

Statistical Analysis Plan for

Official Title of Study

**A Phase 3, Randomized Study of Nivolumab plus Ipilimumab in Combination with
Chemotherapy vs Chemotherapy alone as First Line Therapy in Stage IV Non-Small Cell Lung
Cancer (NSCLC)
(CheckMate 9LA, CHECKpoint pathway and nivoluMAB clinical Trial Evaluation 9LA)**

NCT03215706

06-Aug-2019

**STATISTICAL ANALYSIS PLAN
FOR CLINICAL STUDY REPORT**

**A PHASE 3, RANDOMIZED STUDY OF NIVOLUMAB PLUS IPILIMUMAB IN
COMBINATION WITH CHEMOTHERAPY VS CHEMOTHERAPY ALONE AS FIRST
LINE THERAPY IN STAGE IV NON-SMALL CELL LUNG CANCER (NSCLC)**

PROTOCOL CA209-9LA

VERSION # 1.0

DATE: 06-AUG-2019

TABLE OF CONTENTS

STATISTICAL ANALYSIS PLAN FOR CLINICAL STUDY REPORT 1

TABLE OF CONTENTS 2

LIST OF TABLES 4

LIST OF FIGURES 5

█ [REDACTED]

2 STUDY DESCRIPTION 7

2.1 Study Design: 7

2.2 Treatment Assignment 8

2.3 Blinding and Unblinding 9

2.4 Protocol Amendments 9

2.5 Data Monitoring Committee 10

2.6 Blinded Independent Radiology Central Review 11

3 OBJECTIVES 11

3.1 Primary 11

3.2 Secondary 11

█ [REDACTED]

4 ENDPOINTS 12

4.1 Efficacy Endpoints 12

4.1.1 Overall Survival 12

4.1.2 Progression-Free Survival 13

4.1.2.1 Primary Definition of Progression-Free Survival (Accounting for Subsequent Therapy) 13

4.1.2.2 Secondary Definition of Progression Free Survival (Irrespective of Subsequent Therapy) 15

4.1.3 Objective Response Rate 16

4.1.3.1 Time to Response 16

4.1.3.2 Duration of Response 16

4.1.4 PFS2 17

4.2 Safety endpoints 18

█ [REDACTED]

4.3.2.1 PD-L1 Expression 18

4.3.2.2 Tumor Mutation Burden (TMB) 18

█ [REDACTED]

5 SAMPLE SIZE AND POWER 21


6 STUDY PERIODS, TREATMENT REGIMENS AND POPULATIONS FOR ANALYSES 22



6.1 Study Periods 22

6.2 Treatment Regimens 23

6.3 Populations for Analyses 23

7 STATISTICAL ANALYSES 24

7.1	General Methods	24
7.1.1	<i>Adverse Events, Serious Adverse Events, Multiple events, Select Adverse Events, Other Events of Special Interest and Immune-Mediated Adverse Events</i>	24
7.1.1.1	<i>Select Adverse Events (EU/ROW Submissions)</i>	25
7.1.1.2	<i>Other Events of Special Interest</i>	26
7.1.1.3	<i>Immune-Mediated Adverse Events (US Submission)</i>	26
7.1.2	<i>Laboratory Tests</i>	26
		
7.2	Study Conduct	27
7.3	Study Population	27
7.3.1	<i>Subject Disposition</i>	27
7.3.2	<i>Demographics and Other Baseline Disease Characteristics</i>	28
7.3.3	<i>Medical History</i>	29
7.3.4	<i>Prior Therapy Agents</i>	29
7.3.5	<i>Physical Examinations</i>	29
7.3.6	<i>Baseline Physical Measurements</i>	29
7.4	Extent of Exposure	29
7.4.1	<i>Administration of Study Therapy</i>	29
7.4.2	<i>Modifications of Study Therapy</i>	32
7.4.2.1	<i>Dose Delays</i>	32
7.4.2.2	<i>Infusion Interruptions and Rate Changes</i>	33
7.4.2.3	<i>Dose Escalations</i>	33
7.4.2.4	<i>Dose Reductions</i>	33
7.4.2.5	<i>Partial Discontinuation of Ipilimumab</i>	34
		
