Statistical Analysis Plan for End of Dose Escalation Phase

Clinical Trial Protocol Identification No.	EMR 100070-001
Title:	A Phase I, open-label, multiple-ascending dose trial to investigate the safety, tolerability, pharmacokinetics, biological and clinical activity of avelumab (MSB0010718C) in subjects with metastatic or locally advanced solid tumor and expansion to selected indications
Trial Phase	Phase I
Investigational Medicinal Product(s)	Avelumab
Clinical Trial Protocol Version	23 July 2015/Version 13.0 (Amendment 12)
Statistical Analysis Plan Author	Jiali Tang
Statistical Analysis Plan Date and Version	12 October 2015/Final Version 1.0
Statistical Analysis Plan Reviewers	PPD

Statistical Analysis Plan Reviewers

PPD

1 Signature Page

Statistical Analysis Plan: EMR 100070-001

A Phase I, open-label, multiple-ascending dose trial to investigate the safety, tolerability, pharmacokinetics, biological and clinical activity of avelumab (MSB0010718C) in subjects with metastatic or locally advanced solid tumor and expansion to selected indications

PPD

_____ _____

Signature Date of Signature

Trial Biostatistician

PPD

PPD

_____ _____

Signature Date of Signature

Senior Biostatistics Review

PPD

PPD

_____ _____

Signature Date of Signature

Programming Team Lead

PPD

PPD

_____ _____
Signature Date of Signature
Project Statistician
Merck KGaA
Frankfurter Str. 250
64293 Darmstadt, Germany

PPD

_____ _____
Signature Date of Signature
Medical Responsible
EMD Serono, Inc.
45 Middlesex Turnpike
Billerica, MA 01821, USA

2 Table of Contents

3 List of Abbreviations and Definition of Terms

ACC	Adrenocortical Carcinoma
ACTH	Adrenocorticotropic Hormone
ADR	Adverse Drug Reaction
AE	Adverse Event
ALP	Alkaline Phosphatase
ALT	Alanine Aminotransferase
ANA	Anti-Nuclear Antibody
ANC	Absolute Neutrophils Count
ANCA	Anti-Neutrophil Cytoplasmic Antibody
AST	Aspartate Aminotransferase
aPTT	Activated Partial Thromboplastin Time
ATC	Anatomical Therapeutic Chemical
BMI	Body Mass Index
BOR	Best Overall Response
CALCIO	Corrected Calcium and Ionized Calcium
CI	Confidence Interval
Cmin	Trough Concentration
CPK	Creatine Kinase
CR	Complete Response
CRC	Colorectal Cancer
CRPC	Castrate Resistant Prostate Cancer
CTCAE	Common Terminology Criteria for Adverse Events
CTR	Clinical Trial Report
CV	Coefficient of Variation

DLT	Dose Limiting Toxicity
DRM	Database Review Meeting
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCcr	Estimated Creatinine Clearance Rate
eCRF	Electronic Case Report Form
eDISH	Evaluation of Drug-Induced Serious Hepatotoxicity
Free T4	Free Thyroxine
GEJ	Gastroesophageal Junction
GeoCV	Geometric Coefficient of Variation
GeoMean	Geometric Mean
GGT	Gamma Glutamyl Transferase
CCI [redacted]	
HB	Hemoglobin
HNSCC	Head and Neck Squamous Cell Carcinoma
HR	Heart Rate
ICF	Informed Consent Form
IERC	Independent Endpoint Review Committee
IMP	Investigational Medical Product
irAE	Immune Related Adverse Event
irBOR	immune-related BOR
irCR	immune-related Complete Response
irPD	immune-related Progressive Disease
irPR	immune-related Partial Response

IRR	Infusion Related Reaction
irRC	immune-related Response Criteria
irSD	immune-related Stable Disease
i.v.	Intravenous
LDH	Lactate Dehydrogenase
LLOQ	Lower Limit of Quantification
LogFC	Log Fold Change from Baseline
MBC	Metastatic Breast Cancer
MCH	Mean Corpuscular Hemoglobin
MCHC	Mean Corpuscular Hemoglobin Concentration
MCV	Mean Corpuscular Volume
MedDRA	Medical Dictionary for Regulatory Activities
MTD	Maximum Tolerated Dose
NCA	Non-Compartment Analysis
NCI	National Cancer Institute
NE	Not Assessable
NOS	Not Otherwise Specified
NSCLC	Non-Small Cell Lung Cancer
PBMC	Peripheral Blood Mononuclear Cell
PCSA	Potentially Clinically Significant Abnormalities
PD	Progressive Disease
PD-1	Programmed Death 1
PD-L1	Programmed Death Ligand 1
RF	Rheumatoid Factor

PK	Pharmacokinetics
PLT	Platelet Count
PR	Partial Response
PT	Preferred Term
QTc	Corrected QT
QTcB	QT Interval Correction based on Bazett's Formula
QTcF	QT Interval Correction based on Fridericia's Formula
RBC	Red Blood Cell
RCC	Renal Cell Carcinoma
RECIST	Response Evaluation Criteria in Solid Tumors
SAE	Serious Adverse Event
SAF	Safety
SAP	Statistical Analysis Plan
SD	Stable Disease
SEM	Standard Error of Mean
SMC	Safety Monitoring Committee
SOC	System Organ Class
TEAE	Treatment Emergent Adverse Event
TNM	Tumor Node Metastasis Classification of Malignant Tumors
TSH	Thyroid-Stimulating Hormone
ULN	Upper Limit of Normal

4 Modification History

Unique Identifier for SAP Version	Date of SAP Version	Author	Changes from the Previous Version
Final v1.0	12Oct2015	Jiali Tang	NA

5 Purpose of the Statistical Analysis Plan

The purpose of this statistical analysis plan (SAP) is to document technical and detailed specifications for the final analysis of data collected for the Dose Escalation phase under protocol EMR 100070-001. Results of the analyses described in this SAP will be included in the Clinical Trial Report (CTR). Additionally, the planned analyses identified in this SAP will be included in regulatory submissions or future manuscripts. Any post-hoc, or unplanned analyses performed to provide results for inclusion in a CTR but not identified in this prospective SAP will be clearly identified in the CTR.

6 Summary of Clinical Trial Features

Trial Objectives	**Primary** • To assess the safety and tolerability of avelumab and to determine the maximum tolerated dose (MTD) of avelumab in subjects with metastatic or locally advanced solid tumors. **Secondary** • To characterize the pharmacokinetic (PK) profile of avelumab and to correlate exposure with target occupancy. • To evaluate the immunogenicity of avelumab and to correlate it to exposure and biological activity. • To assess the best overall response (BOR) according to Response Evaluation Criteria in Solid Tumors version 1.1 (RECIST 1.1, 1). • To assess the immune-related BOR (irBOR) using the modified Immune-Related Response Criteria (irRC, 2), derived from RECIST 1.1. • To evaluate biological responses to avelumab in blood/serum. • To characterize changes in soluble factors (e.g., cytokine profiles, soluble programmed death 1 [PD-1], and soluble PD-L1) and immune cell profiling (e.g., natural killer cells, neutrophils, lymphocytes). The objectives associated with expansion phase were included into the expansion SAP.

Trial Endpoints	**Primary** • Occurrence of dose limiting toxicities (DLTs) during the first 3 weeks of treatment in the dose escalation part. **Secondary** • Number, severity, and duration of treatment-emergent adverse events (TEAEs) according to the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) v4.0. • Number, severity, and duration of treatment-related adverse events (AEs) according to NCI-CTCAE v4.0. • PK profile. • irBOR and BOR according to modified irRC and to RECIST 1.1 criteria, respectively, per investigator assessment. • Pharmacodynamic profile • Serum titers of anti-avelumab antibodies. The endpoints associated with expansion phase were included into the expansion SAP.
Trial Design	This is a Phase I, open-label, dose-escalation trial with consecutive parallel group expansion in non-small cell lung cancer (NSCLC), metastatic breast cancer (MBC), gastric and GEJ cancer, colorectal cancer (CRC), castrate resistant prostate cancer (CRPC), melanoma, ovarian cancer, head and neck squamous cell carcinoma (HNSCC), adrenocortical carcinoma (ACC), renal cell carcinoma (RCC), mesothelioma, and urothelial carcinoma. **Dose escalation phase** Cohorts of 3 subjects with metastatic or locally advanced solid tumors will receive avelumab at escalating dose levels. At each dose level, subjects will receive avelumab as a 1-hour intravenous (i.v.) infusion once every 2 weeks until confirmed progression, unacceptable toxicity, or any criterion for withdrawal from the trial or the investigational medicinal product (IMP) occurs. Subjects who have experienced a confirmed complete response (CR) should be treated for a maximum of 24 months after confirmation, at the discretion of the investigator. If the investigator believes that a subject may benefit from treatment beyond 24 months, it may be permissible after discussion with the sponsor. Subjects who experienced a CR and have already stopped treatment can resume treatment with avelumab at the same dose and schedule. Subjects re-initiating treatment should be assessed according to the Schedule of Assessments. **Dose escalation (3+3 design) will be performed at the following dose levels**

- 1.0 mg/kg
- 3.0 mg/kg
- 10.0 mg/kg

Once 1 subject has experienced a DLT at a dose below 10.0 mg/kg, dose escalation will be reduced as described in Section 5.1.4.2 of Protocol Amendment 12.

The first subject of each cohort should be observed for 16 days (i.e., 48 hours after the second dose) for the occurrence of DLT before the second subject is administered the trial medication. Thereafter, within each cohort of the dose escalation phase, subjects may only be consecutively dosed with an interval of at least 48 hours. However, after 3 subjects have been treated at 10 mg/kg and no DLT has been observed, the other 3 subjects required to complete this cohort can be enrolled without sequential dosing (i.e., not required to wait until 48 hours). If no more than one DLT has been observed in these 6 subjects, the safety of 10 mg/kg will have been established.

Each subject will stay on the dose level assigned at trial entry (only adaptations for weight changes are needed as described in Section 5.1.7.1 of Protocol Amendment 12).

The decision to escalate to the next dose level will be based on safety assessments after all subjects of a cohort have reached Day 21 (DLT evaluation period). In order to assess the safety of avelumab, a safety monitoring committee (SMC), responsible for dose escalation decisions, will be established.

Once the MTD or maximum dose to be investigated is reached, the respective dose level cohort will be filled to a total of 6 subjects. Once the dose of 10 mg/kg is established as safe, 10 additional subjects at 3 mg/kg and 10 mg/kg each may be enrolled, for the purpose of generating additional safety, PK and receptor occupancy data, if agreed with the SMC.

Once 6 subjects treated at 10 mg/kg have completed the DLT observation period and the safety of 10 mg/kg is established, a dose level of 15 mg/kg (if 1 DLT was observed) or 20 mg/kg (if no DLT was observed) dosing every 2 weeks will be initiated. In this 20 mg/kg dose level, the safety, PK, receptor occupancy, and pharmacodynamic activity of the IMP will be evaluated using the methodology that was used for the other cohorts. Accrual in these dose levels will be completed using a "3+3" method, the same methodology that was used for the completion of the previous dose levels. Once the safety of the 15 and/or 20 mg dose level has been established (i.e., no more than 1 DLT out of 6 subjects treated), up to 15 additional subjects will be enrolled at 15 or 20 mg/kg without sequential dosing (i.e., not required to wait until

	48 hours between 2 subjects). This additional cohort will have the purpose of generating safety data, PK data and receptor occupancy data at the respective dose. **Definition of DLT** A DLT is specifically defined as any Grade ≥ 3 toxicity that is possibly, probably, or definitely related to avelumab, occurring during the DLT evaluation period (21 days after administration of avelumab), except for any of the following: • Grade 3 infusion-related reaction resolving within 6 hours and controlled with medical management. • Transient (≤ 6 hours) Grade 3 flu-like symptoms or fever, which is controlled with medical management. • Transient (≤ 24 hours) Grade 3 fatigue, local reactions, headache, nausea, emesis that resolves to ≤ Grade 1. • Grade 3 diarrhea, Grade 3 skin toxicity, or Grade 3 liver function test (ALT, AST, or GGT) increase that resolves to ≤ Grade 1 in less than 7 days after medical management (e.g., immunosuppressant treatment) has been initiated. • Single laboratory values out of normal range that are unlikely related to trial treatment according to the investigator, do not have any clinical correlate, and resolve to ≤ Grade 1 within 7 days with adequate medical management. • Tumor flare phenomenon defined as local pain, irritation, or rash localized at sites of known or suspected tumor. DLTs requiring treatment discontinuation are described in Section 5.1.7.2 of Protocol Amendment 12. Subjects who do not complete the DLT observation period for reasons other than a DLT will be replaced. **Expansion phase** This section was included into the Expansion SAP.
Number of Subjects	Dose escalation phase: 18 up to 60 subjects. The final sample size, however, will depend on the total number of dose levels to be tested, and subject replacement for DLT evaluation if applicable.
Trial Product	Avelumab will be administered as a 1-hour (-10 minutes /+20 minutes, i.e., 50-80 minutes) i.v. infusion. Subjects will receive avelumab once every 2

	weeks until confirmed progression, unacceptable toxicity, or any criterion for withdrawal from the trial or IMP occurs. The dose of avelumab will be calculated based on the weight of the subject determined on the day prior to or the day of each drug administration. Premedication with an antihistamine and with paracetamol (acetaminophen) approximately 30 to 60 minutes prior to each dose of avelumab is mandatory (for example, 25-50 mg diphenhydramine and 500-650 mg paracetamol [acetaminophen] i.v. or oral equivalent). This regimen may be modified based on local treatment standards and guidelines, as appropriate. Immediate access to intensive care unit or equivalent environment and appropriate medical therapy (including i.v. epinephrine, corticosteroids, antihistamines, bronchodilators, and oxygen) must be in place for use in the treatment of potential infusion-related reactions. Infusion of avelumab will be stopped in case of Grade ≥2 infusion-related, allergic, or anaphylactic reactions (according to NCI-CTCAE v4.0). Following avelumab infusions, subjects must be observed for 2 hours post infusion for potential infusion-related reactions. Relevant clinical laboratory results essential for patient management decisions (hematology, biochemistry, liver function tests) must be available and reviewed before administration of avelumab.
Treatment and Trial Duration	The planned treatment duration is until unacceptable toxicity, or any criterion for withdrawal from the trial or IMP occurs.
Schedule of Visits	**For dose escalation cohorts** 1. Screening (days -18 to first treatment) 2. Treatment phase Visits will take place on Days 1, 2, 3, 15, 29, 43, and every 2 weeks thereafter. 3. Discontinuation visit and end-of-treatment visit All subjects who discontinue trial treatment prematurely for an AE should have a full safety evaluation at the time of discontinuation of trial treatment (discontinuation visit). In addition, all subjects will have an end-of-treatment visit scheduled 4 weeks after the last administration of avelumab but before any new therapy is started, if possible. 4. Post-treatment follow-up

	All subjects will have a subsequent visit scheduled 10 weeks after the last administration of avelumab. The visit will include a full assessment of safety parameters. AEs will be documented until the end of treatment visit. After the end of treatment visit only treatment related AEs have to be documented until the post-treatment safety follow-up visit. Subjects with a serious AE ongoing at the post-treatment safety follow-up must be monitored and followed up by the investigator until stabilization or until the outcome is known, unless the subject is documented as "lost to follow-up". Subjects without progressive disease at the end-of-treatment visit will be followed up for disease progression (CT / MRI scans every 12 weeks) up to 1 year. In addition, subjects will be followed for any AE suspected to be related to trial treatment, especially for the occurrence of new autoimmune events up to 3 months after the last dose of avelumab. After the end-of-treatment visit, subjects will be followed quarterly for survival (including assessment of any further tumor therapy). The survival follow-up will continue until 1 year after the last subject receives the last dose of avelumab. **For dose expansion cohorts** This section was included into the Expansion SAP.
Randomization and Blinding	Not applicable.

CCI [redacted]

8 Overview of Planned Analyses

There will be one planned analysis after the end of dose escalation phase. This SAP will only address the analysis for the dose escalation phase. Additional SAPs for SMC analyses and for all other analyses (interim, primary, and final) for expansion cohorts have been developed separately.

9 Changes to the Planned Analyses in the Clinical Trial Protocol

The data cut-off date was changed to Nov 20, 2015 from 6 weeks after the last subject of the dose escalation part has received its first administration of avelumab in order to provide a more comprehensive evaluation of the dose escalation phase in the first interim CTR for this study.

10 Analysis Sets

The following analysis sets will be defined for the dose escalation part:

- Screening analysis set: all subjects who signed the informed consent form (ICF).
- DLT analysis set: all subjects with data used for implementing the dose-escalation schedule. These subjects should:
  - Have received all study treatment administrations during the DLT evaluation period (first 21 days of treatment) or
  - Have stopped treatment because of DLTs during the DLT evaluation period.
- Safety (SAF) analysis set: all subjects who have received at least 1 dose of trial treatment.
- PK analysis set: all subjects who receive at least one dose of avelumab, and provide at least one measurable post-dose concentration.

Unless otherwise stated, DLT analysis set will be used for the summary of the incidence of DLT, and PK analysis set will be used for the PK analysis, other analyses including safety analyses will be based on SAF analysis set.

11 General Specifications for Statistical Analyses

- The endpoint data will be summarized by dose level and overall, unless otherwise specified.
- All data recorded during the study will be presented in individual data listings. Data collected after re-initiation of treatment will not be included for safety summaries except for disposition, but will be included in the data listings.
- All data will be evaluated as observed, and no imputation method for missing values will be used, unless otherwise specified.
- Duration will be calculated as stop date – start date + 1, unless otherwise specified.

- The first day (Day 1) of study treatment is defined as the day of the first administration of avelumab, unless otherwise stated. The last dose date of study treatment is defined as the day of the last administration of avelumab, prior to the re-initiation of study treatment if applicable.
- Baseline is defined as the last non-missing observation prior to the administration of first dose of trial treatment. Additionally, baseline for HR and QT/corrected QT (QTc) assessments will be derived from the visit where both HR and QT are not missing. If duplicate or triplicate electrocardiograms (ECGs) are collected, baseline for each ECG measurement is the average of the pre-dose replicate measurements on the baseline day. QT interval correction based Fridericia's or Bazett's formula (QTcF/QTcB) will be derived based on HR and QT. The average of the replicate measurements should be determined after the derivation of the individual parameter at each time point.
- If the laboratory assessments are not done for scheduled visit but they are available for unscheduled visit on Day 1 from a different laboratory, the unscheduled visit will be included for the derivation of baseline; if there are multiple non-missing assessments on Day 1, the assessment from scheduled visit will be used for the derivation of baseline.
- On-treatment period is defined as the time from the first dose of study treatment to the last dose date + 30 days, or the earliest date of subsequent anti-cancer drug therapy – 1 day, whichever occurs first. If the earliest date of subsequent anti-cancer drug therapy is a partial date and only day is missing, it will be imputed as the last day of the month. If both day and month are missing, no imputation should be performed. The imputed date will be used for defining on-treatment period as well as confirming immune-related progressive disease (irPD).
- All statistical analyses will be performed using SAS® Version 9.1.3 or higher.
- There will be no difference between scheduled and un-scheduled visits except for by-visit analysis of safety data and baseline derivation.
- The assignments of visit windows are described in Table 1 for the purpose of by-visit analyses of safety data:
  - Baseline will be derived as described above.
  - No visit windowing will be performed at discontinuation, end of treatment, or safety follow-up visits for laboratory, vital sign, and ECG data, and 2hr post dose assessment on Week 1 Day 1 for ECG data. Instead, the earliest non-missing observation among the unscheduled or scheduled assessments for each visit (discontinuation, end of treatment, or safety follow-up) will be used for the analysis. For 2hr post dose assessment on Week 1 Day 1 ECG data, the earliest non-missing observation on Week 1 Day 1 will be used for the analysis.
  - Scheduled and unscheduled assessments are included for visit windowing. Assessments on or after re-initiation of treatment are not be included for visit windowing.
  - If there are multiple assessments for any specified visit and some of them are from scheduled visits, the assessment from scheduled visit with the closest distance to the planned study day will be used for analysis.
  - If there are multiple assessments for any specified visit and none of them are from scheduled visits, the assessment with the closest distance to the planned study day will be used for analysis.

- If there are two or more unscheduled assessments with same distance to the planned study day such as (-1/+1 day), the assessment prior to the planned study day such as -1 day will be used for windowing.
- There is no difference for visit windowing between tests from core serum chemistry panel and tests from full serum chemistry panel. Some subjects had non-core serum chemistry tests assessed at the scheduled visits only intended for core serum chemistry. Multiple protocol amendments is also taken into the consideration, as the full serum chemistry and hematology panels were assessed weekly until week 7 and bi-weekly thereafter prior to the approval of Protocol Amendment 7.
- For ECG assessment associated with study drug dose, only assessments where time point (prior to infusion or 2 hr after infusion) are not missing will be considered for the analysis.

Table 1. Visit Window Definition for Safety Assessment

Assigned Study Day (Inclusive)		**Planned Study Day (AWTARGET)**	**Analysis Visit (N) (AVISITN)**	**Analysis Visit (AVISIT)**	**Assessment**
From (AWLO)	**To (AWHI)**				
~	1		1	Baseline	Lab, Vital Sign, ECG
1	1	1	2	Week 1 Day 1*	ECG
5	11	8	3	Week 2 Day 5-11	Lab, Vital Sign
12	18	15	4	Week 3 Day 12-18	Lab, Vital Sign
5	18	15	4	Week 3 Day 5-18	ECG
19	25	22	5	Week 4 Day 19-25	Lab, Vital Sign
19	25	22	5	Week 4 Day 19-25	ECG
26	32	29	6	Week 5 Day 26-32	Lab, Vital Sign
26	36	29	6	Week 5 Day 26-36	ECG
33	39	36	7	Week 6 Day 33-39	Lab, Vital Sign
40	50	43	8	Week 7 Day 40-50	Lab, Vital Sign
37	50	43	8	Week 7 Day 37-50	ECG
51	64	57	10	Week 9 Day 51-64	Lab, Vital Sign, ECG
65	78	71	12	Week 11 Day 65-78	Lab, Vital Sign, ECG
79	92	85	14	Week 13 Day 79-92	Lab, Vital Sign
79	106	85	14	Week 13 Day 79-106	ECG
93	106	99	16	Week 15 Day 93-106	Lab, Vital Sign

107	120	113	18	Week 17 Day 107-120	Lab, Vital Sign
121	134	127	20	Week 19 Day 121-134	Lab, Vital Sign
107	148	127	20	Week 19 Day 107-148	ECG
135	148	141	22	Week 21 Day 135-148	Lab, Vital Sign
149	162	155	24	Week 23 Day 149-162	Lab, Vital Sign
163	176	169	26	Week 25 Day 163-176	Lab, Vital Sign
149	190	169	26	Week 25 Day 149-190	ECG
177	190	183	28	Week 27 Day 177-190	Lab, Vital Sign
191	204	197	30	Week 29 Day 191-204	Lab, Vital Sign
205	218	211	32	Week 31 Day 205-218	Lab, Vital Sign
191	232	211	32	Week 31 Day 191-232	ECG

* only applies to 2 hr post dose.

- Presentation of continuous and qualitative variables:
  - Continuous variables will be summarized using descriptive statistics i.e., number of non-missing values and number of missing values, [i.e. n (missing)], mean, median, standard deviation, minimum, maximum and first and third quartile (Q1 and Q3). Confidence interval (CI) may be estimated for some of the endpoints, if appropriate.
  - Qualitative variables will be summarized by counts and percentages. Unless otherwise stated, the calculation of proportions will include the missing category. Therefore counts of missing observations will be included in the denominator and presented as a separate category.

12 Trial Subjects

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

12.1 Disposition of Subjects and Discontinuations

Summary of analysis sets will be tabulated using frequency and percentage by dose level (if applicable) and overall based on all the subjects who signed ICF, the number of subjects in SAF analysis set will be used as the denominator:

- All subjects who signed ICF
- Number of subjects in the SAF analysis set
- Number of subjects in the DLT analysis set

One table will provide the reasons for permanent discontinuation of study treatment and for end of study as collected on the Treatment Termination, and End of Study (if data is available) electric

Case Report Forms (eCRFs), respectively. The number and percentage of subjects in each disposition category will be presented in the table based on the SAF analysis set:

- Number of subjects in the SAF analysis set
- Number of subjects still on treatment
- Number of subjects off-treatment
- Reasons off-treatment
  - Adverse event
  - Lost to follow-up
  - Protocol non-compliance
  - Death
  - Disease progression
  - Withdrew consent
  - Other
- Number of subjects in follow-up
- Number of subjects off-study
- Reasons off-study
  - Study reached predefined end
  - Lost to follow-up
  - Death
  - Withdrew consent
  - Other

The listing of subject disposition will include all subjects who signed ICF (i.e. including screening failures). The listing will include the following information (if applicable): subject identifier, date of informed consent, included in the trial, reason for inclusion/exclusion, first/last dosing date, reason off-treatment, date and reason off-study, flags for SAF and DLT analysis sets, and be sorted by dose level and subject identifier.

A secondary listing for reason for end of treatment due to AEs will also be provided. The listing will be restricted to the SAF analysis set subjects who discontinued study treatment for the primary reason of an AE, and will include the following information: subject identifier, first/last dosing date, date off-treatment, and the relevant AE system organ classes (SOCs), preferred terms (PTs) and AE relationship to the study treatment.

12.2 Protocol Deviations

12.2.1 Minor Protocol Deviation

A minor protocol deviation can be defined as any deviation from the study protocol that does not materially affect the safety of the subjects and/or the conduct of the study and/or its evaluation. An example of a minor protocol deviation would include a missed PK blood sample.

12.2.2 Major Protocol Deviation

A major and/or serious protocol deviation (otherwise known as a protocol violation) is one that materially affects the safety of the subjects and/or the evaluation of primary or key secondary endpoints of the study.

Current ICH and EU GCP guidelines list the major protocol deviations that must be listed in the clinical report. These include:

- subjects that are dosed on the study despite not satisfying the inclusion criteria;
- subjects that develop withdrawal criteria whilst on the study but are not withdrawn;
- subjects that receive the wrong treatment or an incorrect dose;
- subjects that receive an excluded concomitant medication.

Major protocol deviations will be based upon clinical database and determined for all subjects by either medical review processes or programming based on the inclusion/exclusion criteria or other criteria presented in the protocol. The results will be included into SDTM, if identified by means of medical review. The ADAM datasets will include both, those identified by medical review and those identified by programming.

Major protocol deviations are specified in Appendix 20.1, and will be reported in a data listing.

13 Demographics and Other Baseline Characteristics

The demographics and other baseline characteristics will be summarized on the SAF analysis set.

13.1 Demographics

The demographics and baseline characteristics table will include descriptive statistics for the following variables:

- Age (in years)
- Age category (<65/≥65 years)
  - 65-74,
  - 75-84,
  - ≥ 85

- Sex
- Race
- Height (cm)
- Weight (kg)
- Body Mass Index (BMI) (kg/m^2)
- Eastern Cooperative Oncology Group (ECOG) Performance Status
- Nicotine use status (Never used/ Regular user/ Occasional user/ Former user)

Baseline weight and height will be the last non-missing values prior to the first dose of study treatment from the Vital Signs eCRF page while baseline ECOG will be derived from the data collected on the ECOG eCRF page. Nicotine use status will be extracted from Nicotine Consumption eCRF page.

Age and BMI will be derived as:

- Age(year) = (Date of Informed Consent – Date of Birth + 1)/365.25.
  - In case of missing day only: Age (years) = (year/month of given informed consent – year/month of birth)/12
  - In case only year of birth is given: Age (years) = (year of given informed consent - year of birth)
- BMI(kg/m^2) = Weight(kg)/[Height(m)]^2.

The integer part of the calculated age will be used for report purpose. The above table will be generated by dose level and overall.

The listing of demographics and baseline characteristics will include the following information: dose level, subject identifier, age, sex, race, height (cm), weight (kg), BMI (kg/m^2), and ECOG.

The listing of nicotine consumption will be produced with the following data: nicotine use status, frequency of nicotine use, start/end date of nicotine consumption, nicotine consumption habit, and duration of consumption (years).

13.2 Medical History

Medical history will be coded using the latest available version of Medical Dictionary for Regulatory Activities (MedDRA). Medical history will be summarized as the numbers and percentages of subjects by MedDRA PT as event category and MedDRA SOC as summary category, and sorted by SOC and PT in alphabetical order. Each subject will be counted only once within each PT or SOC.

