Official Title of Study:

A Randomized, Open-Label, Two-arm, Comparative Study in Chinese Subjects with Chemotherapy Naïve Stage IV Melanoma Receiving Ipilimumab (3 mg/kg) vs. Dacarbazine.

NCT Number: NCT02545075

Document Date (Date in which document was last revised): 17 January 2019

STATISTICAL ANALYSIS PLAN
FOR CLINICAL STUDY REPORT

A RANDOMIZED, OPEN-LABEL, TWO-ARM, COMPARATIVE STUDY IN CHINESE SUBJECTS WITH CHEMOTHERAPY NAÏVE STAGE IV MELANOMA RECEIVING IPILIMUMAB (3 MG/KG) VS. DACARBAZINE.

PROTOCOL(S) CA184248

VERSION # 2.0

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

[redacted]

[redacted]

[redacted]

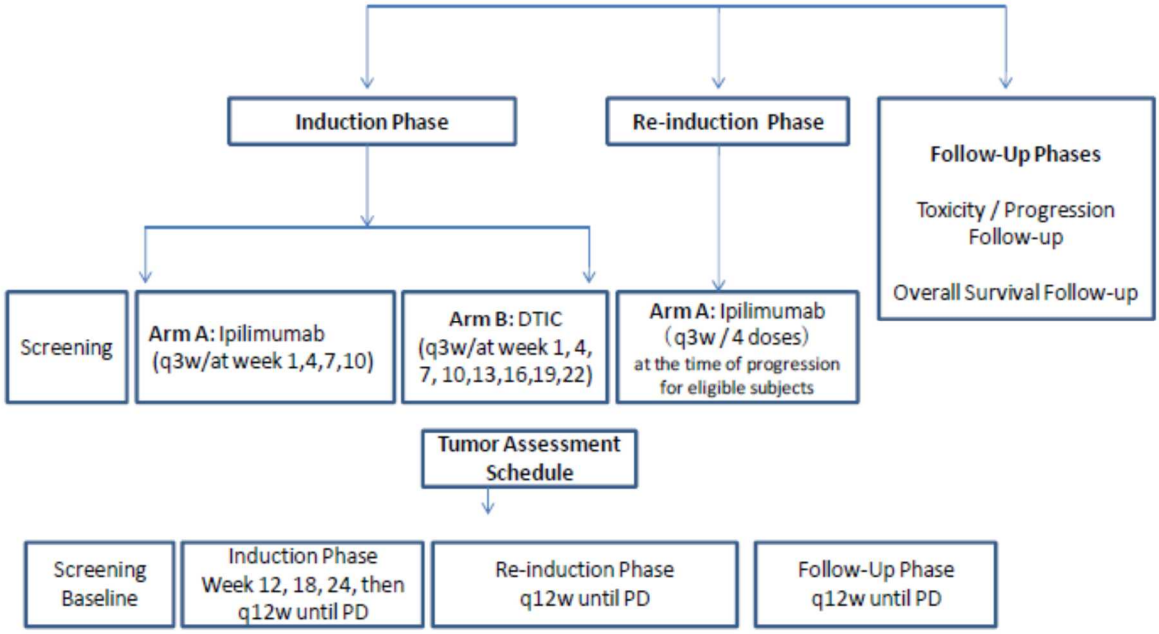
2 STUDY DESCRIPTION

2.1 Study Design

This SAP addresses the randomized, open-label study, CA184248, for which the planned sample size is approximately 180 randomized subjects with 2:1 allocation to ipilimumab 3mg/kg or DTIC.

An overview of the administration of study drug and assessment schedule is presented in Figure 2.1-1. Further details of the design are given in the Protocol.

Figure 2.1-1: Study Design Schematic



2.2 Treatment Assignment

After informed consent has been obtained, the site will enroll a subject into the study at the time of eligibility screening by accessing a call-in Interactive Voice Response System (IVRS). A subject number will be assigned by the IVRS at this time.

Following completion of all Screening evaluations, all eligible subjects will be randomized to one of two arms in a 2:1 ratio (Ipilimumab:DTIC) using a stratified permuted block randomization method with the following stratification actors:

- M-substage (M1a+M1b vs. M1c);
- BRAF mutation status (yes vs. no).

2.3 Blinding and Unblinding

Not applicable (open-label study).

2.4 Protocol Amendments

This SAP reflects the protocol amendments described in Table 2.4-1:

Table 2.4-1: Protocol Amendments

Amendments	Date of Issue	Summary of Major Changes
Revised Protocol 1	08-JUL-2015	- Update the dose of DTIC per China label requirement. - Other changes per program level standards and model document.
Revised Protocol 2	23-DEC-2015	- Remove the use of a Data Monitoring Committee (DMC since: 1) This is an open-label study; 2) This is a single country study (China) and not a MRCT (Multi Regional Clinical Trial) and therefore an independent DMC is not required as per the Chinese regulation. - Further clarification with regards to: 1) The infusion pump used for the administration of ipilimumab; 2) The T&E schedule (Endocrine assessments and collection of PK/ADA)
Revised Protocol 3	20-DEC-2018	-Change in primary objective and primary endpoint to 2-year OS rate from OS -Change in allowable minimum scan time from baseline for determination of stable disease from 12 weeks to 8 weeks per RECIST v1.1 guidelines

3 OBJECTIVES

3.1 Primary

- To compare the survival rate at 2 years in chemotherapy naive Chinese subjects with Stage IV melanoma who have been randomized to ipilimumab (3mg/kg) vs. DTIC (250mg/m² - Day 1 - 5).

3.2 Secondary

- To estimate survival rates at 1 year for each treatment arm;
- To compare overall survival between the two arms
- To compare PFS between the two arms by RECIST v1.1;
- To compare disease control rate between the two arms by RECIST v1.1;

- To estimate best overall response rate, duration of response, and duration of stable disease for each treatment arm by RECIST v1.1.

[redacted]

4 ENDPOINTS

4.1 Primary Endpoint

4.1.1 Two-Year Survival Rate

Two-year survival rate is defined as the probability that a subject is alive at 2 years following the randomization date and will be estimated via the Kaplan-Meier (KM) method.

4.2 Secondary Endpoints

Secondary response-based endpoints of PFS, DCR, DoR, DoSD, and BORR will be captured using the RECIST v1.1 criteria.

4.2.1 One-Year Survival Rate

Survival rate at 1 year is defined as the probability that a subject is alive at 1 year following the randomization date and will be estimated via the Kaplan-Meier (KM) method.

4.2.2 Overall Survival

OS is defined for each subject as the time between randomization date and the date of death (of any cause). If a subject has not died, the subject will be censored at the time of last contact (last known alive date).

A modified variable will be examined for the purposes of sensitivity analysis:

- OS accounting for subsequent PD-1/PD-L1 therapy will be defined through addition of the following restriction to the primary definition: subjects who receive subsequent PD-1/PD-L1 cancer therapy will be censored at the start date of subsequent PD-1/PD-L1 therapy.

4.2.3 Progression Free Survival

PFS is defined for each subject as the time between randomization date and the date of progression or death, whichever occurs first. The date of progression will be taken from the investigator assessment of progression on the CRF. A subject who dies without reported prior progression will be considered to have progressed on the date of death if no subsequent therapy prior to death date. For those who remain alive and have not progressed, PFS will be censored on the date of last evaluable tumor assessment (TA) prior to subsequent therapy. For those who remain alive and have no recorded evaluable TAs post baseline, PFS will be censored on the date of randomization.

For subjects without baseline tumor assessment, PFS will also be censored on the date of randomizaion.

