A PHASE 1 PHARMACOKINETIC– PHARMACODYNAMIC STUDY OF AVELUMAB (MSB00100718C) IN PATIENTS WITH PREVIOUSLY TREATED ADVANCED STAGE CLASSICAL HODGKIN'S LYMPHOMA

JAVELIN HODGKIN'S

STATISTICAL ANALYSIS PLAN - B9991007

Compounds:	MSB0010718C
Compound Name:	Avelumab
Version:	V3
Date:	09-Apr-2018

[redacted]

TABLE OF CONTENTS

LIST OF TABLES

APPENDICES

1. VERSION HISTORY

This Statistical Analysis Plan (SAP) for study B9991007 is based on the protocol amendment 4 dated 13NOV2017.

Table 1. Summary of Major Changes in SAP Amendments

Version	Version Date	Summary of Changes
3	09-Apr-2018	Section 3.4.1 "Study drug, study treatment and baseline definitions" - Added the definition of baseline for biomarker analyses. Section 5.1.2 "Decision Rules" - added safety stopping rules for the expansion phase, as per Protocol Amendment 4. Section 5.3.3.1 "Date of last contact" - Withdrawal of consent date was added to the derivation of date of last contact. Section 6.1.2.1 "Primary Analysis" - Added analysis summary of concordance between BICR and Investigator for objective response and best overall response. Section 6.2.2.5 "Progression Free Survival" - Expanded the analysis of Time of Follow-Up for PFS to include Kaplan-Meier estimates Section 6.2.2.6 "Overall Survival" - Expanded the analysis of Time of Follow-Up for OS to include Kaplan-Meier estimates Section 6.2.3 "Pharmacokinetic endpoints" - Clarified the PK analysis for Lead-in and Expansion phases. Section 6.2.5 "Biomarker endpoints" – Added details on how data from patients with more than one result at a visit for a specific biomarker analyte will be handled. Section 6.2.6 "Endpoints for immunogenicity data of avelumab" – provided further details for the analyses that will be performed including summaries of SAEs, AEs with grade ≥3, AEs leading to dose reduction of avelumab and AEs leading to discontinuation of avelumab by anti-drug antibody (ADA) status and, if data permits, neutralizing antibody (nAb) status. Added to Section 3.4.1 the baseline definition for immunogenicity analyses. Section 6.5.3.3 "Dose Delays" – Modified the grouping of delays for categorical analysis to be aligned with dosing window allowed per protocol. Section 6.6.1 "Adverse events" and subsections – given the nature of immune-related AEs (irAEs) and infusion related reactions (IRRs), removed the summaries of treatment-related irAEs and treatment-related IRRs; added summaries of AEs leading to dose reduction, AEs leading to interruption of study treatment and AEs leading to both dose reduction and interruption of study treatment. Appendix 1 "Immune-related Adverse Events" – provided additional ATC codes for concomitant medications in Step 4 of the case definition for irAEs; ATC code H02C was deleted from Step 4 since it reflects a group of drugs that is "anti-corticosteroids" used to treat Cushing's syndrome rather than "corticosteroids". Minor editorial and consistency changes throughout the document.
2	03-Aug-2017	Updated study design and analysis for expansion phase as per Protocol Amendment 3. Section 2 "Introduction" – Changed the cut-off date for the primary analysis from 12 months to 24 months after the last patient receives the first dose of study drug.

		Section 2.1 "Study Objectives" – Added primary, secondary CCI objectives for the expansion phase according to Protocol Amendment 3. Section 2.2 "Study Design" – Added design for expansion phase according to Protocol Amendment 3. Section 3.1 "Primary Endpoints" – added primary endpoint for expansion phase according to Protocol Amendment 3. Section 3.2 "Secondary Endpoints" – added secondary safety, efficacy, PK, immunogenicity, biomarker endpoints for expansion phase according to Protocol Amendment 3. CCI Section 3.4.1 "Study drug, study treatment and baseline definitions" – Added Cohort F for expansion phase (avelumab 70 mg IV Q2W) according to Protocol Amendment 3; provided further detail for definition of start and end dates of study treatment; added definition of baseline for the non-randomized expansion phase of the study; changed heart rate to RR in the references for derivations since the ECG CRF collects RR rather than heart rate. Section 4.1 "Full Analysis Set" – Added definition of full analysis set for the non-randomized expansion phase as per Protocol Amendment 3. Section 4.3.1. "Target Occupancy analysis set" – Changed the definition of the analysis set to include only patients with at least one receptor occupancy sample both pre- and post- Cycle 1 Day 1 dose. Section 4.3.3. "Biomarker analysis set" – provided further details for the selection of patients to be included in each of the biomarker analysis sets. Section 5.1.1. "Hypotheses and sample size determination" – Modified according to specifications in Protocol Amendment 3. Section 5.2. "General Methods" - Added Cohort F for expansion phase (avelumab 70 mg IV) according to Protocol Amendment 3 and provided the details for the summaries across treatment groups and phases for each endpoint (Table 5); the associated wording is also added to each subsection of Section 6 "Analyses and Summaries" for the relevant endpoint. Section 5.2.5. "Definition of start of new anti-cancer drug therapy" – Section added. Section 5.2.6. "Definition of start of new anti-cancer therapy" – Section added. Section 5.2.8. "Standard derivations and reporting conventions" – Removed derivation of BSA since no drug in this study is administered based on BSA. Section 5.2.10. "Adequate baseline tumor assessment" – Section added. Section 5.2.11. "Adequate post-baseline tumor assessment" – Section added. Section 5.3.2.4. "Exposure" – Imputation rules updated with regards to comparison with start date of study drug. Section 5.3.3.2. "Death date" – Imputation rules simplified to be based on the derived date of last contact. Section 5.3.3.4. "Date of start of new anti-cancer therapy" – Section added. Section 6.1.2. "Objective Response by BICR (Expansion Phase)" – Section added according to Protocol Amendment 3.

		Section 6.2.2. "Efficacy endpoints" – Section updated to reflect the different criteria of response evaluation in the lead-in and expansion phase, as per Protocol Amendment 3. Section 6.2.2.1. "Tumor shrinkage" – Index Lesions changed to Largest Dominant Masses, to correctly reflect the updated CRF (as per Cheson 2007). Section 6.2.2.3. "Duration of response" – 18 months added to the pre-specified time points where DR rate will be estimated given the increase in the minimum follow-up for all patients at the time of the cut-off for the primary analysis. Section 6.2.2.5. "Progression-free survival" – updated to reflect the tumor assessment schedule in the expansion phase, as per Protocol Amendment 3. Outcome and Event Dates for PFS and DR Analyses as well as Reason for Censoring have been updated with more details; 18 months added to the pre-specified time points where PFS rate will be estimated given the increase in the minimum follow-up for all patients at the time of the cut-off for the primary analysis Section 6.2.2.6. "Overall Survival" – Section added to reflect the additional endpoint in the expansion phase, as per Protocol Amendment 3. CCI [redacted] Section 6.4. "Subset Analyses" – Added subgroup analyses of efficacy endpoints by prior stem cell transplant type. Section 6.5.1.1. "Demographic characteristics" – Updated Geographic Regions, removed BSA. Section 6.5.1.2 "Medical history" – Added a listing for prior stem cell transplant details. Section 6.5.1.3 "Disease characteristics" – variables updated to reflect the updated CRF, including prior transplant details. Section 6.5.1.5. "Prior and concomitant immunosuppressive therapies" – Section added. Section 6.5.2. "Study conduct and patient disposition" – Study phases will not be analyzed together. Section 6.5.2.1. "Patient disposition" – Added the analyses for the non-randomized expansion phase. Section 6.5.3. "Study treatment compliance and exposure" and subsections – Added Cohort F according to Protocol Amendment 3 and provided further details for derivations. Section 6.5.3.3. "Dose delays" – Added formula for the calculation of dose delays and updated delays categories. Section 6.5.3.5. "Infusion interruptions" – Section added. Section 6.5.4 "Concomitant medications and non-drug treatments" – added summaries for pre-medications. Section 6.5.5. "Subsequent anti-cancer therapies" – Updated references to CRF forms to reflect the updated CRF. Section 6.6. "Safety Summaries and Analyses" and subsections – updated the definition and analyses for IRRs and irAEs and added Appendix 1 and 2 with the algorithms for derivation of IRRs and irAEs. Section 6.6.1.1. "All adverse events"

		– Added analysis of GVHD as per Protocol Amendment 3. Section 6.6.5 "Laboratory data" and subsections – provided the formula for corrected calcium in SI units and corrected the descriptors associated with the eDISH plot. Section 6.6.6. "Vital signs" – Schedule of assessment for expansion phase added as per Protocol Amendment 3. Removed confirmation of response throughout the document, as per Cheson Criteria 2007 and Protocol Amendment 3. Minor additional editorial changes.
1	18- Sep-2015	Not applicable (N/A)

2. INTRODUCTION

This SAP provides the detailed methodology for summary and statistical analyses of the data collected in study B9991007. This document may modify the plans outlined in the protocol; however, any major modifications of the primary endpoint definition or its analysis will also be reflected in a protocol amendment.

Statistical analyses will be performed using cleaned eCRF data as well as non-CRF data (i.e., biopsy data from the central laboratory, pharmacokinetic (PK) data, and biomarker data). The primary analysis of Target Receptor Occupancy (TO) will include all data up to a analysis cut-off date corresponding to 1 cycle of treatment (2-week or 3-week, depends on treatment cycle) after the last patient is randomized. The primary analysis for the Clinical Study Report will include all data up to an analysis cut-off date corresponding to 24 months after the last patient receives the first dose of study drug. The final analysis of the data will be performed after last patient last visit (LPLV).

Additional analyses of the data may be performed for publication or regulatory reporting purposes.

Throughout this document 'start date' refers to date of randomization for patients enrolled in the lead-in phase, and first dose of study treatment for patients enrolled in the expansion phase.

2.1. Study Objectives

Primary Objectives

Lead-in Phase

- To characterize the pharmacokinetics (PK) of different dosing regimens of avelumab and its relation to target occupancy (TO) in peripheral blood of patients with Classic Hodgkin's Lymphoma (cHL).

Expansion Phase

- To evaluate the objective response rate (ORR) of avelumab in patients with relapsed or refractory cHL who have previously been treated with an allogeneic Hematopoietic Stem Cell Transplant (HSCT).

Secondary Objectives

Lead-in Phase

- To evaluate the overall safety and tolerability of different dosing regimens of avelumab.
- To assess the immunogenicity of different dosing regimens of avelumab.
- To evaluate the effect of different dosing regimens of avelumab on pharmacodynamics biomarkers of tumor immunophenotype and anti-tumor immune response.
- To evaluate the anti-tumor activity of avelumab in patients with cHL.

Expansion Phase

- To evaluate the overall anti-tumor activity of avelumab.
- To evaluate the overall safety profile of avelumab.
- To evaluate the incidence and severity of acute and chronic GVHD.
- To characterize the pharmacokinetics of avelumab.
- To assess the immunogenicity of avelumab.
- To evaluate the phenotype and quantity of TILs and correlate these findings with antitumor activity.

CCI

[redacted]

CCI

2.2. Study Design

This is a Phase 1b, open-label, multi-center study comprising a lead-in phase and an expansion phase in patients with relapsed/refractory cHL.

Lead-in Phase

The lead-in is a multiple-dose, randomized, parallel-arm, pharmacokinetic and pharmacodynamic study of avelumab administered intravenously (IV) as a single agent in adult patients with cHL. Patients enrolled in the lead-in phase of the study are required to have relapsed following a prior autologous or allogeneic HSCT, or to be ineligible for HSCT.

In the lead-in phase of the study, a total of approximately 30 patients will be randomized across 5 treatment cohorts in a ratio of 1:1:1:1:1, with 6 patients per treatment cohort. The 5 treatment cohorts will be 70 mg every 2 weeks (Q2W) (Cohort A), 350 mg Q2W (Cohort B), 500 mg every 3 weeks (Q3W) (Cohort C), 500 mg Q2W (Cohort D), and 10 mg/kg Q2W (Cohort E). The goal of the lead-in phase is to determine the doses and schedules of avelumab that provide greater than 90% TO over the dosing interval and to identify a dose regimen for use in the dose expansion phase.

Expansion Phase

Based on the preliminary TO, safety, and efficacy results from the lead-in phase, the expansion phase will evaluate the anti-tumor activity and safety of single-agent avelumab utilizing an intra-patient dose escalation paradigm based on two of the dosing regimens studied in the lead-in phase in 40 cHL patients in whom an allogeneic HSCT has failed.

In the expansion phase all patients will commence dosing at 70 mg Q2W and will be monitored for safety and efficacy. Following 3 cycles of treatment at 70 mg Q2W,

- patients who achieve a CR at the 6-week tumor assessment will continue dosing at 70 mg Q2W;
- patients who achieve a PR at the 6-week tumor assessment will continue dosing at 70 mg Q2W for an additional 3 cycles (6 weeks);
  - if the 12-week tumor assessment still shows a PR, patients will be dose escalated to 500 mg Q2W;
- patients who achieve a SD at the 6-week tumor assessment, will be dose escalated to 500 mg Q2W.

3. ENDPOINTS AND BASELINE VARIABLES: DEFINITIONS AND CONVENTIONS

3.1. Primary Endpoints

Lead-in Phase

- Percent TO by dose/schedule in peripheral blood immune cells, including CD14+ monocytes and CD3+ T cells.

  Percent TO is defined as the percentage of target receptors occupied by avelumab in peripheral blood immune cells, including, but not limited to monocytes.

- Pharmacokinetic parameters of avelumab including, but not limited to, C_{max}, T_{max}, AUC_{last}, T_{last}, $AUC_{sd,\tau}$, $t_{1/2}$, $AUC_{sd,inf}$, CL, and Vz as data permit. Multiple Dose (MD) - $C_{ss,max}$, $T_{ss,max}$, $AUC_{ss,\tau}$, $t_{1/2}$, $C_{ss,min}$, $C_{ss,av}$, CL and V_{ss} ,R_{ac} ($AUC_{ss,\tau}$ /$AUC_{sd,\tau}$) and R_{ss} ($AUC_{ss,\tau}$ /$AUC_{sd,inf}$) as data permit.

