding notable values and abnormality findings will be included in the listings for end of combination 1B dose escalation analysis.

Additionally data will be displayed by cohort, dose level, in a listing for final analysis. Abnormal values of ECG will be flagged according to the criteria presented in Table 6.

Table 6 Potentially Clinically Significant Abnormalities Criteria for ECG

Parameter	Potentially Clinically Significant Abnormalities (PCSA) Criteria
Heart rate	≤ 50 bpm and decrease from baseline ≥ 20 bpm. ≥ 120 bpm and increase from baseline ≥ 20 bpm.
PR interval	≥ 220 ms and increase from baseline ≥ 20 ms.
QRS	≥ 120 ms.
QTcF absolute	Borderline: 430-450 ms (Male), 451-470 ms (Female) Prolonged: > 450 ms (Male), >470 ms (Female) ≥ 500 ms

QTcF change from baseline	Borderline: Increase from baseline ≥ 30 ms and ≤ 60 ms Prolonged: Increase from baseline > 60 ms

Concentration-QTcF analysis.

Time-matched, replicate ECGs and PK samples collected will be used to analyze for QTcF responses using slope analysis of exposure/response. This will be part of a separate cross-study analysis and will be described in a separate SAP.

ECG Endpoints by Local Read (including single safety ECGs)

Information on ECGs by local read will be presented in the same way as the triplicate 12 lead ECG stated above. In case data 12 lead data are available at end of dose escalation, only the 12 lead ecg data will be displayed at end of dose escalation. If not available, only the local read data will be displayed. For final analysis both types of ECG data (if available) are to be analyzed

TGF-Beta mediated skin lesions

A listing will summarize the data gathered on TGF-beta mediated skin lesions that occurred after start of treatment at the final analysis.

16 Analyses of Other Endpoints

16.1 Pharmacokinetics

Pharmacokinetics analyses will only be performed at final analysis (except limited analyses using preliminary PK data for SMCs and a line plot of concentration over time (C1 including C2D1), different lines by dose and regimen, at interim end of mono dose escalation (1A))

Pharmacokinetic parameters for M6223 in Plasma and Bintrafusp alfa (where applicable) will be calculated by Fortrea, using standard non-compartmental methods and the actual administered dose. In cases where the actual sampling time is missing, calculations will be performed using the scheduled time.

Further population PK modelling may be conducted separately and reported outside of the CSR. Non-compartmental (NCA) computation of pharmacokinetic parameters will be performed using the computer program Phoenix® WinNonlin® version 6.3, or higher (Certara, L.P., 1699 S Hanley Road, St Louis, MO 63144, USA).

The following PK parameters will be calculated by for M6223, where possible, from Day 1 of treatment of Cycles 1, 2 and 4, zero hours to 14 days postdose.

Parameters	Units[a]	Definition
AUC_{0-t}	ng.h/mL	area under the concentration time curve (AUC) from zero to the last sampling time at which the concentration is at or above the LLOQ ($_{tlast}$)[b]
AUC_{τ}	ng.h/mL	area under the concentration time curve (AUC) over the dosing interval[b]
$AUC_{0-\infty}$	ng.h/mL	area under the concentration-time curve from time 0 extrapolated to infinity, based (first dose only)[c]
$AUC_{extra\%}$	%	percentage of $AUC_{0-\infty}$ that is due to extrapolation from the last measurable concentration (first dose only).
t_{max}	h	time to maximum observed concentration
C_{max}	ng/mL	maximum observed concentration
C_{trough}	ng/mL	concentration observed immediately before the next dose (corresponding to the predose concentration for multiple dosing).

λz	1/h	apparent terminal first order elimination rate constant (first dose only)
$t_{1/2}$	h	apparent terminal half-life (first dose only)
RA_{AUC}	Not applicable	accumulation ratio based upon AUC over the dosing interval (multiple dose only)
RA_{Cmax}	Not applicable	accumulation ratio based upon C_{max} (multiple dose only)
LR	Not applicable	linearity ratio (multiple dose only)
CL	L/h	total plasma clearance following intravenous dosing (first dose only)
V_{ss}	L	volume of distribution at steady state (first dose only)
V_z	L	volume of distribution during the terminal elimination phase (first dose only)

[a] Units are based on concentration units (provided by bioanalytical lab) and dose units used in the study.
[b] AUC will be calculated using the linear trapezoidal rule for increasing concentrations and the logarithmic rule for decreasing concentrations (linear up/log down rule).
[c] Based on predicted Clast

Additional pharmacokinetic parameters will be calculated where appropriate.

Dose normalised parameters for M6223 will be calculated for $AUC_{0-\infty}$, AUC_{0-t}, AUC_{τ} and C_{max} by dividing the original parameter by the dose.

The following pharmacokinetic parameters will be calculated for bintrafusp alfa where possible, from Day 1 of treatment of Cycles 1 and 4 to 14 days postdose.

Parameters	Units[a]	Definition
t_{max}	h	Time to maximum observed concentration
C_{max}	ng/mL	Maximum observed concentration
C_{trough}	ng/mL	Concentration observed immediately before the next dose (corresponding to the predose concentration for multiple dosing.

The EOI sample will be assumed to be t_{max}.

C_{max} and t_{max} will be obtained directly from the plasma concentration-time profiles.

For multiple peaks, the highest postdose concentration will be reported as C_{max}. In the case that multiple peaks are of equal magnitude, the earliest t_{max} will be reported.

Accumulation ratios will be determined from single dose and multiple dose (at steady state) C_{max}, AUC_{τ}(for which τ= 336 hours for Q2W and 504 hours for Q3W regimens) as follows

$$RA_{AUC} = AUC_{\tau,\ multiple\ dose}/AUC_{\tau,\ single\ dose}$$

$$RA_{Cmax} = C_{max,\ multiple\ dose}/C_{max,\ single\ dose}$$

$$LR = AUC_{\tau,\ multiple\ dose}/AUC_{0-\infty,\ single\ dose}$$

The Phoenix WinNonlin NCA Core Output will be provided in a separate listing.

16.1.1 Criteria for Handling Concentrations Below the Limit of Quantification in Pharmacokinetic Analysis

Concentration values that are below the level of quantification (BLQ) will be set to zero.

16.1.2 Criteria for calculation of the apparent terminal elimination rate constant and half-life

At least 3 data points will be included in the regression analysis and should not include C_{max}. The last quantifiable concentration should be included.

When assessing terminal elimination phases, the coefficient of correlation for exponential fit (R^2) will be used as a measure of the goodness of fit of the data points to the determined line. Where the, R^2 value for the regression line is ≤ 0.8 the $t_{1/2}$ and $t_{1/2}$ dependent parameters will be calculated and flagged in the PK parameter output. Any flags will be included in the study specific SDTM and the exclusion of parameters agreed with the sponsor

The start of the terminal elimination phase for each subject will be defined by visual inspection and generally will be the first point at which there is no systematic deviation from the log-linear decline in concentrations.

Where $t_{1/2}$ is estimated over a time period of less than 2 half-lives the robustness of the value should be discussed in the CSR.

The following regression-related diagnostic PK parameters will be determined:

Parameter	Units	Definition
λ_z Upper	h	end of exponential fit
λ_z Lower	h	start of exponential fit
λ_z N	Not applicable	number of data points included in the log-linear regression
λ_z Span Ratio	Not applicable	time period over which λ_z was determined as a ratio of $t_{1/2}$
R^2	Not applicable	coefficient for determination of exponential fit
Rsq, adj	Not Applicable	adjusted coefficient of determination of exponential fit

16.1.3 Calculation of Area Under the Concentration-time Curve

Area under the concentration ‑ time curve (AUC) will be calculated using the mixed log-linear trapezoidal rule (linear up - log down).

The minimum requirement for the calculation of AUC will be the inclusion of at least 3 consecutive plasma concentrations above the lower limit of quantification (LLOQ), with at least one of these concentrations following C_{max}

For any partial AUC determination (i.e. AUC over a dosing interval), nominal time will generally be used for the end of the interval. Actual times for partial AUC intervals may be used if the sample was taken within +/-10% of the protocol defined time after sponsor agreement.

Data will not be automatically excluded on the basis of percentage extrapolated. However, potentially unreliable values will be discussed with the sponsor.

If $AUC_{0-\infty}$ cannot be determined for all participants or all dose levels, an alternative AUC measure, such as AUC to a fixed time point, may be used in the assessment of dose proportionality.