Modified variables will be examined for the purposes of sensitivity analysis:

- *PFS accounting for early TAs in DTIC group* is defined in the same manner as PFS, except that for subjects randomized to DTIC with radiographic, photographic or clinical evidence of progression before Week 12, subjects will be considered to have PD at Week 12.

4.2.4 Disease Control Rate

Primary disease control rate (DCR) definition is defined as the number of subjects in the arm with Best Overall Response (BOR) of complete response (CR), partial response (PR), or stable disease (SD), divided by the total number of randomized subjects in the arm.

Per protocol DCR definition is defined as the number of response evaluable subjects in the arm with BOR of complete response (CR), partial response (PR), or stable disease (SD), divided by the total number of response evalubale subjects in the arm.

4.2.5 Best Overall Response Rate

Primary BORR definition is defined as the number of subjects in the arm with a BOR of CR or PR, divided by the total number of randomized subjects in the arm.

Per protocol BORR definition is defined as the number of response evaluable subjects in the arm with a BOR of CR or PR, divided by the total number of response evaluable subjects in the arm.

4.2.6 Duration of Response

Primary duration of response (DoR) definition for the randomized subjects whose BOR is CR or PR is defined as the time between the date of response of CR or PR (whichever occurs first) and the first date of progressive disease (PD) or the date of death (whichever occurs first).

Per protocol DoR definition for the response evaluable subjects whose BOR is CR or PR is defined as the time between the date of response of CR or PR (whichever occurs first) and the first date of progressive disease (PD) or the date of death (whichever occurs first).

For both definitions, for subjects who underwent tumor resection following response but prior to disease progression, DoR will be censored on the date of the last evaluable TA prior to resection. For subjects who remain alive and have not progressed following response, DoR will be censored on the date of last evaluable TA prior to subsequent therapy.

4.2.7 Duration of Stable Disease

Primary duration of stable disease (DoSD) definition for the randomized subjects whose BOR is SD is defined as the time between the randomization date and the first date of PD or the date of death (whichever occurs first).

Per protocol DoSD definition for the response evaluable subjects whose BOR is SD is defined as the time between the randomization date and the first date of PD or the date of death (whichever occurs first).

For both definitions, for subjects who underwent tumor resection following SD but prior to disease progression, DoSD will be censored on the date of the last evaluable TA prior to resection. For subjects who remain alive and have not progressed following SD, DoSD will be censored on the date of last evaluable TA prior to subsequent therapy.

[redacted]

[redacted]

[redacted]

5 SAMPLE SIZE AND POWER

Approximately 180 Chinese subjects will be randomized to ipilimumab and DTIC at a 2:1 ratio and followed at least 2 years after randomization to ensure 80% power through simulation using a chi-squared test with a Type I error rate of 0.20 to reject the null hypothesis of no difference in 2-year survival rate between treatment groups, assuming a HR [ipilimumab vs DTIC] of 0.67. Under an exponential distribution, this hazard ratio (HR) would correspond to an 12 month median OS for the ipilimumab arm with an 8 month median OS in the DTIC arm[8](or equivalently 25% and 12.5% in the ipilimumab and DTIC arm). The accrual duration is approximately 18 months, assuming a monthly rate of 10 subjects per month. The primary analysis will be performed about 42 months after the first subject is randomized.

6 STUDY PERIODS, TREATMENT REGIMENS AND POPULATIONS FOR ANALYSES

6.1 Study Periods

- Baseline period:
  - Baseline evaluations or events will be defined as evaluations or events that occur before the date and time of the first dose of study treatment. Evaluations on the same date and time of the first dose of study treatment will be considered as baseline evaluations.
  - In cases where the time (onset time of event or evaluation time and dosing time) is missing or not collected, the following definitions will apply:
    - Pre-treatment AEs will be defined as AEs with an onset date prior to but not including the day of the first dose of study treatment.
    - Baseline evaluations (laboratory tests, etc.) will be defined as evaluations with a date on or prior to the day of first dose of study treatment.

If there are multiple valid assessments, the assessment that is closest to day (and time if collected) of the first dose of study treatment will be used as the baseline in the analyses. If multiple assessments are collected on the same date (and time, if collected), the assessment with the latest database entry date (and time, if collected) will be considered as baseline, unless mentioned otherwise.

- Post baseline period:
  - On-treatment AEs will be defined as AEs with an onset date and time on or after the date and time of the first dose of study treatment (or with an onset date on or after the day of first dose of study treatment if time is not collected or is missing). For subjects who are off study treatment, AEs will be included if event occurred within a safety window of 90 days after the last dose of study treatment. No 'subtracting rule' will be applied when an AE occurs both pre-treatment and posttreatment with the same preferred term and grade.
  - On-treatment evaluations (laboratory tests, etc.) will be defined as evaluations taken after the day (and time, if collected and not missing) of first dose of study treatment. For subjects who are off study treatment, evaluations should be within a safety window of 90 days after the last dose of study treatment.

6.2 Treatment Regimens

The treatment group "**as randomized"** will be retrieved from the IVRS system

- Arm A: ipilimumab
- Arm B: DTIC

The treatment group "**as treated**" will be the same as the arm randomized by IVRS. However, if a subject received the incorrect drug for **the entire period** of treatment, the subject's treatment group will be defined as the incorrect drug the subject actually received.

6.3 Populations for Analyses

The statistical analyses in this study will be defined with respect to the following data sets.

All Enrolled Subjects: All subjects who signed an informed consent form and were registered into the IVRS. Subject disposition will be based on all enrolled subjects.

All Randomized Subjects: All enrolled subjects who were randomized to a treatment arm (ipilimumab or DTIC). Unless otherwise indicated, analyses of most of study population and efficacy endpoints will be based on all randomized subjects and will be performed using the treatment group as randomized (i.e. on an intent-to-treat basis). Analyses of QoL will also be based on the all randomized subjects population: subjects with both baseline and on-study measurements will contribute to the analyses of change at on-study time-points.

Response Evaluable Subjects: All randomized subjects who have measurable disease (i.e., at least one measurable lesion) at baseline.

All Treated Subjects: All subjects who received at least one dose of treatment with ipilimumab or DTIC. Unless otherwise specified, analyses of safety and extent of exposure will be based on

all treated subjects. Analyses of all treated subjects will be performed using the treatment group as treated.

Pharmacokinetic Subjects: All available pharmacokinetic data from subjects who received ipilimumab will be included.

ADA Evaluable Subjects: All treated subjects with baseline and at least one post-baseline ipilimumab immunogenicity assessment

7 STATISTICAL ANALYSES

7.1 General Methods

No interim analyses are planned for this study. The final analysis will be performed when a minimum of 2 year follow-up is reached by all randomized subjects.

Unless otherwise noted, the bulleted titles in the following subsections describe tabulations of discrete variables, by the frequency and proportion of subjects falling into each category, grouped by treatment (with total). Percentages given in these tables will be rounded to the first decimal and, therefore, may not always sum to 100%. Percentages less than 0.1 will be indicated as '< 0.1'. Continuous variables will be summarized by treatment group (with total) using the mean, standard deviation, median, minimum and maximum values. Subject listings will be produced to accompany the tabulations. Time-to-event variables (e.g. time-to resolution) will be analyzed using the Kaplan-Meier technique. When specified, the median will be reported along with 80% CI using Brookmeyer and Crowley method[9] (using log-log transformation for constructing the confidence intervals). The conventions to be used for imputing partial dates for analyses requiring dates are described in Section 8.

