Official Title of Study:

A Phase 3, Randomized Study of Adjuvant Immunotherapy with Nivolumab Combined with Ipilimumab Versus Nivolumab Monotherapy after Complete Resection of Stage IIIb/c/d or Stage IV Melanoma

(CheckMate 915: CHECKpoint pathway and nivoluMAb clinical Trial Evaluation 915)

NCT Number: NCT03068455

Document Date (Date in which document was last revised): August 23, 2019

STATISTICAL ANALYSIS PLAN
FOR CLINICAL STUDY REPORT

A PHASE 3, RANDOMIZED STUDY OF ADJUVANT IMMUNOTHERAPY WITH NIVOLUMAB COMBINED WITH IPILIMUMAB VERSUS NIVOLUMAB MONOTHERAPY AFTER COMPLETE RESECTION OF STAGE IIIB/C/D OR STAGE IV MELANOMA

PROTOCOL(S) CA209915

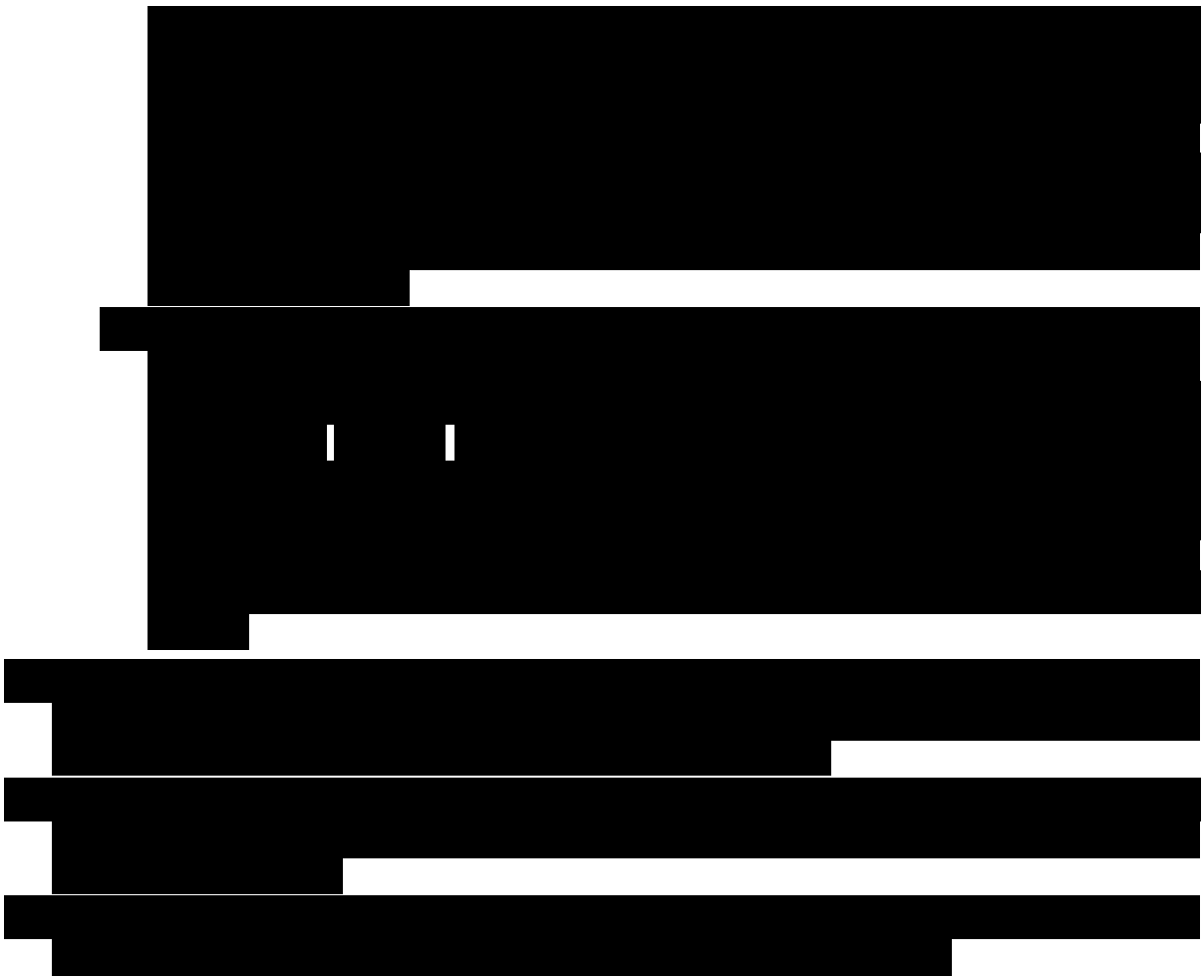
VERSION # 1.0

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

[redacted]



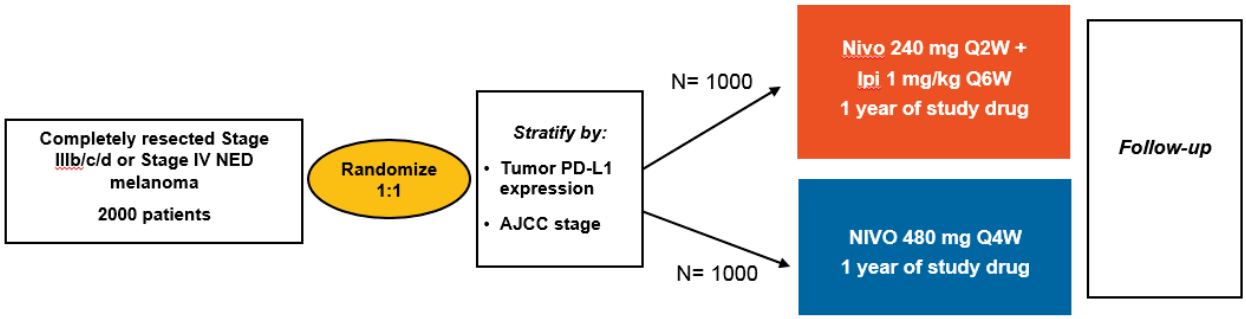
2 STUDY DESCRIPTION

2.1 Study Design

CA209915 is a Phase 3, randomized, double-blind study of nivolumab plus ipilimumab vs nivolumab monotherapy in participants (≥ 12 years) with completely resected stage IIIb/c/d or stage IV no evidence of disease (NED) melanoma.

The study design schematic is presented in Figure 2.1-1.

Figure 2.1-1: Study Design Schematic



The participants will be treated in both arms until disease recurrence, unacceptable toxicity, or participant withdrawal of consent with a maximum of 1-year total duration of study medication.

The original study design included an ipilimumab monotherapy arm. Randomization into Arm C was discontinued upon implementation of Amendment 06. Sites who have randomized participants to the ipilimumab arm were unblinded to that information and the participants were allowed to continue as open-label on either ipilimumab 10mg/kg IV Q3 weeks for 4 doses, then Q12 weeks starting at Week 24 or switch to the nivolumab 480mg IV Q4 weeks in 8 weeks after the last dose of ipilimumab.

2.2 Treatment Assignment

After the participant's initial eligibility is established and informed consent has been obtained, the participant must be enrolled into the study through IRT web to obtain the participant number. Participants meeting all eligibility criteria will be randomized in a 1:1 ratio to nivolumab + ipilimumab or nivolumab stratified by the following factors:

- PD-L1 evaluable status
  - PD-L1 expression < 1% or indeterminate,
  - PD-L1 expression 1% - < 5%
  - PD-L1 expression ≥ 5%
- AJCC Stage, 8th edition
  - stage IIIb
  - stage IIIc/d
  - stage IV

2.3 Blinding and Unblinding

Blinding of treatment assignment is critical to the integrity of this clinical study. However, in the event of a medical emergency or pregnancy in an individual participant in which knowledge of the investigational product is critical to the participant's management, the blind for that participant may be broken by the investigator. The participant's safety takes priority over any other considerations in determining if a treatment assignment should be unblinded.

Before breaking the blind of an individual participant's treatment, the investigator should determine that the unblinded information is necessary, i.e., that it will alter the participant's immediate management. In many cases, particularly when the emergency is clearly not related to the investigational product, the problem may be properly managed by assuming that the participant is receiving active product. It is highly desirable that the decision to unblind treatment assignment be discussed with the Medical Monitor, but the investigator always has ultimate authority for the decision to unblind. The Principal Investigator should only call in for emergency unblinding AFTER the decision to discontinue the participant has been made.

For this study, the method of unblinding for emergency purposes is IRT.

In cases of accidental unblinding, contact the Medical Monitor and ensure every attempt is made to preserve the blind.

Any request to unblind a participant for non-emergency purposes must be discussed with the Medical Monitor.

In case of an emergency, the investigator(s) has unrestricted access to randomization information via the Interactive Response Technology system (IRT) and is capable of breaking the blind through the IRT system without prior approval from sponsor. Following the unblinding the Investigator shall notify the medical monitor and/or study director.

Designated staff of BMS Research & Development may be unblinded prior to database lock to facilitate the bioanalytical analysis of pharmacokinetic samples and immunogenicity. A bioanalytical scientist in the Bioanalytical Sciences department of BMS Research & Development (or a designee in the external central bioanalytical laboratory) will be unblinded to the randomized treatment assignments in order to minimize unnecessary bioanalytical analysis of samples. The unblinded data will not impact the data integrity of the study. The pharmacist at the site and/or designate will be unblinded to the randomized treatment assignments in order to dispense treatment from bulk supplies, as needed. Except as noted above, other members of BMS Research and Development will remain blinded.

2.4 Protocol Amendments

Amendments incorporated in the protocol (Version 3.0, dated 11-Mar-2019) are described in in Table 2.4-1.

Table 2.4-1: Protocol Amendments

Amendments/ Revised Protocol	Date of Issue	Summary of Major Changes
Revised Protocol 03	11-Mar-2019	Updates to this protocol include, additional medical monitor, continued treatment for malignant melanoma in situ, modification of the statistical design to evaluate RFS after a sufficient period of minimum follow-up, evaluation of the surrogate endpoint Progression-Free Survival 2 (PFS2), and assessment of quality of life for 4 years in survival follow up.
Revised Protocol 02	12-Dec-2017	The purpose of this revised protocol is to modify the study design based on changes in the knowledge in the management of melanoma, and possible similarities in the pattern of efficacy by PD-L1 expression in adjuvant and metastatic settings. [redacted]
Amendment 06	20-Jul-2017	This purpose of this amendment is: • Remove treatment Arm C for Ipilimumab 10mg/kg. • Reduce infusion time for Ipilimumab from 90 minutes to 30 minutes • Clarify the duration of post-treatment contraception

2.5 Data Monitoring and Other External Committees

A Data Monitoring Committee (DMC) is established to provide oversight of safety and efficacy considerations in protocol CA209915. Additionally, the DMC will provide advice to the sponsor regarding actions the committee deems necessary for the continuing protection of participants enrolled in the study. The DMC will be charged with assessing such actions in light of an acceptable benefit/risk profile for nivolumab in combination with ipilimumab, nivolumab and ipilimumab The DMC will act in an advisory capacity to BMS and will monitor participant safety and evaluate the available efficacy data for the study. The oncology therapeutic area of BMS has primary responsibility for design and conduct of the study.

3 OBJECTIVES

3.1 Primary

To compare the efficacy, as measured by RFS, provided by nivolumab plus ipilimumab versus nivolumab monotherapy in participants with completely resected stage IIIb/c/d or stage IV no evidence of disease (NED) melanoma (in all randomized participants with PD-L1 expression level < 1%. and all randomized participants)

3.2 Secondary

- To compare the overall survival provided by nivolumab plus ipilimumab versus nivolumab monotherapy in participants with completely resected stage IIIb/c/d or stage IV NED

melanoma (in all randomized participants with PD-L1 expression level < 1% and all randomized participants)

- To evaluate the association between PD-L1 expression and RFS
- To evaluate investigator-assessed outcomes on next-line therapies

[redacted]