Table 2. PK Parameters to be Determined for Avelumab

Parameter	Definition	Method of Determination
AUC_{last}	Area under the plasma concentration-time profile from time zero((pre-dose) to the time of the last quantifiable concentration (C_{last})	Linear/Log trapezoidal method
$AUC_{sd,\tau}$ $AUC_{ss,\tau}$	Area under the plasma concentration-time profile from time zero ((pre-dose) to the next dose (after single dose(sd) and at steady state (ss))	Linear/Log trapezoidal method
C_{max}	Maximum observed plasma concentration	Observed directly from data
T_{max}	Time for C_{max}	Observed directly from data as time of first occurrence
$T_{ss,max}$	Time for Cmax at steady state	Observed directly from data as time of first occurrence at steady state

Parameter	Definition	Method of Determination
$t_{½}$[a]	Terminal half-life	$Log_e(2)/k_{el}$, where k_{el} is the terminal phase rate constant calculated by a linear regression of the log-linear concentration-time curve. Only those data points judged to describe the terminal log-linear decline were used in the regression.
C_{trough}	Pre-dose concentration during multiple dosing	Observed directly from data
$C_{ss, av}$	Pre-dose concentration during multiple dosing	AUCss / τ
CL	Clearance	Dose / AUCτ for steady state
V_z[a]	Volume of distribution	Dose / (AUCτ ·kel) for steady state
R_{ac}	Accumulation ratio	$AUC_{ss,\tau}$ / $AUC_{sd,\tau}$
R_{ss}	Steady state accumulation ratio	$AUC_{ss,\tau}$/$AUC_{sd,inf}$
C_{max}(dn)	Dose normalized C_{max}	C_{max} / Dose
T_{last}	The last time point of the last quantifiable concentration (Clast)	Observed directly from data
$C_{ss,max}$	Maximum observed plasma concentration at steady state	Observed directly from data
$AUC_{sd,inf}$	Area under the plasma concentration-time profile from time zero to infinity, after single dose.	Linear/Log trapezoidal method
$C_{ss,max}$	Maximum observed plasma concentration at steady state	Observed directly from data
$C_{ss,min}$	Minimum observed plasma concentration at steady state	Observed directly from data

[a] If data permit

Expansion Phase

- Objective response defined by the Lugano Classification (Cheson et al., 2014) as evaluated by the blinded independent central review (BICR).

  OR is defined as complete response (CR) or partial response (PR) according to the Lugano Classification (Cheson et al., 2014) from 'start date' until documented disease progression.

3.2. Secondary Endpoints

3.2.1. Safety endpoints

Lead-in Phase and Expansion Phase

- Adverse Events as characterized by type, severity (as graded by National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) v.4.03), timing, seriousness, and relationship to study therapy.

  AEs will be graded by the investigator according to the Common Terminology Criteria for Adverse Events (CTCAE) version 4.03 and coded using the Medical Dictionary for Regulatory Activities (MedDRA)

- Laboratory abnormalities as characterized by type, severity (as graded by NCI CTCAE v. 4.03), and timing.

Expansion Phase

- Acute GVHD as defined by the modified Seattle Glucksberg criteria (Consensus Conference on Acute GVHD Grading Criteria) and Chronic GVHD as defined by the NIH Consensus Development Project.

3.2.2. Efficacy endpoints

Lead-in Phase

- Objective response according to Response Criteria for Malignant Lymphoma (Cheson et al., 2007) per Investigator assessment.

- Disease control (DC), time to tumor response (TTR), duration of response (DR), progression-free survival (PFS) per Investigator assessment.

Expansion Phase

- Objective response as defined by the Lugano Classification (Cheson et al., 2014) and evaluated by Investigator's assessment.

- Time to tumor response (TTR), duration of response (DR), Disease Control (DC) and progression-free survival (PFS) according to the Lugano Classification (Cheson et al., 2014) by BICR and by Investigator's assessment, as well as overall survival (OS).

  OR is defined as complete response (CR) or partial response (PR) according to Response Criteria for Malignant Lymphoma (Cheson et al., 2007) (Lead-in Phase) or Lugano Classification (Cheson et al., 2014) (Expansion Phase) from 'start date' until first documentation of progressive disease (PD).

  DR is defined, for patients with OR, as the time from first documentation of objective response (CR or PR) to the date of first documentation of PD or death due to any cause.

DC is defined as CR, PR or SD. Criterion for SD must have been met at least 6 weeks after the 'start date'.

TTR is defined, for patients with an OR, as the time from the 'start date' to the first documentation of objective response (CR or PR).

PFS is defined as the time from 'start date' to the date of the first documentation of PD or death due to any cause, whichever occurs first.

OS is defined as the time from 'start date' to the date of death due to any cause.

3.2.3. Pharmacokinetic endpoints

Expansion Phase

- Pharmacokinetic parameters of avelumab including, but not limited to, C_{max}, T_{max}, $AUC_{sd,\tau}$, $t_{1/2}$, CL, and V_z as data permit. Multiple Dose (MD) - $C_{ss,max}$, $T_{ss,max}$, $AUC_{ss,\tau}$, $t_{1/2}$, $C_{ss,min}$, $C_{ss,av}$, CL, and V_{ss} as data permit.

  Definition of PK parameters is provided in Table 2.

3.2.4. Immunogenicity endpoints

Lead-in Phase and Expansion Phase

- Anti-drug antibodies (ADAs; neutralizing antibodies) and serum titers against avelumab.

3.2.5. Biomarker endpoints

Lead-in Phase

- Relative expression of transcripts associated with immune activation and regulation in tumor biopsy tissue by gene expression profiling.

- Phenotype, relative proportions, activation state and PD-L1 expression of peripheral blood T cell subsets by flow cytometry.

Lead-in Phase and Expansion Phase

- Phenotype, quantity, and localization of tumor infiltrating lymphocytes (TILs) in tumor biopsy tissue by immunohistochemistry (IHC).

Table 3. Biomarker Definition and Determination

Parameter	Definition	Method of Determination
Peripheral blood target occupancy	Occupancy of PD-L1 in peripheral blood immune cells by avelumab	Flow cytometry
Biopsy CD8, FoxP3, PD-L1 expression	Expression of immune- and target-related markers on immune cells, tumor cells and tumor vs stroma	Immunohistochemistry

Biopsy expression of RNA transcripts	Expression of immune- and pathway-related RNA transcripts	RNA Sequencing
Peripheral blood immune cell phenotypes	Relative proportion of T cell subsets	Flow cytometry

CCI

[redacted]

3.4. Baseline Variables

3.4.1. Study drug, study treatment and baseline definitions

In this study, '**study drug**' refers to avelumab and '**study treatment**' (or '**treatment group**') refers to one of the following:

Lead-in Phase

- Cohort A: avelumab 70 mg IV Q2W.
- Cohort B: avelumab 350 mg IV Q2W.
- Cohort C: avelumab 500 mg IV Q3W.
- Cohort D: avelumab 500 mg IV Q2W.
- Cohort E: avelumab 10 mg/kg IV Q2W.

Expansion Phase

- Cohort F: avelumab 70 mg IV Q2W or avelumab 70 mg IV Q2W followed by avelumab 500 mg Q2W

  As described in Section 2.2, patients enrolled in Cohort F initiate dosing with avelumab at 70 mg Q2W but intra-patient dose escalation to 500 mg Q2W may occur.

Start and end dates of study treatment:

The date/time of first dose of study treatment is the earliest date/time of non-zero dosing of the study drug.

The date/time of last dose of study treatment is the latest date/time of non-zero dosing of the study drug.

Definition of baseline:

Definition of baseline for efficacy analyses in the lead-in phase (randomized)

The last measurement prior to randomization will serve as the baseline measurement for efficacy analyses. If such a value is missing, the last measurement prior to the first dose of study treatment will be used as the baseline measurement except for analyses of tumor assessments data where the baseline assessment would be considered as missing.

Definition of baseline for immunogenicity analyses

The last available assessment prior to the start of treatment with avelumab is defined as 'baseline' result or 'baseline' assessment. If an assessment is planned to be performed prior to the first dose of avelumab in the protocol and the assessment is performed on the same day as the first dose of avelumab, it will be assumed that it was performed prior to avelumab administration, if assessment time point is not collected or is missing.

Definition of baseline for biomarker analyses

The last available assessment prior to the start of study treatment is defined as 'baseline' value or 'baseline' assessment for biomarker analyses. For biomarkers that are planned to be measured on Cycle 1 Day 1, if the assessment time point is not collected or is missing, it will be assumed that the measurement was performed prior to first dose of study treatment.

Definition of baseline for efficacy analyses in the expansion phase (non-randomized) and for safety analyses

The last available assessment prior to the start of study treatment is defined as 'baseline' value or 'baseline' assessment for safety and efficacy (for the non-randomized phase) analyses. If an assessment is planned to be performed prior to the first dose of study treatment in the protocol and the assessment is performed on the same day as the first dose of study treatment, it will be assumed that it was performed prior to study treatment administration, if assessment time point is not collected or is missing. If assessment time points are collected, the observed time point will be used to determine pre-dose on study day 1 for baseline calculation. Unscheduled assessments will be used in the determination of baseline. However, if time is missing, an unscheduled assessment on study day 1 will be considered to have been obtained after study treatment administration.

Patients who start treatment and discontinue from the study on the same day may have two different sets of data collected on study day 1 (one during study and one in the End of Treatment (EOT) visit. Data reported at the EOT visit are not eligible for baseline selection.

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

Baseline for RR and QT/QTc interval assessments will be derived from the visit where both RR and QT are not missing. Triplicate ECGs are collected in the study and the baseline for each ECG measurement is the average of the pre-dose replicate measurements on the baseline day. Unscheduled assessments will not be included in the calculation of the average. QTcB and QTcF will be derived based on RR and QT. The average of the replicate measurements will be determined after the derivation of the individual parameter at each time point.

3.4.2. Baseline characteristics

Baseline characteristics (including demographics, physical measurements, disease history and prior anti-cancer therapies) are described in Section 6.5.1. These baseline characteristics are not planned to be included as stratification variables or covariates in statistical models unless otherwise specified in Section 6.

3.5. Safety Endpoints

3.5.1. Adverse events

Treatment-Emergent Adverse Events

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.

On-treatment period is defined as the time from the first dose of study treatment through minimum (30 days + last dose of study treatment, start day of new anti-cancer drug therapy – 1 day). The start day of new anti-cancer drug therapy after the first dose of study treatment is derived as outlined in Section 5.2.5.

Adverse Events of Special Interest (AESIs)

AESIs are immune-related adverse events (irAE) and infusion-related reactions (IRRs). The criteria for classification of an AE as an irAE or IRR are described in Appendices 1 and 2, respectively.

4. ANALYSIS SETS

Data for all patients will be assessed to determine if patients meet the criteria for inclusion in each analysis population prior to releasing the database and classifications will be documented per Pfizer's standard operating procedures.

Only patients who signed informed consent will be included in the analysis sets below.

4.1. Full Analysis Set

For the lead-in phase (randomized): The full analysis set (FAS) will include all randomized patients. Patients will be classified according to the study treatment assigned at randomization.

For the expansion phase (non-randomized): The FAS will include all patients who receive at least one dose of study drug. Patients will be classified according to the study treatment actually received. If a patient receives more than one treatment the patient will be classified according to the first study treatment received.

4.2. Safety Analysis Set

The safety analysis set will include all patients who receive at least one dose of study drug. Patients will be classified according to the study treatment assigned at randomization unless the incorrect treatment(s) was/were received throughout the dosing period in which case patients will be classified according to the first study treatment received. In the expansion phase, since non-randomized, the FAS and the safety analysis set are identical.

4.3. Other Analysis Set

4.3.1. Target Occupancy analysis set

The TO analysis set includes patients in the safety analysis set who have at least one Receptor Occupancy blood sample collected both pre and post Cycle 1 Day 1 dose.

4.3.2. PK analysis set

The PK concentration analysis set is 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 (LLQ) for avelumab.

The PK parameter analysis set is a subset of the safety analysis set and will include patients who have at least one of the PK parameters of interest for avelumab.

4.3.3. Biomarker analysis set

The biomarker analysis set is a subset of the safety analysis set and will include patients who have at least one baseline biomarker sample.

- Whole Blood Biomarker Analysis Set: subjects in the biomarker analysis set who have at least one whole blood biomarker sample collected
- Plasma Biomarker Analysis Set: subjects in the biomarker analysis set who have at least one plasma biomarker sample collected
- Tumor Biopsy Biomarker Analysis Set: subjects in the biomarker analysis set who have provided at least one evaluable tumor biopsy sample.

4.3.4. Immunogenicity analysis set

The immunogenicity analysis set is a subset of the safety analysis set and will include patients who have at least one ADA/nAb sample collected for avelumab.

5. GENERAL METHODOLOGY AND CONVENTIONS

5.1. Hypotheses and Decision Rules

5.1.1. Hypotheses and sample size determination

Lead-in Phase

There is no formal hypothesis testing in this phase. Approximately 30 patients will be randomized in a 1:1:1:1:1 ratio across 5 treatment groups. The 6 patients per treatment group in the lead-in phase will be used to enable the initial estimation of TO; based on the historical data, with an observed mean TO of 90%, the corresponding 95% confidence interval (CI) will be 88.98% to 91.02%, assuming a standard deviation of 1.27%.

Expansion Phase

There is no formal hypothesis testing in this phase. Approximately 40 patients will be enrolled. With 40 treated patients, ORR can be estimated with a standard error not exceeding 0.079.

Table 4 provides 95% CIs for ORR based on different observed responses in the post-allogeneic HSCT cohort.

Table 4. Sample Size and Exact 95% CIs for ORR - Post-allogeneic HSCT cohort F

Number of patients in Cohort F	Number of observed responders	Observed ORR	95% CI for True ORR
40	16	40%	(24.86%, 56.67%)
	20	50%	(33.80%, 66.20%)
	24	60%	(43.33%, 75.14%)
	28	70%	(53.47%, 83.44%)
	32	80%	(64.35%, 90.95%)

The study plans to enroll approximately 70 patients in total.

5.1.2. Decision rules

Lead-in Phase

A dose regimen studied in the lead-in phase may be selected for the expansion phase if, based on the data from the lead-in phase:

- greater than a mean 90% TO after 1 treatment cycle is achieved, and
- at least 3 objective responses per Response Criteria for Malignant Lymphoma are observed.

Expansion Phase

The safety stopping rules will be based on the observed incidence of acute GVHD Grade ≥3, acute GVHD-related mortality and Grade ≥4 irAEs, within 16 weeks from start of study treatment for each patient in the expansion phase and each patient with prior allogeneic HSCT in the lead-in phase, according to the following clinical thresholds:

1. Acute GVHD Grade ≥3 <33% (Grade 3 liver GVHD must be confirmed by biopsy).
2. Acute GVHD-related mortality ≤20%.
3. irAEs Grade ≥4 <15%.

Using a Beta(0.5, 0.5) prior, the trial will be stopped if given the data the posterior probability that the toxicity rate is above the clinical threshold for any of the event categories is ≥ 70%. The following stopping rules are obtained using the beta-binomial posterior.

Table 5. Safety Stopping Rules

Number of Patients	Gr ≥3 acute GVHD *	GVHD-related deaths	Gr ≥4 irAEs (excluding acute GVHD)
15	6/15 (40%)	4/15 (27%)	3/15 (20%)
20	8/20 (40%)	5/20 (25%)	4/20 (20%)
25	10/25 (40%)	6/25 (24%)	5/25 (20%)
30	12/30 (40%)	8/30 (27%)	6/30 (20%)
35	13/35 (37%)	9/35 (26%)	7/35 (20%)
40	15/40 (38%)	10/40 (25%)	8/40 (20%)

* Grade 3 liver GVHD must be confirmed by biopsy.

Only patients meeting eligibility criteria will be included in the evaluation. Events from patients who at the time of the formal evaluation have not yet been followed for 16 weeks from first dose of study treatment will also be considered in the evaluation. The study will stop if the stopping boundary is crossed for any of the 3 event categories.

Two formal evaluations of safety will be performed:

- 16 weeks after 15 patients have received the first dose of study treatment.
- 16 weeks after 30 patients have received the first dose of study treatment.