Listing of medical history data by subject will include coded terms and all the relevant data fields as collected on the Medical History eCRF page.

13.3 Disease History

Disease histories are collected on Disease History eCRF page. Partial date will be imputed as described in the Section 18.1.

The disease history table will include descriptive statistics for the following variables:

- Site of primary tumor
  - Lung, not otherwise specified (NOS)
  - Breast, NOS
  - ….
- Tumor histopathologic/ cytologic type of interest
  - Adrenocortical
  - Anal
  - Bladder
  - ……
  - Other
- Time since first diagnosis (years), defined as (the first dosing date – the date of first diagnosis)/365.25
- Time since metastatic or locally advanced disease (months), defined as (the first dosing date – the date of first occurrence of metastatic or locally advanced disease)/30.4375
- Time since last disease progression (months), defined as (the first dosing date - the date of last progression of disease)/30.4375
- Tumor Node Metastasis Classification of Malignant Tumors (TNM) at initial diagnosis
  - TX
  - T0
  - N1
  - ….
- TNM at study entry
  - TX
  - T0
  - N1
  - …

Listing of disease history will be provided with all relevant data (primary site of tumor, sub-site, initial diagnosis date, first occurrence of metastatic or locally advanced disease, date of last disease

progression, TNM classification, tumor histopathologic/ cytologic type of interest etc.) and derived variables used in the above table.

14 Prior and Concomitant Medications/Procedures

Prior and concomitant anti-cancer therapy / other medications will be coded using the latest available version of WHO Drug Dictionary, and summarized by dose level and overall based on SAF analysis set.

14.1 Prior Anti-Cancer Therapies/Procedures

The prior anti-cancer treatments and procedures are collected under the Prior Anti-Cancer Drug Therapies Details, Prior Anti-Cancer Radiotherapies Details and Prior Anti-Cancer Surgeries Details eCRF pages.

The overall summary of presence of prior anti-cancer treatments table will include: the number and percentage of subjects by type of treatment, i.e.

- Number of subjects with at least one type of prior anti-cancer treatment
- Number of subjects with at least one prior anti-cancer drug therapy
- Number of subjects with at least one prior anti-cancer radiotherapy
- Number of subjects with at least one prior anti-cancer surgery

Summary of prior anti-cancer drug therapy will include the following variables:

- Number of subjects with at least one prior anti-cancer drug therapy
- Number of any prior anti-cancer therapy lines: missing / 1 / 2 / 3 / ≥4
- Number of any prior anti-cancer therapy lines as continuous variable
- Number of prior anti-cancer therapy lines for metastatic or locally advanced disease: missing / 0 / 1 / 2 / 3 / ≥4. If the intent of therapy is metastatic, locally advanced, or palliative, it will be counted into therapy lines for metastatic or locally advanced disease.
- Number of prior anti-cancer therapy lines for metastatic or locally advanced disease as continuous variable
- Type of prior anti-cancer therapy: chemotherapy / antibody therapy / kinase inhibitors / hormonal therapy / vaccines / bone marrow transplant / lymphocyte infusion / other
- Intent of therapy: adjuvant / neo-adjuvant / metastatic / locally advanced / palliative
- Best response: CR / partial response (PR) / progressive disease (PD) / stable disease (SD) / unknown / not assessable (NE) / not applicable. Best response is derived from the last treatment regimen.

The prior anti-cancer drugs will also be extensively detailed with the number and percentage of subjects by ATC classification level 2 and PT in a table. A subject will be counted only once within a given drug class and within a given drug name, even if he/she received the same medication at

different times. If any prior anti-cancer medication is classified into multiple Anatomical Therapeutic Chemical (ATC) classes, the medication will be summarized separately under each of these ATC classes. In case any specific medication does not have ATC classification level 2 coded term, it will be summarized under "Unavailable ATC classification" category. The summary will be sorted on decreasing frequency of drug class and decreasing frequency of drug name in a given drug class, based on the incidence in the "Overall" column. In case of equal frequency regarding drug class (respectively drug name), alphabetical order will be used.

The listings of prior anti-cancer treatments and procedures will also be provided: a) listing of prior anti-cancer drug therapies, b) listing of prior anti-cancer radiotherapies, and c) listing of prior anti-cancer surgeries. These will include subject identifier and all the relevant collected data-fields on the corresponding eCRF pages.

14.2 Prior and Concomitant Medications/Procedures

Prior and concomitant procedures are collected on the Concomitant Procedures Details eCRF page. Prior and concomitant medications are collected on the Concomitant Medications Details eCRF page.

Medications started prior to first dose date of study treatment and continued into the on-treatment period as well as those started during on-treatment period are referred to as concomitant medications. Prior medications are defined as the medications started and stopped prior to the first dose date of trial treatment. Post medications are defined as any medications started after on-treatment period.

Summary of concomitant medications will include the number and percentage of subjects by ATC classification level 2 and PT. A subject will be counted only once within a given drug class and within a given drug name, even if he/she received the same medication at different times. If any concomitant medication is classified into multiple ATC classes, the medication will be summarize separately under each of these ATC classes. In case any specific medication does not have ATC classification level 2 coded term, it will be summarized under "Unavailable ATC classification" category. The summary of concomitant medications will be sorted on decreasing frequency of drug class and decreasing frequency of drug name in a given drug class, based on the incidence in the "Overall" column. In case of equal frequency regarding drug class (respectively drug name), alphabetical order will be used.

Prior and concomitant medication data will be listed from the Concomitant Medications eCRF page. The following variables will be included in the prior and concomitant medication listing: subject identifier, prior/ concomitant/ post medication, and all corresponding data fields on the corresponding eCRF page.

Prior and concomitant procedures data will be listed from the Concomitant Procedures Details eCRF page. Subject identifier and all collected data-fields on the corresponding eCRF page will be included in the listing.

14.3 Subsequent Anti-Cancer Therapies/Procedures

Anti-cancer treatment after discontinuation of treatment with avelumab will be provided in a data listing with data retrieved from Anti-Cancer Treatment After Discontinuation, Radiotherapy After Discontinuation, and Surgery After Discontinuation eCRF pages.

15 Treatment Compliance and Exposure

Analysis of exposure will be based on the calculated actual dose levels (total dose administered/weight, mg/kg). The last available weight of the subject on or prior to the day of dosing will be used for the calculation.

The summary of treatment exposure and compliance based on the SAF analysis set will include the following variables per subject (a cycle refers to the planned dosing interval of two weeks):

- Treatment duration (in weeks), defined as (the last dose date – the first dose date + 14)/7
- Number of administrations as continuous variable
- Cumulative dose (mg/kg), defined as sum of actual dose levels
- Dose intensity (mg/kg/cycle), defined as cumulative dose (mg/kg) / (0.5 * treatment duration (week))
- Relative dose intensity (%), defined as actual dose intensity (mg/kg/cycle) * 100/ planned dose level (mg/kg/cycle).
- Relative dose intensity by the following categories:
  - >0.9
  - >0.8-0.9
  - <=0.8

Individual relative dose intensity (%) is calculated as actual dose (kg)/ planned dose (kg) × 100 for each infusion of study medication. A dose reduction is defined as actual non-zero dose < 90% of planned dose, or individual relative dose intensity < 90%. A table based on SAF analysis set will be prepared to summarize the number and percentage of subjects with at least one dose reduction, and a breakdown by the number of dose reductions (1/2/3/≥4).

Per protocol, avelumab will be administered as 1-hour i.v. infusion. Subjects will receive the study treatment once every 2 weeks. Dose delays will be grouped into the following categories based on the deviation of the actual to the planned treatment administration day (relative to the previous non-zero dose date): no delay (including 1-2 days delays), 3-6 days delay, 7 or more days delay. For example, if one subject receives the study drug on day 1, then the next study drug administration date will be on day 15; however, if the subject receives the study drug at day 16 or 17, this is considered as 'no delay'. Any zero dose prior to the last treatment administration is considered as a dose interruption.

The summary of dose delays will be based on the SAF analysis set and include the following categories:

- No delay
- 3-6 days delay
- 7 or more days delay

The categorization is based on the maximum length of delay, i.e. the worst case of delay if subjects have multiple dose delays.

A listing of study drug administration will include subject identifier, study day, # of days relative to prior treatment, infusion rate, most recent body weight prior to infusion, actual/planned dose, batch ID, dose reduction/ dose delay or interruption, and other relevant information collected on the Cohort Treatment MSB0010718C Administration Details eCRF page.

A listing of treatment exposure and compliance will include subject identifier, assigned dose level, and above derived variables summarized in the tables.

In order to mitigate infusion-related reactions, a premedication regimen of 25 to 50 mg diphenhydramine and 650 mg acetaminophen (i.v. or oral equivalent) is mandatory 30 to 60 minutes prior to each dose of study drug starting on January 29, 2014. The use of premedication will be summarized as numbers of subjects with 0, 1, 2, 3, 4 doses among the first 4 treatment administrations that were administered without pre-medication. For example: if a subject discontinued after 3 doses, and 2 of them were administered with premedication, the number for that subject would be 1.

A listing of premedication will include subject identifier, reported medication term, date/time of premedication, and dose (unit). A listing containing subject identifier, visit, and unique study drug batch ID will also be created.

16 Endpoint Evaluation

The subsections in this section include specifications for analyzing clinical trial endpoints specified in the Clinical Trial Protocol to meet the trial objectives for dose escalation, as well as any endpoints not identified in the Clinical Trial Protocol. Endpoints identified for expansion part in the Clinical Trial Protocol are included into other SAPs.

16.1 Primary Endpoint Analysis

The primary endpoint in the dose escalation part of this trial is the occurrence of DLTs during the DLT evaluation period.

A DLT is defined as a Grade >=3 adverse drug reaction according to the NCI-CTCAE v4.0, occurring during the DLT evaluation period confirmed by the SMC to be relevant for the study drug treatment based on the criteria specified in Protocol.

The summary of DLTs will include the following variables by dose level and overall for the DLT analysis set:

- the number and percentage of subjects who experienced a DLT during the DLT evaluation period
- number of DLT per subject, 1 /2 /≥3
- DLT by SOC and PT

The number and percentage of subjects who experienced a TEAE by SOC and PT during the DLT evaluation period will also be summarized by dose level and overall for the DLT analysis set. A listing of DLT AEs will include dose level, subject identifier, SOC/PT, and all relevant variables, such as start/stop date, toxicity grade, relationship to the study treatment, outcome etc., from AE eCRF page.

The MTD or maximum dose will be determined according to the dose-escalation plan described in section 5.1.4.2 of trial protocol. The MTD or maximum dose is defined as the highest dose level at which no more than one out of 6 subjects treated in a cohort and evaluable for DLT determination experiences a DLT.

16.2 Secondary Endpoint Analyses

The Secondary endpoints for dose escalation are:

- Number, severity, and duration of TEAEs according to the NCI-Common Terminology Criteria for Adverse Events (CTCAE) v4.0.
- Number, severity, and duration of treatment-related AEs according to NCI-CTCAE v4.0.
- PK profile.
- irBOR and BOR according to modified irRC and to RECIST 1.1, respectively, per investigator assessment.
- Pharmacodynamic profile.
- Serum titers of anti-avelumab antibodies.

TEAE related endpoints will be discussed in the Section 17.1. Other secondary endpoints such as cytokine profile, receptor occupancy, PK profile, immunogenicity will be addressed in Sections 16.3, 16.4, 16.5, or 16.6, respectively. No efficacy summary will be performed due to small sample size and heterogeneous population in dose escalation part, but the endpoints at subject level will be derived as described in this section and displayed in the listings.

Efficacy evaluations include lesion assessments for target, non-target, and new lesions, tumor response according to RECIST 1.1 (1) and irRC (2). The confirmed BOR is defined as the best response obtained among all tumor assessment visits after first dose of study treatment until documented disease progression, taking into account the following requirements for confirmation:

- CR or PR needs to be confirmed at a subsequent tumor assessment, preferably at the regularly scheduled 6-week assessment interval, but no sooner than 4 weeks after the initial documentation of CR or PR.
- PR can be confirmed at an assessment later than the next assessment after the initial documentation of PR.

- The minimum duration for a BOR of SD is defined as at least 37 days after start of study treatment accounting for permitted deviations from the tumor assessment visit schedule.

Table 2 summarized the derivation rules for the BOR when confirmation from subsequent assessment is needed (1). It is reasonable to consider a subject with time point response of PR-SD-PR, PR-NE-PR or CR-NE-CR as a confirmed response as long as the second CR or PR is at least 28 days away from the first time point.

Table 2 Best Overall Response when Confirmation of CR/PR Required

Initial overall response	Subsequent overall response	Confirmed time point overall response
CR	CR	CR provided subsequent CR is >=28 days away from the first time point
CR	PR	SD provided minimum criteria for SD duration met; otherwise, PD
CR	SD	SD provided minimum criteria for SD duration met; otherwise, PD
CR	PD	SD provided minimum criteria for SD duration met; otherwise, PD
CR	NE	SD provided minimum criteria for SD duration met, otherwise, NE
PR	CR	PR provided subsequent CR is >=28 days away from the first time point
PR	PR	PR provided subsequent PR is >=28 days away from the first time point
PR	SD	SD
PR	PD	SD provided minimum criteria for SD duration met, otherwise, PD
PR	NE	SD provided minimum criteria for SD duration met, otherwise, NE
SD	Any	SD provided minimum criteria for SD duration met, otherwise, NE
PD	Any	PD
NE	NE	NE

CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease, and NE = Not assessable.
If a subject only has a response value of NE or the only response value is SD and is within 36 days of the first dose date, the best response will be NE.

For irBOR, immune-related complete response (irCR), immune-related partial response (irPR), and irPD need to be confirmed by a second, consecutive assessment at least 4 weeks apart as described in Table 3 (1, 2). It is reasonable to consider a subject with time point response of irPR-irSD-irPR, irPR-NE-irPR or irCR-NE-irCR as a confirmed response as long as the second irCR or irPR is at least 28 days away from the first time point. irPD is also considered to be confirmed if the following event occurs:

- If subject is assessed with time point response of irPD-NE-irPD as long as the second irPD is at least 28 days away from the first time point; or
- If subject dies within 84 days after the initial observation of irPD; or

- If subject receives subsequent anti-cancer drug therapy within 84 days after the initial observation of irPD; or
- If subject experiences clinical deterioration as assessed by investigator and recorded as reason for treatment discontinuation prior to or within 84 days after the assessment of irPD.

In case a subject with a confirmed CR relapses within 1 year after stopping treatment, one re-initiation of treatment is allowed according to Protocol. If the subject has irPD assessed at the same visit as the PD prior to the initiation of treatment, the irPD will be considered as confirmed. The minimum duration of immune-related stable disease (irSD) is defined as at least 37 days from Day 1.

Table 3 Immune-related BOR when Confirmation of irCR, irPR, irPD Required

Initial overall response	Subsequent overall response	Confirmed time point overall response
irCR	irCR	irCR provided subsequent irCR is >=28 days away from the first time point
irCR	irPR	irSD provided minimum criteria for irSD duration met; otherwise, irPD
irCR	irSD	irSD provided minimum criteria for irSD duration met; otherwise, irPD
irCR	irPD	irSD provided minimum criteria for irSD duration met; otherwise, irPD
irCR	NE	irSD provided minimum criteria for irSD duration met, otherwise, NE
irPR	irCR	irPR provided subsequent irCR is >=28 days away from the first time point
irPR	irPR	irPR provided subsequent irPR is >=28 days away from the first time point
irPR	irSD	irSD
irPR	irPD	irSD provided minimum criteria for irSD duration met, otherwise, irPD
irPR	NE	irSD provided minimum criteria for irSD duration met, otherwise, NE
irSD	Any	irSD provided minimum criteria for irSD duration met, otherwise, NE
irPD	irPD	irPD provided subsequent irPD is >=28 days away from the first time point, death or take subsequent anti-cancer drug therapy within 84 days after the first time point, or clinical deterioration.
irPD	Missing	irPD if subject dies or takes subsequent anti-cancer drug therapy within 84 days after the initial observation of irPD, or clinical deterioration, otherwise, NE
NE	NE	NE

irCR = immune-related complete response, irPR = immune-related partial response, irSD = immune-related stable disease, irPD = immune-related progressive disease, and NE = Not evaluable.

If a subject only has a response value of NE or the only response value is irSD and is within 36 days of the first dose date, the best response will be NE

These evaluations will be presented in data listings with detailed information, such as lesion type, site, size, time point tumor response by type collected per eCRF pages as well as BOR and irBOR, for all the subjects from SAF analysis set.

16.3 Cytokine Profile

The following parameters are assessed for the cytokines profile based on SAF analysis set:

- Rantes, Interleukin-17, IL-6, MCP-1, IL-1beta, IL-2, IL-4, IL-5, IL-10, IL-12p70, IL-13, TNF-Alpha, IFN-Gamma.

These parameters will be summarized by visit and dose level using the following variables:

- Absolute value
- Log fold change from baseline

If values are indicated to be below the lower limit of quantification (LLOQ), they will be imputed as 0.5* LLOQ for summary analysis. If multiple measurements per time point per subject exist, a single signal value will be derived by averaging.

Log fold change from baseline (logFC) is defined as

- logFC (cytokine at time t) = log(cytokine signal at time t) – log(cytokine signal at baseline):

The cytokine data will be displayed graphically by dose level for each parameter:

- Boxplot of absolute value by visit (logarithmic and raw scale)
- Boxplot of log fold change from baseline by visit
- Line plot of per-subject data

The cytokine listing will be produced with relevant variables such as dose level, subject identifier, visit, parameter names and results.

16.4 Receptor Occupancy

The receptor occupancy was assessed by flow cytometry on PBMC (peripheral blood mononuclear cell) samples collected on Day 1 before start of the infusion, at 4 and 48 hrs after the start of infusion, and before the start of each infusion on Days 15, 29, 43, and 85, if applicable. Quality of PBMCs such as cell viability was determined after thawing, ≥ 85% cell viability is required for reliable receptor occupancy assessment. Additional or alternative criteria might be used if a review of the data indicates that this is necessary. Data generated from samples with inadequate quality will be excluded from the analysis.

This data will be summarized using descriptive statistics (n, mean, standard deviation, etc.) by visit and by dose level and overall based on SAF analysis set. Line plots will be created to display the data graphically across visit per subject by dose level, or to display the mean and standard deviation across visit by dose level.

A listing of receptor occupancy will include variables such as subject identifier, visit, date of assessment, test parameter, and values etc.

16.5 Pharmacokinetic Profile

16.5.1 Missing PK Data

All PK analyses will be based on the PK analysis set.

- **Concentrations below the limit of assay quantitation**

PK concentrations below LLOQ are taken as zero for descriptive statistics.

PK concentrations below LLOQ, which are before the last quantifiable data point, will be taken as zero for calculating the AUC of single dose profiles. Concentration below LLOQ, which occur after the last quantifiable data point, will not be considered in the calculation of the terminal first order rate constant (λ_z).

- **Deviations, missing concentrations and anomalous values**

Concentrations will be set to missing in summary tables if the value is reported as no result.

If a PK parameter cannot be derived from a subject'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 subject discontinues the treatment). For statistical analyses (i.e. analysis of variance), PK parameters coded as NC will be set to missing.

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

Relevant decisions on subject inclusion in the PK analysis set will be made before database lock in the Database Review Meeting (DRM).

16.5.2 Descriptive PK Analysis

Avelumab concentrations in serum will be displayed in listing and descriptively summarized by treatment, cohort, day and nominal time using the number of non-missing observations, number of missing observations, arithmetic mean, standard deviation, coefficient of variation (CV%), standard error of the mean (SEM), minimum, median and maximum. The pre-dose samples will be considered as if they had been taken simultaneously with the administration. Additional tables listing trough concentrations (Cmin) per cohort/day/administration will also be provided, and summarized descriptively using the aforementioned statistics.

Descriptive statistics of PK parameters will additionally show the geometric mean (GeoMean), the geometric coefficient of variation (GeoCV) and the 95% confidence interval for the GeoMean.

For the calculation of descriptive statistics, values as presented in the data listings will be used. Minimum and maximum will be presented to the same decimal precision as collected. Mean, median and GeoMean will be presented to one decimal place more than the precision of the data collected. SD and SEM will be presented to two decimal places more than the precision of the data collected. The coefficient of variation (CV%), GeoCV and the 95% CI will be reported to 1 decimal place.

16.5.3 Pharmacokinetic Non-Compartment Analysis

Pharmacokinetic parameters for avelumab will be evaluated according to standard non-compartmental analysis (NCA) by the PK/pharmacodynamic data processing group of QPD, Merck Serono, Darmstadt, Germany, using the validated software tool Phoenix®/WinNonlin 6.3® (or later).

The PK parameters listed below will be calculated for avelumab using the actual time elapsed from dosing (or using scheduled time if actual time is not available).

Cmax	Maximum observed concentration in serum
tmax	Time to reach Cmax
AUC0-t	Area under the serum concentration-time curve from time zero to the last quantifiable concentration
Cmin	The minimum observed concentration during a complete dosing interval (serum trough concentration)

When applicable the following parameters will also be calculated:

AUC0-∞	Area under the serum concentration-time curve from time zero extrapolated to infinity. $AUC_{0-\infty} = AUC_{0-t} + AUC_{extra}$, where $AUC_{extra} = C_{last\ pred} / \lambda_z$
AUCextra	The AUC from time tlast extrapolated to infinity given as a percentage of AUC0-∞
t1/2	Apparent terminal half-life, calculated by ln2/ λz
λz	Apparent terminal elimination rate constant determined by log-linear regression analysis of the measured serum concentrations of the terminal log-linear phase.
CL	Apparent total body clearance of drug from serum. CL = Dose i.v / $AUC_{0-\infty}$
Vz	Apparent volume of distribution during terminal phase. VZ = Dose/($AUC_{0-\infty} * \lambda_z$)
AUC0-t/Dose	Dose-Normalized AUC0-t. Normalized using actual dose.
Cmax/Dose	Dose-Normalized Cmax, Normalized using actual dose.

The following PK parameters will be calculated for diagnostic purposes and listed, but will not be summarized:

- The time interval (h) of the log-linear regression to determine λz (t½, Interval).
- Number of data points (t½, N) included in the log-linear regression analysis to determine λz.
- Goodness of fit statistic (Rsq) for calculation of λz.
- Percentage of AUC0-∞ obtained by extrapolation (%AUCextra), calculated by (1－[AUC0-t/AUC0-∞])×100. If %AUCextra is greater than 20%, AUC0-∞, CL/f, and Vz/f will be included in the Phoenix® WinNonlin® parameter outputs but will be flagged.

The regression analysis (determination of λz) should contain as many data points as possible (but excluding Cmax) and has to include concentration data from at least 3 different time points, consistent with the assessment of a straight line (the terminal elimination phase) on the log-transformed scale. The observation period over which the regression line is estimated should be at least twofold the resulting t½ itself.

The calculation of the AUCs will be performed using the mixed log-linear trapezoidal method (linear up, log down). Extrapolated areas will always be computed using the predicted last

concentration that is estimated using the linear regression from terminal rate constant determination.

If AUCextra>20% then AUC0-∞ and all related parameters will be excluded from descriptive statistics.

16.5.4 Figures

Individual serum concentration-time profiles (linear and semi-logarithmic scales) will be plotted by dose level. Mean serum concentrations per administration will also be plotted on linear (±SD) and semi-logarithmic scales using schedule time points.

Dose proportionality will be evaluated for avelumab using data from subjects with a full PK profile, and will be presented graphically as follows:

- Boxplots of dose-normalized PK parameters (AUC0-t/Dose, AUC0-∞/Dose and Cmax/Dose) by dose level.
- Scatterplots with regression lines for individual AUC0-t, AUC0-∞ and Cmax values (where applicable) and dose normalized AUC0-t, AUC0-∞ and Cmax versus dose on a linear scale.

CCI

CCI

[redacted]

17 Safety Evaluation

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

All safety analyses will be performed using the SAF analysis set, unless otherwise specified.

All safety parameters will be summarized by dose level and overall, unless otherwise stated.

17.1 Adverse Events

The severity of adverse events will be graded using the NCI-CTCAE, version 4.0 except where CTCAE grades are missing. No imputation of missing grades will be performed. Adverse events will be coded according to the latest available version of MedDRA.

- **Treatment Emergent Adverse Event:** TEAEs are those events with onset dates occurring during the on-treatment period for the first time, or if the worsening of an event is during the on-treatment period as defined in Section 11.
- **Related Adverse Event:** adverse events with relationship to study treatment (relationship with study treatment = related) reported by the investigator and those of unknown relationship (i.e. no answer to the question "Relationship with study treatment").
- **Serious Adverse Event (SAE):** serious adverse events (as recorded on the AE eCRF page, serious adverse event = yes).
- **Adverse Event Leading to Treatment Discontinuation:** adverse events leading to permanent discontinuation of study treatment (as recorded on the AE eCRF page, action taken with study treatment = drug withdrawn).
- **Adverse Event Leading to Death:** adverse event leading to death (as recorded on the AE eCRF page, outcome = fatal or toxicity grade = 5).
- **Infusion Related Reaction (IRR):** IRRs are identified by 3 PTs (infusion related reaction, drug hypersensitivity, anaphylactic reaction) with onset on study drug dosing date (not prior to infusion of study drug) or the day following study drug infusion.
- **Immune Related Adverse Event (irAE):** irAEs are identified according to a pre-specified search list of MedDRA PTs, documented in a version-controlled repository maintained by the Sponsor and finalized for analysis prior to database lock.

AEs will be summarized using the MedDRA PT as event category and MedDRA primary SOC as summary category. All AE tables will be restricted to TEAEs unless otherwise specified.

Each subject will be counted only once within each PT or SOC and recording period. If a subject experiences more than one AE within a PT or SOC for the same recording period, only the AE with the strongest relationship or the worst severity, as appropriate, will be included in the summaries of relationship and severity. AEs with missing classifications regarding relationship to study treatment and start date greater or equal to start of study treatment will be considered as related to study treatment. AEs with missing toxicity grade will be counted into 'any grade' in the summarization by toxicity grade.

The AE tables will include the number and percentage of subjects with at least one TEAE, by MedDRA primary SOC (sorted by decreasing SOC frequencies in the overall column) and by PT (sorted by decreasing PT frequencies in overall column within SOC), unless otherwise stated.

17.1.1 All Adverse Events

All adverse events will be tabulated in the following tables by dose level and overall.

- The overall summary of AEs table will include the following variables:
  - TEAEs
  - TEAEs, grade ≥ 3
  - Related TEAEs
  - Related TEAEs, grade ≥ 3
  - TEAEs leading to permanent treatment discontinuation
  - Related TEAEs leading to permanent treatment discontinuation
  - Serious TEAEs
  - Related serious TEAEs
  - TEAEs leading to death
  - Related TEAEs leading to death
  - Treatment emergent irAEs
  - Related treatment emergent irAEs
  - Treatment emergent IRRs
  - Related treatment emergent IRRs
- Incidence of TEAEs by SOC, PT, and worst grade
- Incidence of related TEAEs by SOC, PT, and worst grade
- Incidence of TEAEs by SOC and PT: displaying in separate columns All TEAEs / Related TEAEs / Grade ≥3 TEAEs / Related Grade ≥3 TEAEs
- Incidence of TEAEs leading to death by SOC and PT
- Incidence of related TEAEs leading to death by SOC and PT
- Incidence of non-serious TEAEs by SOC and PT

Listing of AEs including all relevant information such as AE SOC/PT, start/stop date, toxicity grade, relationship to the study treatment, action taken with study treatment, and outcome etc., will be provided. A separate listing of TEAEs started or worsened after on-treatment period will also be provided.

17.1.2 Adverse Events Leading to Treatment Discontinuation

Following summary tables will be produced:

- Incidence of TEAEs leading to permanent treatment discontinuation by SOC and PT
- Incidence of related TEAEs leading to permanent treatment discontinuation by SOC and PT

The listing of AEs leading to permanent treatment discontinuation will also be provided with the relevant information such as AE SOC/PT, start/stop date, toxicity grade, relationship to the study treatment, and outcome etc.

17.1.3 Serious Adverse Events

Following summary tables will be produced:

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

The listings of SAEs will also be provided with the relevant information such as AE SOC/PT, start/stop date, toxicity grade, relationship to the study treatment, action taken with study treatment, and outcome etc.