For multiple dose the pre-dose concentration may be duplicated and used as the 336-hour (Q2W) and 504 hour (Q3W) concentration, to allow AUC_τ to be calculated (assuming steady state) if the following predose sample is missing. Similarly, the 336 hour (Q2W) or 504 hour (Q3W) concentration may be imputed for missing predose concentrations.

16.1.4 Anomalous Values

Handling of anomalous values will be discussed with the sponsor. PK concentrations which are erroneous due to a protocol violation (as defined in the CTP), documented handling error or analytical error (as documented in the bioanalytical report) may be excluded from the PK analysis if agreed upon prior to performing a statistical analysis. In this case the rationale for exclusion must be provided in the Clinical Trial Report (CTR). Any other PK concentrations that appear implausible to the Pharmacokineticist/PKPD Data Analyst must not be excluded from the analysis. Any implausible data will be documented in the Clinical Trial Report (CTR).

Positive predose values at Day 1 >5% of C_{max} may be excluded from the PK analysis, summary statistics and statistical analysis at the discretion of the pharmacokineticist.

16.1.5 Dose Proportionality Assessment

Dose proportionality for M6223 will be provided separately for $AUC_{0-\infty}$, $AUC_{0-tlast}$,AUC_{τ}, and C_{max} on study Cycle 1, Cycle 2 and Cycle 4 for Part 1A, monotherapy and Part 1B, combination therapy using a power model. The power model will have the form:

$$Y = a*(dose)^b$$

where Y is the PK parameter, and a and b are the coefficient and exponent, respectively, of the power equation. By taking the natural logarithm (ln), the power model can be analysed using linear regression and has the form:

$$\ln(Y) = \ln(a)+b*\ln(dose) + error$$

$$=\alpha + \beta*\ln(dose) + error,$$

where α is the intercept, and β is the slope, and ln(dose) is based on the dose size for each subject. For each PK parameter, the point estimate of slope β and its 95% CI will be provided. In addition, the estimated regression line will be overlaid on the individual points in the corresponding scatter plot (log-log scale).

Exploratory dose proportionality will be declared if the 90% confidence interval (CI) for β is entirely within the critical region,

$$(1+ \frac{\ln(0.50)}{\ln(r)}, 1 + \frac{\ln(2.00)}{\ln(r)})$$

where r is the ratio of the highest and the lowest dose in a given part of the study. Dose proportionality for participants who were administered the study intervention will be tested for C_{max} and $AUC_{0-\infty}$, AUC_{τ}, on Cycle 1, Cycle 2 and Cycle 4.

Since the study is neither optimally designed nor powered to confirm the presence or absence of meaningful departures from dose proportionality, these results should be interpreted with caution. In addition, graphical assessment would also be considered to understand dose proportionality.

If the samples remain above LLOQ for the whole dosing interval, then $AUC_{0-tlast}$ will be the same as AUC_{τ}.

16.1.6 Presentation of Pharmacokinetic Data

16.1.6.1 Descriptive Statistics of PK Concentration Data

Available individual M6223 and bintrafusp alfa (Part 1B) serum concentrations will be listed and summarized by Part, Cycle, dose, analyte and scheduled collection time, including summary statistics where data allows, using: number of non-missing observations (N), arithmetic mean (Mean), standard deviation (SD), coefficient of variation (CV%), minimum (Min), median (Median) and maximum (Max). Pharmacokinetic parameters will be summarized additionally, using the geometric mean (GeoMean), the geometric coefficient of variation (GeoCV) and the two-sided 95% confidence interval for the GeoMean (LCI 95% GM, UCI 95% GM). For time to reach maximum observed concentration (t_{max}), only n, Min, Median, and Max will be reported.

Missing statistics, e.g. when they cannot be calculated, should be presented as “NC”. “ND” will be used if a PK sample was not taken.

Descriptive statistics will only be calculated for N>2 in which a measurement of <LLOQ represents a valid measurement set to 0.

Descriptive statistics of PK concentration data will be calculated using values with the same precision as the source data and rounded for reporting purposes only. In export datasets, as well as in the SDTM PC domain, PK concentrations will be provided with full precision and will not be rounded.

Concentration data collected outside of the protocol specified window may be excluded from summary statistics with sponsor agreement.

The following conventions will be applied when reporting descriptive statistics of PK concentration data:

N, n	0 decimal places
Mean, Min, Median, Max,:	3 significant digits
SD	4 significant digits
CV%, GeoCV%:	1 decimal place

16.1.6.2 Descriptive Statistics of PK Parameter Data

Pharmacokinetic parameters will be summarized, using: number of subjects (N), number of non-missing observations (n), arithmetic mean (Mean), standard deviation (SD), coefficient of variation (CV%), minimum (Min), median (Median), maximum (Max), using the geometric mean (GeoMean), the geometric coefficient of variation (GeoCV) and the two-sided 95% confidence interval for the GeoMean (LCI 95% GM, UCI 95% GM). For PK parameters related to observed time (t_{max}, t_{last}), only n, Min, Median, and Max will be reported.

Descriptive statistics will be calculated for pharmacokinetic parameter where N>2. All no result (NR) and not done (nd) values will be set to missing.

PK parameters read directly from the measurements (i.e. C_{max}) will be reported with the same precision as the source data. All other PK parameters will be reported to 3 significant figures. In export datasets, as well as in the SDTM PP domain, PK parameters will be provided with full precision and will not be rounded. Descriptive statistics of PK parameter data will be calculated using full precision values and rounded for reporting purposes only.

The following conventions will be applied when reporting descriptive statistics of PK data:

N, n	0 decimal places
Mean, Min, Median, Max, GeoMean, two-sided 95% CI:	3 significant digits
SD:	4 significant digits
CV%, GeoCV%: 1 decimal place	

16.1.6.3 Listings and Tables

Individual M6223 serum concentrations will be reported by Part, Cycle, dose and scheduled collection time, including summary statistics where data allows, based upon the PK analysis set.

- Statistical assessment of dose proportionality, separate tables for Part A and Part B. Do not include Q3W data.
- Descriptive statistics of concentration of M6223 in serum by Part, Cycle and dose level for Cycles 1, 2 and 4.
- Descriptive statistics of C_{trough} (predose) concentrations of M6223 in serum by Part, Cycle and dose level for all available cycles.
- Descriptive statistics of concentration of bintrafusp alfa (Part B only) in serum by Part, Cycle and dose level for all available cycles.
- Descriptive statistics of serum PK parameters for M6223 by Part, Cycle, and dose level (do not include diagnostic parameters) for cycles 1, 2 and 4.
- Descriptive statistics of serum PK parameters for bintrafusp alfa by Part, Cycle and dose level for cycles 1, 2 and 4.

The following PK listings will be produced based upon the safety analysis set.

- Plasma M6223 PK sampling date, actual time, nominal time, deviation from time, percentage time deviation, concentration by participant, part and dose level status sorted in chronological order
- Individual plasma concentrations of M6233 by Part, Cycle and dose level for cycles 1 to 4
- Individual plasma concentrations of M6233 by Part, Cycle and dose level for cycles > 4
- Individual plasma concentrations of bintrasup alfa by Part, Cycle and dose level
- Individual plasma pharmacokinetic parameters for M6233 by Part, Cycle and dose level (cycles 1, 2 and 4)
- Individual plasma pharmacokinetic parameters for bintrafusp alfa by Part, Cycle and dose level (cycles 1, 2 and 4)
- Plasma diagnostic parameters for M6223 by Part, Cycle and dose level

- Phoenix WinNonlin core output

16.1.6.4 Graphical Summaries and Individual Plots (PK Analysis Set)

Both individual and arithmetic mean concentration-time profiles, will be plotted on linear and semi-logarithmic scales as follows:

- Arithmetic mean concentration vs scheduled time profile showing M6223 grouped by cycle with dose levels overlaid, separate plots for Part A and Part B. On linear plots ± SD will be included.
- Arithmetic mean concentration vs scheduled -time profile showing M6223 grouped by dose level with Cycles overlaid, separate plots for Part A and Part B. On linear plots ± SD will be included
- Arithmetic mean concentration vs scheduled vs time profile showing M6223 plot of Part 1A and 1B at the same dose overlaid, by dose level with cycles overlaid. On linear plots ±SD will be included.
- Arithmetic mean Ctrough concentrations vs scheduled time with dose levels overlaid (linear scale only)
- Combined indivdual plots showing M6223 serum concentration vs actual time grouped by Part, dose and Cycle with individual participants overlaid
- Combined indivdual plots showing M6223 serum Ctrough concentration vs scheduled time grouped by Part and dose level with individual participants overlaid (linear plot only)
- Individual M6223 serum concentrations grouped Part, dose level and participant with cycles overlaid

The following plots will also be provided, for key PK parameters:

- $_{max,}$ $AUC_{0-\infty}$, $AUC_{0-tlast}$ and AUC_{τ} versus dose as a scatter plot. Any regression line should not be fixed through the origin. Linear – linear and semi-logarithmic – semi-logarithmic scale. Separate plots for Part A and Part V

- Dose normalized $C_{max,}$ $AUC_{0-\infty}$, $AUC_{0-tlast}$ and AUC_{τ} versus dose as a scatter plot showing the geometric mean value.