5.2. General Methods

As described in Section 3.4, in this study '**treatment group**' refers to one of the following:

Lead-in Phase

- Cohort A: avelumab 70 mg IV Q2W.
- Cohort B: avelumab 350 mg IV Q2W.
- Cohort C: avelumab 500 mg IV Q3W.
- Cohort D: avelumab 500 mg IV Q2W.
- Cohort E: avelumab 10 mg/kg IV Q2W.

Expansion Phase

- Cohort F: avelumab 70 mg IV Q2W or avelumab 70 mg IV Q2W followed by avelumab 500 mg Q2W

  As described in Section 2.2, patients enrolled in Cohort F initiate dosing with avelumab at 70 mg Q2W but intra-patient dose escalation to 500 mg Q2W may occur

Unless otherwise specified (see Table 6), all data will be analyzed by study phase and treatment group, and treatment groups in the lead-in phase will also be pooled together.

Table 6. Summaries for each endpoint

Endpoints	Lead-in by treatment group	Lead-in treatment groups pooled	Expansion	Lead-in and expansion treatment groups pooled
Efficacy	✓	✓	✓	
Target Occupancy	✓			
Safety	✓	✓	✓	
PK	✓		✓	
Immunogenicity	✓	✓	✓	✓
Biomarkers	✓	✓	✓	✓
Exposure	✓		✓	
Demographic characteristics	✓	✓	✓	✓
Disease characteristics, prior anti-cancer therapies, prior and concomitant therapies	✓	✓	✓	
Disposition	✓	✓	✓	

5.2.1. Data handling after the cut-off date

Data after the cut-off date may not undergo the cleaning process and will not be displayed in any listings or used for summary statistics, statistical analyses or imputations.

5.2.2. Pooling of centers

In order to provide overall estimates of treatment effects, data will be pooled across centers. The 'center' factor will not be considered in statistical models or for subgroup analyses due to the high number of participating centers in contrast to the anticipated small number of patients randomized/treated at each center.

5.2.3. 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 (SD), minimum, maximum and first and third quartile (Q1 and Q3).

Qualitative variables will be summarized by frequency counts and percentages. Unless otherwise specified, 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.

In case the analysis refers only to certain visits, percentages will be based on the number of patients still present in the study at that visit, unless otherwise specified.

5.2.4. Definition of study day

Start day of study treatment is the day of the first dose of study treatment.

The study day for assessments occurring on or after the start of study treatment (e.g., adverse event onset, tumor measurement) will be calculated as:

Study day = Date of the assessment/event - start of study treatment + 1.

The study day for assessments occurring prior to the first dose of study treatment (e.g., baseline characteristics, medical history) will be negative and calculated as:

Study day = Date of the assessment/event - start of study treatment.

The study day will be displayed in all relevant data listings.

5.2.5. Definition of start of new anti-cancer drug therapy

Start date of new anti-cancer drug therapy is used to determine the end of the on-treatment period (see Section 5.2.7).

The start date of new anti-cancer drug therapy is the earliest start date of anti-cancer drug therapy recorded in the ‘Follow-up Cancer Therapy’ eCRF pages that is after the first dose of study treatment. When start date of anti-cancer drug therapy is missing or partially missing, the imputation rules described in Section 5.3.3.4 should be applied using only data from the ‘Follow-up Cancer Therapy’ eCRF pages.

5.2.6. Definition of start of new anti-cancer therapy

Start date of new anti-cancer therapy (drug, radiation, surgery) is used for censoring in efficacy analyses (see Section 6.1.2 and Section 6.2.2).

The start date of new anti-cancer therapy is the earliest date after the first dose of study treatment for the expansion phase or after the date of randomization for the lead-in phase amongst the following:

- Start date of anti-cancer drug therapy recorded in the 'Follow-up Cancer Therapy' eCRF pages
- Start date of radiation therapy recorded in 'Concomitant Radiation Therapy', and 'Follow-up Radiation Therapy' eCRF pages with 'Treatment Intent' = 'Curative in intent'
- Surgery date recorded in 'Concomitant Surgery', and 'Follow-up Surgery' eCRF pages when 'Surgery Outcome' = 'Resected' or 'Partially Resected'.

When start date of anti-cancer therapy is missing or partially missing, the imputation rules described in Section 5.3.3.4 should be applied using 'Follow-up Cancer Therapy', 'Concomitant Radiation Therapy', 'Follow-up Radiation Therapy', 'Concomitant Surgery', and 'Follow-up Surgery' eCRF pages.

5.2.7. Definition of on-treatment period

Safety endpoints will be summarized based on the on-treatment period unless otherwise specified.

On-treatment period is defined as the time from the first dose of study treatment through minimum (30 days + last dose of study treatment, start day of new anti-cancer drug therapy – 1 day).

Safety data collected outside the on-treatment period as described above will be listed and flagged in listings but not summarized.

5.2.8. Standard derivations and reporting conventions

The following conversion factors will be used to convert days into weeks, months or years: 1 week = 7 days, 1 month = 30.4375 days, 1 year = 365.25 days.

Demographics and physical measurements:

- Age [years]:
  - (date of given informed consent - date of birth + 1) / 365.25
  - In case of missing day, day only: Age [years]: (year/month of given informed consent – year/month of birth)
  - In case only year of birth is given: Age [years]: (year of given informed consent - year of birth)

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

- BMI (kg/m^2) = weight (kg)/[height (m)]^2

For reporting conventions, mean and median should generally be displayed one more decimal place than the raw data and standard deviation should be displayed to two more decimal places than the raw data. Percentages will be reported to 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.

5.2.9. Unscheduled visits

Generally, data collected at unscheduled visits will be included and analyzed for both safety and efficacy analyses in the same fashion as the data collected at scheduled visits except where otherwise noted in the sections that follow. Descriptive statistics (mean, SD, median, minimum, maximum, quartiles) by nominal visit or time point for safety endpoints such as laboratory measurements, ECGs and vital signs will include only data from scheduled visits.

5.2.10. Adequate baseline tumor assessment

Adequate baseline is defined using the following criteria:

- All baseline assessments must be within 28 days prior to and including the 'start date'.
- All documented lesions must have non-missing assessments (i.e. non missing measurements for target lesions and non-missing lesions assessment status at baseline for non-target lesions).

5.2.11. Adequate post-baseline tumor assessment

An adequate post-baseline assessment is defined as an assessment where a response of CR, PR, SD, or PD can be determined (see Section 6.1.2.1). Time points where the response is not evaluable (NE) or no assessment was performed will not be used for determining the censoring date.

5.3. Methods to Manage Missing Data

5.3.1. Missing data

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

In all patient data listings imputed values will be presented. In all listings imputed information will be flagged.

Missing statistics, e.g. when they cannot be calculated, should be presented as 'ND' or 'NA'. For example, if N=1, the measure of variability (SD) cannot be computed and should be presented as 'ND' or 'NA'.

5.3.1.1. Pharmacokinetic concentrations

Concentrations Below the Limit of Quantification

For all calculations, figures and estimation of individual pharmacokinetic parameters, all concentrations assayed as below the level of quantification (BLQ) will be set to zero. In log-linear plots these values will not be represented. The BLQ values will be excluded from calculations of geometric means and their CIs. A statement similar to 'All values reported as

BLQ have been replaced with zero' should be included as a footnote to the appropriate tables and figures.

Deviations, Missing Concentrations and Anomalous Values

In summary tables and plots of median profiles, concentrations will be set to missing if one of the following cases is true:

1. A concentration has been reported as ND (i.e., not done) or NS (i.e., no sample);
2. A deviation in sampling time is of sufficient concern or a concentration has been flagged as anomalous by the clinical pharmacologist.

Summary statistics will not be presented at a particular time point if more than 50% of the data are missing. For analysis of pharmacokinetic concentrations, no values will be imputed for missing data.

5.3.1.2. Pharmacokinetic parameters

Actual PK sampling time will be used for the derivation of PK parameters. If a PK parameter cannot be derived from a patient's concentration data, the parameter will be coded as NC (i.e., not calculated). NC values will not be generated beyond the day that a patient discontinues.

In summary tables, statistics will be calculated by setting NC values to missing. Statistics will not be presented for a particular treatment if more than 50% of the data are NC. For statistical analyses (i.e., analysis of variance), PK parameters coded as NC will also be set to missing.

If an individual patient has a known biased estimate of a PK parameter (due for example to a deviation from the assigned dose level), this will be footnoted in summary tables and the parameter will not be included in the calculation of summary statistics or statistical analyses.

5.3.2. Handling of incomplete dates

5.3.2.1. Disease history

Incomplete dates for disease history (e.g., initial diagnosis date, date of documented, locally advanced, inoperable or metastatic disease diagnosis, date of response or progression in prior treatment) will be imputed as follows:

- If the day is missing, it will be imputed to the 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.

5.3.2.2. Adverse events

Incomplete AE-related dates will be imputed as follows:

- If the AE onset date is missing completely, then the onset date will be replaced by the start of study treatment.
- If only the day part of the AE onset date is missing, but the month and year are equal to the start of study treatment, then the AE onset date will be replaced by the start of study treatment. For example, if the AE onset date is --/JAN/2015, and study treatment start date is 15/JAN/2015, then the imputed AE onset date will be 15/JAN/2015.
- If both the day and month of the AE onset date are missing but the onset year is equal to the start of study treatment, then the onset date will be replaced by the start of study treatment. For example, if AE onset date is --/---/2014, and study treatment start date is 19/NOV/2014, then the imputed AE onset date will be 19/NOV/2014.
- In all other cases the missing onset day or missing onset month will be replaced by 1.
- Incomplete stop date will be replaced by the last day of the month (if day is missing only), if not resulting in a date later than the date of patient's death. In the latter case the date of death will be used to impute the incomplete stop date.
- In all other cases the incomplete stop date will not be imputed. If stop date of AE is after the date of cut-off outcome of AE is ongoing at cut-off.

5.3.2.3. Prior and concomitant medications

Incomplete prior/concomitant medication dates will be imputed as follows:

- If the medication date is missing completely, then the medication date will be replaced by the start of study treatment.
- If the day of medication date is missing, but the month and year are equal to the start of study treatment, then the medication date will be replaced by the start of study treatment. For example, if the medication start date is --/JAN/2015, and study treatment start date is 15/JAN/2015, then the imputed medication start date will be 15/JAN/2015.
- If both the day and month of medication start date are missing but the start year is equal to the start of study treatment, then the medication date will be replaced by the start of study treatment. For example, if the medication start date is --/---/2014, and study treatment start date is 19/NOV/2014, then the imputed medication start date will be 19/NOV/2014.
- In all other cases the missing medication day or missing medication month will be replaced by 1.
- Incomplete stop date will be replaced by the last day of the month (if day is missing only), if not resulting in a date later than the date of patient's death. In the latter case the date of death will be used to impute the incomplete stop date.

- In all other cases the incomplete medication stop date will not be imputed.

5.3.2.4. Exposure

No imputation will be done for first dose date. Date of last dose of study drug, if unknown or partially unknown, will be imputed as follows:

- If the last date of study drug is completely missing and there is no End of Treatment eCRF page and no death date, the patient should be considered to be ongoing and use the cut-off date for the analysis as the last dosing date
- If the last date of study drug is completely or partially missing and there is EITHER an End of Treatment eCRF page OR a death date available (within the cut-off date), then imputed last dose date is:

  = 31DECYYYY, if only Year is available and Year < Year of min (EOT date, death date)

  = Last day of the month, if both Year and Month are available and Year = Year of min (EOT date, death date) and Month < the month of min (EOT date, death date)

  = min (EOT date, death date), for all other cases.

5.3.3. Imputation rules for date of last contact and efficacy assessments

5.3.3.1. 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 'Survival Follow-up' eCRF (do not use date of survival follow-up assessment unless status is 'alive')
- Study drug start and end dates
- Randomization date
- Withdrawal of consent date
- 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.

5.3.3.2. Death date

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

- If the date is missing it will be imputed as the day after the date of last contact
- If the day or both day and 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

5.3.3.3. Tumor assessments

All investigation dates (e.g., X-ray, CT scan) must be completed with day, month and year.

If there are multiple scan dates associated with an evaluation, i.e., radiological assessments occur over a series of days rather than the same day, the choice of date of assessment could impact the date of progression and/or date of response. If there are multiple scan dates associated with an evaluation, the earliest of the scan dates associated with the evaluation will be used as the date of assessment.

If one or more investigation dates for an evaluation are incomplete but other investigation dates are available, the incomplete date(s) are not considered for calculation of the assessment date and assessment date is calculated as the earliest of all investigation dates (e.g., X-ray, CT-scan).

If all measurement dates for an evaluation have no day recorded, the 1st of the month is used.

If the month is not completed, for any of the investigations for an evaluation, the respective assessment will be considered to be at the date which is exactly between the previous and the following assessment. If both a previous and following assessments are not available, this assessment will not be used for any calculations.

5.3.3.4. Date of start of new anti-cancer therapy

Incomplete dates for start date of new anti-cancer therapy (drug therapy, radiation, surgery) will be imputed as follows and will be used for determining censoring dates for efficacy analyses and in the derivation of the end of on-treatment period. PD date below refers to PD date by investigator assessment.

- The end date of new anti-cancer therapy will be included in the imputations for start date of new anti-cancer therapy. If the end date of new anti-cancer therapy is
  - completely missing then it will be ignored in the imputations below
  - partially missing with only year (YYYY) available then the imputations below will consider 31DECYYYY as the end date of the new anti-cancer therapy

- partially missing with only month and year available then the imputations below will consider the last day of the month for MMMYYYY as the end date of the new anti-cancer therapy

- For patients who have not discontinued study treatment at the analysis cut-off date, last dose of study treatment is set to the analysis cut-off date in the imputations below.

- If the start date of new anti-cancer therapy is completely or partially missing then the imputed start date of new anti-cancer therapy is derived as follows:

  - Start date of new anti-cancer therapy is completely missing

    Imputed start date = min [max(PD date + 1, last dose of study treatment + 1), end date of new anti-cancer therapy]

  - Only year (YYYY) for start of anti-cancer therapy is available

    IF YYYY < Year of min [max(PD date + 1, last dose of study treatment + 1), end date of new anti-cancer therapy] THEN imputed start date = 31DECYYYY;

    ELSE IF YYYY = Year of min [max(PD date + 1, last dose of study treatment + 1), end date of new anti-cancer therapy]

    THEN imputed start date = min [max(PD date + 1, last dose of study treatment + 1), end date of new anti-cancer therapy]

    ELSE IF YYYY > Year of min [max(PD date + 1, last dose of study treatment + 1), end date of new anti-cancer therapy]

    THEN imputed start date = 01JANYYYY

  - Both Year (YYYY) and Month (MMM) for start of anti-cancer therapy are available

    IF

    YYYY = Year of min [max(PD date + 1, last dose of study treatment + 1), end date of new anti-cancer therapy], AND

    MMM < Month of min [max(PD date + 1 day, last dose of study treatment + 1 day), end date of new anti-cancer therapy]

    THEN

    imputed start date = DAY (Last day of MMM) MMM YYYY ;

    ELSE IF

    YYYY = Year of min [max(PD date + 1, last dose of study treatment + 1), end date of new anti-cancer therapy], AND

    MMM = Month of min [max(PD date + 1 day, last dose of study treatment + 1 day), end date of new anti-cancer therapy]

    THEN

    imputed start date = min [max(PD date + 1 day, last dose of study treatment + 1 day), end date of new anti-cancer therapy]);

ELSE IF

YYYY = Year of min [max(PD date + 1, last dose of study treatment + 1), end date of new anti-cancer therapy], AND

MMM > Month of min [max(PD date + 1 day, last dose of study treatment + 1 day), end date of new anti-cancer therapy]

THEN

imputed start date = 01 MMM YYYY;

ELSE IF

YYYY < Year of min [max(PD date + 1, last dose of study treatment + 1), end date of new anti-cancer therapy]

THEN

imputed start date = DAY (Last day of MMM) MMM YYYY;

ELSE IF

YYYY > Year of min [max(PD date + 1, last dose of study treatment + 1), end date of new anti-cancer therapy]

THEN

imputed start date = 01 MMM YYYY.