17.1.4 Infusion Related Reactions

IRRs will be summarized by the follow variables:

- Number of subjects with at least one event by the worst toxicity grade (grade 1/ grade 2/ grade 3/ grade 4/ grade 5/ missing grade)
- Number of subjects with IRR leading to permanent treatment discontinuation
- Time related to first onset (infusion 1/ infusion 2/ infusion 3/ infusion 4 or later). The events should be assigned to the actual drug infusions that the subject received, not to the planned dates. An IRR is assigned to a drug infusion if its onset is at the same date (but not before dosing) or the following day of drug infusion.
- Number of subjects with at least one event by the worst toxicity grade that occurred in the presence of premedication (grade 1/ grade 2/ grade 3/ grade 4/ grade 5/ missing grade). The denominator will be the number of subjects with at least one dose administered in the presence of premedication. The maximum toxicity will be derived among those IRRs that occurred in the presence of premedication.
- Number of subjects with at least one event by the worst toxicity grade that occurred in the absence of premedication (grade 1/ grade 2/ grade 3/ grade 4/ grade 5/ missing grade). The denominator will be the number of subjects with at least one dose administered in the absence of premedication. The maximum toxicity will be derived among those IRRs that occurred in the absence of premedication.

The listing of IRRs will be provided with the relevant information such as AE SOC/PT, start/stop date, toxicity grade, relationship to the study treatment, action taken with study treatment, outcome, premedication, and study medication batch ID etc.

17.1.5 Immune Related Adverse Events

irAEs will be summarized using the following variables:

- The overall summary of treatment emergent irAEs will include the following categories:
  - irAEs
  - Related irAEs
  - Grade ≥ 3 irAE
  - Related grade ≥ 3 irAE
  - irAEs leading to permanent treatment discontinuation
  - Related irAEs leading to permanent treatment discontinuation
  - Serious irAEs
  - Related serious irAEs
  - irAEs leading to death
  - Related irAEs leading to death
- Incidence of treatment emergent irAEs by SOC and PT.

A listing containing all the PTs used to identify irAEs and a listing containing all the irAEs in the study will be generated as well.

17.1.6 Subgroup Analysis of Adverse Events

Subgroup analyses of AEs (TEAEs, irAEs, SAEs, IRRs) based on immunogenicity responses will be performed using summary statistics for qualitative comparison only. The number and percentage of subjects with at least one TEAE, irAE, SAE, or IRR will be compared between HAHA ever positive versus HAHA never positive subjects, and between subjects with treatment emergent immunogenicity response and the rest of the subjects.

17.2 Deaths

All deaths will be tabulated and listed for SAF analysis set by dose level and overall. The deaths table will include the following information:

- Number of subjects who died
  - Primary reason for death
    - Disease progression
    - Adverse event related to study treatment
    - Adverse event not related to study treatment
    - Other
    - Unknown

- Number of subjects who died within 30 days of the last study treatment administration
  - Primary reason for death
    - Disease progression
    - Adverse event related to study treatment
    - Adverse event not related to study treatment
    - Other
    - Unknown

The listing of deaths will be provided with all the relevant information such as death date and reason for death. The death data will be ascertained from the dedicated Report of Death eCRF form.

17.3 Clinical Laboratory Evaluation

Laboratory abnormalities are classified according to NCI-CTCAE toxicity grading version 4.0 or based on normal ranges collected from laboratories. The toxicity grading is only related to the lab values itself and does not respect the non-numeric information as described in the CTC grading definition. CTCAE gradable parameters and associated toxicities are listed in Table 6.

Table 6 CTCAE Gradable Parameters and Associated Toxicities

Panel	Parameter (test)	Low direction toxicity	High direction toxicity
Hematology	Hemoglobin (HB)	anemia	hemoglobin increased
	Leukocytes	white blood cell decreased	
	Lymphocytes	lymphocyte count decreased	lymphocyte count increased
	Neutrophils/ absolute neutrophils count (ANC)	neutrophil count decreased	
	Platelet count (PLT)	platelet count decreased	
Chemistry	Albumin	hypoalbuminemia	
	Alkaline phosphatase (ALP)		alkaline phosphatase increased
	Alanine aminotransferase (ALT)		ALT increased
	Amylase		serum amylase increased
	Aspartate aminotransferase (AST)		AST increased
	Bilirubin (total)		blood bilirubin increased
	Cholesterol		cholesterol high
	Creatinine		creatinine increased
	Creatinine clearance	chronic kidney disease	
	Creatine kinase (CPK)		CPK increased
	Potassium	hypokalemia	hyperkalemia
	Sodium	hyponatremia	hypernatremia
	Magnesium	hypomagnesemia	hypermagnesemia
	Calcium	hypocalcemia	hypercalcemia

	Glucose	hypoglycemia	hyperglycemia
	Gamma glutamyl transferase (GGT)		GGT increased
	Lipase		lipase increased
	Phosphates	hypophosphatemia	
	Triglycerides		hypertriglyceridemia
Coagulation	Prothrombin time INR		INR increased
	Activated Partial thromboplastin time (aPTT)		aPTT prolonged

Grade 1 and 2 hypoglycemia/ hyperglycemia are based on fasting glucose, they will not be graded for this study because blood samples are taken from non-fasted subjects.

For calcium, CTCAE grading is based on corrected calcium and ionized calcium (CALCIO), if available. Corrected calcium is calculated from albumin and calcium as follows

Corrected calcium (mg/dL) = Calcium (mg/dL) – 0.8 [Albumin (g/dL) - 4].

Chronic kidney disease will be graded based on estimated creatinine clearance rate (eCcr, ml/min), which is derived using Cockcroft-Gault formula:

$$eCcr = \frac{(140 - Age) \times Weight(kg) \times Constant}{Serum\ Creatinine\ (\mu mol/L)}$$

where the constant is 1.23 for men and 1.04 for women.

The following are non-CTCAE gradable parameters collected in this study, their abnormalities are assessed as low, high, normal based on the comparison of observed values with normal ranges.

- Hematology: hematocrit, red blood cell (RBC), reticulocytes, mean corpuscular hemoglobin (MCH), mean corpuscular volume (MCV), mean corpuscular hemoglobin concentration (MCHC).
- Serum Chemistry: chlorine, C-reactive protein, lactate dehydrogenase (LDH), total protein, total urea, uric acid.
- Hormone: adrenocorticotropic hormone (ACTH), anti-nuclear antibody (ANA), anti-neutrophil cytoplasmic antibody (ANCA), rheumatoid factor (RF), free thyroxine (Free T4), thyroid-stimulating hormone (TSH).

Laboratory abnormalities will be summarized using the worst grade during the on-treatment period. For these parameters which are graded with two toxicities such as potassium (hypokalemia/ hyperkalemia), the toxicities will be summarized separately. Low direction toxicity (eg. 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 change and percent change from baseline will be derived for parameters with numeric results.

17.3.1 Hematology, Chemistry, and Hormone Parameters

CTCAE gradable parameters

The laboratory toxicities will be tabulated by the worst on-treatment CTCAE grade or the shift of CTCAE grade from baseline to worst grade during on-treatment period using descriptive statistics (count and percentage). The highest CTCAE grade during the on-treatment period is considered as the worst grade for the summary.

- The worst grade during the on-treatment period will be summarized considering only patients with post baseline laboratory samples: Laboratory tests by NCI-CTCAE grade (0, 1, 2, 3, 4, and missing).
- The shift table will summarize baseline CTCAE grade vs. the worst on-treatment CTCAE grade (grade = 0, 1, 2, 3, 4, missing).

Non-CTCAE gradable parameters

Hematology, chemistry, and hormone evaluations which can't be graded per CTCAE will be summarized as:

- Shift from baseline value (low, normal, high) to above normal during on-treatment period
- Shift from baseline value (low, normal, high) to below normal during on-treatment period

Quantitative data will also be examined for trends using descriptive statistics (n, missing, mean, SD, median, Q1, Q3, minimum, and maximum) of actual values, absolute changes and percent changes from baseline to each visit over time. This summarization will apply to hematology and chemistry parameters with numeric results assessed at baseline, post-baseline, discontinuation, end of treatment, and/or safety follow-up visits based on visit windows specified in Section 11.

The listings (hematology, chemistry, and hormone) will include all the laboratory parameters as available in the database with the relevant information such as visit, assessment date, parameter, value, normal ranges etc. Listings will be sorted by subject identifier, group variable, parameter, assessment date or visit.

Liver function parameters

ALT, AST, ALP, and total bilirubin are used individually or together to assess possible drug induced liver toxicity. The ratio of test result over upper limit of normal (ULN) for individual test or combined tests will be calculated and classified for these parameters during the on-treatment period.

Summary of liver function tests will include the following categories. The number and percentage of subjects with each of the following during the on-treatment period will be summarized by dose level and overall:

- AST ≥3*ULN / ≥5*ULN / ≥10*ULN / ≥20*ULN.
- ALT ≥3*ULN / ≥5*ULN / ≥10*ULN / ≥20*ULN.
- (ALT or AST) ≥ 3×ULN/ ≥ 5×ULN/ ≥ 10×ULN/ ≥ 20×ULN

- Total bilirubin ≥2*ULN.
- Concurrent ALT ≥ 3×ULN and TBILI ≥ 2×ULN
- Concurrent AST ≥ 3×ULN and TBILI ≥ 2×ULN
- (ALT or AST) ≥3*ULN concurrently with total bilirubin ≥2*ULN.
- (ALT or AST) ≥3*ULN concurrently with total bilirubin ≥2*ULN and ALP >2*ULN.
- (ALT or AST) ≥3*ULN concurrently with total bilirubin ≥2*ULN and (ALP ≤2*ULN or missing).

Concurrent measurements are those occurring on the same date.

Categories will be cumulative, i.e. a subject with an elevation of AST ≥10*ULN will also appear in the categories ≥5*ULN and ≥3*ULN. Liver function elevation and possible Hy's Law cases will be summarized using frequency and percentage by dose level and overall.

An evaluation of Drug-Induced Serious Hepatotoxicity (eDISH) plot will also be created to graphically display

- peak serum ALT(/ULN) vs. concurrent total bilirubin (/ULN) including reference lines at ALT >3×ULN and total bilirubin >2×ULN.
- peak serum AST(/ULN) vs. concurrent total bilirubin (/ULN).

Listing of subjects with ALT or AST ≥3*ULN or total bilirubin ≥2*ULN will include variables subject identifier, visit, date of collection, study day, parameter (ALT, AST, ALP, total bilirubin), result, unit, result/ULN, CTCAE grade.

17.3.2 Other laboratory parameters

All other parameters collected on the eCRF will be listed in dedicated listings presenting all corresponding collected data-fields on the eCRF.

- Coagulation: aPTT, prothrombin time INR
- Urinalysis: all urinalysis parameters
- Other parameters: such as immunology, soluble factor
- Pregnancy test
- Serology

17.4 Vital Signs

Summary of vital signs will be based on the SAF analysis set. The potentially clinically significant changes from baseline as below in vital signs will be derived and summarized with subject incidence and percentage during the on-treatment period:

- ≥ 10% weight increase

- ≥ 10% weight decreases
- ≤ 95 mmHg and decrease from baseline ≥ 20 mmHg in systolic blood pressure
- ≥ 160 mmHg and increased from baseline ≥ 20 mmHg in systolic blood pressure
- ≤ 45 mmHg and decrease from baseline ≥ 10 mmHg in diastolic blood pressure
- ≥ 110 mmHg and increased from baseline ≥ 10 mmHg in diastolic blood pressure
- ≤ 50 bpm and decrease from baseline ≥ 20 bpm in pulse rate
- ≥ 120 bpm and increase from baseline ≥ 20 bpm in pulse rate

Quantitative data will also be examined for trends using descriptive statistics (n, missing, mean, SD, median, Q1, Q3, minimum, and maximum) of actual values, absolute changes and percent changes from baseline to each visit based on visit windows specified in Section 11. This summarization will apply to weight, blood pressure, respiratory rate, pulse, and temperature assessed at baseline, post-baseline, discontinuation, end of treatment, and/or safety follow-up visits.

Data listing of all vital signs will be provided with all relevant information such as visit, assessment date, parameter, and results.

17.5 Other Safety or Tolerability Evaluations

The incidence and percentage of subjects with potentially clinically significant abnormalities (PCSA) for 12-lead ECG parameters will be summarized during the on-treatment period based on SAF analysis set. The PCSA criteria are provided in the Table 7.

Table 7 Potentially Clinically Significant Abnormalities Criteria for ECG

Test	Potentially Clinically Significant Abnormalities (PCSA) Criteria
Heart Rate (HR)	≤ 50 bpm and decrease from baseline ≥ 20 bpm ≥ 120 bpm and increased from baseline ≥ 20 bpm
PR Interval	≥ 220 ms and increase from baseline ≥ 20 ms
QRS	≥ 120 ms
QTcF, QTcB absolute	interval >450 - ≤ 480 ms interval >480 - ≤ 500 ms interval >500 ms
QTcF, QTcB change from baseline	Increase from baseline > 30 - ≤ 60 ms Increase from baseline > 60 ms

QT will be corrected based on Fridericia's formula (QTcF= QT/ $\sqrt[3]{RR}$) and RR=60/HR, and Bazett's formula (QTcB= QT/ √RR) and RR=60/HR, if possible. Baseline QTcF/QTcB will be derived from the visit that QT and HR are flagged as baseline. If there are multiple assessments at the same visit and time point, the average will be calculated for each parameter and used for the analysis. Quantitative data will also be examined for trends using descriptive statistics (n, missing, mean, SD, median, Q1, Q3, minimum, and maximum) of actual values (with 95% CI of the mean), absolute changes from baseline (with 90% CI of the mean) to each visit based on visit windows

specified in Section 11. This summarization will apply to heart rate, QRS interval, QT interval, PR interval, QTcB, and QTcF assessed at baseline, post-baseline, discontinuation, end of treatment, and/or safety follow-up visits.

Listings of 12-lead ECGs will be provided with all relevant information such as visit, date/time of assessment, parameter, and results.

The ECOG shift from baseline to highest on-treatment score will be summarized by dose level and overall based on SAF analysis set. Missing category will be included and the number of subjects in each dose level will be used as the denominator.

ECOG performance status will also be presented in a data listing.

18 Reporting Conventions

- Reporting will require placement of decimals, and this will depend on the raw data collected. Generally, mean and median should be displayed one more decimal place than the raw data and standard deviation should be displayed two more decimal place than the raw data. Percentages will be reported as one decimal place.
- The rounding will be performed to closest integer / first decimal using the common mid-point between the two consecutive values. e.g. 5.1 to 5.4 will be rounded to an integer of 5, and 5.5 to 5.9 will be rounded to an integer of 6.
- The following conversion factors will be used to convert days to months or years, where applicable: 1 month = 30.4375 days and 1 year = 365.25 days.
- Data listings will be sorted by dose level, subject identifier, visit or date (if applicable), or parameters, as appropriate.

18.1 Incomplete dates

Missing or partial start dates for adverse events will be imputed as following:

- When the start Date of the AE is missing (but Month & Year is available), then the AE date will be imputed to the "1st Date of the reported Month" (e.g. if reported date is --/JAN/09, imputed date will be 01/JAN/09). If the reported AE Month = the Month of the First Dosing date, then the AE date will be imputed to the "1st Dosing date" (e.g. if AE reported date is --/JAN/09, and the First doing date is 13/JAN/09, then the AE imputed date will be 13/JAN/09).
- When the Date & the Month of the AE is missing (but Year is available), then the AE date will be imputed to the "1st Date of the 1st Month of the reported Year" (e.g. if reported date is --/--/09, imputed date will be 01/JAN/09). If the reported AE Year = the Year of the First Dosing date, then the AE date & month will be imputed to the "1st Dosing date" (e.g. if AE reported date is --/--/09, and the First doing date is 13/APR/09, then the AE imputed date will be 13/APR/09).

- When the date is completely missing, no imputation will be performed and the AE will be considered as treatment emergent, unless there is rational to clarify otherwise, e.g. AE stop date is prior to the first dose date.

Missing or partial dates for concomitant medications will be imputed as following:

- If the start day of medication is missing, it will be imputed to the first day of the month; if the stop day is missing, it will be imputed to the last day of the month. If both day and month are missing, the start date will be imputed as the first day of the year and stop date will be imputed as the last day of the year. If the start or stop date is completely missing, no imputation will be performed and the determination of prior medication or post medication will be based on non-missing stop or start date, respectively; otherwise, the medication will be considered as concomitant.

Missing or partial dates for disease history (initial diagnosis date, first occurrence of metastatic or locally advanced disease, date of last progression of disease) will be imputed as following:

- If the day is missing, it will be imputed to the 15th day of the month; if both day and month are missing and the year is prior to the year of the first study treatment, the month and day will be imputed as July 1st; if both day and month are missing and the year is same as the year of the first study treatment, the month and day will be imputed as January 1st. If the date is completely missing, no imputation will be performed.

Partial dates for prior anti-cancer drug therapies will be imputed as following:

- If the day is missing, it will be imputed to the first day of the month; if both day and month are missing, no imputation will be performed.

No other dates will be imputed.

19 References

1. Eisenhauer EA. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). Eur J Cancer 2009;45:228-47.
2. Wolchok et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. Clin Cancer Res 2009;15(23):7412-20.

20 Appendices

20.1 Programmed Major Protocol Deviations

	Category of Protocol Deviation	Description of Protocol Deviation	Deviation Code	Protocol Section	Variable [dataset]	Proposed check / comment
Inclusion criteria: For the subject to be eligible for inclusion, each criterion must be checked 'YES':						
Criterion1: Signed written informed consent.	Eligibility and Entry Criteria	Subject did not meet inclusion criterion 1 (signed inform consent form).	PDEV01	Section 5.3.1	DM. RFICDTC	list if DM. RFICDTC is missing or if DM.RFICDTC<Earliest date of SV.SVSTDTC. Medical Review Required.
Criterion2: Male or female subjects aged >=18 years.	Eligibility and Entry Criteria	Subject did not meet inclusion criterion 2 (age >= 18 years).	PDEV02	Section 5.3.1	DM.AGE	List if DM.AGE<18.
Criterion3: Histologically or cytologically proven metastatic or locally advanced solid tumor or related prior anti-cancer therapy.	Eligibility and Entry Criteria	Subject did not meet inclusion criterion 3 (histologically or cytologically proven metastatic or locally advanced solid tumor).	PDEV03	Section 5.3.1		Medical review required

	Category of Protocol Deviation	Description of Protocol Deviation	Deviation Code	Protocol Section	Variable [dataset]	Proposed check / comment
Criterion5: Disease must be measurable at least 1 dimension except for CRPC or MBC (dose escalation).	Eligibility and Entry Criteria	Subject did not meet inclusion criterion 5 (measurable disease at least one dimension except for CRPC and escalation MBC subjects).	PDEV04	Section 5.3.1	TR.TRORRES, TRTESTCD; SUPPTR.QNAM, QVAL.	list if TR.TRORRES is missing or not assessable where TRTESTCD=' SUMDIAM' and TRCAT= 'RECIST' and SUPPTR.QVAL='SCREENING' where QNAM='ASMNTVIS'. Do not check this for subjects where SUPPDM.QVAL = ' CASTRATE-RESISTANT PROSTATE CANCER' where QNAM='COHORT' or SUPPDM.QVAL=' DOSE ESCALATION AND DLT/MTD COHORTS' where QNAM='COHORT' and MHLOC in ('NIPPLE' 'CENTRAL PORTION OF BREAST' 'UPPER-INNER QUADRANT OF BREAST (UIQ)' 'LOWER-INNER QUADRANT OF BREAST (LIQ)' 'UPPER-OUTER QUADRANT OF BREAST (UOQ)' 'LOWER-OUTER QUADRANT OF BREAST (LOQ)' 'AXILLARY TAIL OF BREAST' 'OVERLAPPING LESION OF BREAST 'BREAST, NOS')
Criterion9: Effective contraception for both male and female subjects if the risk of conception exists.	Eligibility and Entry Criteria	Subject did not meet inclusion criterion 9 (effective contraception).	PDEV05	Section 5.3.1		Medical review required
Exclusion criteria: For the subject to be eligible for inclusion, each criterion must be checked 'NO':						
1. Concurrent treatment with a non-permitted drug.	Eligibility and Entry Criteria	Subject met exclusion criterion 1 (non-permitted drug).	PDEV06	Section 5.3.2		Medical review required

	Category of Protocol Deviation	Description of Protocol Deviation	Deviation Code	Protocol Section	Variable [dataset]	Proposed check / comment
2. Prior therapy with any antibody/drug targeting T cell co-regulatory proteins.	Eligibility and Entry Criteria	Subject met exclusion criterion 2 (prior therapy with any antibody/drug targeting T cell co-regulatory proteins).	PDEV07	Section 5.3.2		Medical review required
3. Concurrent anticancer treatment, major surgery, concurrent systemic therapy with steroids or other immunosuppressive agents, or use of any investigational drug within 28 days before the start of trial treatment.	Eligibility and Entry Criteria	Subject met exclusion criterion 3 (concurrent anti-cancer therapy or surgery, concurrent systemic therapy with steroids or other),	PDEV08	Section 5.3.2		Medical review required
4. Previous malignant disease other than the target malignancy to be investigated in this trial within the last 5 years with the exception of basal or squamous cell carcinoma of the skin or cervical carcinoma in situ	Eligibility and Entry Criteria	Subject met exclusion criterion 4 (previous malignant disease other than the target malignancy to be investigated in this trial within the last 5 years).	PDEV09	Section 5.3.2		Medical review required
12. Pregnancy or lactation period.	Eligibility and Entry Criteria	Subject met exclusion criterion 12 (pregnancy or lactation period).	PDEV10	Section 5.3.2	PREG.PRRLTCD	Medical review required.

	Category of Protocol Deviation	Description of Protocol Deviation	Deviation Code	Protocol Section	Variable [dataset]	Proposed check / comment
Non-permitted concomitant medication during the study	Concomitant Medication Criteria	Subject took prohibited medication during the study.	PDEV11	Section 6.5.2	CMED.CMTERM	Medical review required
Subjects that developed withdrawal criteria whilst on the study but were not withdrawn;	Other Criteria	Subject became pregnant, but continued on the study.	PDEV12	Section 5.5.2	LB.LBORRES, BTESTCD; DS.DSSTDTC, DSSCATE;	Medical review required.
Subjects that developed withdrawal criteria whilst on the study but were not withdrawn;	Other Criteria	Subjects had ECOG>=3, did not resolved to <=2 by day 14 of next cycle, and continued on the study.	PDEV13	Section 5.1.7.2 and 5.5.2	XP.XPORRES, XPDY; DS.DSSTDTC, DSSCATE;	Medical review required.
Subjects that developed withdrawal criteria whilst on the study but were not withdrawn;	Other Criteria	Subject developed grade 4 AE, but continued on the study.	PDEV14	Section 5.1.7.2 and 5.5.2	ADAE.ATOXGRN, TRTEMFL, AREL; SDTM.AE.AEACNOTH; SUPPAE.QVAL, QNAM	List if TOXGRN=4 and AREL='Related' and TRTEMFL='Y' and (SUPPAE.QVAL^='Drug Withdrawn' where QNAM in (ACNMSB, ACNMSB3) and SDTM.AEACNOTH^='LED TO STUDY TERMINATION')
Subjects that developed withdrawal criteria whilst on the study but were not withdrawn;	Other Criteria	Subject developed grade 3 AE, but continued on the study.	PDEV15	Section 5.1.7.2 and 5.5.2	ADAE.ASEVN, TRTEMFL, AEDECOD, AEDUR. DS.DSSTDTC, DSSCATE;	Medical review required.
Subjects overdosed (>=110% of assigned dose)	IP Compliance	Subject was overdosed.	PDEV16	Medical defined.	ADEXSUM.AVAL where PARAMACD=' RDOSINT'	list if cumulative relative dose intensity or individual relative dose intensity >=110

	Category of Protocol Deviation	Description of Protocol Deviation	Deviation Code	Protocol Section	Variable [dataset]	Proposed check / comment
NA	Other Criteria	Other protocol deviation	PDEV99	Medical defined		Medical review required.

Statistical Analysis Plan for Expansion Phase

Clinical Trial Protocol Identification No.	EMR 100070-001
Title:	A Phase I, open-label, multiple-ascending dose trial to investigate the safety, tolerability, pharmacokinetics, biological and clinical activity of avelumab (MSB0010718C) in subjects with metastatic or locally advanced solid tumor and expansion to selected indications
Trial Phase	Phase I
Investigational Medicinal Product(s)	Avelumab
Clinical Trial Protocol Version	02 March 2017/Version 18.0 (Amendment 17)
Statistical Analysis Plan Author	PPD
Statistical Analysis Plan Date and Version	31 August 2017/Version 9.0
Statistical Analysis Plan Reviewers	PPD

PPD

[redacted]

1 Signature Page

Statistical Analysis Plan: EMR 100070-001

A Phase I, open-label, multiple-ascending dose trial to investigate the safety, tolerability, pharmacokinetics, biological and clinical activity of avelumab in subjects with metastatic or locally advanced solid tumor and expansion to selected indications

PPD

PPD

PPD

Signature

Date of Signature

Trial Biostatistician

PPD

PPD

PPD

Signature

Date of Signature

Trial Statistician
Merck KgaA
Frankfurter Str. 250
64293 Darmstadt, Germany

PPD

Signature

Date of Signature

PPD

EMD Serono
45A Middlesex Turnpike
Billerica, MA 01821, USA

2 Table of Contents

3 List of Abbreviations and Definition of Terms

ACC	Adrenocortical Carcinoma
ACTH	Adrenocorticotropic Hormone
ADA	Anti-Drug Antibody (referred to as HAHA on CRF and SDTM)
AE	Adverse Event
ALP	Alkaline Phosphatase
ALK	Anaplastic Lymphoma Kinase
ALT	Alanine Aminotransferase
ANA	Anti-nuclear Antibody
ANCA	Anti-neutrophil Cytoplasmic Antibody
ANC	Absolute Neutrophils Count
AST	Aspartate Aminotransferase
aPTT	Activated Partial Thromboplastin Time
ATC	Anatomical Therapeutic Chemical
BMI	Body Mass Index
BOR	Best Overall Response
BRCA1/2	Breast Cancer 1 or 2
BSA	Body Surface Area
CA125	Cancer Antigen 125
CALCIO	Corrected Calcium and Ionized Calcium
CI	Confidence Interval
Cmin	Trough Concentration
CPK	Creatine Kinase
CR	Complete Response

CRC	Colorectal Cancer
CRPC	Castrate Resistant Prostate Cancer
CRO	Contract Research Organization
CT	Center of Tumor
CTCAE	Common Terminology Criteria for Adverse Events
CSR	Clinical Study Report
CV	Coefficient of Variation
DR	Duration of Response
DRM	Data Review Meeting
EBV	Epstein-Barr Virus
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic Case Report Form
eCcr	Estimated Creatinine Clearance Rate
eDISH	Evaluation of Drug-Induced Serious Hepatotoxicity
EFF	Efficacy Analysis Set
CCI	[redacted]
EOI	End of Each Infusion
ER	Estrogen Receptor
FAS	Full Analysis Set
Free T4	Free Thyroxine
GEJ	Gastroesophageal Junction
GeoCV	Geometric Coefficient of Variation
GeoMean	Geometric Mean

GGT	Gamma Glutamyl Transferase
CCI	[redacted]
HB	Hemoglobin
CCI	[redacted]
HNSCC	Head and Neck Squamous Cell Carcinoma
HPV	Human Papillomavirus
HR	Heart Rate
ICF	Informed Consent Form
IERC	Independent Endpoint Review Committee
IM	Invasive Margin
IMP	Investigational Medical Product
irAE	Immune Related Adverse Event
irBOR	immune-related BOR
irCR	immune-related Complete Response
irORR	immune-related Objective Response Rate
irPD	immune-related Progressive Disease
irPFS	immune-related Progression Free Survival
irPR	immune-related Partial Response
IRR	Infusion Related Reaction
irRC	immune-related Response Criteria
irSD	immune-related Stable Disease
i.v.	Intravenous
CCI	[redacted]
LDH	Lactate Dehydrogenase

LLOQ	Lower Limit of Quantification
MBC	Metastatic Breast Cancer
MCH	Mean Corpuscular Hemoglobin
MCHC	Mean Corpuscular Hemoglobin Concentration
MCV	Mean Corpuscular Volume
MedDRA	Medical Dictionary for Regulatory Activities
MSI	Microsatellite Instability
nAb	Neutralizing Antibodies
NCI	National Cancer Institute
NE	Not Assessable
NSCLC	Non-Small Cell Lung Cancer
ORR	Objective Response Rate
OS	Overall Survival
PCSA	Potentially Clinically Significant Abnormalities
PFS	Progression Free Survival
PLT	Platelet Count
PD	Progressive Disease
PD-1	Programmed Death 1
PD-L1	Programmed Death Ligand 1
PK	Pharmacokinetics
PKAS	PK Analysis Set
PO	Per Oral
PR	Partial Response
PRF	Pathology Report Form

