Integrated Analysis Plan

Clinical Study Protocol Identification No.	MS200647_0020
Title	A Phase II, Multicenter, Open Label Study of Bintrafusp alfa (M7824) Monotherapy in Participants with HMGA2-expressing Triple Negative Breast Cancer
Study Phase	Phase II
Investigational Medicinal Product(s)	Bintrafusp alfa (M7824)
Clinical Study Protocol Version	11 December 2019/Version 1.0
Integrated Analysis Plan Author	**Coordinating Authors** PPD PPD
Integrated Analysis Plan Date and Version	21 December 2021/Version 2
Integrated Analysis Plan Reviewers	

Function	Name
PPD Merck Healthcare	PPD
PPD PPD	
PPD PPD	
PPD , PPD	
PPD , EMD Serono	
PPD EMD Serono	
PPD , Merck	
PPD Merck	
PPD Merck	
PPD PPD	
PPD Merck	
PPD PPD	
PPD PPD	
PPD Merck	
PPD , EMD Serono	

Confidential
This document is the property of Merck KGaA, Darmstadt, Germany, or one of its affiliated companies.
It is intended for restricted use only and may not - in full or part - be passed on, reproduced, published or used without express permission of Merck KGaA, Darmstadt, Germany or its affiliate.
Copyright © 2020 by Merck KGaA, Darmstadt, Germany or its affiliate. All rights reserved.

CCI
CCI

Approval Page

Integrated Analysis Plan: MS200647_0020

A Phase II, Multicenter, Open Label Study of Bintrafusp alfa (M7824) Monotherapy in Participants with HMGA2 expressing Triple Negative Breast Cancer

Approval of the IAP by all Merck Data Analysis Responsible has to be documented within BREEZE VEEVA VAULT system 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.

CCI
CCI

1 Table of Contents

CCI
CCI

2 List of Abbreviations and Definition of Terms

ADA	Anti-Drug Antibodies
AE	Adverse Event
AESI	Adverse Event of Special Interest
AIDS	Acquired Immune Deficiency Syndrome
ALP	Alkaline Phosphatase
ALT	Alanine Aminotransferase
aPTT	Activated Partial Thromboplastin Time
AST	Aspartate Aminotransferase
ATC	Anatomical Therapeutic Chemical classification
BM	Biomarker Analysis Set
BMI	Body Mass Index
BOR	Best Overall Response
CDISC	Clinical Data Interchange Standards Consortium
CI	Confidence Interval
CR	Complete Response
CSR	Clinical Study Report
CT	Computed Tomography
DCR	Disease Control Rate
DI	Dose Intensity
DOR	Duration of Response
DRR	Durable Response Rate
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic Case Report Form
eDISH	Evaluation of Drug-Induced Serious Hepatotoxicity
EOT	End of Treatment
EudraCT	European Union Drug Regulating Authorities Clinical Trials
FAS	Full Analysis Set
HIV	Human Immunodeficiency Virus
HMGA2	High Mobility Group AT-Hook 2

CCI
CCI

HR	Hazard Ratio
HRQOL	Health-Related Quality of Life
IAP	Integrated Analysis Plan
ICH	International Conference on Harmonization
IMM	Immunogenicity Analysis Set
IPD	Important protocol deviation
ir	Immune-Related
IRC	Independent Review Committee
IRR	Infusion-Related Reaction
KM	Kaplan-Meier
Max	Maximum
Med	Median
MedDRA	Medical Dictionary for Regulatory Activities
MH	Mantel-Haenszel
Min	Minimum
MRI	Magnetic Resonance Imaging
NA	Not Applicable
NCI-CTCAE	National Cancer Institute – Common Terminology Criteria for Adverse Events
nd	Not Determined
NE	Not Evaluable
NR	No Result
ORR	Objective Response Rate
OS	Overall Survival
PD	Progressive Disease or Protocol Deviation
PD-L1	Programmed Death-Ligand 1
PFS	Progression Free Survival
PK	Pharmacokinetics
PKAS	Pharmacokinetic Analysis Set
PR	Partial Response
PT	Preferred Term
Q1	25th Percentile
Q3	75th Percentile

RDI	Relative Dose Intensity
RECIST	Response Evaluation Criteria in Solid Tumors
SAE	Serious Adverse Event
SAF	Safety Analysis Set
SCR	Screening Analysis Set
SD	Stable Disease
SDTM	Study Data Tabulation Model
SMQ	Standardised MedDRA Queries
SOC	System Organ Class
StD	Standard Deviation
TEAE	Treatment-Emergent Adverse Event
TGFβ	Transforming Growth Factor Beta
TLF	Tables, Listings, and Figures
TMB	Tumor Mutational Burden
TNBC	Triple Negative Breast Cancer
TNM	Tumor Node Metastasis
ULN	Upper Limit of Normal
UNK	Unknown
WHO-DD	World Health Organization Drug Dictionary

CCI
CCI

3 Modification History

Unique Identifier for Version	Date of IAP Version	Author	Changes from the Previous Version
1.0	03 August 2020	PPD	Not applicable.
2.0	18 November 2021	PPD	Replaced "TGFβ-mediated" with "TGFβ inhibition mediated" across the whole document. Added section 4.1 about early discontinuation of the study. Added COVID-related wording in section 7.1. Table 2: removed BM population. Section 8.2: added details for race subgroup analyses. Section 12.3: replaced "Concomitant Medication Details" with "Concomitant Procedures Details". Medication listing: removed "medication type" and added "regimen name" as per data collected in the latest eCRF. Section 15.2.3.4: added a bleeding adverse events listing. Section 15.2.3.5: updated terms to identify anemias. Table 12: updated directionality of Glucose and added Lactate dehydrogenase. Table 13: updated parameter code for Urea Nitrogen and Prothrombin Time/Standard Prothrombin Time. Section 15.4: removed "oxygen saturation" and "timepoint" from listing. Section 16.4: updated analysis population from BM to FAS. Appendix 1: updated to v2. Appendix 2: updated to v2.

4 Purpose of the Integrated Analysis Plan

The purpose of this IAP is to document technical and detailed specifications for the interim, primary and final analysis of data collected for protocol MS200647_0020. 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 study protocol and is prepared in compliance with ICH E9. It describes analyses planned in the protocol.

4.1 Early Discontinuation of the Study

Due to the decision to early discontinue the study, not all the analyses described in this IAP will be provided. Especially, no analyses for efficacy, pharmacokinetics. immunogenicity, or biomarkers will be performed. Due to the low number of participant listings will be delivered for all other endpoints instead of summary tables as deemed appropriate and will be included in an abbreviated CSR.

5 Objectives and Endpoints

Table 1 Study Objectives and Endpoints

Objectives	Endpoints (Outcome Measures)	IAP section
Primary		
To evaluate clinical efficacy of bintrafusp alfa in participants with TNBC with high HMGA2 expression, based on ORR	Confirmed objective response according to RECIST 1.1 assessed by an IRC	14.1
Secondary		
To evaluate clinical efficacy of bintrafusp alfa based on DOR	DOR according to RECIST 1.1 assessed by an IRC	14.6
To evaluate clinical efficacy of bintrafusp alfa based on DRR	Durable response of at least 6 months assessed by an IRC	14.6
To evaluate clinical efficacy based on PFS	PFS according to RECIST 1.1 assessed by an IRC	14.4
To evaluate ORR, DOR, DRR, and PFS by Investigator read	Objective response, DOR, Durable response, and PFS according to RECIST 1.1 as assessed by the Investigator	14.2 14.6 14.4
To evaluate clinical efficacy based on OS	OS	14.7
To evaluate clinical safety of bintrafusp alfa	Occurrence of TEAEs and treatment-related AEs including AEs of special interest	15.1 15.2
To characterize the PK profile of Bintrafusp alfa	• The concentration observed immediately at the end of infusion (C_{EOI}) of bintrafusp alfa • The concentration observed immediately before next dosing (corresponding to predose or trough concentration [C_{trough}] for multiple dosing) of bintrafusp alfa	16.1 16.2
To characterize the immunogenicity of bintrafusp alfa	Immunogenicity of bintrafusp alfa as measured by ADA assay from Screening through Safety Follow-up Visit up to 28 days after last treatment	16.3
CCI [redacted]		
ADA=antidrug antibody, AE=adverse event, DOR=duration of response, DRR=durable response rate, HMGA2=High Mobility Group AT-Hook 2, ir=immune-related, IRC=Independent Review Committee, ORR=Objective response rate, OS=overall survival, PD-L1=programmed death-ligand 1, PFS=progression-free survival, PK=pharmacokinetic, RECIST 1.1=Response Evaluation Criteria in Solid Tumors Version 1.1, TEAE=Treatment-emergent adverse event, TMB=tumor mutational burden, TNBC=triple negative breast cancer.		

6 Overview of Planned Analyses

This IAP addresses the interim, primary and final analyses. If the study continues after the cut-off date for final analysis, subsequent analyses may be performed but are not described in this IAP.

Statistical analyses will be performed on the basis of CDISC SDTM data. These SDTM data contain as clean as possible electronic Case Report Forms (eCRF) data as well as external data including laboratory data, biomarker data and tumor assessment results by the Independent Review Committee (IRC). All data used in the analysis will be included up to a data cut-off point defined in the following subsections.

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 Interim Analysis

The interim analysis on unconfirmed response as assessed by the Investigator will be performed 4 months after the accrual of the first 15 participants. CCI [redacted]

For this interim analysis all data will be included up to a clinical cut-off date corresponding to 4 months after the accrual of the first 15 participants.

6.2 Primary Analysis

This analysis will be the main analysis and all planned analyses identified in the Clinical Study Protocol and in this IAP will be performed. All data will be included up to a clinical cut-off date corresponding to 8 months after the accrual of the last participant.

6.3 Final Analysis

An extract of the primary analyses as specified in the TLF table of content will be done in the final follow-up analysis for evaluation of OS. All data will be included up to a clinical cut-off date corresponding to 24 months after accrual of last participant.

7 Changes to the Planned Analyses in the Clinical Study Protocol

7.1 COVID-19 Impact

No changes to the planned analysis of the efficacy endpoints will be performed due to the impact of COVID-19 outbreak.

Instead, given the low number of subjects included in the study, additional tags and labels will be used to identify Protocol Deviations (PDs) and AEs related with COVID in listings. See sections 10.2 and 15.1.1 for more details.

8 Analysis Populations and Subgroups

8.1 Definition of Analysis Populations

Prescreening Analysis Set (PRESCR)

The Prescreening analysis set population includes all participants, who provided prescreening informed consent.

Screening Analysis Set (SCR)

The Screening analysis set population includes all participants, who provided informed consent, regardless of the participant's study intervention status in the study.

Full Analysis Set (FAS)

The Full analysis set will include all participants, who were administered at least one infusion of bintrafusp alfa.

Safety Analysis Set (SAF)

The Safety analysis set will include all participants, who were administered at least one infusion of bintrafusp alfa.

Pharmacokinetic Analysis Set (PKAS)

The Pharmacokinetic analysis set will include all participants who complete at least 1 infusion of bintrafusp alfa, and who provide at least one sample with a measurable concentration of bintrafusp alfa.

Immunogenicity Analysis Set (IMM)

The Immunogenicity analysis set will include all participants, who were administered at least one infusion of bintrafusp alfa and have at least one valid ADA result.

Table 2 displays the use of each analysis set in the different analysis.

Table 2 Overview of the Analysis Set Used in for each Analysis

Analysis	PRESCR	SCR	FAS	SAF	PKAS	IMM
Disposition	✓	✓				
Demographics			✓			
Baseline Assessments			✓			

CCI [redacted] CCI

Analysis	PRESCR	SCR	FAS	SAF	PKAS	IMM
Previous and Concomitant Therapies			✓			
Compliance and Exposure				✓		
Efficacy			✓			
Safety and Tolerability				✓		
Pharmacokinetics					✓	
Patient-Reported Outcome			✓			
Immunogenicity						✓
CCI [redacted]			✓			

PRESCR=prescreening, SCR=screening analysis set, FAS=full analysis set, SAF=safety analysis set, PKAS=Pharmacokinetics analysis set, IMM=immunogenicity analysis set.

8.2 Subgroup Definition and Parameterization

Subgroup analyses may be performed on primary and key secondary efficacy endpoints as defined below. All subgroup analyses will be exploratory and no adjustment for multiplicity will be performed.

For the definition of subgroup level, data as documented in the electronic case report form (eCRF) will be taken. The category "missing" will not be included in any subgroup analysis.

In case of low number of participants within a category (<5 participants), categories will be pooled when meaningful. Subgroup analyses will be omitted where pooling would result in only one remaining category.