6. ANALYSES AND SUMMARIES

Refer to Section 4 for definitions of analysis sets and Section 5.2 for general methodology.

Unless otherwise specified, all data will be analyzed by study phase and treatment group, and treatment groups in the lead-in phase will also be pooled together.

6.1. Primary Endpoints

6.1.1. Percent TO (Lead-in Phase)

6.1.1.1. Primary analysis

The following analyses will be based on the TO analysis set for patients in the lead-in phase. Percent TO will be summarized by treatment group; treatment groups in the lead-in phase will not be pooled together.

Percent PD-L1 TO will be summarized descriptively (n, mean, SD, CV, median, minimum, maximum, geometric mean, its associated CV, and 95% CI) by treatment group, cycle, day and nominal time. Mean PD-L1 TO before the second cycle (at the expected C_{trough}), together with safety and clinical activity, will be used to decide which dose to move forward into the expansion phase of the study.

Presentation of Percent PD-L1 TO data will include:

- Descriptive statistics (n, mean, SD, %CV, median, minimum, maximum) of percent TO will be presented in tabular form by treatment group, cycle, day and nominal time.

- Linear-linear plots of mean and median percent TO by nominal time will be presented for TO sampling days by treatment group, cycle, and study day. Similar plots will be presented for each individual patient percent TO. Patients who have undergone intra-patient dose reduction or escalation will be excluded from the median percent TO plots.

- Box plots for percent TO will be generated. Individual data points, the geometric mean and the median of the parameter in each treatment will be overlaid on the box plots. If a treatment group has limited evaluable TO data (n<4), matchstick plots showing changes in percent TO in individual patients will then be generated. The geometric mean of the parameter in each treatment will be overlaid in the plots.

Percent TO will be plotted for each treatment group using a box-whisker plot by cycle and day within cycle in order to assess the attainment of a steady-state pharmacodynamic effect.

6.1.2. Objective Response by BICR Assessment (Expansion Phase)

6.1.2.1. Primary analysis

The following analyses will be based on the FAS. Assessment of response will be made using the Lugano Classification (Cheson et al., 2014). Assessments below refer to BICR assessment. Similar analyses will be repeated using investigator assessment as secondary analysis.

Best overall response (BOR) will be assessed based on reported overall lesion responses at different evaluation time points from the 'start date' until the first documentation of PD, according to the following rules. Only tumor assessments performed on or before the start of any further anti-cancer therapies will be considered in the assessment of BOR. Clinical deterioration will not be considered as documentation of disease progression.

BOR:

- CR = one objective status of CR documented before first documentation of PD.

- PR = one objective status of PR documented before first documentation of PD (and not qualifying for CR).

- SD = at least one SD assessment (or better) ≥ 6 weeks after 'start date' and before first documentation of PD (and not qualifying for CR or PR).

- PD = first documentation of PD ≤ 12 weeks after 'start date' (and not qualifying for CR, PR, SD). NE: all other cases.

- NE: all other cases.

An objective status of PR or SD cannot follow one of CR.

Objective Response (OR) is defined as BOR of CR or PR according to the Lugano Classification (Cheson et al., 2014).

Patients who do not have a post-baseline radiographic tumor assessment due to early progression, who receive anti-cancer therapies other than the study treatments prior to reaching a CR or PR, or who die, progress, or drop out for any reason prior to reaching a CR or PR will be counted as non-responders in the assessment of OR. Each patient will have an objective response status (0: no OR; 1: OR). OR rate (ORR) is the proportion of patients with OR in the analysis set.

ORR by treatment group will also be calculated along with the 2-sided 95% CI using the Clopper-Pearson method (exact CI for a binomial proportion as computed by default by the FREQ procedure using the EXACT option).

In addition, the frequency (number and percentage) of patients with a BOR of CR, PR, SD, PD, and NE will be tabulated. Patients with BOR of NE will be summarized by reason for having NE status. The following reasons will be used:

- No baseline assessment
- No post-baseline assessments due to death
- No post-baseline assessments due to other reasons
- All post-baseline assessments have overall response NE
- New anti-cancer therapy started before first post-baseline assessment
- SD of insufficient duration (<6 weeks after the 'start date')
- PD too late (>12 weeks after the 'start date')

Special and rare cases where BOR is NE due to both SD of insufficient duration and late PD will be classified as 'SD too early' (ie, SD of insufficient duration).

BICR vs Investigator Assessment:

Table 7 outlines the possible BOR outcomes by investigator and BICR.

Table 7. Possible BOR Outcomes for Investigator vs BICR

BOR		BICR Assessment				
		CR	**PR**	**SD**	**PD**	**NE**
Investigator Assessment	**CR**	n_{11}	n_{12}	n_{13}	n_{14}	n_{15}
	PR	n_{21}	n_{22}	n_{23}	n_{24}	n_{25}
	SD	n_{31}	n_{32}	n_{33}	n_{34}	n_{35}
	PD	n_{41}	n_{42}	n_{43}	n_{44}	n_{45}
	NE	n_{51}	n_{52}	n_{53}	n_{54}	n_{55}

$\sum_{i=1}^{6}(n_{ii})$ is the number of agreements on BOR between BICR and Investigator

$\sum_{i,j=1}^{6}(n_{ij})$ for i ≠ j is the number of disagreements on BOR between BICR and Investigator

N= $\sum_{i,j=1}^{6}(n_{ij})$

The following measures of concordance will be calculated for each treatment group:

- Concordance rate for BOR = $\sum_{i=1}^{6}(n_{ii})/N$
- Concordance rate for response = $[\sum_{i,j=1}^{2}(n_{ij}) + \sum_{i,j=3}^{6}(n_{ij})]/N$

Concordance rates are calculated for each treatment group and, for each metric, the difference in concordance between the experimental and control groups are used to evaluate potential bias. If the concordance is similar across the treatment groups then this suggests the absence of evaluation bias favoring a particular treatment group.

6.2. Secondary Endpoint(s)

6.2.1. Safety endpoints

Refer to Section 6.6.

6.2.2. Efficacy endpoints

The following analyses will be based on the FAS. Data will be analyzed by study phase and treatment group; treatment groups in the lead-in phase will also be pooled together. Assessment of response will be made using the Response Criteria for Malignant Lymphoma (Cheson et al., 2007) for the Lead-in Phase and the Lugano Classification (Cheson et al., 2014) for the expansion phase. Tumor-related endpoints will be analyzed based on investigator assessment for the Lead-in Phase, and separately based on BICR assessments and based on investigator assessment for the Expansion Phase.

6.2.2.1. Tumor shrinkage from baseline

Tumor shrinkage will be summarized as the percent change from baseline in the Sum of the Product of the Diameters (SPD) for Largest Dominant Masses per time point. It will be derived as:

- ((SPD at week XX – SPD at baseline)/ SPD at baseline) × 100

The maximum reduction in SPD from baseline will be derived across all the post-baseline assessments until documented disease progression, excluding assessments after start of subsequent anti-cancer therapy, as:

- Minimum of ((SPD at week XX – SPD at baseline)/ SPD at baseline) × 100

A waterfall plot of maximum percent reduction in the SPD for Largest Dominant Masses from baseline will be created by treatment group. These plots will display the best percentage change from baseline in SPD for each patient with measurable disease at baseline and at least one post-baseline assessment.

6.2.2.2. Disease control

Disease Control (DC) is defined as BOR of CR, PR or SD. DC rate (DCR) is the proportion of patients with DC.

DCR will be summarized by frequency counts and percentages.

6.2.2.3. Duration of response

Duration of Response (DR) is defined, for patients with OR, as the time from the first documentation of objective response (CR or PR) to the date of first documentation of PD or death due to any cause. If a patient has not had an event (PD or death), DR is censored at the date of last adequate tumor assessment. The censoring rules for DR are as described for PFS in Table 8.

DR (months) = [date of event or censoring– first date of OR +1]/30.4375

Kaplan-Meier estimates (product-limit estimates) will be presented by treatment group together with a summary of associated statistics including the median DR time with 2-sided 95% CIs. In particular, the DR rate at 3, 6, 12, and 18 months will be estimated with corresponding 2-sided 95% CIs. The CIs for the median will be calculated according to Brookmeyer and Crowley and the CIs for the survival function estimates at the time points defined above will be derived using the log-log transformation according to Kalbfleisch and Prentice (conftype=loglog default option in SAS Proc LIFETEST) with back transformation to a CI on the untransformed scale. The estimate of the standard error will be computed using Greenwood's formula.

DR will be displayed graphically and analyzed using Kaplan-Meier methodology. If the number of patients with OR is small, the Kaplan-Meier method may not provide reliable estimates. In this case, only descriptive statistics or listings will be provided.

6.2.2.4. Time to response

Time to response (TTR) is defined, for patients with OR, as the time from the 'start date' to the first documentation of objective response (CR or PR).

TTR (in months) = [first date of OR – 'start date' +1]/30.4375

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

6.2.2.5. Progression-free survival

Progression-Free Survival (PFS) is defined as the time from the 'start date' to the date of the first documentation of PD or death due to any cause, whichever occurs first.

PFS data will be censored on the date of the last adequate tumor assessment for patients who do not have an event (PD or death), for patients who start a new anti-cancer therapy prior to an event (see Section 5.2.6) or for patients with an event after 2 or more missing tumor assessments. Patients who do not have an adequate baseline tumor assessment or who do not have an adequate post-baseline tumor assessment will be censored on the 'start date' unless death occurred on or before the time of the second planned tumor assessment (i.e. ≤ 12 weeks after the 'start date') in which case the death will be considered an event.

In the lead-in phase antitumor activity will be assessed through radiological tumor assessments conducted at screening, 6 weeks, 12 weeks, and at 12-week intervals thereafter until PD regardless of initiation of subsequent anti-cancer therapy. In the expansion phase, antitumor activity will be assessed through radiological tumor assessments conducted at screening, 6 weeks, 12 weeks, and at 12-week intervals thereafter for one year from the start of study treatment, and every 24 weeks thereafter until PD by BICR assessment, regardless of the initiation of subsequent anti-cancer therapy.

The censoring and event date options to be considered for the PFS and DR analysis are presented in Table 8.

PFS (months) = [date of event or censoring– 'start date' +1]/30.4375

Table 8. Outcome and event dates for PFS and DR analyses

Scenario	Date of event/censoring	Outcome
No adequate baseline assessment	'start date' [a]	Censored [a]
PD or death - After at most one missing or inadequate post-baseline tumor assessment, OR - ≤ 12 weeks after the 'start date'	Date of PD or death	Event
PD or death - after 2 or more missing or inadequate post-baseline tumor assessments	Date of last adequate tumor assessment [b] documenting no PD before new anti-cancer therapy is given or missed tumor assessments	Censored
No PD and no death	Date of last adequate tumor assessment [b] documenting no PD before new anti-cancer therapy is given or missed tumor assessments	Censored
Treatment discontinuation due to 'Disease progression' without documented progression	Not applicable	Information is ignored. Outcome is derived based on documented progression only.
New anti-cancer therapy given	Date of last adequate tumor assessment [b] documenting no PD before new anti-cancer therapy is given or missed tumor assessments	Censored

[a] However if the patient dies ≤12 weeks after the 'start date' the death is an event with date on death date

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

Kaplan-Meier estimates (product-limit estimates) will be presented by treatment group together with a summary of associated statistics including the median PFS time with 2-sided 95% CIs. In particular, the PFS rate at 3, 6, 9, 12, and 18 months will be estimated with corresponding 2-sided 95% CIs. The CIs for the median will be calculated according to Brookmeyer and Crowley and the CIs for the survival function estimates at the time points defined above will be derived using the log-log transformation according to Kalbfleisch and Prentice (conftype=loglog default option in SAS Proc LIFETEST) with back transformation to a CI on the untransformed scale. The estimate of the standard error will be computed using Greenwood's formula.

Frequency (number and percentage) of patients with each event type (PD or death) and censoring reasons will be presented by treatment group. Reasons for censoring will be summarized according to the categories in Table 9 following the hierarchy shown.

Table 9. PFS Censoring Reasons and Hierarchy

Hierarchy	Condition	Censoring Reason
1	No adequate baseline assessment	No adequate baseline assessment
2	Start of new anti-cancer therapy.	Start of new anti-cancer therapy
3	Event after 2 or more missing or inadequate post-baseline tumor assessments/'start date'	Event after missing assessments [a]
4	No event and [withdrawal of consent date ≥ 'start date' OR End of study (EOS) = Subject refused further follow-up]	Withdrawal of consent
5	No event and lost to follow-up in any disposition page	Lost to follow-up
6	No event and [EOS present OR disposition page for any epoch after screening says patient will not continue into any subsequent phase of the study] and no adequate post-baseline tumor assessment	No adequate post-baseline tumor assessment
7	No event and none of the conditions in the prior hierarchy are met	Ongoing without an event

[a] 2 or more missing or inadequate post-baseline tumor assessments.

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

Time of Follow-Up for PFS

A plot will be generated to compare planned and actual relative day of tumor assessments by treatment group. A Kaplan-Meier plot for PFS follow-up duration will also be generated to assess the follow-up time in the treatment groups reversing the PFS censoring and event indicators. Kaplan-Meier estimates (product-limit estimates) will be presented by treatment group together with a summary of associated statistics including the median time of follow-up for PFS with 2-sided 95% CIs. In particular, the rate at 3, 6, 9, 12 and 18 months will be estimated with corresponding 2-sided 95% CIs.

6.2.2.6. Overall Survival

Overall survival (OS) is defined as the time from the 'start date' to the date of death due to any cause. Patients last known to be alive will be censored at date of last contact.

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

Kaplan-Meier estimates (product-limit estimates) will be presented by treatment group together with a summary of associated statistics including the median OS time with 2-sided 95% CIs. In particular, the OS rate at 3, 6, 9, 12, 18 and 24 months will be estimated with corresponding 2-sided 95% CIs. The CIs for the median will be calculated according to Brookmeyer and Crowley and the CIs for the survival function estimates at the time points defined above will be derived using the log-log transformation according to Kalbfleisch and Prentice (conftype=loglog default option in SAS Proc LIFETEST) with back transformation

to a CI on the untransformed scale. The estimate of the standard error will be computed using Greenwood's formula.

Frequency (number and percentage) of patients with an event (death) and censoring reasons will be presented by treatment group. Reasons for censoring will be summarized according to the categories in Table 10 following the hierarchy shown.