PT	Preferred Term
QIMS	QuintilesIMS
QTcB	QT Interval Corrected Using Bazett's Formula
QTcF	QT Interval Corrected Using Fridericia's Formula
RBC	Red Blood Cell
RCC	Renal Cell Carcinoma
RECIST	Response Evaluation Criteria in Solid Tumors
RF	Rheumatoid Factor
SAE	Serious Adverse Event
SAF	Safety Analysis Set
SAP	Statistical Analysis Plan
SCRN	Screening Analysis Set
SD	Stable Disease
SD	Standard Deviation
SMC	Safety Monitoring Committee
SOC	System Organ Class
CCI [redacted]	
TEAE	Treatment Emergent Adverse Event
CCI [redacted]	
TTR	Time to Response
ULN	Upper Limit of Normal

4 Modification History

Unique Identifier for SAP Version	Date of SAP Version	Author	Changes from the Previous Version
Final 1.0	24Jun2014	Jiali Tang	Not Applicable
Final 2.0	17Mar2015	Jiali Tang	Updates based on Protocol V8.0 (Amendment 7)
Final 3.0	15May2015	Jiali Tang	Updates based on Protocol V12.0 (Amendment 11)
Final 4.0	22Sep2015	Jiali Tang	1)Added CD8 and associated analyses; 2)Updated based on Protocol V13.0 (Amendment 12); 3)Add 20 subjects interim analysis for HNSCC; 4)Updated baseline and safety based on Master Avelumab SAP.
Final 5.0	28Oct2015	Jiali Tang	1)Updated based on Protocol V14.0 (Amendment 13) 2)Added safety analysis for interim CSR 3)Added PK/HAHA analyses 4)Updated per Harmonized Avelumab Master SAP draft v6.0
Final 6.0	12Jan2016	Danielle Lamy	1) Updated definitions of PD-L1 positive analysis sets for the urothelial carcinoma efficacy expansion cohort 2) Updated 'Changes to the Planned Analyses' section (added interim analyses and noted no interim analysis after 30 patients for UC efficacy cohort and added additional primary analysis updates for secondary UC cohort) 3) Updated IRR section, date of last contact, imputation of missing or partial death date according to program-level standard per Harmonized Avelumab Master SAP final V1 4) Removed hypoglycemia from the set of parameters that are not categorized for grades 1 and 2 based on fasting status 5) Updated overall survival section to specify that data after the cut-off date is not used in the analysis of OS
Final 7.0	29Apr2016	Deborah Templin	1)Added confirmed BOR based on IERC as a secondary endpoint for secondary Urothelial Carcinoma cohort. 2)Harmonized statistical testing approach for efficacy expansion cohorts. 3)Added interim analysis for 109 efficacy urothelial carcinoma subjects with 6 months follow-up period. 4)Added analysis for time-to-response. 5)Added summary of baseline albumin and hemoglobin, eligibility of platinum-based therapy. 6)Added subgroup analyses by albumin and hemoglobin, eligibility of platinum-based therapy, and HAHA status, and CCI [redacted] for gastric cancer, ovarian cancer or MBC cohorts. 7)Updated summary of subject death. 8)Updated language to apply IERC assessment to secondary urothelial carcinoma cohort and related analyses based on IERC data. 9) Updated NCI-CTCAE to v4.03 for laboratory toxicity grading

Unique Identifier for SAP Version	Date of SAP Version	Author	Changes from the Previous Version
			10) Updated language to apply to most recent ICH E3 guideline for important protocol deviations
Final 8.0			16.4.4 Relation of Pharmacokinetics to Efficacy in Urothelial Carcinoma Cohorts Exposure response analyses supplemented Additional PD-L1 cut-offs supplemented Wording for presence of metastases at baseline updated. RCC cohort updates. PFS and OS Reasons for Censoring alignment with Avelumab Harmonized SAP.
Final 9.0	10Feb2017	Jie Wang	1) Section 8.1 added the subgroup for RCC first-line. 2) Section 10, added PD-L1 positive (1% cut-off) FAS for NSCLC cohorts. 3) Section 11, added an algorithm on how to calculate time since an event. 4) Section 13.1, added Baseline Bellmunt Score, Baseline eCcr, and Time Since Last Prior Anti-Cancer Chemotherapy. 5) Section 13.4, added the definition of 'negative' for PD-L1 expression status based on the secondary cut-off for tumor cells. Added PD-L1 scoring for gastric. 6) Section 16.1, updated BOR derivation. 7) Section 16.2, updated BOR reasons for NE to align with the Harmonized SAP. Updated BOR and irBOR derivation. 8) Section 16.2.3, updated the censoring reasons to be consistent with PFS censoring scenarios. 9) Section 16.2.6, added subgroup analysis by Baseline Bellmunt Score, Time Since Last Prior Anti-cancer Chemotherapy, Gastric specific PD-L1, and Region (Asia, Non-Asia). 10) Section 16.5, added nAb TLFs 11) Section 17.1 added the new process on how to identify and produce irAEs and clarified original and updated definitions. 12) Section 17.1.6, added nAb categorization and expanded safety tables for ADA subgroup analysis 13) Section 17.3, Added the calculation of corrected eCcr. 14) Updated the terminology HAHA to the more accurate terminology: ADA 15) Added Appendix II for irAEs and IRRs.

5 Purpose of the Statistical Analysis Plan

The purpose of this statistical analysis plan (SAP) is to document technical and detailed specifications for the interim, primary, and final analyses of data collected for the expansion phase under protocol EMR 100070-001. Results of interim, primary, and final analyses described in this SAP will be included in the interim or final Clinical Study Reports (CSRs) or addenda thereto. Additionally, the planned interim, primary, and final analyses identified in this SAP will be included in regulatory submissions or future manuscripts. Any post-hoc, or unplanned analyses performed to provide results for inclusion in a CSR or report, but not identified in this prospective SAP, will be clearly identified in the CSR or report.

6 Summary of Clinical Trial Features

Trial Objectives	Trial objectives specifically for dose escalation were included into the dose escalation SAP. **Primary** • To assess the best overall response (BOR) according to Response Evaluation Criteria in Solid Tumors version 1.1 (RECIST 1.1, 1) in the efficacy expansion cohorts (ovarian cancer, platinum refractory and prior liposomal doxorubicin; urothelial carcinoma, platinum ineligible or progressed after at least 1 line of platinum-based therapy; gastric and gastroesophageal junction [GEJ] cancer, third-line; head and neck squamous cell carcinoma [HNSCC], platinum ineligible or progressed after at least 1 line of platinum-based therapy). **Secondary** • To characterize the pharmacokinetic (PK) profile of avelumab and to correlate exposure with target occupancy. • To evaluate the immunogenicity of avelumab and to correlate it to exposure and biological activity. • To assess the BOR and progression-free survival time (PFS) according to RECIST 1.1. • To assess the immune-related BOR (irBOR) and immune-related PFS (irPFS) using the modified Immune-Related Response Criteria (irRC, 2), derived from RECIST 1.1. • To assess overall survival (OS) time. • To evaluate biological responses to avelumab in blood/serum. • To evaluate the association between tumor programmed death ligand 1 (PD-L1) expression and BOR. • To characterize changes in soluble factors (e.g., cytokine profiles, soluble programmed death 1 [PD-1], and soluble PD-L1) and immune cell profiling (e.g., natural killer cells, neutrophils, lymphocytes). CCI [redacted]
Trial Endpoints	Trial endpoints specifically for dose escalation were included into the dose escalation SAP.

	Primary
	• The confirmed BOR, per RECIST 1.1, as adjudicated by an Independent Endpoint Review Committee (IERC) for subjects enrolled in the efficacy expansion cohorts only.
	
	Secondary
	• Number, severity, and duration of treatment-emergent adverse events (TEAEs) for all cohorts according to the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) v4.0.
	• Number, severity, and duration of treatment-related adverse events (AEs) according to NCI-CTCAE v4.0.
	• PK profile.
	• irBOR and BOR according to modified irRC and to RECIST 1.1, respectively, per investigator assessment.
	• The confirmed BOR, per RECIST 1.1, as adjudicated by an IERC, for subjects enrolled in the secondary urothelial carcinoma cohort.
	• irPFS time and PFS time according to modified irRC and to RECIST 1.1, respectively, per investigator assessment.
	• OS time.
	• Pharmacodynamic (PD) profile
	• Serum titers of anti-avelumab antibodies (ADA).
	• Expression of PD-L1 on tumor tissue.
	• For the primary expansion cohorts only: Unconfirmed response at Week 13 according to RECIST 1.1.
	• Duration of response (DR) according to modified irRC and to RECIST 1.1, respectively.
	• For the efficacy expansion cohorts and secondary urothelial carcinoma cohort:
	  o PFS time, according to RECIST 1.1, per IERC
	  o DR according to RECIST 1.1, per IERC.

Trial Design	This is a Phase I, open-label, dose-escalation trial with consecutive parallel group expansion in non-small cell lung cancer (NSCLC), metastatic breast cancer (MBC), gastric and GEJ cancer, colorectal cancer (CRC), castrate resistant prostate cancer (CRPC), melanoma, ovarian cancer, HNSCC, adrenocortical carcinoma (ACC), renal cell carcinoma (RCC), mesothelioma, and urothelial carcinoma. **Dose escalation phase** This section was included into the Dose Escalation SAP.

Expansion phase

After determination of the avelumab dose and regimen for further investigation, enrolment in several expansion cohorts will be opened in selected tumor indications to determine the safety and clinical activity of avelumab. Subject eligibility will be confirmed by the contract research organization (CRO) / Sponsor for each subject before the first administration of the study treatment during the expansion phase.

Based on data generated in the dose escalation phase, the dose of avelumab to be used in the expansion phase was determined to be 10 mg/kg. In addition, with the emergence of promising efficacy data, expansion cohorts have been expanded and divided into:

- 4 primary cohorts of:
  1. NSCLC, post platinum doublet (N=150);
  2. NSCLC, first-line, does not carry an epidermal growth factor receptor (EGFR) activating mutation or anaplastic lymphoma kinase (ALK) re-arrangements (N=150);
  3. Gastric and GEJ cancer (N=150); and
  4. MBC (N=150)
- 8 secondary cohorts of:
  1. CRC (N=20),
  2. CRPC (N=20),
  3. ACC (N=50),
  4. Melanoma (N=50),
  5. Mesothelioma (N=50),
  6. Urothelial carcinoma (N=50; note: enrollment is being stopped [N=44] due to the opening of a urothelial efficacy expansion cohort),
  7. Ovarian cancer (N=120), and
  8. RCC, second-line, (N=20 with expansion of 60 first-line).
- 4 efficacy expansion cohorts of:
  1. Ovarian cancer, platinum refractory, prior liposomal doxorubicin (N=100);
  2. Urothelial carcinoma, platinum ineligible or progressed after at least 1 line of platinum-based therapy (N=200);
  3. Gastric and GEJ cancer, third line (N=150);
  4. HNSCC, platinum ineligible or progressed after at least 1 line of platinum-based therapy (N=150);

Subjects in the NSCLC (post platinum doublet), CRC, and CRPC cohorts will be enrolled in the USA only.

	For subjects enrolled in the efficacy expansion cohorts and the secondary urothelial carcinoma cohort, an IERC will perform a blinded determination as to whether the criteria for tumor response or progression according to RECIST 1.1 have been met. Subjects will receive avelumab intravenously as a 1-hour infusion once every 2 weeks until confirmed progression, unacceptable toxicity, or any reason for withdrawal from the trial or Investigational Medical Product (IMP) occurs. Subjects who have experienced a confirmed complete response (CR) should be treated for a maximum of 24 months after confirmation, at the discretion of the investigator. If the investigator believes that a subject may benefit from treatment beyond 24 months, it may be permissible after discussion with the sponsor. Subjects who experienced a CR and have already stopped treatment can resume treatment with avelumab at the same dose and schedule. Subjects re-initiating treatment should be assessed according to the Schedule of Assessments. For subjects who achieve a CR on avelumab therapy and then subsequently develop disease progression after stopping therapy, but prior to the end of the trial, one re-initiation of treatment at the same dose and schedule is allowed at the discretion of the investigator and agreement of the trial Medical Monitor. In order to be eligible for retreatment, the subject must not have experienced any toxicity that led to treatment discontinuation of the initial avelumab therapy. Prior to re-initiation of the study treatment, malignant disease needs to be radiologically re-staged to assess all known sites of the disease and to establish a new baseline for subsequent tumor measurements. Relevant safety laboratory results must be available and verified prior to re-initiating of treatment. Subjects who re-initiate treatment will stay on study and will be treated and monitored according to the protocol and the “until progression” schedule in the Schedule of Assessments.
Number of Subjects	Expansion phase: 1610 subjects.
Trial Product	Avelumab will be administered as 1-hour intravenous (i.v.) infusion. Subjects will receive avelumab once every 2 weeks until confirmed progression, unacceptable toxicity, or any criterion for withdrawal from the trial or IMP occurs. The dose of avelumab will be calculated based on the weight of the subject determined on the day of each drug administration. Premedication with an antihistamine and with paracetamol (acetaminophen) approximately 30 to 60 minutes prior to each dose of

	avelumab is mandatory (for example, 25-50 mg diphenhydramine and 500-650 mg paracetamol [acetaminophen] i.v. or oral equivalent). This regimen may be modified based on local treatment standards and guidelines, as appropriate. Immediate access to intensive care unit or equivalent environment and appropriate medical therapy (including intravenous epinephrine, corticosteroids, antihistamines, bronchodilators, and oxygen) must be in place for use in the treatment of potential infusion reactions. Infusion of avelumab will be stopped in case of ≥ Grade 2 infusion-related, allergic, or anaphylactic reactions (according to NCI-CTCAE v4.0). Following avelumab infusion, subjects must be observed for 2 hours post infusion for potential infusion-related reactions (IRRs).
Treatment and Trial Duration	The planned treatment duration is until unacceptable toxicity, or any criterion for withdrawal from the trial or IMP occurs.
Schedule of Visits	1. Screening/Baseline assessment (day -18 to first treatment) 2. Treatment phase Visits will be conducted weekly until week 7 and every 2 weeks thereafter prior to the approval of Protocol Amendment 7. For subjects participating in the expanded PK sampling, visits at Days 2 and 3 will be required. Following the approval of Protocol Amendment 7, the visits at Week 2, 4, and 6 are no longer required and subjects will not be required to attend these visits. 3. Discontinuation visit and end-of-treatment visit All subjects who discontinue study treatment prematurely for an AE should have a full safety evaluation at the time of discontinuation (discontinuation visit). For all subjects who have completed treatment, an end-of-treatment visit should be scheduled 4 weeks after the last administration of avelumab. The end-of-treatment visit is scheduled 4 weeks after the last administration of avelumab but before any new therapy is started, if possible. The visit will comprise a full assessment of safety parameters, immunogenicity assessment, and tumor response assessment as appropriate. 4. Post-treatment follow-up (safety follow-up visit and survival follow-up) All subjects will have a subsequent visit scheduled 10 weeks after the last administration of avelumab. The visit will include a full assessment of safety parameters. AEs will be documented until the end of treatment visit. After the end of treatment visit only treatment related AEs have to be documented until the post-treatment safety follow-up visit. Subjects with a serious AE ongoing

	at the post treatment safety follow-up must be monitored and followed up by the investigator until stabilization or until the outcome is known, unless the subject is documented as "lost to follow-up". Subjects without progressive disease (PD) at the end-of-treatment visit will be followed up for disease progression (CT / MRI scans every 12 weeks) up to 1 year. In addition, subjects will be followed for any AE suspected to be related to study treatment, especially for the occurrence of new autoimmune events up to 3 months after the last dose of avelumab. After the end-of-treatment visit, subjects will be followed quarterly for survival (including assessment of any further tumor therapy). The survival follow-up will continue until 1 year after the last subject receives the last dose of avelumab. Schedule of Assessments can be found in Section 12, Appendix I, of the protocol.
Randomization and Blinding	Not applicable.

CCI

CCI

[redacted]

CCI

[redacted]

8 Overview of Planned Analyses

This SAP will only address analyses for the expansion phase. Additional SAPs for Safety Monitoring Committee (SMC) analyses and for the final analysis of dose escalation phase have been developed separately. Trial endpoints specifically for dose escalation have been included into the dose escalation SAP.

The planned analyses for the expansion phase are summarized in Table 2. An additional interim analysis may be conducted 13 weeks after the start of treatment of the last subject in that cohort for selected primary and secondary expansion cohorts. Depending on subject enrollment or planning of future study, a separate primary analysis may be performed for selected secondary cohorts. Those analyses, as well as an interim safety analysis for interim CSR, are also included into Table 2.

Table 2 Planned Analyses for Expansion Cohorts

Cohort	Interim Analysis	Primary Analysis	Interim CSR and Final Analysis
NSCLC post platinum doublet	60 subjects 75 subjects all subjects	X	Pooled expansion cohorts for interim CSR and all expansion cohorts together for final analysis
NSCLC first-line	30 subjects 75 subjects all subjects	X	
MBC	75 subjects	X	
Gastric and GEJ cancer primary	75 subjects all subjects	X	
CRPC		X	
CRC		X	
2nd Ovarian cancer cohort	20 subjects 75 subjects all subjects	X	
ACC	20 subjects all subjects	X	
Melanoma	20 subjects all subjects	X	
Mesothelioma	20 subjects all subjects	X	
2nd Urothelial carcinoma cohort	20 subjects all subjects	X*	
RCC second-line	20 subjects	X	
RCC first-line	30 subjects	X	
Ovarian cancer efficacy	30 subjects 60 subjects	X	
Urothelial carcinoma efficacy	30 subjects 109 subjects*	X	
Gastric and GEJ cancer efficacy	30 subjects 90 subjects	X	
HNSCC	30 subjects 90 subjects	X	

X = Primary analysis to be conducted with a data cut-off date of the last subject first dose date + 6 months for each cohort.
* = Interim analysis will be performed after the 109th dosed subject from efficacy urothelial carcinoma cohort has reached 6 months follow-up period.

8.1 Interim Analysis

For the NSCLC (post platinum doublet and first-line), gastric/GEJ cancer, and MBC expansion cohorts, an interim analysis will be performed after the 75th subject in each cohort has reached the

time point of the second post-baseline tumor assessment scheduled in Week 13, i.e. 13 weeks after start of treatment of the 75th subject. Efficacy in this 75 subject subset of the cohort will be analyzed in terms of the unconfirmed response at Week 13. If the rate of unconfirmed response at Week 13 (according to RECIST 1.1) in the efficacy population defined as all treated subjects with measurable disease at baseline is less than 5%, enrollment in the given cohort may be stopped.

Based on a comprehensive review of the efficacy and safety data it may be considered whether recruitment in a subgroup of the study population of the given indication, defined by PD-L1 expression status, might be resumed by means of a substantial Protocol Amendment.

Statistical considerations related to this futility rule:

Under different assumptions on the true response rate in the overall population, the probabilities of observing a response rate of less than 5% in this analysis (i.e., 3 or less responders out of 75 subjects) are noted in Table 3.

Table 3 Probability of Observing a Response Rate of Less Than 5% in Interim Analysis

True response rate in overall population	Probability of 3 or less responders in 75 patients
0.03	0.81
0.05	0.48
0.10	0.05
0.15	0.002

Interim analyses on the first 75 NSCLC post platinum doublet or MBC subjects included the baseline characteristics, efficacy (excluding DR, TTR and OS), and safety evaluations. Interim analyses on the first 75 gastric/GEJ cancer or NSCLC first-line subjects will include the baseline characteristics and efficacy (excluding DR, TTR and OS).

In the NSCLC post platinum doublet cohort only, 2 additional interim analyses of efficacy parameters are planned for internal planning purposes at the following time points:

- 13 weeks after start of treatment of the 60th subject. This interim analysis was performed on demographic, disease history, and efficacy endpoints.
- 13 weeks after start of treatment of the last subject. This interim analysis will be performed on disposition, demographic and baseline characteristics, treatment exposure, and efficacy endpoints. The subgroup analyses on efficacy endpoints will be performed on all dosed subjects and on PD-L1 positive subjects based on tertiary cut-off as defined in Section 13.4.

In the first-line NSCLC primary expansion cohort, an interim analysis of response will be conducted 13 weeks after start of treatment of the 30th subject.

In the efficacy expansion cohorts, interim analyses for efficacy are planned 13 weeks after the start of treatment of the 30th subject in all cohorts, 13 weeks after start of treatment of the 60th subject in the ovarian cohort, and 13 weeks after the start of treatment of 90th subject in the gastric / GEJ

and HNSCC cohorts. The interim analyses after 60/90 subjects aim to demonstrate efficacy as specified in Section 16.1. No futility rule is foreseen because the clinical activity of anti-PD-1 / anti-PD-L1 agents in these tumor types is established, and the patient populations are characterized by a high unmet medical need. If efficacy criteria are met at the interim analysis, enrollment will continue to the planned full number of subjects in order to collect further data on the primary and secondary endpoints, especially on the association between PD-L1 expression and efficacy endpoints.

In addition, in the secondary cohorts that plan to enroll more than 20 subjects, i.e., the ACC, melanoma, mesothelioma, ovarian cancer, and urothelial carcinoma cohorts, an interim analysis of response will be performed 13 weeks after the start of treatment of the 20th subject. Accrual in each cohort may be paused during the interim analysis. If no unconfirmed response according to RECIST 1.1 is observed in a given cohort in the interim analysis, accrual in that cohort will be stopped. The data included into the interim analysis are demographic, disease history, prior anti-cancer therapy, PD-L1 expression status, lesion assessment, BOR, ORR by PD-L1 expression status, and PFS overall and by PD-L1 expression status. In addition, for the ovarian cancer secondary expansion cohort, an interim analysis of response will be performed for internal planning purposes 13 weeks after the start of treatment of the 75th subject.

Enrollment of first-line RCC subjects was opened after 2 documented objective responses among the 20 subjects enrolled in the second-line RCC cohort were observed by RECIST 1.1 (2 PRs), and justified further evaluation in this patient population. In the RCC cohort, interim analyses of response will be performed 13 weeks after the start of treatment of the 20th subject of second-line RCC; and after the 30th subject of the first-line RCC.

The sequence of statistical analyses planned for urothelial cancer subjects will consider the objective to evaluate the association between tumor PD-L1 expression and BOR prospectively. In a first step, the secondary urothelial carcinoma cohort served as a "training set" for the identification of a PD-L1 expression cut-off (5%) that is most likely to identify a subset of the subject population with enhanced clinical benefit. The PD-L1 expression cut-off (5%) was specified prior to any statistical analysis of the PD-L1 expression data from the urothelial carcinoma efficacy expansion cohort (cut-off determined 11Mar2016). In the next step, the cut-off will be verified by conducting an interim evaluation with data from subjects of the efficacy expansion cohort at 6 months after the last subject's first dose of study treatment for the 109 subjects enrolled in the urothelial carcinoma efficacy expansion cohort prior to Protocol Amendment 13. The cut-off date for this analyses was 19Mar2016. In case the cut-offs are not mutually supportive in terms of clinical efficacy endpoints, the cut-off could be refined and the remainder of approximately 100 subjects of the efficacy expansion cohort will serve as the "validation set" to qualify the tumor PD-L1 expression cut-off. Otherwise, data from subjects of the urothelial carcinoma expansion cohorts will be pooled for the final efficacy analyses of the expansion cohorts at 6 months after start of treatment of the last subject enrolled in the efficacy expansion cohort.

For each primary and secondary expansion cohort, an additional interim analysis may be conducted 13 weeks after the start of treatment of the last subject in that cohort. In general, interim

analyses at time points that are not specified in the protocol may be performed for internal planning purposes.

8.2 Safety Analysis for Interim CSR

The purpose of analysis is to support comprehensive safety review of avelumab in the ongoing trials, the results will be included into an interim CSR.

- The analysis for the interim CSR will be based on a data cut-off date of 20Nov2015.
- The data to be included into the first interim CSR are demographics and baseline characteristics, prior anti-cancer therapy, concomitant medication and procedure, drug exposure and compliance, premedication, safety parameters (AEs, lab, vital signs, electrocardiograms [ECGs], and Eastern Cooperative Oncology Group [ECOG]), immunogenicity, and PK.
- The data will be pooled across all the expansion cohorts for the analysis.
- The safety analysis set (SAF) used for the analysis will contain all the expansion subjects who receive at least one dose of study treatment by the data cut-off date.

8.3 Primary and Final Analyses

Primary analyses will include all the baseline characteristics, efficacy, and safety evaluations. Final analysis will contain efficacy evaluations and AEs. PK, pharmacodynamics, biomarker, and immunogenicity data may also be included for the primary or final analyses depending upon the purpose of reporting.

9 Changes to the Planned Analyses in the Clinical Trial Protocol

- Interim analysis for ovarian cancer cohort was performed on the first 23 dosed subjects because of the over-enrollment in the first stage of this cohort, with a data cut-off date on 17Jul2014 for the purpose of internal planning.
- Addition of a combined interim analysis performed on the first 90 NSCLC and 75 MBC dosed subjects with a data cut-off date on 17Jul2014 for the purpose of internal planning.
- The first interim analysis for secondary urothelial carcinoma cohort was performed on the first 26 subjects due to using the same cut-off date as 75 subject interim analysis for primary gastric/GEJ cancer cohort. This was an acceptable deviation from the planned interim analysis on the first 20 subjects. A second interim analysis will be performed on 44 subjects from secondary urothelial carcinoma cohort due to addition of efficacy urothelial carcinoma cohort in Protocol Amendment 10.
- Addition of analyses to explore the association between PD-L1 expression status and histology in NSCLC cohorts.
- Addition of analysis related to premedication.
- Addition of CD8+ and associated exploratory analyses in Section 16.3.
- The cut-off date for the interim analysis for the HNSCC cohort is September 09, 2015, defined as the date at which 20 subjects have been followed-up for at least 13 weeks (instead of 30

subjects as planned in the protocol). This analysis will include summary of baseline characteristics, efficacy and safety data. All dosed subjects until the data cut-off date will be included for baseline and safety analyses.

- Addition of an analysis on all the safety data including PK and immunogenicity to be included into an interim CSR.
- For the purpose of internal planning and for reporting to regulatory authorities, the primary analysis results of the urothelial carcinoma secondary cohort will be updated based on a data cut-off date on Oct 7, 2015, as well as on the cut-off dates for the urothelial carcinoma efficacy expansion cohort, i.e. i) 4 months after start of treatment of the 109th subject (Jan 18, 2016); ii) 6 months after start of treatment of the 109th subject (Mar 19, 2016), and iii) 6 months after start of treatment of the last subject enrolled in the efficacy expansion cohort. The efficacy endpoints such as BOR, PFS, OS, and safety endpoints such as the occurrence of TEAE will be included for these additional analyses.
- The interim analysis 13 weeks after the start of treatment of the 30th subject in the urothelial carcinoma efficacy cohort was not performed. Interim analyses will be conducted at 4 and 6 months after the last subject's first dose of study treatment for the 109 subjects enrolled in the urothelial carcinoma efficacy expansion cohort prior to Protocol Amendment 13. The results of these analyses may be subject to reporting to regulatory authorities. The interim analysis is considered positive if the lower limit of the 95% CI of the confirmed BOR exceeds 10%. The following subsets of the cohort are used for the interim analysis when 109th subject has been followed up for 4 months:
    - All subjects first dosed prior to or up to the cut-off date (analysis of safety)
    - All dosed subjects with at least 4 months follow-up as of the cut-off date (analysis of BOR, PFS, OS)
- The following subsets of the cohort will be used for the interim analysis when 109th subject has been followed up for 6 months:
    - All subjects first dosed prior to or up to the cut-off date (analysis of safety)
    - All dosed subjects with at least 13 weeks follow-up as of the cut-off date (BOR, PFS, OS, DR, TTR)
    - All dosed subjects with at least 6 months follow-up as of the cut-off date (BOR, PFS, OS, DR, TTR)
- Analyses on the corresponding PD-L1 positive subsets will be performed depending on PD-L1 data availability at the time of the analysis.
- CTCAE version 4.03 was utilized for laboratory toxicity grading in place of version 4.0 for outputs relating to the Interim Escalation and Expansion CSR.
- Addition of TTR according to modified irRC and to RECIST 1.1 criteria, respectively, per investigator assessment.
- For the efficacy expansion cohorts and the urothelial carcinoma secondary cohort:
    - TTR according to RECIST 1.1, per IERC

10 Analysis Sets

The following analysis sets are defined for the dose expansion phase and summarized in Table 4:

- Screening analysis set (SCRN): all subjects who signed informed consent form (ICF).
- PK analysis set (PKAS): a subset of the safety analysis set and will include patients who have at least one post-dose concentration measurement above the lower limit of quantitation (LLOQ) for avelumab.
- Safety analysis set (SAF): all subjects who have received at least 1 dose of study treatment.
- Full analysis set (FAS): all subjects who have received at least 1 dose of study treatment.
  - PD-L1 positive (5% cut-off) FAS (urothelial carcinoma efficacy expansion cohort): all PD-L1+ subjects (defined as those with at least 5% of the tumor cells showing PD-L1 membrane staining ≥1+ assessed by immunohistochemistry) who have received at least 1 dose of study treatment.
  - PD-L1 positive (1% cut-off) FAS (NSCLC cohorts): all PD-L1+ subjects (defined as those with at least 1% of the tumor cells showing PD-L1 membrane staining ≥1+ assessed by immunohistochemistry) who have received at least 1 dose of study treatment.
- Efficacy analysis set (EFF, efficacy expansion cohorts): all subjects who have received at least 1 dose of study treatment and have measurable disease at baseline according to IERC assessment.
  - PD-L1 positive EFF (urothelial carcinoma efficacy expansion cohort): all PD-L1+ subjects (defined as those with at least 5% of the tumor cells showing PD-L1 staining ≥1+ assessed by immunohistochemistry) who have received at least 1 dose of study treatment and have measurable disease at baseline according to IERC assessment.
- Efficacy analysis set (primary and secondary expansion cohorts): all subjects who have received at least 1 dose of study treatment and have measurable disease at baseline according to investigator assessment.