The following subgroups will be defined:

- Age: < 65 years (reference level) / ≥ 65 years
- Number of prior treatment lines (neoadjuvant, adjuvant or metastatic/locally advanced prior anti-cancer drug therapies): 0 (reference level) / 1 / 2 / 3 / ≥ 4
- ADA status: Ever positive (reference level) / Never positive
- PD-L1 expression at baseline: negative (reference level) / positive / not evaluable
- Race: White (reference level) / Asian / Black / others **
- ECOG at day 1 PS 0 (reference level)/ 1 / >1
- Prior treatment: Taxane yes (reference level) / no
- Prior treatment: Anthracycline yes (reference level) / no
- Prior treatment: Radiotherapy yes (reference level) / no
- Prior treatment: Neoadjuvant or adjuvant yes (reference level)/ no
- Number of metastatic site: 0-3 (reference level)/ >=4
- Site of metastatic disease Lung No (reference level) / yes
- Site of metastatic disease Liver No (reference level) / yes

- Site of metastatic disease Bone No (reference level) / yes
- Site of metastatic disease Brain No (reference level) / yes

** In case of race subgroups, the following categorization of the data collected in the eCRF will be applied:

- White: "White"
- Asian: "Asian"
- Black: "Black or African American"
- Others: "American Indian or Alaska Native", "Native Hawaiian or Other Pacific Islander", "Other", "More than one race"
- Note that records marked as "Not collected at this site" in eCRF will be treated as missing race

9 General Specifications for Data Analyses

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

The "start date" for this study is the start date of treatment.

Continuous (non-PK) variables will be summarized using descriptive statistics, i.e., the number of participants with non-missing values (n), arithmetic mean (Mean), standard deviation (StD), median (Med), 25th percentile (Q1) and 75th percentile (Q3), minimum (Min), and maximum (Max).

Qualitative variables will be summarized by frequency counts and percentages. Unless otherwise stated the calculation of proportions will be based on the number of participants of the analysis set of interest. Therefore, counts of missing observations will be included in the denominator and presented as a separate category.

Descriptive statistics by nominal visit or time point, will include only data from scheduled visits, unless stated otherwise. Unscheduled visits will be included in the derivation of baseline or worst on-treatment values.

The overall significance level is 2.5% one-sided. If confidence intervals (CIs) are to be calculated, these will be two-sided with a confidence probability of 95%, unless otherwise specified in this IAP.

In order to provide overall estimates of treatment effects, data will be pooled across sites. The "site" factor will not be considered in statistical models or for subgroup analyses due to the high number of participating sites in contrast to the anticipated small number of participants treated at each site.

Pharmacokinetics (PK) variables (concentrations and parameters) will also be summarized as described in Section 16.1. Refer also to Section 16.1 for PK data handling/analysis details.

All statistical analyses will be performed using SAS® Software Version 9.4 or higher in the SAS Grid environment (Statistical Analysis System, SAS-Institute, Cary NC, USA). The computer program Phoenix® WinNonlin® Version 8.0, or higher (Certara, L.P., Princeton, New Jersey, USA) could be used for PK data.

9.1 Data Handling After Cut-off Date

By its nature, data after cut-off may be incomplete and subject to further change and will not be used for summary statistics, statistical analyses, listings or imputations.

Stop dates are not affected by this rule, e.g. a stop date of AEs, which starts prior to the cut-off, but stops after date of cut-off, will not be changed.

9.2 Definition of Baseline and Change from Baseline

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

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 assessment time is collected, the observed time as well as time of first dose will be used to determine pre-dose on study day 1 for baseline calculation.

Unscheduled assessments will be used in the determination of baseline. However, if time is missing, an unscheduled assessment on Study Day 1 will be considered to have been obtained after study intervention.

Absolute and percent changes from baseline are defined as:

$$Absolute\ change\ (unit) = Visit\ value - Baseline\ value$$

$$Percent\ change\ (\%) = 100 \times \frac{(Visit\ value - Baseline\ value)}{Baseline\ value}$$

9.3 Study Day / Study Treatment Day

Day 1 is the day of start of study treatment, the day before is Day -1 (no Day 0 is defined). Study treatment day is defined relative to Day 1.

9.4 Definition of Duration and 'time since' Variables

Durations in days will be calculated by the difference of start and stop date + 1 (e.g. survival time (days) = date of death – date of first study intervention + 1) if not otherwise specified.

CCI [redacted] CCI

Date of first study intervention is defined as the date of the first administration of bintrafusp alfa during the study.

The time since an event (e.g. time since first diagnosis) will be calculated as reference date minus date of event. In general, the reference date will be the date of first study intervention.

The time to an event will be calculated by the difference between the time of event and the reference date + 1 if not otherwise specified. For example, survival time (days) = date of death - date of first study intervention + 1.

9.5 Conversion Factors

The following conversion factors will be used to convert days into months or years:

1 week = 7 days, 1 month = 30.4375 days, 1 year = 365.25 days

1 inch = 0.0254 meters

9.6 Date of last Contact

The date of last contact will be derived for participants not known to have died at the analysis cut-off using the latest complete date prior to or at the data cut-off date among the following:

- All participant assessment dates (blood draws [laboratory, PK], vital signs, performance status, electrocardiogram [ECG], tumor assessments)
- Start and end dates of anti-cancer therapies administered after study treatment discontinuation
- AE start and end dates
- Last known to be alive date collected on the 'Subject Status / Survival Follow-Up' eCRF page and status = "Alive"
- Study drug start and end dates (including treatment reinitiation)
- Date of discontinuation on disposition eCRF pages (do not use if reason for discontinuation is lost to follow-up)

Only dates associated with actual examinations of the participant reported in the eCRF will be used in the derivation. Dates associated with a technical operation unrelated to participant status such as the date a blood sample was processed will not be used. Assessment dates after the cut-off date will not be applied to derive the last contact date.

Data collected after reinitiated treatment will be considered in the derivation of the last known to be alive date.

9.7 Time Window

Not applicable.

CCI
CCI

9.8 Definition of On-treatment Period

The on-treatment period is defined as the time from the first study intervention administration to the last study intervention administration date + 30 days, or the cut-off date or death or the earliest date of subsequent anticancer drug therapy minus 1 day, whichever occurs first, unless otherwise stated.

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.

Any systemic anticancer therapy, any anticancer surgery and any anticancer radiotherapy as documented in the 'Anti-cancer treatment after discontinuation details', 'Radiotherapy after discontinuation' and 'Surgery after discontinuation' eCRF pages will be considered as subsequent anticancer therapy.

The on-treatment period will include the initial treatment period as well as the reinitiation of treatment period, as applicable. Whether the participant reinitiates treatment (following the rules as outlined in the protocol) or not, the on-treatment period is defined as the time from the first study intervention administration to the last study intervention administration date + 30 days or the earliest date of subsequent anticancer drug (anticancer therapy, anticancer surgery and anticancer radiotherapy) therapy minus 1 day, whichever occurs first, unless otherwise stated.

For immune-related AEs as listed in section 15.2.3.2, an expanded on-treatment period will be used as a default for any analysis, defined as follows:

Time from the first study intervention to the last study intervention date + 90 days, death OR to the earliest date of subsequent anticancer therapy minus 1 day, whichever occurs first, unless otherwise stated.

9.9 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 'not determined'. For example, if n = 1, the measure of variability (StD) cannot be computed and should be presented as 'nd'.

The following table for imputation rules for incomplete dates is to be considered:

Table 3 Imputation Rules for Incomplete Dates

Item	Imputation rules
Age calculation	Incomplete dates (date of informed consent, date of birth) for the calculation of age will be imputed as follows: • In case of missing day for at least one date, but month and year available for both dates: the day of informed consent and the day of birth will be set to 1. • In case of missing month for at least one date, but year available for both dates, the day and the month of informed consent and the day and month of birth will be set to 1. • In all other cases, the incomplete dates will not be imputed.
Disease history	Incomplete dates for disease history (e.g. initial diagnosis date, date of documented, locally advanced, inoperable or metastatic disease diagnosis, date of response or progression in prior treatment) will be imputed as follows: • If the day is missing, it will be imputed to the 1st day of the month. • If both day and month are missing, the month and day will be imputed as January 1st. • If the date is completely missing, no imputation will be performed.
Adverse events	Incomplete AE-related dates will be imputed as follows: • In case the onset date is missing completely or missing partially but the onset month and year, or the onset year are equal to the start of study treatment then the onset date will be imputed by the minimum of start of study treatment and AE resolution date (if not missing). • In all other cases, the missing onset day or missing onset month will be imputed by 1. • Incomplete stop dates will be imputed 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.

For identification of previous or concomitant medications/procedures, rules presented in Table 4 and Table 5 below will be used to define if a medication/procedure is considered as a previous, concomitant or both previous and concomitant medication/procedure.

Table 4 Stopping rules for medication/procedure end dates

End date of medication/procedure			Stopping 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

CCI
CCI

End date of medication/procedure			Stopping rule
Day	Month	Year	
< Treatment start (complete date)			Before treatment start
>= Treatment start (complete date)			After treatment start

UNK = Unknown

Table 5 Rules to define previous and/or concomitant medication/procedure

Start date of medication/procedure			Stopping rule (see Table 4)	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

Other dates imputation rules are defined in Table 6 below:

Table 6 Other Dates Imputation Rules

Item	Imputation rules
Dates of study treatment	Start date of study treatments: • No imputation will be done. End date of study treatments:

CCI
CCI

Item	Imputation rules
	• In case the last date of study drug is missing or incomplete the date of last administration of study drug will be taken from the treatment termination eCRF pages. • If the last date of study drug is completely missing and there is no End of Treatment eCRF page and no death date the participant should be considered to be ongoing and use the cut-off date for the analysis as the last dosing date. • If the last date of study drug is completely or partially missing and there is EITHER an End of Treatment eCRF page OR a death date available (within the cut-off date) then imputed last dose date is: = 31DECYYYY, if only Year is available and Year < Year of min (EOT date, death date) = Last day of the month, if both Year and Month are available and Year = Year of min (EOT date, death date) and Month < the month of min (EOT date, death date) = min (EOT date, death date), for all other case
Death date	For the purpose of survival analyses (PFS and OS) 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. In listings, the imputed date of death will be presented with a flag indicating the level of imputation.
Tumor assessments	All investigation dates (MRI scan, CT scan) must be completed with day, month and year. If there are multiple scan dates associated with an evaluation, i.e., radiological assessments occur over a series of days rather than the same day, the choice of date of assessment could impact the date of progression and/or date of response. If there are multiple scan dates associated with an evaluation, the earliest of the scan dates associated with the evaluation will be used as the date of assessment. If one or more investigation dates for an evaluation are incomplete but other investigation dates are available, the incomplete date(s) are not considered for calculation of the assessment date and assessment date is calculated as the earliest of all investigation dates (e.g. X-ray, CT scan). If all measurement dates for an evaluation have no day recorded, the 1st of the month is used. If the month is not completed, for any of the investigations for an evaluation, the respective assessment will be considered to be at the date which is exactly between the previous and the following assessment. If both a previous and following assessments are not available, this assessment will not be used for any calculations.
Dates of subsequent anti-cancer therapy	Incomplete dates for start date of subsequent anti-cancer therapy (drug therapy, radiation, surgery) will be imputed as follows and will be used for determining censoring dates for efficacy analyses and in the derivation of the end of on-treatment period. • If only day is missing, it will be imputed as the last day of the month unless the end date of subsequent anticancer therapy is before that date.

CCI
CCI [redacted]

Item	Imputation rules
	In that case, the incomplete anticancer therapy start date will be imputed as the end date of the anticancer therapy. • If both day and month are missing, no imputation will be performed. Incomplete subsequent anti-cancer therapy stop dates will not be imputed.

CT=computed tomography, eCRF=electronic case report form, EOT=end of treatment, min=minimum, MRI=magnetic resonance imaging, OS=overall survival, PFS=progression-free survival.

9.10 Scoring of HRQOL Data

Not applicable.

9.11 Further Specifications

Table 7 Further specifications

Re-screened participants	Re-screened participants will be only counted once in the screening analysis set, considering the latest screening (screening with latest informed consent).
Data collected after reinitiated treatment	Data collected after reinitiation of treatment will be included in the summary statistics. A data listing will include AE data for all reinitiated patients.
Preferred term for analysis of WHO-DD coded data	For data coded according to WHO Drug B3 (e.g., concomitant medications), summaries will be done on the preferred term level where the preferred term is corresponding to codes ending in 01001. With this approach, variations of salt forms of active ingredients will be analyzed under the same term, e.g. diphenhydramine and diphenhydramine hydrochloride will be analyzed as the same preferred term diphenhydramine.
Categorization of participant for COVID-19 impact assessment	For the assessment of COVID-19 impact on this study, participants will be categorized as being potentially affected by COVID-19 based on the COVID-19 pandemic start date, defined as the minimum of the first COVID-19 death date per country and 11 March 2020 (WHO-start of world-wide pandemic). First death from COVID-19 occurred per country is determined according to the published data by European Centre for Disease Prevention and Control (status of 26th June 2020).