7.4.3.1	<i>Immune modulating medication</i>	35
7.4.3.2	<i>Subsequent Cancer Therapy</i>	36
7.5	Efficacy	36
7.5.1	<i>Analysis of Overall Survival</i>	37
7.5.2	<i>Supportive Analyses of Overall Survival</i>	38
7.5.3	<i>Subset Analyses of Overall Survival</i>	39
7.5.4	<i>Interim Analysis of Overall Survival</i>	40
7.5.5	<i>Analysis of Progression-Free Survival</i>	40
7.5.6	<i>Supportive Analyses of Progression-Free Survival</i>	41
7.5.7	<i>Subset Analyses of Progression-Free Survival</i>	42
7.5.8	<i>Current Status of PFS and OS Follow-up</i>	43
7.5.9	<i>Analysis of Objective Response Rate</i>	43
7.5.10	<i>Subset Analyses of Objective Response</i>	44
7.5.11	<i>Time to Tumor Response and Duration of Response</i>	45
7.5.12	<i>PFS2</i>	46
7.6	Safety	46
7.6.1	<i>Deaths</i>	46
7.6.2	<i>Serious Adverse Events</i>	46
7.6.3	<i>Adverse Events Leading to Discontinuation of Study Therapy</i>	46
7.6.4	<i>Adverse Events Leading to Dose Modification</i>	47

7.6.5	<i>Adverse Events</i>	47
7.6.6	<i>Select Adverse Events (EU/ROW Submissions)</i>	47
7.6.6.1	<i>Incidence of Select AE</i>	47
7.6.6.2	<i>Time-to Onset of Select AE</i>	48
7.6.6.3	<i>Time-to Resolution of Select AE</i>	48
7.6.7	<i>Immune-Mediated Adverse Events (US Submission)</i>	49
7.6.8	<i>Other Events of Special Interest</i>	49
7.6.9	<i>Multiple Events</i>	50
7.6.10	<i>Laboratory Parameters</i>	50
7.6.10.1	<i>Hematology</i>	50
7.6.10.2	<i>Serum Chemistry</i>	50
7.6.10.3	<i>Electrolytes</i>	51
7.6.10.4	<i>Additional Analyses</i>	51
7.6.11	<i>Vital Signs and Pulse Oximetry</i>	52
7.6.12	<i>Physical Measurements</i>	52
7.6.13	<i>Non-Protocol Medical Procedures</i>	52
7.6.14	<i>Immunogenicity Analysis</i>	52
7.6.15	<i>Pregnancy</i>	53
7.6.16	<i>Adverse Events By Subgroup</i>	53
		
7.8.1	<i>Distribution of PD-L1 Expression</i>	54
7.8.2	<i>Tumor Mutation Burden (TMB)</i>	54
7.8.2.1	<i>Tumor Tissue TMB</i>	54
7.8.2.2	<i>Blood TMB</i>	55
7.8.2.3	<i>Correlation between Tissue TMB and Blood TMB</i>	56
7.9	<i>Clinical Outcomes Assessments</i>	56
7.9.1	<i>EuroQol EQ-5D-3L</i>	56
7.9.2	<i>LCSS</i>	56
8	CONVENTIONS	57
9	CONTENT OF REPORTS	58
10	DOCUMENT HISTORY	58
		

LIST OF TABLES

Table 2.4-1:	<i>Protocol Amendments</i>	9
--------------	----------------------------------	---

Table 4.1.2.1-1: Censoring Scheme used in Primary Definition of PFS 14
Table 4.1.2.2-1: Censoring Scheme for Secondary definition of PFS..... 16
[REDACTED]
Table 7.4.1-1: Administration of study therapy - Nivolumab and Ipilimumab..... 29
Table 7.4.1-2: Administration of study therapy - Carboplatin/Paclitaxel 30
Table 7.4.1-3: Administration of study therapy - Carboplatin and Pemetrexed..... 31
Table 7.4.1-4: Administration of study therapy - Cisplatin/Pemetrexed 31
Table 7.4.2.4-1: Dose Modifications of Chemotherapeutic Agents 33
[REDACTED]
Table 10-1: Document History 58
Table 10-2: Derivation of clustered AE 60