The definition of the SAF and the FAS are identical in this non-randomized study; the SAF will be used for the safety analysis and the FAS will be used for efficacy analysis. The PD-L1 positive FAS will be the primary analysis population for the primary endpoint of BOR by IERC in the urothelial carcinoma efficacy expansion cohort, whereas for the other efficacy expansion cohorts the primary analysis population will be the FAS.

Table 4 Summary of Analysis Sets and Associated Analyses

Analysis Set	Data/Endpoints	Cohorts
SAF	Disposition, baseline, safety, PD-L1 and biomarker, CD8, exposure, immunogenicity	All expansion cohorts
PKAS	PK data	All expansion cohorts
FAS	Efficacy (BOR, PFS, TTR, OS, subgroup analysis, lesion assessment)	All expansion cohorts
PD-L1 positive (5% cut-off) FAS	Efficacy (BOR, PFS, TTR, DR, OS, subgroup analysis, lesion assessment)	Urothelial carcinoma efficacy cohort
PD-L1 positive (1% cut-off) FAS	Efficacy (BOR, PFS, TTR, DR, OS, subgroup analysis, lesion assessment)	NSCLC cohorts
EFF	Efficacy (sensitivity analysis)	All expansion cohorts
PD-L1 positive EFF	Efficacy (sensitivity analysis)	Urothelial carcinoma efficacy cohort
EFF	BOR at 13 weeks (75 subjects interim analysis)	Primary cohorts

For interim analysis, all dosed subjects up to and including the data cut-off date will be included for the analysis. Additionally, two subsets of the population will be defined for interim analysis, dosed subjects with ≥ 6 or ≥ 13 weeks follow-up period up to the data cut-off date. The follow-up period is calculated from the first dose date to the data cut-off date. The analysis of BOR or associated subgroup analysis by PD-L1 status will be based on the two subset populations. The subset populations may also be used for the summary of demographic or other baseline data.

11 General Specifications for Statistical Analyses

- Statistical analyses will be performed using electronic case report form (eCRF) data obtained until a clinical cut-off date.
  - The primary data cut-off for expansion cohorts is 6 months after the last subject started the treatment.
  - The interim data cut-off is 13 weeks after the pre-specified number of subjects started the treatment. For example, an interim analysis will be conducted after the first 75 subjects in each primary cohort have reached the time point of the second post-baseline tumor assessment scheduled in Week 13, i.e. 13 weeks after start of treatment of the 75th subject. There may be minor deviation from this requirement due to combined data transfer used for multiple interim analyses, i.e. a few subjects may not reach the time point for the second post-baseline tumor assessment scheduled in Week 13, which is considered as acceptable.
  - Final data cut-off will be 1 year after the last subject in the expansion phase receives his or her last dose of avelumab.
- Analysis results will be summarized by cohort(s) and/or pooled across cohorts, depending upon the purpose of the reporting. Primary analysis for a cohort will include data summarized by that cohort. Final analysis will include data from all the expansion cohorts.
- Analysis of primary gastric and GEJ cancer cohort will be primarily stratified by the status of disease progression after first line chemotherapy.

- Data collected after re-initiation of treatment will not be included for safety and efficacy analyses except for PK, overall survival, and disposition.
- All data will be evaluated as observed, and no imputation method for missing values will be used, unless otherwise specified.
- Duration will be calculated as stop date – start date + 1, unless otherwise specified.
- The time since an event (e.g. time since first diagnosis, time since last dose) will be calculated as reference date minus date of event.
- The first day (Day 1) of study treatment is defined as the day of the first administration of avelumab, unless otherwise stated. The last dose date of study treatment is defined as the day of the last administration of avelumab, prior to the re-initiation of study treatment if applicable.
- Baseline is defined as the last non-missing observation prior to the administration of first dose of study treatment. Additionally, baseline for HR and QT/corrected QT (QTc) assessments will be derived from the visit where both HR and QT are not missing. If duplicate or triplicate ECGs are collected, baseline for each ECG measurement is the average of the pre-dose replicate measurements on the baseline day. QT interval correction based Fridericia's or Bazett's formula (QTcF/QTcB) will be derived based on HR and QT. The average of the replicate measurements should be determined after the derivation of the individual parameter at each time point.
- If the laboratory assessments are not done for scheduled visit but they are available for unscheduled visit on Day 1 from a different laboratory, the unscheduled visit will be included for the derivation of baseline; if there are multiple non-missing assessments on Day 1, the assessment from scheduled visit will be used for the derivation of baseline.
- On-treatment period will be defined as the time from the first dose of study treatment to min(last dose date + 30 days, earliest date of subsequent anti-cancer drug therapy – 1 day). If the earliest date of subsequent anti-cancer drug therapy is a partial date and only day is missing, it will be imputed as the last day of the month. If both day and month are missing, no imputation should be performed. The imputed date will be used for defining on-treatment period as well as confirming immune-related progressive disease (irPD).
- All statistical analyses will be performed using SAS® Version 9.1.3 or higher.
- There will be no difference between scheduled and un-scheduled visits except for by-visit analysis of safety analyses and baseline derivation.
- The assignments of visit windows are described in Table 5 for the purpose of by-visit analyses of safety data:
  - Baseline will be derived as described above.
  - No visit windowing will be performed at discontinuation, end of treatment, or safety follow-up visits for laboratory, vital sign, and ECG data, and 2hr post dose assessment on Week 1 Day 1 for ECG data. Instead, the earliest non-missing observation among the unscheduled or scheduled assessments for each visit (discontinuation, end of treatment, or safety follow-up) will be used for the analysis. For 2hr post dose assessment on Week 1 Day 1 ECG data, the earliest non-missing observation on Week 1 Day 1 will be used for the analysis.

- Scheduled and unscheduled assessments are included for visit windowing. Assessments on or after re-initiation of treatment are not be included for visit windowing.
- If there are multiple assessments for any specified visit and some of them are from scheduled visits, the assessment from scheduled visit with the closest distance to the planned study day will be used for analysis.
- If there are multiple assessments for any specified visit and none of them are from scheduled visits, the assessment with the closest distance to the planned study day will be used for analysis.
- If there are two or more unscheduled assessments with same distance to the planned study day such as (-1/+1 day), the assessment prior to the planned study day such as -1 day will be used for windowing.
- There is no difference for visit windowing between tests from core serum chemistry panel and tests from full serum chemistry panel. Some subjects had non-core serum chemistry tests assessed at the scheduled visits only intended for core serum chemistry. Multiple protocol amendments are also taken into the consideration, as the full serum chemistry and hematology panels were assessed weekly until week 7 and bi-weekly thereafter prior to the approval of Protocol Amendment 7.
- For ECG assessment associated with study treatment dose, only assessments where time point (prior to infusion or 2 hr after infusion) are not missing will be considered for the analysis.

Table 5 Visit Window Definition for Safety Assessment

Assigned Study Day (Inclusive)		**Planned Study Day (AWTARGET)**	**Analysis Visit (N) (AVISITN)**	**Analysis Visit (AVISIT)**	**Assessment**
From (AWLO)	**To (AWHI)**				
~	1		1	Baseline	Lab, Vital Sign, ECG
1	1	1	2	Week 1 Day 1*	ECG
5	11	8	3	Week 2 Day 5-11	Lab, Vital Sign
12	18	15	4	Week 3 Day 12-18	Lab, Vital Sign
5	18	15	4	Week 3 Day 5-18	ECG
19	25	22	5	Week 4 Day 19-25	Lab, Vital Sign
19	25	22	5	Week 4 Day 19-25	ECG
26	32	29	6	Week 5 Day 26-32	Lab, Vital Sign
26	36	29	6	Week 5 Day 26-36	ECG
33	39	36	7	Week 6 Day 33-39	Lab, Vital Sign

40	50	43	8	Week 7 Day 40-50	Lab, Vital Sign
37	50	43	8	Week 7 Day 37-50	ECG
51	64	57	10	Week 9 Day 51-64	Lab, Vital Sign, ECG
65	78	71	12	Week 11 Day 65-78	Lab, Vital Sign, ECG
79	92	85	14	Week 13 Day 79-92	Lab, Vital Sign
79	106	85	14	Week 13 Day 79-106	ECG
93	106	99	16	Week 15 Day 93-106	Lab, Vital Sign
107	120	113	18	Week 17 Day 107-120	Lab, Vital Sign
121	134	127	20	Week 19 Day 121-134	Lab, Vital Sign
107	148	127	20	Week 19 Day 107-148	ECG
135	148	141	22	Week 21 Day 135-148	Lab, Vital Sign
149	162	155	24	Week 23 Day 149-162	Lab, Vital Sign
163	176	169	26	Week 25 Day 163-176	Lab, Vital Sign
149	190	169	26	Week 25 Day 149-190	ECG
177	190	183	28	Week 27 Day 177-190	Lab, Vital Sign
191	204	197	30	Week 29 Day 191-204	Lab, Vital Sign
205	218	211	32	Week 31 Day 205-218	Lab, Vital Sign
191	232	211	32	Week 31 Day 191-232	ECG

* only applies to 2 hr post dose, AWLO = Analysis Window Beginning Timepoint, AWHI = Analysis Window Ending Timepoint, AWTARGET = Analysis Window Target, AVISIT = Analysis Visit, AVISITN = Analysis Visit (N).

- Presentation of continuous and qualitative variables:
  - Continuous variables will be summarized using descriptive statistics i.e., number of non-missing values and number of missing values, [i.e. n (missing)], mean, median, standard deviation, minimum, maximum and first and third quartile (Q1 and Q3). CI may be estimated for some of the endpoints, if appropriate.
  - Qualitative variables will be summarized by counts and percentages. Unless otherwise stated, the calculation of proportions will include the missing category. Therefore counts of missing observations will be included in the denominator and presented as a separate category.

12 Trial Subjects

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

12.1 Disposition of Subjects and Discontinuations

Summary of analysis sets will be tabulated using frequency and percentage by cohort(s) and/or pooled across cohorts, depending upon the purpose of the reporting, on all the subjects who signed ICF, the number of subjects in SAF will be used as the denominator:

- All subjects who signed ICF
- Number of subjects in the safety analysis set
- Number of subjects in the full analysis set
- Number of subjects in the PD-L1 positive full analysis set, if applicable
- Number of subjects in the efficacy analysis set, if applicable
- Number of subjects in the PD-L1 positive efficacy analysis set, if applicable
- Number of subjects in the full analysis set/PD-L1 positive full analysis set with 6/13 weeks (or 4 or 6 months) follow-up period, if applicable

One table will provide the reasons for permanent discontinuation of study treatment and for end of study as collected on the Treatment Termination, and End of Study (if data is available) eCRF pages, respectively. The number and percentage of subjects in each disposition category will be presented in the table based on the SAF analysis set:

- Number of subjects in the SAF analysis set
- Number of subjects still on treatment
- Number of subjects off-treatment
- Reasons off-treatment
  - Adverse event
  - Lost to follow-up
  - Protocol non-compliance
  - Death
  - Disease progression
  - Withdrew consent
  - Other
- Number of subjects in follow-up
- Number of subjects who discontinued from the study
- Reasons off-study
  - Study reached predefined end
  - Lost to follow-up
  - Death

- Withdrew consent
- Other

- Number of subjects with treatment reinitiated

The number and percentage of screened and/or dosed subjects may be summarized by geographic region, country and clinical site or by expansion cohort depending on the purpose of reporting.

The follow-up time (weeks) in the study will be calculated as (the analysis cut-off date – the first dose date + 1)/7. Summary statistics (mean, standard deviation, median etc.) will be presented in a table.

The listing of subject disposition will include all subjects who signed ICF (i.e. including screening failures). The listing will include the following information (if applicable): subject identifier, date of informed consent, included in the trial, reason for inclusion/exclusion, first/last dosing date, reason off-treatment, date and reason off-study, flags for SAF, FAS, and EFF (or PD-L1 positive FAS, and PD-L1 positive EFF for Urothelial carcinoma efficacy cohort), and flags to identify if subjects are included for each interim analysis (if applicable). The reason off-treatment will be retrieved from the End of Treatment eCRF page.

A secondary listing for reason for end of treatment due to AEs will also be provided. The listing will be restricted to the SAF subjects who discontinued study treatment for the primary reason of an AE, and will include the following information: subject identifier, first/last dosing date, date off-treatment, and the relevant AE system organ classes (SOCs), preferred terms (PTs) and AE relationship to the study treatment.

12.2 Protocol Deviations

12.2.1 Minor Protocol Deviation

A minor protocol deviation can be defined as any deviation from the study protocol that does not materially affect the safety of the subjects and/or the conduct of the study and/or its evaluation. An example of a minor protocol deviation would include a missed PK blood sample.

12.2.2 Important Protocol Deviation

Important (previously used terminology major) protocol deviations is one that materially affects the safety of the subjects and/or the evaluation of primary or key secondary efficacy endpoints of the study.

Current ICH and EU GCP guidelines list the important protocol deviations that must be listed in the clinical report. These include:

- subjects that are dosed on the study despite not satisfying the inclusion criteria;
- subjects that develop withdrawal criteria whilst on the study but are not withdrawn;
- subjects that receive the wrong treatment or an incorrect dose;

- subjects that receive an excluded concomitant medication.
- deviation from GCP.

Important protocol deviations will be based upon the eCRF database and determined for all subjects by either medical review processes or programming based on the inclusion/exclusion criteria or other criteria presented in the protocol. The results will be included into SDTM, if identified by means of medical review. The ADaM datasets will include both, those identified by medical review and those identified by programming.

Important protocol deviations are specified in Appendix 20.1, and will be summarized in a table and presented in a data listing. All protocol deviations included into SDTM will be presented in a data listing.

13 Demographics and Other Baseline Characteristics

The demographics and other baseline characteristics will be summarized on the SAF.

13.1 Demographics

The demographics and baseline characteristics table will include descriptive statistics for the following variables:

- Age (in years)
- Age category (<65/≥65 years)
  - 65-<75
  - 75-<85
  - ≥ 85
- Sex
- Race
- Pooled Geographical region
  - North America
  - Europe
  - Asia
  - Rest of the World (Australia and/or Latin America will be included as additional pooled geographical regions if including > 10% of the overall randomized population)

North America contains subjects from United States, Europe contains subjects from Belgium, Czech Republic, Germany, France, United Kingdom, Hungary, and Poland, and Asia contains subjects from South Korea and Taiwan.

- Height (cm)

- Weight (kg)
- Body Mass Index (BMI) (kg/m2)
- ECOG performance status
- Nicotine use status (Never used/ Regular user/ Occasional user/ Former user)

Baseline weight and height will be the last non-missing values prior to the first dose of study treatment from the Vital Signs eCRF page while baseline ECOG will be derived from the data collected on the ECOG eCRF page. Nicotine use status will be extracted from Nicotine Consumption eCRF page.

Age and BMI will be derived as:

- Age (year) = (date of informed consent – date of birth + 1)/365.25.
  - In case of missing day only: Age (years) = (year/month of given informed consent – year/month of birth)/12
  - In case only year of birth is given: Age (years) = (year of given informed consent - year of birth)
- BMI (kg/m^2) = weight(kg)/[height(m)]^2.

The integer part of the calculated age will be used for reporting purpose.

The listing of demographics and baseline characteristics will include the following information: subject identifier, age, sex, race, country/geographic region, height (cm), weight (kg), BMI (kg/m2), and ECOG.

The listing of nicotine consumption will be produced with the following data: nicotine use status, frequency of nicotine use, start/end date of nicotine consumption, nicotine consumption habit, and duration of consumption (years).

Lesion assessments at screening will be grouped into the following categories and summarized using descriptive statistics (count and percentage) in a table:

- Tumor size at baseline: The sum of target lesion diameters ≥ vs < median in each cohort.
- Presence of metastases at baseline (present, absent). Target or non-target lesions that are categorized as ‘metastasis’ are classified as metastases. This only applies to urothelial carcinoma cohorts.

Baseline albumin and hemoglobin will be classified as follows and summarized in a table using descriptive statistics (frequency and percentage).

- Albumin (< 35 g/L vs. ≥ 35 g/L)
- Hemoglobin (< 100 g/L vs. ≥ 100 g/L)

Baseline Bellmunt Score will be classified as 0, 1, 2, or 3 as a sum of the sub-scores of baseline ECOG, baseline Hgb, and baseline liver mets, defined as follows and summarized in a table using

descriptive statistics (frequency and percentage; for urothelial carcinoma secondary and efficacy cohorts).

	Baseline value	Bellmunt sub-score
Baseline ECOG	0	0
	>0	1
Baseline Hemoglobin	>=100 g/L	0
	< 100 g/L	1
Baseline Liver Metastasis	N	0
	Y	1

- Baseline eCcr (≤30, 30 - ≤50, 50 - <60, ≥60; for urothelial carcinoma secondary and efficacy cohorts)
- Time since last prior anti-cancer chemotherapy (<3, ≥3 - <6, ≥6; for urothelial carcinoma secondary and efficacy cohorts)

13.2 Medical History

Medical history will be coded using the latest available version of Medical Dictionary for Regulatory Activities (MedDRA). Medical history will be summarized as the numbers and percentages of subjects by MedDRA PT as event category and MedDRA SOC as summary category, and sorted by SOC and PT in alphabetical order. Each subject will be counted only once within each PT or SOC.

Listing of medical history data by subject will include coded terms and all the relevant data fields as collected on the Medical History eCRF page.

13.3 Disease History

Disease histories are collected on NSCL Cancer Diagnosis, Colorectal Cancer Diagnosis, Gastric and GEJ Cancer, Castrate-Resistant Prostate Cancer, Melanoma Diagnosis, Ovarian Cancer, Metastatic Breast Cancer, Adrenocortical Carcinoma Diagnosis, Mesothelioma Cancer Diagnosis, Renal Cell Carcinoma Diagnosis, Head and Neck Cancer Diagnosis, and Urothelial Carcinoma Diagnosis eCRF pages. Partial date will be imputed as described in the Section 18.1.

The disease history table will include descriptive statistics for the following variables:

- Sub-site of tumor

Summary of tumor sub-sites applies to NSCLC, MBC, Gastric and GEJ cancer, CRC, ovarian cancer, melanoma cancer, ACC, mesothelioma, urothelial carcinoma, HNSCC where pre-defined sub-sites or standard terms will be captured on the eCRF pages.

- Time since first diagnosis (years), defined as (the first dosing date – the date of first diagnosis)/365.25

- Time since metastatic or locally advanced disease (months), defined as (the first dosing date – the date of first occurrence of metastatic or locally advanced disease)/30.4375
- Time since last disease progression (months), defined as (the first dosing date - the date of last progression of disease)/30.4375
- Tumor Node Metastasis Classification of Malignant Tumors (TNM) at initial diagnosis
- TNM at study entry

Listing of disease history will be provided with all relevant data (tumor sub-site, initial diagnosis date, first occurrence of metastatic or locally advanced disease, date of last disease progression, TNM classification) and derived variables used in the above table.

Eligibility for platinum-based therapy is collected on Platinum Ineligibility eCRF page for urothelial carcinoma and HNSCC cohorts. It will be summarized for the following categories:

- Eligibility for platinum-based therapy
  - Yes
  - No
    - Impaired renal function
    - Hearing loss (25 decibels at 2 contiguous frequencies)
    - Peripheral neuropathy
    - Other

Listing of eligibility for platinum-based therapy will be provided with all relevant data (eligibility, reason for ineligibility, and date ineligibility defined).

13.4 PD-L1 Expression Status and Biomarker

PD-L1 expression status will be collected using Pathology Report Form. The percentages of viable tumor cells that exhibit PD-L1 membrane staining at any intensity are evaluated. No staining should be scored as “0”, weak staining as “1+”, moderate staining as “2+”, and strong staining as “3+”.

PD-L1 expression status will be classified as positive or negative based on the following cut-offs:

- For tumor cells:
  - Subjects will be considered PD-L1 expression positive (negative) if at least (less than) 5% of the tumor cells show PD-L1 membrane staining ≥1+, respectively. This will be used as the primary cut-off.
  - Subjects will be considered PD-L1 expression positive (negative) if at least (less than) 25% of the tumor cells show PD-L1 membrane staining ≥2+, respectively. This will be considered as secondary cut-off.

- Subject will be considered PD-L1 expression positive (negative) if at least (less than) 1% of the tumor cells show PD-L1 membrane staining ≥1+, respectively. This will be used as the tertiary cut-off.
- Subject will be considered PD-L1 expression positive (negative) if at least (less than) 50% of the tumor cells show PD-L1 membrane staining ≥1+, respectively. This will be used as the '50% cut-off'.
- Subject will be considered PD-L1 expression positive (negative) if at least (less than) 80% of the tumor cells show PD-L1 membrane staining ≥1+, respectively. This will be used as the '80% cut-off'.

- For Immune cells:

Subjects will be considered PD-L1 expression positive with regard to immune cell expression if tumor has 'PD-L1 hotspots' with at least 10% PD-L1 expressing immune cells. Subjects with an evaluable specimen not meeting this criterion are considered PD-L1 expression negative with regard to immune cell expression. The PD-L1 expression negative subjects will contain subjects from the following categories:

- Negative 1: if there are no tumor associated immune cells (TAICs) present in the specimen.
- Negative 2: if there are TAICs present but no 'PD-L1 hotspots'.
- Negative 3: if tumor has 'PD-L1 hotspots' with <1% PD-L1 expressing immune cells.
- Negative 4: if tumor has 'PD-L1 hotspots' with 1-9% PD-L1 expressing immune cells.

PD-L1 expression status will be summarized using the following variables:

- PD-L1 expression status based on tertiary cut-off for tumor cells (positive/ negative/ not evaluable)
- PD-L1 expression status based on primary cut-off for tumor cells (positive/ negative/ not evaluable)
- PD-L1 expression status based on secondary cut-off for tumor cells (positive/ negative 1, negative 2 / not evaluable)
  - Negative 1: subjects with <5% tumor cells with staining ≥1+
  - Negative 2: subjects with ≥5% tumor cells with staining ≥1+, but <25% tumor cells with staining ≥2+
  - Negative: combined subjects from Negative 1 and Negative 2 as mentioned above
- PD-L1 expression status based on '50% cut-off' for tumor cells (positive/ negative/ not evaluable)
- PD-L1 expression status based on '80% cut-off' for tumor cells (positive/ negative/ not evaluable)
- PD-L1 expression status based on immune cells (positive/ negative1, negative 2, negative 3, negative 4 / not evaluable if data is collected, or positive/ negative/ not present)

- % of tumor cells with any staining (grade ≥1+) as continuous variable
- % of tumor cells with at least 2+ staining as continuous variable

The percentage of tumor cells with any staining (grade ≥1+) stratified by PD-L1 expression status based on immune cells will be displayed graphically using a boxplot. The association between PD-L1 expression and NSCLC histology (adenocarcinoma, squamous cell carcinoma, others) will be summarized using frequency and percentage in a table.

For the gastric and GEJ cancer cohorts, PD-L1 expression is in addition scored with alternative scoring methodologies, where three scoring methods to evaluate PD-L1 expression status will be used as follows:

- Conventional tumor cell (TC) (cell number-based percentage) scoring method
- Immune cell (IC) rescoring method
- Aggregated (TC + IC) scoring method

  Aggregate PD-L1 expression score combines TC and IC scores, and the results will be categorized as follows:
  - <1% for both TC and IC vs ≥1% for either TC or IC
  - <5% for both TC and IC vs ≥5% for either TC or IC
  - <25% for both TC and IC vs ≥25% for either TC or IC

Additional details of these scoring algorithms are specified in the Pathology Report Form (PRF).

PD-L1 expression status (positive vs. negative) at baseline with different cutoff values will be summarized in total for each scoring method. PD-L1 assay status of percentage of PD-L1 positive tumor cells will be summarized as below:

- <1%
- ≥1% to <5% (cut-off is 1%)
- ≥5% to <25%, (cut-off is 5%)
- ≥25% (cut-off is 25%)

The dates of sample collection for PD-L1 expression analysis will be summarized using the following variables. All subjects with valid PD-L1 expression results will be included for the analysis.

- Time from sampling date to first dose date (months), defined as (the first dosing date – the date of tissue sampling)/30.4375
- Timing related to the first date of prior anti-cancer therapy for metastatic or locally advanced disease (before/ after). If the sampling date is prior to the first date of prior anti-cancer therapy for metastatic or locally advanced disease, it is considered as ‘before’; otherwise, it is considered as ‘after’.

CCI

[redacted]

CCI

[redacted]

CCI

[redacted]

14 Prior and Concomitant Medications/Procedures

Prior and concomitant anti-cancer therapy/ other medications will be coded using the latest available version of WHO Drug Dictionary, and summarized based on SAF.

14.1 Prior Anti-Cancer Therapies/Procedures

The prior anti-cancer treatments and procedures are collected under the Prior Anti-Cancer Drug Therapies Details, Prior Anti-Cancer Radiotherapies Details and Prior Anti-Cancer Surgeries Details eCRF pages.

The overall summary of prior anti-cancer treatments will include: the number and percentages of subjects by type of treatment, i.e.

- Number of subjects with at least one type of prior anti-cancer treatment
- Number of subjects with at least one prior anti-cancer surgery
- Number of subjects with at least one prior anti-cancer drug therapy
- Number of subjects with at least one prior anti-cancer radiotherapy

Summary of prior anti-cancer drug therapy will include the following variables for all the cohorts with exceptions for ovarian cancer cohort:

- Number of subjects with at least one prior anti-cancer drug therapy
- Number of any prior anti-cancer therapy lines: missing/ 1/ 2/ 3/ ≥4
- Number of any prior anti-cancer therapy lines as continuous variable
- Number of prior anti-cancer therapy lines for metastatic or locally advanced disease: missing/ 0/ 1/ 2/ 3/ ≥4. If the intent of therapy is metastatic, locally advanced, or palliative, it will be counted into therapy lines for metastatic or locally advanced disease
- Number of prior anti-cancer therapy lines for metastatic or locally advanced disease as continuous variable
- Type of prior anti-cancer therapy: chemotherapy/ antibody therapy/ kinase inhibitors/ hormonal therapy/ vaccines/ bone marrow transplant/ lymphocyte infusion/ other
- Intent of therapy: adjuvant / neo-adjuvant / metastatic / locally advanced / palliative
- Best response: CR/ PR/ PD/ stable disease (SD)/ unknown/ not assessable (NE)/ not applicable. Best response is derived from the last treatment regimen

For secondary ovarian cancer cohort, one additional category will be defined for the following two variables:

- Number of any prior anti-cancer therapy lines: missing/ 1/ 2/ 3/ 4/ ≥5
- Number of prior anti-cancer therapy lines for metastatic or locally advanced disease: missing/ 0/ 1/ 2/ 3/ 4/ ≥5.