Table 10. OS Censoring Reasons and Hierarchy

Hierarchy	Condition	Censoring Reason
1	No event and [withdrawal of consent date ≥ 'start date' OR End of study (EOS) = Subject refused further follow-up]	Withdrawal of consent
2	No event and [lost to follow-up in any disposition page OR data cut-off date – last contact date > 14 weeks]	Lost to follow-up
3	No event and none of the conditions in the prior hierarchy are met	Alive

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

Time of Follow-Up for OS

A Kaplan-Meier plot for OS follow-up duration will also be generated to assess the follow-up time in the treatment groups reversing the OS censoring and event indicators. Kaplan-Meier estimates (product-limit estimates) will be presented by treatment group together with a summary of associated statistics including the median time of follow-up for OS with 2-sided 95% CIs. In particular, the rate at 3, 6, 9, 12, 18, and 24 months will be estimated with corresponding 2-sided 95% CIs.

6.2.3. Pharmacokinetic endpoints

The following pharmacokinetic analyses will be based on the PK analyses set. Data will be analyzed by study phase and treatment group; treatment groups in the lead-in phase will not be pooled together.

Additionally, in the expansion phase, PK data will be summarized separately by dose (i.e. PK data following 70 mg Q2W will be summarized separately from PK data following 500 mg Q2W).

C_{trough} and C_{max} for avelumab will be summarized descriptively (n, mean, SD, CV, median, minimum, maximum, geometric mean, its associated CV, and 95% CI) by treatment group, cycle, and day. Other standard parameters will be calculated as data permit, including, but not limited to: following Single Dose (SD): T_{max}, AUC_{last}, T_{last}, $AUC_{sd,\tau}$, $t_{1/2}$, CL, and V_z, and following multiple Dose (MD) - $T_{ss,max}$, $AUC_{ss,\tau}$, $t_{1/2}$, $C_{ss,av}$, CL, and V_{ss}. In the lead-in phase, parameters R_{ac} ($AUC_{ss,\tau}/AUC_{sd,\tau}$) and R_{ss} ($AUC_{ss,\tau}/AUC_{sd,inf}$) will also be determined, as data permit.

Dose normalized parameters (e.g., CDN- C_{max}, CDN- C_{trough}) will be reported as appropriate. The trough concentrations for avelumab will be plotted for each dose using a box whisker plot by cycle and day in order to assess the attainment of steady state.

Pharmacokinetic parameters for avelumab will be taken from observed values or derived from plasma concentration-time data as described in Section 3.2.3.

Presentation of pharmacokinetic data will include:

- Descriptive statistics (n, mean, SD, %CV, median, minimum, maximum) of plasma concentrations will be presented in tabular form by treatment group, dose level, cycle, day and nominal time. Additionally similar descriptive statistics will also be generated for dose-normalized avelumab pharmacokinetic parameters.

- Linear-linear and log-linear plots of mean and median plasma concentrations by nominal time for avelumab will be presented for PK sampling days by treatment group, cycle, and study day. Similar plots will be presented for each individual patient's concentrations. In the lead-in phase, patients who have undergone intrapatient dose reduction or escalation will be excluded from the median plasma concentration-time plots. In the expansion phase, mean and median concertation-time plots will be generated separately for data from patients who stay on the 70 mg Q2W dosing regimen, those who dose-escalate to 500 mg Q2W in Cycle 3 and those who dose escalate to 500 mg Q2W in Cycle 7.

- Pharmacokinetic parameters for avelumab will be listed and summarized by treatment group/dose level, cycle and study day using descriptive statistics (n, mean, SD, %CV, median, minimum, maximum, geometric mean and its associated %CV, and 95% CI). For T_{max}, the range (min, max) will also be provided. PK parameters with zero values will be excluded from the calculation of geometric means and its associated %CV. In the lead-in phase, if an intrapatient dose escalation or reduction occurs, dose-dependent PK parameters (AUC and C_{max}) for that patient may be dose-normalized when it is known that the drug exhibits linear PK within the dose range and other PK parameters will be reported as estimated; or may only be included in descriptive statistics and summary plots up to the time of the dose change. In addition, dose-normalized C_{max} and AUC parameters will be summarized (as described above) using data pooled across treatment groups in which different avelumab doses were administered. In the expansion phase, summary statistics will be presented separately by cycle and day for patients receiving 70 mg Q2W and those receiving 500 mg Q2W.

- Box plots for AUC and C_{max} for avelumab will be generated. Individual data points, the geometric mean and the median of the parameter in each treatment will be overlaid on the box plots. If a treatment group has limited evaluable PK data (n<4), matchstick plots showing changes in AUC and C_{max} for each drug in individual patients will then be generated. The geometric mean of the parameter in each treatment will be overlaid in the plots. In the lead-in phase, box plots for dose-normalized avelumab C_{max} and AUC parameters will be created using data pooled across treatment groups in which different avelumab doses were administered. Box plots for AUC and C_{max} from patients in the expansion phase will only be presented on data up to Cycle 3.

- C_{trough} and C_{max} for avelumab will be plotted for each treatment group using a box-whisker plot by cycle and day within cycle in order to assess the attainment of steady-state. For the lead-in phase, data from patients who undergo intrapatient dose escalation or reduction will be excluded from these plots. For the expansion cohort, C_{trough} data will be presented separately for patients receiving 70 mg Q2W and those receiving 500 mg Q2W.

6.2.4. Population pharmacokinetic endpoints

Pharmacokinetic and pharmacodynamic data from this study may be analyzed using modeling approaches and may also be pooled with data from other studies to investigate any association between avelumab exposure and biomarkers or significant safety/efficacy endpoints. The results of these analyses, if performed, may be reported separately.

6.2.5. Biomarker endpoints

The following biomarker analyses will be based on the biomarker analyses set. Data will be analyzed by study phase and treatment group; treatment groups in the lead-in phase will also be pooled together. In addition, biomarker endpoints will be presented pooling together the lead-in and expansion phases.

If a patient has more than one result at a visit for a specific biomarker analyte, then:

- For continuous data, the duplicate results will be averaged, and the average used in the analysis;
- For non-continuous data (eg, identified genes), the study team will select the record appropriate for analysis. A flag will be added to the data sets indicating which record was selected for analysis.

6.2.5.1. IHC Analysis of PD-L1, CD8 and FoxP3 Expression in Tumor Biopsy

- The absolute number and percent positive of cells expressing PD-L1, CD8 and FoxP3 in the tumor, in the stroma and the tumor/stroma ratio of these marker positive cells by IHC will be summarized using descriptive statistics .
- The ratio of cells expressing biologically relevant pairs of these markers in each region by IHC (for example, the ratio of CD8+/FoxP3+ cells in the tumor) will be summarized using descriptive statistics.
- Correlation of the absolute numbers and percent positive cells by region, the ratio of expression by region and the ratio of marker pairs at baseline with efficacy measures such as tumor shrinkage, time to response or PFS will be evaluated.
- Correlation of the change in absolute numbers and percent positive cells by region, the ratio of expression by region and the ratio of marker pairs from baseline to time of on study tumor biopsy with efficacy measures such as tumor shrinkage, time to response or PFS will be evaluated.

6.2.5.2. RNA Expression Profiling of Tumor Biopsy Tissue by RNA Sequencing

- Abundance of immune- and pathway-related RNA transcripts and enrichment of related gene sets at baseline as well as correlation of gene set enrichment with efficacy measures such as tumor shrinkage, time to response or PFS will be evaluated.
- Change in expression of immune- and pathway-related RNA transcripts and enrichment of related gene sets between baseline and time of on study tumor biopsy as well as correlation of change in gene expression and gene set enrichment with efficacy measures such as tumor shrinkage, time to response or PFS will be evaluated.

6.2.5.3. Relative Proportion of T Cell Subsets in the Peripheral Blood by Flow Cytometry

- Relative proportion of naïve, memory and regulatory T cell subsets at baseline will be summarized using descriptive statistics.
- Change in proportions of T cell subsets in response to dosing with avelumab will be summarized using descriptive statistics.
- Correlation of baseline and changes in relative proportions of T cell subsets with efficacy measures such as tumor shrinkage, time to response or PFS will be evaluated.

6.2.6. Endpoints for Immunogenicity Data of Avelumab

All analyses described below are performed by study phase and treatment group; treatment groups in the lead-in phase will also be pooled together. In addition, immunogenicity endpoints will be presented pooling together the lead-in and expansion phases.

In the lead-in phase, blood samples for avelumab immunogenicity testing will be collected pre-dose on Day 1 of Cycles 1, 2, 3, 4, 6 and 8, and every 12 weeks up to cycle 20. Additional samples for anti-avelumab antibodies will be collected at the End of Treatment visit. Follow up samples are only collected in patients who are ADA positive at end of treatment. In the dose-expansion phase, blood samples for avelumab immunogenicity testing will be collected pre-dose on Day 1 of Cycles 1, 2, 3, 4, 7 and 13, and at Day 30, 60 and 90 follow up visits. Additional samples for anti-avelumab antibodies will be collected at the End of Treatment visit.

Samples positive for ADA will be analyzed for titer and may be analyzed for nAb. As of the finalization of this SAP, the nAb assay is not yet available, therefore the analyses of nAb data described in the following sections will only be conducted contingent upon assay and data availability at the time of reporting.

Patients will be characterized into different ADA categories based on the criteria defined in Table 11.

Table 11. Patients Characterized Based on Anti-Drug Antibody Results (ADA Status)

Category	Definition	Subjects at Risk (Denominator for Incidence)
ADA never-positive	No positive ADA results at any time point; ADA-negative patients (titer < cutpoint)	Number of patients with at least one valid ADA result at any time point
ADA ever-positive	At least one positive ADA result at any time point; ADA-positive patients (titer ≥ cutpoint)	Number of patients with at least one valid ADA result at any time point
Baseline ADA positive	A positive ADA result at baseline	Number of patients with valid baseline ADA result
Treatment-boosted ADA	A positive ADA result at baseline and the titer ≥ 8×baseline titer at least once after treatment with avelumab	Number of patients with valid baseline ADA results and at least one valid post-baseline ADA result
Treatment-induced ADA	Patient is ADA-negative at baseline and has at least one positive post-baseline ADA result; or if patient does not have a baseline sample, the patient has at least one positive past-baseline ADA result	Number of patients with at least one valid post-baseline ADA result and without positive baseline ADA result (including missing, NR)
Transient ADA response	If patients with treatment-induced ADA have (a single positive ADA result or duration between first and last positive result <16 weeks) and ADA result at the last assessment is not positive.	Number of patients with at least one valid post-baseline ADA result and without positive baseline ADA result (including missing, NR)
Persistent ADA response	If patients with treatment-induced ADA have duration between first and last positive ADA result ≥16 weeks or a positive ADA result at the last assessment	Number of patients with at least one valid post-baseline ADA result and without positive baseline ADA result (including missing, NR)

ADA: anti-drug antibody, NR = not reportable.

Patients 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 12. Patients Characterized Based on Neutralizing Antibody Results (nAb Status)

Category	Definition	Subjects at Risk (Denominator for Incidence)
nAb never-positive	No positive nAb results at any time point	Number of patients with at least one valid ADA result at any time point
nAb ever-positive	At least one positive nAb result at any time point	Number of patients with at least one valid ADA result at any time point
Baseline nAb positive	A positive nAb result at baseline	Number of patients with valid baseline ADA result
Treatment-induced nAb	Patient is not nAb positive at baseline and has at least one positive post-baseline nAb result; or if patient does not have a baseline sample, the patient has at least one positive past-baseline ADA result	Number of patients with at least one valid post-baseline ADA result and without positive baseline nAb result (including missing, NR)
Transient nAb response	If patients with treatment-induced nAb have (a single positive nAb result or duration between first and last positive result <16 weeks) and nAb result at the last assessment is not positive.	Number of patients with at least one ADA valid post-baseline result and without positive baseline nAb result (including missing, NR)
Persistent nAb response	If patients with treatment-induced nAb have duration between first and last positive nAb result ≥16 weeks or a positive nAb result at the last assessment	Number of patients with at least one valid post-baseline ADA result and without positive baseline nAb result (including missing, NR)

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

The number and percentage of patients in each ADA and nAb category will be summarized.

6.2.6.1. Time to and Duration of ADA and nAb response

The ADA and nAb analyses described below will include patients with treatment-induced ADA or nAb, respectively.

Time (weeks) to ADA response is defined as:

(Date of first positive ADA result – date of first dose of avelumab + 1)/7.

Time to ADA response will be summarized using simple descriptive statistics (mean, SD, median, min, max. Q1, Q3).

Duration (weeks) of ADA response is defined as:

(Date of last positive ADA result – date of first positive ADA result + 1)/7.

Duration of ADA response will be censored if:

- the last ADA assessment is positive AND patient is ongoing treatment with avelumab, or

- the last ADA assessment is positive AND patient discontinued treatment with avelumab AND the last planned ADA assessment (End of Treatment visit) is after the cut-off date.

Time to nAb response and duration of nAb response are defined similarly based on first and last positive nAb result.

Kaplan-Meier estimates (product-limit estimates) will be presented together with a summary of associated statistics including the median ADA response time with 2-sided 95% CIs. ADA response rates at different timepoints will be estimated with corresponding 2-sided 95% CIs. The CIs for the median will be calculated according to Brookmeyer and Crowley (1982) and the CIs for the survival function estimates will be derived using the log-log transformation according to Kalbfleisch and Prentice (2011) (conftype=loglog default option in SAS Proc LIFETEST) with back transformation to a CI on the untransformed scale. The estimate of the standard error will be computed using Greenwood's formula.

Duration of ADA response will be displayed graphically and analyzed using Kaplan-Meier methodology. If the number of patients with ADA response is small, the Kaplan-Meier method may not provide reliable estimates. In this case, only descriptive statistics or listings will be provided

As data permit, the analyses described above will be repeated for patients with treatment-induced nAb.

6.2.6.2. ADA titer

For patients who are ADA ever positive, the maximum observed ADA titer for a patient will be summarized, overall and by ADA subcategories (baseline ADA positive, treatment-boosted ADA, treatment-induced ADA, transient ADA response, persistent ADA response) of patients having each discrete maximum titer value will be tabulated. The denominator to calculate the percentages will be the total number of patients in the associated ADA subcategory.

For patients with treatment-induced ADA, a cross tabulation of duration of ADA response and maximum ADA titer will be provided. The following categories for duration of ADA response will be used: ≤1, >1 to ≤3, >3 to ≤5, >5 to ≤7, >7 to ≤13, >13 to ≤16, >16 to ≤25, >25 weeks. In this categorization, the censoring in duration of ADA response is ignored.

6.2.6.3. Analysis of PK, safety and efficacy by immunogenicity status

The following ADA and nAb status will be used for the analyses described below.

ADA

- ADA ever-positive versus ADA never-positive

- ADA: treatment-induced ADA versus ADA never-positive or baseline ADA positive

nAb

- nAb ever-positive versus nAb never-positive
- nAb: treatment-induced nAb versus nAb never-positive or baseline nAb positive

Data listings will include immunogenicity data together with relevant PK, safety and efficacy data.