4 ENDPOINTS

4.1 Primary Endpoint(s)

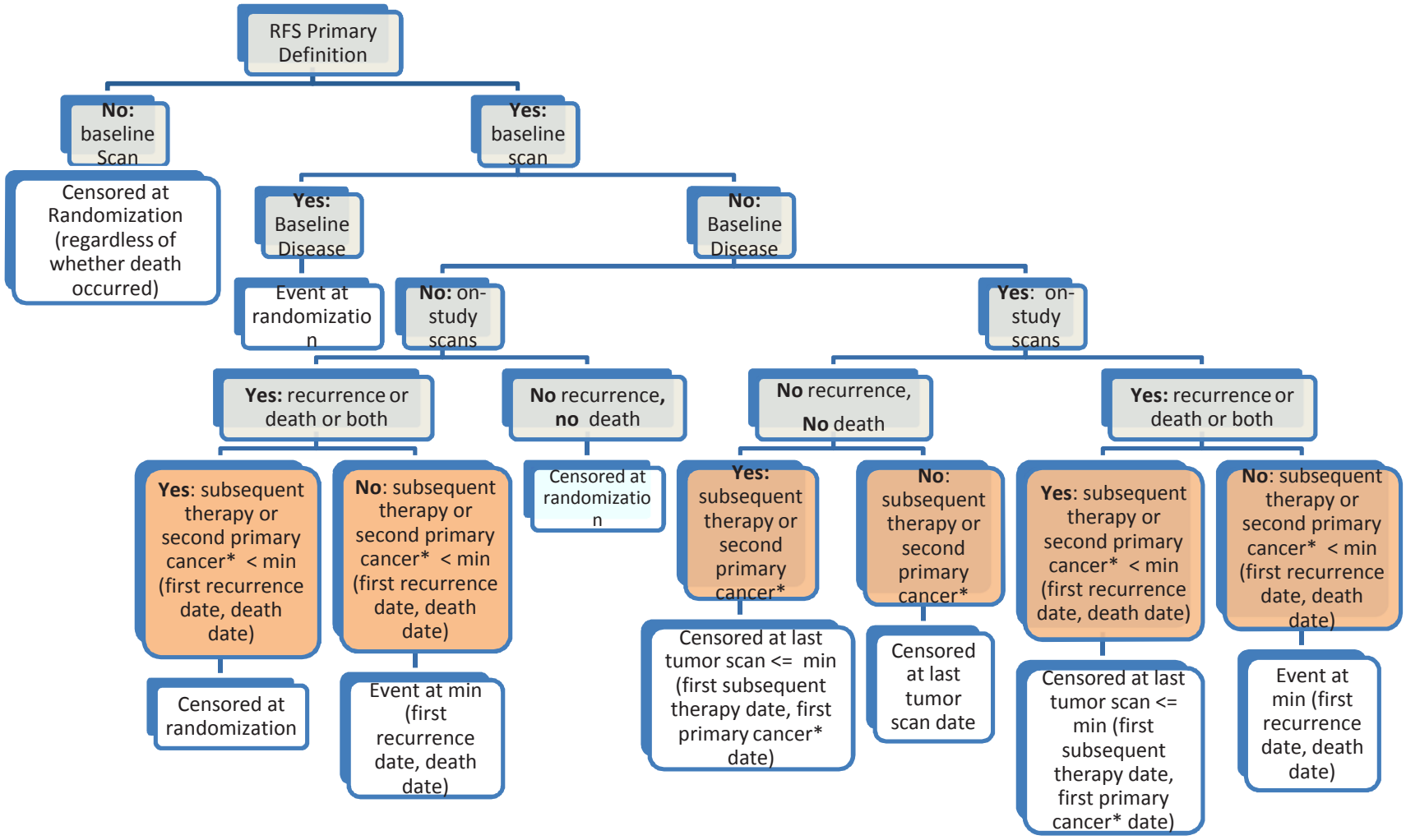
The primary endpoint is RFS. The primary endpoint of RFS will be programmatically determined based on the disease recurrence date provided by the investigator and is defined as the time between the date of randomization and the date of first recurrence (local, regional or distant metastasis), new primary melanoma (including melanoma in situ), or death (whatever the cause), whichever occurs first. A participant who dies without reported recurrence will be considered to have recurred on the date of death. For participants who remain alive and whose disease has not recurred, RFS will be censored on the date of last evaluable disease assessment. For those participants who remain alive and have no recorded post-randomization disease assessment, RFS

will be censored on the day of randomization. Censoring rules for the primary analysis of RFS are presented in Table 4.1-1 and Figure 4.1-1.

Table 4.1-1: Censoring Scheme for Primary Definition of RFS

Situation	Date of Event or Censoring	Outcome
Recurrence (local, regional, distant, new primary melanoma, including melanoma in situ)	Date of first recurrence	Event
Death without recurrence	Date of death	Event
Disease at baseline	Date of randomization	Event
No baseline disease assessment	Date of randomization	Censored
No on-study disease assessments and no death	Date of randomization	Censored
No recurrence and no death	Date of last evaluable disease assessment	Censored
New anticancer therapy, tumor-directed radiotherapy, or tumor-directed surgery received without recurrence reported prior to or on the same day of disease assessment	Date of last evaluable disease assessment prior to or on the same date of initiation of subsequent therapy	Censored
Second non-melanoma primary cancer reported prior or on the same day of disease assessment	Date of last evaluable disease assessment prior to or on the same date of diagnosis of second non-melanoma primary cancer	Censored

Figure 4.1-1: Graphic Display of RFS Primary Definition



* non-melanoma primary cancer

4.2 Secondary Endpoint(s)

4.2.1 Overall Survival

The first secondary endpoint (OS) is defined as the time between the date of randomization and the date of death, from any cause. For participants without documentation of death, OS will be censored on the last date the participant was known to be alive. Overall survival will be censored at the date of randomization for subjects who were randomized but had no follow-up.

Survival will be followed continuously while participants are on the study drug and every 3 months via in-person or phone contact after participants discontinue the study drug.

4.2.2 RFS by PD-L1 Expression Level

The second secondary objective (to evaluate the association between PD-L1 expression and RFS) will be measured by the RFS endpoint based on PD-L1 expression level.

PD-L1 expression is defined as the percent of tumor cells membrane staining in a minimum of 100 evaluable tumor cells per validated Dako PD-L1 IHC assay. This is referred to as quantifiable PD-L1 expression. If the PD-L1 staining could not be quantified, it is further classified as:

1) Indeterminate: Tumor cell membrane staining hampered for reasons attributed to the biology of the tumor tissue sample and not because of improper sample preparation or handling.
2) Not evaluable: Tumor tissue sample was not optimally collected or prepared and PD-L1 expression is neither quantifiable nor indeterminate. Not evaluable can be determined from H&E process before the tumor biopsy specimen is sent for PD-L1 evaluation or from the H&E process during PD-L1 evaluation.

Participants with missing PD-L1 expression are participants with no tumor tissue sample available for evaluation.

In order to be randomized, a participant must have PD-L1 expression available or be PD-L1 indeterminate, as determined by a central laboratory during the screening period.

PD-L1 expression will be collected in the Interactive Response Technology (IRT) system as well as in the clinical database. Statistical analysis using PD-L1 expression will be solely based on PD-L1 expression data from clinical database, except for when we would perform a stratified analysis, where the PD-L1 stratification factor would be used as collected in the IRT.

Evaluation of different thresholds for PD-L1 positivity will also be assessed as exploratory analyses, at 10% tumor cell expression cut-off.

4.2.3 Investigator-Assessed Outcomes on Next-Line Therapies

Investigator outcomes on subsequent systemic therapies are defined on two levels:

- the primary definition takes into account the regimen,

- the secondary definition takes into account the line of therapy.

Progression-free survival on next-line systemic therapy is defined as:

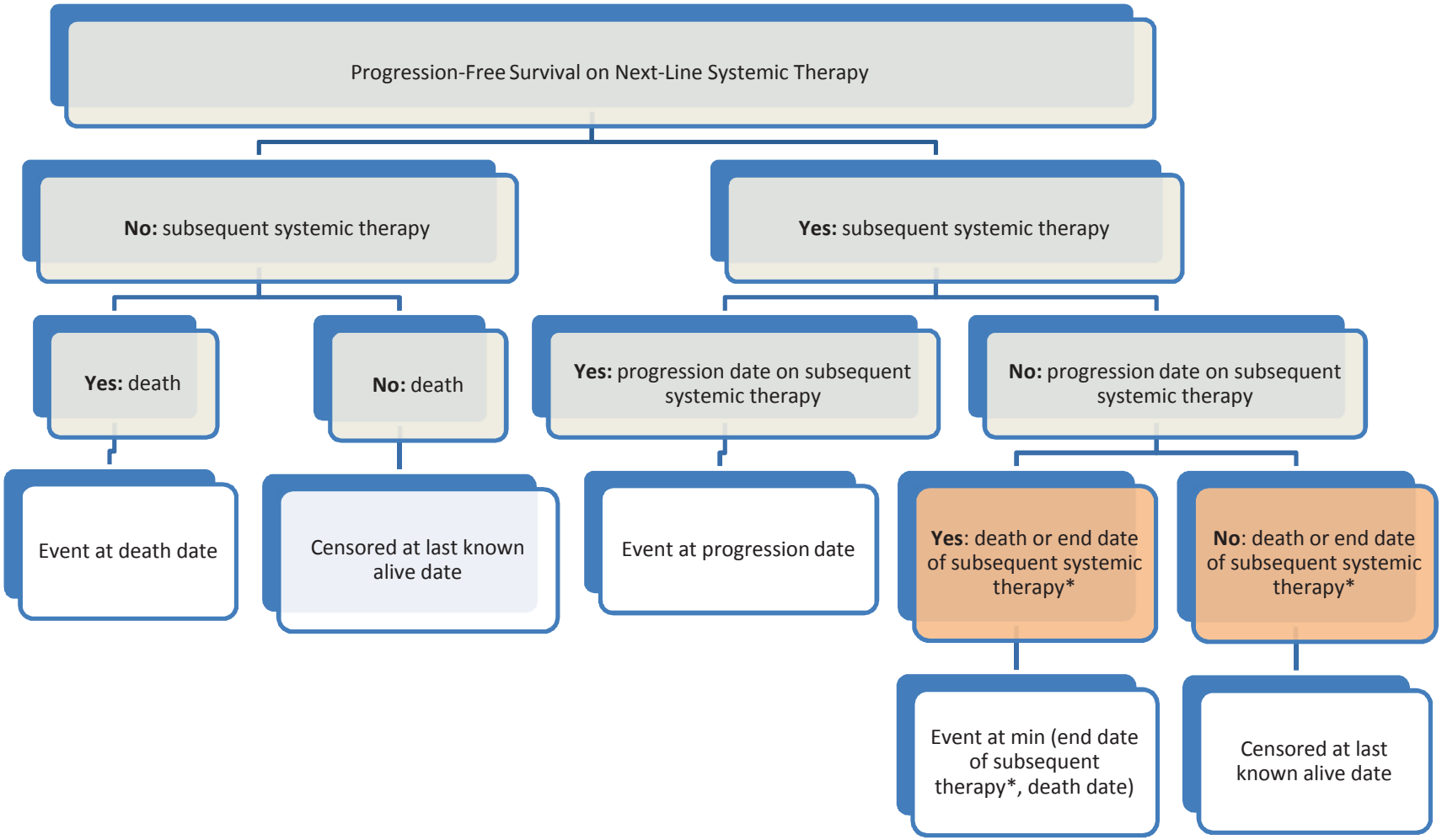
Time from randomization to:

- in case of event (i.e., progression date on next subsequent systemic therapy available and/or end date of next subsequent systemic therapy available and/or death):
  - progression date on next subsequent systemic therapy, if available
  - else minimum (end date of next subsequent systemic therapy OR death from any cause)
- in case of no event (censored, i.e., 1) no subsequent systemic therapy and no death or 2) subsequent systemic therapy but no progression date and no end date available and no death): last known alive date

In case no end date of next systemic therapy was reported and a second next systemic therapy was administered, the next systemic therapy is assumed to have ended on the day before the second next systemic therapy was started.

Progression-free survival on next-line systemic therapy (in months) = (date as defined above - randomization date + 1) / 30.4375.

Figure 4.2.3-1: Graphic Display of Progression-Free Survival on Next-Line Systemic Therapy



* In case no end date of next systemic therapy was reported and a second next systemic therapy was administered, the next systemic therapy is assumed to have ended on the day before the second next systemic therapy was started.

Time to next treatment is defined as the time from the date of randomization to the start date of next systemic therapy. Participants who do not receive next treatment will be censored at the last known alive date.

Time to second next treatment is defined as the time from the date of randomization to the start date of second next systemic therapy. Participants who do not receive second next treatment will be censored at the last known alive date.

Time from next treatment to second next treatment is defined as the time from the start date of next systemic therapy to start date of second next systemic therapy. No censoring rules apply here as analysis will only be performed on subset of participants who received second next treatment.

[redacted]

[redacted]

5 SAMPLE SIZE AND POWER

The sample size of the study is based on a comparison of the RFS distribution between participants randomized to nivolumab plus ipilimumab and participants randomized to nivolumab. RFS will be evaluated for treatment effect using the following testing strategy: RFS will be compared first in the all randomized participants with PD-L1 expression level < 1% subgroup with an alpha allocation of 0.03 (two-sided); and if significant, the alpha allocated to this subgroup will be recycled to the treatment comparison in the overall population (all randomized participants). Thus RFS will be compared in all randomized participants with an alpha allocation of 0.05 (two-sided). If the treatment comparison in all randomized participants with PD-L1 expression level < 1% is not significant, RFS will be compared in all randomized participants with an alpha allocation based on the method published by Spiessens and Debois[1]. This method takes into account the correlation between the test statistics for the overall and the subgroup analysis, and is essentially the same as the method used when dealing with interim analyses, i.e., group sequential method. The alpha allocated to the overall analysis will be calculated based on the fraction of information in the subgroup, relative to the overall population, which is determined by the ratio of events observed in the two groups. For example, if 257 events are observed in the PD-L1 expression level < 1% subgroup, and 651 events are observed in the overall population (i.e., about 40% of all events in the PD-L1<1% subgroup), the alpha allocated to the overall analysis is 0.0265.

Simulation models incorporating aspects of immunotherapy like delayed separation (observed as late separation of survival curves between the experimental and SOC arms) and long term survival benefits (observed as a long-lasting plateau towards the tail of the survival curve) were developed

for sample size estimation. Sample size calculations for this study design were done using EAST 6 (v 6.3.1) and R.

The sample size is driven by the comparison of RFS between all randomized participants with PD-L1 expression level < 1% randomized to receive nivolumab plus ipilimumab vs. nivolumab. Approximately 600 participants with PD-L1 expression level < 1% will be randomized in a 1:1 ratio to nivolumab plus ipilimumab and nivolumab respectively. Approximately 257 RFS events in all randomized participants with PD-L1 expression level < 1% will provide approximately 90% power to detect an average hazard ratio (HR) of 0.65 with an alpha level of 0.03 (two-sided). In case the 257 RFS events would be reached at a timepoint earlier than 13 months of minimum follow-up (from the randomization of the last participant), analysis of RFS will be conducted when there would be a minimum follow up of 13 months (i.e., at the time when approximately 257 events are reached or a minimum follow-up of 13 months, whatever occurs later).

Accrual will be stopped when approximately 600 participants with PD-L1 expression level < 1% will be randomized in a 1:1 ratio to nivolumab plus ipilimumab and nivolumab respectively. The prevalence of participants with PD-L1 expression level < 1% is expected to be around 30% of all randomized participants. Therefore, it is estimated that approximately 2000 participants will be randomized in a 1:1 ratio to nivolumab plus ipilimumab and nivolumab respectively. No more than 1400 participants with PD-L1 expression level ≥ 1% will be randomized to nivolumab plus ipilimumab and nivolumab respectively. Approximately 700 participants with PD-L1 expression level < 1% were randomized in a 1:1 ratio to nivolumab plus ipilimumab and nivolumab respectively.