For the primary gastric and GEJ cancer cohort, the following will be summarized for Not Progressed First Line Cancer (FLC) per drug category:

- Duration of latest prior anti-cancer regimen (Days)
- Duration of latest prior anti-cancer regimen start date to first study treatment (Days)

The prior anti-cancer drugs will also be extensively detailed with the number and percentage of subjects by the Anatomical Therapeutic Chemical (ATC) class level 2 and PT in a table. A subject will be counted only once within a given drug class and within a given drug name, even if he/she received the same medication at different times. If any prior anti-cancer medication is classified into multiple ATC classes, the medication will be summarized separately under each of these ATC classes. In case any specific medication does not have ATC classification level 2 coded term, it will be summarized under "Unavailable ATC classification" category. The summary will be sorted on decreasing frequency of drug class and decreasing frequency of drug name in a given drug class, based on the incidence in the "Overall" column. In case of equal frequency regarding drug class (respectively drug name), alphabetical order will be used.

The listings of prior anti-cancer treatments and procedures will also be provided: a) listing of prior anti-cancer drug therapies, b) listing of prior anti-cancer radiotherapies, and c) listing of prior anti-cancer surgeries. These will include subject identifier and all the relevant collected data-fields on the corresponding eCRF pages.

14.2 Prior and Concomitant Medications/Procedures

Prior and concomitant procedures are collected on the Concomitant Procedures Details eCRF page. Prior and concomitant medications are collected on the Concomitant Medications Details eCRF page.

Medications started prior to first dose date of study treatment and continued into the on-treatment period as well as those started during on-treatment period are referred to as concomitant medications. Prior medications are defined as the medications started and stopped prior to the first dose date of study treatment. Post medications are defined as any medications started after on-treatment period.

Summary of concomitant medications will include the number and percentage of subjects by ATC classification level 2 and PT. A subject will be counted only once within a given drug class and within a given drug name, even if he/she received the same medication at different times. If any concomitant medication is classified into multiple ATC classes, the medication will be summarize separately under each of these ATC classes. In case any specific medication does not have ATC classification level 2 coded term, it will be summarized under "Unavailable ATC classification" category. The summary of concomitant medications will be sorted on decreasing frequency of drug class and decreasing frequency of drug name in a given drug class, based on the incidence in the

"Overall" column. In case of equal frequency regarding drug class (respectively drug name), alphabetical order will be used.

Prior and concomitant medication data will be listed from the Concomitant Medications eCRF page. Following variables will be included in the prior and concomitant medication listing: subject identifier, prior/ concomitant/ post medication, and all corresponding data field on the corresponding eCRF page.

Prior and concomitant procedures data will be listed from the Concomitant Procedures Details eCRF page. Subject identifier and all collected data-field on the corresponding eCRF page will be included in the listing.

14.3 Subsequent Anti-Cancer Therapies/Procedures

Anti-cancer treatment after discontinuation will be provided in a data listing with data retrieved from Anti-Cancer Treatment After Discontinuation, Radiotherapy After Discontinuation, and Surgery After Discontinuation eCRF pages. A table for anti-cancer treatment after discontinuation will be added for primary analysis.

15 Treatment Compliance and Exposure

Analysis of exposure will be based on the calculated actual dose levels (total dose administered/weight, mg/kg). The last non-missing weight of the subject on or prior to the day of dosing will be used for the calculation.

The summary of treatment exposure and compliance based on the SAF analysis set will include the following variables per subject (a cycle refers to the planned dosing interval of two weeks):

- Treatment duration (in weeks), defined as (the last dose date – the first dose date + 14)/7
- Number of administrations as continuous variable
- Cumulative dose (mg/kg), defined as sum of actual dose levels
- Dose intensity (mg/kg/cycle), defined as cumulative dose (mg/kg) / (0.5 * treatment duration (week))
- Relative dose intensity (%), defined as actual dose intensity (mg/kg/cycle) * 100/ planned dose level (mg/kg/cycle).
- Relative dose intensity by the following categories:
  - >0.9
  - >0.8-0.9
  - <=0.8

Individual relative dose intensity (%) is calculated as actual dose level (mg/kg)/ planned dose level (mg/kg) × 100 for each administration of study medication. A dose reduction is defined as actual non-zero dose < 90% of planned dose, or individual relative dose intensity < 90%. A table based

on SAF will be prepared to summarize the number and percentage of subjects with at least one dose reduction, and a breakdown by the number of dose reductions (1/2/3/≥4).

Per protocol, avelumab will be administered as 1-hour i.v. infusion. Subjects will receive the study treatment once every 2 weeks. Dose delays will be grouped into the following categories based on the deviation of the actual to the planned treatment administration day (relative to the previous non-zero dose date): no delay (including 1-2 days delays), 3-6 days delay, 7 or more days delay. For example, if one subject receives the study treatment on day 1, then the next study treatment administration date will be on day 15; however, if the subject receives the study treatment at day 16 or 17, this is considered as 'no delay'. Any zero dose prior to the last treatment administration is considered as a dose interruption.

The summary of dose delays will be based on the SAF and include the following categories:

- No delay
- 3-6 days delay
- 7 or more days delay

The categorization is based on the maximum length of delay, i.e. the worst case of delay if subjects have multiple dose delays.

A listing of study treatment administration will include subject identifier, study day, # of days relative to prior treatment, infusion rate, most recent body weight prior to infusion, actual/planned dose, batch ID, dose reduction/dose delay or interruption, and other relevant information collected on the Cohort Treatment MSB0010718C Administration Details eCRF page. A subset of this listing will be created for subjects with at least one dose reduction.

A listing of treatment exposure and compliance will include subject identifier, assigned dose level, and above derived variables summarized in the tables.

In order to mitigate infusion-related reactions, a premedication regimen of 25 to 50 mg diphenhydramine and 650 mg acetaminophen (i.v. or oral equivalent) is mandatory 30 to 60 minutes prior to each dose of study treatment starting on January 29, 2014. The compliance to this requirement will be summarized as numbers of subjects with 0, 1, 2, 3, 4 doses among the first 4 treatment administrations that were administered without pre-medication. For example: if a subject discontinued after 3 doses, and 2 of them were administered with premedication, the number for that subject would be 1.

A listing of premedication will include subject identifier, reported medication term, the relative time to the start of infusion, date/time of premedication, and dose (unit). A listing containing subject identifier, visit, and unique study treatment batch ID will also be created.

16 Endpoint Evaluation

The subsections in this section include specifications for analyzing clinical trial endpoints specified in the Clinical Trial Protocol to meet the trial objectives, as well as any endpoints not identified in the Clinical Trial Protocol.

The primary endpoint for expansion phase is

- The confirmed BOR, per RECIST 1.1, as adjudicated by an IERC for subjects enrolled in the efficacy expansion cohorts only, which is addressed in the Section 16.1.

The secondary endpoints (excluding AE) for expansion phase are:

- irBOR and BOR according to modified irRC and to RECIST 1.1 criteria, respectively, per investigator assessment.
- The confirmed BOR, per RECIST 1.1, as adjudicated by an IERC for subjects enrolled in the urothelial carcinoma secondary cohort.irPFS time and PFS time according to modified irRC and to RECIST 1.1 criteria, respectively, per investigator assessment.
- OS time.
- For the primary expansion cohorts only: Unconfirmed response at Week 13 according to RECIST 1.1 criteria, per investigator assessment.
- DR according to modified irRC and to RECIST 1.1 criteria, respectively, per investigator assessment.
- TTR according to modified irRC and to RECIST 1.1 criteria, respectively, per investigator assessment.
- For the efficacy expansion cohorts and the urothelial carcinoma secondary cohort:
    - PFS time, according to RECIST 1.1, per IERC
    - DR according to RECIST 1.1, per IERC
    - TTR according to RECIST 1.1, per IERC
- PK profile.
- Pharmacodynamic profile
- Serum titers, isotypes, and neutralizing capacity of anti-avelumab antibodies.
- Expression of PD-L1 on tumor tissue.

Secondary efficacy endpoints are addressed in Section 16.2. The secondary endpoints PK and immunogenicity are addressed in Section 16.4 and 16.5, respectively. CCI [redacted]

16.1 Primary Endpoint Analyses

The primary endpoint in the efficacy expansion cohorts is the confirmed BOR according to RECIST 1.1 and as adjudicated by an IERC, defined as the best response obtained among all tumor assessment visits after start of study treatment until documented disease progression (taking into account the requirement for confirmation). Only tumor assessments performed before the start of any further anti-cancer treatment will be considered in the assessment of BOR.

For the gastric / GEJ cancer, HNSCC, and ovarian cancer efficacy expansion cohorts, the primary analysis of the BOR by IERC will be conducted in the FAS, defined as all treated subjects. The number and proportion of BOR (defined as CR + PR) will be tabulated. The ORR will be determined as the proportion of subjects with a confirmed BOR of PR or CR. An exact binomial test will be performed at a 1-sided alpha level of 0.025. The primary analysis is planned 6 months after start of treatment of the last subject in the given cohort. Interim analyses will be conducted after 60% of the subjects in the given cohort have been followed up for 13 weeks. Analyses are considered positive if the lower limit of the 95% confidence interval exceed 10%. Confidence intervals will be constructed using the Clopper-Pearson method.

For the urothelial carcinoma efficacy expansion cohort, the analysis of the BOR by IERC will be conducted in the PD-L1 positive FAS followed by the FAS. The number and proportion of BOR (defined as CR + PR) will be tabulated. The ORR will be determined as the proportion of subjects with a confirmed BOR of PR or CR. An exact binomial test will be performed in the PD-L1 positive FAS and in the FAS to determine whether the null hypothesis of an ORR $\leq$ 10% can be rejected at the 1-sided alpha level of 0.025. Interim analyses will be conducted at 4 and 6 months after the first dose date of last subject of the 109 subjects enrolled in the urothelial carcinoma efficacy expansion cohort prior to Protocol Amendment 13. Respective cut-off dates are 18Jan2016 and 19Mar2016. Analyses are considered positive if the lower limit of the 95% CI of the confirmed BOR exceeds 10%. Confidence intervals will be constructed using the Clopper-Pearson method.

The tumor response will be based on the IERC assessment of overall response at each time point. Details of determination of tumor response are provided in Imaging Review charter document. A separate Imaging Data Management Plan and Data Transfer Plan will be created to summarize the details of the data structure and data delivery schedule of IERC assessment results.

The following are the requirements for confirmation of CR or PR or of minimum SD duration:

- CR or PR needs to be confirmed at a subsequent tumor assessment, preferably at the regularly scheduled 6-week assessment interval, but no sooner than 4 weeks after the initial documentation of CR or PR.
- The minimum duration for a BOR of SD is defined as at least 37 days after start of study treatment accounting for permitted deviations from the tumor assessment visit schedule.

Table 6 summarizes the derivation rules for the BOR when confirmation from subsequent assessment is needed (1). It is reasonable to consider a subject with time point response of PR-SD-

PR, PR-NE-PR or CR-NE-CR as a confirmed response as long as the second CR or PR is >= 28 days away from the first time point.

Table 6 Best overall response when confirmation of CR/PR required

Initial overall response	Subsequent overall response	Confirmed time point overall response
CR	CR	CR provided subsequent CR is >=28 days away from the first time point
CR	PR	SD provided minimum criteria for SD duration met; otherwise, PD
CR	SD	SD provided minimum criteria for SD duration met; otherwise, PD
CR	PD	SD provided minimum criteria for SD duration met; otherwise, PD
CR	NE	SD provided minimum criteria for SD duration met, otherwise, NE
PR	CR	PR provided subsequent CR is >=28 days away from the first time point
PR	PR	PR provided subsequent PR is >=28 days away from the first time point
PR	SD	SD
PR	PD	SD provided minimum criteria for SD duration met, otherwise, PD
PR	NE	SD provided minimum criteria for SD duration met, otherwise, NE
SD	Any	SD provided minimum criteria for SD duration met, otherwise, NE
PD	Any	PD
NE	NE	NE

CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease, and NE = Not evaluable.
If a subject only has a response value of NE or the only response value is SD and is within 36 days of the first dose date the best response will be NE.

In addition, the unadjusted 95% CIs of the ORR will also be calculated with the Clopper-Pearson method at the interim and primary analyses in order to provide comparable results based on the FAS and subgroup analysis. The number and percentage of subjects with BOR of CR, PR, SD, Non-CR/Non-PD, PD, and NE will be tabulated.

16.2 Secondary Endpoint Analyses

Clinical efficacy parameters will be analyzed descriptively in the FAS, and, in addition, for the urothelial carcinoma efficacy expansion cohort, in the PD-L1 positive FAS. Pooling of data from subjects with at least 6 months follow-up period in the urothelial carcinoma efficacy and secondary cohorts may be considered to enhance precision of estimates if data is sufficiently homogeneous. The following efficacy endpoints will be considered: The confirmed and unconfirmed BOR, per RECIST 1.1, as adjudicated by an IERC and per investigator assessment; irBOR according to modified irRC per investigator assessment, DR, TTR, and PFS time according to RECIST 1.1 criteria per IERC and investigator assessment; DR, TTR, and irPFS time according to irRC criteria

per investigator assessment; OS time. Subgroup analyses specified in Section 16.2.6 and summary of demographics will also be presented for pooled population.

The percent change in target lesions from baseline will be derived as:

- ((Sum of target lesions at week XX – sum of target lesions at baseline)/sum of target lesions at baseline)* 100%

The maximum reduction in target lesions from baseline will be derived across all the post-baseline assessments as:

- Minimum of ((sum of target lesions at week XX – sum of target lesions at baseline)/sum of target lesions at baseline) * 100%

The tumor shrinkage will be calculated based on investigator assessment for all the expansion cohorts and, in addition, based on IERC data for efficacy expansion cohorts and the urothelial carcinoma secondary cohort. The percent change from baseline in target lesions per time point as well as other relevant information will be presented in a data listing. The percent change from baseline in target lesions as well as the first occurrence of new lesion and subject off treatment overall and by PD-L1 expression status will be displayed against time point (weeks) in spider plots. The maximum reduction from baseline in the sum of target lesion diameters overall and by PD-L1 expression status will be presented per subject in waterfall plots. The PD-L1 expression status will be displayed using different colors or line styles in the plots. The time point of a tumor assessment per investigator is defined as the earliest scan date of the respective visit.

BOR per investigator assessment will be determined according to RECIST 1.1 for all the expansion cohorts. CR/PR will be confirmed per Table 5 in Section 16.1. irBOR is defined as the best result obtained among all tumor assessment visits from baseline until immune related disease progression (i.e., confirmed irPD), and will be determined according to modified irRC per investigator assessment, taking confirmation requirements into account as detailed below. Only tumor assessments performed before the start of any further anti-cancer treatment will be considered in the assessment of BOR/irBOR. In addition to the confirmed BOR/ irBOR, the unconfirmed BOR will be derived for the interim and/or primary analyses. The date of unconfirmed BOR / confirmed BOR / irBOR will be the date of the best result that is first observed, or first confirmed if the confirmation is required. In case of different dates of scans within the same tumor assessment visit, the earliest scan date should be used as the date of tumor assessment.

Subjects with a BOR of NE will be summarized by reason for having NE status and displayed in a listing with relevant data. The following reasons will be used:

- No post-baseline assessments
- All post-baseline assessments have overall response NE
- New anticancer therapy started before first post-baseline assessment
- SD of insufficient duration (<6 weeks after start date without further evaluable tumor assessment)
- PD too late (>12 weeks after start date)

A contingency table will be created to compare the following tumor assessment results between IERC and investigators:

- BOR (NE / PD /(Non-CR/Non-PD) / SD / CR / PR / CR+PR)
- Disease Progression (No event / PD / Death)

For irBOR, the response of immune-related complete response (irCR), immune-related partial response (irPR), and irPD need to be confirmed by a second, consecutive assessment at least 4 weeks apart as described in Table 7 (1, 2). It is reasonable to consider a subject with time point response of irPR-irSD-irPR, irPR-NE-irPR or irCR-NE-irCR as a confirmed response as long as the second irCR or irPR is more than 28 days away from the first time point. irPD is also considered to be confirmed if the following event occurs:

- If subject is assessed with time point response of irPD-NE-irPD as long as the second irPD is more than 28 days away from the first time point; or
- If subject dies within 84 days after the initial observation of irPD; or
- If subject receives subsequent anti-cancer drug therapy within 84 days after the initial observation of irPD; or
- If subject experiences clinical deterioration as assessed by investigator and recorded as reason for treatment discontinuation prior to or within 84 days after the assessment of irPD.

In case a subject with a confirmed CR relapses within 1 year after stopping treatment, one re-initiation of treatment is allowed according to Protocol. If the subject has irPD assessed at the same visit as the PD prior to the initiation of treatment, the irPD will be considered as confirmed. Immune-related stable disease (irSD) duration is required to be no less than 37 days from Day 1.

Table 7 Immune-related BOR when confirmation of irCR, irPR, irPD required

Initial overall response	Subsequent overall response	Confirmed time point overall response
irCR	irCR	irCR provided subsequent irCR is >=28 days away from the first time point
irCR	irPR	irSD provided minimum criteria for irSD duration met; otherwise, irPD
irCR	irSD	irSD provided minimum criteria for irSD duration met; otherwise, irPD
irCR	irPD	irSD provided minimum criteria for irSD duration met; otherwise, irPD
irCR	NE	irSD provided minimum criteria for irSD duration met, otherwise, NE
irPR	irCR	irPR provided subsequent irCR is >=28 days away from the first time point
irPR	irPR	irPR provided subsequent irPR is >=28 days away from the first time point
irPR	irSD	irSD
irPR	irPD	irSD provided minimum criteria for irSD duration met, otherwise, irPD
irPR	NE	irSD provided minimum criteria for irSD duration met, otherwise, NE
irSD	Any	irSD provided minimum criteria for irSD duration met, otherwise, NE
irPD	irPD	irPD provided subsequent irPD is >=28 days away from the first time point, death or take subsequent anti-cancer drug therapy within 84 days after the first time point, or clinical deterioration.

Initial overall response	Subsequent overall response	Confirmed time point overall response
irPD	Missing	irPD if subject dies or takes subsequent anti-cancer drug therapy within 84 days after the initial observation of irPD, or clinical deterioration, otherwise, NE
NE	NE	NE

irCR = immune-related complete response, irPR = immune related partial response, irSD = immune related stable disease, irPD = immune related progressive disease, and NE = Not evaluable.

If a subject only has a response value of NE or the only response value is irSD and is within 36 days of the first dose date the best response will be NE

In general, efficacy analyses on secondary endpoints will be performed as follows:

- Unconfirmed response at week 13 only applies to the interim analyses performed on the first 75 subjects for all the primary cohorts.
- Unconfirmed BOR and PFS per investigator assessment applies to all the interim analyses for all the expansion cohorts.
- Unconfirmed/confirmed BOR and PFS per IERC apply to the interim analyses for the efficacy expansion cohorts and the urothelial carcinoma secondary cohort if data is available.
- All the endpoints (unconfirmed/confirmed BOR, irBOR, PFS/irPFS, DR/immune-related DR, TTR, OS) per investigator assessment except for unconfirmed response at week 13 apply to primary analysis for all the expansion cohort. PFS, TTR and DR per IERC will also be analyzed for efficacy expansion cohort and the urothelial carcinoma secondary cohort if data is available.

The details will be provided in the Table, Listing, and Figure Shells document.

16.2.1 Objective Tumor Response According to RECIST 1.1 or Modified irRC (per investigator assessment)

Objective tumor response is defined as having a BOR assessment of unconfirmed/confirmed CR/PR, or having an irBOR assessment of irCR/irPR. Objective tumor response will be evaluated by ORR or immune-related objective response rate (irORR) for each cohort, defined as the number of subjects reached a best overall response of CR/PR (irCR/irPR) divided by the number of subjects in the FAS (PD-L1 positive FAS)/ EFF (PD-L1 positive EFF). Time to objective response will be calculated as:

(Date of the first documented objective response (PR or CR) – date of the first dose +1)/7 (weeks)

Two-sided 80% and 95% exact CIs for ORR or irORR will be estimated using Clopper-Pearson method for each cohort. Additionally, the number and percentage of subjects with BOR of CR, PR, SD, PD, and NE or irBOR of irCR, irPR, irSD, irPD, and NE will be tabulated.

These evaluations will be presented in data listings with detailed information collected per eCRF pages as well as BOR and irBOR for all the subjects from the FAS subjects. Time to response, time of progression, and duration of study treatment will be displayed in a bar chart for subjects with objective response.

16.2.2 Duration of Response According to RECIST 1.1 or Modified irRC

DR is measured from the time measurement criteria are first met for CR/PR or irCR/irPR (whichever is first recorded) until the first date that progressive disease or death within 84 days of last tumor assessment is objectively documented. The analysis of DR will be performed among the subjects who had unconfirmed/confirmed CR/PR or irCR/irPR.

DR will be censored in the following scenarios:

- Subjects who have not experienced an event (PD or death) will be right-censored on the date of their last evaluable tumor assessment.
- If death without previously documented PD is observed after more than 84 days (12 weeks) of last tumor assessment, subject will be right-censored at the date of the last evaluable tumor assessment.

DR = (date of PD or death/censoring – date of first documented objective response + 1)/30.4375 (months).

The Kaplan-Meier method will be used to estimate parameters for duration of response. Kaplan-Meier estimates (product-limit estimates) will be presented together with a summary of associated statistics including corresponding two-sided 95% CIs. In particular, the proportion of duration of responses at 6 and 12 months will be estimated with corresponding two-sided 95% CIs. The CIs for the median will be calculated according to Brookmeyer and Crowley (4) and CIs for the survival function estimates at the above defined time points will be derived using the log-log transformation according to Kalbfleisch and Prentice (5) (CONFTYPE = loglog default option in SAS PROC LIFETEST). The estimate of the standard error will be computed using Greenwood's formula. A listing with pertinent information will be provided.

16.2.3 Progression-Free Survival According to RECIST 1.1 or Modified irRC

PFS time is defined as the time from first administration of study treatment until date of the first documentation of PD or death by any cause (whichever occurs first), when death occurs within 84 days of last tumor assessment or first administration of study treatment (whichever is later).

PFS will be censored in the following scenarios:

- Subjects who do not have any post-baseline tumor assessment, or do not have a baseline tumor assessment, and die (when applicable) more than 84 days after initial avelumab dose will be right-censored on the date of first dose of study treatment.
- Subjects who have not experienced an event (PD or death) will be right-censored on the date of their last evaluable tumor assessment.
- If death without previously documented PD is observed after more than 84 days of last evaluable tumor assessment, subject will be right-censored at the date of the last evaluable tumor assessment.

PFS = (date of PD or death/censoring - date of the first dose + 1)/30.4375 (months).

The analysis of PFS will be performed with a Kaplan-Meier method with the same approach as for duration of response described in Section 16.2.2. Kaplan-Meier estimates (product-limit estimates) will be presented together with a summary of associated statistics including corresponding two-sided 95% CIs. In particular, the proportion of PFS at 3, 6, and 12 months will be estimated with corresponding two-sided 95% CIs. Kaplan-Meier plots of PFS time and listing of PFS will be provided as well.

Frequency (number and percentage) of patients with each event type (PD or death) and censoring reasons will be presented. Censoring reasons are as follows:

- Ongoing in the study without an event (PD or death)
- No baseline assessment
- No adequate post-baseline assessments
- No documented PD and death more than 84 days after last evaluable tumor assessment.

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

16.2.4 Overall Survival

OS is defined as the time from first dose to death due to any cause. For subjects who are still alive at the time of data analysis or who are lost to follow up, OS time will be censored at the date of last contact as specified in Section 18.1.

OS = (date of death/censoring – date of the first dose + 1)/30.4375 (months).

The analysis of OS time will be performed with a Kaplan-Meier method with the same approach as for duration of response described in Section 16.2.2. Kaplan-Meier estimates (product-limit estimates) will be presented together with a summary of associated statistics including corresponding two-sided 95% CIs. In particular, the proportion of overall survival at 6, 12, and 24 months will be estimated with corresponding two-sided 95% CIs.

The Kaplan-Meier plots of OS time and listing of OS will provided as well.

Frequency (number and percentage) of patients with an event type (death) and censoring reasons will be presented. Censoring reasons are as follows:

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

Lost to follow-up will include the following subjects:

- Lost to follow-up status is collected on the eCRF page prior to the analysis cut-off;

- Subjects with the last contact date > 14 weeks prior to the analysis cut-off date (duration of 14 weeks is based on the assessment schedule of every 12 week for survival follow-up interval + 2 week window).

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

16.2.5 Time to Response per RECIST v1.1 or Modified irRC

TTR will be analyzed for subjects with confirmed CR/PR based on investigator and/or IERC assessments and subjects with irCR/irPR based on investigator assessment.

TTR (in weeks) = (first date of objective response – first dose date +1)/7

TTR will be summarized using simple descriptive statistics (mean, SD, median, min, max. Q1, Q3).

16.2.6 Subgroup Analyses

Subgroup analyses will be performed according to the following parameters:

- PD-L1 expression status based on:
  - Tumor cell (positive, negative)
    - Positive: ≥1% of the tumor cells show PD-L1 membrane staining intensity ≥1+. Negative: <1% of the tumor cells show PD-L1 membrane staining intensity ≥1+. Evaluable: positive + negative.
    - Positive: ≥5% of the tumor cells show PD-L1 membrane staining intensity ≥1+. Negative: <5% of the tumor cells show PD-L1 membrane staining intensity ≥1+. Evaluable: positive + negative.
    - Positive: ≥25% of the tumor cells show PD-L1 membrane staining intensity ≥2+. Negative 1: <5% of the tumor cells show PD-L1 membrane staining intensity ≥1+. Negative 2: ≥5% of the tumor cells show PD-L1 membrane staining intensity ≥1+ but <25% of the tumor cells show PD-L1 membrane staining intensity ≥2+. Negative: negative 1 + negative 2. Evaluable: positive + negative.
    - Positive: ≥50% of the tumor cells show PD-L1 membrane staining intensity ≥1+. Negative: <50% of the tumor cells show PD-L1 membrane staining intensity ≥1+. Evaluable: positive + negative.
    - Positive: ≥80% of the tumor cells show PD-L1 membrane staining intensity ≥1+. Negative: <80% of the tumor cells show PD-L1 membrane staining intensity ≥1+. Evaluable: positive + negative.

Immune cell (positive, negative, evaluable (positive + negative)). Subjects will be considered PD-L1 positive if tumor has hotspots with at least 10% PD-L1 expressing immune cells (see also Section 13.4).

- Demographics
  - Age (<65, ≥65 years)
  - Sex (male, female)
  - Race (white, others)
  - ECOG (0, ≥1)
- Pooled Geographical region
  - North America
  - Europe
  - Asia
- Region (Asia, Non-Asia) for gastric and GEJ cohorts
- Smoking status: never smoked, ever smoked (containing regular user, occasional user, and former user).
- Histology (adenocarcinoma, squamous cell carcinoma, others). This only applies to NSCLC cohorts.
- Tumor sub-site (oral cavity, oropharynx and hypopharynx, larynx). This only applies to HNSCC cohorts.
- Tumor size at baseline: sum of target lesion diameters ≥ vs < median
- Presence of metastases at baseline (present, absent). Target or non-target lesions that are categorized as 'metastasis' are classified as metastases.
- Prior anti-cancer drug therapy
  - # of any prior anti-cancer therapy lines (≤1, 2, ≥3).
  - # of prior anti-cancer therapy lines for metastatic or locally advanced disease (≤1, 2, ≥3).
  - # of any prior anti-cancer therapy lines (≤1, >1: for NSCLC post platinum doublet).
  - # of prior anti-cancer therapy lines for metastatic or locally advanced disease (≤1, >1: for NSCLC post platinum doublet).
- NSCLC biomarker
  - EGFR (normal, abnormal, unknown)
  - KRAS (normal, abnormal, unknown)
  - ALK (normal, abnormal, unknown)
  - EGFR/ALK (normal, abnormal, unknown)
- MBC biomarker: HER2- (ER- and PR-), HER2- (ER+ or PR+), HER2+, or unknown

- Gastric and GEJ cancer biomarker: HER2+, HER2-, or unknown
- HNSCC biomarker: HPV+, HPV-, or unknown
- Gastric and GEJ cancer and ovarian cancer biomarker: MSI (low vs. stable vs. high)
- Gastric and GEJ cancer biomarker: EBV (positive vs. negative)
- MBC and ovarian cancer biomarker: BRCA1/2, mutant (if at least one mutation in both genes) vs. wildtype (no mutation in BRCA1 and BRCA2)
- Gastric and GEJ cancer biomarker: PD-L1 expression status according to each of the three alternative scoring methods (conventional tumor cell (TC) scoring, immune cell (IC) scoring, aggregated (TC + IC) scoring): PD-L1 assay status at cut-off value as follows:
    - < 1% vs. >=1%
    - < 5% vs. >=5%
    - < 25% vs. >=25%
- Eligibility for platinum-based therapy (yes, no) for urothelial carcinoma cohort
- Baseline laboratory assessment (only presented for urothelial carcinoma efficacy and secondary cohorts)
    - Albumin (< 35 g/L vs. ≥ 35 g/L)
    - Hemoglobin (< 100 g/L vs. ≥ 100 g/L)
- Baseline Bellmunt Scores (only presented for urothelial carcinoma efficacy and secondary cohorts)
- Time since last prior anti-cancer chemotherapy (only presented for urothelial carcinoma efficacy and secondary cohorts)
- Sub-site of tumor (only presented for urothelial carcinoma efficacy and secondary cohorts)
    - Upper tract defined as ureter or renal pelvis
    - Lower tract defined as bladder or urethral

Frequency of objective response, percentage, and 95% CI of ORR will be estimated for all the subgroup analyses on BOR or irBOR. CI will be based on exact Clopper-Pearson method. For subgroup analyses of BOR or irBOR based on PD-L1 expression status, the p-value and 80% CI will be provided as well. The p-value will be based on Fisher's exact test for association between PD-L1 status (positive vs. negative) and ORR.