For all linear PK plots, the x and y axes should both begin at (0,0).

16.2 Pharmacodynamics

16.2.1 TIGIT target occupancy profile:

Analysis will be by study part (1A, 1B) except otherwise specified. All analyses will be provided for end of combination 1B dose escalation analysis and final analysis, unless otherwise specified.

Blood samples for TIGIT target occupancy will be collected at baseline and on-treatment time points. Samples not fulfilling quality criteria will be excluded, post-baseline samples of participants where baseline sample was excluded need to be excluded also. The results will be presented as % TIGIT target occupancy (ratio on-treatment to baseline). Summary statistics of TIGIT target occupancy at visit timepoints C1D2, C2D1 and C4D1 will be performed on the pharmacodynamics analysis set. Corresponding listing will be provided only for final analysis.

Visit windows applicable for this analysis will be:

Treatment day	Window Definition
C1D2	- 1 day(but after treatment start), + 5 days
C2D1	+/- 7 days
C4D1	+/- 10 days

Line plots will show the percent TIGIT target occupancy versus time, one plot for each regimen/dose level. Each participant in a cohort will be differentiated from the others by using a participant-specific symbol. Data from the first two single-patient cohorts will be illustrated on a single plot. A value of a patient that is M6223 ADA positive at the time of the TO value will be signified with an open symbol (e.g filled circle normal symbol, not-filled circle symbol at time participant is M6223 ADA positive)

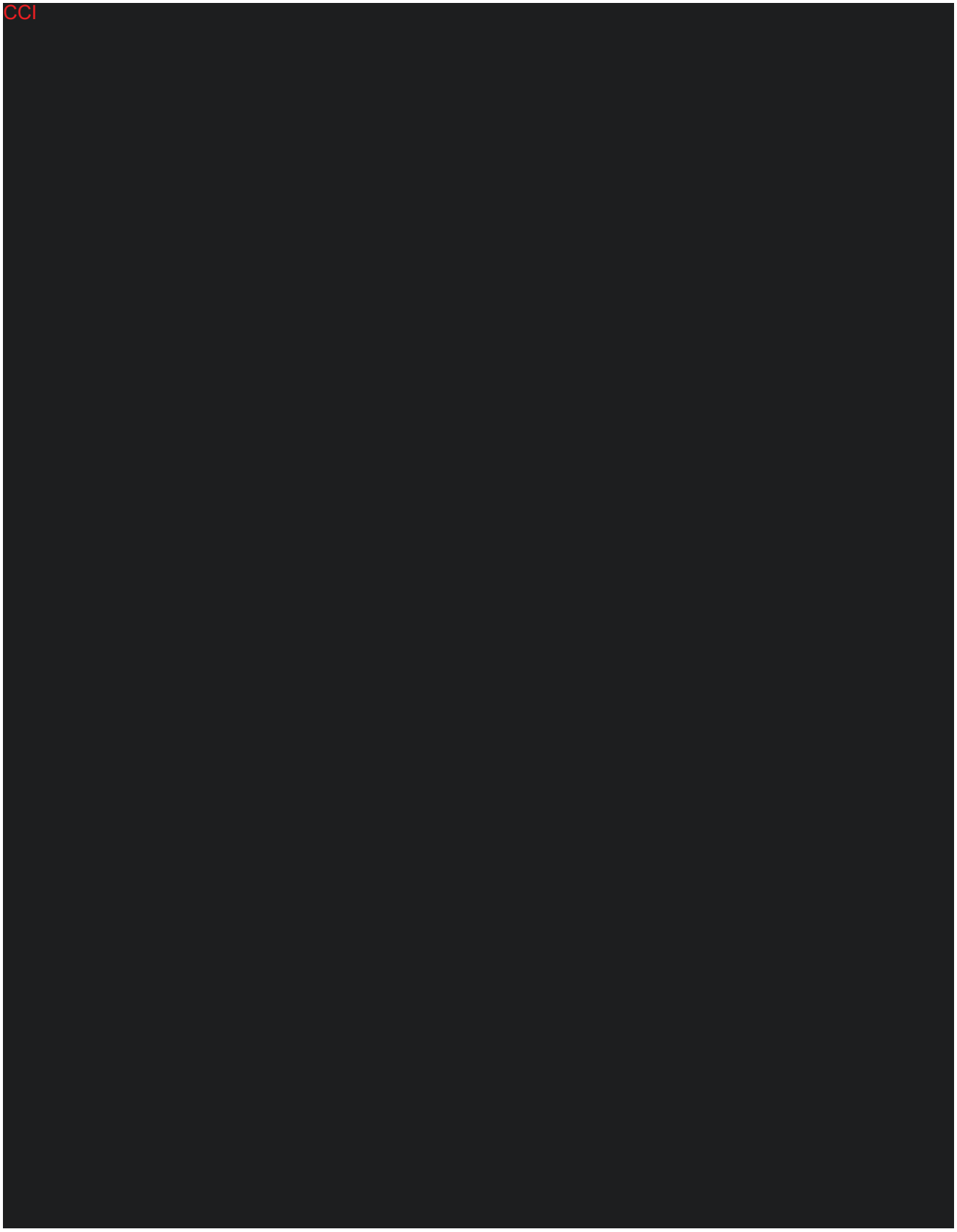
A line plot will show the mean TIGIT target occupancy for each regimen/dose level over time (different colors to distinguish dose, different line types to distinguish different regimen).

Dot plots will show TIGIT target occupancy data of individual participants versus dose (x-axis dose, y-axis % TIGIT target occupancy) on C2D1 at all analysis. Separate additional plots may be generated for C1D2 and C4D1 in final analysis. For end of monotherapy dose escalation and final analysis the participant symbols (“dots”) will be open (not filled) at times when a patient is M6223 ADA positive.