Frequency tables of subject characteristics and subset analyses of efficacy endpoints (including those by M-substage and BRAF mutation status) will be based on the assessment recorded on the baseline CRF. However, stratified analyses will be based on the variables used by the IVRS to randomize the subject.

Assessments taken during re-induction and subsequent progression follow-up phase will not contribute to derived endpoints, such as BOR or PFS, since these assessments are necessarily following an initial progression. However exploratory version of these endpoints corresponding to the administration of re-induction therapy may be performed (if there are sufficient data) - details will be provided in the DPP.

All statistical analyses will be carried out using SAS software (SAS Institute, Cary, North Carolina, USA), unless otherwise noted.

7.2 Study Conduct

Summaries of study conduct will be based on all randomized subjects.

7.2.1 Accrual

- Accrual by country and site
- Accrual by region
- Accrual by month randomized

7.2.2 Protocol Deviations

Each subject's data will be examined for deviation from protocol criteria (including violation of eligibility criteria and on-study deviations). 'Significant' deviations include those that could have an impact on the primary results of the study, those of ethical concern and those which pose a safety risk to the subject. The 'Relevant' deviations are those 'Significant" deviations that significantly affect *interpretability of the study results*. (Subjects who never received their

randomized treatment [at least one dose] will also be considered relevant deviations.) Subjects with relevant deviations will be determined programmatically (based on eCRF or central lab data). *Only programmable relevant deviations will be summarized and listed by the statistical team.* Programmable relevant deviation criteria are given in Table 7.2.2-1.

Table 7.2.2-1: Relevant Protocol Deviation Criteria

	Timeframe	Deviation Type	Protocol Deviation Specification
E1	Pretreatment	Eligibility	Subjects with wrong diagnosis: Disease Stage at study entry is not Stage IV malignant melanoma
E2	Pretreatment	Eligibility	Subjects with ECOG PS not equal to 0 or 1 during screening period.
E3	Pretreatment	Eligibility	Subject has no lesion (measurable or non-measurable) at screening
P1	On treatment	Not Discontinued	On-treatment laboratory test results that do not meet protocol-specified re-treatment criteria: ALT, AST, TBILI
P2	On treatment	Not Discontinued	Treated subjects whose study drug dosing continued after experiencing drug-related AEs requiring permanent discontinuation per protocol.
P3	On treatment	Incorrect Dosing	Dosing not delayed for Gr 3 or higher skin AE regardless of causality
P4	On treatment	Incorrect Dosing	Dosing not started within 3 days of randomization
P5	On treatment	Incorrect Dosing	Treatment assignment as randomized different from actual treatment given (as per Section 6.2)
P6	[redacted]	[redacted]	[redacted]

Relevant deviations will be summarized and will be listed by subject.

7.3 Study Population

Summaries of study population will be based on all randomized subjects except: subject disposition (based on all enrolled subjects), time from advanced disease diagnosis to first dose of study therapy (based on all treated subjects).

7.3.1 Subject Disposition

- Disposition of enrolled subjects (number of subjects randomized, treated, not treated [with reason])

7.3.2 Demography and Subject Characteristics

- Age (< 65, ≥ 65; summary statistics)
  - The age of a subject is the truncated difference in years between the date of the informed consent and the date of birth, plus 1 day.
- Height (summary statistics)
- Weight (summary statistics)
- Gender (male, female)
- ECOG performance status (0, 1)
- M-substage at study entry (M1a, M1b, M1c)
- Baseline LDH (≤ ULN, > ULN)
- Baseline LDH (≤ 2 × ULN, > 2 × ULN)
- BRAF mutation status (positive, negative, unknown)

The following summaries of factors used in IVRS to stratify subjects will be produced.

- M-substage stratification factor (M1a/M1b, M1c) x BRAF mutation status (yes, no)
- M-substage: stratification factor compared to CRF (M1a/M1b, M1c)
- BRAF mutation status: stratification factor compared to CRF (yes, no)

7.3.3 Medical History

- General medical history (head, eyes, ears, nose, throat; cardiovascular; peripheral vascular; respiratory; gastrointestinal; hepatobiliary; renal; genitourinary; endocrine-metabolic; hematologic-lymphatic; musculoskeletal; dermatologic; neurologic; psychiatric; immunology/allergies; neoplasia; alcohol use; tobacco use; drug abuse)

7.3.4 Disease Characteristics

- Time from advanced disease diagnosis to first dose of study therapy (summary statistics)
- Baseline TA (site of disease, number of disease sites)
- Baseline target lesions (number and site of target lesions, procedures for evaluating target lesions; summary statistics and quartiles of sum of reference diameters of target lesions)
  - TAs on or prior to first dose date will be considered as baseline. The latest assessment for each lesion will be considered.
- Melanoma subtype (mucosal, cutaneous, acral, ocular/uveal, and other)

7.3.5 Prior Therapy

- Prior therapy (systemic cancer, radiotherapy, surgery related to cancer)
- Duration of prior systemic therapy (summary statistics)

7.3.6 Clinical Complaints and Findings

Events with onset prior to the date of first dose will be examined (whether recorded on the pre-treatment events CRF or on the on-treatment adverse events CRF).

- Pre-treatment events (worst CTC grade of any event and of events by SOC)

7.3.7 Baseline Laboratory Examinations

- Baseline hematology (CTC grade for hemoglobin, WBC, ANC, platelet count, ALC)
- Baseline liver function (CTC grade for ALT, AST, total bilirubin, ALP)
- Baseline renal function (CTC grade for creatinine)
- Baseline pancreatic function: (CTC grade for amylase)

7.4 Extent of Exposure

Summaries of exposure will be based on all treated subjects, except for the summary for subsequent therapies (based on all randomized subjects).

7.4.1 Administration of Study Therapy

- Time from randomization to first dose of ipilimumab (0 to 3 days, > 3 to 7, > 7 to 14, > 14 to 21, > 21 to 28, > 28)
- Number of doses of study therapy (summary statistics)
- Cumulative dose of study therapy (summary statistics)

  Cumulative dose of ipilimumab/dacarbazine in mg is defined as the sum of all calculated doses (mg) at each dosing. Cumulative dose of ipilimumab in mg/kg and dacarbazine in mg/m2 is defined as the sum of all doses (in mg/kg and mg/m2 respectively) received by the subject.

  Ipilimumab dose in mg/kg is calculated as the dose in mg divided by the most recent weight in kg prior to infusion. If the subject's weight is missing, the last observation carried forward approach will be used.

  Dacarbazine dose in mg/m2 is calculated as the dose in mg divided by the most recent body surface area (BSA) prior to infusion. BSA is derived from most recent weight and baseline height. If the subject's weight is missing, the last observation carried forward approach will be used.

Cumulative dose will be listed by subject.

7.4.2 Discontinuation from Induction Phase

Failure to complete the (re-)induction phase will be tabulated.

- Number of subjects who completed or failed to complete (with reason) the (re-)induction phase.

- Discontinuation from study therapy in induction phase (subjects with study therapy discontinued, reason for discontinuation).
- Discontinuation from ipilimumab in re-induction phase (subjects with study therapy discontinued, reason for discontinuation).

7.4.3 Interruption of Infusion or Delay of Study Therapy

An infusion interruption refers to an interruption of the infusion of study therapy, whether or not the infusion is resumed. A dose delay refers to a delay in the administration of ipilimumab (but not a discontinuation thereof) relative to the protocol schedule. A dose omission refers to an omission in the administration of DTIC (but not a discontinuation thereof) relative to the protocol schedule.