Same statistics will be provided for the subgroup analyses on PFS, irPFS, and OS as those provided for non-subgroup analyses on PFS, irPFS, and OS, respectively. For subgroup analysis by PD-L1 status, the hazard ratios and their associated 95% CIs will be estimated by Cox Proportional Hazards model using PD-L1 status (negative as reference) as covariate.

The subgroup analyses are exploratory for all the cohorts and will be primarily performed on IERC data for efficacy expansion cohorts and the urothelial carcinoma secondary cohort. Subgroup

analysis will not be performed if the largest subgroup covers ≥ 90% of the subjects. Subgroup analysis to be performed depends on the cohorts and type of analysis, but in general:

- Subgroup analysis by PD-L1 status applies to interim and primary analyses for all the expansion cohorts if PD-L1 data is available.
- Subgroup analyses by demographic and prior anti-cancer therapy line apply to all the expansion cohorts. Subgroup analysis by sex is excluded for MBC, CRPC, and ovarian cancer cohorts.
- Subgroup analysis by geographic region applies to all the cohorts with subjects enrolled from different regions.
- Subgroup analysis of tumor response (confirmed/unconfirmed BOR, irBOR) by baseline tumor size applies to all expansion cohorts.
- Subgroup analysis of tumor response (confirmed/unconfirmed BOR, irBOR) and PFS by the status of metastases at baseline applies to urothelial carcinoma cohorts only.
- Subgroup analysis by biomarker applies to interim and primary analyses for NSCLC, MBC, gastric and GEJ cancer, HNSCC cohorts based on the above definitions.
- Subgroup analysis by histology or tumor sub-site applies to NSCLC or HNSCC cohorts, respectively.
- Subgroup analysis by smoking status applies to NSCLC and HNSCC cohorts.
- Subgroup analyses by eligibility of platinum-based therapy and baseline albumin and hemoglobin apply to urothelial carcinoma cohorts only.

Subgroup analyses may be also performed on selected sub-population such as PD-L1+ subjects based on tertiary cut-off defined in Section 13.4. The details will be provided in the Table, Listing, and Figure Shells document.

The association between unconfirmed/confirmed objective response and PD-L1 status (based on tertiary cut-off) will be stratified by PD-L1 sampling time related to the first date of prior anti-cancer therapy for metastatic or locally advanced disease (before, after) and summarized in a table. This will apply to post-platinum doublet NSCLC cohort.

The maximum target lesion percentage reduction vs. percentage of tumor cells with any staining (≥1+) or with ≥2+ staining will be displayed graphically using scatterplot.

Subgroup information and BOR, PFS, or OS will be presented in three separate data listings.

16.2.7 Sensitivity Analyses

If there is a more than 5% difference in the number of subjects between FAS and EFF analysis set or between PD-L1 positive FAS and PD-L1 positive EFF analysis set, tumor response based on BOR and irBOR per IERC or investigator assessment will be repeated using EFF analysis set or PD-L1 positive EFF analysis set.

CCI

[redacted]

CCI

[redacted]

16.4 Pharmacokinetics and Pharmacodynamics

PK analysis will be based on the PK analysis set.

16.4.1 Missing PK Data

Concentrations below the limit of assay quantitation

PK concentrations below the lower limit of quantification (<LLOQ) are taken as zero for descriptive statistics.

PK concentrations below the lower limit of quantification (<LLOQ), which are before the last quantifiable data point, will be taken as zero for calculating the AUC of single dose profiles. Concentration below LLOQ, which occur after the last quantifiable data point, will not be considered in the calculation of the terminal first order rate constant (λ_z).

Deviations, missing concentrations and anomalous values

Concentrations reported as no result will be set to missing in summary tables.

If a PK parameter cannot be derived from a subject'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 subject discontinues the treatment). For statistical analyses (i.e. analysis of variance), PK parameters coded as NC will be set to missing.

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

Relevant decisions on subject inclusion in the PK analysis set will be made before database lock in the Database Review Meeting (DRM), as far as possible. Remaining decisions will be made

prior to the performance of a descriptive analysis. PK concentrations which are erroneous due to a protocol violation (as defined in the CTP), documented handling error or analytical error (as documented in the bioanalytical report) may be excluded from the PK analysis if agreed upon prior to performing a statistical analyses. In this case the rational for exclusion must be provided in the Clinical Trial Report (CTR). Any other PK concentrations that appear implausible to the Pharmacokineticist/PK/PD Data Analyst must not be excluded from the analysis. Any implausible data will be documented in the Clinical Trial Report (CTR). Any exclusions will be listed and flagged.

16.4.2 Descriptive PK Analysis

Avelumab serum concentrations will be provided in listings and descriptively summarized by day and nominal time using the number of non-missing observations (N), arithmetic mean (Mean), standard deviation (SD), coefficient of variation (CV%), minimum (Min), median (Median) and maximum (Max). The pre-dose samples will be considered as if they had been taken simultaneously with the administration.

PK concentrations at the end of each infusion (C_{EOI}) and trough concentrations (C_{trough} - concentration at the end of a dosing interval) will be summarized descriptively by nominal day on scheduled visits for all tumor types. Descriptive statistics for C_{trough} and C_{EOI} will additionally show the geometric mean (GeoMean), the geometric coefficient of variation (GeoCV) and the 95% confidence interval for the GeoMean (Lower CI 95% GeoMean, Upper CI 95% GeoMean).

Descriptive statistics of PK concentration data will be calculated using values with the same precision as the source data, and rounded for reporting purposes only. The following conventions will be applied when reporting descriptive statistics of PK/PD concentration data:

Mean, Min, Median, Max:	3 significant digits
SD:	4 significant digits
CV%:	1 decimal place

The following conventions will be applied when reporting descriptive statistics of C_{EOI} and C_{trough} data:

Mean, Min, Median, Max, GeoMean, 95% CI:	3 significant digits
SD:	4 significant digits
CV%, GeoCV%:	1 decimal place

Individual C_{trough} and C_{EOI} values will be plotted against actual time points on a linear scale, for all subjects by group (cohort and dose level). Arithmetic mean C_{trough} ±SD and arithmetic mean C_{EOI} ±SD will also be plotted by group (cohort and dose level), on a linear scale.

16.4.3 Population Pharmacokinetic Analysis

Sampled PK profiles from study EMR100070-001 will be analyzed jointly with data from other studies by non-linear mixed effect approach, in order to describe the PK concentration time profile followed by multiple dose infusion of avelumab, to identify covariates explaining (part of) the between patient PK variability and to estimate the residual PK inter-individual variability. The PK analysis set will be used. More details will be given in a separate Data Analysis Plan for Population Pharmacokinetic Analysis. The results will be reported separately.

16.4.4 Relation of Pharmacokinetics to Efficacy in Urothelial Carcinoma Cohorts

The exposure-response analysis will include subjects with urothelial carcinoma from the Full Analysis Set (secondary cohort of urothelial carcinoma and efficacy expansion cohort of urothelial carcinoma). The analyses will be based on a data cut-off for the urothelial carcinoma efficacy expansion cohort, i.e. 6 months after start of treatment of the 109th subject.

The objectives of this exploratory population exposure response analysis are: To assess the relationships between objective response (OR), overall survival (OS), and progression-free survival (PFS) with avelumab exposure (e.g., single dose trough concentration (C_{trough_sd}), steady-state C_{trough_SS},), and steady-state AUC_ss, and single dose $AUC_{\_sd}$) in the presence of other relevant covariates.

For OR, the patients will be classified as responder or non-responder based on best overall response (BOR) according to RECIST 1.1 following IERC assessment (responder: CR and PR; non-responder: stable disease [SD], progressive disease [PD], non-evaluable [NE]). OS and PFS will be reported in months.

Population PK parameters, as determined by the final population modeling analysis will be used to derive individual exposure metrics including, but are not limited to C_{trough_sd}, C_{trough_SS}, AUC_ss and $AUC_{\_sd}$. In addition, observed concentration values (i.e. observed single dose C_{trough}, observed stead-state C_{trough}) may be used to represent exposures.

Multivariable regression models will be employed with a stepwise variable selection process. The covariates will be inserted subsequently to the regression model starting with the covariate linked to the smallest p value (corresponding to the Wald chi-square statistic, given the covariates degrees-of-freedom) not yet included to the model if the corresponding p value is below or equal the inclusion threshold of 15%. After each forward step (i.e. inclusion of a new covariate) the new larger regression model passes a backward step, in which subsequently all covariates are excluded (starting with the covariate linked to the highest p value) if their Wald chi-square p-values lie above the exclusion threshold of 40%. The stepwise procedure stops when no further effect can be added or the previously added effect in the forward step was removed in the next backward step.

The interpretation and final inclusion of the covariate(s) in the final model will depend on biological plausibility, dataset attributes and clinical significance of the covariate in question.

Covariates to be tested consider potential prognostic factors including the following in Table 8:

Table 8 Table of potential covariates for regression models

Category	Covariates
Demographics	race, sex, age, baseline body weight
Prior Treatments	Prior radiotherapy (Y/N), Number of prior anti-cancer drug therapies, Prior adjuvant therapy
Safety laboratory information at baseline	Albumin, alkaline phosphatase, bilirubin, alanine aminotransferase, aspartate aminotransferase, platelet count, lactate dehydrogenase at baseline, eGFR*, hemoglobin
Disease-related laboratory values	PD-L1 status
Disease Status/Treatment	Baseline ECOG Status (0/1/2) Concomitant medication of Corticosteroids for systemic use (Y/N) Metastatic disease** at enrolment (study entry) – liver metastasis, other metastasis, no metastasis Number of non-target lesions at baseline Tumor burden at baseline (mm) Tumor sub-site (Lower tract / Upper tract) Eligibility of platinum-based therapy *** (Y/N)

*eGFR is calculated as follows:

eGFR=32788*CREAT**(-1.154)*AGE**(-0.203) + SEX*0.742+MRACE*1.210

where CREAT is serum creatinine (µmol/L), AGE is age (years), Sex is a flag for sex (0 for males, 1 for females) and MRACE is a flag for race (1 for Black or African American, 0 otherwise). CREAT(mg/dL)*88.4=CREAT(µmol/L)

** as evaluated by local or central tumor assessments

*** Subjects are considered eligible for platinum-based therapy in case they receive cisplatin as a prior anti-cancer treatment.

Missing time-invariant categorical covariates will be set to be a separate category distinct from other categories. If the percentage of missing values is below 10%, the missing values should be imputed based on modeler's decision to balance the distribution of the categorical covariate. Missing time-invariant continuous covariates will be to the population median or excluded from the analysis if the percentage of missing values is equal or larger than 25%.

The exposure-response relationship for BOR will be performed using R or NONMEM 7.3. Exposure-response calculations regarding PFS and OS will be performed using the software SAS 9.3 or R.

16.4.4.1 Objective Response: Logistic Regression Model

Multivariate logistic regression of the following form will be used to evaluate potential relationships between OR and exposure:

$$logit(P) = log\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 \cdot C + \beta_2 \cdot X_2 + \beta_3 \cdot X_3 + \cdots + \beta_n \cdot X_n \quad (1)$$

Where P represents probability of being responder, β_0 hypothetically represents the odds of the event occurring without any exposure. β_1 represents a linear effect of the exposure. C represents exposure metric. β_j represents the estimate of the effect of an additional covariate where j = 2, . . ., n and for a total of n− 1 covariates (X_n).

16.4.4.2 Objective Response: Nonlinear logistic regression

The classical logistic model assumes linear relationship of the logit with exposure, meaning that probability of response can asymptotically reach 100% for an infinite exposure, which is unrealistic. Therefore, an exploratory nonlinear logistic regression will be performed using an Emax function (Equation 2) instead of the linear function as shown in Equation 1.

$$logit(P) = log\left(\frac{P}{1-P}\right) = \beta_0 + \frac{Emax \times C}{C+EC50} \quad (2)$$

Here, *P* is the probability of response, β_0 hypothetically represents the odds of the event occurring without any exposure, *Emax* is the maximum drug effect, *EC50* is the exposure level where 50% of maximum drug effect is reached. *C* represents exposure metric

16.4.4.3 Overall survival and progression free survival: Cox Proportional Hazards Model

Progression free survival (PFS) and overall survival (OS) were captured as events and non-events and therefore the drug exposure response model will be developed using time-to-event analysis. A Cox proportional hazards model will be used to assess the relationship for OS and PFS versus avelumab exposure as well as to explore prognostic factors (covariates) for each endpoint. The exposure metrics will include, but are not limited to, C_{trough_sd}, C_{trough_SS} and AUC_ss, $AUC_{\_sd}$ Validity of the assumption of proportionality of hazards for the covariates could be verified by residual analysis (plots). Maximum likelihood estimates of the parameters will include hazard ratios and 95% confidence intervals.

The Cox proportional hazards model is the most commonly used multivariate approach for analyzing survival time data in medical research. It is a survival analysis regression model, which describes the relation between the event incidence, as expressed by the hazard function and a set of covariates. Mathematically, the Cox model is written as:

$$h(t) = h_0(t) * \exp(\beta_1 \cdot C + \beta_2 \cdot X_2 + \beta_3 \cdot X_3 + \cdots + \beta_n \cdot X_n) \quad (3)$$

Where the hazard function $h(t)$ is dependent on (or determined by) a set of n covariates (C, $x_2, \ldots, x_n$) including exposure metric (C), whose impact is measured by the size of the respective

regression coefficients ($\beta_1, \ldots., \beta_n$). An appealing feature of the Cox model is that the baseline hazard function $h_0(t)$ is estimated non-parametrically, and so unlike most other statistical models, the survival times are not assumed to follow a particular statistical distribution.

The Cox model is essentially a multiple linear regression of the logarithm of the hazard on the covariables, with the baseline hazard being an 'intercept' term that varies with time. The covariates then act multiplicatively on the hazard at any point in time, and this provides us with the key assumption of the proportional hazards model: the hazard of the event in any group is a constant multiple of the hazard in any other. In case of ties, the discrete logistic likelihood

$$L(\beta_1, \ldots, \beta_n) = \prod_{i=1}^{k} \frac{\exp(\beta_1 \sum_{j \in D_i} C_j + \cdots + \beta_n \sum_{j \in D_i} X_{nj})}{\sum_{q \in Q_i} \exp(\beta_1 \sum_{l=1}^{d_i} C_{q_l} + \cdots + \beta_n \sum_{j \in D_i} X_{nq_l})} \quad (4)$$

is maximized for the estimation of regression coefficients. The term D_i in equation (4) denotes the set of subjects failing at time point i for distinct event time points $i = 1, \ldots, k$ (with d_i as number of subjects failing at time point i). The set Q_i includes all d_i-tuples $(q_1, \ldots, q_{d_i})$ of subjects, which are under risk at time point i.

Correlation plots will be produced to explore co-linearity of relevant covariates to exposure and correlation coefficients will be reported.

16.5 Immunogenicity

ADA (referred to as HAHA in CRF and SDTM) was assessed before the study treatment start, and on Days 15, 29, 43, 57, 71, 85, 127, and 169 prior to the start of infusion, and at the end of treatment visit. If the sample is positive for ADA, it will be re-analyzed to determine the titer and nAb. The ADA results will be derived based on the algorithm in Table 9. Subjects will be characterized into different ADA categories based on the criteria in Table 10.

Table 9 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.

Table 10 Subjects Characterized based on ADA Results

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

Samples with a reportable ADA titer will also be tested in the neutralizing antibody (nAb) assay. NAb results are positive or negative in a single assay and only derived when not performed because ADA was negative (see Table 11). Subjects will be characterized into different nAb categories based on the criteria in Table 12. For nAb, treatment-boosted is not applicable since no titer result is available.

Table 1 Algorithm for the Derivation of nAb Results

ADA Confirmatory Result	nAb Result	Derived nAb Result
Negative	NA	Negative
NR	NA	NR
NA	NA	Negative
Positive	NR	NR
Positive	Positive	Positive
Positive	Negative	Negative

ADA = antidrug antibody, NA = not applicable, nAb = neutralizing antibody, NR = no result.

Table 2 Subjects Characterized based on nAb Results

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

ADA = antidrug antibody, nAb = neutralizing antibody, NR = no result.

The frequency and percentage of each ADA and nAb category will be presented in tables by cohort. Listings of ADA and nAb results from ADA ever-positive subjects will be prepared.

For the ADA ever-positive subjects, a listing will be prepared by cohort with subject ID, start and stop of treatment, date of onset, time to onset (weeks since treatment start) and last date of ADA positive results, as well as date of onset, time to onset and last date of nAb positive results, confirmed BOR and confirmed BOR date, DOR, PFS time or censoring time and reason for censoring, and OS time or censoring time and reason for censoring. Confirmed BOR will be based upon IERC results for the efficacy expansion cohorts and secondary UC cohort and investigator results otherwise.

For the ADA ever-positive subjects by cohort, the percent change from baseline in target lesions as well as the first occurrence of a new lesion and subject off treatment will be displayed against time point (weeks) in a line plot. Additional symbols will indicate the first and last ADA positive result and, if applicable, the first and last nAb positive result.

16.5.1 Subset Analysis by IgE

Anti-avelumab IgE will be assessed in a subset of samples covering ADA results and IgE positive and negative subjects from SAF analysis set. This analysis will only be performed for CSRs with integration across all dose expansion cohorts.

A listing of IgE result, ADA result, pre-dose avelumab concentration (mg/L), and IRR status will be prepared for each sample tested. Avelumab concentrations, IgE assessments, ADA results, and

IRR status will be matched by visit/ dosing date for each subject. IRR status is derived as positive or negative if subjects have at least one or no treatment emergent IRR started on the dosing date or the following day, respectively. IgE status is derived as negative if the non-missing value is < 0.1, or positive if the value is ≥ 0.1. A descriptive (count and percentage) summary table will be presented for IgE status by IRR status and ADA result, the denominator will be the number of samples with valid ADA and IgE results from subjects dosed at the corresponding visit.

Two scatterplots will be created to graphically display the data, one for IgE results vs. ADA titer, and one for IgE results vs. avelumab concentration (µg/mL). ADA results of negative or positive-TNR, IgE value or avelumab concentration below LLOQ will be set to 0 in the plot.

17 Safety Evaluation

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

All safety analyses will be performed using the SAF, unless otherwise specified.

All safety parameters will be summarized by cohort(s) and/or pooled across cohorts, depending upon the purpose of the reporting. Primary and final analyses for a cohort will include safety data summarized by that cohort unless integrated safety analysis is performed on the same data cut.

17.1 Adverse Events

The severity of AEs will be graded using the NCI-CTCAE, version 4.0 except where CTCAE grades are missing. No imputation of missing grades will be performed. AEs will be coded according to latest available version of MedDRA. AEs of special interest include IRRs and immune related adverse events (irAEs), which are detailed in Section 17.1.4 and 17.1.5, respectively.

- **Treatment Emergent Adverse Events:** TEAEs are those events with onset dates occurring during the on-treatment period for the first time, or if the worsening of an event is during the on-treatment period as defined in Section 11.
- **Related Adverse Events:** adverse events with relationship to study treatment (relationship with study treatment = related) reported by the investigator and those of unknown relationship (i.e. no answer to the question “Relationship with study treatment”).
- **Serious Adverse Events (SAE):** serious adverse events (as recorded on the AE eCRF page, serious adverse event = yes).
- **Adverse Events Leading to Treatment Discontinuation:** adverse events leading to permanent discontinuation of study treatment (as recorded on the AE eCRF page, action taken with study treatment = drug withdrawn).
- **Adverse Events Leading to Death:** adverse event leading to death (as recorded on the AE eCRF page, outcome = fatal or toxicity grade = 5).

- **Original Definition Immune Related Adverse Events (irAE):** irAEs are identified according to a pre-specified search list of MedDRA PTs, documented in a version-controlled repository maintained by the Sponsor and finalized for analysis prior to database lock. The original definition irAEs are utilized when the process for irAE medical review is either not followed or has not yet been completed for the delivery.
- **Updated Definition Immune Related Adverse Events:** irAEs according to case definition classified by medical review. Details are included in Table 15 in Appendix II.
- **Updated Definition Infusion Related Reaction:** IRRs are identified based on a list of MedDRA PTs. The detailed criteria of the timing relationship to infusion are specified in Table 16 in Appendix II.

AEs will be summarized using th**e** MedDRA PT as event category and MedDRA primary SOC as summary category. All AE tables will be restricted to TEAEs unless otherwise specified.

Each subject will be counted only once within each PT or SOC and recording period. If a subject experiences more than one AE within a PT or SOC for the same recording period, only the AE with the strongest relationship or the worst severity, as appropriate, will be included in the summaries of relationship and severity. AEs with missing classifications regarding relationship to study treatment and start date greater or equal to start of study treatment will be considered as related to study treatment. AEs with missing toxicity grade will be counted into 'any grade' in the summarization by toxicity grade.

The AE tables will include the number and percentage of subjects with at least one TEAE, by MedDRA primary SOC (sorted by decreasing SOC frequencies within the overall column) and by PT (sorted by decreasing PT frequencies in overall column within SOC), unless otherwise stated.

17.1.1 All Adverse Events

All AEs will be tabulated in the following tables.

- The overall summary of AEs table will include the following summaries:
  - TEAEs
  - TEAEs, grade ≥ 3
  - Related TEAEs
  - Related TEAEs, grade ≥ 3
  - TEAEs leading to permanent treatment discontinuation
  - Related TEAEs leading to permanent treatment discontinuation
  - Serious TEAEs
  - Related serious TEAEs

- TEAEs leading to death
- Related TEAEs leading to death
- Treatment emergent irAEs
- Related treatment emergent irAEs
- Treatment emergent IRRs
- Related treatment emergent IRRs

- Incidence of TEAEs by SOC, PT, and worst grade
- Incidence of related TEAEs by SOC, PT, and worst grade
- Incidence of TEAEs by SOC and PT: displaying in separate columns the All TEAEs / Related TEAEs / Grade ≥3 TEAEs / Related Grade ≥3 TEAEs
- Incidence of TEAEs leading to death by SOC and PT
- Incidence of related TEAEs leading to death by SOC and PT
- Incidence of non-serious TEAEs by SOC and PT

Listing of AEs including all relevant information such as AE SOC/PT, start/stop date, duration of AE, toxicity grade, relationship to the study treatment, action taken with study treatment, and outcome etc., will be provided. A separate listing of AEs started or worsened after on-treatment period will also be provided.

17.1.2 Adverse Events Leading to Treatment Discontinuation

Following summary tables will be produced:

- Incidence of TEAEs leading to permanent treatment discontinuation by SOC and PT
- Incidence of related TEAEs leading to permanent treatment discontinuation by SOC and PT

The listing of AEs leading to permanent treatment discontinuation will also be provided with the relevant information such as AE SOC/PT, start/stop date, toxicity grade, and outcome etc.

17.1.3 Serious Adverse Events

Following summary tables will be produced:

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

The listings of SAEs will also be provided with the relevant information such as AE SOC/PT, start/stop date, toxicity grade, relationship to the study treatment, action taken with study treatment, and outcome etc.

17.1.4 Infusion Related Reaction

IRRs will be summarized by the follow variables:

- Number of subjects with at least one event by the worst toxicity grade (grade 1/ grade 2/ grade 3/ grade 4/ grade 5/ missing grade)
- Number of subjects with IRR leading to permanent treatment discontinuation
- Time related to first onset (infusion 1/ infusion 2/ infusion 3/ infusion 4 or later). The events should be assigned to the actual drug infusions that the subject received, not to the planned dates. An IRR is assigned to a drug infusion if its onset is at the same date (but not before dosing) or the following day of drug infusion.
- Number of subjects with at least one event by the worst toxicity grade that occurred in the presence of premedication (grade 1/ grade 2/ grade 3/ grade 4/ grade 5/ missing grade). The denominator will be the number of subjects with at least one dose administered in the presence of pre-medication. The maximum toxicity will be derived among those IRRs that occurred in the presence of premedication.
- Number of subjects with at least one event by the worst toxicity grade that occurred in the absence of premedication (grade 1/ grade 2/ grade 3/ grade 4/ grade 5/ missing grade). The denominator will be the number of subjects with at least one dose administered in the absence of pre-medication. The maximum toxicity will be derived among those IRRs that occurred in the absence of premedication.

The listing of IRRs will be provided with the relevant information such as AE SOC/PT, start/stop date, toxicity grade, relationship to the study treatment, action taken with study treatment, outcome, premedication, and study medication batch ID etc.

17.1.5 Immune Related Adverse Event

Treatment emergent irAEs will be summarized using the following variables:

- The overall summary of treatment emergent irAEs will include the following categories:
  - irAEs
  - Related irAEs
  - irAEs, grade ≥3
  - Related irAEs, grade ≥3
  - irAEs leading to permanent treatment discontinuation
  - Related irAEs leading to permanent treatment discontinuation
  - Serious irAEs
  - Related serious irAEs

- irAEs leading to death
- Related irAEs leading to death

- Incidence of irAEs by SOC and PT: displaying in separate columns the All irAEs / Related irAEs / Grade ≥3 irAEs / Related Grade ≥3 irAEs
- Incidence of irAEs by SOC and PT and worst grade
- Incidence of related irAEs by SOC and PT and worst grade
- Incidence of irAEs leading to permanent treatment discontinuation by SOC and PT
- Incidence of related irAEs leading to permanent treatment discontinuation by SOC and PT
- Incidence of irAEs leading to death
- Incidence of related irAEs leading to death

A listing containing all the PTs used to identify irAEs and a listing containing all the irAEs in the study will be generated as well. A separate listing of irAEs started or worsened after on-treatment period will also be provided

17.1.6 Subgroup Analysis of Adverse Events

A listing of immunogenicity data and TEAEs will be provided containing subject ID, cohort (if more than one included in CSR), age, gender, study treatment start and stop date, all dates with positive ADA result, all dates with positive nAb results, preferred term of TEAE, TEAE start date, stop date, CTCAE toxicity grade and flags for immune-related adverse event or infusion related reaction or reason for permanent treatment discontinuation, as well as ADA status group.

Only subjects which are pre-existing positive, transient treatment-emergent positive, or persistent treatment-emergent positive will be listed.

17.2 Deaths

All deaths will be tabulated and listed for the SAF subjects. The death table will include the following information:

- Number of subjects who died
  - Primary reason for death
    - Disease progression
    - Adverse event related to study treatment
    - Adverse event not related to study treatment
    - Other
    - Unknown
- Number of subjects who died within 30 days of the last study treatment administration

- Primary reason for death
    - Disease progression
    - Adverse event related to study treatment
    - Adverse event not related to study treatment
    - Other
    - Unknown

- Number of subjects who died within 60 days of the first study treatment administration
    - Primary reason for death
        - Disease progression
        - Adverse event related to study treatment
        - Adverse event not related to study treatment
        - Other
        - Unknown

The listing of deaths will be provided with all the relevant information such as death date and reason for death. The death data will be ascertained from the dedicated Report of Death eCRF form.