LIST OF FIGURES

Figure 2.1-1: Study Design Schematic 7
Figure 4.1.2.1-1: PFS Primary Definition 14
Figure 4.1.2.2-1: PFS Secondary Definition 15
Figure 4.1.4-1: PFS2 Definition..... 17



2 STUDY DESCRIPTION

2.1 Study Design:

Adult (≥ 18 years) male and female participants, with stage IV NSCLC, previously untreated for advanced disease are eligible for enrollment, irrespectively of PD-L1 expression. Tumor tissue is required for study enrollment.

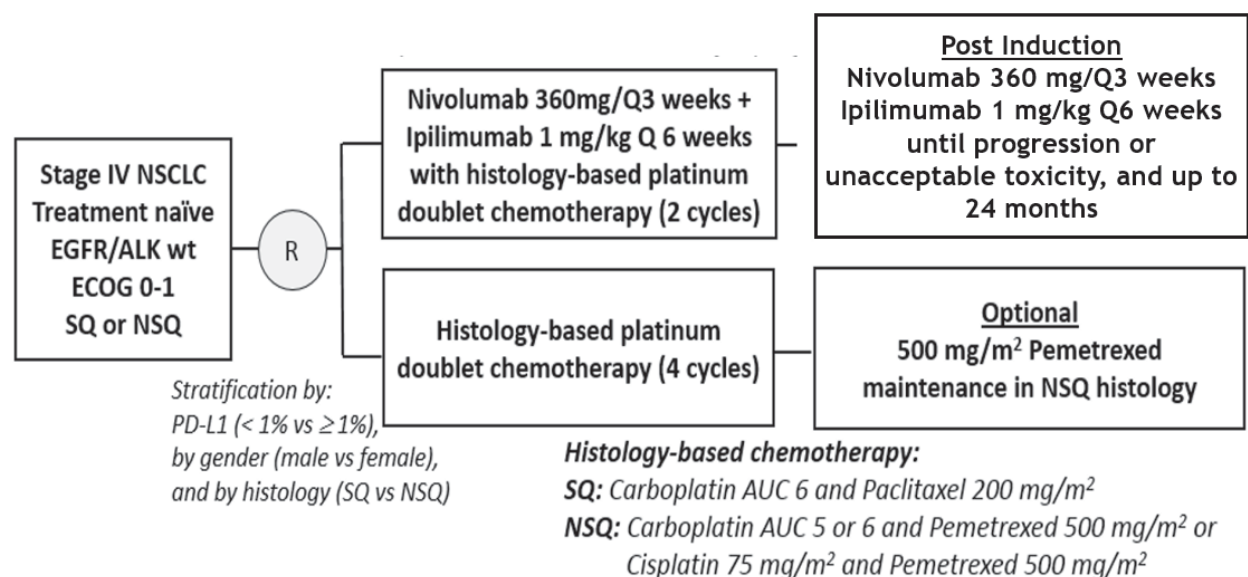
Participants will be assessed by PD-L1 expression, and categorized into 3 groups (PD-L1 positive, PD-L1 negative, and PD-L1 not quantifiable). PD-L1 status will be determined by Dako PD-L1 IHC 28-8 pharmDx test for immunohistochemical (IHC) staining of PD-L1 protein in submitted tumor sample.

- PD-L1 positive ($\geq 1\%$ tumor cell membrane staining in a minimum of a hundred evaluable tumor cells)
- PD-L1 negative ($< 1\%$ tumor cell membrane staining in a minimum of a hundred evaluable tumor cells)
- PD-L1 not quantifiable (participants with tumor biopsy specimens without quantifiable PD-L1 expression)

Participants are to have tumor tissue sample available for PD-L1 IHC testing performed by a central laboratory during screening period. PD-L1 status is required for randomization.