- Infusion interruption in induction phase (subjects with at least one interrupted dose, reason for first dose interruption, number of interruptions).
- Infusion rate reduction in induction phase (subjects with at least one infusion with IV rate reduced, reason for first infusion IV rate reduced, number of infusions with IV rate reduced).
- Dose delay of ipilimumab in induction phase (subjects with at least one delayed dose, reason for first dose delay, number of delays).
- Dose omission of DTIC in induction phase (subjects with at least one omitted dose, reason for first dose omitted, number of omissions).
- Infusion interruption of ipilimumab in re-induction phase (subjects with at least one interrupted dose, reason for first dose interruption, number of interruptions).
- Infusion rate reduction of ipilimumab in re-induction phase (subjects with at least one infusion with IV rate reduced, reason for first infusion IV rate reduced, number of infusions with IV rate reduced).
- Dose delay of ipilimumab in re-induction phase (subjects with at least one delayed dose, reason for first dose delay, number of delays).

[redacted]

7.4.5 Other On-Study Therapy

Non-protocol-specified medical treatment procedures which occur on-study (i.e. on or after the first day of study therapy and within 90 days following the last dose of study therapy) will be summarized and listed.

- On-study surgery (with reason)
- On-study radiotherapy (with reason)

7.4.6 Subsequent Therapy

Therapies received subsequent to discontinuation of study therapy will be summarized and listed.

- Frequency of subsequent systemic therapies by medication class and generic name
- Frequency of subsequent systemic therapy type (any, and classes, e.g.: immunotherapy only, chemotherapy only, immunotherapy and chemotherapy, other)
- Frequency of subsequent surgery
- Frequency of subsequent radiotherapy
- Time from randomization to first subsequent therapy (median, range)
- Listing of subsequent therapies (including dates of therapy)

7.5 Efficacy

All hypothesis testing will be two-sided and will be based upon a significance level of 0.20.

No type I error rate multiplicity adjustment will be made for secondary endpoint testing. P-values associated with the comparisons of OS, DCR and PFS will be purely descriptive in nature. All efficacy analyses will be based on all randomized subjects, as randomized, unless indicated.

7.5.1 Survival Rates at Specified Time Points

Yearly survival rates will be calculated for each treatment group using the KM product-limit method, along with their corresponding log-log transformed 80% CIs[10].

Two-year survival rate is the primary endpoint. The survival rate at 2 years will be compared between treatment group using a chi-square test stratified by M substage (M1a+M1b versus M1c) and BRAF mutation status (yes vs no), as recorded in the IVRS. The stratified chi-square test is defined by the following formula:

$$\chi^2 = \frac{\left(\sum_{s=1}^{m} \hat{S}_{1s}(t) - \hat{S}_{2s}(t)\right)^2}{\sum_{s=1}^{m} \hat{V}\left(\hat{S}_{1s}(t)\right) + \hat{V}\left(\hat{S}_{2s}(t)\right)},$$

where $\hat{S}_{ks}(t)$ and $\hat{\mathrm{V}}\left(\hat{\mathrm{S}}_{\mathrm{ks}}(\mathrm{t})\right)$ be the Kaplan-Meier estimate and variance based on the Greenwood formula for the *s*th stratum in the *k*th treatment group at time *t*.

7.5.1.1 Subgroup Analyses

Two-year survival rates along with two-sided 80% CI based on the log-log transformation will be presented within subsets defined by information recorded on the baseline CRF.

This summary and comparison based on the unstratified chi-square test will be provided for the following subgroups:

- by age (< 65, ≥ 65)
- by gender (male, female)
- by M-Substage at study entry (M1a/M1b, M1c)
- by ECOG performance status (0, 1)

- by baseline LDH (≤ULN, > ULN)
- by baseline LDH (≤ 2 × ULN, > 2 ×ULN)
- by BRAF mutation status (positive, negative, unknown)
- by subtype (acral, cutaneous, other)

In case a subgroup includes less than 10 subjects, the analysis for the given subgroup will not be carried out (or the possibility of combining subgroups will be considered).

7.5.1.2 Sensitivity Analyses

The following analyses will be conducted

- An unstratified chi-square test will be performed to compare survival rate at 2 years after randomization between treatment groups
- The stratified and unstratified chi-square tests based on the modified OS accounting for subsequent PD-1/PD-L1 therapy.

7.5.2 Overall Survival

OS within treatment group, estimated using the KM product-limit method, will be plotted. Summary tables will report the median OS together with a two-sided 80% confidence interval (CI) for the median, calculated using the method of Brookmeyer and Crowley. All available events (deaths) in the database will contribute to the analysis of OS.

The extent of follow-up within treatment group, defined as the time between randomization and last known alive date whether status is alive or dead, will be summarized, as will be the currentness of follow-up, defined as the time between last known alive date and data cut-off date (defined by last patient last visit date). Subjects who died and subjects with last known date alive on or after data cut-off date will be considered as current.

- Summary of OS
- Extent of follow-up
- Currentness of follow-up

7.5.2.1 Comparison between Treatment Groups

OS is the key secondary efficacy endpoint. The distribution of OS will be compared between treatment groups using a two-sided log-rank test stratified by M substage (M1a+M1b versus M1c) and BRAF mutation status (yes vs no), as recorded in the IVRS. The HR, and its associated two-sided 80% CI, of ipilimumab to DTIC will be estimated using a Cox proportional hazards model, stratified by the factors described above, and with treatment as the single covariate.

- Comparison of OS

7.5.2.2 Subgroup Analyses

OS will be summarized within subsets defined by information recorded on the baseline CRF (with two-sided 80% CI for the median OS calculated via the method of Brookmeyer and Crowley). For informal comparisons of OS within these subsets, the HR, and its 80% two-sided CI, of ipilimumab

to DTIC will be computed using an unstratified Cox proportional hazards model with treatment as the single covariate.

This summary and comparison will be provided for the following subgroups:

- by age (< 65, ≥ 65)
- by gender (male, female)
- by M-Substage at study entry (M1a/M1b, M1c)
- by ECOG performance status (0, 1)
- by baseline LDH (≤ ULN, > ULN)
- by baseline LDH (≤ 2 × ULN, > 2 × ULN)
- by BRAF mutation status (positive, negative, unknown)
- by subtype (acral, cutaneous, other)

  In case a subgroup includes less than 10 subjects, the analysis for the given subgroup will not be carried out (or the possibility of combining subgroups will be considered).

7.5.2.3 Sensitivity Analyses

An unstratified log-rank test will be performed to compare OS between treatment groups. The HR of ipilimumab to DTIC and the corresponding two-sided 80% CI will also be estimated using an *unstratified*, multivariate Cox proportional hazards model, with covariates for stratification factors (from CRF), as well as for treatment group.

- Comparison of OS using unstratified log-rank test
- Comparison of OS using multivariate Cox model
- Furthermore, the following analyses will be presented for the modified OS (defined in Section 4.2.2):
  - Summary of OS accounting for subsequent PD-1/PD-L1 therapy
  - Comparison of OS accounting for subsequent PD-1/PD-L1 therapy
  - KM plot of OS accounting for subsequent PD-1/PD-L1 therapy

7.5.3 Progression-Free Survival

PFS is one of the secondary endpoints. Similar methods as defined for OS will be used for PFS.

PFS within treatment group, estimated using the KM product-limit method, will be plotted. The plot will display the median PFS, together with a two-sided 80% CI for the median, calculated using the method of Brookmeyer and Crowley.

For the comparison of PFS between the treatment groups, a two-sided log-rank test, stratified by M substage (M1a+M1b versus M1c) and BRAF mutation status (yes vs no), as recorded in the IVRS, will be used.

The HR, and its two-sided 80% CI, of ipilimumab to DTIC will be estimated using a Cox proportional hazards model, stratified by the factors described above, and with treatment as the single covariate.