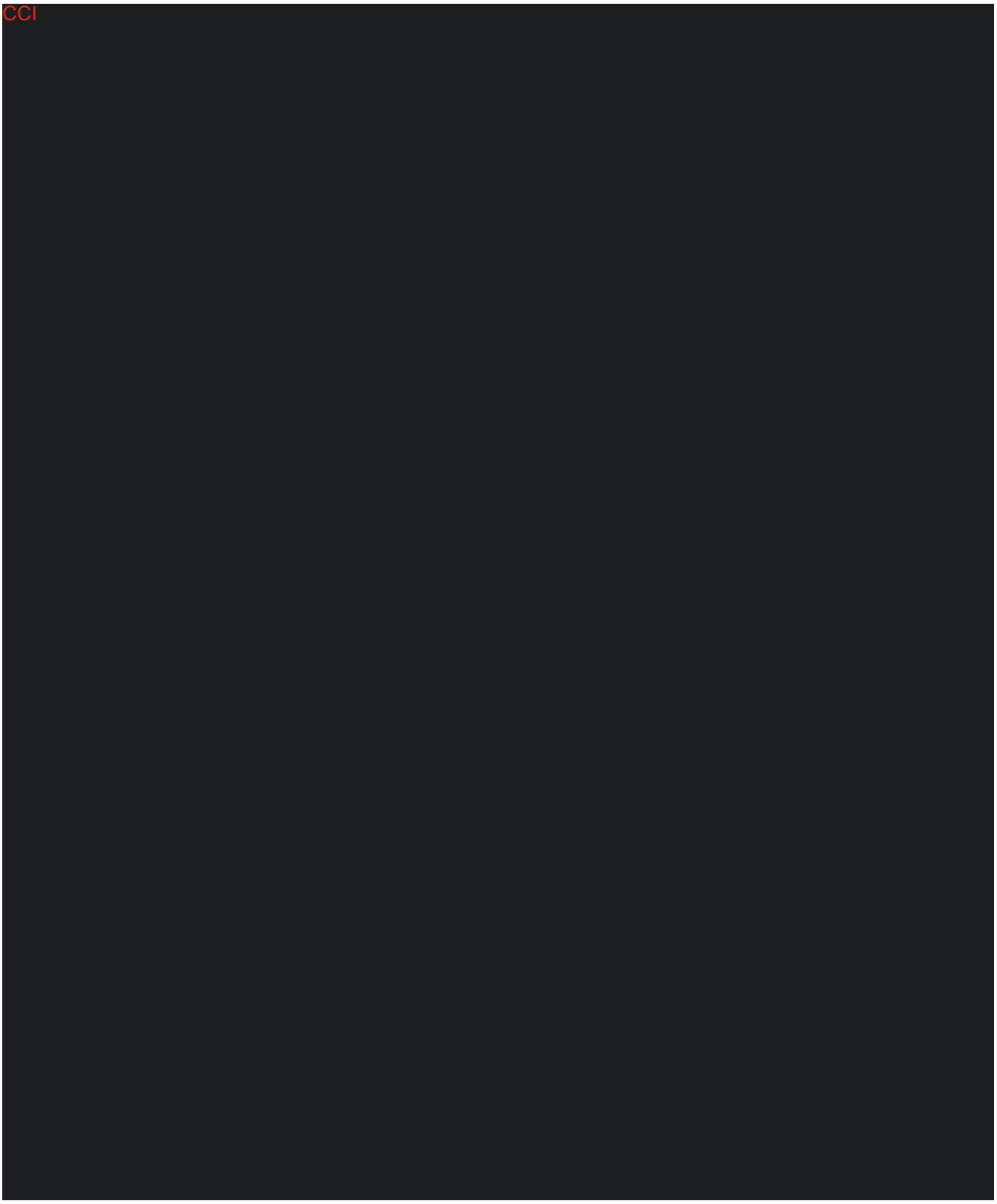
CCI

CCI

CCI



CCI



16.3 Immunogenicity (ADA)

Immunogenicity data will be collected for all participants. Participants having at least one valid antidrug antibody (ADA) result will be included in the Immunogenicity analysis set.

The immunogenicity classification of each study participant will be evaluated according to the criteria described in Table 7 separately for M6223 and BA (the latter for patients in combination). If screening value is negative, ADA result is negative. If screening result is positive, ADA result is positive, if confirmation result is also positive, if confirmation result is negative, ADA result is negative. If screening is positive but confirmation is having no result, the ADA result is “no result”

The immunogenicity classification (per table 7) will be listed for all participants in the Immunogenicity analysis set together with the ADA result at baseline and the last available measurement. The same listing will be prepared for the combination cohorts containing the BA ADA results.

An additional listing will show all measurements (results as per table 7 and titers) for all subjects with at least one positive result (including baseline) for M6223 ADAs and one listing for Bintrafusp alfa ADAs (up to cycle 19 only, including EOT and Safety follow-up visit).

Summaries may be done in-house at Merck, e.g. summarizing the number of participants in per category or per broader category (e.g. “ever positive”, “treatment emergent”) and calculating percentages of specific categories based on the number of patients that could fall in this category based on baseline or on-treatment results.

ADA outputs will only be produced for end of combination 1B dose escalation and final analysis.

Table 7 Immunogenicity classification

ADA result at Baseline	ADA results on treatment	Classification
Negative	Negative at all measurements	Never positive
Negative or missing	• Only one positive result post-baseline (not at EOT) or • duration between first and last positive result <16 weeks and last ADA assessment not positive	Treatment-emergent transient positive
Negative or missing	• At least one positive result post baseline and duration between first and last positive result ≥16 weeks or • last ADA assessment positive	Treatment-emergent persistent positive
Positive	Negative, or positive with titer lower or equal to baseline	Pre-existing positive
Positive	At least one positive post-baseline result with titer >= 8 fold higher than baseline	Treatment-boosted
Missing	Negative	Not evaluable
Negative or positive	Missing	Not evaluable
Missing	Missing	Not evaluable

16.4 Exploratory Biomarker Analyses

Exploratory biomarker analyses will depend on the availability of samples and data collected, and could be described in a separate exploratory analysis plan as needed.

16.4.1 PD-L-1 expression

PD-L1 expression in baseline tumor tissue will be scored in tumor cells, immune cells separately or combined. The association of PD-L1 expression to response will be analyzed. See inclusion of PD-L1 status in waterfall plot (described in section 14.3). Additionally, a dot plot will display the PD-L1 expression by BOR and by tumor type and BOR. A listing will display the tumor type, the PD-L1 expression level and BOR. These analyses will be done at final analysis.

Further exploratory analysis will be described in additional analysis plan if needed.

17 References

Brookmeyer R, Crowley J. A Confidence Interval for the Median Survival Time. Biometrics 1982; 38, 29–41. DOI: 10.2307/2530286

Kalbfleisch JD, Prentice LP. The Statistical Analysis of Failure Time Data. Chapter 8: Competing Risks and Multistate Models. Wiley Series in Probability and Statistics, 2nd edition, 2011. Print ISBN: 9780471363576, Online ISBN: 9781118032985 (http://onlinelibrary.wiley.com/book/10.1002/9781118032985)

18 Appendices

18.1 Appendix 1 – IAP for SMC

Integrated Analysis Plan for SMCs

Clinical Study Protocol Identification No.	MS201430-0001
Title	Phase I, First-in-Human, Open-Label, Multiple Ascending Dose Study to Investigate the Safety, Tolerability, Pharmacokinetics, Pharmacodynamics and Clinical Activity of M6223 (an Inhibitor of TIGIT) as Single Agent and in Combination with Bintrafusp alfa (anti-PDL1/TGFß Trap in Participants with Metastatic or Locally Advanced Solid Unresectable Tumors
Study Phase	I
Investigational Medicinal Product(s)	M6223
Clinical Study Protocol Version	14 August 2019 / Version 1.0
Integrated Analysis Plan Author	**Coordinating Author** Biostatistics, Merck — PPD **Function** — **Author(s) / Data Analyst(s)** Quantitative Pharmacology, Merck Healthcare KGaA — PPD
Integrated Analysis Plan Date and Version	10 September 2020 / Version 1.0
Integrated Analysis Plan Reviewers	**Function** — **Name** Medical — PPD GPS Biomarker Biostatistics PPD Programming

M6223
MS201430-0001
First in Human Study of M6223
v1.0

Approval Page

Integrated Analysis Plan: MS201430-0001

Phase I, First-in-Human, Open-Label, Multiple Ascending Dose Study to Investigate the Safety, Tolerability, Pharmacokinetics, Pharmacodynamics and Clinical Activity of M6223 (an Inhibitor of TIGIT) as Single Agent and in Combination with Bintrafusp alfa (anti-PDL1/TGFß Trap in Participants with Metastatic or Locally Advanced Solid Unresectable Tumors

Approval of the IAP by all Merck Data Analysis Responsible has to be documented within ELDORADO via eSignature. With the approval, the Merck responsible for each of the analysis also takes responsibility that all reviewers' comments are addressed adequately.

By using eSignature, the signature will appear at the end of the document.

M6223
MS201430-0001

First in Human Study of M6223
v1.0

1 Table of Contents

Approval Page 2
1 Table of Contents 3
2 List of Abbreviations and Definition of Terms 5
3 Modification History 6
4 Purpose of the Integrated Analysis Plan 6
5 Objectives and Endpoints 6
6 Overview of Planned Analyses 8
7 Changes to the Planned Analyses in the Clinical Study Protocol 8
8 Analysis Populations and Subgroups 9
8.1 Definition of Analysis Populations 9
8.2 Subgroup Definition and Parameterization 10
9 General Specifications for Data Analyses 10
9.1 Study Treatment Day 12
9.2 Definition of Duration and 'time since' Variables 12
9.3 Conversion Factors 12
9.4 Imputation of Missing Data 12
10 Study Participants 13
10.1 Disposition of Participants and Discontinuations 13
10.2 Protocol Deviations / Exclusion from Analysis Populations 14
10.2.1 Important Protocol Deviations 14
10.2.2 Reasons Leading to the Exclusion from an Analysis Population 14
11 Demographics and Other Baseline Characteristics 14
11.1 Demographics 14
11.2 Medical History 14
11.3 Other Baseline Characteristics 14
12 Previous or Concomitant Medications/Procedures 14
13 Study Treatment: Compliance and Exposure 14
14 Efficacy Analyses 15
15 Safety Analyses 16
15.1 DLT analysis 16
15.2 Patient profiles 18