Approximately 651 RFS events in all randomized participants will provide approximately 90% power to detect an average hazard ratio (HR) of 0.76 with an alpha level of 0.0265 (two-sided).

The analysis of RFS in the all randomized participants with PD-L1 expression level < 1% subgroup will be performed at the time when approximately 257 RFS events are observed in this subgroup or a minimum follow-up of 13 months is reached, whatever occurs later.

If the treatment comparison in the all randomized participants with PD-L1 expression level < 1% subgroup is significant, RFS will be compared in all randomized participants with an alpha allocation of 0.05. In this case, two analyses, an interim and a final, will be conducted in all randomized participants. The interim analysis will take place at the same time as the RFS analysis for the PD-L1 expression level < 1% subgroup. The final analysis will take place when approximately 560 RFS events are observed in the overall population. The stopping boundaries at the interim and final analyses will be derived based on the exact number of RFS events using Lan-DeMets alpha spending function with O'Brien-Fleming boundaries.

If the treatment comparison in all randomized participants with PD-L1 expression level < 1% subgroup is not significant, RFS will be compared in the overall population with an alpha allocation based on the method published by Spiessens and Debois[1]. In this case, no interim analysis is planned. The final RFS analysis for the overall population will be conducted when

approximately 651 RFS events are observed (assuming allocated alpha is 0.0265). In case the events come slower than projected, final analysis of RFS for the overall population will be conducted when the minimum follow up from randomization of the last participant reaches 30 months.

Given actual accrual rates so far and estimated accrual rates of 157 participants per month for the coming months, it is estimated that accrual of 2000 participants (i.e., 1000 participants in a 1:1 ratio in the nivolumab plus ipilimumab and nivolumab arms each) and the participants in the ipilimumab arm that has already taken place, will take approximately 17 months. Given actual accrual rates, accrual took approximately 13.5 months.

5.1 RFS

The primary objective of the study is to compare RFS of nivolumab plus ipilimumab to nivolumab monotherapy in participants with completely resected stage IIIb/c/d or stage IV no evidence of disease (NED) melanoma. The number of events and power were calculated assuming a non-proportional hazards model with a 3-month delayed treatment effect and a cure rate in each of the treatment arms.

5.1.1 RFS in All Randomized Participants with PD-L1 Expression Level < 1%

For this comparison of RFS between nivolumab plus ipilimumab and nivolumab in all randomized participants with PD-L1 expression level < 1%, at least 257 RFS events would be required with PD-L1 expression level < 1% in the two respective treatment arms for a two-sided experiment-wise alpha = 0.03 log-rank test, to show a statistically significant difference in RFS between the treatment arms with at least 90.0% power when the average hazard ratio (HR) of the nivolumab plus ipilimumab arm to the nivolumab arm is 0.65. A cure rate of 0.42 is assumed in the nivolumab treatment arm. Under the assumptions for accrual and RFS distribution as per CA209238, Section 7.3 of the Clinical Study Report, nivolumab estimates and assumed HR as stated above, it would take approximately 29 months from the randomization of the first participant to observe the required number of RFS events. It is projected that an observed HR of 0.76 or less would result in a statistically significant improvement at the final analysis of RFS.

5.1.2 RFS in All Randomized Participants

For this comparison of RFS between nivolumab plus ipilimumab and nivolumab in all randomized participants, at least 651 RFS events would be required in the two respective treatment arms for a two-sided alpha =0.0265 log-rank test, to show a statistically significant difference in RFS between the treatment arms with approximately 90% power when the average hazard ratio (HR) of the nivolumab plus ipilimumab arm to the nivolumab arm is 0.76. A cure rate of 0.49 is assumed in the nivolumab treatment arm. Under the assumptions for accrual and RFS distribution as per CA209238, Section 7.3 of the Clinical Study Report, nivolumab estimates and assumed HR as stated above, it would take approximately 44 months from the randomization of the first participant

to observe the required number of RFS events. It is projected that an observed HR of 0.84 or less would result in a statistically significant improvement at the final analysis of RFS.

For this comparison of RFS between nivolumab plus ipilimumab and nivolumab in all randomized participants, at least 560 RFS events would be required in the two respective treatment arms for a two-sided alpha = 0.05 log-rank test, to show a statistically significant difference in RFS between the treatment arms with approximately 90% power when the average hazard ratio (HR) of the nivolumab plus ipilimumab arm to the nivolumab arm is 0.76. A cure rate of 0.49 is assumed in the nivolumab treatment arm. Under the assumptions for accrual and RFS distribution as per CA209238, Section 7.3 of the Clinical Study Report, nivolumab estimates and assumed HR as stated above, it would take approximately 30 months from the randomization of the first participant to observe the required number of RFS events. It is projected that an observed HR of 0.85 or less would result in a statistically significant improvement at the final analysis of RFS.

5.2 Overall Survival

The secondary objective of the study is to compare OS of nivolumab plus ipilimumab to nivolumab monotherapy in participants with completely resected stage IIIb/c/d or stage IV no evidence of disease (NED) melanoma. The number of events and power were calculated assuming a non-proportional hazards model with a cure rate in each of the treatment arms.

5.2.1 OS in All Randomized Participants with PD-L1 Expression Level < 1%

For this comparison of OS between nivolumab plus ipilimumab and nivolumab in all randomized participants with PD-L1 expression level < 1%, at least 250 deaths would be required in the two respective treatment arms for a two-sided experiment-wise alpha =0.05 adjusted log-rank test, to show a statistically significant difference in OS between the treatment arms with at least 80% power when the average hazard ratio (HR) of the nivolumab plus ipilimumab arm to the nivolumab arm is 0.70. It is projected that an observed HR of 0.78 or less would result in a statistically significant improvement at the final analysis of OS.

5.2.2 OS in All Randomized Participants

For this comparison of OS between nivolumab plus ipilimumab and nivolumab in all randomized participants, at least 630 deaths would be required in the two respective treatment arms for a two-sided experiment-wise alpha =0.05 adjusted log-rank test, to show a statistically significant difference in OS between the treatment arms with at least 80% power when the average hazard ratio (HR) of the nivolumab plus ipilimumab arm to the nivolumab arm is 0.80. It is projected that an observed HR of 0.85 or less would result in a statistically significant improvement at the final analysis of OS.

For this comparison of OS between nivolumab plus ipilimumab and nivolumab in all randomized participants, at least 753 deaths would be required in the two respective treatment arms for a two-sided experiment-wise alpha=0.0265 adjusted log-rank test, to show a statistically significant

difference in OS between the treatment arms with at least 80% power when the average hazard ratio (HR) of the nivolumab plus ipilimumab arm to the nivolumab arm is 0.80. It is projected that an observed HR of 0.85 or less would result in a statistically significant improvement at the final analysis of OS.

6 STUDY PERIODS, TREATMENT REGIMENS AND POPULATIONS FOR ANALYSES

6.1 Study Periods

6.1.1 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. Events (AEs) on the same date and time of the first dose of study treatment will not be considered as pre-treatment events. For participants who are randomized but not treated, baseline evaluation or events will be defined as those that occur before the date and time of randomization.

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 and vital signs) 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 at the same date (and time if collected), the assessment with the latest database entry date (and time if collected) will be considered as baseline.

If more than one tumor biopsy specimen is available, baseline PD-L1 expression is determined from the most recently collected specimen before first dose date (randomization date in case the participant is randomized but not treated) with a quantifiable result. If all specimens for a given participant are either indeterminate or not evaluable, then the PD-L1 expression will be considered indeterminate as long as at least one specimen is indeterminate. Otherwise, PD-L1 expression will be considered not evaluable.

6.1.2 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 participants who are off study treatment, AEs will be included if the event occurred within a safety window of 30 days (or 100 days depending

on the analysis) after the last dose of study treatment. No "subtracting rule" will be applied when an AE occurs both pre-treatment and post-treatment with the same preferred term and grade.

On-treatment evaluations (laboratory tests, pulse oximetry, and vital signs) will be defined as evaluations taken after the day (and time, if collected and not missing) of first dose of study treatment. For participants who are off study treatment, evaluations should be within a safety window of 30 days (or 100 days depending on the analysis) after the last dose of study treatment.

Late emergent drug-related AEs will be defined as drug-related AEs with an onset date greater than 100 days after the last dose of study treatment in participants who are off study treatment.

6.2 Treatment Regimens

The **treatment group "as randomized"** corresponds to the treatment group assigned by the IRT.

The **treatment group "as treated"** will be the same as the treatment group "as randomized" by IRT unless a participant received the incorrect study treatment for the entire period of treatment, in which case the participant's treatment group "as treated" will be defined as the incorrect study treatment.

Unless otherwise specified, the safety analysis will be based on the treatment group "as treated".

Unless otherwise specified, the efficacy analysis will be based on the treatment group "as randomized".

6.3 Populations for Analyses

Table 6.3-1: Populations for Analyses

Population	Description
All enrolled participants	All participants who signed an informed consent form and were registered into the IRT.
All randomized participants	All participants who were randomized to any treatment arm in the study.
All randomized participants with PD-L1 expression level < 1%	All randomized participants with baseline tumor sample testing per IHC result < 1% membranous staining in tumor cell.
All treated participants	All participants who received at least one dose of study drug.
All treated participants with PD-L1 expression level < 1%	All participants with baseline tumor sample testing per IHC result < 1% membranous staining in tumor cell who received at least one dose of study drug.
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]

Table 6.3-1: Populations for Analyses

Population	Description
[redacted]	[redacted]
All PD-L1 tested participants	All randomized participants who had a tumor biopsy specimen assessed for PD-L1 expression.
All PD-L1 evaluable participants	All PD-L1 tested participants with quantifiable PD-L1 expression.
[redacted]	[redacted]

7 STATISTICAL ANALYSES

7.1 General Methods

Unless otherwise noted, the following subsections describe tabulations of discrete variables, by the frequency and proportion of participants falling into each category, grouped by treatment (with total). Percentages given in these tables will be rounded 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.

Time to event distribution will be estimated using Kaplan Meier techniques. This will be done for endpoints of recurrence-free survival, overall survival, and distant metastasis-free survival. Median survival time along with 95% CI will be constructed based on a log-log transformed CI for the survivor function S(t)[2,3]. Rates at fixed time points will be derived from the Kaplan Meier estimate and corresponding confidence interval will be derived based on Greenwood formula[4] for variance derivation and on log-log transformation applied on the survivor function S(t)[5].

Unless otherwise specified, the hazard ratio of Nivolumab plus Ipilimumab to Nivolumab, and its associated CI, will be obtained by fitting a stratified Cox model with the treatment group variable as the sole covariate using stratification factor information recorded in the IRT.

P-values from sensitivity analyses for efficacy endpoints are for descriptive purpose only and there will be no multiplicity adjustment for these analyses.

Unless otherwise specified, analyses will be presented for all participants with PD-L1 expression level < 1% as well as for all participants.

For participants in the ipilimumab arm, their data will be included in the main tables for disposition presenting "all enrolled participants" and will be presented separately otherwise. The following data will be summarized for participants in the ipilimumab arm:

- Study conduct:

- accrual

- Study population:
  - participant disposition,
  - baseline and disease characteristics
  - prior therapy: prior radiotherapy, surgery and systemic therapy
- Extent of exposure:
  - number of doses, cumulative dose and relative dose intensity
  - duration of therapy,
  - subsequent cancer therapy
- Efficacy:
  - KM plot of RFS, [redacted], OS,
  - RFS, [redacted], OS rates
  - efficacy by PD-L1
  - outcome on subsequent therapy
- safety:
  - from the first dose of blinded ipilimumab or first dose of open-label ipilimumab or nivolumab (for subjects not dosed with blinded ipilimumab) to 100 days post last dose of blinded ipilimumab (for subjects not entering the open-label phase) or 100 days post last dose of open-label ipilimumab or nivolumab (for subjects entering the open-label phase)
  - AEs / drug-related AEs
  - SAEs / drug-related SAEs
  - AEs /drug-related AEs leading to discontinuation
  - Deaths

7.1.1 Adverse Events, Serious Adverse Events, Multiple events, Select Adverse Events, Other Events of Special Interest and Immune-Mediated Adverse Events

Drug-related AEs are those events with relationship to study drug "Related", as recorded on the CRF. If the relationship to study drug is missing, the AE will be considered as drug-related.

Serious adverse events consist of AEs deemed serious by the Investigator and flagged accordingly in the CRF and clinical database.

Adverse events leading to study drug discontinuation are AEs with action taken regarding study drug(s) = "Drug was discontinued".

Adverse events leading to dose delay are AEs with action taken regarding study drug(s) = "Drug was delayed".

Adverse events leading to dose reduction are AEs with action taken regarding study drug(s) = "Dose was reduced".

Adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA), and the most recent version of the dictionary at the time of the database lock will be used. Adverse events results will be graded for severity using NCI Common Terminology Criteria for Adverse Events (CTCAE) and the most recent version of the criteria at the time of the database lock will be used.

In the AE summary tables, unless otherwise specified, subjects will be counted only once at the Preferred Term (PT), only once at the System Organ Class (SOC), and only once at subject level for the counting of total number of subjects with an AE. The AE tables will be sorted by the SOCs and then PTs. SOC will be ordered by descending frequency overall and then alphabetically. PTs will be ordered within SOC by descending frequency overall and then alphabetically. The sorting will be done based on the 'Any Grade' column of the experimental arm when arms are presented side-by-side.

Unless otherwise specified, the AE summary tables will be restricted to on-treatment events regardless of the causality.