17.3 Clinical Laboratory Evaluation

Laboratory abnormalities are classified according to NCI-CTCAE toxicity grading version 4.03 or based on normal ranges collected from laboratories. The toxicity grading is only related to the lab values itself and does not respect the non-numeric information as described in the CTC grading definition. CTCAE gradable parameters and associated toxicities are listed in Table 13.

Table 13 CTCAE Gradable Parameters and Associated Toxicities

Panel	Parameter (test)	Low direction toxicity	High direction toxicity
Hematology	Hemoglobin (HB)	Anemia	hemoglobin increased
	Leukocytes	white blood cell decreased	
	Lymphocytes	lymphocyte count decreased	lymphocyte count increased
	Neutrophils/ absolute neutrophils count (ANC)	neutrophil count decreased	
	Platelet count (PLT)	platelet count decreased	
Chemistry	Albumin	Hypoalbuminemia	
	Alkaline phosphatase (ALP)		alkaline phosphatase increased
	Alanine aminotransferase (ALT)		ALT increased
	Amylase		serum amylase increased
	Aspartate aminotransferase (AST)		AST increased

	Bilirubin (total)		blood bilirubin increased
	Cholesterol		cholesterol high
	Creatinine		creatinine increased
	Creatinine clearance	chronic kidney disease	
	Creatine kinase (CPK)		CPK increased
	Potassium	Hypokalemia	hyperkalemia
	Sodium	Hyponatremia	hypernatremia
	Magnesium	Hypomagnesemia	hypermagnesemia
	Calcium	Hypocalcemia	hypercalcemia
	Glucose	Hypoglycemia	hyperglycemia
	Gamma glutamyl transferase (GGT)		GGT increased
	Lipase		lipase increased
	Phosphates	Hypophosphatemia	
	Triglycerides		hypertriglyceridemia
Coagulation	Prothrombin time INR		INR increased
	Activated Partial thromboplastin time (aPTT)		aPTT prolonged

Grade 1 and 2 hyperglycemia are based on fasting glucose, they will not be graded for this study because blood samples are taken from non-fasted subjects.

For calcium, CTCAE grading is based on corrected calcium and ionized calcium (CALCIO), if available. Corrected calcium is calculated from albumin and calcium as follows

Corrected calcium (mg/dL) = Calcium (mg/dL) – 0.8 [Albumin (g/dL) - 4], or

Corrected calcium (mmol/L) = Calcium (mmol/L) + 0.02 (40 - Serum albumin [g/L]).

Chronic kidney disease will be graded based on estimated creatinine clearance rate (eCcr, ml/min), which is derived using Cockcroft-Gault formula:

$$eCcr = \frac{(140 - Age) \times Weight(kg) \times Constant}{Serum\ Creatinine\ (\mu mol/L)}$$

where the constant is 1.23 for men and 1.04 for women.

The corrected eCcr is derived as:

Corrected eCcr (ml/min/1.73m^2) = eCcr × 1.73 / BSA

where BSA (body surface area, m^2) = ([baseline height (cm) × weight (kg)] / 3600)$^{0.5}$.

The following are non-CTCAE gradable parameters collected in this study, their abnormalities are assessed as low, high, normal based on the comparison of observed values with normal ranges.

- Hematology: hematocrit, red blood cell (RBC), reticulocytes, mean corpuscular hemoglobin (MCH), mean corpuscular volume (MCV), mean corpuscular hemoglobin concentration (MCHC).
- Serum Chemistry: chlorine, C-reactive protein, lactate dehydrogenase (LDH), total protein, total urea, uric acid.
- Hormone: adrenocorticotropic hormone (ACTH), anti-nuclear antibody (ANA), anti-neutrophil cytoplasmic antibody (ANCA), rheumatoid factor (RF), free thyroxine (Free T4), thyroid-stimulating hormone (TSH).

Laboratory abnormalities will be summarized using the worst grade during the on-treatment period. For these parameters which are graded with two toxicities such as potassium (hypokalemia/ hyperkalemia), the toxicities will be summarized separately. Low direction toxicity (eg. 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 change and percent change from baseline will be derived for parameters with numeric results.

17.3.1 Hematology and Clinical Chemistry Parameters

CTCAE gradable parameters

The laboratory toxicities will be tabulated by the worst on-treatment CTCAE grade or the shift of CTCAE grade from baseline to worst grade during on-treatment period using descriptive statistics (count and percentage). The highest CTCAE grade during the on-treatment period is considered as the worst grade for the summary.

- The worst grade during the on-treatment period will be summarized considering only patients with post baseline laboratory samples: Laboratory tests by NCI-CTCAE grade (0, 1, 2, 3, 4, and missing).
- The shift table will summarize baseline CTCAE grade vs. the worst on-treatment CTCAE grade (grade = 0, 1, 2, 3, 4, missing).

Non-CTCAE gradable parameters

Hematology, chemistry, and hormone evaluations which can't be graded per CTCAE will be summarized as:

- Shift from baseline value (low, normal, high) to above normal during on-treatment period
- Shift from baseline value (low, normal, high) to below normal during on-treatment period

Quantitative data will also be examined for trends using descriptive statistics (n, missing, mean, SD, median, Q1, Q3, minimum, and maximum) of actual values, absolute changes and percent changes from baseline to each visit over time. This summarization will apply to hematology and chemistry parameters with numeric results assessed at baseline, post-baseline, discontinuation, end of treatment, and/or safety follow-up visits based on visit windows specified in Section 11.

The listings (hematology, chemistry, and hormone) will include all the laboratory parameters as available in the database with the relevant information such as visit, assessment date, parameter, value, normal ranges etc. Listings will be sorted by subject identifier, group variable, parameter, assessment date or visit.

Liver function parameters

ALT, AST, ALP, and total bilirubin are used individually or together to assess possible drug induced liver toxicity. The ratios of test result over upper limit of normal (ULN) for individual test or combined tests will be calculated and classified for these parameters during the on-treatment period.

Summary of liver function tests will include the following categories. The number and percentage of subjects with each of the following during the on-treatment period will be summarized by cohort, if applicable:

- AST ≥3*ULN / ≥5*ULN / ≥10*ULN / ≥20*ULN.
- ALT ≥3*ULN / ≥5*ULN / ≥10*ULN / ≥20*ULN.
- (ALT or AST) ≥ 3*ULN/ ≥ 5*ULN/ ≥ 10*ULN/ ≥ 20*ULN
- Total bilirubin ≥2*ULN
- Concurrent ALT ≥ 3*ULN and TBILI ≥ 2*ULN
- Concurrent AST ≥ 3*ULN and TBILI ≥ 2*ULN
- (ALT or AST) ≥3*ULN concurrently with total bilirubin ≥2*ULN.
- (ALT or AST) ≥3*ULN concurrently with total bilirubin ≥2*ULN and ALP >2*ULN.
- (ALT or AST) ≥3*ULN concurrently with total bilirubin ≥2*ULN and (ALP ≤2*ULN or missing).

Concurrent measurements are those occurring on the same date.

Categories will be cumulative, i.e a subject with an elevation of AST ≥10*ULN will also appear in the categories ≥5*ULN and ≥3*ULN. Liver function elevation and possible Hy's Law cases will be summarized using frequency and percentage by cohort.

An evaluation of Drug-Induced Serious Hepatotoxicity (eDISH) plot will also be created to graphically display

- peak serum ALT(/ULN) vs. peak total bilirubin (/ULN) including reference lines at ALT ≥3×ULN and total bilirubin ≥2×ULN.
- peak serum AST(/ULN) vs. peak total bilirubin (/ULN).

Listing of subjects with ALT or AST ≥3*ULN or total bilirubin ≥2*ULN will include variables subject identifier, visit, date of collection, study day, parameter (ALT, AST, ALP, total bilirubin), result, unit, result/ULN, CTCAE grade.

17.3.2 Other Laboratory Parameters

All other parameters collected on the eCRF will be listed in dedicated listings presenting all corresponding collected data-fields on the eCRF.

- Coagulation: aPTT, prothrombin time INR
- Urinalysis: all urinalysis parameters
- Other parameters: such as immunology, soluble factor
- Pregnancy test
- Serology

17.4 Vital Signs

Summary of vital signs will be based on the SAF. The potentially clinically significant changes from baseline as below in vital signs will be derived and summarized with subject incidence and percentage during the on-treatment period:

- ≥ 10% weight increase
- ≥ 10% weight decreases
- ≤ 95 mmHg and decrease from baseline ≥ 20 mmHg in systolic blood pressure
- ≥ 160 mmHg and increased from baseline ≥ 20 mmHg in systolic blood pressure
- ≤ 45 mmHg and decrease from baseline ≥ 10 mmHg in diastolic blood pressure
- ≥ 110 mmHg and increased from baseline ≥ 10 mmHg in diastolic blood pressure
- ≤ 50 bpm and decrease from baseline ≥ 20 bpm in pulse rate
- ≥ 120 bpm and increase from baseline ≥ 20 bpm in pulse rate

Quantitative data will also be examined for trends using descriptive statistics (n, missing, mean, SD, median, Q1, Q3, minimum, and maximum) of actual values, absolute changes and percent changes from baseline to each visit based on visit windows specified in Section 11. This summarization will apply to weight, blood pressure, respiratory rate, pulse, and temperature assessed at baseline, post-baseline, discontinuation, end of treatment, and safety follow-up visits.

Data listing of all vital signs will be provided with all relevant information such as visit, assessment date, parameter, and results.

17.5 Other Safety or Tolerability Evaluations

The incidence and percentage of subjects with potentially clinically significant abnormalities (PCSA) for 12-lead ECG parameters will be summarized for scheduled visits during the on-treatment period based on the SAF. The PCSA criteria are provided in the Table 14.

Table 14 Potentially Clinically Significant Abnormalities Criteria for ECG

Test	Potentially Clinically Significant Abnormalities (PCSA) Criteria
Heart Rate (HR)	≤ 50 bpm and decrease from baseline ≥ 20 bpm ≥ 120 bpm and increased from baseline ≥ 20 bpm
PR Interval	≥ 220 ms and increase from baseline ≥ 20 ms
QRS	≥ 120 ms
QTcF, QTcB absolute	interval >450 - ≤ 480 ms interval >480 - ≤ 500 ms interval >500 ms
QTcF, QTcB change from baseline	Increase from baseline > 30 - ≤ 60 ms Increase from baseline > 60 ms

QT interval will be corrected based on Fridericia's formula (QTcF= QT/ $\sqrt[3]{RR}$) and RR=60/HR, Bazett's formula (QTcB= QT/ √RR) and RR=60/HR, and lineal regression, if possible. The correction be linear regression (QTcP) will be performed as:

- Fit a linear regression model QT = a + b * RR to baseline QT and RR data of the SAF subjects.
- Use the estimated slope, $\hat{b}$, to correct QT
- Corrected QT will be computed as QTcP = QT + $\hat{b}$ * (1- RR)

Baseline QTcF/QTcB/QTcP will be derived from the visit that QT and HR are flagged as baseline. If there are multiple assessments at the same visit and time point, the average will be calculated for each parameter and used for the analysis.

Quantitative data will also be examined for trends using descriptive statistics (n, missing, mean, SD, median, Q1, Q3, minimum, and maximum) of actual values (with 95% CI of the mean), absolute changes from baseline (with 90% CI of the mean) to each visit based on visit windows specified in Section 11. This summarization will apply to heart rate, QRS interval, QT interval, PR interval, QTcB, and QTcF assessed at baseline, post-baseline, discontinuation, end of treatment, and/or safety follow-up visits.

Listings of 12-lead ECGs will be provided with all relevant information such as visit, date/time of assessment, parameter, and results.

The ECOG shift from baseline to highest on-treatment score will be summarized by cohort based on SAF analysis set. Missing category will be included and the number of subjects in each cohort will be used as the denominator.

ECOG performance status will also be presented in a data listing.

Data of subject status and survival follow-up will be provided in a listing.

18 Reporting Conventions

- Reporting will require placement of decimals, and this will depend on the raw data collected. Generally, mean and median should be displayed one more decimal place than the raw data and standard deviation should be displayed two more decimal place than the raw data. Percentages will be reported as one decimal place.
- The rounding will be performed to closest integer / first decimal using the common mid-point between the two consecutive values. E.g. 5.1 to 5.4 will be rounded to an integer of 5, and 5.5 to 5.9 will be rounded to an integer of 6.
- The following conversion factors will be used to convert days to months or years, where applicable: 1 month = 30.4375 days and 1 year = 365.25 days.
- Data listings will be sorted by dose level, subject identifier, visit or date (if applicable), or parameters, as appropriate.

18.1 Date of Last Contact

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

- All patient assessment dates (blood draws (laboratory, PK), vital signs, performance status, ECG, tumor assessments)
- Start and end dates of anti-cancer therapies administered after study treatment discontinuation.
- AE start and end dates
- Last date of contact collected on the 'Subject Status/Survival Follow-up' eCRF (do not use date of survival follow-up assessment unless status is 'alive')
- Study treatment start and end dates
- 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 patient will be used in the derivation. Dates associated with a technical operation unrelated to patient 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.

18.2 Incomplete Dates

Missing or partial start dates for adverse events will be imputed as following:

- When the start Date of the AE is missing (but Month & Year is available), then the AE date will be imputed to the "1st Date of the reported Month" (e.g. if reported date is --/JAN/09, imputed date will be 01/JAN/09). If the reported AE Month = the Month of the First Dosing date, then the AE date will be imputed to the "1st Dosing date" (e.g. if AE reported date is --/JAN/09, and the First doing date is 13/JAN/09, then the AE imputed date will be 13/JAN/09).

- When the Date & the Month of the AE is missing (but Year is available), then the AE date will be imputed to the "1st Date of the 1st Month of the reported Year" (e.g. if reported date is --/--/09, imputed date will be 01/JAN/09). If the reported AE Year = the Year of the First Dosing date, then the AE date & month will be imputed to the "1st Dosing date" (e.g. if AE reported date is --/--/09, and the First dosing date is 13/APR/09, then the AE imputed date will be 13/APR/09).
- When the date is completely missing, no imputation will be performed and the AE will be considered as treatment emergent, unless there is rational to clarify otherwise, eg. AE stop date is prior to the first dose date.

Missing or partial dates for concomitant medications will be imputed as following:

- If the start day of medication is missing, it will be imputed to the first day of the month; if the stop day is missing, it will be imputed to the last day of the month. If both day and month are missing, the start date will be imputed as the first day of the year and stop date will be imputed as the last day of the year. If the start or stop date is completely missing, no imputation will be performed and the determination of pre-medication or post-medication will be based on non-missing stop or start date, respectively; otherwise, the medication will be considered as concomitant.

Missing or partial dates for disease history (initial diagnosis date, first occurrence of metastatic or locally advanced disease, date of last progression of disease) will be imputed as following:

- If the day is missing, it will be imputed to the 15th day of the month; if both day and month are missing and the year is prior to the year of the first study treatment, the month and day will be imputed as July 1st; if both day and month are missing and the year is same as the year of the first study treatment, the month and day will be imputed as January 1st. If the date is completely missing, no imputation will be performed.

Partial dates for prior anti-cancer drug therapies will be imputed as following:

- If the day is missing, it will be imputed to the first day of the month; if both day and month are missing, no imputation will be performed.

Missing or partial death dates will be imputed based on the last contact date:

- If the date is missing it will be imputed as day after date of last contact from the CRF survival page
- If the day or month is missing, death will be imputed to the maximum of the full (non-imputed) day after the date of last contact and the following:
  - Missing day: 1st day of the month and year of death
  - Missing day and month: January 1st of the year of death

If the day is missing from the date of last contact it will be imputed to 1st day of the month and year of last contact only if derived from the survival page.

No other dates will be imputed.

19 References

1. Eisenhauer EA. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). Eur J Cancer 2009;45:228-47.
2. Wolchok et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. Clin Cancer Res 2009;15(23):7412-20.
3. Jennison C, Turnbull BW. In: Group Sequential Methods with Applications to Clinical Trials. Chapman & Hall/CRC, Boca Raton, 2000.
4. Brookmeyer R, Crowley J. A Confidence Interval for the Median Survival Time. Biometrics 1982;38:29–41.
5. Kalbfleisch JD, Prentice RL. The Statistical Analysis of Failure Time Data, New York: John Wiley & Sons 1980.

20 Appendices

Appendix I Important Protocol Deviations

	Category of Protocol Deviation	Description of Protocol Deviation	Deviation Code	Protocol Section	Variable [dataset]	Proposed check / comment
Inclusion criteria: For the subject to be eligible for inclusion, each criterion must be checked 'YES':						
Criterion1: Signed written informed consent.	Eligibility and Entry Criteria	Subject did not meet inclusion criterion 1 (signed inform consent form).	PDEV01	Section 5.3.1	DM. RFICDTC	list if DM. RFICDTC is missing or if DM.RFICDTC<Earliest date of SV.SVSTDTC. Medical Review Required.
Criterion2: Male or female subjects aged >=18 years.	Eligibility and Entry Criteria	Subject did not meet inclusion criterion 2 (age >= 18 years).	PDEV02	Section 5.3.1	DM.AGE	List if DM.AGE<18.

	Category of Protocol Deviation	Description of Protocol Deviation	Deviation Code	Protocol Section	Variable [dataset]	Proposed check / comment
Criterion3: Histologically or cytologically proven metastatic or locally advanced solid tumor or related prior anti-cancer therapy.	Eligibility and Entry Criteria	Subject did not meet inclusion criterion 3 (histologically or cytologically proven metastatic or locally advanced solid tumor or prior anti-cancer therapy for metastatic or locally advanced disease).	PDEV03	Section 5.3.1	XX.XXDTC where XXTESTCD='TBPALL'; SUPPXX.QNAM(FRESHDTC, FRESHTB), QVAL; CM.CMTRT, CMINDC	Medical review required or programming. List if subject did not meet all of the following requirements (depending on cohort): A. Tumor tissue is available - 1) there is at least one record in XX (XX.XXDTC ne missing where XXTESTCD='TBPALL' and XXCAT='TUMOR BIOPSY OR BLOCK') for all cohorts or at least one record (missing < XX.XXDTC < TRTSDT where XXTESTCD='TBPALL' and XXCAT= 'UNSCHEDULED TUMOR BIOPSY OR BLOCK') for expansion subjects ; or 2) there is at least one records in SUPPXX (SUPPXX.QVAL='Y' where QNAM= 'FRESHTB') for all cohorts; B. At least one prior anti-cancer therapy for NSCLC and ACC subjects - 1) there is at least one record in CM where CMCAT='PRIOR ANTI-CANCER TREATMENT' and CMSCAT='DRUG THERAPY' and CMTRT ne missing and CMINDC=('METASTATIC', 'LOCALLY ADVANCED', 'PALLIATIVE') for subjects where ACTARMCD=(MSB_EXP_NSCLC, MSB_EXP_ACC); C. No prior anti-cancer therapy for first line NSCLC subjects: - 1) there is NO record in CM where CMCAT='PRIOR ANTI-CANCER TREATMENT' and CMSCAT='DRUG THERAPY' and CMTRT ne missing and CMINDC= ('METASTATIC', 'LOCALLY ADVANCED', 'PALLIATIVE') for subjects where ACTARMCD=(MSB_EXP_FLNSCLC).

	Category of Protocol Deviation	Description of Protocol Deviation	Deviation Code	Protocol Section	Variable [dataset]	Proposed check / comment
Criterion5: Disease must be measurable at least 1 dimension except for CRPC or MBC (dose escalation).	Eligibility and Entry Criteria	Subject did not meet inclusion criterion 5 (measurable disease at least one dimension except for CRPC and escalation MBC subjects).	PDEV04	Section 5.3.1	TR.TRORRES, TRTESTCD; SUPPTR.QNAM, QVAL.	list if TR.TRORRES is missing or not assessable where TRTESTCD=' SUMDIAM' and TRCAT= 'RECIST' and SUPPTR.QVAL='SCREENING' where QNAM='ASMNTVIS'. Do not check this for subjects where SUPPDM.QVAL = ' CASTRATE-RESISTANT PROSTATE CANCER' where QNAM='COHORT' or SUPPDM.QVAL=' DOSE ESCALATION AND DLT/MTD COHORTS' where QNAM='COHORT' and MHLOC in ('NIPPLE' 'CENTRAL PORTION OF BREAST' 'UPPER-INNER QUADRANT OF BREAST (UIQ)' 'LOWER-INNER QUADRANT OF BREAST (LIQ)' 'UPPER-OUTER QUADRANT OF BREAST (UOQ)' 'LOWER-OUTER QUADRANT OF BREAST (LOQ)' 'AXILLARY TAIL OF BREAST' 'OVERLAPPING LESION OF BREAST 'BREAST, NOS')
Criterion9: Effective contraception for both male and female subjects if the risk of conception exists.	Eligibility and Entry Criteria	Subject did not meet inclusion criterion 9 (effective contraception).	PDEV05	Section 5.3.1		Medical review required
Exclusion criteria: For the subject to be eligible for inclusion, each criterion must be checked 'NO':						
1. Concurrent treatment with a non-permitted drug.	Eligibility and Entry Criteria	Subject met exclusion criterion 1 (non-permitted drug).	PDEV06	Section 5.3.2		Medical review required

	Category of Protocol Deviation	Description of Protocol Deviation	Deviation Code	Protocol Section	Variable [dataset]	Proposed check / comment
2. Prior therapy with any antibody/drug targeting T cell co-regulatory proteins.	Eligibility and Entry Criteria	Subject met exclusion criterion 2 (prior therapy with any antibody/drug targeting T cell co-regulatory proteins).	PDEV07	Section 5.3.2		Medical review required
3. Concurrent anticancer treatment, major surgery, concurrent systemic therapy with steroids or other immunosuppressive agents, or use of any investigational drug within 28 days before the start of trial treatment.	Eligibility and Entry Criteria	Subject met exclusion criterion 3 (concurrent anti-cancer therapy or surgery, concurrent systemic therapy with steroids or other),	PDEV08	Section 5.3.2		Medical review required
4. Previous malignant disease other than the target malignancy to be investigated in this trial within the last 5 years with the exception of basal or squamous cell carcinoma of the skin or cervical carcinoma in situ	Eligibility and Entry Criteria	Subject met exclusion criterion 4 (previous malignant disease other than the target malignancy to be investigated in this trial within the last 5 years).	PDEV09	Section 5.3.2		Medical review required
12. Pregnancy or lactation period.	Eligibility and Entry Criteria	Subject met exclusion criterion 12 (pregnancy or lactation period).	PDEV10	Section 5.3.2	PREG.PRRLTCD	Medical review required.

	Category of Protocol Deviation	Description of Protocol Deviation	Deviation Code	Protocol Section	Variable [dataset]	Proposed check / comment
Non-permitted concomitant medication during the study	Concomitant Medication Criteria	Subject took prohibited medication during the study.	PDEV11	Section 6.5.2	CMED.CMTERM	Medical review required
Subjects that developed withdrawal criteria whilst on the study but were not withdrawn;	Other Criteria	Subject became pregnant, but continued on the study.	PDEV12	Section 5.5.2	LB.LBORRES, BTESTCD; DS.DSSTDTC, DSSCATE;	Medical review required.
Subjects that developed withdrawal criteria whilst on the study but were not withdrawn;	Other Criteria	Subjects had ECOG>=3, did not resolved to <=2 by day 14 of next cycle, and continued on the study.	PDEV13	Section 5.1.7.2 and 5.5.2	XP.XPORRES, XPDY; DS.DSSTDTC, DSSCATE;	Medical review required.
Subjects that developed withdrawal criteria whilst on the study but were not withdrawn;	Other Criteria	Subject developed grade 4 AE, but continued on the study.	PDEV14	Section 5.1.7.2 and 5.5.2	ADAE.ATOXGRN, TRTEMFL, AREL; SDTM.AE.AEACNOTH; SUPPAE.QVAL, QNAM	List if TOXGRN=4 and AREL='Related' and TRTEMFL='Y' and (SUPPAE.QVAL^='Drug Withdrawn' where QNAM in (ACNMSB, ACNMSB3) and SDTM.AEACNOTH^='LED TO STUDY TERMINATION')
Subjects that developed withdrawal criteria whilst on the study but were not withdrawn;	Other Criteria	Subject developed grade 3 AE, but continued on the study.	PDEV15	Section 5.1.7.2 and 5.5.2	ADAE.ASEVN, TRTEMFL, AEDECOD, AEDUR. DS.DSSTDTC, DSSCATE;	Medical review required.

	Category of Protocol Deviation	Description of Protocol Deviation	Deviation Code	Protocol Section	Variable [dataset]	Proposed check / comment
Subjects overdosed (>=110% of assigned dose)	IP Compliance	Subject was overdosed.	PDEV16	Medical defined.	ADEXSUM.AVAL where PARAMACD=' RDOSINT' or 'TRTCMP'	list if individual or cumulative relative dose intensity >=110
Deviation from GCP	Other Criteria	GCP deviation		Section 3.7		Medical review required.
NA	Other Criteria	Other protocol deviation	PDEV99	Medical defined		Medical review required.

Appendix II Description of the Case Review for Assessment of Immune-Related AEs and Definition of Infusion Related Reactions

In order to thoroughly and consistently analyze potential immune-mediated adverse events (AEs), a two-level approach is proposed including:

1. A MedDRA Preferred Term (PT) query is proposed for each event category (i.e., immune-mediated rash, colitis, pneumonitis, hepatitis, nephritis and renal dysfunction, endocrinopathies and other immune-mediated adverse reactions).

2. AEs identified by the MedDRA PT queries will then be medically reviewed using predefined case definitions for immune-mediated adverse reactions.

Level 1:

To identify potentially immune-mediated AEs, the MedDRA PT queries will be used to search for AEs of interest in the clinical database. The proposed event categories such as:

Immune-mediated rash, Immune-mediated colitis, Immune-mediated pneumonitis, Immune-mediated hepatitis, Immune-mediated nephritis and renal dysfunction, Immune-mediated endocrinopathies (Thyroid disorders: Hypothyroidism, Hyperthyroidism, and Thyroiditis), Immune-mediated endocrinopathies (Adrenal insufficiency, Immune-mediated endocrinopathies (Type 1 Diabetes Mellitus), Immune-mediated endocrinopathies (Pituitary dysfunction), Immune-mediated endocrinopathies (Hypogonadism), Other immune-mediated adverse events. Further details e.g. MedDRA PT queries are regularly updated based on the current MedDRA version.

In order to standardize the MedDRA PT queries as much as possible, High Level Terms (HLT) and Standardized MedDRA Queries (SMQ) were used whenever a choice, that was considered reflective of the events of interest, was available.

Level 2:

In a second level (medical review), the potential immune-mediated AEs identified from the search performed at Level 1, will be reviewed by qualified medical personnel to determine whether the AE meets the criteria (case definition) for an immune-mediated adverse reaction based on the following algorithm:

Table 15 Algorithm for immune-related adverse reactions

Criteria	Description
Onset	AE onset after 1st avelumab administration until up to 90 days after last dose
Duration	AE does not spontaneously resolve (i.e., without corticosteroids/ immunosuppressant treatment) within 7 days after onset
Immunosuppressive therapy	AE treated with corticosteroid or other immunosuppressant therapy. *For endocrinopathies only*: AE required hormone replacement* and /or (corticosteroid or other immunosuppressive therapy)

Etiology	No other clear etiology or Histopathology/biopsy consistent with immune-mediated event
All criteria listed in the left column need to be fulfilled for an event to meet the case definition of immune-mediated reaction.	
*Hormone replacement will be evaluated for specific endocrinopathy disorders only as follows: • Thyroid disorders (HLT): Thyroid therapy (ATC codes (H03A, H03B)) • Diabetes mellitus (including hyperglycemia): Insulin (ATC code A10A)	

Infusion related reactions are identified based on a list of MedDRA PTs and criteria on the timely relationship according to Table 16.

Table 16 Criteria for infusion related reactions

Infusion related reactions	**Reactions - Considered when onset is on the day of avelumab infusion (during or after the infusion) or the day after the avelumab infusion (irrespective of resolution date):** • Infusion related reaction • Drug hypersensitivity • Anaphylactic reaction • Hypersensitivity • Type 1 hypersensitivity **Signs and Symptoms - occurring on the day of avelumab infusion (during or after the infusion) and resolved with end date within 2 days after onset** • Pyrexia • Chills • Flushing • Hypotension • Dyspnea • Wheezing • Back pain • Abdominal pain • Urticaria