WHO-DD: World health organization drug dictionary.

9.11.1 Unscheduled Assessments

As per database definition, the safety unscheduled assessments are always linked to a scheduled timepoint (each unscheduled assessment is linked to the previous scheduled timepoint). Safety data retrieved from an unscheduled timepoint (vital signs, ECG and laboratory data) will be analyzed according to the following scenario:

- For shift table, they will be taken into account in the definition of the worst assessment during study.
- For description at each timepoint post-baseline, the first available result (in chronological order) per timepoint will be taken into account in the analysis in case of multiple values.
- For description at baseline, the last available result before first study intervention will be taken into account in the analysis in case of multiple values.

CCI [redacted] CCI

For immunogenicity analysis, unscheduled visits will also be taken into account in the analysis following the same rules as detailed above.

For PK analysis, unscheduled samples will not be linked to a scheduled timepoint and will be excluded from summaries.

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

The number and percentage of participants in each of the below disposition categories will be presented based on the data collected in the eCRF. Percentages will be presented with respect to the number of treated participants.

- Total number of prescreened participants
- Number of participants who did not continue beyond prescreening (overall and by reason)
- Total number of participants screened (i.e. participants who gave informed consent)
- Number of re-screened participants
- Number of participants who discontinued from the study prior to treatment overall and grouped by the main reason (e.g. the failed specific inclusion or exclusion criteria, progressive disease, adverse event, lost to follow-up, death, withdrawal of consent and other)
- Number of participants who received at least one dose of study intervention (FAS/SAF analysis set)

The end of treatment status will be summarized by:

- Number and percentage of participants with treatment ongoing at the data cut-off date
- Number and percentage of participants off study treatment, grouped by main reason (treatment completed as per protocol, progressive disease, death, adverse event, lost to follow-up, protocol non-compliance, withdrew consent, other)

With regards to treatment reinitiation status, it will be summarized by:

- Number and percentage of participants who reinitiated the study treatment
- Number and percentage of participants with treatment ongoing after reinitiation
- Number and percentage of participants who discontinued treatment after reinitiation

The end of study status will be summarized by:

- Number and percentage of participants with study ongoing at cut-off date (including participants with treatment ongoing and participants who are off treatment and in follow-up)
- Number and percentage of participants who completed/discontinued the study participation, with the associated main primary reason (study completed according to protocol, adverse event, lost to follow-up, protocol non-compliance, death, withdrew consent, other)

In addition, the number of participants screened, and enrolled in each analysis population defined in Section 8.1 will be summarized, overall and by region (North America and Europe), country within region and site.

A listing containing participant disposition information will be displayed, including all participants (i.e. including screening failures, but not re-screened participants (at their screen failure time) which will be listed in a specific listing). The listing will include the following information: participant identifier, date of informed consent, included in the study (if not reason for exclusion), first/last study intervention date, date and reason off-treatment, date and reason off-study, population flags. If the reason for exclusion from trial, off-treatment or off-study are categorized as "Other, specify" or "Withdrew consent from treatment, specify", the verbatim text as entered in the eCRF will be presented in the listing. This listing will be sorted by participant ID.

In addition, a listing of participants for which study treatment has been reinitiated will be provided with the following information: participant identifier, date of first study intervention, date of last study intervention, reason for treatment termination, date of last reinitiation of study treatment administration date and status at end of treatment reinitiation including reason for treatment discontinuation, as applicable. If the status at end of treatment is categorized as "Other, display the verbatim text as entered in the eCRF. This listing will be sorted by participant ID.

If any re-screened participants are observed, they will be presented in a specific listing which will include: participant identifier (identifier at inclusion in the study), date of informed consent at inclusion, date of first study treatment administration, initial participant identifier (identifier at screen failure), date of informed consent at screen failure, date and reason of screen failure. Note in case participants have been screened several times, all screening attempts will be listed. This listing will be sorted by participant ID.

10.2 Protocol Deviations/Exclusion from Analysis Populations

10.2.1 Important Protocol Deviations

Analysis Set: FAS

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.

A full list of potential protocol deviations including definition and categorization is maintained in Appendix 1 – Definition of Important Protocol Deviations.

Important protocol deviations will be summarized for:

- Deviations from the inclusion and exclusion criteria
- Deviations post inclusion

A frequency table for IPDs, separated for such pre-/post inclusion deviations, as well as a listing will be provided. The listing will include: participant identifier, type of protocol deviation (e.g. inclusion/exclusion), category of the deviation (code assigned), and a description of the deviation and will be sorted by participant identifier.

COVID-related protocol deviations will contain the wording "COVID", "COVID-19", "SARS-COV-2" or similar at the beginning of the PD description in order to easily identify them.

Considerations for Pharmacokinetics

Protocol deviations or events will be reviewed by the study pharmacokineticist and biostatistician to identify deviations or events which have the potential to affect the PK results. Deviations, changes to the procedures, or events resulting in non-evaluable PK results, and thus exclusion from PK summaries, will be documented as reasons for exclusion from these summaries, and will be described within the CSR body text. These changes or events will include any circumstances that will alter the evaluation of the PK. Examples of important protocol deviations or important events for PK in terms of this study may include, but may not be limited to the following:

- Dose delayed outside the allowed window
- Actual dosing time not recorded
- Dose change, missed dose, or incomplete/inaccurate dose
- Pre-dose sample collected after the actual start of infusion
- End-of-infusion sample collected before the actual end of infusion
- Documented sample processing errors or analytical errors that may lead to inaccurate bioanalytical results

For the above important protocol deviations or important events for PK, the relevant PK data will be excluded from summaries based on the PKAS. Other changes to the procedures or events which do not impact the quality of the PK data will not be considered important protocol deviations for PK. Common examples of minor (non-important) protocol deviations are a missed sample or minor deviations from sample collection times/windows.

Refer to section 16.1 for more details of protocol deviations and handling relevant to PK.

10.2.2 Reasons Leading to the Exclusion from an Analysis Population

A frequency table per listing of excluded PK data and reasons of for exclusion from summaries/analyses based on the PKAS as well as a listing will be provided.

11 Demographics and Other Baseline Characteristics

Analysis Set: FAS

11.1 Demographics

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

The following demographic characteristics will be included:

- Sex: male, female
- Ethnicity: Hispanic or Latino/Not Hispanic or Latino, Japanese/Not Japanese
- Race:
    - For participants reporting one race only: White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, Not collected at the site, Other.
    - For participants reporting several races, all combinations will be reported under 'More than one race" category.
- Age (years)
- Age categories:
    - < 65 years: 18-49 and 50-64 years
    - ≥ 65 years: 65-74, 75-84 and ≥85 years
- Pooled Geographic Region:
    - North America
    - Europe

Demographic characteristics including participant identifier, sex, race (including all reported races in case of "multiple" races, and details in case of "other" race), ethnicity, geographic region, age (years and category), height, weight and body mass index (BMI) at baseline will be presented in a listing and sorted by participant ID.

Specifications for computation:

- Age [years]: (date of given informed consent - date of birth + 1) / 365.25
    - In case of missing or partial dates see age imputation rules in Table 3.
    - The integer part of the calculated age will be used for reporting purposes.
- Site codes will be used for the determination of the participant's geographic region.
- BMI (kg/m^2) = weight (kg)/[height (m)]^2.

CCI [redacted] CCI

11.2 Medical History

The medical history will be summarized from the "Medical History Details" eCRF page, using the most recent Medical Dictionary for Regulatory Activities (MedDRA) version at time of database lock, preferred term (PT) as event category and system organ class (SOC) body term as Body System category.

Medical history will be displayed in terms of frequency tables: ordered by primary SOC and PT in alphabetical order. Each participant will be counted only once within each PT or SOC. Categories for human immunodeficiency virus (HIV) infection, multi-drug resistance preventing effective antiretroviral therapy, multi-drug resistance not preventing effective antiretroviral therapy and acquired immune deficiency syndrome (AIDS) defining opportunistic infection will be also displayed in a separate table.

Listing of medical history including participant identifier, age, sex, race, PT, reported medical history term, start/end dates, related study condition, ongoing at screening and toxicity grade (when medical history is ongoing) will be presented. This listing will be sorted by participant identifier, start date, end date and PT.

11.3 Other Baseline Characteristics

11.3.1 Disease History

Information on disease characteristics collected on the "Disease History" eCRF page will be summarized as follows:

- Histological diagnosis: lobular, ductal, paget, other
- Time since initial cancer diagnosis (months)
- Time since documented, inoperable or metastatic disease diagnosis (months)
- Time since last progression of disease prior to study entry (months)
- Tumor Node Metastasis (TNM) classification at initial diagnosis: each T, N, M category will be described (TX, T0, N1, etc.)

Listing will also be provided with the following information: participant identifier, age, sex, race, histological diagnosis, date of initial cancer diagnosis (months), date of documented, inoperable or metastatic disease (months), date of last progression of disease prior to study entry (months), and TNM classification at initial diagnosis.

A second listing will be provided containing the molecular abnormalities information, including: participant identifier, age, sex, race, molecular abnormality (estrogen receptor, progesterone receptor, human epidermal growth factor 2, and other molecular abnormality), date of sample (months before first bintrafusp alfa administration), result and percentage. For other molecular abnormalities, tumor genomic marker and codon and kit used for mutation assessment will also be displayed. This listing will be sorted by participant ID.

11.3.2 ECOG Performance Status at Baseline

The ECOG Performance Status will be described from the data collected on the "ECOG Performance Status" eCRF page. It will be described at baseline by the frequency and percentage of participants in each category:

- 0: Fully active, able to carry on all pre-disease performance without restriction.
- 1: Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g., light housework, office work.
- 2: Ambulatory and capable of all selfcare but unable to carry out any work activities. Up and about more than 50% of waking hours.
- 3: Capable of only limited selfcare, confined to bed or chair more than 50% of waking hours.
- 4: Completely disabled. Cannot carry on any selfcare. Totally confined to bed or chair.
- 5: Dead.

11.3.3 Vital signs at Baseline

The following vital signs at baseline will be collected from the "Vital signs" eCRF page and will be summarized:

- Height (cm)
- Weight (kg)
- Body Mass Index (BMI) (kg/m^2)

Height, weight and BMI at baseline will be listed in the demographics listing (see section 11.1, also for BMI computation).

Other vital signs parameters such as systolic and diastolic blood pressure, pulse rate and respiration rate at baseline will be included in the summary tables by timepoint.

11.3.4 Skin status history

Skin status history is collected on the "Skin Status History" eCRF page and will be summarized by the frequency and percentage of participants having the following history of:

- Frequent sunburn (Yes, No, Unknown)
- Easy sunburn (Yes, No, Unknown)
- Skin cancer (Yes, No, Unknown)
- Significant UV exposure (Yes, No, Unknown)
- Photosensitivity due to skin disorder (Yes, No, Unknown)
- Photosensitivity due to medication (Yes, No, Unknown)

- Family history of skin cancer in first degree relative (i.e. parents, siblings and/or children) (Yes, No, Unknown)
- Number of participants having history of the skin conditions above (No condition, 1 condition, 2 conditions, 3 or more conditions)

A listing of skin status history will be provided.

11.3.5 PD-L1 test

All tests collected on the "PD-L1 Tumor Testing" eCRF page will be summarized. This includes the Roche Ventana PD-L1 SP142 assay and other testing. All tests will be described using the following categories: positive, negative, not evaluable (NE) or other.

Listing will also be provided and include: participant identifier, age, sex, race, visit, collection date, date of PD-L1 testing, unique sample identifier, test performed (SP142 assay or other), result (positive, negative, NE or other), tissue used (specimen type and location), if testing performed for clinical trial inclusion, if there is membrane tumor cell staining (% of tumor cell stained), if there is immune cell staining (staining score) and if the combined tumor cell and tumor associated immune cells score available (score). This listing will be sorted by participant identifier.

At the end of the study, retrospective testing on tumor samples will be performed by a central lab. Analyses to be done are described in section 16.4.

11.3.6 HIV testing

The result from HIV testing will be reported in a listing, including participant identifier, age, sex, race, test performed, test name, testing date and result. This listing will be sorted by participant identifier.

11.3.7 Tumor Biopsy

Tumor biopsy is collected on the "Tumor Biopsy (Block or Slides)" eCRF page, a listing of tumor biopsy will be provided including: participant identifier, age, sex, race, visit, collection date, sample type (fresh biopsy or archival tissue sample), sample identifier, sample type for tumor block and for tumor tissue the number of slides collected. In case tumor is assessed as not evaluable for HMGA2 expression by the central laboratory, if an additional tumor specimen has been submitted, its date of collection and additional specimen ID will be displayed. This listing will be sorted by participant identifier and collection date.