15.3 Adverse Events 19
15.3.1 All Adverse Events 20
15.3.2 Adverse Events Leading to Discontinuation of Study Treatment 20
15.4 Deaths, Other Serious Adverse Events, and Other Significant Adverse Events 21
15.4.1 Deaths 21
15.4.2 Serious Adverse Events 21
15.4.3 Other Significant Adverse Events 21
15.5 Clinical Laboratory Evaluation 21
15.6 Vital Signs 22
15.7 Other Safety or Tolerability Evaluations 22
CCI [redacted] 23
16.1 Pharmacokinetics 23
CCI [redacted] 23
17 References 24
18 Appendices 25
18.1 Appendix 1 – Mock TLFs 26

2 List of Abbreviations and Definition of Terms

ADaM	Analysis Data Model
AE	Adverse Event
AESI	Adverse Events of Special Interest
CDISC	Clinical Data Interchange Standards Consortium
CI	Confidence Interval
CR	Complete Response
(e)CRF	(electronic) Case Report Form
CSR	Clinical Study Report
DLT	Dose Limiting Toxicity
ECG	Electrocardiogram
GBS	Global Biostatistics
IAP	Integrated Analysis Plan
MedDRA	Medical Dictionary for Regulatory Activities
MTD	Maximum Tolerated Dose
NA	Not Applicable
NCI-CTCAE	National Cancer Institute – Common Terminology Criteria for Adverse Events
ORR	Objective Response Rate
Pd	Pharmacodynamics
PT	Preferred Term
PGx	Pharmacogenetics/Pharmacogenomics
PK	Pharmacokinetics
PP	Per Protocol
PR	Partial Response
SAE	Serious Adverse Event
SCR	Screening analysis population
SD	Stable Disease or Standard Deviation
SDTM	Study Data Tabulation Model
SMC	Safety Monitoring Committee
SOC	System Organ Class

TEAE	Treatment-Emergent Adverse Event
TLF	Tables, Listings, and Figures
WHO-DD	World Health Organization Drug Dictionary

3 Modification History

Unique Identifier for Version	Date of IAP Version	Author	Changes from the Previous Version
1.0	10 Sep 2020	PPD	Initial version

4 Purpose of the Integrated Analysis Plan

The purpose of this Integrated Analysis Plan (iAP) is to document technical and detailed specifications for the analyses performed for the Safety Monitoring Committee (SMC) reviews of data collected for protocol MS201430-0001. Results of the analyses described in this iAP will be used (amongst other data) by the SMC to decide upon dose of future cohorts. This iAP describes the Bayesian two-parameter logistic regression model analysis methods used to make a recommendation to the SMC for dose escalation. Additionally, this iAP will describe preliminary PK and Pd analyses for M6223, as well as patient profiles and Tables, Listings, and Figures (TLFs) summarizing safety data. TLFs will only be produced if a SMC is taking place and data of more than 4 subjects treated in the respective study part are available (for part 1A expected to begin at SMC after cohort 3 (cohort 1 and 2 are expected to be n=1) and at the 2nd SMC for part 1B). Biomarker analyses will be provided for SMC by biomarker representative, PK analyses by the Clin Pharm representative.

The IAP is based upon Section 9 (Statistics) of the trial protocol and protocol amendments and is prepared in compliance with ICH E9.

5 Objectives and Endpoints

Objectives	Endpoints (Outcome Measures)	IAP section
Primary		
To determine safety and tolerability and (if observed) the maximum tolerated dose (MTD) of M6223 as a single agent (Part 1A) and of M6223 combined with bintrafusp alfa (Part 1B) in participants with advanced solid tumors who are refractory to or have progressed under standard treatment and have no other treatment options known to confer clinical benefit.	DLTs, TEAEs, changes in lab values, changes in ECG and changes in ECOG	
To determine the recommended dose of M6223 monotherapy (Part 1A) and of M6223 combined with bintrafusp alfa (Part 1B) for further exploratory clinical development. Primary endpoints (see above) and available data for secondary and exploratory endpoints (see below).	DLTs, TEAEs, changes in lab values, changes in ECG and changes in ECOG, PK, Tigit target occupancy at Cycle 1, 2 and 4.	
Secondary		
To characterize the PK profile of M6223 when given alone (Part 1A) or when combined with bintrafusp alfa (Part 1B).	PK parameters of M6223 in terms of AUC_{0-t}, $AUC_{0-\infty}$, AUC_T, C_{max}, C_{trough}, t_{max}, $t_{1/2}$ and λ_z (Part 1A and Part 1B) on Day 1 of Treatment of Cycle 1, 2 and 4 from time zero to 14 days postdose.	
To characterize the PK of bintrafusp alfa in combination with M6223 (Part 1B) by sparse sampling.	PK parameters of bintrafusp alfa in terms of C_{max} and C_{trough} on Day 1 of Treatment of Cycle 1 and 4 from time zero to 14 days postdose.	
To characterize immunogenicity of M6223 when given alone and immunogenicity of M6223 and bintrafusp alfa when given in combination and its relationship to drug exposure.	Immunogenicity of M6223 (Part 1A and Part 1B) and bintrafusp alfa (Part 1B), as measured by AntiDrug-Antibody (ADA) assays, from predose of Cycle 1 Day 1 through 30 days safety follow-up Visit.	
To characterize the relationship between exposure and QT interval of M6223 alone and in combination with bintrafusp alfa.	QT interval from Cycle 1 to Cycle 7.	
To characterize preliminary clinical activity parameters of M6223 in monotherapy (Part 1A) and in combined with bintrafusp alfa (Part 1B) using RECIST 1.1.	OR, DOR, time to response, disease control (DC), progression free survival (PFS) using RECIST 1.1, per Investigator, and overall survival (OS).	
Tertiary/Exploratory		
To assess the peripheral TIGIT target occupancy with M6223 alone (Part 1A) and combined with bintrafusp alfa (Part 1B) and exposure/ target occupancy relationship.	Target occupancy at Cycle 1, 2 and 4.	

CCI [redacted]

6 Overview of Planned Analyses

The following analyses are planned for this trial:

- SMC analyses (this iAP, it will be added as an appendix of the main iAP, once available)
- End of monotherapy dose escalation analysis (part 1A) (part of main iAP)
- End of combination dose escalation analysis (part 1B) (part of main iAP)
- Primary analysis of trial (main iAP)

Once all patients of the respective cohort have completed the DLT period or discontinued from trial prematurely, a data snapshot will be taken for provision of SMC outputs. The usual data snapshot is taken at the end of this day. In cases where enrollment of the last subject in a dosing cohort is delayed, the SMC may decide (based on available data) upon enrollment and dose for the next dosing cohort before all subjects in a cohort have completed the dosing cycle and thus the snapshot is set to an earlier date. The SMC can also determine any other date, e.g. for an ad-hoc SMC.

There will be no data cut-off applied (all data in the data transfer are considered).

The patient profiles and TLFs on safety and safety lab data for SMCs will be produced on a raw data export. Patient profiles will be created by Xcellerate, TLFs will be based on SAS data exports from the EDC. No SDTM transfers are foreseen for analyses for SMCs.

7 Changes to the Planned Analyses in the Clinical Study Protocol

Definition of TEAE differs from what is mentioned in protocol. The protocol mentions : "from Screening up to the 30 days Safety Follow-up Visit". The definition used is "from start of treatment until (including) 30 days after the end of treatment).

Definition of Pd analysis set differs, as the condition regarding the need of a predose sample has been added: "in addition to a predose measurement". The full definition now reads: All participants, who receive at least one dose of study intervention, have no clinically important protocol deviations or important events affecting Pd and provide at least one measurable Pd endpoint postdose in addition to a predose measurement.