PK parameters and immunogenicity status

The following analyses will include patients in both the immunogenicity analysis set and in the PK parameter analysis set. The PK endpoints pertinent to the immunogenicity analyses are C_{trough} and C_{max}.

Blood samples for avelumab PK will be collected where noted in the Schedule of Activities section of the protocol.

C_{trough} and C_{max} will be summarized descriptively (n, mean, SD, CV, median, minimum, maximum, geometric mean, its associated CV, and 95% CI) by nominal time and ADA status. Linear-linear and log-linear plots of mean and median for C_{trough} and C_{max} over nominal time and by ADA status will be presented.

Among patients with treatment-induced ADA, analyses will be conducted to assess whether C_{trough} and C_{max} have any changes before and after the first positive ADA assessment. To be included in this analysis, patients must have the same PK parameter available both before and after the first positive ADA assessment. Relative PK day will be calculated as:

(PK assessment nominal day) – (first positive ADA assessment nominal day).

Nominal day is the protocol scheduled timing for an assessment. For example, if C_{trough} is collected on Day 1 of Cycle 2 and the first positive ADA result is observed on Day 1 of Cycle 3, then the relative PK day for this C_{trough} is -14 for 2-week cycles and -21 for 3-week cycles. Linear-linear and log-linear plots of mean and median for C_{trough} and C_{max} over relative PK day will be presented.

As data permit, the analyses described above will be repeated for nAb.

Safety and immunogenicity status

The following analyses will include patients in the immunogenicity analysis set.

The frequency (number and percentage) of patients with each of the following will be presented by ADA status.

- TEAEs, by SOC and PT
- TEAEs leading to dose reduction of avelumab, by SOC and PT

- TEAEs leading to discontinuation of avelumab, by SOC and PT
- Grade ≥ 3 TEAEs, by SOC and PT
- SAEs, by SOC and PT
- IRRs, by PT

For patients who had at least one IRR and have treatment-induced ADA, time related to first onset of an IRR (infusion 1, infusion 2, infusion 3, infusion 4 or later) will be summarized taking into account whether the IRR occurred on or after the first ADA positive assessment or whether the IRR occurred before the first ADA positive assessment.

As data permit, the analyses described above will be repeated for nAb.

Efficacy and immunogenicity status

For the ADA ever-positive patients, a listing will be prepared with patient ID, start and stop of avelumab treatment, date of first positive ADA result, time to ADA response, duration of ADA response, date of last ADA positive result, BOR, DR, PFS time or censoring time and reason for censoring, and OS time or censoring time and reason for censoring. If applicable, date of first positive nAb result, time to nAb response, duration of nAb response, date of last nAb positive result will also be presented. In the expansion phase, tumor-related endpoints will be presented based on BICR assessment and based on Investigator assessment.

For the ADA ever-positive patients, the percent change from baseline in target lesions as well as the first occurrence of a new lesion and patient off avelumab 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. In the dose-expansion phase, plot will be presented separately based on BICR assessment and based on Investigator assessment.

CCI

[redacted]

6.4. Subset Analyses

In the lead-in phase, OR and DR (if applicable) will be summarized by prior HSCT type (Allogeneic or Syngeneic vs. other) for each treatment group and for all treatment groups pooled together.

6.5. Baseline and Other Summaries and Analyses

6.5.1. Baseline summaries

The following analyses will be based on the FAS. Unless otherwise specified, data will be analyzed by study phase and treatment group; treatment groups in the lead-in phase will also be pooled together.

6.5.1.1. Demographic characteristics

Demographic characteristics and physical measurements will be summarized by study phase and treatment group using the following information from the 'Screening/Baseline Visit' eCRF pages. Demographic characteristics will also be presented pooling together treatment groups in the lead-in phase and the lead-in and expansion phases.

- Demographic characteristics
  - Gender: Male, Female
  - Race: White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, Other, Unknown
  - Ethnic origin: Hispanic/Latino (Yes/No)
  - Age (years): summary statistics
  - Age categories :
    - < 65 years, ≥ 65 years
    - < 65, 65-<75, 75-<85, ≥ 85 years
  - Pooled Geographical Region (as applicable):
    - North America
    - Europe
    - Asia
    - Rest of the World (Australasia, Latin America, Africa and/or Middle East will be included as additional pooled geographical regions if including > 10% of the overall randomized/treated population)
  - Geographic Region (as applicable):
    - North America
    - Latin America
    - Western Europe
    - Eastern Europe

    - Middle East
    - Australasia
    - Asia
    - Africa
  - Eastern Cooperative Oncology Group (ECOG) Performance Status: 0, 1, 2, 3, and 4
- Physical measurements
  - Height (cm)
  - Weight (kg)
  - Body Mass Index (BMI) (kg/m^2)

Center codes will be used for the determination of the patient's geographic region.

The listing of demographics and baseline characteristics will include the following information: patient identifier, treatment group, age, sex, race, ethnicity, height (cm), weight (kg), BMI (kg/m^2), and ECOG performance status.

6.5.1.2. Medical history

Medical history will be coded using the most current available version of Medical Dictionary for Regulatory Activities (MedDRA) and will be summarized from the 'Medical History' eCRF page. Medical history will be summarized as the numbers and percentages of patients by MedDRA preferred term (PT) as event category and MedDRA primary system organ class (SOC) as summary category. Each patient will be counted only once within each PT or SOC.

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

Prior stem cell transplant details will only be listed.

6.5.1.3. Disease characteristics

Information on disease characteristics collected on 'Primary Diagnosis', 'Substance Use' and 'Transplant Details' eCRF pages will be summarized. Summary statistics will be presented for the following.

From the 'Primary Diagnosis' eCRF page:

- Site of primary tumor
- Primary diagnosis (summarize all categories collected in the 'Primary Diagnosis' eCRF page)

- Time since initial diagnosis to 'start date' (months), defined as ('start date' – date of initial diagnosis)/30.4375
- Time since diagnosis of local/regional recurrence of disease (months), defined as ('start date' – date of diagnosis of local/regional recurrence of disease)/30.4375
- Hodgkin's Lymphoma stage and substage at initial diagnosis
- Current Hodgkin's Lymphoma stage and substage

From the 'Substance Use' eCRF page:

- Smoking history
  - Never smoker vs current vs former smoker
  - Smoking exposure (pack-years): 0, <20, 20-<40, ≥40 and summary statistics
  - Years since quitting: never smoker, current smoker, <5, 5-<10, ≥10 and summary statistics

  Specifications for computation:

  - Cigarette equivalents are calculated as follows: one cigar is regarded equivalent to 5 cigarettes and 1 pipe is regarded equivalent to 3 cigarettes
  - Duration of nicotine consumption [years]:

    (end of nicotine consumption – start of nicotine consumption + 1) / 365.25
  - Pack-years:
    - calculate cigarette equivalents per day using the conversion factors given above
    - convert to packs per day where 20 cigarettes are regarded as 1 pack
    - pack-years = packs per day × duration of nicotine consumption [years]

From the 'Transplant Details' eCRF page:

- Transplant Type (Allogeneic or Syngeneic, Autologous, No Prior Stem Cell Transplant, Unknown)
- Donor-Receptor Relatedness (Related, Unrelated, Unknown)
- HLA Compatibility (HLA-matched, HLA-unmatched, HLA-haploidentical, Unknown)
- Stem Cell Source (Bone marrow, Peripheral blood, Cord Blood, Unknown)
- Type of Conditioning (Myeloablative, Reduced-intensity, Unknown)

Listing of disease history will be provided with all relevant data (as collected on the 'Primary Diagnosis' and 'Substance Use' eCRF pages) and derived variables as above.

6.5.1.4. Prior anti-cancer therapies

The prior anti-cancer therapies are collected under the 'Prior Cancer Therapy', 'Prior Radiation Therapy' and 'Prior Surgery' eCRF pages.

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

- Patients with at least one type of prior anti-cancer therapy
- Patients with at least one prior anti-cancer drug therapy
- Patients with at least one prior anti-cancer radiotherapy
- Patients with at least one prior anti-cancer surgery

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

- At least one prior anti-cancer drug therapy
- Number of prior anti-cancer drug therapy regimens: missing, 1, 2, 3, ≥4
- Prior anti-cancer immune therapy (including PD-1, PD-L1, anti-CTLA4, others)
- Best response: CR, PR, SD, PD, Unknown, Not applicable. Best response is derived from the last treatment regimen.

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

Prior anti-cancer therapies will be included in the listings that follow with a flag to identify prior therapies. These will include the patient identification number, and all the relevant collected data-fields on the corresponding eCRF pages.

- Listing of anti-cancer drug therapies
- Listing of anti-cancer radiotherapy
- Listing of anti-cancer surgeries

6.5.1.5. Prior and concomitant immunosuppressive therapies

The prior and concomitant immunosuppressive therapies are collected under the 'Prior Immunosuppressive Therapy' and 'Concomitant Immunosuppressive Therapy' eCRF pages.

The immunosuppressive therapies will be summarized based on the number and percentage of patients by the drug class and preferred term. A patient will be counted only once within a

given drug class and within a given drug name, even if he/she received the same medication at different times. The summary will be sorted on decreasing frequency of drug class and decreasing frequency of drug name in a given drug class. In case of equal frequency regarding drug class (respectively drug name), alphabetical order will be used.

Immunosuppressive therapies will be included in the listing of immunosuppressive therapies with a flag to identify prior and concomitant therapies. These will include the patient identification number, and all the relevant collected data-fields on the corresponding eCRF pages.

6.5.2. Study conduct and patient disposition

The following analyses will be performed based on the FAS. Data will be analyzed by study phase and treatment group; treatment groups in the lead-in phase will also be pooled together.

6.5.2.1. Patient disposition

Lead-in Phase

The percentages below will be calculated based on the number of patients in the FAS.

- Total number of patients screened overall
- Number of patients who discontinued from the study prior to randomization overall and by the main reason for discontinuation
- Number and percentage of randomized patients in each of the analysis sets defined in Section 4
- Number and percentage of randomized patients with study drug ongoing
- Number and percentage of randomized patients who discontinued study drug overall and by the main reason for discontinuation of study drug
- Number and percentage of patients who entered follow-up
- Number and percentage of patients who discontinued follow-up overall and by the main reason for discontinuation
- Number and percentage of patients who entered long-term follow-up
- Number and percentage of patients who discontinued long-term follow-up overall and by the main reason for discontinuation.

The results of the randomization algorithm (according to IRT) will be summarized as follows:

- Number and percentage of randomized patients overall, by region (Europe, EEA (required by EudraCT), North America, Latin America, Middle East, Asia, Australasia, Africa), by country within region

- Number and percentage of randomized patients by center
- Cross tabulation: patients randomized (Cohort A/ Cohort B/ Cohort C/ Cohort D/ Cohort E/ not randomized) vs. patients treated (Cohort A/ Cohort B/ Cohort C/ Cohort D/ Cohort E/ not treated)

Expansion Phase

The percentages below will be calculated based on the number of patients in the FAS.

- Total number of patients screened overall
- Number of patients who discontinued from the study prior to treatment with study drug overall and by the main reason for discontinuation
- Number and percentage of treated patients in each of the analysis sets defined in Section 4
- Number and percentage of patients with study drug ongoing
- Number and percentage of patients who discontinued study drug overall and by the main reason for discontinuation of study drug
- Number and percentage of patients who entered follow-up
- Number and percentage of patients who discontinued follow-up overall and by the main reason for discontinuation
- Number and percentage of patients who entered long-term follow-up
- Number and percentage of patients who discontinued long-term follow-up overall and by the main reason for discontinuation

In addition the following will be summarized:

- Number and percentage of treated patients overall, by region (Europe, EEA (required by EudraCT), North America, Latin America, Middle East, Asia, Australasia, Africa), by country within region
- Number and percentage of treated patients by center

6.5.2.2. Protocol deviations

All protocol violations that impact the safety of the patients and/or the conduct of the study and/or its evaluation will be reported. These include:

- Patients who are dosed on the study despite not satisfying the inclusion criteria
- Patients who develop withdrawal criteria whilst on the study but are not withdrawn

- Patients who receive the wrong treatment or an incorrect dose
- Patients who receive an excluded concomitant medication
- Deviations from GCP.

The identification of these and other CSR-reportable deviations will be based on the inclusion/exclusion criteria or other criteria presented in the protocol.

6.5.3. Study treatment compliance and exposure

The following analyses will be based on the safety analysis set by treatment group.

For Cycle X, actual cycle start date for each patient is

- the earliest start date of dosing in the Cycle X day 1 visit eCRF exposure page, if the patient received study treatment on that visit (i.e., any study drug with dose>0 at that visit)
- the first day of assessments in the Cycle X day 1 visit, if the patient did not receive study treatment on that visit (i.e., all study drugs had dose=0 at that visit). Use start date in the exposure page if available; if start date is not available then use date of collection of vital signs on Cycle X day 1 visit.

Actual cycle end date for each patient is,

- for all cycles X except the last cycle, actual cycle end date = actual cycle (X+1) start date – 1 day;
- for the last cycle, actual cycle end date = actual cycle start date + d (in days) – 1 day

  where d=14 for Q2W schedule, and d=21 for Q3W schedule.

Cycle duration (weeks) = (actual cycle end date – actual cycle start date + 1)/7

When summarizing exposure for each study drug, only cycles from first dose of study treatment until the last cycle with non-zero dose of study drug should be included.

Exposure may be summarized (overall) as dose received (cumulative dose, actual dose intensity) and as dose received relative to intended dose (relative dose intensity [RDI]).

The derivations below are provided assuming 1 cycle = 3 weeks for Cohort C and 1 cycle = 2 weeks for all the other cohorts, and for the following study drugs (administered alone or in combination):

- Cohort A: Avelumab administered as a 1-hour IV infusion at a dose of 70 mg once every 2 weeks in 2-week cycles.
- Cohort B: Avelumab administered as a 1-hour IV infusion at a dose of 350 mg once every 2 weeks in 2-week cycles.

- Cohort C: Avelumab administered as a 1-hour IV infusion at a dose of 500 mg once every 3 weeks in 3-week cycles.

- Cohort D: Avelumab administered as a 1-hour IV infusion at a dose of 500 mg once every 2 weeks in 2-week cycles.

- Cohort E: Avelumab administered as a 1-hour IV infusion at a dose of 10 mg/kg once every 2 weeks in 2-week cycles.

- Cohort F: Avelumab administered as a 1-hour IV infusion at a dose of 70 mg once every 2 weeks with the possibility of intra-patient dose escalation to avelumab administered as a 1-hour IV infusion at a dose of 500 mg once every 2 weeks in 2-week cycles (see Section 2.2).

6.5.3.1. Exposure to avelumab

The dose level for avelumab is calculated as actual dose administered/weight (mg/kg) for Cohort E, and as actual dose administered (mg) for all the other cohorts. The last available weight of the patient on or prior to the day of dosing will be used.

Intended duration of treatment with avelumab (weeks) =

(end date–date of first dose of study drug +1)/7,

where end date = start date of last cycle with non-zero dose of study drug + d – 1

and where d=14 for Q2W schedule, and d=21 for Q3W schedule.

Duration of exposure to avelumab (weeks) =

(last dose date of avelumab – first dose date of avelumab + d)/7

where d=14 for Q2W schedule, and d=21 for Q3W schedule.