11.4 Prior Anticancer Therapy

The prior anti-cancer therapies are collected under the "Prior Anti-Cancer Drug Therapies Details" and "Prior Anti-Cancer Surgeries Details" eCRF pages.

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

- Participants with at least one type of prior anticancer treatment or procedure (i.e. drug therapy, radiotherapy or procedure)
- Participants with at least one prior anticancer drug therapy
- Participants with at least one prior anticancer radiotherapy
- Participants with at least one prior anticancer surgery

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

- Number of prior lines of therapy for metastatic/locally advanced disease: 0 / 1 / 2 / 3 / ≥4
- Intent of Therapy: Neoadjuvant / Adjuvant / Metastatic or Locally advanced
- Time since documented progression disease while receiving prior anticancer drug therapy.
- Best response of last treatment regimen: Complete Response (CR) / Partial Response (PR) / Stable Disease (SD) / Progressive Disease (PD) / Non-Complete Response/Non-Progressive Disease (Non-CR/Non-PD) / Not evaluable / Unknown. Best response of last treatment regimen will be identified based on previous anticancer drug therapies with non-missing start date.

The percentages of prior anticancer drug therapies will be based on the number of participants who had received at least one previous anticancer drug therapy regimen.

Additionally, separate listings for prior anticancer drug therapies, radiotherapy and surgeries will be displayed. These listings will include the participant identifier, age, sex, race and all the relevant collected data-fields on the corresponding eCRF pages as follows:

- For drug therapies: regimen ID, PT, medication name, start date, end date, intent of therapy, best response and date of progression . This listing will be sorted by participant identifier, regimen ID, start date, end date (note that missing end date will be considered as ongoing and will be displayed after non-missing date in case of same start date) and PT.
- For radiotherapies: record ID, stat and stop dates and location of radiotherapy. This listing will be sorted by participant identifier, record ID, start date and end date (note that missing end date will be considered as ongoing and will be displayed after non-missing date in case of same start date).
- For surgeries: date of surgery, name and location of surgery, curative intent of surgery (Y/N), and outcome of surgery. This listing will be sorted by participant identifier, date of surgery and name of surgery.

12 Previous or Concomitant Medications/Procedures

Analysis Set: FAS

CCI [redacted]

12.1 Previous and concomitant medications

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

Concomitant treatments are medications, other than study treatment, which are taken by participants any time during the on-treatment period (see section 9.8).

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 "Concomitant Medication Details" eCRF. ATC-2nd level and PT will be tabulated as given from the WHO-DD dictionary most current version at the time of the database lock. 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.

Specific rules will be used to define if a medication/procedure is considered as a previous, concomitant or both previous and concomitant medication/procedure as detailed in section 9.9.

Previous and concomitant medications will be presented in listing and will include: participant identifier, age, sex, race, PT, medication name as provided by the Investigator, start date, end date, dose, dose units, frequency, route, reason for the medication. This listing will be sorted by participant identifier, start date, end date (note that missing end date will be considered as ongoing and will be displayed after non-missing date in case of same start date) and PT.

12.2 Premedications prior to bintrafusp alfa administration

Premedications are medications administered per protocol on the same day as, but prior to, the study intervention to mitigate potential infusion-related reactions.

As per protocol, premedication prior to bintrafusp alfa administration is optional and at the discretion of the Investigator. If Grade $\geq$ 2 IRRs are seen during the first 2 infusions, premedication should be continued/implemented for future infusions.

Participants receiving premedication during the study will be summarized for each treatment visit based on "Premedication Bintrafusp Alfa Details" eCRF page. Percentages will be calculated on the number of participants who received an infusion at the associated visit.

Additionally, a listing will be provided including: participant identifier, age, sex, race, medication name, visit, date/time of study intervention, dose, dose units, and route. Premedications occurring

CCI
CCI

after reinitiation of treatment will be flagged. This listing will be sorted by participant identifier, date and time of administration and medication name.

12.3 Concurrent procedures

Concurrent procedures are reported according to the "Concomitant Procedures Details" eCRF page.

A listing will be provided including: participant identifier, age, sex, race, name of procedure (as provided by the Investigator), start date, end date, indication, reason for procedure and type of specimen collected. A flag will be displayed to identify each procedure as prior to treatment and on-treatment. In case the procedure has been conducted during the reinitiation phase it will also be flagged. This listing will be sorted by participant identifier, start date, end date (note that missing end date will be considered as ongoing and will be displayed after non-missing date in case of same start date) and procedure name.

12.4 Subsequent anticancer treatments and procedures

Anticancer therapy after end or discontinuation of study treatment will be summarized according to "Anti-Cancer Treatment After Discontinuation Details" for anticancer drug therapies, "Radiotherapy After Discontinuation Details" and "Surgery After Discontinuation Details" eCRF pages. The earliest date of start of new anti-cancer drug therapy will be used for the definition of the on-treatment period and the earliest date of start of new anti-cancer therapy will be used for censoring for efficacy analyses.

The number and percentage of participants in each of the following anticancer treatment categories will be tabulated:

- Participants with at least one subsequent anticancer treatment (i.e. drug therapy, radiotherapy or surgery)
- Participants with at least one subsequent anticancer drug therapy
- Participants with at least one subsequent anticancer radiotherapy
- Participants with at least one subsequent anticancer surgery

If collected in the eCRF, the best response of first subsequent anticancer drug therapy will also be described (Complete Response (CR) / Partial Response (PR) / Stable Disease (SD) / Progressive Disease (PD) / Non-Complete Response/Non-Progressive Disease (non-CR/non-PD) / Not Evaluable / Unknown).

In addition, the anticancer treatment after discontinuation of study treatment will be provided in two listings:

- For medication: participant identifier, age, sex, race, PT/medication name, regimen name, start date, end date, route and best response (if collected in the eCRF). This listing will be sorted by participant identifier, start date, end date, PT/medication name and regimen name.

CCI
CCI

For radiotherapy and surgery: participant identifier, age, sex, race, start date, end date, radiotherapy site or name of surgery/location. For surgery only: outcome and was the surgery curative in intent (Y/N).

13 Study Treatment: Compliance and Exposure

Analysis set: SAF

Participants will receive an intravenous infusion of bintrafusp alfa at a dose of 1200 mg over 1 hour (-10 minutes/+ 20 minutes, i.e., over 50 to 80 minutes) once every 2 weeks as detailed in the Table 1 in the study protocol, until confirmed disease progression (PD), unacceptable toxicity, or occurrence of any criterion for withdrawal from the study.

All dosing calculations and summaries will be based on "Bintrafusp Alfa Administration Details" eCRF page. Data collected during the treatment reinitiation phase will be included in the summary statistics and described in listings.

For the analysis of exposure, a dose is regarded to be administered if the actual dose received is > 0 mg.

Imputation for incomplete start and end dates of study treatments are described in section 9.9.

The **duration** of bintrafusp alfa treatment in weeks during the study is defined as:

$$\text{Duration of treatment (weeks)} = \left(\frac{\text{date of last dose - date of first dose + 14}}{7}\right)$$

The **cumulative dose** overall of bintrafusp alfa per participant is the sum of the actual dose that the participants received during the study (i.e., total dose administered [mg]).

The **dose intensity** of treatment (DI) (mg/week) of bintrafusp alfa during the study is defined as:

$$\text{DI of treatment (mg/week)} = \left(\frac{\text{Cumulative dose of treatment (mg)}}{\text{Duration of bintrafusp alfa treatment (weeks)}}\right)$$

The **relative dose intensity** (RDI) is defined as the actual dose intensity divided by the planned dose intensity during the study and expressed in percentage:

$$\text{RDI of treatment (\%)} = 100\times\left(\frac{\text{DI of treatment (mg/week)}}{\text{planned dose level (mg/week)}}\right)$$

The following summary tables will be provided:

- Duration of therapy (weeks)
- Total number of infusions received
- Cumulative dose (mg)

- Dose intensity (mg/week)
- Relative dose intensity (%) as continue variable, and categorized as:
  - < 80%
  - 80%-90%
  - > 90%

Two listings will be presented:

- A listing of study intervention which will provide: participant identifier, age, sex, race, visit, infusion start date and time, infusion end date and time, infusion rate (mL/hr), actual dose (mg), route, administration modification and reason for modification, change in administration detail, treatment delay (days). Data collected during the treatment reinitiation phase will be flagged.

- An additional listing of treatment exposure and compliance which will include participant identifier, age, sex, race, duration of therapy (weeks), total number of infusions received, cumulative dose of therapy (mg), dose intensity (mg/week), and relative dose intensity (%).

Dose Modification

Dose modification is not allowed per protocol. No summaries will be provided.

Infusion Rate Reductions

Infusion rate reductions as recorded on the "Bintrafusp Alfa Administration Details" eCRF page will be used for analysis. Number of participants with at least one infusion rate reduction, reason for infusion rate reductions (adverse event or other), as well as a categorization of the number of infusion rate reductions (1 / 2 / ≥ 3) will be summarized.

Therapy Delays

Delays of therapy can be calculated only for participants having at least two treatment administrations. It will be derived for each infusion as the number of days since last infusion – 14:

$$Therapy\ delay = start\ date\ of\ current\ infusion - start\ date\ of\ the\ previous\ infusion - 14$$

If the result is > 0 day, then this will be classed as a delay. A participant may have more than one treatment delay throughout the course of treatment.

The following will be summarized in a table:

- Number of participants with at least one delay
- Number of participants with no delay
- Number of delays per participant (1 delay, 2 delays, 3 delays, ≥ 4 delays)
- Longest delay per participant (1-2 days, 3-8 days, 9-15 days, ≥ 16 days)

CCI
CCI

Infusion Temporary Interruptions

Study drug infusion temporarily interrupted as recorded on the "Study Treatment Administration Details" page of the eCRF will be used for analysis. Number of participants with at least one study drug temporary interruption, reason for study drug temporary interruption (adverse event or other), as well as a categorization of the number of study drug temporary interruptions (1 / 2 / ≥ 3) will be summarized.

14 Efficacy Analyses

Analysis Set: FAS

14.1 Primary Endpoint: Confirmed Objective Response

Primary endpoint is the confirmed objective response according to Response Evaluation Criteria in Solid Tumors Version 1.1 (RECIST 1.1) assessed by an IRC.

14.1.1 Primary Objective: Derivation and Analysis of the Primary Endpoint Confirmed Objective Response

Best Overall Response

Best overall response (BOR) will be assessed based on reported overall responses at different evaluation time points from the treatment start date until documented disease progression in accordance to RECIST v1.1, 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.

BOR can be defined as confirmed or unconfirmed, both definitions are provided below.

Confirmed Best Overall Response (cBOR)

- CR = at least two determinations of CR at least 4 weeks apart (with no PD in between).
    - Note: Assessments with results CR-NE-CR, with NE as not evaluable, or CR-PR-CR will be considered as CR as long as the second CR is more than 28 days away from the first timepoint.
- PR = at least two determinations of PR or better (PR followed by PR or PR followed by CR) at least 28 days apart (and not qualifying for a CR), with no PD in between.
    - Note: Assessments with results PR-NE-PR or PR-SD-PR will be considered PR as long as the second PR is more than 28 days away from the first timepoint.

CCI [redacted]

- SD (applicable only to participant with measurable disease at baseline) = at least one SD assessment (or better) ≥ 6 weeks, after start date and not qualifying for CR or PR.
- Non-CR/non-PD (applicable only to participants with non-measurable disease at baseline) = at least one non-CR/non-PD assessment (or better) ≥ 6 weeks after start date (and not qualifying for CR or PR).
- PD = PD ≤ 12 weeks after start date (and not qualifying for CR, PR, non-CR/non-PD or SD).
- Not Evaluable (NE): all other cases.

Unconfirmed Best Overall Response (uBOR)

uBOR according to RECIST 1.1 as adjudicated by IRC will be assessed according to the following rules:

- CR = at least one assessment of objective status of CR documented before progression.
- PR = at least one assessment of objective status of PR documented before progression (and not qualifying for CR).
- SD (applicable only to participant with measurable disease at baseline) = at least one SD assessment (or better) ≥ 6 weeks after start date and before progression (and not qualifying for CR or PR).
- Non-CR/non-PD (applicable only to participants with non-measurable disease at baseline) = at least one non-CR/non-PD assessment (or better) ≥ 6 weeks after start date and before progression (and not qualifying for CR or PR).
- PD = progression ≤ 12 weeks after start date (and not qualifying for CR, PR, SD or non-CR/non-PD).
- NE: all other cases.

Objective Response (OR)

The Objective Response (OR) is defined as a BOR of CR or 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').

Objective Response Rate (ORR)

The ORR is defined as the number of participants having an OR assessment of CR or PR (responders), out of the total number of participants in the analysis set.

Confirmed ORR will be determined as the proportion of participants in the analysis set with a confirmed OR of PR or CR (having at least two determinations of CR/PR at least 4 weeks apart and before progression).