The enrollment will end when approximately 700 participants are randomized. The Interactive Response Technology (IRT) system will be used to track the enrollment number. Testing of PD-L1 through a central laboratory is required prior to randomization. Participants tested but with non quantifiable PD-L1 will stratify with PD-L1 negative participants. PD-L1 not quantifiable population will be capped to 10% of total randomized population. The study design schematic is presented in [Figure 2.1-1](#)

Figure 2.1-1: Study Design Schematic



Approximately 700 participants will be randomized in a 1:1 ratio to the treatment arm or control arm. The stratification factors for randomization are PD-L1 level ($\geq 1\%$ vs $< 1\%$), histology (squamous vs non-squamous), and sex (male vs female).

The investigator must decide prior to randomization if participant with non-squamous histology will receive cisplatin therapy, based on cisplatin eligibility criteria.

Treatment Arm: Nivolumab will be administered with ipilimumab, plus 2 cycles of histology-based platinum doublet chemotherapy:

- Squamous histology: carboplatin AUC 6 + paclitaxel 200 mg/m²
- Non-squamous histology: carboplatin AUC 5 or 6 + pemetrexed 500 mg/m² or cisplatin 75 mg/m² + pemetrexed 500 mg/m²

The dosing schedule of nivolumab and ipilimumab during induction will be determined by safety lead-in phase of study CA209568. After 2 cycles of induction treatment, nivolumab and ipilimumab treatment (nivolumab 360 mg every 3 weeks and ipilimumab 1 mg/kg every 6 weeks) will continue until progression, unacceptable toxicity, or other reasons specified in the protocol. In this study CA2099LA, treatment with nivolumab and ipilimumab will be given for up to 24 months in the absence of disease progression or unacceptable toxicity.

Treatment beyond initial investigator assessed RECIST 1.1 defined progression is permitted in treatment arm only, if the participant as specified in Protocol Section 7.1.3 has investigator assessed clinical benefit and is tolerating nivolumab and ipilimumab. Participants receiving nivolumab plus ipilimumab beyond investigator assessed progression must also continue tumor assessments until further progression at subsequent tumor assessment as indicated in Protocol Section 7.1.3. A maximum of 24 months of treatment also applies to treatment beyond progression.

Control Arm: In the control arm, histology-based platinum doublet chemotherapy options will be selected by the investigator and administered on day 1 every 3 weeks for 4 cycles; participants with non-squamous histology may also receive optional maintenance therapy with 400 mg/m² pemetrexed alone on day 1 of each 3 weeks until disease or unacceptable toxicity.

Histology-based platinum doublet chemotherapy:

- Squamous histology: carboplatin AUC 6 + paclitaxel 200 mg/m²
- Non-squamous histology: carboplatin AUC 5 or 6 + pemetrexed 500 mg/m² or cisplatin 75 mg/m² + pemetrexed 500 mg/m²

On-study tumor assessments will begin at week 6 post first dose date (± 7 days) and be performed every 6 weeks (± 7 days) until Week 48. After Week 48, tumor assessments will be performed every 12 weeks (± 7 days) until Blinded Independent Central Review (BICR) assessed progression.

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 by calling an IRT system to obtain the participant

number. Every participant that signs the informed consent form must be assigned a participant number in the IRT system.

Participants enrolled will be grouped according to PD-L1 status (positive, negative, or not quantifiable, using membranous staining in $\geq 1\%$ tumor cells vs membranous staining in $< 1\%$ tumor cells). PD-L1 expression data will be transferred directly from the analyzing lab to the IRT system. The IRT system will be used to track the enrollment number.

Once enrolled in IRT, enrolled participants that have met all eligibility criteria will be randomized through the IRT system. Randomization will be stratified by PD-L1 level ($\geq 1\%$ vs $< 1\%$), histology (squamous vs non-squamous), and sex (male vs female). Participants will be randomized in a 1:1 ratio to treatment and control arms.