8 Analysis Populations and Subgroups

8.1 Definition of Analysis Populations

Analysis Set	Description
Screening (SCR)	All participants, who provided informed consent.
Safety (SAF)	All participants, who were administered any dose of any study intervention. Analyses will consider participants as treated.
Dose Limiting Toxicity (DLT)	The DLT analysis set will include all participants who received at least one dose of study intervention and meet at least one of the following criteria: • Experienced at least one DLT during the DLT period, regardless of the administered number of doses of study treatment/completion of the DLT period. • Received at least 80% of the planned cumulative dose of each treatment during the DLT period and completed the DLT period. • Additionally, participants that did not receive 80% of the planned total dose of study treatment during the DLT period, but at least 80% dosing of a different (lower) dose cohort during the completed DLT period are eligible for the DLT analysis set to be analyzed in the highest dose cohort for which they received 80% of dosing.
Pharmacokinetics (PK)	All participants, who receive at least one dose of study intervention, have no clinically important protocol deviations or important events affecting PK, and provide at least one measurable postdose concentration. Participants will be analyzed per the actual study intervention they received. The PK population will include all participants: • Who have completed all study periods without any relevant protocol deviations and factors likely to affect the comparability of PK results. • With adequate study intervention compliance. • With evaluable PK data, i.e., non-missing values for primary endpoints in each study period. If participants received prohibited concomitant therapy or medicines, as specified in Section 6.5, they will be excluded from the PK population. Relevant decisions will be made before database lock. All PK analyses will be based on this analysis set.
Pharmacodynamic (Pd)	All participants, who receive at least one dose of study intervention, have no clinically important protocol deviations or important events affecting Pd and provide at least one measurable Pd endpoint postdose in addition to a predose measurement. Participants will be analyzed per the actual study intervention they received. All Pd analyses will be based on this analysis population.
Immunogenicity	All participants who receive at least one dose of study intervention and have at least one valid antidrug antibody (ADA) result. All ADA analyses will be based on this analysis set.
Electrocardiogram (ECG)	All participants, who receive at least one dose of study intervention, have no clinically important protocol deviations or important events affecting ECG, and provide at least one measurable postdose ECG endpoint.

The Bayesian model will be based on the DLT set, all other safety analyses for SMCs on the SAF set. Preliminary PK and Pd analysis for SMCs will also be based on safety analysis set.

The patient profiles will be provided for all participants in the SCR set.

8.2 Subgroup Definition and Parameterization

As is the nature of dose escalation studies, analyses will be performed by dose (and regimen if regimen is changed).

9 General Specifications for Data Analyses

All analyses will be separate for each part (dose escalation in monotherapy (1A) and dose escalation in combination (1B)), and so will decisions by the SMC.

Analyses will be displayed separately by regimen (if different regimens are administered) and dose.

Significance level:

There will be no statistical tests performed in these SMC analyses. If confidence or credibility intervals are to be calculated, these will be two-sided with a confidence probability of 95%, unless otherwise specified in this IAP.

Presentation of continuous and qualitative variables:

Continuous variables will be summarized using descriptive statistics, i.e.

- number of participants, number of participants with non-missing values
- mean, standard deviation
- median, 25th Percentile - 75th Percentile (Q1-Q3)
- minimum, maximum

If there are fewer than 5 observations summarized, only the number of subjects (N), number of subjects with non-missing values, the mean, and the values themselves will be given.

Qualitative variables will be summarized by counts and percentages.

Unless otherwise stated the calculation of proportions will be based on the number of subjects in the analysis set of interest. Therefore, counts of missing observations will be included in the denominator and presented as a separate category.

M6223
MS201430-0001
First in Human Study of M6223
v1.0

Deviations from this definition might apply to the PK analysis (see section **Error! Reference source not found.** for definitions)

Definition of baseline:

In general, the last non-missing measurement prior to the first study treatment administration will be used as the baseline measurement.

If an assessment that is planned to be performed before treatment per protocol is performed on the same day as the start of treatment, respectively, but the assessment time is not available, it will be assumed that it was performed prior and will be considered as baseline.

If a scheduled pre-dose measurement actually occurred post-dose, then the corresponding measurement will be analyzed similar to an unscheduled post-dose measurement.

Definition of change from baseline

Absolute change from baseline = visit value – baseline value

Percent Change from Baseline = 100 * (visit value – baseline value) / baseline value

Software

All analyses (except DLT plot, calculations of posterior distributions, recommended next dose and associated outputs, PK outputs and patient profiles) will be performed using SAS® Software version 9.2 or higher. For the outputs of the Bayesian two-parameter logistic model updates, R (version 3.5.1 or higher [1]) and the R packages bcrm [2] or crmpack [3], or SAS proc MCMC will be used.

Patient profiles will be generated using Xcellerate.

The computer program Phoenix® WinNonlin® version 6.4, or higher (Certara, L.P., 1699 S Hanley Road, St Louis, MO 63144, USA) will be used to derive PK parameters applying Non-compartmental analysis (NCA).

Pharmacodynamic profiles will be visually generated and explored using R (version 3.5.1 or higher [1]). .

9.1 Study Treatment Day

Day 1 is the day of start of study treatment administration (M6223 in Part 1A and M6223 or bintrafusp-alpha in part 1B), the day before is Day -1 (no Day 0 is defined). Study treatment day is defined relative to Day 1.

9.2 Definition of Duration and 'time since' Variables

If not otherwise specified, duration will be calculated by the difference of start and stop date + 1 (e.g. survival time (days) = date of death – date of first study drug administration + 1).

The time since an event (e.g. time since first diagnosis) will be calculated as reference date minus date of event.

9.3 Conversion Factors

The following conversion factors will be used to convert days into months or years: 1 month = 30.4375 days, 1 year = 365.25 days.

9.4 Imputation of Missing Data

Unless otherwise specified (Sections 15 and 16) all data will be evaluated as observed, and no imputation method for missing values will be used.

In all participant data listings, imputed values will be presented, and imputed information will be flagged.

Missing statistics, e.g. when they cannot be calculated, should be presented as "nd".

Where tables are presented over different time points, the total of missing and non-missing observations at each time-point should reflect the population still in the trial at that time. This does not apply when imputations are made beyond trial withdrawal. For example, if a subject is still in the trial at the time-point but with missing data, they should be counted in the number of missing observations.

- Further information after data transfer (such as fatal outcome) might be taken from the Safety database or reported by the investigator at the SMC meeting..

Adverse events	• No imputation will be done for SMC outputs

Dates of study treatment	For SMC outputs, no imputation will be done

10 Study Participants

The subsections in this section include specifications for reporting participant disposition and study treatment/study discontinuations. Additionally, procedures for reporting protocol deviations are provided.

10.1 Disposition of Participants and Discontinuations

A table of disposition (mock shell T_SMCDE.DISP) should display the following (per study part, regimen (if a change occurred), dose level and in total):

- Number of subjects screened (in total only)
- Number of subjects who received at least one dose of study intervention (safety analysis set)
- Number of subjects in preliminary DLT set: All subjects who fulfill the dosing criterion (>=80% of a cohort's dose and at max 5 missed daily dosing) and who completed the DLT period (subjects with any assessment documented on or after day 28) or had DLT per investigator. Due to final DLT decision by SMC this may be different from final DLT Analysis set.
- Number of subjects with a DLT per investigator's judgement as recorded in eCRF
- Number of subjects with treatment ongoing at data extract
- Number of subjects with documented end of treatment (of subjects in safety analysis set) with tabulation of primary reason

A listing of End of treatment reason will be provided (mock L_SMCDE.DISCON).

10.2 Protocol Deviations / Exclusion from Analysis Populations

10.2.1 Important Protocol Deviations

Except for the dosing criterion for the DLT analysis set, this is not applicable for this iAP for SMCs.

10.2.2 Reasons Leading to the Exclusion from an Analysis Population

Subjects who did not receive at least 80% of the planned dose are excluded from the DLT analysis set unless they can be considered in another dose level as described above.

11 Demographics and Other Baseline Characteristics

Not applicable for this iAP, some individual data to appear as part of the patient profiles.

11.1 Demographics

Not applicable for this iAP, some individual data to appear as part of the patient profiles.

11.2 Medical History

Not applicable for this iAP, some individual data to appear as part of the patient profiles

11.3 Other Baseline Characteristics

Not applicable for this iAP, some individual data to appear as part of the patient profiles.

12 Previous or Concomitant Medications/Procedures

Not applicable for this iAP, some individual data to appear as part of the patient profiles.

13 Study Treatment: Compliance and Exposure

Not applicable for this iAP, some individual data to appear as part of the patient profiles and in laboratory line plots.

M6223 First in Human Study of M6223
MS201430-0001 v1.0

14 Efficacy Analyses

Not applicable for this iAP, some individual data to appear as part of the patient profiles

15 Safety Analyses

The subsections in this section include specifications for summarizing safety endpoints that are common across clinical studies such as adverse events, laboratory tests and vital signs.

The primary objective of this study is defining the safety profile, determining the MTD (if available) and MTD. This is the iAP for SMCs therefore safety analyses are the primary analyses.

Safety analyses will be done on the safety analysis population and according to the as-treated principle, except for the Bayesian modelling that will be performed on the DLT analysis set.