The confirmed ORR according to RECIST 1.1 as assessed by the IRC is the primary endpoint of the study. The study aims to estimate the ORR with a sufficient level of precision and that the associated two-sided 95% CI is above a minimal threshold of 10%.

Disease Control Rate (DCR)

Disease Control Rate (DCR) is defined as the proportion of participants with BOR according to evaluation criteria of CR, PR, or SD out of the total number of participants in the analysis set.

Details of analyses

Separate tables will be provided for confirmed and unconfirmed BOR. Each one will provide the number and percentage of participants with cBOR (respectively uBOR) of CR, PR, SD, PD and NE as well as the ORR and the DCR. The ORR and DCR will be provided with a two-sided 95% CI using the Clopper-Pearson method (exact CI for a binomial proportion as computed by default by the FREQ procedure using the EXACT option).

The individual percentage of change in the sum of diameter since baseline will be displayed over time on a spider plot, together with the first occurrence of new lesion, the participant off treatment and the start of subsequent anticancer treatment. The curves will be color-coded to differentiate responder vs non-responder subjects.

The change in the sum of diameters between baseline and the best post-baseline assessment (i.e. maximum tumor shrinkage) will be displayed on a waterfall graph. The sum of diameters includes all target lesions (longest diameter for non-nodal lesions and short axis for nodal target lesions).

For spider plots and waterfall plots, the percent change from baseline in the sum of diameters will be displayed for valid timepoint assessments, only. For the purpose of this analysis, a valid timepoint assessment is defined as a complete assessment of all target lesions reported at baseline. Further, split and coalesced lesions have to be taken into account appropriately to determine if a timepoint assessment is valid to derive the percent change from baseline in sum of diameters. All sum of diameters will be used, including the ones beyond first PD (for waterfall plot, the best post-baseline sum of diameters will be used even if occurring beyond PD). For waterfall plot, the percent change from baseline to 8-weeks assessment, as well as the percent change from baseline to the best post-baseline sum of diameters will be displayed for each participant.

Two listings of tumor assessment will be provided with the following information:

- First listing will display participant identifier, age, sex, race, first and last date of treatment, date of start of subsequent anticancer therapy, date of death when death occurs, visit, date(s) of imaging, description of target lesions (size, site, type, method), non-target lesions (status, site, type, method), and new lesions (site, type, method) and if the lesion split or merged. This listing will be sorted by participant identifier, visit, date of imaging, lesion type (TL, NTL, NEW) and lesion ID.

CCI
CCI

- Second listing will present participant identifier, age, sex, race, first and last date of treatment, cBOR, uBOR, date of start of subsequent anticancer therapy, date of death when death occurs, target response, non-target response, new lesion, sum of diameters, % of change in sum of diameters since baseline and overall response. This listing will be sorted by participant identifier, visit and date of imaging.

In addition, a summary table of the reasons for non-evaluable confirmed BOR will be provided, the following reasons will be detailed:

- No baseline assessment (if applicable)
- No post-baseline assessments due to death within 8 weeks after the start of study treatment
- No post-baseline assessments due to other reasons
- All post-baseline assessments have overall response 'Non-evaluable'
- New anticancer therapy started before first evaluable post-baseline assessment
- SD of insufficient duration (< 6 weeks after the start of study treatment)
    - Note: Special cases where BOR is NE due to both early SD and late PD will be classified into this category
- No evaluable tumor assessment > 16 weeks followed by PD (i.e. tumor assessment of PD was > 16 weeks after start of study treatment and there was no evaluable tumor assessment in between)
- No IRC review and Not determined categories may also be added if applicable

A listing of reasons for non-evaluable confirmed BOR will also be created including: participant identifier, age, sex, race, date of first and last dose, date(s) of imaging, overall response and the reason for confirmed BOR non-evaluable.

14.1.2 Subgroup Analysis of the Primary Endpoint Confirmed Objective Response

Confirmed BOR, ORR and DCR according to RECIST 1.1 as adjudicated by IRC will also be evaluated on all subgroups as defined in section 8.2.

Participants with missing subgroup category will not be included in the related subgroup analysis. Response rates and corresponding 95% CI of all subgroups will be presented in a forest plot together with the respective results for the primary analysis set. Corresponding efficacy tables will not be provided.

No adjustment for multiplicity will be performed.

14.2 Secondary Endpoint: Objective Response According to RECIST 1.1 as Assessed by Investigator

Data collected in the eCRF for RECIST 1.1 response criteria as assessed by the Investigator ("Assessment of disease based on imaging (according to RECIST 1.1)" eCRF page) will be treated

CCI [redacted] CCI

in the same way as data adjudicated by the IRC. Analyses described in section 14.1.1 will be repeated based on Investigator assessment, i.e.:

- Number and percentage of participants with confirmed BOR of CR, PR, SD, PD and NE and corresponding listing. Analyses will be repeated for unconfirmed BOR
- ORR and DCR with their two-sided 95% CI. Analyses will be repeated for unconfirmed BOR
- Spider and Waterfall plots
- A summary table of the reasons for non-evaluable confirmed BOR and related listing, with the following reasons:
  - No baseline assessment (if applicable)
  - No post-baseline assessments due to death within 8 weeks after the start of study treatment
  - No post-baseline assessments due to other reasons
  - All post-baseline assessments have overall response 'Non-evaluable'
  - New anticancer therapy started before first evaluable post-baseline assessment
  - SD of insufficient duration (< 6 weeks after the start of study treatment)
  - No evaluable tumor assessment > 16 weeks followed by PD (i.e. tumor assessment of PD was > 16 weeks after start of study treatment and there was no evaluable tumor assessment in between)
  - Not determined category may also be added if applicable
- Listing of reasons for non-evaluable confirmed BOR
- Listing of individual tumors assessment as well as of time point tumor response and overall response

In addition, a summary of the BOR (confirmed and unconfirmed) as adjudicated by IRC versus Investigator assessment will be provided including numbers of concordant and discordant assessments, and a listing of inconsistencies will be provided.

CCI [redacted]

CCI

[redacted]

CCI
CCI

CCI

[redacted]

14.4 Secondary Endpoint: Progression-Free Survival as Adjudicated by IRC

Progression-Free Survival (PFS) time in months, is defined as the time from first administration of study treatment until the first documentation of progression of disease (PD) or death due to any cause, whichever occur first. The tumor response will be determined according to RECIST 1.1 and assessed by the IRC. PFS time will be calculated as follows:

$$PFS\ (months) = \frac{(date\ of\ PD\ or\ death\ or\ censoring - date\ of\ fist\ dose\ of\ study\ treatment + 1)}{30.4375}$$

Tumor assessments will be measured at screening and during the intervention period: at week 9 day 57, week 17 day 113, then every 8 weeks (or until disease progression) for 12 months and then every 12 weeks or until disease progression (see Table 1 in the study protocol).

The following censoring rules will also be applied for the PFS computation:

- Participants with no event (PD or death) will be censored on the date of the last adequate tumor assessment
- Participants who do not have a baseline tumor assessment or who do not have any evaluable post-baseline tumor assessments will be censored at the date of first administration of study treatment unless death occurred on or before the time of the second planned tumor assessment (i.e. 17 weeks) in which case the death will be considered an event.
- Participants who start new anticancer treatment prior to an event will be censored on the date of the last evaluable tumor assessment before anticancer therapy is given.

  Note: any systemic anticancer therapy, any anticancer surgery and any anticancer radiotherapy will be considered as new/subsequent anticancer therapy and will lead to censoring.
- Participants with an event after two or more subsequent missing response assessments (i.e. no assessments in 113 days prior to the event during the first 12 months [i.e. 365 days of follow-up] or 168 days prior to the event after the first 12 months of follow-up) will be censored on the date of the last evaluable tumor assessment.

The last tumor assessment date is defined as the last available and evaluable tumor assessment performed prior to the cut-off date (or prior to end of study, i.e. participants lost to follow-up or who withdraw consent) or prior to subsequent anticancer therapy.

Censoring rules are summarized in Table 8 and Table 9.

Table 8 Date of Event/Censoring Definition for PFS Analysis

Status		Censoring	Date of event / censoring
Progressed or died	Within two subsequent scheduled tumor assessments after last response assessment of CR, PR or SD receiving the first dose of study treatment	Event	Minimum (Date of PD, Date of death)
	Otherwise	Censored	Date of last tumor assessment with outcome CR, PR or SD or date of first study drug administration, whatever is later.
Neither progressed nor died		Censored	Date of last tumor assessment with outcome CR, PR or SD or date of first study drug administration, whatever is later.

CR=Complete response, PD=Progressive disease, PR=Partial response, SD=Stable disease.

Table 9 Outcome and Event Dates for PFS

Scenario	Date of event/censoring	Outcome
No baseline assessment or no evaluable post-baseline assessment	Treatment start date	Censored [a]
PD or death ≤ 16 weeks after last adequate tumor assessment or ≤ 16 weeks after start date	Date of PD or death	Event
PD or death after 2 or more missing [b] post-baseline tumor assessments	Date of last adequate tumor assessment [c] documenting no PD before subsequent anti-cancer therapy is given or missed assessments	Censored [b]
No PD	Date of last adequate tumor assessment [c] documenting no PD before [subsequent anti-cancer therapy is given or] missed assessments	Censored
Treatment discontinuation due to 'Disease progression' without documented progression	Not applicable	Information is ignored. Outcome is derived based on documented progression only.
Subsequent anti-cancer therapy given before PD or death event	Date of last adequate tumor assessment [c] documenting no PD before subsequent anti-cancer therapy is given or missed assessments	Censored

PD=Progressive disease.

[a] However if the participant dies ≤ 17 weeks after start date the death is an event with date on death date.

[b] Derivation of 2 or more missing tumor assessments according to the following rules:

Last Evaluable Scan Prior to Event (PD or Death), days since treatment start	Censor if Duration Between Last Scan and Event
[Day 1, Day 336]	>56 days (8 weeks)
>Day 337	>84 days (12 weeks)

CCI [redacted] CCI

Scenario	Date of event/censoring	Outcome

[c] If there are no adequate post-baseline assessments prior to the PD or death, then the time without adequate assessment should be measured from the start date; if the criteria were met the censoring will be on the start date.

The analysis of PFS as measured by the IRC (secondary endpoint) will be performed with a Kaplan-Meier method (product-limit estimates) and a summary of associated statistics will be presented including corresponding two-sided 95% CIs. The CIs for the median will be calculated according to Brookmeyer and Crowley (1982) and CIs for the survival function estimates will be derived using the log-log transformation according to Kalbfleisch and Prentice (1980) (CONFTYPE = loglog default option in SAS PROC LIFETEST). The estimate of standard error will be computed using Greenwood's formula.

Frequency (number and percentage) of participants with and without event and within each event type (PD or death) will be presented as well as the PFS Med, Min, and Max. Number of participants at risk, failed and PFS rates with their CI at 3, 6, 9, 12, 18 and 24 months will also be provided. Censoring reasons will also be described. Censoring reasons are as follows:

- Ongoing in the study without an event
- No baseline assessment
- No evaluable post-baseline assessment
- Start of new anticancer therapy
- Event after 2 or more missing or non-evaluable post-baseline assessments
- Withdrawal of consent
- Lost to follow-up

Lost to follow-up will include the following participants:

- Lost to follow-up status is collected on the eCRF treatment termination page or eCRF study termination page prior to the analysis cut-off
- Participants with the last alive date > 14 weeks prior to the analysis cut-off date (duration of 14 weeks is based on the assessment schedule of every 3 months for survival follow-up interval + 2-weeks window)

PFS will also be presented graphically with Kaplan-Meier figures.

Listing will be provided with the following information: participant identifier, age, sex, race, date of first administration of study treatment, date of last tumor assessment, date of event/censoring, event/censoring reason, time to event. This listing will be sorted by participant identifier.

14.4.1 Secondary Endpoint: Sensitivity Analysis for PFS

PFS as assessed by the Investigator will be described in the same way as the PFS as assessed by IRC using rules for censoring defined in Table 8 and Table 9.

CCI
CCI

These sensitive analyses will be supported with Kaplan-Meier plot and a listing as described in section 14.4.

CCI

[redacted]

14.6 Secondary Endpoint: Duration of Response

Duration of Response (DOR) is defined for participants with a confirmed objective response as the time from first documentation of a confirmed objective response (CR or PR) according to RECIST 1.1 to the date of first documentation of objective progression of disease (PD) or death due to any cause, whichever occurs first. The censoring rules for DOR are as described for PFS (Table 8 and Table 9).

$$DOR\ (months) = \frac{(date\ of\ PD\ or\ death\ or\ censoring - date\ of\ confirmed\ objective\ response + 1)}{30.4375}$$

DOR of confirmed CR/PR according to RECIST 1.1 as measured by the IRC and the Investigator will be described. The date of occurrence of the first CR/PR will be used as date of objective response.