- Summary of PFS
- Comparison of PFS

7.5.3.1 Sensitivity Analyses

An unstratified log-rank test will be performed to compare PFS between treatment groups. The HR of ipilimumab to DTIC and the corresponding two-sided 80% CI will also be estimated using an *unstratified*, multivariate Cox proportional hazards model, with covariates for stratification factors (from CRF), as well as for treatment group.

- Comparison of PFS using unstratified log-rank test
- Comparison of PFS using multivariate Cox model
- Furthermore, the following analyses will be presented for the modified PFS variables (defined in Section 4.2.3):
  - Summary of PFS accounting for early TAs in DTIC group
  - Comparison of PFS accounting for early TAs in DTIC group
  - KM plot of PFS accounting for early TAs in DTIC group

7.5.4 Disease Control Rate

Exact, two-sided 80% CI for DCR within treatment group will be calculated using the method of Clopper and Pearson[11].

For the comparison of DCR between treatment groups, a Cochran-Mantel-Haenszel (CMH) test with an associated odds ratio estimate and exact two-sided 80% CI, stratified by M substage (M1a+M1b versus M1c) and BRAF mutation status (yes vs. no), as recorded in the IVRS, will be applied.

- Summary of DCR (primary definition)
- Comparison of DCR (primary definition)
- Summary of DCR (per protocol definition)
- Comparison of DCR (per protocol definition)

7.5.5 Best Overall Response

- Summary of BOR (primary definition)
- Summary of BOR (per protocol definition)

7.5.6 Best Overall Response Rate

An exact, two-sided 80% CI for BORR within treatment group will be calculated using the method of Clopper and Pearson.

- Summary of BORR (primary definition)
- Summary of BORR (per protocol definition)

7.5.7 Duration of Response

The DoR in subjects with a BOR of CR or PR, estimated using the KM product-limit method, will be plotted. The plot will display the median duration, together with a two-sided 80% CI for the median, calculated using the method of Brookmeyer and Crowley.

- Summary of DoR (primary definition)
- Summary of DoR (per protocol definition)

7.5.8 Duration of Stable Disease

The DoSD in subjects whose BOR is SD, estimated using the KM product-limit method, will be plotted. The plot will display the median duration, together with a two-sided 80% CI for the median, calculated using the method of Brookmeyer and Crowley.

- Summary of DoSD (primary definition)
- Summary of DoSD (per protocol definition)

7.6 Safety

Summaries of safety will be based on all treated subjects by treatment arm, except for the resource utilization (based on all randomized subjects).

7.6.1 Adverse Events

Adverse Events, Serious Adverse Events

The following summary tables of AEs and SAEs will be restricted to events occurring in the on-study phase (see Section 4.3.1.1).

- On-study phase AEs (worst CTC grade [Any grade, 3-4, 5] of any event and of events by SOC)
- On-study phase AEs (worst CTC grade [1, 2, 3, 4, 5, unknown, any grade] of any event and of events by SOC)
- On-study non-serious AEs occurring in at least 5% of subjects (worst CTC grade [Any grade, 3-4, 5] of any event and of events by SOC)
- On-study phase SAEs (worst CTC grade [Any grade, 3-4, 5] of any event and of events by SOC)
- On-study phase SAEs (worst CTC grade [1, 2, 3, 4, 5, unknown, any grade] of any event and of events by SOC)
- On-study phase drug-related AEs (worst CTC grade [Any grade, 3-4, 5] of any event and of events by SOC)
- On-study phase drug-related AEs (worst CTC grade [1, 2, 3, 4, 5, unknown, any grade] of any event and of events by SOC)
- On-study phase drug-related SAEs (worst CTC grade [Any grade, 3-4, 5] of any event and of events by SOC)
- On-study phase drug-related SAEs (worst CTC grade [1, 2, 3, 4, 5, unknown, any grade] of any event and of events by SOC)

- On-study phase AEs leading to discontinuation of ipilimumab (worst CTC grade [Any grade, 3-4, 5] of any event and of events by SOC)
- On-study phase AEs leading to discontinuation of ipilimumab (worst CTC grade [1, 2, 3, 4, 5, unknown, any grade] of any event and of events by SOC)

Listings by subject will be produced for all AEs (including drug-related AEs and those leading to discontinuation of ipilimumab), all AEs leading to discontinuation of ipilimumab (including drug-related AEs), all SAEs (including drug-related SAEs).

Immune-Related Adverse Events

Summaries irAEs will be based on all treated subjects and will use the pre-defined list of MedDRA high-level group terms, high-level terms and preferred terms that is current at the time of analysis. (APPENDIX 1 contains the most recent version of the list, version 20.0)

The following summary tables of irAEs will be restricted to events occurring in the on-study phase (see Section 4.3.1.1).

- On-study phase irAEs (worst CTC grade [Any grade, 3-4, 5] of any event and of events by SOC)
- On-study phase irAEs (worst CTC grade [1, 2, 3, 4, 5, unknown, any grade] of any event and of events by SOC)
- On-study phase irAEs by subcategory (worst CTC grade [Any grade, 3-4, 5] by subcategories of GI, liver, skin, endocrine, neurological, other)
- On-study phase irAEs by subcategory (worst CTC grade [1, 2, 3, 4, 5, unknown, any grade] by subcategories of GI, liver, skin, endocrine, neurological, other)
- Time to onset of on-study phase grade 3-5 irAEs by subcategory (GI, liver, skin, endocrine, neurological).
- Time to resolution of on-study phase grade 3-4 irAEs by subcategory (GI, liver, skin, endocrine, neurological)
- Number of episodes of on-study phase grade 3-4 irAEs by subcategory (GI, liver, skin, endocrine, neurological)

Listings by subject will be produced for all on-study and post-study irAEs (including steroid therapy received and details of resolution).

Immune-Mediated Adverse Reactions

Summaries of imARs will be based on all treated subjects.

The following summaries for on-study phase imARs will be provided (for subcategories as specified in the core imAR SAP):

- Summary of subjects with worst grade of grade 2 on-study phase imAR
- Summary of subjects with worst grade of grade 3-4 on-study phase imAR