15.1 DLT analysis

The primary objective of the SMC meetings is to regularly monitor the overall safety of the subjects enrolled in the trial. An important part of that is reviewing DLTs. Besides individual medical judgement on DLTs, a summary analysis of DLTs is performed using a Bayesian two-parameter logistic modeling approach to model the relation of dose to the occurrence of DLTs. The results from updating this model will assist the SMC in their dosing decisions.
A listing of DLTs as flagged by investigator in eCRF will be provided. It is possible that there is a discrepancy between final SMC decision of SMC and investigator flag in eCRF. Therefore the title of this listing needs to clearly state "DLTs as per investigator flag in eCRF". The listing will be sorted by regimen, dose level and subject ID. This listing will additionally contain age, sex, site of primary tumor, end of treatment date, SOC, PT, investigator term, start date (+treatment day), stop date (+treatment day), relatedness to M6223 (yes/no), to bintrafusp alpha (yes/no) (if applicable, in part 1B only), grade, action taken with M6223, outcome, SAE (yes/no), DLT per investigator (yes/no). See mock table L_SMCDE.AE1
A DLT profile plot of all subjects in the SAF Analysis Set with DLT decisions from previous SMCs will be produced (per regimen). This will show an open square for all subjects who did not have a DLT, a closed square for those who experienced a DLT, and an open circle for those who were excluded from the DLT Analysis Set. This plot will have Cohort Number or subject index on the x-axis and dose level (mg) on the y-axis. See mock figure F_SMCDE.DLTplot

Bayesian two-parameter logistic model

Summary statistics of the posterior probability distribution of the DLT rate (2.5%, 25%, 50%, 75%, and 97.5% quantiles) for each predefined dose level will be updated by estimation according to the logistic model. Using data from all subjects evaluable for DLT or who experienced a DLT at the completion of a new cohort (and data from all previous cohorts: DLT analysis set), the Bayesian logistic model provides a recommended dose level for the next cohort based on minimal loss. This recommendation will be shared with the SMC after the SMC has made the decision upon DLTs. In preparation of the SMC, the model will be updated with potential scenarios' data (e.g. 3 subjects evaluable on current dose level, 0 DLTs, the same with 1 DLT, 2 subjects evaluable 0 DLT etc.…), to have the results ready at SMC.

M6223
MS201430-0001
First in Human Study of M6223
v1.0

Recommendation is based on a loss function (probabilities of being in 1 of the 4 intervals will be multiplied with a loss term as follows:

0 * P(target dosing (estimated DLT rate in (0.2-0.35])) + 1 * P(under dosing (estimated DLT rate in [0-0.2])) +1* P(for over dosing (estimated DLT rate in (0.35-0.6]) + 2 * P(excessive dosing (estimated DLT rate in (0.60-1])).

The recommended dose level resulting from the model for the next cohort is the dose with the lowest loss function. This Bayesian escalation approach will be used to assist the SMC to select the next dose from a predicted set of acceptable doses, i.e., 10 mg, 30 mg, 100 mg, 300 mg, 900mg, 1600mg. It is possible to choose a dose(s) not within the pre-specified dose-escalation plan. In this case the estimated posterior probabilities of the selected dose will also be provided to the SMC.

The model will continue to be updated at each SMC until the SMC has decided to stop dose escalation.

Prior distribution and likelihood are used to calculate the posterior probabilities based on Bayes theorem.

The likelihood is defined based on a binomial distribution, modelling the rate of subjects with at least 1 DLT.

The relationship between dose and toxicity rate is defined by

$$P(DLT|d_j,\alpha,\beta) = \frac{\exp\left(\alpha + \exp(\beta) * \log\left(\frac{d_j}{d_{ref}}\right)\right)}{1 + \exp\left(\alpha + \exp(\beta) * \log\left(\frac{d_j}{d_{ref}}\right)\right)}$$

Where $d_j \in \{10\text{ mg}, 30\text{ mg}, 100\text{ mg}, 300\text{ mg}, 900\text{ mg}, 1600\text{ mg}\}$ (or different) and (α,β) are bivariate normally distributed. Reference dose d_{ref} is 2000 mg. The chosen prior for monotherapy escalation has the following parameters

CCI [redacted]

Handling of changes in regimen

For different regimens, separate models will be used. In general the same specifications outlined above will be used for each of these models, applicable to the dose given on the treatment days, unless otherwise specified. This means that in a every two weeks schedule regimen, 2000 mg as

reference dose refers to the dosing of 2000 mg every two weeks. In a every three weeks schedule the 2000 mg reference dose refers to a dosing of 2000 mg on C1D1, C3D1 etc. The approach is not considering weekly cumulative dosing.

Rationale: It is assumed that changes in regimen occur due to the fact that the observed dose-toxicity relationship was not as a priori expected and the change should lead to coming closer to the a priori expected dose-toxicity relationship.

15.2 Patient profiles

The SMC will receive patient profiles containing:

- Subject disposition (still in trial, or withdrawn with reason for withdrawal)
- Demographics and baseline characteristics (e.g., cancer diagnosis, staging)
- Medical history
- History of disease under study
- Previous and concomitant medications
- Prior anti-cancer drug therapies
- Prior anti-cancer radiotherapy
- Concomitant procedures
- Study drug administration, and dose adjustments
- All serious and non-serious AEs (with details like e.g. grade, start and stop date), including but not limited to:
  - DLTs
  - AESIs
  - AEs leading to dose reduction or temporary discontinuation
  - AEs leading to permanent treatment discontinuation
  - AEs leading to death
- Laboratory data (hematology, coagulation, biochemistry, urinalysis)
- ECG results
- Vital signs

Patient profiles will be provided for the current cohorts under review and updated patient profiles will be provided for subjects from previous cohorts that had changes from last SMC meeting.

15.3 Adverse Events

The severity of adverse events will be graded using the NCI-CTCAE version 5.0, except where CTCAE grades are missing. No imputation of missing grades will be performed. Adverse events will be coded according to the latest MedDRA version at the time of the data cut-off.

- **TEAEs**: Any AEs that are reported (serious and non-serious) will be considered treatment emergent adverse events (TEAEs), with the exception of those that started prior to the first dose of study treatment (unless a worsening of the event is recorded after the first dosing, in which case the event will be counted as a TEAE), or AEs starting more than 30 days after the last dose of study treatment.
- **Related Adverse Events**: AEs with relationship missing, unknown or yes. Relationship is judged separately for M6223, and Bintrafusp-alpha.
- **SAEs**: Serious adverse events (as recorded on the AE eCRF page, Serious Adverse Event = 'Yes').
- **AEs Leading to Treatment Discontinuation**: AEs leading to permanent discontinuation of study treatment (as recorded on the AE eCRF page, Action taken with study treatment = 'Drug withdrawn'). Treatment discontinuation is recorded separately for M6233, and dexamethasone.
- **AEs Leading to Death**: AEs leading to death (as recorded on the AE eCRF page, Outcome = 'Fatal').
- **Adverse Events of Special Interest**: AEs for which the investigator ticked "YES" on: "Is this an adverse event of special interest"

AEs will be summarized by MedDRA PT as event category and MedDRA primary SOC as summary category. In general, each subject will be counted only once within each PT or SOC. If a participant experiences more than one AE within a SOC or PT for the same summary period, only the AE with the strongest relationship or the worst severity, as appropriate, will be included in the summaries of relationship and severity

AEs with missing classifications regarding relationship to study treatment, and those with start date on or after the start of study treatment, will be considered as related to the study treatment

Treatment-emergent adverse events (TEAE) are those events with onset dates occurring within the on-treatment periods from start of treatment up to (including) 30 days after end of treatment.

.

15.3.1 All Adverse Events

Adverse events will be summarized by worst severity (according to NCI-CTCAE version 5.0)] per participant, using the latest version of MedDRA preferred term (PT) as event category and MedDRA primary system organ class (SOC) as body system category.

In case a participant had events with missing and non-missing grades, the maximum of the non-missing grades will be displayed. No imputation of missing grades will be performed.

Incomplete AE-related dates will be not be imputed for SMC outputs.