Analyses that take into account the multiple occurrences of a given adverse event will be conducted (see Section 7.6.9). To prepare these analyses, the CRF data will be processed according to standard BMS algorithms[6] in order to collapse adverse event records into unique records based on the preferred term. These data will be presented as the rate per 100 person-years of exposure. These analyses will take into account all on-treatment events (allowing more than 1 event per subject) and the total exposure time. The person-year exposure will be computed as the sum over the subjects' exposure expressed in years where the exposure time is defined as

- (Date of last dose of study treatment - date of first dose of study treatment + 31 days)/365.25, for subject who are off study treatment and were followed for at least 30 days (or 100 days, depending on the analysis) after last dose of study treatment.
- (Last known alive date - date of first dose of study treatment +1)/365.25, for subjects who are still on-treatment or who are off study treatment and were followed less than 30 days after last dose of study treatment.

7.1.1.1 Select Adverse Events (EU Submission)

The select Adverse Events (select AEs) consist of a list of preferred terms grouped by specific category (e.g., pulmonary events, gastrointestinal events categories, etc.). AEs that may differ from or be more severe than AEs caused by non-immunotherapies and AEs whose early recognition and management may mitigate severe toxicity are included as select AEs. Categories of select AEs may include subcategories (e.g., adrenal disorders, diabetes, pituitary disorders, and thyroid disorders are subcategories of the endocrine event category).

The list of MedDRA preferred terms used to identify select adverse events is revisited quarterly and updated accordingly. The preferred terms used for the selection at the time of the database lock will be provided by categories/subcategories.

In addition to the frequency and worst severity of select AEs, time-to onset, time-to resolution, and time-to resolution where immune modulating medication was initiated will be analyzed for each specific category/subcategory of drug-related select AEs when applicable.

Further details on the definitions of select adverse event, time-to onset and time-to resolution are described in APPENDIX 1.

7.1.1.2 Other Events of Special Interest

Other events of special interest (OEOSI) consist of a list of preferred terms grouped by specific category (e.g. Myositis Event, Myocarditis Event, Demyelination Event, Guillain-Barre Syndrome, Pancreatitis Event, Uveitis Event, Encephalitis Event, Myasthenic Syndrome, Rhabdomyolysis Event, Graft Versus Host Disease). The list of MedDRA preferred terms used to identify OEOSI is revisited quarterly and updated accordingly. The preferred terms used for the selection at the time of the database lock by categories will be provided.

7.1.1.3 Immune-Mediated Adverse Events (US Submission)

In order to further characterize AEs of special clinical interest, analysis of immune-mediated AEs (IMAE) will be conducted. IMAEs are specific events (or groups of PTs describing specific events) that include pneumonitis, diarrhea/colitis, hepatitis, nephritis/renal dysfunction, rash, endocrine (adrenal insufficiency, hypothyroidism/thyroiditis, hyperthyroidism, diabetes mellitus, and hypophysitis), and other specific events, considered as potential immune-mediated events by investigator that meet the definition summarized below:

- those occurring within 100 days of the last dose,
- regardless of causality,
- treated with immune-modulating medication (of note, endocrine AEs such as adrenal insufficiency, hypothyroidism/thyroiditis, hyperthyroidism, diabetes mellitus, and hypophysitis are considered IMAEs regardless of immune-modulating medication use, since endocrine drug reactions are often managed without immune-modulating medication).
- with no clear alternate etiology based on investigator assessment, or with an immune-mediated component

The list of MedDRA preferred terms used to identify IMAEs is revisited quarterly and updated accordingly. The preferred terms used for the selection at the time of the database lock by categories will be provided.

7.1.2 Laboratory Tests

Clinical laboratory parameters (hematology, serum chemistry and electrolytes) will be evaluated.

Laboratory tests will be graded using the NCI Common Terminology Criteria, and the most recent version of the criteria at the time of the database lock will be used.

Clinical laboratory data will be first analyzed using International System of Units (SI).

Analyses will be repeated using US conventional units.

In the laboratory summary tables, unless otherwise specified, subjects will be counted only once for each lab parameter according to their worst on treatment CTC grade (worst being the highest CTC grade). The laboratory tables and listings will be sorted by laboratory category, laboratory subcategory and laboratory test code sequence number.

7.1.3 Immunogenicity Data

Blood samples for immunogenicity analysis will be collected from subjects assigned to the experimental treatment group(s) according to the protocol schedule. Samples will be evaluated for development of Anti-Drug Antibody (ADA) by a validated electrochemiluminescent (ECL) immunoassay.

7.2 Study Conduct

7.2.1 Accrual

The following will be presented on the enrolled participants population (not split by participants with PD-L1 expression level < 1% and all participants):

- Number of participants accrued by country and investigational site
- Number of participants accrued by month

A by-participant listing of accrual will be produced.

7.2.2 Relevant Protocol Deviations

The following programmable deviations will be considered as relevant protocol deviations and be summarized by treatment group and overall for all randomized participants. Non-programmable relevant eligibility and on-treatment protocol deviations, as well as significant (both programmable and non-programmable) eligibility and on-treatment protocol deviations will be reported through ClinSIGHT listings.

<u>At Entrance</u>:

- No histologically documented stage IIIb or stage IIIc or stage IIId or stage IV melanoma as per AJCC staging
- Documented/confirmed disease at baseline
- Participant with baseline ECOG performance status > 1
- The last intervention demonstrating that the participant is free of disease is more than 13 weeks prior to randomization
- Participant received prior systemic anti-cancer therapy (prior treatment with adjuvant interferon is allowed if completed > 6 months prior to randomization)
- Participant with ocular/uveal melanoma

On-study:

- Participants receiving systemic anti-cancer therapy (chemotherapy, immunotherapy, standard or investigational agents for treatment of cancer) while on study therapy
- Participants treated differently than as randomized (participants who received the wrong treatment, excluding the never treated)

A by participant listing will be produced.

7.3 Study Population

7.3.1 Participant Disposition

The total number of participants enrolled (randomized or not) will be presented along with the reason for not being randomized. This summary will be on all enrolled participants pooled together (and not split by participants with PD-L1 expression level < 1% and all participants).

Number of participants randomized but not treated along with the reason will be tabulated by treatment group as randomized for all randomized participants.

Number of participants who discontinued study treatment along with corresponding reason will be tabulated by treatment group as treated for all treated participants. Reason for discontinuation will be derived from the participant status CRF page. For participants in the ipilimumab 10mg/kg group who were further treated with open-label ipilimumab or nivolumab, number of participants who discontinued open-label study treatment along with corresponding reason will be tabulated by open-label treatment group.

A participant listing for all treated participants will be provided showing the participant's randomization date, first and last dosing date, off study treatment date and reason for going off-study treatment. A participant listing for participants not randomized will also be provided, showing the participant's race, gender, age, consent date and reason for not being randomized.

7.3.2 Demographics and Other Baseline Disease Characteristics

Descriptive statistics of the following baseline characteristics will be summarized for all randomized participants.

- Age (descriptive statistics, years);
- Age category I (< 65, ≥ 65)
- Age category II (< 65, ≥ 65 - < 75, ≥ 75)
- Gender (male, female)
- Race (white, black, asian, other)
- Region (US and Canada, Western Europe, Eastern Europe, Australia, ROW)
- Baseline ECOG PS (0, 1, 2, ...)
- Baseline LDH (≤ ULN, > ULN)
- Baseline LDH (≤ 2 x ULN, > 2 x ULN)

- Baseline weight (descriptive statistics, kg)
- PD-L1 status (< 1%, ≥ 1%, indeterminate) (source: clinical database)
- PD-L1 status (< 5%, ≥ 5%, indeterminate) (source: clinical database)
- PD-L1 status (< 10%, ≥ 10%, indeterminate) (source: clinical database)
- BRAF mutation status (Wildtype, Mutant)
- Time from Surgical Resection to Randomization (in weeks):
  - descriptive statistics
  - < 3, 3 - < 6, 6 - < 9, 9 - < 12, 12 - < 15, 15 - < 18, 18 - < 21, ≥ 21
- Completion lymph node dissection (CLND) (no, yes)
- Disease Stage (Stage IIIb, Stage IIIc, Stage IIId, Stage IV: source: eCRF)
- M-status in Stage IV participants (M1a, M1b, M1c, M1d: Source: eCRF)
- Equivocal lymph nodes present (no, yes)
- Tumor origin (primary, recurrent)
- Location of primary tumor
- Melanoma sub-type (mucosal, cutaneous, acral, ocular/uveal, other)
- Tumor thickness (NO EVIDENCE OF PRIMARY MELANOMA, <0.8 mm, 0.8-1.0 mm, >1.0-2.0 mm, >2.0-4.0 mm, >4.0 mm) in Stage III participants
- Tumor ulceration status (present, absent, unknown) in Stage III participants
- Lymph node involvement I (clinically occult only, clinically detected only, clinically detected and clinically occult) in Stage III participants
- Lymph node involvement II (clinically occult only, clinically detected only + clinically detected and clinically occult) in Stage III participants
- Tumor ulceration status by lymph node involvement I (present + clinically occult only, present + clinically detected only, present + clinically detected and clinically occult, absent + clinically occult only, absent + clinically detected only, absent + clinically detected and clinically occult) in Stage III participants
- Tumor ulceration status by lymph node involvement II (present + clinically occult only, present + clinically detected only + clinically detected and clinically occult, absent + clinically occult only, absent + clinically detected only + clinically detected and clinically occult) in Stage III participants
- Total number of tumor involved regional lymph nodes (0, 1, 2-3, 4+) in Stage III participants
- Number of Clinically Occult Nodes (0, 1, 2-3, 4+) in Stage III participants
- Number of Clinically Detected Nodes (0, 1, 2-3, 4+) in Stage III participants
- Classification of Nodes (IN-TRANSIT/SATELLITE/MICRO-SATELLITE METS PRESENT WITH NO TUMOR-INVOLVED NODES, IN-TRANSIT/SATELLITE/MICRO-SATELLITE METS PRESENT WITH TUMOR-INVOLVED NODES, MATTED NODES, NOT APPLICABLE) in Stage III participants

Similarly the following IRT data will be summarized by treatment group as randomized for all randomized participants.

- Disease Stage (stage IIIb vs stage IIIc/d vs stage IV)
- PD-L1 status (< 1% or indeterminate, 1% - < 5%, ≥ 5%)

7.3.3 Medical History

General medical history will be listed by participant for all randomized participants.

Pre-treatment events will be summarized by treatment group for all randomized participants.

7.3.4 Prior Therapy

The following will be summarized by treatment group for all randomized participants:

- Prior radiotherapy (yes, no)
- Prior systemic therapy (yes, no)
- Prior surgery (yes, no)
- Prior/current non-study medication classified by anatomic and therapeutic classes. Agents and medication will be reported using the generic name.

Listings by participants will also be provided.

7.3.5 Baseline Examinations

Participants with abnormal baseline physical examination will be tabulated by examination criteria and by treatment group for all randomized participants.

7.3.6 Discrepancies Between IRT and CRF Information

Summary tables (cross-tabulations) of stratification factors for all randomized participants by treatment group will be provided to show any discrepancies between what was reported through IRT vs. CRF data or clinical database (baseline).

- M Stage at Study Entry (IRT vs. CRF data)
- PD-L1 status (IRT vs. clinical database)

7.4 Extent of Exposure

Analyses will be performed by treatment group “as treated” in all treated participants, unless otherwise specified.

7.4.1 Administration of Study Therapy

The following parameters will be summarized (descriptive statistics) by treatment group:

- Time from randomization to first dose of study therapy (days; 0 to 3, > 3 to 7, > 7 to 14, > 14 to 21, > 21 to 28, > 28)
- Number of doses received (nivolumab, ipilimumab, placebo)
  - summary statistics
  - 1, 2, 3, ...
- Cumulative dose in mg/kg and/or mg (nivolumab, ipilimumab)
- Relative dose intensity (%) using the following categories: < 50%; 50 - < 70%; 70 - < 90%; 90 - < 110%; ≥ 110% (nivolumab, ipilimumab)
- Duration of study therapy (in months):
  - duration of treatment will be presented using a Kaplan-Meier curve whereby the last dose date will be the event date for participants who discontinued study therapy. Participants who are still on study therapy will be censored on their last dose date. Median duration of treatment and associated 95% CI will be provided.
  - duration of treatment will also be summarized in a table with descriptive statistics (mean, median, minimum, and maximum). The percentage of participants with study therapy duration > 3 months, > 6 months, > 9 months, and > 12 months will be tabulated.

A by-participant listing of dosing of study medication (record of study medication, infusion details, and dose changes) and a listing of batch numbers will be also provided.

For participants in the ipilimumab 10mg/kg arm, the entire study therapy duration from first dose of blinded ipilimumab or first dose of open-label ipilimumab or nivolumab (for subjects not dosed with blinded ipilimumab) to last dose of blinded ipilimumab (for subjects not entering the open-label phase) or last dose of open-label ipilimumab or nivolumab (for subjects entering the open-label phase) will also be plotted:

- duration of treatment will be presented using a Kaplan-Meier curve whereby the last dose date will be the event date for participants who discontinued study therapy. Participants who are still on study therapy will be censored on their last dose date. Median duration of treatment and associated 95% CI will be provided.

For participants in the ipilimumab 10mg/kg group who were further treated with open-label ipilimumab or nivolumab, the following parameters will be summarized (descriptive statistics):

- Number of doses received (open-label nivolumab, open-label ipilimumab)
  - summary statistics
  - 1, 2, 3, ...
- Cumulative dose in mg/kg and/or mg (open-label nivolumab, open-label ipilimumab)
- Relative dose intensity (%) using the following categories: < 50%; 50 - < 70%; 70 - < 90%; 90 - < 110%; ≥ 110% (open-label nivolumab, open-label ipilimumab)

- Duration of open-label study therapy (in months):
  - duration of open-label treatment will be presented using a Kaplan-Meier curve whereby the last dose date will be the event date for participants who discontinued study therapy. Participants who are still on study therapy will be censored on their last dose date. Median duration of treatment and associated 95% CI will be provided.