- Summary of subjects with worst grade of grade 5 on-study phase imAR
- Summary of subjects with worst grade of grade 2-5 on-study phase imAR and outcome of hospitalization by subcategory (worst CTC grade [2+, 3-4, 5] by subcategories)
- Summary of subjects with worst grade of grade 2-5 on-study phase imAR who received blood transfusion by subcategory (worst CTC grade [2+, 3-4, 5] by subcategories)
- Time to first onset (any grade) of imARs for subjects with worst grade of grade 2-5 on-study phase imAR by subcategory
- Time to resolution of imARs for subjects with worst grade of grade 2-5 on-study phase imARs by outcome of resolution by subcategory
- Time to resolution of imARs for subjects with worst grade of grade 2-5 on-study phase imARs by use of systemic corticosteroids and outcome of resolution by subcategory
- Time to resolution of imARs for subjects with worst grade of grade 2-5 on-study phase imARs by use of high dose systemic corticosteroids and outcome of resolution by subcategory
- Time to resolution of imARs for subjects with worst grade of grade 2-5 on-study phase dermatitis imARs by use of topical corticosteroids and outcome of resolution
- imARs leading to study drug discontinuation by use of high dose systemic corticosteroids for subjects with worst grade of grade 2-5 imARs, by subcategory
- Distribution of outcome (last reported grade) of imARs for subjects with worst grade of grade 2-5 on-study phase imARs by subcategory
- Distribution of outcome (last reported grade) of imARs for subjects with worst grade of grade 2-5 on-study phase imARs by use of systemic corticosteroids by subcategory
- Distribution of outcome (last reported grade) of imARs for subjects with worst grade of grade 2-5 on-study phase imARs by use of high dose systemic corticosteroids by subcategory
- Distribution of outcome (last reported grade) of imARs for subjects with worst grade of grade 2-5 on-study phase dermatitis imARs by use of topical corticosteroids
- Distribution of outcome (last reported grade) of imARs for subjects with worst grade of grade 3-5 on-study phase imARs by subcategory
- Duration of administration of systemic corticosteroids for imARs for subjects with worst grade of grade 2-5 on-study phase imARs by subcategory
- Duration of administration of systemic corticosteroids for imARs for subjects with worst grade of grade 2-5 on-study phase imARs by high dose systemic corticosteroids, by subcategory
- Duration of administration of topical corticosteroids for imARs for subjects with worst grade of grade 2-5 on-study phase dermatitis imARs
- Distribution of immunosuppression administration for imARs for subjects with worst grade of grade 2-5 imARs by subcategory
- Distribution of administration of hormonal therapy for immune-mediated endocrinopathies for subjects with worst grade of grade 2-5 on-study phase immune-mediated endocrinopathies by use of high dose of systemic corticosteroids by subcategory
- Summary of Dosage of Systemic Corticosteroids for Subjects with Worst Grade of Grade 2-5 imARs, by subcategory

- Number of Subjects Receiving High Dose Systemic Corticosteroids for any imAR for Subjects with Worst Grade of Grade 2-5 imARs
- Cross-tabulation of imAR Grade at First Onset Versus Worst imAR Grade, by Class
- Time From First Non-Severe (Grade 1 or 2) to First Severe (Grade 3-5) Onset for imARs, by Class

For endocrinopathies only:

- Time on Hormonal Replacement Therapy for Endocrinopathy imARs, by type of therapy (adrenal replacement therapy [identified as systemic corticosteroids (identified by the BMS physicians) used for adrenal replacement (as indicated by the investigator on the concomitant medication CRF, use=adrenal insufficiency)], thyroid medication, testosterone medication) and by medication status (therapy discontinued, therapy ongoing)
- Distribution of Outcome (Last Reported Grade) of Endocrinopathy imARs for Subjects with Worst Grade of Grade 2-5 imARs, by Hormonal Replacement Therapy Status (adrenal replacement therapy discontinued, adrenal replacement therapy ongoing, hormonal replacement therapy other than adrenal replacement therapy, no hormonal replacement therapy)
- Serious Endocrinopathy imARs for Subjects with Worst Grade of Grade 2-5 Endocrinopathy imARs, by Hormonal Replacement Therapy Status (adrenal replacement therapy discontinued, adrenal replacement therapy ongoing, hormonal replacement therapy other than adrenal replacement therapy, no hormonal replacement therapy)
- Listing of Endocrinopathy imARs and Hormonal Replacement Therapy status for subjects with on-study endocrinopathy imARs
- Time on Systemic Corticosteroid for Endocrinopathy imARs, by medication status (therapy discontinued, therapy ongoing)

  Present the table for 3 categories:

  – ANY SYSTEMIC CORTICOSTEROID
  – SYSTEMIC CORTICOSTEROIDS USED FOR ADRENAL REPLACEMENT
  – SYSTEMIC CORTICOSTEROIDS FOR USE OTHER THAN ADRENAL REPLACEMENT
- Listing of Time on Systemic Corticosteroids for Endocrinopathy imARs, for subjects with on-study endocrinopathy imARs, by medication status (therapy discontinued, therapy ongoing)

  Present the listing for 2 categories:

  – SYSTEMIC CORTICOSTEROIDS USED FOR ADRENAL REPLACEMENT
  – SYSTEMIC CORTICOSTEROIDS FOR USE OTHER THAN ADRENAL REPLACEMENT
- Distribution of Outcome (Last Reported Grade) of Endocrinopathy imARs by Use of Systemic Corticosteroids for Subjects with Worst Grade of Grade 2-5 Endocrinopathy imARs

  Present the table for 3 categories:

  – SYSTEMIC CORTICOSTEROIDS USED FOR ADRENAL REPLACEMENT

  - SYSTEMIC CORTICOSTEROIDS FOR USE OTHER THAN ADRENAL REPLACEMENT
  - NO SYSTEMIC CORTICOSTEROIDS
- Time to Resolution for Endocrinopathy imARs by use of Systemic Corticosteroids and by Outcome of Resolution for Subjects with Worst Grade of Grade 2-5 Endocrinopathy imARs

  Present the table for 3 categories:
  - SYSTEMIC CORTICOSTEROIDS USED FOR ADRENAL REPLACEMENT
  - SYSTEMIC CORTICOSTEROIDS FOR USE OTHER THAN ADRENAL REPLACEMENT
  - NO SYSTEMIC CORTICOSTEROIDS

The following listings will be produced for all class imARs:

- Listing of On-Study and Post-Study imARs
  - For events with imAR=Yes or Unknown or No, and including reasons for imAR=Yes/No
  - Post induction events at class level (i.e., events that occur for the first time after the induction period and prior to the start of the re-induction period without any events in the same class that occurred during the induction period) will be flagged.
  - Novel re-induction events will be flagged.
  - This listing will include all classes: enterocolitis, dermatitis, hepatitis, endocrinopathies, neuropathies, other.
  - Present all treatment regimens.
- Listing of Time to Onset and Resolution of on-study imAR
  - This listing will include classes: enterocolitis, dermatitis, hepatitis, endocrinopathies, and neuropathies.
  - Present only ipilimumab containing regimens.
- Listing of imARs Initiated After 90 Days of Last Dose Corresponding to on-study imARs Class
  - This listing will include classes: enterocolitis, dermatitis, hepatitis, endocrinopathies, and neuropathies.
  - Present only ipilimumab containing regimens.
- Listing of New imARs Initiated During High Dose Systemic Corticosteroid Treatment of Early Onset imARs for subjects with on-study imAR
  - This listing will include all classes: enterocolitis, dermatitis, hepatitis, endocrinopathies, neuropathies, and other.
  - Present only ipilimumab containing regimens.

The following summaries for <u>novel re-induction</u> imARs will be provided:

- Summary of subjects with worst grade of grade 2 imAR
- Summary of subjects with worst grade of grade 3-4 imAR
- Summary of subjects with worst grade of grade 5 imAR

Multiple Events Analyses

For the multiple event analyses, the following summary tables will be provided:

- A table showing the total number and rate (exposure-adjusted) of occurrences for all on-study AEs occurring in at least 5% of the subjects treated
- A table showing the total number and rate (exposure-adjusted) of occurrences for all on-study irAEs
- A table showing the number of subjects experiencing an on-study AE once or multiple times

No formal comparisons/statistical testing will be performed. A listing will be produced that displays the unique instances of all AEs, i.e. after duplicates have been eliminated and overlapping and contiguous occurrences of the same event have been collapsed.

7.6.2 Clinical Laboratory Evaluations

Laboratory parameters will be analyzed based on all treated subjects with an on-study test conducted. For the list of laboratory parameters, see Section 7.3.7.

- On-study hematology (worst CTC grade)
- Baseline vs. on-study hematology (worst CTC grade)
- On-study liver function (worst CTC grade)
- Baseline vs. on-study liver function (worst CTC grade)
- On-study renal function (worst CTC grade)
- Baseline vs. on-study renal function (worst CTC grade)
- On-study pancreatic function (worst CTC grade)
- Baseline vs. on-study pancreatic function (worst CTC grade)

7.6.3 Deaths

Deaths will be analyzed based on all treated subjects.