CCI
CCI
[redacted]

Frequency (number and percentage) of participants with an event will be presented as well as the DOR time Med, Min, and Max. Number of participants at risk, failed and DOR rates with their CI at 3, 6, 9, 12, 18 and 24 months will also be provided.

The analysis of DOR will be performed with a Kaplan-Meier method (product-limit estimates) and a summary of associated statistics will be presented including corresponding two-sided 95% CIs. The CIs for the Med will be calculated according to Brookmeyer and Crowley (1982) and CIs for the survival function estimates will be derived using the log-log transformation according to Kalbfleisch and Prentice (1980) (CONFTYPE = loglog default option in SAS PROC LIFETEST). The estimate of standard error will be computed using Greenwood's formula.

The time to and duration of response per participant having a confirmed objective response will be displayed in swimmer graphs. Kaplan-Meier figures will also be provided.

A listing for responders will be provided with the following information: participant identifier, age, sex, race, date of first study treatment administration, date of first response, date of last tumor assessment, censored (Y/N), date of event/censoring, event/censoring reason, duration of response and if the response is ongoing (administrative censoring for subject ongoing in the study without an event). This listing will be sorted by participant identifier.

Exploratory subgroup analyses for DOR, using the subgroups defined in section 8.2 will be performed using Kaplan-Meier figures. Participants with missing subgroup category will not be included in the related subgroup analysis.

Durable Response Rate (DRR) is defined as the number of participants having a DOR of at least 6 months, out of the total number of participants. Participants for whom the DOR is censored will be treated as not having a DOR of at least 6 months.

The number of participants with DRR will be tabulated and a two-sided 95% CI for percentage using the Clopper-Pearson method (exact CI for a binomial proportion as computed by default by the FREQ procedure using the EXACT option).

DRR will be described according to RECIST 1.1 as measured by the IRC and the Investigator.

14.7 Secondary Endpoint: Overall Survival

Overall survival (OS) is defined as the time from first administration of study treatment to the date of death due to any cause:

$$OS\ (months) = \frac{(date\ of\ event\ or\ censoring - date\ of\ first\ dose + 1)}{30.4375}$$

For participants alive at the time of data cut-off date or who are lost to follow up, OS will be censored at the last date known to be alive. The date of event/censoring is defined in Table 10.

CCI
CCI
[redacted]

Table 10 Date of event/censoring definition for OS analysis

Survival Status		Source	Censoring	Date of event/censoring
Died	Before cut-off	Death eCRF	Event	Date of death
	After cutoff	Death eCRF	Censored	Date of cutoff
Alive (no date of death)	Alive after cut-off	FU eCRF	Censored	Date of cut-off
	Otherwise	Date of last known to be alive, see Section 9.6	Censored	Date of last contact
eCRF: electronic case report form, FU: follow-up.				

The analysis of OS time will be performed with a Kaplan-Meier method with the same approach as for PFS described in section 14.4. Kaplan-Meier estimates (product-limit estimates) will be presented together with a summary of associated statistics including corresponding two-sided 95% CIs. OS rates with their CI at 3, 6, 9,12, 18 and 24 months will be presented, as well as the number of participants at risk and failed. Censoring reasons will also be described. Censoring reasons are as follows:

- Alive at cut-off date
- Withdrawal of consent
- Lost to follow-up

OS will also be presented graphically with Kaplan-Meier figures.

A participant listing will provide the following information: participant identifier, age, sex, race, date of first study intervention, date of event/censoring, event/censoring reason, time to event. This listing will be sorted by participant identifier.

CCI

[redacted]

CCI

[redacted]

15 Safety Analyses

Analysis set: SAF

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.

15.1 Adverse Events

Treatment-emergent adverse events (TEAEs) 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 (see on-treatment definition in section 9.8).

Changes in toxicity grade, seriousness or outcome of AEs are recorded as separate entries in the eCRF. Such entries reporting the same event in such immediately consecutive periods will be considered as one event in the analysis (using the field "AE id for new grade" in the eCRF). These events will be kept as separate records in the database in order to maintain the full detailed history of the events. The start date of the initial record in the sequence is taken as start date of the entire event, similarly the end date of the last event in the sequence is taken as end date of the entire event. The overall outcome of the adverse event is the outcome of the last event in the sequence. Duration of the AE and the TEAE flag will be adjusted accordingly in the analysis.

Incomplete AE-related dates will be handled as stated in section 9.9.

Related Adverse Events are those events with relationship to study treatment (as recorded on the "Adverse Events Details" eCRF page, Relationship with Bintrafusp alfa= Related) reported by the Investigator and those of missing or unknown relationship.

Serious Adverse Events (SAE) are those events reported on the "Adverse Events Details" eCRF page, with the "Serious Adverse Event" field ticked "Yes".

CCI
CCI [redacted]

Adverse Events Leading to Temporary Discontinuation are those events leading to temporary discontinuation of study treatment (answer to the question "Action(s) taken with Bintrafusp alfa" = "Drug interrupted" on "Adverse Event Details" eCRF page).

Adverse Events leading to Permanent Treatment Discontinuation are those events leading to permanent discontinuation of study treatment (answer to the question "Action(s) taken with Bintrafusp alfa" = "Drug withdrawn" on "Adverse Event Details" eCRF page).

Adverse Events leading to Death are those events leading to death (as recorded on the "Adverse Event Details" eCRF page, change in grade = "No" and outcome = "Fatal", or Grade = "Grade 5 or death related to AE" or Serious adverse event = "Yes" and seriousness criteria include "Results in death").

Adverse Events of Special Interest (AESI): AESI are identified according to a pre-specified search list of MedDRA Preferred Terms (PTs) which will be finalized before the database lock. Categories of AESI include:

- Infusion-Related Reactions (IRRs)
- Immune-Related Adverse Events (irAEs)
- Skin AEs possibly related to TGF-β inhibition
- Anemia

Bleeding events are those events belonging to the MedDRA SMQ Haemorrhage terms (excluding laboratory terms).

All analyses described in this section will be based on TEAEs if not otherwise specified. The AE listings will include all AEs. AEs outside the on-treatment period (prior or after) will be flagged in listings. AE occurring during the reinitiation phase will be considered as TEAE and will be reported in the summary tables and will be flagged in the listings.

15.1.1 All Adverse Events

Adverse events will be summarized by worst severity (according to National Cancer Institute - Common Terminology Criteria for Adverse Events [NCI-CTCAE] version 5.0) per participant, using the latest available version of MedDRA PT at the time of the database lock as event category and MedDRA primary SOC body term as Body System category.

Unless otherwise stated adverse events will be displayed in terms of frequency tables: primary SOCs and PTs in alphabetical order.

If an adverse event is reported for a given participant more than once during treatment, the worst severity and the worst relationship to study treatment 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.

A table presenting the overall summary of AEs will be presented including the frequency (number and percentage) of participants within each of the following categories:

- TEAEs
- Related TEAEs
- Serious TEAEs
- Related Serious TEAEs
- TEAEs NCI-CTCAE severity grade (≥3, ≥4)
- Related TEAEs NCI-CTCAE severity grade (≥3, ≥4)
- TEAEs leading to death
- Related TEAEs leading to death
- AEs of special interest:
    - Infusion-related reactions (IRRs)
    - Immune-related AEs (irAEs)
    - Potential TGFβ inhibition mediated skin Aes
    - Bleeding events
    - Anemia
- Related AEs of special interest:
    - Infusion-related reactions (IRRs)
    - Immune-related Aes (irAEs)
    - Potential TGF-β-mediated skin Aes
    - Bleeding events
    - Anemia

Tables for TEAEs frequency corresponding to each category in the overview table above will be provided by:

- MedDRA primary SOC (ordered alphabetically) and PT (ordered alphabetically), (except for AE of special interest and TEAE with NCI-CTCAE severity grade ≥3 or ≥4).

TEAEs and related TEAEs by worst grade will also be summarized, and the most frequent PTs (at least 5%) will be presented graphically by worst grade and PT with bar chart figures.

Clinicaltrials.gov and EudraCT - requirements

Summary table for non-serious TEAEs excluding SAEs applying frequency threshold of 5% will be provided.

Listings of adverse events will contain the following information: participant identifier, age, sex, race, first and last date of study intervention, PT, reported term for the AE, start date, end date, duration of AE (in days), day relative to the first infusion, day relative to the most recent infusion prior to AE onset, relationship to study treatment, toxicity grade, action(s) taken, outcome,

seriousness (Y/N), AESI infusion-related (Y/N), AESI immune-related (Y/N), potential TGF-β-mediated skin AESI (Y/N), bleeding event AESI (Y/N), anemia AESI (Y/N) and AE occurring after COVID-19 pandemic start (Y/N). TEAEs and AEs occurring during the reinitiation period specifically will be flagged. This listing will be sorted by participant identifier, start and end date and PT.

The following listings will be provided with the relevant information:

- Listing of all AEs (whether treatment-emergent or not) (TEAEs will be flagged; AEs occurring during reinitiation phase will be flagged)
- Listing of TEAEs
- Listing of non-TEAEs for AEs starting and not worsening after enrollment (date of first signature of informed consent/date of first signature of first informed consent) but prior to the first dose of study intervention.
- Listing of AEs with onset or worsening after the on-treatment period (AEs occurring during reinitiation phase will be flagged)
- Listing of all AEs for reinitiated participants (AEs occurring during reinitiation phase will be flagged)

Evaluation of Potential Effect of ADA on Bintrafusp alfa Safety

The overall summary of TEAEs table, as described above will also be provided by ADA status (ever positive, never positive). See section 16.3 and Table 15 for ADA status details.

Listings of all AEs and all IRRs for ever-positive ADA participants (pre-existing, transient treatment-emergent, persistent-treatment emergent) will be prepared including participant identifier and showing the date(s) of the positive ADA result together with the AEs or IRRs. For the AEs and IRRs, start and stop date will be shown along with grade. Adverse events recorded during the period of 2-weeks prior to the positive ADA value till two weeks after the positive ADA value will be flagged.

15.1.2 Adverse Events Leading to Discontinuation of Study Treatment

Frequency tables summarizing the following actions taken with study treatment will be presented by PT and primary SOC in alphabetical order:

- TEAEs leading to temporary treatment interruption
- Related TEAEs leading to temporary treatment interruption
- TEAEs leading to permanent treatment discontinuation
- Related TEAEs leading to permanent treatment discontinuation

- TEAEs leading to infusion rate reduction
- Related TEAEs leading to infusion rate reduction

In addition, the incidences for above items will be summarized in an overview table.

The listing of TEAEs leading to permanent treatment discontinuation will also be provided with the relevant information.

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

15.2.1 Deaths

All deaths, deaths within 30 days after last dose of study treatment (in case of reinitiated participants, the last dose will be the last dose of the reinitiation phase), death within 60 days after first dose of study treatment (for all participants, the first dose will be the first dose of the first treatment phase) as well as the primary reason for death will be tabulated based on information from the "Death" eCRF pages.

The following summaries will be provided:

- Number of deaths (including deaths during reinitiation phase)
- Number of deaths within 30 days after last dose of study treatment
- Number of deaths within 60 days after 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

In addition, date and cause of death will be provided in an individual participant data listing together with following dosing information: participant identifier, age, sex, race, date of first/last study intervention, number of infusions, day relative to the first and the last infusion, primary reason, autopsy (Y/N/U), AEs with fatal outcome (list PTs of AEs with outcome = Fatal, as well as grade 5 or serious resulting in death), flag for death within 30 days of last dose of study treatment and flag for death within 60 days of first dose of study treatment.

15.2.2 Serious Adverse Events

The following overall frequency tables will be prepared for serious adverse events (SAEs):

- Incidence of serious AEs by SOC and PT
- Incidence of related serious AEs by SOC and PT

The listings of SAEs will also be provided with the relevant information with a flag for TEAEs and AEs occurring during reinitiation phase (see description of listing in section 15.1.1).

15.2.3 Adverse Events of Special Interest

Adverse events of special interest (AESI) are serious or nonserious AEs specific to the known mechanism of action of the study intervention that are of clinical interest. The AESIs considered on this study are: infusion-related reactions including immediate hypersensitivity, immune-related adverse events, skin adverse events, bleeding events and anemia.

15.2.3.1 Infusion-Related Reaction including Immediate Hypersensitivity

Infusion-Related Reactions (IRRs) are defined as adverse events with PTs according to a pre-specified MedDRA search list which will be finalized before database lock, and are divided into two subcategories: "Reactions" and "signs and symptoms" based on criteria on the timely relationship as detailed below:

Reactions of IRR: will be considered when onset is on the day of infusion (during or after the infusion) or the day after the infusion (irrespective of resolution date) for any infusion-related reaction, drug hypersensitivity, anaphylactic reaction, hypersensitivity and/or Type 1 hypersensitivity.