Table 7.4.1-1: Administration of Study Therapy: Definition of Parameters - Nivolumab Arm

	Nivolumab	**Nivolumab**
Dosing schedule per protocol	**480mg every 4 weeks**	**6mg/kg every 4 weeks (for adolescents between 12 and < 18 years)**
Dose	*Dose (mg) is defined as: the nominal dose (mg) * total volume infused (mL) /total volume prepared (mL) nominal dose = 480mg*	*Dose (mg/kg) is defined as: the minimum of [vial strength (mg/mL) *total volume infused (mL)] /most recent weight (kg) and nominal dose (mg/kg) * total volume infused (mL) /total volume prepared (mL) Vial strength = 10mg/mL Dose (mg) is defined as dose (mg/kg) * most recent weight (kg)*
Cumulative Dose	*Cum dose (mg) is the sum of the doses (mg) administered to a participant during the treatment period.*	*Cum dose (mg) is the sum of the doses (mg) administered to a participant during the treatment period.*
Relative dose intensity (%)	*Cum dose (mg) / [(Last dose date - Start dose date + 28) × 480 / 28] × 100*	*Cum dose (mg/kg) / [(Last dose date - Start dose date + 28) × 6 / 28] × 100*
Duration of treatment	*Last dose date - Start dose date + 1*	*Last dose date - Start dose date + 1*

Table 7.4.1-2: Administration of Study Therapy: Definition of Parameters - Nivolumab+Ipilimumab Arm

	Nivolumab	**Nivolumab**	**Ipilimumab**
Dosing schedule per protocol	**240mg every 2 weeks**	**3mg/kg every 2 weeks (for adolescents between 12 and < 18 years)**	**1mg/kg every 6 weeks**
Dose	*Dose (mg) is defined as: the nominal dose (mg) * total volume infused (mL) /total volume prepared (mL)*	*Dose (mg/kg) is defined as: the minimum of [vial strength (mg/mL) *total volume infused (mL)] /most recent weight (kg) and*	*Dose (mg/kg) is defined as the minimum of [vial strength (mg/mL) *total volume infused (mL)] /most recent weight (kg) and nominal dose (mg/kg) **

Table 7.4.1-2: Administration of Study Therapy: Definition of Parameters - Nivolumab+Ipilimumab Arm

	Nivolumab	Nivolumab	Ipilimumab
	nominal dose = 240mg	*nominal dose (mg/kg) * total volume infused (mL) /total volume prepared (mL)* *Vial strength = 10mg/mL* *Dose (mg) is defined as dose (mg/kg) * most recent weight (kg)*	*total volume infused (mL) /total volume prepared (mL)* *Vial strength = 5mg/mL* *Dose (mg) is defined as dose (mg/kg) * most recent weight (kg)*
Cumulative Dose	*Cum dose (mg) is the sum of the doses (mg) administered to a participant during the treatment period.*	*Cum dose (mg) is the sum of the doses (mg) administered to a participant during the treatment period.*	*Cum dose (mg/kg) is the sum of the doses (mg/kg) administered to a participant during the treatment period.* *Cum dose (mg) is the sum of the doses (mg) administered to a participant during the treatment period.*
Relative dose intensity (%)	*Cum dose (mg) / [(Last dose date - Start dose date + 14) × 240 / 14] × 100*	*Cum dose (mg/kg) / [(Last dose date - Start dose date + 14) × 3 / 14] × 100*	*Cum dose (mg/kg) /[(Last dose date - Start dose date + 42) x 1 / 42] x 100*
Duration of treatment	*Last dose date - Start dose date +1*	*Last dose date - Start dose date +1*	*Last dose date - Start dose date +1*

Table 7.4.1-3: Administration of Study Therapy: Definition of Parameters - Ipilimumab Arm

	Ipilimumab	Ipilimumab Open-Label	Nivolumab Open-Label
Dosing schedule per protocol	**10mg/kg every 3 weeks for 4 doses, then every 12 weeks as from Week 24**	**10mg/kg every 3 weeks for 4 doses, then every 12 weeks as from Week 24**	**480mg every 4 weeks**
Dose	*Dose (mg/kg) is defined as the minimum of [vial strength (mg/mL) *total volume infused (mL)] /most recent weight (kg) and nominal dose (mg/kg) * total volume infused (mL)*	*Dose (mg/kg) is defined as dose delivered (mg) / most recent weight (kg)*	*Dose (mg) is defined as dose delivered (mg)*

Table 7.4.1-3: Administration of Study Therapy: Definition of Parameters - Ipilimumab Arm

	Ipilimumab	Ipilimumab Open-Label	Nivolumab Open-Label
	/total volume prepared (mL) *Vial strength = 5mg/mL* *Dose (mg) is defined as dose (mg/kg) * most recent weight (kg)*	*Dose (mg) is defined as dose delivered (mg)*	
Cumulative Dose	*Cum dose (mg/kg) is the sum of the doses (mg/kg) administered to a participant during the treatment period.* *Cum dose (mg) is the sum of the doses (mg) administered to a participant during the treatment period.*	*Cum dose (mg/kg) is the sum of the doses (mg/kg) administered to a participant during the treatment period.* *Cum dose (mg) is the sum of the doses (mg) administered to a participant during the treatment period.*	*Cum dose (mg) is the sum of the doses (mg) administered to a participant during the treatment period.*
Relative dose intensity (%)	*Cum dose/[10 x (Last dose date - Start dose date + 21) / 21] x 100, if the participant's last dose is in induction phase (i.e., prior to or on visit C07 (Week 10))* *Cum dose/[10 x 4 + 10 x (Last dose date - first dose date - 161 + 84) / 84] x 100 , if the participant's last dose is in maintenance phase (i.e., on or after visit C15 (Week 24))*	-	*Cum dose (mg) / [(Last dose date - Start dose date + 28) × 480 / 28] × 100*
Duration of treatment	*Last dose date - Start dose date +1*	*Last dose date - Start dose date +1*	*Last dose date - Start dose date +1*

7.4.2 Modifications of Study Therapy

7.4.2.1 Dose Delays

Each nivolumab and ipilimumab infusion may be delayed. A dose will be considered as actually delayed if the delay is exceeding 3 days (i.e. greater than or equal to 4 days from scheduled dosing date) for nivolumab and ipilimumab. All study drugs must be delayed until treatment can resume. Reason for dose delay will be retrieved from CRF dosing pages.

The following parameters will be summarized by treatment group:

- Number of participants with at least one dose delayed
- Number of doses delayed per participant
- Total number of doses delayed / total number of doses received
- length of delay (days): on time, 4-7, 8-14, 15-42, > 42 days
- Reason for dose delay

7.4.2.2 Dose Omission

Dose omissions are not allowed in this study.

7.4.2.3 Dose Modifications

There will be no dose escalations or reductions of nivolumab nor ipilimumab allowed.

Each nivolumab or ipilimumab infusion can be interrupted and/or the IV infusion rate can be reduced. This information will be retrieved from CRF dosing pages.

The following will be summarized by treatment group:

- Number of participants with at least one dose infusion interrupted, number of infusions interrupted per participant, total number of infusions interrupted/total number of doses received, and the reason for the interruptions
- Number of participants with at least one infusion with IV infusion rate reduced, number of infusions with IV rate reduced per participant, total number of infusions with IV rate reduced/total number of doses received, and the reason of the IV rate reduction

[redacted]

7.5 Efficacy

RFS will be evaluated for treatment effect using the following testing strategy: RFS will be compared first in the all randomized participants with PD-L1 expression level < 1% subgroup with an alpha allocation of 0.03 (two-sided); and if significant, the alpha allocated to this subgroup will be recycled to the treatment comparison in the overall population (all randomized participants). Thus RFS will be compared in all randomized participants with an alpha allocation of 0.05 (two-sided). If the treatment comparison in all randomized participants with PD-L1 expression level < 1% is not significant, RFS will be compared in all randomized participants with an alpha allocation based on the method published by Spiessens and Debois[1]. This method takes into account the correlation between the test statistics for the overall and the subgroup analysis, and is essentially the same as the method used when dealing with interim analyses, i.e., group sequential method. The alpha allocated to the overall analysis will be calculated based on the fraction of information in the subgroup, relative to the overall population, which is determined by the ratio of events observed in the two groups. For example, if 257 events are observed in the PD-L1 expression level < 1% subgroup, and 651 events are observed in the overall population (i.e., about 40% of all events in the PD-L1<1% subgroup), the alpha allocated to the overall analysis is 0.0265.

If the treatment comparison in the all randomized participants with PD-L1 expression level < 1% subgroup is significant, RFS will be compared in all randomized participants with an alpha allocation of 0.05. In this case, two analyses, an interim and a final, will be conducted in all randomized participants. The interim analysis will take place at the same time as the RFS analysis for the PD-L1 expression level < 1% subgroup. The final analysis will take place when approximately 560 RFS events are observed in the overall population. The stopping boundaries at the interim and final analyses will be derived based on the exact number of RFS events using Lan-DeMets alpha spending function with O'Brien-Fleming boundaries.

If the treatment comparison in the all randomized participants with PD-L1 expression level < 1% subgroup is not significant, RFS will be compared in the overall population with an alpha allocation based on the method published by Spiessens and Debois[1]. In this case, no interim analysis is planned. The final RFS analysis for the overall population will be conducted when approximately 651 RFS events are observed. (assuming allocated alpha is 0.0265). In case the events come slower than projected, final analysis for the overall population will be conducted when the minimum follow up from randomization of the last participant reaches 30 months.

If both hypotheses for the RFS endpoint (for the PD-L1 < 1% subgroup and the overall population) are rejected, OS will be compared in all randomized participants with PD-L1 expression level < 1% with an alpha allocation of 0.05 (two-sided); and if significant, then OS will be compared in all randomized participants with an alpha allocation of 0.05 (two-sided). One formal OS interim analysis will be conducted at the time that approximately 162 deaths (65% information fraction) have been reached among all randomized participants with PD-L1 expression level < 1%. This formal comparison of OS will allow for early stopping for superiority. The stopping boundaries at the interim and final analyses will be derived based on the exact number of deaths using Lan DeMets alpha spending function with O'Brien-Fleming boundaries.

If the treatment comparison of RFS in the all randomized participants with PD-L1 expression level < 1% subgroup is not significant, but the treatment comparison of RFS in the all randomized participants population is significant (at alpha of 0.0265 or calculated), OS will be compared in all randomized participants at an alpha allocation of 0.0265 or calculated (two-sided).

No multiplicity adjustment for other secondary analyses will be made.

7.5.1 Recurrence-Free Survival (RFS)

The primary objective of the study is to compare the recurrence-free survival of nivolumab plus ipilimumab to nivolumab in all randomized participants with PD-L1 expression level < 1% and all randomized participants.

The primary RFS analyses will be conducted in hierarchical testing order as defined in Section 7.5 above using a stratified two-sided log-rank test. Stratification factor to be used in the analysis of all randomized participants with PD-L1 expression level < 1% is stage at screening (per IRT). Stratification factors to be used in the analysis of all randomized participants are PD-L1 status and stage at screening (per IRT). The two-sided log-rank p-value will be reported. The estimate of the RFS hazard ratio, of nivolumab plus ipilimumab to nivolumab, will be calculated using a stratified Cox proportional hazards model, with treatment as the single covariate, stratified by the above stratification factors. A two-sided 100x(1-adjusted α)% CI for the hazard ratio will also be presented.

The RFS distribution for each treatment group will be estimated using Kaplan-Meier techniques. Median RFS along with 95% CI will be constructed based on a log-log transformed CI for the survivor function. Rates at fixed time points (6, 12, 18, 24, 30, 36, 42, 48 months, depending on the minimum follow-up) will be derived from the Kaplan Meier estimate and corresponding confidence interval will be derived based on Greenwood formula for variance derivation and on log-log transformation applied on the survivor function.

The source of RFS event will be summarized:

- recurrence
  - disease at baseline
  - local recurrence
  - regional recurrence (in-transit metastasis or regional node recurrence)
  - distant metastasis
  - new primary melanoma
  - melanoma in situ
- death

The status of participants who are censored in the RFS KM analysis will be tabulated using following categories:

- Censored on randomization date

  - no baseline disease assessment
  - no on-study disease assessment with either no recurrence/death or recurrence/death with prior subsequent therapy or second non-melanoma primary cancer
- Censored on date of last disease assessment on-study or last assessment prior to subsequent anticancer therapy or second non-melanoma primary cancer
  - Received subsequent anti-cancer therapy
  - Second non-melanoma primary cancer
  - Still on treatment
  - In follow-up
  - Off study
    - lost to follow-up
    - participant withdrew consent
    - other

7.5.1.1 Sensitivity Analyses of RFS

Sensitivity analyses of RFS will also be performed:

- *RFS stratified analysis using stratification factors as obtained from the baseline CRF pages* (instead of IRT). This analysis will be performed only if at least one stratification factor at randomization (as per IRT) and baseline are not concordant for at least 10% of all randomized participants.
- *RFS analysis using a 2-sided, un-stratified log-rank tes*t and an un-stratified Cox proportional hazards model with treatment as the single covariate will be conducted.
- *RFS analysis for participants with no relevant deviation*. This analysis will be conducted only if there are more than 10% participants with relevant protocol deviations.
- *RFS accounting for assessment on/after subsequent therapy or on/after second non-melanoma primary cancer:* RFS will be defined similarly to the primary definition except that events (recurrence or death) and disease assessments that occurred on or after subsequent anti-cancer therapy or on or after second non-melanoma primary cancer will be considered (no time point truncation).
- *RFS accounting for missing disease assessments prior to RFS event (recurrence or death)*: This analysis will be performed only if at least 10% of RFS events have missing prior disease assessments. It will apply the following restriction to the primary definition. In case a participant has two or more missing disease assessments, the participant will be censored at the last disease assessment date prior the RFS event.
- A comparison of RFS between the two treatment arms using a 2-sided, stratified log-rank test will be conducted in which recurrence-free participants who are lost to follow-up for any cause will be considered as having an event at the time of the last tumor assessment date prior to loss to follow-up.