- All deaths (with reason)
- Deaths within 30 days of last dose received
- Deaths within 90 days of last dose received

All deaths, together with the reason, will be listed by subject; deaths occurring within 30 days following the last day of study treatment, and within 90 days following the last day of study treatment will be indicated.

7.6.4 Resource Utilization

The number of subjects hospitalized, with primary reason for hospitalization, will be tabulated, as will be type of non-protocol-specified visit. Summary statistics (N, mean with SD, median, minimum, maximum) will be provided by treatment group for number of hospitalizations per subject, length of stay, number of non-protocol-specified visits per subject.

- Summary of hospitalization

- Summary of non-protocol-specified visits

7.6.5 Immunogenicity

A listing will be provided of all available ADA assessments. The frequency of ADA positive subjects and of neutralizing ADA positive subjects will be tabulated.

7.7 Pharmacokinetic Analyses

PK parameters will be summarized and listed. PK data obtained from this study may be pooled with data from other studies to perform an integrated population PK analysis (including assessment of covariate effects on PK), as well as exposure-response analysis for selected safety and efficacy endpoints. These analyses will be described in (a) separate report(s).

7.8 Health-Related Quality of Life

At each time point for analysis, the number of subjects who completed the questionnaires will be reported as a percentage of eligible subjects. Change from baseline analyses will be based on all randomized subjects who have a baseline measurement and an on-study assessment for the time-point of interest.

7.8.1 EORTC QLQ-C30

The following descriptive analyses will be conducted:

- EORTC QLQ-C30 questionnaire completion rate, defined as the proportion of questionnaires actually received out of the expected number (i.e., number of subjects on treatment or in follow up), will be calculated and summarized for each assessment time point by treatment group.
- Mean scores and post-baseline score changes will be summarized using descriptive statistics (N, mean with SD and 80% CI, median, first and third quartiles, minimum, maximum) for all scales at each assessment time point.
- A line graph summarizing the mean post-baseline score changes will be produced for all scales.

7.8.2 EuroQolEQ-5D-3L

The following descriptive analyses will be conducted:

- EQ-5D-3L questionnaire completion rate, defined as the proportion of questionnaires actually received out of the expected number (i.e. number of subjects on treatment or in follow up), will be calculated and summarized for each assessment time point by treatment group.
- A by-subject listing of the level of problems in each dimension, corresponding EQ-5D-3L health state (i.e., 5-digit vector), EQ-5D-3L utility index score, and EQ-5D-3L VAS score will be provided.
- Proportion of subjects reporting problems for the 5 EQ-5D-3L dimensions at each assessment timepoint will be summarized by level of problem and by treatment group. Percentages will be based on number subjects assessed at assessment time point.
- For the EQ-5D-3L utility index and VAS scores, separately:

- Mean score and mean change from baseline at each assessment time point will be summarized by treatment group using descriptive statistics (N, mean with SD and 80% CI, median, first and third quartiles, minimum, maximum).
- A line graph summarizing the mean post-baseline score changes will be produced

8 CONVENTIONS

Unless otherwise noted, the following conventions should be understood to apply. Further conventions may be detailed in the Data Presentation Plan.

The duration between two dates will be calculated as [later date] – [earlier date] + 1 day. Study day will be calculated as

$$Day = \begin{cases} [assessment date] - [first medication date], & if\ assessment before\ first\ medication, \\ [assessment date] - [first medication date] + 1\ day, & if\ assessment on\ or\ after\ first\ medication. \end{cases}$$

Assessment times will be calculated relative to the first date of study medication, i.e. Day 1.

Events will be considered as occurring within X days (e.g. deaths within 30 days, deaths and AEs within 90 days) of last dose of study medication if [event date] – [last dosing date] ≤ X.

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

Conventions for Partial/Missing Dates

Unless specified otherwise the following conventions will be used for imputing partial dates:

- If only the day of the month is missing, the 15th of the month will be used to replace the missing day.
- If both the day and the month are missing, "July 1" will be used to replace the missing information.
- If a date is completely missing, it will be considered as missing.

For date of death:

- If only the day of the month is missing, in principle the 1st of the month will be used to replace the missing day - but the imputed day of death will still be compared to the last known alive date (date of censoring for survival), and may be adjusted to ensure consistency.
- If both the day and the month are missing, in principle "Jan 1" will be used to replace the missing information - but the imputed day of death will still be compared to the last known alive date (date of censoring for survival), and may be adjusted to ensure consistency.
- If date is missing, death date will be imputed as the last known alive date + 1 day.

For date of progression:

- If only the day of the month is missing, the 1st of the month will be used to replace the missing day(*).

- If the day and month are missing or a date is completely missing, it will be considered as missing.

(*) In case the date of death is present (complete date), the imputed progression date will be compared with the date of death. The minimum of the (imputed progression date, date of death) will be considered as the date of progression.

For date of last tumor assessment:

- If only the day of the month is missing, the 1st of the month will be used to replace the missing day(*).
- If the day and month are missing or a date is completely missing, it will be considered as missing.

(*) In case the date of death is present (complete date), the imputed date will be compared with the date of death. The minimum of the (imputed date, date of death) will be considered as the date of last tumor assessment.

Otherwise, missing values will not in general be imputed. Regardless of missing data, all subjects will contribute to the calculation of the following tumor response endpoints (as detailed in Sections 4.2.3-4.2.6): BORR and disease control rate (by inclusion in the denominators); and in the Kaplan-Meier analyses (by censoring, as appropriate) of DoR (for subjects with confirmed CR or PR), DoSD (for subjects whose BOR is SD) and PFS.

Conventions for EQ-5D:

The EQ-5D descriptive system will be converted into a single summary index score which provides a simple measure of health on a scale of 0 to 1 (0 = death and 1 = perfect health) as follows[12]:

X = 1 - 0.146016*MO2 - 0.557685*MO3 - 0.1753425*SC2 - 0.4711896*SC3 - 0.1397295*UA2 - 0.3742594*UA3 - 0.1728907*PD2 - 0.5371011*PD3 - 0.156223*AD2 - 0.4501876*AD3 + 0.1395949*D1 - 0.0106868*I2^2 + 0.1215579*I3 + 0.0147963*I3^2,

where

- MO2, SC2, UA2, PD2, and AD2 are dummy variables for level 2 in dimensions mobility, self-care, usual activities, pain/discomfort, and anxiety/depression, respectively;
- MO3, SC3, UA3, PD3, and AD3 are dummy variables for level 3 in dimensions mobility, self-care, usual activities, pain/discomfort, and anxiety/depression, respectively;
- D1 is an ordinal variable: the number of movements away from Full health (1) beyond the first (i.e. it took on values ranging from 0 to 4);
- I2 is an ordinal variable: the number of dimensions at level 2 beyond the first;
- I3 is an ordinal variable: the number of dimensions at level 3 beyond the first.

9 CONTENT OF REPORTS

Study reports will be written corresponding to the final analysis of the study. The table, figure and listing outputs that will be produced are described in the Data Presentation Plan for the study.

10 DOCUMENT HISTORY

Table 10-1: Document History

Version	Author	Summary of Changes
1.0	[redacted]	Original version
2.0	[redacted]	-Incorporate protocol amendment (revprot03) to update 2-year OS rate as primary endpoint and OS as secondary endpoint -Add clarifications in health related quality of life -General updates for typos, inconsistencies, etc

APPENDIX 1 IMMUNE-RELATED ADVERSE EVENT SEARCH CRITERIA VERSION 20.0 (APRIL 2017)

The following MedDRA terms (**MedDRA version 20.0)** will be used for generating irAE AEs/SAEs among all adverse events reported by investigators as possibly, probably or certainly related or with unknown causality (PTs from primary SOC will be used for multi-category PTs with exception as noted).