Signs and symptoms of IRRs and hypersensitivity/allergic reactions: will be considered when onset is on the day of infusion (during or after the infusion) and resolved completely with the end date on the same day of the infusion or the day after for any of the following: pyrexia, chills, flushing, hypotension, dyspnea, wheezing, back pain, abdominal pain and urticaria.

IRR, overall and by subcategories, will be summarized by the following variables:

- Number of participants with at least one IRR by the worst NCI-CTCAE toxicity grade (grade 1/ grade 2/ grade 3/ grade 4/ grade 5/ missing grade)
- Number of participants with IRR leading to permanent treatment discontinuation.
- Time related to first onset (infusion 1/ infusion 2/ infusion 3/ infusion 4 or later). The events will be assigned to the actual drug infusions that the participant received, not to the planned dates. An IRR is assigned to a drug infusion if its onset is at the same date and time is missing or same or after dosing (but not before dosing when time is recorded) or the following day of drug infusion.

If at least 10 participants received premedication during the study, IRR will be summarized by premedication (Y/N) subgroups. In this table, the worst grade for each subject will be displayed

for all the IRR started after the use of a premedication and the worst grade for each subject for all the IRR started without the use of a premedication.

The frequency table of IRR and drug related IRR AEs by worst grade, SOC, and PT will also be provided.

The listing of IRRs will be provided with the relevant information (see description of listing in section 15.1.1). One additional listing will display the study drug administration details together with the infusion-related adverse event including administration date (day) /time, reason for modification, type of modification, modification start time, use of pre-medication, IRR AE Preferred Term, IRR AE grade, IRR AE start day /stop day, IRR AE time related to infusion.

15.2.3.2 Immune-Related Adverse Events

Immune-related adverse events (irAEs) will be identified programmatically. AEs which satisfy all of the following criteria will be flagged as immune-related:

1) The AE PT matches a PT on the list of pre-selected MedDRA terms.

2) The AE onset or worsening occurs after the first study drug administration and no more than 90 days after last dose.

3) On the "Adverse Events Details" eCRF page, the question "Were Corticosteroids, Immunosuppressants, or hormonal therapy (e.g. Thyroid) or Insulin applied?" has the answer "Yes" selected.

4) On the "imAE Specific Questions" eCRF page, either:

> a. The question "Does any of the following provide a clear etiology for the event?" has the answer "No" selected, indicating that the AE is not attributable to underlying cancer disease/PD, prior or concomitant medications/procedures, nor another medical condition such as an infection or pre-existing disease.
>
> OR
>
> b. The "imAE Specific Questions" eCRF page indicates that a biopsy was performed and the question "Is the histopathology/biopsy consistent with an immune-mediated event?" has the answer "Yes" selected.

In the case that criteria (1) through (3) are met, and entries for condition (4) are missing, the following rules will be applied:

- If the answer to "Does any of the following provide a clear etiology for the event?" (4a) is missing, the event will be considered as irAE (irrespective of biopsy results).
- If the answer to "Is the histopathology/biopsy consistent with an immune-mediated event?" (4b) is missing, or if no biopsy was performed, and condition (4a) is not satisfied, i.e. "Yes" is selected (i.e. at least one (clear) etiology of the event is provided) as the answer to the

CCI [redacted] CCI

question "Does any of the following provide a clear etiology for the event?", the event will be considered as a non-irAE.

PTs will be compiled into categories: Immune-mediated rash, Immune-mediated colitis, Immune-mediated pneumonitis, Immune-mediated hepatitis, Immune-mediated nephritis and renal dysfunction, Immune-mediated endocrinopathies (Adrenal insufficiency, Hypogonadism, Pituitary dysfunction, Type 1 Diabetes Mellitus, Thyroid disorders), Other immune-mediated myositis, Other immune-mediated adverse events. PTs belonging to the category "Immune related endocrinopathies – Thyroid disorders" will also be compiled into sub-subcategories: Hyperthyroidism, Hypothyroidism, Thyroiditis.

Immune-related adverse events (irAEs) will be summarized by the following variables:

- Any irAEs
- irAEs by the worst grade
- irAEs leading to permanent treatment discontinuation
- Serious irAEs

The frequency table of immune-related AEs by worst grade, category, subcategory (for Immune-mediated endocrinopathies), sub-subcategory (for Immune-mediated endocrinopathies – thyroid disorders) and PT will also be provided.

The listing of irAE will also be provided with the relevant information, including additional interventions for irAE (e.g. biopsies, surgical procedures, medical procedures) (see description of listing in section 15.1.1).

15.2.3.3 Potential TGFβ Inhibition Mediated Skin Adverse Events

To identify potential skin AEs possibly related to TGFβ inhibition, MedDRA PT queries will be used to search for skin AEs of interest in the clinical database. A listing containing these pre-specified PT search terms will be generated. PTs will be compiled into categories: Narrow definition, and Broad definition. Further details (e.g. MedDRA PT queries) are regularly updated based on the current MedDRA version.

Narrow definition:

- Keratoacanthoma
- Squamous cell carcinoma of skin

Broad definition has additional PTs:

- Hyperkerathosis
- Actinic keratosis

CCI [redacted] CCI

- Basal cell carcinoma
- Lip squamous cell carcinoma
- Bowen´s disease

The overall summary of potential TGF Beta-mediated skin TEAE will include the following categories for narrow and broad definition:

- All skin TEAE
- All skin TEAE by worst grade
- Skin TEAE leading to permanent treatment discontinuation
- Serious skin TEAEs

Table for skin TEAEs and treatment related skin TEAEs frequency will be provided by MedDRA PTs (including both narrow and broad definition PTs).

A listing of skin AEs will also be provided. This listing will also provide, for the AE identified from the PT list, the number of lesions, if a biopsy or an excision was done and if it confirmed the diagnosis and the lesion location coming from "TGFβ Mediated Skin Reaction" eCRF page.

15.2.3.4 Bleeding Events

Bleeding events are those events belonging to the MedDRA SMQ Haemorrhage terms (excluding laboratory terms). Bleeding events and trial drug related bleeding events will be summarized in a frequency table presenting SOC and PT sorted by alphabetical order. The worst grade per participant, per SOC and per PT will be reported: any grade (including missing grade), grade 1, grade 2, grade 3, grade 4 and grade 5. Additionally, a listing of bleeding AEs will be delivered.

15.2.3.5 Anemia

To identify potential anemia AEs MedDRA PT queries will be used to search for these AEs of interest in the clinical database. A listing containing these pre-specified PT search terms provided with the relevant information (see description of listing in section 15.1.1).

The list of PTs includes:

- Anemias NEC (HTL)
- Anemia haemolytic immune (HLT)
- Anemia haemolytic NEC (HLT)
- Haemoglobin decreased (PT)

CCI
CCI [redacted]

15.3 Clinical Laboratory Evaluation

Baseline and on-treatment laboratory values (including corresponding normal ranges), converted in standard unit, will be used for boxplots and shift tables.

Laboratory results will be classified according to the NCI-CTCAE criteria version 5.0 and as specified in Appendix 2 – Definition of NCI-CTCAE grading. Additional laboratory results that are not part of NCI-CTCAE will be categorized as follows: below normal limits, within normal limits, and above normal limits (according to the original laboratory normal ranges).

The following figures will be provided for hematology and biochemistry parameters:

- Boxplots of the laboratory values by timepoint
- Boxplots of the change from baseline by timepoint

Boxplots for laboratory parameters where toxicity grades are defined based on the ratio of the parameter values and the upper limit of normal (ULN) will not be displayed using the unit of measurement but instead using the ratio of the measured value over ULN. This comprises alkaline phosphatase (ALP), alanine aminotransferase (ALT), aspartate aminotransferase (AST), activated partial thromboplastin time (aPTT), bilirubin, and creatinine.

Laboratory parameters with NCI-CTC grades available

Laboratory parameters with NCI-CTC grades available will be analyzed with their respective NCI-CTC name and direction of abnormality. For parameters which are graded with both low and high values 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.

The following summaries will be displayed:

- Number and percentage of participants by worst on-treatment grade (>= 1, >= 3, >= 4)
- Shift in toxicity grading from baseline to highest on-treatment toxicity grade

The definitions of the NCI-CTCAE toxicity grading version 5.0 for each parameter are provided in Appendix 2 – Definition of NCI-CTCAE grading of this IAP.

Table 12 NCI-CTC Gradable parameters

Parameter	Parameter code	Name in NCI-CTC	Direction of abnormality
Biochemistry			
Alanine Aminotransferase	ALT	Alanine aminotransferase increased	High
Albumin	ALB	Hypoalbuminemia	Low

Parameter	Parameter code	Name in NCI-CTC	Direction of abnormality
Alkaline Phosphatase	ALP	Alkaline phosphatase increased	High
Amylase	AMYLASE	Serum amylase increased	High
Aspartate Aminotransferase	AST	Aspartate aminotransferase increased	High
Bilirubin total	BILI	Blood bilirubin increased	High
Calcium[a]	CA	Hypercalcemia/Hypocalcemia[a]	High/Low
Creatinine	CREAT	Creatinine increased	High
Glucose	GLUC	Hypoglycemia	Low
Lipase	LIPASET	Lipase increased	High
Potassium	K	Hyperkalemia/Hypokalemia	High/Low
Sodium	SODIUM	Hypernatremia/Hyponatremia	High/Low
Lactate dehydrogenase	LDH	Blood lactate dehydrogenase increased	High
Hematology			
Absolute lymphocyte	LYM	Lymphocyte count decreased/Lymphocyte count increased	High/Low
Absolute eosinophils	EOS	Eosinophilia	High
Absolute neutrophils	NEUT	Neutrophil count decreased	Low
Hemoglobin	HGB	Anemia/Hemoglobin increased	Low/High
Leukocytes (WBC)	WBC	Leukocytosis/White blood cell decreased	High/Low
Platelets count	PLAT	Platelet count decreased	Low
Coagulation			
Activated Partial Thromboplastin Time[b]	APTT	Activated partial thromboplastin time prolonged	High
Prothrombin International Normalized Ratio[b]	INR	INR increased	High
Activated PTT/Standard[b]	APTTSTND	Activated partial thromboplastin time prolonged	High

[a] based on corrected calcium (see Appendix 2 – Definition of NCI-CTCAE grading).
[b] reported on the "Coagulation" eCRF page.

For calcium, CTCAE grading is based on corrected calcium. Corrected calcium is calculated from albumin and calcium as follows based on the International System of Units (SI):

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

For **WBC differential counts** (neutrophil, lymphocyte 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)

Laboratory parameters with NCI-CTC grades not available

Table 13 Non-NCI-CTC Gradable Parameters

Parameter (LBTEST)	Parameter code
Biochemistry	
Bilirubin direct	BILDIR
Bilirubin Indirect	BILIND
Chloride	CL
C-Reactive Protein	CRP
Total Protein	PROT
Urea	UREA
Urea Nitrogen	UREAN
Hematology	
Absolute Basophils	BASO
Absolute Monocytes	MONO
Absolute Reticulocytes	RETI
Basophils/Leukocytes	BASOLE
Eosinophils/Leukocytes	EOSLE
Erythrocytes (RBC)	RBC
Hematocrit	HCT
Lymphocytes/Leukocytes	LYMLE
Mean Corpuscular Hemoglobin	MCH
Mean Corpuscular HGB Concentration	MCHC
Mean Corpuscular Volume	MCV
Monocytes/Leukocytes	MONOLE
Neutrophils/Leukocytes	NEUTLE
Reticulocytes/Erythrocytes	RETIRBC
Coagulation	
Prothrombin Time/Standard Prothrombin Time*	PTS

* reported on the "Coagulation" eCRF page.

For all non-gradable parameters, the following summaries will be displayed:

- Number and percentage of participants by lowest on-treatment value (classified as normal, high, low)
- Number and percentage of participants by highest on-treatment value (classified as normal, high, low)
- Shift from baseline to highest/lowest on-treatment value (classified as normal, high, low)

Separate listings of hematology (including coagulation) and biochemistry will be created. Each listing will include: participant identifier, age, sex, race, first dose date, last dose date, laboratory parameter (units), visit, date, SI value, change from baseline value, lower limit of normal, ULN,

indicator of normal range (low, normal, high), toxicity grade according to NCI-CTCAE (when applicable) and highest/lowest on treatment value flag. Baseline and post-baseline values after the on-treatment period will be flagged. These listings will be sorted by participant identifier, parameter and laboratory measurement date.

Liver function tests

ALT, AST and total bilirubin are used to assess possible drug induced liver toxicity. The ratios of test result over ULN will be calculated and classified for these three parameters during the on-treatment period.