Estimates of the hazard ratio, its two-sided 100 x (1-adjusted α)% CI and p-value will be presented.

- A *multivariate Cox regression model* will be used in order to estimate the treatment effect after adjustment for the following factors, which are all measured at baseline:
  - Age (≥ 65 vs. < 65)
  - Gender (Male vs. Female)
  - ECOG PS (1 vs. 0)
  - Disease Stage (Stage IIIc vs. Stage IIIb, Stage IIId vs. Stage IIIb, Stage IV vs. Stage IIIb; source: eCRF)
  - PD-L1 status (≥ 1% vs. <1%, indeterminate vs. < 1%) (source: clinical database)
  - Time from Surgical Resection to Randomization (in weeks) (≥ 6 weeks vs. < 6 weeks)

Estimates of the hazard ratio, its two sided 95% CI and p-value will be presented.

7.5.1.2 Consistency of Treatment Effect on RFS in Subsets

To assess consistency of treatment effects in different subsets, a forest plot of the RFS un-stratified hazard ratio (and 95% CI) will be produced for the following subgroups.

- Age category I (< 65, ≥ 65)
- Age category II (< 65, ≥ 65 - < 75, ≥ 75)
- Gender (Male, Female)
- Race (white, black, Asian, other)
- Stage category I (Stage IIIb, Stage IIIc-IIId, Stage IV) (source: CRF)
- Stage category II (Stage IIIb, Stage IIIc, Stage IIId, Stage IV) (source: CRF)
- Stage category III (Stage III, Stage IV) (source: CRF)
- Stage category IV (Stage IIIb, Stage IIIc, Stage IIId, Stage IV M1a, Stage IV M1b, Stage IV M1c-M1d) (source: CRF)
- Tumor thickness (NO EVIDENCE OF PRIMARY MELANOMA, < 0.8 mm, 0.8 - 1.0mm, > 1.0 - 2.0mm, > 2.0 - 4.0mm, > 4.0mm) in Stage III participants
- Tumor ulceration status (present, absent, unknown) in Stage III participants
- Lymph node involvement I (clinically occult only, clinically detected only, clinically detected and clinically occult) in Stage III participants
- Lymph node involvement II (clinically occult only, clinically detected only + clinically detected and clinically occult) in Stage III participants
- Tumor ulceration status by lymph node involvement I (present + clinically occult only, present + clinically detected only, present + clinically detected and clinically occult, absent + clinically occult only, absent + clinically detected only, absent + clinically detected and clinically occult) in Stage III participants
- Tumor ulceration status by lymph node involvement II (present + clinically occult only, present + clinically detected only + clinically detected and clinically occult, absent + clinically occult only, absent + clinically detected only + clinically detected and clinically occult) in Stage III participants

- Total number of tumor involved regional lymph nodes (0, 1, 2-3, 4+) in Stage III participants
- Number of Clinically Occult Nodes (0, 1, 2-3, 4+) in Stage III participants
- Number of Clinically Detected Nodes (0, 1, 2-3, 4+) in Stage III participants
- Classification of Nodes (IN-TRANSIT/SATELLITE/MICRO-SATELLITE METS PRESENT WITH NO TUMOR-INVOLVED NODES, IN-TRANSIT/SATELLITE/MICRO-SATELLITE METS PRESENT WITH TUMOR-INVOLVED NODES, MATTED NODES, NOT APPLICABLE) in Stage III participants
- Completion lymph node dissection (CLND) (no, yes)
- PD-L1 Status (≥ 1% expression, < 1% expression/indeterminate) (source: clinical database)
- PD-L1 Status (≥ 5% expression, < 5% expression/indeterminate) (source: clinical database)
- PD-L1 Status (≥ 10% expression, < 10% expression/indeterminate) (source: clinical database)
- PD-L1 Status (< 1%, 1-< 5%, ≥ 5%, indeterminate) (source: clinical database)
- Melanoma subtype (mucosal, cutaneous, acral, ocular/uveal, other)
- BRAF mutation status (BRAF mutant, BRAF wildtype)
- Region (US and Canada, Western Europe, Eastern Europe, Australia, ROW)

If a subset category has less than 10 participants per treatment group, HR will not be computed/displayed. Number of events and median RFS along with 95% CI will be displayed for each treatment group.

Kaplan-Meier plots of RFS as well as RFS rates at fixed time points (6, 12, 18, 24, 30, 36, 42, 48 months, depending on the minimum follow-up) will be produced by:

- disease stage:
  - Stage category I (Stage IIIb, Stage IIIc, Stage IIId, Stage IV) (source: CRF)
  - Stage category II (Stage III, Stage IV) (source: CRF)
- BRAF mutation status (BRAF mutant, BRAF wildtype)

Further analyses of RFS and PD-L1 are described in Section 7.8.

Participant listings will be produced.

7.5.1.3 Participant Follow-up for RFS

The currentness of follow-up for RFS, defined as the time between last disease assessment (i.e., regardless of initiation of subsequent therapy) or randomization date (if the participant did not have on-study disease assessments) and data cut-off date, will be summarized by treatment group for all randomized participants. Participants who have an RFS event (regardless of initiation of subsequent therapy prior to RFS event) before data cut-off date will automatically have zero value for currentness of follow-up. Participants who have an RFS event (regardless of initiation of subsequent therapy prior to RFS event) and participants with last disease assessment date on or after data cut-off date will have zero value for currentness of follow-up. The currentness of follow-

up will be categorized into the following categories: 0, >0 - < 3 months, 3 - < 6 months, 6 - < 9 months, 9 - < 12 months, 12 - < 18 months, 18 - < 24 months, 24 - < 30 months, 30 - < 36 months, and ≥ 36 months.

7.5.2 Overall Survival (OS)

The primary OS analyses will be conducted using a two-sided log-rank test. Stratification factor to be used in the analysis of all randomized participants with PD-L1 expression level < 1% is stage at screening (per IRT). Stratification factors to be used in the analysis of all randomized participants are PD-L1 status and stage at screening (per IRT). The two-sided log-rank p-value will be reported. The estimate of the OS hazard ratio, of nivolumab plus ipilimumab to nivolumab, will be calculated using a stratified Cox proportional hazards model, with treatment as the single covariate, stratified by the above stratification factors. A two-sided (1 - adjusted α)% CI for the hazard ratio will also be presented.

The OS distribution for each treatment group will be estimated using Kaplan-Meier techniques. Median OS along with 95% CI will be constructed based on a log-log transformed CI for the survivor function. Rates at fixed time points (6, 12, 18, 24, 30, 36, 42, 48 months, depending on the minimum follow-up) will be derived from the Kaplan Meier estimate and corresponding confidence interval will be derived based on Greenwood formula for variance derivation and on log-log transformation applied on the survivor function.

The number of participants who are censored in the OS KM analysis and their status will be tabulated using following categories:

- Still on treatment (No recurrence, Recurrence)
- In follow-up
- Off study
  - lost to follow-up
  - participant withdrew consent
  - other

7.5.2.1 Sensitivity Analyses of OS

Sensitivity analyses of OS will also be performed:

- *OS stratified analysis using stratification factors as obtained from the baseline CRF pages* (instead of IRT). This analysis will be performed only if at least one stratification factor at randomization (as per IRT) and baseline are not concordant for at least 10% of all randomized participants.
- *OS analysis using a 2-sided, un-stratified log-rank tes*t and an un-stratified Cox proportional hazards model with treatment as the single covariate will be conducted.
- *OS analysis for participants with no relevant deviation*. This analysis will be conducted only if there are more than 10% participants with relevant protocol deviations.

Estimate of the hazard ratio, its two sided (1 - adjusted α)% CI and p-value will be presented.

- A *multivariate Cox regression model* will be used in order to estimate the treatment effect after adjustment for the following factors, which are all measured at baseline:
  - Age (≥ 65 vs. < 65)
  - Gender (Male vs. Female)
  - ECOG PS (0 vs. 1)
  - Disease Stage (Stage IIIc vs. Stage IIIb, Stage IIId vs. Stage IIIb, Stage IV vs. Stage IIIb; source: eCRF)
  - PD-L1 status (≥ 1% vs. <1%, indeterminate vs. <1%) (source: clinical database)
  - Time from Surgical Resection to Randomization (in weeks) (≥ 6 weeks vs. < 6 weeks)

Estimate of the hazard ratio, its two sided 95% CI and p-value will be presented.

7.5.2.2 Consistency of Treatment Effect on OS in Subsets

To assess consistency of treatment effects in different subsets, a forest plot of the OS un-stratified hazard ratio (and 95% CI) will be produced for the following subgroups.

- Age category I (< 65, ≥ 65)
- Age category II (< 65, ≥ 65 - < 75, ≥ 75)
- Gender (Male, Female)
- Race (white, black, Asian, other)
- Stage category I (Stage IIIb, Stage IIIc-IIId, Stage IV) (source: CRF)
- Stage category II (Stage IIIb, Stage IIIc, Stage IIId, Stage IV) (source: CRF)
- Stage category III (Stage III, Stage IV) (source: CRF)
- Stage category IV (Stage IIIb, Stage IIIc, Stage IIId, Stage IV M1a, Stage IV M1b, Stage IV M1c-M1d) (source: CRF)
- Tumor thickness (NO EVIDENCE OF PRIMARY MELANOMA, < 0.8mm, 0.8 - 1.0mm, > 1.0 - 2.0mm, > 2.0 - 4.0mm, > 4.0mm) in Stage III participants
- Tumor ulceration status (present, absent, unknown) in Stage III participants
- Lymph node involvement I (clinically occult only, clinically detected only, clinically detected and clinically occult) in Stage III participants
- Lymph node involvement II (clinically occult only, clinically detected only + clinically detected and clinically occult) in Stage III participants
- Tumor ulceration status by lymph node involvement I (present + clinically occult only, present + clinically detected only, present + clinically detected and clinically occult, absent + clinically occult only, absent + clinically detected only, absent + clinically detected and clinically occult) in Stage III participants
- Tumor ulceration status by lymph node involvement II (present + clinically occult only, present + clinically detected only + clinically detected and clinically occult, absent + clinically occult

only, absent + clinically detected only + clinically detected and clinically occult) in Stage III participants

- Total number of tumor involved regional lymph nodes (0, 1, 2-3, 4+) in Stage III participants
- Number of Clinically Occult Nodes (0, 1, 2-3, 4+) in Stage III participants
- Number of Clinically Detected Nodes (0, 1, 2-3, 4+) in Stage III participants
- Classification of Nodes (IN-TRANSIT/SATELLITE/MICRO-SATELLITE METS PRESENT WITH NO TUMOR-INVOLVED NODES, IN-TRANSIT/SATELLITE/MICRO-SATELLITE METS PRESENT WITH TUMOR-INVOLVED NODES, MATTED NODES, NOT APPLICABLE) in Stage III participants
- Completion lymph node dissection (CLND) (no, yes)
- PD-L1 Status (≥ 1% expression, < 1% expression/indeterminate) (source: clinical database)
- PD-L1 Status (≥ 5% expression, < 5% expression/indeterminate) (source: clinical database)
- PD-L1 Status (≥ 10% expression, < 10% expression/indeterminate) (source: clinical database)
- PD-L1 Status (< 1%, 1-< 5%, ≥ 5%, indeterminate) (source: clinical database)
- Melanoma subtype (mucosal, cutaneous, acral, ocular/uveal, other)
- BRAF mutation status (BRAF mutant, BRAF wildtype)
- Region (US and Canada, Western Europe, Eastern Europe, Australia, ROW)

If a subset category has less than 10 participants per treatment group, HR will not be computed/displayed. Number of events and median OS along with 95% CI will be displayed for each treatment group.

Kaplan-Meier plots of OS as well as OS rates at fixed time points (6, 12, 18, 24, 30, 36, 42, 48 months, depending on the minimum follow-up) will be produced by:

- disease stage:
  - Stage category I (Stage IIIb, Stage IIIc, Stage IIId, Stage IV) (source: CRF)
  - Stage category II (Stage III, Stage IV) (source: CRF)
- BRAF mutation status (BRAF mutant, BRAF wildtype)

Further analyses of OS and PD-L1 are described in Section 7.8.