GI IRAEs

HLGT-Gastrointestinal haemorrhages NEC - except PT-mesenteric hematoma, PT mesenteric haemorrhage
HLGT-Gastrointestinal inflammatory conditions (1 new PT: Lymphocytic oesophagitis)
HLGT-Gastrointestinal ulceration and perforation (2 new PTs: :Lower GI perforation, Upper GI perforation)
HLT-Diarrhoea (excl infective)
HLT-Stomatitis and ulceration
PT - Peritonitis
PT - Gastrointestinal wall thickening

Skin IRAEs

PT-Rash papular
PT- Rash pustular
PT - Macule
PT - Papule
PT - Folliculitis
PT - Pruritus genital
PT - Vulvovaginal pruritus
PT - Dermatitis acneiform
PT- Anal pruritus
PT- Ear pruritus
PT- Lip pruritus
PT- Alopecia
PT- Alopecia areata
HLT- Rashes, eruptions and exanthems NEC
HLT- Skin vasculitides
HLT-Pruritus NEC
HLT-Bullous conditions
HLT-Dermatitis and eczema
HLT-Dermatitis ascribed to specific agent (except Periarticular thenar erythema with onycholysis, Symmetrical drug-related intertriginous and flexural exanthema)
HLT-Erythemas
HLT-Exfoliative conditions
HLT-Granulomatous and deep cutaneous inflammatory conditions (except Annular elastolytic giant cell granuloma)
HLT-Hypopigmentation disorders
HLT-Pigmentation changes NEC (except Haemosiderin stain)
HLGT- Angioedema and urticaria
PT-Drug reaction with eosinophilia and systemic symptoms
PT Paraneoplastic dermatosis

Hepatic IRAEs

HLT-Cholestasis and jaundice
HLT-Hepatic enzymes and function abnormalities
HLT-Hepatic failure and associated disorders

HLT-Hepatocellular damage and hepatitis NEC
HLT-Liver function analyses
PT- Liver sarcoidosis
PT-Liver function test increased

Endocrine IRAEs

HLT-Adrenal cortical hypofunctions (1 new PT: Cortisol deficiency)
HLT-Adrenal cortical hyperfunctions (except PTs-Nelson's syndrome, pseudoaldosteronism, alcoholic pseudocushing's syndrome)
HLT-Anterior pituitary hypofunction (except PT-postpartum hypopituitarism) (1 new PT: Thyroid stimulating hormone deficiency)
HLT-Acute and chronic thyroiditis
HLT-Thyroid hyperfunction disorders
HLT-Thyroid hypofunction disorders
PT-Hypoparathyroidism
PT-endocrine disorder
PT-Hypophysitis
PT-Lymphocytic hypophysitis
HLGT-Endocrine investigations (incl sex hormones)
HLGT-Endocrine disorders of gonadal function
HLGT-Gonadotrophin and sex hormone changes
PT - DIABETES MELLITUS
PT - DIABETES MELLITUS INADEQUATE CONTROL
PT - DIABETES WITH HYPEROSMOLARITY
PT - INCREASED INSULIN REQUIREMENT
PT - PANCREATOGENOUS DIABETES
PT - INSULIN-REQUIRING TYPE 2 DIABETES MELLITUS
PT - LATENT AUTOIMMUNE DIABETES IN ADULTS
PT - TYPE 1 DIABETES MELLITUS
PT - TYPE 2 DIABETES MELLITUS
PT - FULMINANT TYPE 1 DIABETES MELLITUS
PT - ACQUIRED LIPOATROPHIC DIABETES
PT - GLUCOSE TOLERANCE IMPAIRED
PT - HYPERGLYCAEMIA
PT - IMPAIRED INSULIN SECRETION
PT - IMPAIRED FASTING GLUCOSE
PT - HYPOINSULINAEMIA
PT - BLOOD GLUCOSE ABNORMAL
PT - BLOOD GLUCOSE INCREASED
PT - GLUCOSE TOLERANCE DECREASED
PT-Inappropriate thyroid stimulating hormone secretion
PT-Malignant exophthalmos
PT-Premature menarche

Neurological IRAEs

PT- Cerebral sarcoidosis
PT-Meningitis
PT-Meningitis aseptic
PT-Meningism
PT-Meningitis noninfective
PT-Peripheral motor neuropathy
PT-Peripheral sensory neuropathy
PT -Autoimmune neuropathy
PT- Cranial nerve disorder

PT - Neuropathy peripheral
PT - Neuralgia
PT - Hyperaesthesia
PT - Hypoaesthesia
PT - Paraesthesia
PT - Facial paresis
PT - Facial spasm
PT - Motor dysfunction
PT - Radiculopathy
PT - Ocular myasthenia
PT- Encephalitis
PT- Noninfective encephalitis
PT- Noninfective encephalomyelitis
PT- NEUROSENSORY HYPOACUSIS
PT - ENCEPHALITIS AUTOIMMUNE
HLT-Acute polyneuropathies
HLT-Neuromuscular junction dysfunction
PT - Chronic lymphocytic inflammation with pontine perivascular enhancement responsive to steroids
PT - Morvan syndrome
PT - Autoimmune demyelinating disease
PT - Autoimmune encephalopathy
PT - Toxic leukoencephalopathy
PT-Neuromyelitis optica spectrum disorder
PT-Neurosarcoidosis
PT-Intensive care unit acquired weakness

Other IRAEs
PT-Systemic inflammatory response syndrome
PT - Myositis
PT- Necrotising myositis
PT - Polymyositis
PT-Arthritis
PT-Arthritis allergic
PT-Polyarthritis
PT- Autoimmune arthritis
PT-Hyperamylasaemia
PT-Hyperlipasaemia
PT-Optic Neuritis
PT-Episcleritis
PT-Scleritis
PT-Iritis
PT-Uveitis
PT- Intermediate uveitis
PT - Autoimmune uveitis
PT- Retinal depigmentation
PT- Iris hypopigmentation
PT- Noninfective chorioretinitis
PT- Noninfective conjunctivitis
PT- Noninfective retinitis
PT- Lung infiltration
PT-Interstitial lung disease
PT-Acute interstitial pneumonitis
PT- Haemolytic anaemia
PT- Aplastic anaemia
PT- Aplasia pure red cell

PT- Agranulocytosis
PT- Immune thrombocytopenic purpura
PT-eosinophil count increased
PT-Allergic eosinophilia
PT-Cytokine Abnormal
PT-Cytokine test
PT-Immune recovery uveitis
PT-Multiple organ dysfunction syndrome
HLGT-Allergic conditions
HLGT-Autoimmune disorders
HLGT- Immune disorders NEC (except Eosinophilic granulomatosis with polyangiitis)
HLT-Acute and chronic pancreatitis
HLT-Digestive enzymes
HLT-Nephritis NEC
HLT-Glomerulonephritis and nephrotic syndrome
HLT-Lower respiratory tract inflammatory and immunologic conditions (except PT-aspiration pneumonia)
HLT-Eosinophilic disorders,
HLT-Anaemias haemolytic immune
HLT- Noninfectious myocarditis
HLT-Arterial inflammations
PT - Idiopathic orbital inflammation
PT - Pancreatic toxicity
PT - Pancreatic enzyme abnormality
PT - Pancreatic insufficiency