The number and percentage of participants within each of the following liver function categories during on-treatment period will be described:

- ALT < 3×ULN, ALT ≥ 3×ULN, ALT ≥ 5×ULN, ALT ≥ 10×ULN, ALT ≥ 20×ULN
- AST < 3×ULN, AST ≥ 3×ULN, AST ≥ 5×ULN, AST ≥ 10×ULN, AST ≥ 20×ULN
- (ALT and AST) < 3×ULN, (ALT or AST) ≥ 3×ULN, (ALT or AST) ≥ 5×ULN, (ALT or AST) ≥ 10×ULN, (ALT or AST) ≥ 20×ULN
- Total Bilirubin (TBILI) < 2×ULN, TBILI ≥ 2×ULN
- Concurrent ALT ≥ 3×ULN and TBILI ≥ 2×ULN
- Concurrent AST ≥ 3×ULN and TBILI ≥ 2×ULN
- Concurrent (ALT or AST) ≥ 3×ULN and TBILI ≥ 2×ULN
- Concurrent (ALT or AST) ≥ 3×ULN and TBILI ≥ 2×ULN and ALP > 2×ULN
- Concurrent (ALT or AST) ≥ 3×ULN and TBILI ≥ 2×ULN and ALP ≤ 2×ULN or missing

Concurrent measurements are those occurring on the same date.

Categories will be cumulative, i.e., a participant with an elevation of AST ≥ 10 × ULN will also appear in the categories ≥ 5 × ULN and ≥ 3 × ULN.

A plot of peak ALT versus peak total bilirubin, both relative to the ULN will be provided. This eDISH plot (evaluation of drug-induced serious hepatotoxicity) will be divided into 4 quadrants by the lines through ALT ≥ 3×ULN and total bilirubin ≥ 2×ULN. The left lower quadrant is then considered normal or insignificant elevations in liver chemistries, the upper quadrants indicate participants with possible Gilbert's cholestasis; the right upper quadrant are the possible Hy's Law participants; the right lower quadrant is possible Temple's Corollary (participants with ALT ≥ 3×ULN but not satisfying Hy's Law). Same plot will be provided for AST.

In addition, 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.

Urinalysis/urinalysis microscopic evaluation, hormonal tests, tuberculosis test and serology

All test results for urinalysis /urinalysis microscopic evaluation, hormonal tests, serum and serology parameters will also be listed in dedicated listings:

- Urinalysis parameters:
    - Urinalysis full parameters: physical appearance, pH, specific gravity, protein, glucose, ketones, nitrite, leukocyte esterase, blood, urobilinogen, bilirubin, color
    - Urinalysis microscopic parameters: erythrocytes (RBC), leukocytes (WBC), epithelial cells, bacteria, crystals, casts
- Tuberculosis parameters: M. tuberculosis IFN Gamma Response by T-SPOT ELISPOT, Tuberculosis skin test (TST), M. tuberculosis IFN Gamma Response by QuantiFERON TB Gold ELISA
- Hormonal parameters: thyroxine free (Free T4), thyrotropin (Thyroid-Stimulating Hormone; TSH)
- Serology parameters: hepatitis B virus (HBV), hepatitis C virus (HCV), hepatitis B DNA, hepatitis C RNA

Pregnancy test

Results for pregnancy and post-menopausal status as collected on the "Pregnancy Test" eCRF page will also be listed:

- Pregnancy parameters (serum or highly sensitive urine human chorionic gonadotropin [hCG])
- Post-menopausal status: FSH and estradiol parameters

15.4 Vital Signs

All vital sign parameters from the on-treatment period, except oxygen saturation, 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.

The following potentially clinically significant abnormalities will be summarized:

- $\leq$ 95 mmHg and decrease from baseline $\geq$ 20 mmHg in systolic blood pressure
- $\geq$ 140 mmHg and increase from baseline $\geq$ 20 mmHg in systolic blood pressure
- $\leq$ 45 mmHg and decrease from baseline $\geq$ 20 mmHg in diastolic blood pressure
- $\geq$ 90 mmHg and increase from baseline $\geq$ 20 mmHg in diastolic blood pressure
- $\leq$ 50 beats/min and decrease from baseline $\geq$ 20 beats/min in pulse rate
- $\geq$ 100 beats/min and increase from baseline $\geq$ 20 beats/min in pulse rate

CCI [redacted]

- ≤ 20 breaths/min and decrease from baseline ≥ 5 breaths/min in respiratory rate
- ≥ 20 breaths/min and increase from baseline ≥ 5 breaths/min in respiratory rate
- ≥ 10% weight decrease
- ≥ 10% weight increase

A listing of vital signs will be provided including participant identifier, age, sex, race, vital sign parameter, visit, date, time, value, unit and change from baseline. Baseline values and post-baseline values collected after the on-treatment period will be flagged.

15.5 Other Safety or Tolerability Evaluations

ECG

Single 12-lead ECGs will be obtained at screening and repeated during the study if clinically indicated using an ECG machine that automatically calculates the heart rate and measures PR, QRS, QT, QTc Intervals, and ECG parameters will be derived from the data collected on the "Electrocardiogram" eCRF page.

A listing of ECG values will be provided including participant identifier, age, sex, race, ECG parameter and unit, visit, ECG date, value, change from baseline. Baseline values and post-baseline values collected after the on-treatment period will be flagged. Qualitative ECG results will also be provided in the listing.

ECOG Performance Status (ECOG PS)

The ECOG PS will be derived from the data collected on the "ECOG Performance Status" eCRF page.

The ECOG shift from baseline to the highest score during the on-treatment period will be summarized.

ECOG PS will also be presented in a listing at each timepoint.

16 Analyses of Other Endpoints

16.1 Pharmacokinetics

The analyses described in this section will be performed by the Clinical PK/PD group (CPK) of Translational Medicine, Merck Healthcare KGaA, Darmstadt, Germany, or by a Contract Research Organization selected by the Sponsor. Pharmacokinetic listings and individual data will be presented based on the SAF. Summaries and statistical analyses will be based on the PKAS.

Pharmacokinetic concentrations/parameters refer to bintrafusp alfa concentrations/PK parameters.

16.1.1 Missing/non-quantifiable PK Data Handling

Concentrations below the lower limit of assay quantification

Pharmacokinetic concentrations below the lower limit of quantification will be set to zero for calculating parameters and descriptive statistics.

Deviations, missing concentrations, and anomalous values

There will be no imputation of missing data. Concentrations will be set to missing in summary tables if the value is reported as no result. Pharmacokinetic concentrations which are erroneous due to a protocol violation or event (as defined in section 10.2) may be excluded from the PK analysis if agreed upon prior to performing a statistical analysis. Any other PK concentrations that appear implausible to the Pharmacokineticist/PK Data Analyst must not be excluded from the analysis. Any implausible data will be documented in the CSR.

If a PK parameter cannot be derived from a participant's concentration data, the parameter will be coded as NC (not calculated). (Note that NC values will not be generated beyond the day that a participant discontinues the treatment). For statistical analyses, PK parameters coded as NC will be set to missing.

If an individual participant has a known biased estimate of a PK parameter (due for example to a deviation from the assigned dose level), this participant/value will be excluded from the descriptive statistics, and instead the result will be listed in a separate table.

Relevant decisions on participant inclusion in the PKAS will be made before database lock in the Database Review Meeting (DRM).

16.1.2 Descriptive PK Analysis

Presentation of PK Concentration Data

A by-participant listing will present PK sample times, time deviations, and concentrations based on the SAF. Concentrations will be reported with the same precision as the source data.

CCI [redacted]

Presentation of PK Parameter Data

The PK parameters listed below will be taken directly from the observed bintrafusp alfa concentration-time data.

C_{trough}	The concentration observed immediately before next dosing (corresponding to pre-dose or trough concentration for multiple dosing).
C_{EOI}	The concentration observed immediately at the end of infusion.

Individual PK parameters will be listed by nominal study day based on the SAF. Individual PK parameters will be reported with the same precision as the source data.

Pharmacokinetic parameter data will be presented in tables and descriptively summarized by nominal study day using: n, Mean, StD, coefficient of variation (CV%), Min, Med, Max, geometric mean (GeoMean), the geometric coefficient of variation (GeoCV) and the 95% CI for the GeoMean (LCL 95% GM, UCL 95% GM). Summaries will be based on the PKAS.

Additional table(s) will summarize with further stratification by ADA subsets ever positive and never positive.

In export datasets (e.g. from WinNonlin®), 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 parameter data:

- Mean, Min, Med, Max, GeoMean, 95% CI: 3 significant digits
- StD: 4 significant digits
- CV%, GeoCV%: 1 decimal place

Individual PK C_{trough} and C_{EOI} values will be plotted versus actual study day on a linear scale, for all participants. Individual data will be presented based on the SAF.

Arithmetic mean (±StD) and median C_{trough} and C_{EOI} will be plotted versus nominal study day on a linear scale. Summaries will be based on the PKAS.

Additional figure(s) will be presented with further stratification by ADA subsets ever positive and never positive as long as both subsets consist of 5 or more participants.

16.2 Pharmacodynamics

Not applicable.

16.3 Immunogenicity

Analysis Sets: IMM

The ADA results will be derived for each visit based on the algorithm in Table 14.

Table 14 Algorithm for the Derivation of ADA Results

Sample Screen Result	Confirmatory	Titer	ADA Result
Negative	NA	NA	Negative
NR	NA	NA	NR
Positive	Negative	NA	Negative
Positive	NR	NA	NR
Positive	Positive	Number	Number
Positive	Positive	NR	Positive-TNR

NR = no result, NA = not applicable, TNR = titer no result.

Note that samples collected during and after the on-treatment period (e.g. safety follow-up) including the ones collected during the reinitiation period will be included in the analysis. Negative, number, or positive-TNR are valid results while number and positive-TNR are considered as positive. Participants will be characterized into different categories based on the criteria in Table 15.

Table 15 Participants Characterized based on ADA Results

Category	Definition	Participant at Risk (Denominator for Incidence)
Never positive	No positive results at any time point	Number of participants with at least one valid result at any time point
Ever positive	At least one positive result at any time point, including baseline	Number of participants with at least one valid result at any time point
Pre-existing	A positive ADA result prior to treatment with bintrafusp alfa	Number of participants with valid baseline result
Treatment boosted	A positive ADA result prior to treatment with bintrafusp alfa and at least one post-baseline value with ADA titer result ≥ 8*baseline ADA titer result	Number of participants with valid baseline result and at least one valid post-baseline result
Treatment emergent	Not positive prior to treatment with bintrafusp alfa and with at least one positive post-baseline result	Number of participants with at least one valid post-baseline result and without positive baseline result (including missing, NR)
Transient positive	If treatment emergent participants have: - a single positive evaluation and last assessment not positive or - duration between first and last positive result < 16 weeks and last assessment not positive	Number of participants with at least one valid post-baseline result and without positive baseline result (including missing, NR)
Persistent positive	If treatment emergent participants have duration between first and last positive result ≥ 16 weeks or a positive evaluation at the last assessment	Number of participants with at least one valid post-baseline result and without positive baseline result (including missing, NR)

CCI
CCI [redacted]

The following analysis will be described:

- The frequency and percentage of each ADA category, as defined in Table 15, will be tabulated
- Evaluation of potential effect of ADA status on bintrafusp alfa safety (see section 15.1.1)
- Evaluation of potential effect of ADA status on bintrafusp alfa efficacy
    - for best overall response according to RECIST 1.1 as adjudicated by IRC (see section 14.1.1)
    - for progression-free survival according to RECIST 1.1 as adjudicated by IRC (see section 14.4)
    - for overall survival (see section 14.7)

Potential effect of ADA on safety and efficacy will be evaluated on ADA positive status (ever positive, never positive).

A listing of ADA results from ever positive participants will be provided with the following: participant ID, age, sex, race, ADA categories status, visit, date of assessment and results of screening, confirmatory and titer values.

CCI [redacted]

CCI [redacted]

CCI [redacted]

CCI [redacted]

17 References

Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, Dancey J, Arbuck S, Gwyther S, Mooney M, Rubinstein L, Shankar L, Dodd L, Kaplan R, Lacombe D.,Verweij J. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer, National Cancer Institute of Canada-Clinical Trials Group, 10 Stuart Street, Queen's University, Kingston, Ontario, Canada. eeisenhauer@ctg.queensu.ca, 2009;45:228-47

18 Appendices

18.1 Appendix 1 – Definition of Important Protocol Deviations

Refer to: csp-ms200647-0020-iap-appendix-1-ipd-list-v2.docx

18.2 Appendix 2 – Definition of NCI-CTCAE grading

Refer to: csp-ms200647-0020-iap-appendix-2-NCI-CTCAE-grading-v2.xlsx

CCI
CCI

Signature Page for VV-CLIN-282833 v1.0

Approval	PPD 02-Jan-2022 17:13:20 GMT+0000

Signature Page for VV-CLIN-282833 v1.0

CCI
CCI