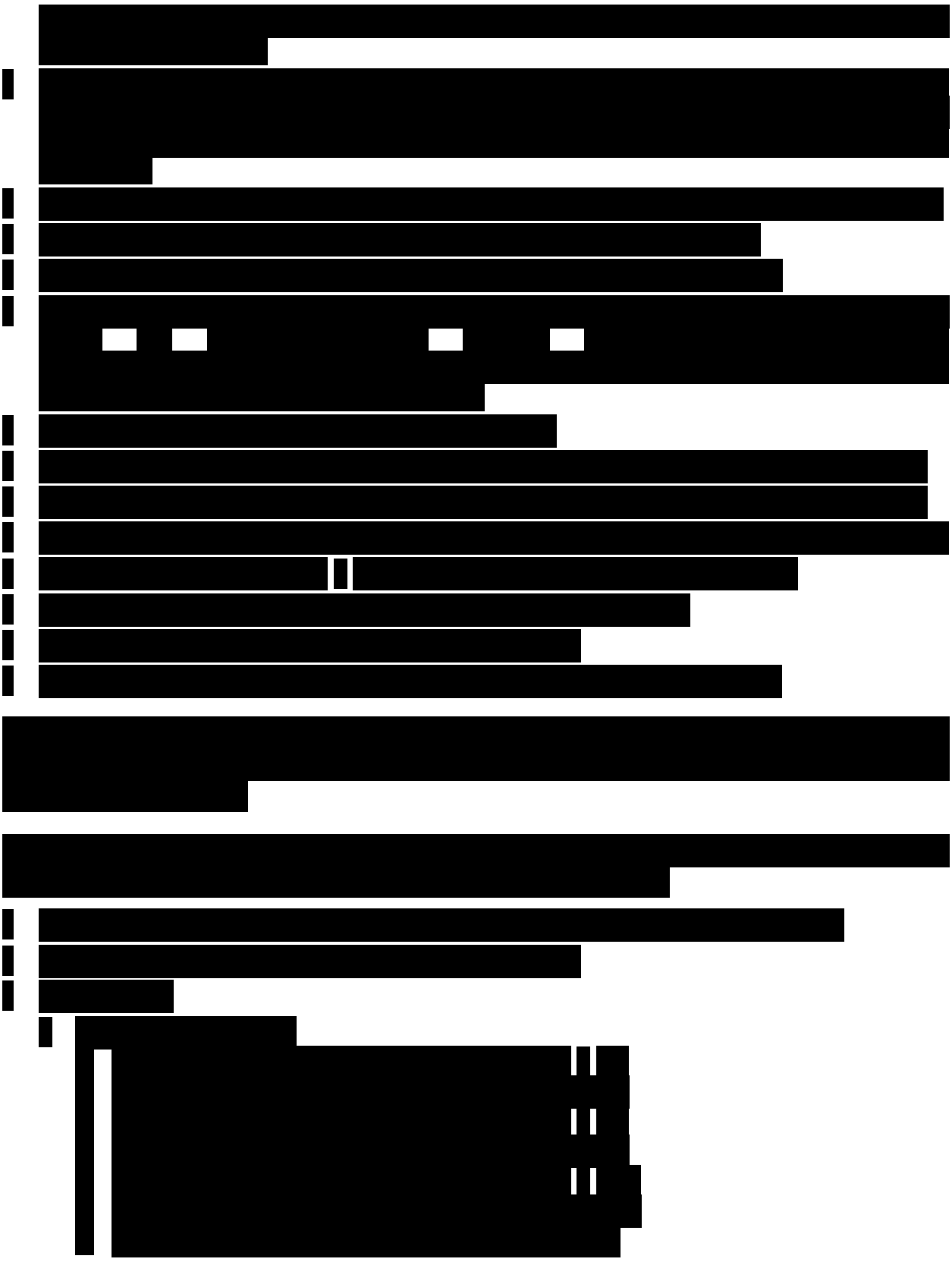
7.5.2.3 Participant Follow-up

The extent of follow-up for survival, defined as the time between randomization date and last known date alive (for participants who are alive) or death date (for participants who died), will be summarized descriptively (mean, standard deviation, median, first and third quartiles, min, max) by treatment group for all randomized participants.

The currentness of follow-up for survival, defined as the time between last OS contact (i.e., last known date alive or death date) and data cut-off date, will be summarized by treatment group for all randomized participants. Participants who died before data cut-off date will automatically have

zero value for currentness of follow-up. Participants who died and participants with last known date alive on or after data cut-off date will have zero value for currentness of follow-up. The currentness of follow-up will be categorized into the following categories: 0 days, 1 day - < 3 months, 3 - < 6 months, 6 - < 9 months, 9 - < 12 months, 12 - < 18 months, 18 - < 24 months, 24 - < 30 months, 30 - < 36 months, and ≥ 36 months.

[redacted]



7.5.4 Outcome on Subsequent Therapy

The following analyses will be performed using the primary and secondary definition defined in Section 4.2.3.

- Kaplan-Meier Plot of Progression-Free Survival on Next-Line Systemic Therapy - All Randomized Participants
- Reason for Event or Censoring, Progression-Free Survival on Next-Line Systemic Therapy
- Next-Line Systemic Cancer Therapy Summary
- Best Response on Next-Line Systemic Cancer Therapy Summary
- Kaplan-Meier Plot of Time to Next-Line Systemic Therapy (primary definition = secondary definition)
- Second Next-Line Systemic Cancer Therapy Summary
- Kaplan-Meier Plot of Time to Second Next-Line Systemic Therapy
- Time from Next-Line Systemic Therapy to Second Next-Line Systemic Therapy

By Participant Listing of Subsequent Systemic Cancer Therapy. By Participant Listing of Progression-Free Survival on Next-Line Systemic Therapy.

7.6 Safety

The safety data will be presented on all treated participants with all treated participants with PD-L1 expression level < 1% as well as on all treated participants.

Some analyses (as defined in Section 7.1) will be presented on all treated participants in the ipilimumab arm.

Analyses in this section will be tabulated by treatment group as treated, unless otherwise specified.

7.6.1 Deaths

Deaths will be summarized by treatment group:

- All deaths, reasons for death.
- Deaths within 30 days of last dose received, reasons for death.
- Deaths within 100 days of last dose received, reasons for death.

A by-subject listing of deaths will be provided for the all enrolled subjects population.

7.6.2 Serious Adverse Events

Serious adverse events will be summarized by treatment group:

- Overall summary of SAEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.
- Overall summary of drug-related SAEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.

All analyses will be conducted using the 30-day safety window.

A by-subject SAE listing will be provided for the "enrolled subjects" population.

7.6.3 Adverse Events Leading to Discontinuation of Study Therapy

AEs leading to discontinuation will be summarized by treatment group:

- Overall summary of AEs leading to discontinuation by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.
- Overall summary of drug-related AEs leading to discontinuation by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.

The analyses will be conducted using the 30-day safety window.

A by-subject AEs leading to discontinuation listing will be provided.

7.6.4 Adverse Events Leading to Dose Modification

AEs leading to dose delay/reduction will be summarized by treatment group:

- Overall summary of AEs leading to dose delay/reduction by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.
- Overall summary of related AEs leading to dose delay/reduction by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.

The analysis will be conducted using the 30-day safety window.

A by-subject AEs leading to dose delay/reduction listing will be provided.

7.6.5 Adverse Events

Adverse events will be summarized by treatment group.

The following analyses will be conducted using the 30 days safety window only:

- Overall summary of any AEs by worst CTC grade (1, 2, 3, 4, 5, not reported, total) presented by SOC/PT.
- Overall summary of any AEs presented by worst CTC grade (any grade, grade 3-4, grade 5) by SOC/PT. This table will be restricted to events with an incidence greater or equal to 5% in any treatment group.

- Overall summary of any non-serious AEs presented by SOC/PT. This table will be restricted to events with an incidence greater or equal to 5% in any treatment group.
- Overall summary of any AEs that required immune modulating medication by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.
- Overall summary of drug-related AEs by worst CTC grade (1, 2, 3, 4, 5, not reported, total) presented by SOC/PT.

The following analyses will be conducted using the 30 days safety window and repeated using the 100 days safety window:

- Overall summary of drug-related AEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT.

A by-subject AE listing will be provided. A by-subject listing of any AE requiring immune modulating medications will also be provided.

7.6.6 Select Adverse Events (EU Submission)

Unless otherwise specified, analyses will be performed by select AE category. Analyses will also be repeated by subcategory of endocrine events.

7.6.6.1 Incidence of Select AE

Select AEs will be summarized by treatment group for each category/subcategory.

The following analyses will be conducted using the 30-day safety window only:

- Overall summaries of any select AEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category or Subcategory/PT.
- Overall summaries of any drug-related select AEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category or Subcategory/PT.
- Overall summaries of any serious select AEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category or Subcategory /PT.
- Overall summaries of drug-related serious select AEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category or Subcategory /PT.
- Overall summaries of any select AEs leading to discontinuation by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category or Subcategory /PT.
- Overall summaries of drug-related select AEs leading to discontinuation by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category or Subcategory /PT.
- Summary of frequency of unique select AEs by Category.

A by-subject select AE listing will be provided.

7.6.6.2 Time-to Onset of Select AE

Time-to onset of drug-related select AEs (any grade, grade 3-5) will be summarized for each category/subcategory by treatment group.

Time-to onset analyses are restricted to treated subjects who experienced at least one drug-related select AE in the category/subcategory. The analyses will be conducted using the 30-day safety window.

Additional details regarding the time-to onset definition are described in time-to onset definition subsection of APPENDIX 1.

7.6.6.3 Time-to Resolution of Select AE

Time-to resolution of the following specific events will be summarized separately for each category/subcategory.

- Time-to resolution of drug-related select AE (any grade, grade 3-5) by treatment group
- Time-to resolution of drug-related select AE (any grade, grade 3-5) where immune modulating medication was initiated, by treatment group

Time-to resolution analyses are restricted to treated subjects who experienced the specific events. Time-to resolution where immune modulating medication was initiated analyses are restricted to treated subjects who experienced the specific events and who received immune modulating medication during the longest select AE.

The analyses will be conducted using the 30-day safety window.

The following summary statistics will be reported: percentage of subjects with resolution of the longest select AE, median time-to resolution along with 95% CI (derived from Kaplan-Meier estimation) and ranges.

See time-to resolution definition subsection of APPENDIX 1 for additional details.

7.6.6.4 Endocrine Select Events

Summary of specific endocrine disease history versus frequency of on-treatment endocrine select AEs using the 100-day safety window.

- Baseline history of hypothyroidism (yes, no, unknown)
  - vs. any grade on-treatment hypothyroidism select AE (yes, no)
  - vs. any grade on-treatment endocrine select AE (yes, no)
  - vs. grade 3-5 on-treatment hypothyroidism select AE (yes, no)
  - vs. grade 3-5 on-treatment endocrine select AE (yes, no)
  - vs. any grade on-treatment drug-related hypothyroidism select AE (yes, no)
  - vs. any grade on-treatment drug-related endocrine select AE (yes, no)
  - vs. grade 3-5 on-treatment drug-related hypothyroidism select AE (yes, no)

  - vs. grade 3-5 on-treatment drug-related endocrine select AE (yes, no)
- Baseline history of hyperthyroidism (yes, no, unknown)
  - vs. any grade on-treatment hyperthyroidism select AE (yes, no)
  - vs. any grade on-treatment endocrine select AE (yes, no)
  - vs. grade 3-5 on-treatment hyperthyroidism select AE (yes, no)
  - vs. grade 3-5 on-treatment endocrine select AE (yes, no)
  - vs. any grade on-treatment drug-related hyperthyroidism select AE (yes, no)
  - vs. any grade on-treatment drug-related endocrine select AE (yes, no)
  - vs. grade 3-5 on-treatment drug-related hyperthyroidism select AE (yes, no)
  - vs. grade 3-5 on-treatment drug-related endocrine select AE (yes, no)
- Baseline history of adrenal insufficiency (yes, no, unknown)
  - vs. any grade on-treatment adrenal insufficiency select AE (yes, no)
  - vs. any grade on-treatment endocrine select AE (yes, no)
  - vs. grade 3-5 on-treatment adrenal insufficiency select AE (yes, no)
  - vs. grade 3-5 on-treatment endocrine select AE (yes, no)
  - vs. any grade on-treatment drug-related adrenal insufficiency select AE (yes, no)
  - vs. any grade on-treatment drug-related endocrine select AE (yes, no)
  - vs. grade 3-5 on-treatment drug-related adrenal insufficiency select AE (yes, no)
  - vs. grade 3-5 on-treatment drug-related endocrine select AE (yes, no)
- Baseline history of type 1 diabetes mellitus (yes, no, unknown)
  - vs. any grade on-treatment diabetes mellitus select AE (yes, no)
  - vs. any grade on-treatment endocrine select AE (yes, no)
  - vs. grade 3-5 on-treatment diabetes mellitus select AE (yes, no)
  - vs. grade 3-5 on-treatment endocrine select AE (yes, no)
  - vs. any grade on-treatment drug-related diabetes mellitus select AE (yes, no)
  - vs. any grade on-treatment drug-related endocrine select AE (yes, no)
  - vs. grade 3-5 on-treatment drug-related diabetes mellitus select AE (yes, no)
  - vs. grade 3-5 on-treatment drug-related endocrine select AE (yes, no)
- Baseline history of type 2 diabetes mellitus (yes, no, unknown)
  - vs. any grade on-treatment diabetes mellitus select AE (yes, no)
  - vs. any grade on-treatment endocrine select AE (yes, no)
  - vs. grade 3-5 on-treatment diabetes mellitus select AE (yes, no)
  - vs. grade 3-5 on-treatment endocrine select AE (yes, no)
  - vs. any grade on-treatment drug-related diabetes mellitus select AE (yes, no)
  - vs. any grade on-treatment drug-related endocrine select AE (yes, no)

- vs. grade 3-5 on-treatment drug-related diabetes mellitus select AE (yes, no)
- vs. grade 3-5 on-treatment drug-related endocrine select AE (yes, no)

7.6.7 Immune-Mediated Adverse Events (US Submission)

IMAEs will be summarized by treatment group for each immune-mediated category / PT using the 100-day safety window:

- Overall summary of non-endocrine IMAEs by worst CTC grade (any grade, grade 3-4, grade 5) where immune modulating medication was initiated presented by Category / PT.
- Overall summary of endocrine IMAEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category / PT.
- Overall summary of non-endocrine IMAEs leading to discontinuation by worst CTC grade (any grade, grade 3-4, grade 5) where immune modulating medication was initiated presented by Category / PT.
- Overall summary of endocrine IMAEs leading to discontinuation by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category / PT.
- Overall summary of non-endocrine IMAEs leading to dose delay or reduction by worst CTC grade (any grade, grade 3-4, grade 5) where immune modulating medication was initiated presented by Category / PT
- Overall summary of endocrine IMAEs leading to dose delay or reduction by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category / PT.
- Summaries of time-to onset and time-to resolution of non-endocrine IMAEs where immune modulating medication was initiated presented by Category.
- Summaries of time-to onset and time-to resolution of endocrine IMAEs presented by Category.