Cumulative dose in a cycle or overall is the sum of the actual doses of avelumab received in a cycle or overall, respectively.

Actual Dose Intensity (DI)

For Cohort E:

- Overall actual DI (mg/kg/2-week cycle) = [overall cumulative dose (mg/kg)] / [intended duration of treatment with avelumab (weeks)/2].

For Cohort C:

- Overall actual DI (mg/3-week cycle) = [overall cumulative dose (mg)] / [intended duration of treatment with avelumab (weeks)/3].

For all the other cohorts:

- Overall actual DI (mg/2-week cycle) = [overall cumulative dose (mg)] / [intended duration of treatment with avelumab (weeks)/2].

Relative Dose Intensity (RDI)

For Cohort E:

- Intended DI (mg/kg/2-week cycle) = [intended cumulative dose per cycle] / [intended number of 2-weeks in a cycle] = [10 (mg/kg)] / [1 (2-week cycle)] = 10 (mg/kg/2-week cycle)
- Overall RDI (%) = 100 × [overall actual DI] / [intended DI]
  = 100 × [overall actual DI] / [10 (mg/kg/2-week cycle)]

For Cohort C:

- Intended DI (mg/3-week cycle) = [intended cumulative dose per cycle] / [intended number of 3-weeks in a cycle] = [500 mg] / [1 (3-week cycle)] = 500 (mg/3-week cycle)
- Overall RDI (%) = 100 × [overall actual DI] / [intended DI]
  = 100 × [overall actual DI] / [500 (mg/3-week cycle)]

For Cohort F:

- For patients that do not undergo dose escalation, then
  - Intended DI (mg/2-week cycle) = [intended cumulative dose per cycle] / [intended number of 2-weeks in a cycle] = [70 mg] / [1 (2-week cycle)] = 70 (mg/2-week cycle)
  - Overall RDI (%) = 100 × [overall actual DI] / [intended DI]
    = 100 × [overall actual DI] / [70 (mg/2-week cycle)]
- For patients that undergo dose escalation, then
  - Intended DI (mg/2-week cycle) = $[(n_1 \times 70\ (mg) + n_2 \times 500\ (mg))/(n_1+n_2)]$ / [intended number of 2-weeks in a cycle] = $[n_1 \times 70\ (mg) + n_2 \times 500\ (mg)]$ / [1 (2-week cycle)] = $[(n_1 \times 70 + n_2 \times 500)/(n_1+n_2)]$ (mg /2-week cycle)
  - Overall RDI (%) = 100 × [overall actual DI] / [intended DI]

  where,

  - n_1 = (date of first administration of the higher dose – 1)/14
  - n_2 = intended duration of treatment with avelumab (weeks)/2 – n1
  - n_1+n_2= intended duration of treatment with avelumab (weeks)/2

For all the other cohorts:

- Intended DI (mg/2-week cycle) = [intended cumulative dose per cycle] / [intended number of 2-weeks in a cycle] = [k mg] / [1 (2-week cycle)] = k (mg/2-week cycle)
- Overall RDI (%) = 100 × [overall actual DI] / [intended DI]
  = 100 × [overall actual DI] / [k (mg/2-week cycle)]

where k=70 for Cohort A, 350 for Cohort B, 500 for Cohort D.

6.5.3.2. Dose reductions

Dose reduction is defined as actual non-zero dose < 90% of the planned dose.

The number and percentage of patients with at least one dose reduction as well as a breakdown of the number of dose reductions (1, 2, 3, ≥4) will be summarized.

6.5.3.3. Dose delays

Dose Delay is the difference between the actual time between two consecutive non-zero doses and the planned time between the same two consecutive non-zero doses.

Dose Delay for Dose x (days) = Date of Dose x – Date of Dose (x-1) – d.

where d=14 for Q2W schedule, and d=21 for Q3W schedule.

Dose delays will be grouped into the following categories:

- No delay
- 1-3 days delay
- 4-6 days delay
- 7 or more days delay

For example, for avelumab administered on a 2-week schedule, if one patient receives avelumab on Day 1, then the next avelumab administration date will be on Day 15; however, if the patient receives avelumab at Day 16 or 17, this is considered as 1-3 days delay.

No delay and 1-3 days delay will also be summarized together.

The number and percentage of patients with delayed study drug administration and maximum length of delay, i.e., the worst case of delay if patients have multiple dose delays will be summarized.

6.5.3.4. Infusion rate reductions

The number and percentage of patients with at least one infusion rate reduction of ≥50% compared to the first infusion rate reported in the eCRF as well as the frequency of patients with 1, 2, 3 or ≥4 infusion rate reductions of ≥50% will be summarized.

6.5.3.5. Infusion interruptions

An infusion interruption is defined as an infusion that is stopped and re-started on the same day (i.e., for a visit more than one infusion start time and infusion end time are recorded).

The number and percentage of patients with at least one infusion interruption as well as the frequency of patients with 1, 2, 3, or ≥4 infusion interruptions will be summarized.

6.5.4. Concomitant medications and non-drug treatments

The following analyses will be based on the safety analysis set. Data will be analyzed by study phase and treatment group; treatment groups in the lead-in phase will also be pooled together.

Concomitant medications are medications, other than study medications, which started prior to first dose date of study treatment and continued on on-treatment period as well as those started during the on-treatment period. **Prior medications** are medications, other than study medications and pre-medications for study drug, which are started before the first dose of study treatment.

Prior and concomitant medications will be summarized from the 'General Concomitant Medications' eCRF page. Pre-medications for study drug will also be summarized separately from the 'Pre-Medication Treatment' eCRF page.

Summary of prior medications, summary of concomitant medications and summary of pre-medications will include the number and percentage of patients by Anatomical Therapeutic Chemical (ATC) Classification level 2 and preferred term. A patient 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 or concomitant medication is classified into multiple ATC classes, the medication will be summarized separately under each of these ATC classes. The summary tables will be sorted on decreasing frequency of drug class and decreasing frequency of drug name in a given drug class. In case of equal frequency regarding drug class (respectively drug name), alphabetical order will be used. In case any specific medication does not have ATC classification level 2 coded term, it will be summarized under 'Unavailable ATC classification' category.

A listing of prior medications and a listing of concomitant medications will be created with the relevant information collected on the 'General Concomitant Medications' eCRF page. A listing of pre-medications will be created with the relevant information collected on the 'Pre-Medication Treatment' eCRF page.

All concurrent procedures, which were undertaken any time during the on-treatment period, will be listed according to the eCRF page 'General Non-drug Treatments'.

A listing of concurrent procedures will be created with the relevant information collected on the 'General Non-drug Treatments' eCRF page.

6.5.5. Subsequent anti-cancer therapies

The following analyses will be based on the FAS. Data will be analyzed by study phase and treatment group; treatment groups in the lead-in phase will also be pooled together.

Anti-cancer treatment will be provided in a data listing with data retrieved from 'Follow-up Cancer Therapy', 'Concomitant Radiation Therapy', 'Follow-up Radiation Therapy', 'Concomitant Surgery', and 'Follow-up Surgery' eCRF pages.

Number and percentage of patients with any anti-cancer therapy after discontinuation will be tabulated overall and by type of therapy based on the data collected from the 'Follow-up Cancer Therapy', 'Follow-up Radiation Therapy' and 'Follow-up Surgery' eCRF pages.

6.6. Safety Summaries and Analyses

The Safety Analysis Set will be the primary population for safety evaluations. Summaries of AEs and other safety parameters will be based on the safety analysis set. Data will be analyzed by study phase and treatment group; treatment groups in the lead-in phase will also be pooled together.

All summaries described below by SOC and PT may further be presented by PT in decreasing frequency based on the frequencies observed in the Cohort A for the Lead-in Phase and in the Cohort F for the Expansion phase.

6.6.1. Adverse events

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 3.5.1.

All analyses described in the following sections will be based on TEAEs (started during the on-treatment period) if not otherwise specified. The AE listings will include all AEs (whether treatment-emergent or not). AEs outside the on-treatment period will be flagged in the listings.

- **Related Adverse Events:** adverse events with relationship to study treatment (as recorded on the AE eCRF page, 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 Dose Reduction:** adverse events leading to dose reduction of study treatment (as recorded on the AE eCRF page, Action taken with study treatment = Dose reduced).
- **Adverse Events Leading to Interruption of Study Treatment:** adverse events leading to interruption of study treatment (as recorded on the AE eCRF page, Action taken with study treatment = Drug interrupted). The eCRF does not allow for a clear separation

between interruption of an infusion and delays of administration for a parenteral drug as both are recorded using the same term on the eCRF ("Drug interrupted"). IRRs will be excluded in the analysis of AEs leading to Drug Interruption in case they only led to an interruption of the infusion.

- **Adverse Events Leading to Permanent 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, as well as AEs of Grade 5).
- **Immune-related Adverse Events (irAE):** irAEs (as identified according to the methodology outlined in Appendix 1 for a pre-specified search list of MedDRA PTs, documented in the Safety Review Plan and finalized for analysis of the current studies data prior to DB lock)
- **Infusion-related Reactions (IRR)**: IRRs (as identified according to the methodology outlined in Appendix 2 for a pre-specified search list of MedDRA PTs documented in the Safety Review Plan and finalized for analysis of the current studies data prior to DB lock.

Unless otherwise specified, AEs will be summarized by number and percentage of patients with the AE in the category of interest as described above, by treatment group, primary SOC and PT in decreasing frequency based on the frequencies observed for the Cohort A for the Lead-in Phase and in the Cohort F for the Expansion phase.

Each patient will be counted only once within each SOC or PT. If a patient experiences more than one AE within a SOC or PT for the same summary period, only the AE with the strongest relationship or the worst severity, as appropriate, will be included in the summaries of relationship and severity.

6.6.1.1. All adverse events

Adverse events will be summarized by worst severity (according to NCI-CTCAE version 4.03) per patient, using the latest version of MedDRA preferred term (PT) as event category and MedDRA primary system organ class (SOC) body term as Body System category.

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

The following tables will be created:

- The overall summary of AEs table will include the frequency (number and percentage) of patients with each of the following by treatment group:
  - TEAEs
  - TEAEs, Grade ≥ 3

  - Related TEAEs
  - Related TEAEs, Grade ≥ 3
  - TEAEs leading to dose reduction of avelumab
  - TEAEs leading to interruption of avelumab
  - TEAEs leading to discontinuation of avelumab
  - Related TEAEs leading to discontinuation of avelumab
  - Serious TEAEs
  - Related Serious TEAEs
  - TEAEs leading to death
  - Related TEAEs leading to death
  - irAEs
  - IRRs
- TEAEs by SOC and PT and worst grade
- Related TEAEs by SOC and PT and worst grade
- TEAEs leading to death by SOC and PT
- Related TEAEs leading to death by SOC and PT
- TEAEs Excluding SAEs, with frequency ≥ 5% in any treatment group by SOC and PT
- Graft Versus Host Disease (GVHD) by SOC and PT. Information recorded in the "Adverse Events" and in the "Graft Versus Host Disease Assessment" eCRF will be combined in the summary tables and listings for GVHD.

6.6.1.2. Adverse events leading to dose reduction

The frequency (number and percentage) of patients with each of the following will be presented for TEAEs leading to dose reduction of each study drug by treatment group:

- TEAEs leading to dose reduction of avelumab by SOC and PT

The listing of all AEs leading to dose reduction will also be provided with the relevant information.

6.6.1.3. Adverse events leading to interruption of study treatment

The eCRF does not allow for a clear separation between interruption of an infusion and delays of administration for a parenteral drug as both are recorded using the same term on the eCRF ("Drug interrupted"). IRRs will be excluded in the analysis of AEs leading to Drug Interruption in case they only led to an interruption of the infusion (ie, did not lead to a dose reduction or a dose delay).

As such, AEs leading to interruption will be defined as AEs identified in the AE eCRF page with an action taken with study treatment of 'drug interrupted' excluding

- IRRs that occurred on the day of infusion with ≥90% of the planned dose given (ie IRRs that did not lead to a dose reduction) and subsequent administration of study drug had no delay (as defined in Section 6.5.3.3). These IRRs will be considered as IRRs leading to interruption of infusion.
- IRRs occurring on the day after infusion and subsequent dose administration had no delay (as defined in Section 6.5.3.3).

The frequency (number and percentage) of patients with each of the following will be presented for TEAEs leading to interruption of each study drug by treatment group:

- TEAEs leading to interruption of avelumab by SOC and PT

The listing of all AEs leading to interruption of study treatment will also be provided with the relevant information.

In addition, the frequency (number and percentage) of patients with each of the following will be presented for TEAEs leading to interruption of each study drug by treatment group:

- TEAEs leading to both interruption and dose reduction of avelumab by SOC and PT

This summary will take into account PTs with both actions as defined in Section 6.6.1, even though the actions may be captured for different PT records (ie, different onset for the PT with action "drug interrupted" and the PT with action "dose reduced".

6.6.1.4. Adverse events leading to discontinuation of study treatment

The frequency (number and percentage) of patients with each of the following will be presented for TEAEs leading to permanent discontinuation of study treatment, by treatment group:

- TEAEs leading to discontinuation of avelumab by SOC and PT
- Related TEAEs leading to discontinuation of avelumab by SOC and PT.

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

6.6.2. Deaths

The frequency (number and percentage) of patients in the safety analysis set who died and who died within 30 days after last dose of study treatment as well as the reason for death, will be tabulated based on information from the 'Notice of Death' and 'Survival Follow-Up' eCRFs, by treatment group.

- All deaths
- Deaths within 30 days after last dose of study treatment

- Reason for Death
  - Disease progression
  - Study treatment toxicity
  - AE not related to study treatment
  - Unknown
  - Other.

In addition, date and cause of death will be provided in individual patient data listing together with selected dosing information (study treatment received, date of first / last administration, dose) and will include the following information:

- AEs with fatal outcome (list preferred terms of AEs with outcome=Fatal, as well as AEs of Grade 5),
- Flag for death within 30 days of last dose of study treatment.

6.6.3. Serious adverse events

The frequency (number and percentage) of patients with each of the following will be presented for treatment-emergent SAEs by treatment group:

- SAEs by SOC and PT
- Related SAEs by SOC and PT

The listings of all SAEs will also be provided with the relevant information with a flag for SAEs with onset outside of the on-treatment period.

6.6.4. Other significant adverse events

The frequency (number and percentage) of patients with each of the following will be presented for irAEs, by treatment group:

- irAEs leading to death, by Cluster and PT
- irAEs, by Cluster and PT
- irAEs, Grade ≥ 3, by Cluster and PT
- irAEs leading to discontinuation of avelumab, by Cluster and PT
- Serious irAEs, by Cluster and PT

The listing of all irAEs will also be provided with the relevant information with a flag for irAEs with onset outside of the on-treatment period.

The frequency (number and percentage) of patients with each of the following will be presented for IRRs, by treatment group:

- IRRs leading to death, by PT
- IRRs, by PT
- IRRs, Grade ≥ 3, by PT
- IRRs leading to discontinuation of avelumab, by PT
- Serious IRRs, by PT
- Time related to first onset of an IRR (infusion 1, infusion 2, infusion 3, infusion 4 or later)

The listing of all IRRs will also be provided with the relevant information with a flag for IRRs with onset outside of the on-treatment period.