AE tables will be restricted to TEAEs only unless otherwise specified. The AE tables will include the number and percentage of subjects with at least one TEAE, by MedDRA SOC and PT (both sorted alphabetically), unless otherwise stated.
All AEs will be tabulated by regimen (if different regimens are tested), dose level and overall in a table showing the incidence of TEAEs, by SOC and PT (both sorted alphabetically). See mock table T_SMCDE.TEAEGR

A listing of all TEAEs will be provided. This listing will be sorted by regimen, dose level and subject. This listing will additionally contain age, sex, site of primary tumor, end of treatment date, SOC, PT, investigator term, start date (+treatment day), stop date (+treatment day), relatedness to M6223 (yes/no), to bintrafusp alpha (yes/no) (if applicable, in part 1B only), grade, action taken with M6223, outcome, SAE (yes/no), DLT per investigator (yes/no). See mock table L_SMCDE.AE1

Additionally, a listing of all adverse events grade >=3 (incl non TEAEs) and SAEs (also including non-treatment emergent SAEs) will be provided: This listing is to be produced twice, once sorted by regimen, dose level and subject ID and once sorted by SOC and PT. These will also include age, sex, site of primary tumor, end of treatment date, SOC, PT, investigator term, start date (+treatment day), stop date (+treatment day), relatedness to M6223 (yes/no), to bintrafusp alpha (yes/no) (if applicable, in part 1B only), grade, action taken with M6223, outcome, SAE (yes/no), DLT per investigator (yes/no). See mock table L_SMCDE.AE1 and L_SMCDE.AE2

15.3.2 Adverse Events Leading to Discontinuation of Study Treatment

A listing of TEAEs leading to treatment discontinuation, interruption, or dose reduction of each study drug will be provided (one for M6223, additionally in part 1B: one for bintrafusp alpha). This listing, sorted by regimen, dose level and subject ID, will also include age, sex, site of primary tumor, end of treatment date, SOC, PT, investigator term, start date (+treatment day), stop date (+treatment day), relatedness to M6223 (yes/no), to bintrafusp alpha (yes/no) (if applicable, in part

1B only), grade, action taken with M6223, outcome, SAE (yes/no), DLT per investigator (yes/no). See mock table L_SMCDE.AE1

15.4 Deaths, Other Serious Adverse Events, and Other Significant Adverse Events

15.4.1 Deaths

Not applicable for this iAP, some individual data to appear as part of the patient profiles.

15.4.2 Serious Adverse Events

SAEs are shown in the patient profiles and part of the listing of all adverse events grade >=3 (incl non TEAEs) and SAEs mentioned in 15.3.1

15.4.3 Other Significant Adverse Events

A listing of AEs of special interest will be provided. AEs for which the investigator ticked "YES" on: "Is this an adverse event of special interest" will be considered AESIs.
This listing, sorted by regimen, dose level and subject ID, will also include age, sex, site of primary tumor, end of treatment date, SOC, PT, investigator term, start date (+treatment day), stop date (+treatment day), relatedness to M6223 (yes/no), to bintrafusp alpha (yes/no) (if applicable, in part 1B only), grade, action taken with M6223, outcome, SAE (yes/no), DLT per investigator (yes/no). See mock table L_SMCDE.AE1

15.5 Clinical Laboratory Evaluation

Laboratory values (including corresponding normal ranges) from the Lab will be used for patient profiles, an eDISH plot, and line plots

Laboratory results will be classified according to the latest NCI-CTC Version [at the moment 5.0]. Some of the toxicity gradings are based on laboratory measurements in conjunction with clinical findings. The classification will be derived from the laboratory results at a given assessment, thus ignoring additional clinical findings, except for the evaluation of blood glucose toxicity grade 1 and 2 the fasting state is required. Ignoring the fasting state might lead to overreporting of grade 1 and 2 events. Therefore, blood glucose grading will focus on grade 3 and 4 reporting only.

If it increases interpretability, graphical displays can be shown on a log-scale.

Line plots (with x-axis time, y-axis lab value), using different colors per regimen, dose level and different line types to identify participants, will be provided for the following lab parameters:

- Neutrophils
- Platelets
- Lymphocytes
- Hemoglobin

These line plots should also depict the on-treatment time of the participants. Inthis study, the patients are receiving infusions on day1 of 2 weeks cycles. If a participant received an infusion, the following 28 days (including the dosing day) are to be considered "on treatment".
If available (and feasible due to potential differences between labs) reference lines for ULN and LLN should be added to the plots. See mock shell F_SMCDE.LBLine

Hepatotoxicity assessment

A plot of peak ALT versus peak total bilirubin, both relative to the upper limit of normal (ULN) will be provided. This eDISH plot (evaluation of drug-induced serious hepatotoxicity) will have reference lines at 3×ULN for ALT and at 2×ULN for total bilirubin. Subjects outside the lower left box will be identified by subject-ID. See mock shell P_SMCEDISH.

15.6 Vital Signs

Not applicable for this iAP, except for display of individual values in patient profiles

15.7 Other Safety or Tolerability Evaluations

Not applicable for this iAP, except for display of individual values in patient profiles

16 Analyses of Other Endpoints

16.1 Pharmacokinetics

Summary NCA pharmacokinetic parameters such as Cmax, AUCtau, AUCss, and CL will be presented for all patients as permitted by the available data. PK profiles will be graphically represented as line plots on linear and logarithmic scale. Serum concentrations for each cohort will be compared to model predicted levels to assess whether the systemic exposure falls within expected range. Nominal time will be used for estimation of all PK parameters for SMC analyses. Actual times will be used for the final analysis.

16.2 Pharmacodynamics

The pharmacodynamics marker analyzed preliminary for the SMCs is TIGIT target occupancy. This is measured on C1D1 predose, C1D2, C2D1 and C4D1. The results will be presented as % TIGIT target occupancy (ratio to baseline).

Line plots will show the TIGIT target occupancy versus time. Each participant in a cohort will be differentiated from the others by using a participant-specific symbol. Data from the first two single-patient cohorts will be illustrated on a single plot. TIGIT target occupancy data from all subsequent cohorts will be individually plotted.

Dot plots will show TIGIT target occupancy data of individual participants versus dose (x-axis dose, y-axis % TIGIT target occupancy), separate plots for C1D2, C2D1 and C4D1. The plot for C2D1 is to be produced, the plots for C1D2, and C4D1 may be produced.

Dot plot will show TIGIT occupancy of individual participants versus drug concentration at the same timepoint (x-axis exposure, y-axis % TIGIT target occupancy). This plot may be produced as pharmacodynamics data become available.

17 References

1 R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

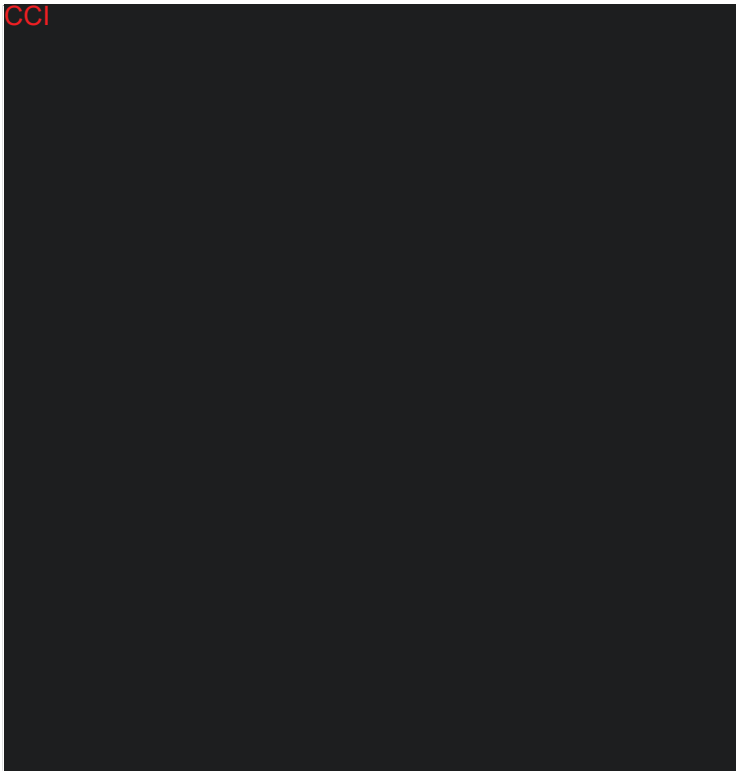
2 Sweeting M, Mander A, Sabin T. (2013) bcrm: Bayesian Continual Reassessment Method Designs for Phase I Dose-Finding Trials. *J Statist Software:* 54(13):1-26. http://www.jstatsoft.org/article/view/v054i13 URL

3 Sabanés Bové, D., Yeung, W., Palermo, G., & Jaki, T. (2019). Model-Based Dose Escalation Designs in R with crmPack. *Journal of Statistical Software, 89*(10), 1 - 22. doi: http://dx.doi.org/10.18637/jss.v089.i10

M6223
MS201430-0001
First in Human Study of M6223
v1.0

18 Appendices

CCI



CCI

CCI

Document Object No. CCI

CCI

Document Object No CCI

CCI

CCI

CCI

[redacted]

CCI

CCI

[redacted]

CCI

CCI

CCI