A by-subject listing of IMAEs will be provided. By-subject listings of time-to resolution for longest IMAEs cluster (any grade and grade 3-5 in separate summaries) will also be provided. For new studies which collect investigator assessment of potential IMAE data, a by-subject listing of AEs considered as immune-mediated events per investigator but not qualified for IMAEs definition will also be provided.

In addition, for all nivolumab treated subjects who experienced at least one IMAE, the following data presentation will be provided:

- Summary of subjects who were re-challenged with nivolumab or ipilimumab by IMAE category, with extended follow-up

For these, re-challenge is considered to have occurred when last nivolumab and/or ipilimumab infusion was administered after the onset of an IMAE.

7.6.8 Other Events of Special Interest

OEOSI will be summarized by treatment group for each category.

The following analyses will be conducted using the 100-day safety window:

- Overall summary of OEOSI by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category / PT
- Overall summary of drug-related OEOSI by worst CTC grade (any grade, grade 3-4, grade 5) presented by Category / PT

A by-subject listing of OEOSI will be provided.

7.6.9 Multiple Events

The following summary tables will be provided:

- A table showing the total number and rate (exposure adjusted) of occurrences for all AEs.
- A table showing the total number and rate (exposure adjusted) of occurrences for AEs occurring in at least 5% of subjects in any treatment group.

In addition, the rate (exposure adjusted) and its 95% CI evaluated for different time intervals will be displayed graphically for each treatment group. This analysis will be limited to the rate of all AEs and all drug-related AEs.

The analyses will be conducted using the 30-day safety window.

A listing displaying the unique instances of all AEs, i.e., after duplicates have been eliminated and overlapping and contiguous occurrences of the same event (i.e. same PT) have been collapsed will be provided. No formal comparisons will be made between treatment groups.

7.6.10 Clinical Laboratory Evaluations

The analysis population for each laboratory test is restricted to treated subjects who underwent that laboratory test.

Lipase and Amylase will be added to the list of lab parameters to be summarized.

A by-subject listing of differences in categorization of SI and US laboratory test results will be provided.

7.6.10.1 Hematology

The following will be summarized by treatment group as worst CTC grade on-treatment per subject and as shift table of worst on-treatment CTC grade compared to baseline CTC grade per subject: hemoglobin (HB), platelets, white blood counts (WBC), absolute neutrophils count (ANC) and lymphocyte count (LYMPH).

The analyses will be conducted using the 30-day safety window.

A by-subject listing of these laboratory parameters will be provided.

7.6.10.2 Serum Chemistry

The following will be summarized by treatment group as worst CTC grade on-treatment per subject and as shift table of worst on-treatment CTC grade compared to baseline CTC grade per subject: ALT, AST, alkaline phosphatase (ALP), total bilirubin and creatinine.

The analyses will be conducted using the 30-day safety window.

A by-subject listing of these laboratory parameters will be provided.

7.6.10.3 Electrolytes

The following will be summarized by treatment group as worst CTC grade on-treatment per subject and as shift table of worst on-treatment CTC grade compared to baseline CTC grade per subject: sodium (high and low), potassium (high and low), calcium (high and low), magnesium (high and low), and Glucose Serum (fasting hyperglycemia and hypoglycemia regardless of fasting status).

The analyses will be conducted using the 30-day safety window.

A by-subject listing of these laboratory parameters will be provided.

7.6.10.4 Additional Analyses

In addition, further analyses on specific laboratory parameters will be performed by treatment group:

Abnormal Hepatic Function Test

The number of subjects with the following laboratory abnormalities from on-treatment evaluations will be summarized by treatment group:

- ALT or AST > 3 x ULN, > 5 x ULN, > 10 x ULN and > 20 x ULN
- Total bilirubin > 2 x ULN
- Concurrent (within 1 day) ALT or AST > 3 x ULN and total bilirubin > 2 x ULN
- Concurrent (within 30 days) ALT or AST > 3 x ULN and total bilirubin > 2 x ULN

The analyses will be conducted using the 30-day safety window.

A by-subject listing of these specific abnormalities will be provided.

Abnormal Thyroid Function Test

The number of subjects with the following laboratory abnormalities from on-treatment evaluations will be summarized by treatment group:

- TSH value > ULN and
  - with baseline TSH value ≤ ULN
  - with at least one FT3/FT4 test value < LLN within 2-week window after the abnormal TSH test

  - with all FT3/FT4 test values ≥ LLN within 2-week window after the abnormal TSH test
  - with FT3/FT4 missing within 2-week window after the abnormal TSH test.

- TSH < LLN and
  - with baseline TSH value ≥ LLN
  - with at least one FT3/FT4 test value > ULN within 2-week window after the abnormal TSH test
  - with all FT3/FT4 test values ≤ ULN within 2-week window after the abnormal TSH test
  - with FT3/FT4 missing within 2-week window after the abnormal TSH test

The analyses will be conducted using the 30-day safety window.

A by-subject listing of these specific abnormalities will be provided.

7.6.11 Vital Signs and Pulse Oximetry

Vital signs collected on the CRF will be provided in a listing.

7.6.12 Non-Protocol Medical Procedures

Non-protocol medical procedures will be listed by subject.

7.6.13 Immunogenicity Analysis

Further details on immunogenicity background and rationale, definitions, population for analyses and endpoints are described in APPENDIX 3.

Incidence of ADA

Number (%) of subjects will be reported for the following parameters based on Evaluable Subjects.

- Baseline ADA-positive
- ADA-positive
  - Persistent Positive (PP)
  - Not PP-Last Sample Positive
  - Other positive
  - Neutralizing Positive
- ADA-negative

A listing of all ADA assessments will be provided.

A spider plot of nivolumab ADA test result (titers) over time will be provided for nivolumab ADA positive subjects.

Clinical implications

Clinical implications of nivolumab immunogenicity will be primarily focused on subjects with persistent ADA-positive relative to ADA-negative. Subjects with any ADA-positive samples after initiation of treatment (relative to baseline) may be used to explore clinical implications.

Effect of immunogenicity on safety will be explored by examining the frequency and type of AEs of special interest such as hypersensitivity/infusion reaction. Summary tables for incidence of overall and each of the preferred terms will be provided, if the number of subjects is of sufficient size (e.g., at least 10 subjects). Otherwise, individual subject's safety profile will be examined and described based on a listing.

Clinical implications on efficacy will also be explored similarly. The following data presentation will be provided:

- Swimmer plot of occurrence of ADA and NAb Occurrence in Relation to RFS and OS

7.6.14 Pregnancy

A by-subject listing of pregnancy tests results will be provided for randomized female subjects.

7.6.15 Adverse Events By Subgroup

Overall summary of any AEs and drug-related AEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT and for each treatment group for the following subgroups:

- Sex (Male vs. Female)
- Race
- Age (< 65 vs. 65 - < 75 vs. ≥ 75 vs. ≥ 65)
- Region (as defined in Section 7.5.1.2)

These analyses will be conducted using the 30-day safety window only.

[redacted]

[redacted]

[redacted]

8 CONVENTIONS

The following conventions may be used for imputing partial dates for analyses requiring dates:

For missing and partial adverse event onset dates, imputation will be performed using the Adverse Event Domain Requirements Specification[7]. Missing and partial Non-Study Medication Domain dates will be imputed using the derivation algorithm described in Section 4.3.3 of BMS Non-Study Medication Domain Requirements Specification[8].

For death dates, the following conventions will be used for imputing partial dates:

- If only the day of the month is missing, the 1st of the month will be used to replace the missing day. The imputed date will be compared to the last known date alive and the maximum will be considered as the death date.
- If the month or the year is missing, the death date will be imputed as the last known date alive
- If the date is completely missing but the reason for death is present the death date will be imputed as the last known date alive

For date of recurrence, the following conventions will be used for imputing partial dates:

- 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 cases where the date of death is present and complete, the imputed recurrence date will be compared to the date of death. The minimum of the recurrence date and date of death will be considered as the date of recurrence.

For other partial/missing dates, the following conventions may be used:

- 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.

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

Duration (e.g. RFS, OS, DMFS, time to onset, time to resolution) will be calculated as follows:

Duration = (Last date - first date + 1)

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

9 CONTENT OF REPORTS

All analyses described in this SAP will be included in the Clinical Study Report except where otherwise noted. Refer to the Data Presentation Plan for mock-ups of all tables and listings.

10 DOCUMENT HISTORY

Table 10-1: Document History

Version	Date	Author	Summary of Changes
1.0	16 July 2019	[redacted]	Original version

APPENDIX 1 TIME-TO ONSET AND TIME-TO RESOLUTION DEFINITION AND CONVENTIONS FOR SELECT ADVERSE EVENTS, IMMUNE-MEDIATED ADVERSE EVENTS AND EVENTS OF SPECIAL INTEREST

Time-to onset definition

Time-to onset of AE (any grade) for a specific category is defined as the time between the day of the first dose of study treatment and the onset date of the earliest AE (of any grade) in this category.

The time-to onset of AE (grade 3-5) for a specific category is defined similarly with an onset date corresponding to a grade 3-5 AE.

Time-to onset of drug-related AE (any grade or grade 3-5) for a specific category is defined similarly but restricted to drug-related AE.

Time-to onset for a specific subcategory is defined similarly but restricted to event of this subcategory.

Time-to resolution definition

In order to derive the time-to resolution, overlapping or contiguous AEs within a specific category or subcategory will be collapsed into what will be termed "clustered" AEs. For example, if a subject (without pre-treatment AE) experienced an AE from 1st to 5th January, another AE (with different PT but within same category) from 6th to 11th January and same AE from 10th to 12th January, these will be collapsed into one clustered AE from 1st to 12th January. Table 10-2 is summarizing key derivation steps for each type of clustered AEs.

Time-to resolution of AE (any grade) for a specific category is defined as the longest time from onset to complete resolution or improvement to the grade at baseline among all clustered AEs experienced by the subject in this category per adverse event criteria category. Events which worsened into grade 5 events (death) or have a resolution date equal to the date of death are considered unresolved. If a clustered AE is considered as unresolved, the resolution date will be censored to the last known alive date. Improvement to the grade at baseline implies that all different events in the clustered adverse event should at least have improved to the corresponding (i.e. with same preferred term) baseline grade. This measure is defined only for subjects who experienced at least one AE in the specific category.

The time-to resolution of AE (grade 3-5) for a specific category is defined similarly with an onset date corresponding to a grade 3-5 AE.

Time-to resolution of drug-related AE (any grade or grade 3-5) for a specific category is defined similarly but restricted to drug-related AE.

The time-to resolution of AE (any grade or grade 3-5, drug-related or all) where immune modulating medication was initiated is defined similarly. For data presentation not restricted to

IMAE, the additional condition that the subject started an immune modulating medication during the longest AE resolution period will be applied.

Time-to resolution for a specific subcategory is defined similarly but restricted to event of this subcategory.

Table 10-2: Derivation of Clustered AE

Type of clustered AE	Derivation
Any grade	Collapse any on-treatment AE from the same category
Drug-related of any grade	Collapse any on-treatment drug-related AE from the same category
Grade 3 - 5	Collapse any on-treatment AE from the same category. Resolution will be based on the onset date of the earliest grade 3 - 5 records (if no grade 3 - 5 record, clustered AE is excluded)
Drug-related of Grade 3 - 5	Collapse any on-treatment drug-related AE from the same category Resolution will be based on the onset date of the earliest grade 3 - 5 record (if no grade 3 - 5 record, clustered AE is excluded)

The algorithm for collapsing adverse event records is using the following conventions:

For each subject and specified category, the corresponding adverse event records will be collapsed when:

1) Multiple adverse event records have the same onset date.
2) The onset date of an event record is either the same day or 1 day later than the resolution date of a preceding event record (contiguous events).
3) The onset date of an event record is after the onset date and prior to or on the resolution date of a preceding event record (overlapping events).

APPENDIX 2 MISSING AND PARTIAL RADIOTHERAPY AND SURGERY DATES IMPUTATION ALGORITHMS

Procedures – Imputation Rules.

If reported procedure start date is a full valid date then set start date equal to the date part of procedure start date.

In case of partial date use imputation rules described below:

- If only day is missing then
  - If month and year of procedure match month and year of first dose date then impute as date of first dose;
  - If month and year of procedure don't match month and year of first dose date then impute as first day of that month and year.
- If both day and month are missing, then impute as maximum between 01JAN of the year and date of the first dose;
- If date is completely missing or invalid then leave missing.

Note: Imputation is not applicable to data where start date is not collected (for example "PRIOR RADIOTHERAPY" CRF). Set start date to missing in this case.

If reported end date is a full valid date then set end date equal to the date part of the reported end date.

In case of partial date use imputation rules described below:

- If reported end date is partial then set end date equal to the last possible reported end date based on the partial entered reported end date.
- If reported end date is missing, continuing, unknown or invalid then set end date equal to the most recent database extraction date.

If end date was imputed then compare end date to the death date or last known alive date if subject is not dead. If posterior then end date should be imputed to death date (or last known alive date if subject not dead).

Note: Imputation of partial dates only applies to data entered on "RADIOTHERAPY" CRF page. For other CRF pages in case of partial dates set end date to missing.

Surgeries – Imputation Rules.

If reported surgery date is a full valid date then set start date equal to the date part of surgery date.

In case of partial date, use one of the two imputation rules described below:

A. For data collected on "PRIOR SURGERY RELATED TO CANCER" CRF page:

- If only day is missing then impute as the first day of the month;
- If both day and month are missing then then impute as 01JAN of the year;
- If date is completely missing or invalid then leave missing.

B. For data collected on other CRF pages (deemed to be on-treatment/subsequent surgeries):

- If only day is missing then
  - If month and year of surgery match month and year of first dose date then impute the missing date as the date of first dose;
  - If month and year of surgery don't match month and year of first dose date then impute as first day of that month and year;
- If both day and month are missing then impute as maximum between 01JAN of the year and date of the first dose;
- If date is completely missing or invalid then leave missing.