6.6.5. Laboratory data

6.6.5.1. Hematology and chemistry parameters

Laboratory results will be classified according to the NCI-CTCAE criteria version 4.03. Non-numerical qualifiers (with the exception of fasting flags) will not be taken into consideration in the derivation of CTCAE criteria (e.g., hypokalemia Grade 1 and Grade 2 are only distinguished by a non-numerical qualifier and therefore Grade 2 will not be derived). Additional laboratory results that are not part of NCI-CTCAE will be presented according to the categories: below normal limit, within normal limits and above normal limit (according to the laboratory normal ranges).

Quantitative data will be summarized using simple descriptive statistics (mean, SD, median, Q1, Q3, minimum, and maximum) of actual values and changes from baseline for each nominal visit over time (unscheduled measurements would therefore not be included in these summaries as described in Section 5.2.9). End of Treatment visit laboratory results will be summarized separately. The changes computed will be the differences from baseline. Qualitative data based on reference ranges will be described according to the categories (i.e., Low, Normal, High).

Abnormalities classified according to NCI-CTCAE toxicity grading version 4.03 will be described using the worst grade. For those parameters which are graded with two toxicities such as potassium (hypokalemia/hyperkalemia), the toxicities will be summarized separately. Low direction toxicity (e.g., hypokalemia) grades at baseline and post baseline will be set to 0 when the variables are derived for summarizing high direction toxicity (e.g., hyperkalemia), and vice versa.

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

$$\text{Derived differential absolute count} = (\text{WBC count}) \times (\text{Differential \%value} / 100)$$

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

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

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

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

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

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

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

- Concurrent (ALT or AST) ≥ 3×ULN and TBILI ≥ 2×ULN and (ALP ≤ 2×ULN or missing)

Concurrent measurements are those occurring on the same date.

Categories will be cumulative, i.e., a patient 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 counts and percentages.

An evaluation of Drug-Induced Serious Hepatotoxicity (eDISH) plot will also be created, with different symbols for different treatment groups, by graphically displaying

- 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) including reference lines at AST=3×ULN and total bilirubin =2×ULN .

In addition, a listing of all TBILI, ALT, AST and ALP values for patients with a post-baseline TBILI ≥ 2×ULN, ALT≥ 3×ULN or AST ≥ 3×ULN will be provided.

Parameters with NCI-CTC grades available:

The laboratory toxicities will be tabulated using descriptive statistics (number of patients and percentages) during the on-treatment period. The denominator to calculate percentages for each laboratory parameter is the number of patients evaluable for CTCAE grading (i.e. those patients for whom a Grade 0, 1, 2, 3 or 4 can be derived).

- The summary of laboratory parameters by CTCAE grade table will include number and percentage of patients with Grade 1, 2, 3, 4, Grade 3/4 and any grade (Grades 1-4), laboratory abnormalities during the on-treatment period.
- The shift table will summarize baseline CTCAE grade versus the worst on-treatment CTCAE grade. The highest CTCAE grade during the on-treatment period is considered as the worst grade for the summary.

The above analyses apply to hematology and chemistry evaluations which can be graded per CTCAE, i.e.:

- Hematology:

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

- Serum Chemistry:

  Albumin (hypoalbuminemia), Alkaline Phosphatase (alkaline phosphatase increased), Alanine Aminotransferase (ALT) (ALT increased), Amylase (serum amylase increased), Aspartate Aminotransferase (AST) (AST increased), Total Bilirubin (blood bilirubin

increased, Cholesterol (cholesterol high), Creatinine (creatinine increased), Creatine Kinase (CPK increased), Potassium (hypokalemia/hyperkalemia),Sodium (hyponatremia/ hypernatremia), Magnesium (hypomagnesemia/hypermagnesemia), Calcium (hypocalcemia/ hypercalcemia), Glucose (hypoglycemia/hyperglycemia), Gamma Glutamyl Transferase (GGT) (GGT increased), Lipase (lipase increased), Phosphates (hypophosphatemia), Triglycerides (hypertriglyceridemia).

Parameters with NCI-CTC grades not available:

Hematology and chemistry evaluations which cannot be graded per CTCAE criteria will be summarized as frequency (number and percentage) of patients with:

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

In this study, these apply to the following parameters:

- Hematology: Absolute Monocytes, Absolute Eosinophils, Absolute Basophils.
- Serum Chemistry: Chloride, Total Urea, Uric Acid, Total Protein, C-Reactive Protein, Lactate Dehydrogenase (LDH).

6.6.5.2. Other laboratory parameters

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

- Coagulation: activated partial thromboplastin time (aPTT) and prothrombin time (INR).
- Urinalysis: all urinalysis parameters.
- Other parameters: hormone, and immunology parameters.
- Pregnancy test.

The listings of laboratory results will be provided for all laboratory parameters. The listings will be sorted by parameters and assessment dates or visits for each patient. Laboratory values that are outside the normal range will also be flagged in the data listings, along with corresponding normal ranges. A listing of CTCAE grading will also be generated for those laboratory tests.

In addition, listings of abnormal values will be provided for hematology, chemistry, urinalysis, coagulation parameters. If there is at least one abnormal assessment for any parameter, all the data for that laboratory parameter will be included into the listing.

For all tests not mentioned above but present in the clinical data, a listing of patients with at least one result for the relevant test will be provided.

6.6.6. Vital signs

Both in the Lead-in and in the Expansion phases, weight for the purposes of dose calculation will be recorded at screening and within 3 days pre-dose Day 1 of each cycle. Weight will also be collected at End of Treatment. Height will be measured at screening only. In the Expansion phase only, weight will be recorded also on Day 7 of Cycles 1 through 3 for all patients, on Day 7 of Cycles 4 thought 6 for patients who escalate to 500 mg Q2W at Cycle 4, and on Day 7 of Cycles 7 through 9 for patients who escalate to 500 mg Q2W at Cycle 7.

Vital sign summaries will include all vital sign assessments from the on-treatment period. All vital sign assessments will be listed, and those collected outside the on-treatment period will be flagged in the listing.

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

6.6.7. Electrocardiogram

ECG summaries will include all ECG assessments from the on-treatment period. All ECG assessments will be listed, and those collected outside the on-treatment period will be flagged in the listing. QTcB and QTcF will be derived based on RR and QT (see below). The average of the replicate measurements should be determined after the derivation of the individual parameter at each time point.

Selecting Primary QT Correction for Heart Rate

The analysis of QT data is complicated by the fact that the QT interval is highly correlated with heart rate. Because of this correlation, formulas are routinely used to obtain a corrected value, denoted QTc, which is independent of heart rate. This QTc interval is intended to represent the QT interval at a standardized heart rate. Several correction formulas have been proposed in the literature. For this analysis we will use some of those methods of correction, as described below. The QT interval corrected for heart rate by the Bazett's formula, QTcB, is defined as

$$QTcB = \frac{QT}{\sqrt{RR}},$$

the QT interval corrected for heart rate by the Fridericia's formula, QTcF, is defined as

$$QTcF = \frac{QT}{\sqrt[3]{RR}},$$

where RR represents the RR interval of the ECG, in seconds, and can be estimated as 60/Heart Rate.

Although Bazett's correction is the historical standard, it does not perform well when heart rate fluctuates. Fridericia's formula may perform better under these conditions. If QTcB and

QTcF methods do not adequately correct for HR and there are a sufficient number of patients (e.g. >30) with baseline ECGs, an alternate correction to achieve the goal of getting uncorrelated QTc and RR is based on a linear regression methods which yields, theoretically, uncorrelated QTc and RR.

Linear regression method:

- Fit a model QT = a + b × RR to baseline data
- Use the estimated slope, $\hat{b}$, to correct QT
- Corrected QT for heart rate will be computed as follows:

$$QTcP = QT + \hat{b} \times (1 - RR)$$

Data will be summarized using QTcF and QTcB. However, if these are not appropriate for the data set due to an observed large correlation between corrected QT and HR using the baseline assessments, the results will also be summarized using QTcP.

ECG Summaries

The following analyses will be performed for each applicable ECG parameters (RR, PR, QRS, QT, ventricular rate -denoted as HR in what follows-, and QTc) by treatment group, during the on-treatment period. The denominator to calculate percentages for each category is the number of patients evaluable for the category.

- Pearson correlation between QT and HR, QTc (QTcB, QTcF and, if applicable, QTcP) and HR using individual (non-averaged) baseline assessments
- For each of the ECG parameters (HR, and QT, QTc, QRS, PR intervals), descriptive statistics at baseline, at each post-baseline time point and changes from baseline at each post-baseline time point
- Frequency (number and percentage) of patients with notable ECG values according to the following categories:
  - QT/QTc increase from baseline >30 ms, >60 ms
  - QT/QTc > 450 ms, > 480 ms, > 500 ms
  - HR ≤ 50 bpm and decrease from baseline ≥ 20 bpm
  - HR ≥ 120 bpm and increase from baseline ≥ 20 bpm
  - PR ≥ 220 ms and increase from baseline ≥ 20 ms
  - QRS ≥ 120 ms

Patients with notable ECG interval values and qualitative ECG abnormalities will be listed for each patient and time point and the corresponding notable values and abnormality findings will be included in the listings.

Unscheduled ECG measurements will not be used in computing the descriptive statistics for change from baseline at each post-baseline time point. However, they will be used in the analysis of notable ECG changes and the shift table analysis of notable QT parameters.

6.6.8. Physical examination

Number and percentage of patients with abnormal findings in physical examination will be summarized by body system.

6.6.9. ECOG performance status

The ECOG shift from baseline to highest score during the on-treatment period will be summarized by treatment group. ECOG performance status with shift from ECOG=0 or 1 to ECOG 2 or higher will also be presented in a data listing.

7. INTERIM ANALYSES

There is no formal interim analysis planned for this study.

7.1. Introduction

Not Applicable.

7.2. Interim Analyses and Summaries

Not Applicable.

8. REFERENCES

1. Brookmeyer R, Crowley JJ. A confidence interval for the median survival time. Biometrics. 38: 29-41, 1982.

2. Cheson BD, Pfistner B, Juweid ME, Gascoyne RD, Specht et al. Revised Response Criteria for Malignant Lymphoma. J Clin Oncol. 2007 25:579-586.

3. Cheson DB, Fisher RI, Barrington SF, Cavalli F, Schwartz LH et al. Recommendations of Initial Evaluation, Staging, and Response Assessment of Hodgkin and Non-Hodgkin Lymphoma: The Lugano Classification. J Clin Oncol. 2014 32:3059-3067.

4. Collet D (1994). Modelling survival data in medical research. London, Chapman & Hall

5. Clopper CJ, Pearson ES (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika; 26, 404-413.

6. Kalbfleisch JD, Prentice, RL. Statistical Analysis of Failure Time Data, 2nd Edition. Hoboken, Wiley Interscience

7. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc. 53: 457-81, 1958.

9. APPENDICES

Appendix 1. Immune-Related Adverse Events

The MedDRA PTs and clusters for irAEs are defined in the Safety Review Plan (SRP) for avelumab.

Immune-related AEs (irAEs) will be programmatically identified as outlined in Table 13. This case definition is hierarchical, i.e., each step is only checked for patients and events that have already met the prior step.

Table 13. Case Definition for irAEs

Step	Selection Criteria	Additional Notes
1	Event selected based on a list of pre-specified MedDRA PTs within clusters. These are included in the SRP as Tier1 events (Immune-mediated xxxx). If AE matches the list then it is in for the next step	
2	AE onset during 1st study drug administration or anytime thereafter through 90 days after last dose of study treatment.	This is regardless of start of new anti-cancer drug therapy and regardless of TEAE classifications
3	Answer in the AE eCRF page to 'Was another treatment given because of the occurrence of the event' is 'YES'	
4	AE treated with corticosteroids or other immunosuppressant therapy. For endocrinopathies only: AE required hormone replacement	Look in the conmed pages for AE identifiers that match the AEs from Step 3. For each of such AEs if A) OR B) below are met then the AE is in for the next step A) conmed ATC code is in (H02A, H02B, D07, A01AC, S01BA, S01BB, L04AA, L04AB, L04AC, L04AD, L04AX, A07EA) and AE PT is in any of the irAE clusters. B) conmed ATC code is in (H03A, H03B) and AE PT is in one of the irAE clusters associated with "Immune-mediated endocrinopathies" C) conmed ATC code is A10A and AE PT is in the irAE cluster associated with "Immune-mediated endocrinopathies: Type I Diabetes Mellitus"

5	A) No clear etiology (other than immune mediated etiology)	A) From the AE eCRF page Is the AE clearly related to an etiology other than immune-mediated etiology? Yes / No If answer is Yes, check all that apply: • Underlying malignancy / progressive disease. • Other medical conditions. • Prior or concomitant medications / procedures. • Other. Specify.
	B) Histopathology / biopsy consistent with immune-mediated event Event is in if [Answer to 5B1 and 5B2 is YES (regardless of answer to 5A)] OR [Answer to 5B1 is YES AND answer to 5B2 is NO AND answer to 5A is NO] OR [Answer to 5B1 is NO AND answer to 5A is NO]	B) From the AE eCRF page B1) Was there a pathology /histology evaluation performed to investigate the AE? Y/N B2) If answer to the above is Yes, does the pathology/histology evaluation confirms an immune mediated mechanism for the AE? Y/N B3) If pathology / histology evaluation performed to investigate the AE, provide summary of relevant findings of the pathology /histology report. (Free Text)

The data set associated with irAEs may be refined based on medical review. The final data set including any changes based on medical review (e.g., addition of cases that are not selected by the programmatic algorithm) will be the basis of the irAE analyses.

Appendix 2. Infusion Related Reactions

For defining an AE as IRR the onset of the event in relation to the infusion of study drug and time to resolution of the event will be considered.

- All AEs identified by the MedDRA PT query describing signs and symptoms will be considered potential IRRs when onset is on the day of study drug infusion (during or after infusion) and the event resolved with end date within 2 days after onset.
- All AEs identified by the MedDRA PTs of Infusion related reaction, Drug hypersensitivity, Anaphylactic reaction, Hypersensitivity, Type 1 hypersensitivity, will be considered potential IRRs when onset is on the day of study drug infusion (during or after the infusion) or the day after the study drug infusion (irrespective of resolution date).

The list of MedDRA PTs for 'IRRs SIGNS and SYMPTOMS' and PTs 'IRRs CORE' are defined in the SRP for avelumab.

Infusion-related reactions (IRRs) will be programmatically identified as outlined in Table 14.

Table 14. Case Definition for IRRs

Condition	Selection criterion
If AE meets [1 AND 2] OR [3 AND (4A OR 4B)] then AE is classified as an IRR	
1	PT is included in the 'IRRs SIGNS and SYMPTOMS' list
2	• AE onset date = date of infusion of study drug AND • AE timing related to study drug ('DURING', 'AFTER') AND • AE outcome in ('RECOVERED/RESOLVED', 'RECOVERED/RESOLVED WITH SEQUELAE', 'RECOVERING/RESOLVING') AND • AE end date – AE onset date <=2
3	PT is included in the 'IRRs CORE' list
4A	• AE onset date = date of infusion of study drug AND • AE timing related to study drug in ('DURING', 'AFTER')
4B	AE onset on the day after infusion