Enrollment will stop once approximately 700 participants have been randomized.

The exact procedures for using the IRT system will be detailed in the IRT manual.

2.3 Blinding and Unblinding

This is an open label study.

2.4 Protocol Amendments

Table 2.4-1: Protocol Amendments

Document	Date of Issue	Rationale/Summary of Changes
Revised Protocol 04	08-Mar-2019	<p>Since the original design of CA2099LA study, several studies assessing PD-1(L1) inhibitor plus chemotherapy in first line NSCLC in nonsquamous and squamous subtypes have been published. Data from these studies have shown that PD-L1 inhibitor plus chemotherapy treatment were superior to chemotherapy alone in most efficacy endpoints. However, most studies shows a delayed effect in terms of overall survival with a late separation of the OS.²³⁴⁵ To ensure sufficient power to detect a survival benefit in such a scenario, the CA2099LA study was revised to update the statistical assumption with a delayed effect on survival, which has led to updated projected timelines and events.</p> <p>[REDACTED]</p> <ul style="list-style-type: none"> Updated the two planned interim analyses to one single interim analysis The interim and final analyses were updated with number of events, power, hazard ratios, and projected timing of events Blood TMB moved to secondary endpoint from exploratory endpoint
Revised Protocol 03	24-Jan-2019	<p>The CA2099LA study was revised to update the protocol with program level language for safety.</p> <p><u>Summary of Changes:</u></p> <ul style="list-style-type: none"> Updated Appendix 3 Adverse Events and Serious Adverse Events: Definitions and Procedures For Recording, Evaluating, Follow Up and Reporting

Table 2.4-1: Protocol Amendments

Document	Date of Issue	Rationale/Summary of Changes
		<ul style="list-style-type: none"> Updated Appendix 4 Women of Childbearing Potential Definitions and Methods of Contraception [REDACTED] Excluded vaccine use
Amendment 02 (site specific)	23-Jan-2019	<p>As the global sample size neared its target number in November 2018, clinical sites in China enrolled < 10 participants. This China-specific amendment adds a China-specific substudy to extend enrollment of participants in China to ensure sufficient sample size for subgroup analysis for safety and efficacy evaluation for participants in China.</p> <p><u>Summary of Changes:</u></p> <ul style="list-style-type: none"> Add a China specific sub-study to extend enrollment of participants from China (up to 70 randomized subjects)
Revised Protocol 02	02-Jul-2018	<p>After completion of 2 years of treatment, retreatment with nivolumab and ipililumab for subsequent disease progression does not appear to re-induce responses or provide clinical benefit derived from continuing I-O treatment beyond two years in advanced tumors. Participant number information was updated to enable adequate power to address the study primary objective based on revised design assumptions. Additionally, program updates were added and internal inconsistencies were corrected.</p> <p><u>Summary of Changes:</u></p> <ul style="list-style-type: none"> Removed 1 year re-treatment after progression [REDACTED] Expanded study sample size; updated study endpoints Updated document with program standards and corrected internal inconsistencies
Revised Protocol 01	10-Aug-2017	<p>[REDACTED]</p> <p><u>Summary of Changes:</u></p> <ul style="list-style-type: none"> Confirmed dosing language in the study [REDACTED] Typographical and formatting errors were corrected
Original Protocol	10-May-2017	Not applicable

2.5 Data Monitoring Committee

A Data Monitoring Committee (DMC) will be utilized to provide general oversight and safety considerations for this study, CA2099LA. The DMC will provide advice to the Sponsor regarding actions the committee deems necessary for the continuing protection of participants enrolled in this study. The DMC will be charged with assessing such actions in light of an acceptable risk/benefit profile for nivolumab/ipilimumab. The DMC will act in an advisory capacity to BMS and will monitor participant safety data for the study.

The DMC will be advisory to the clinical study leadership team. The clinical study leadership will have responsibility for overall conduct of the study including managing the communication of study data. The group will be responsible for promptly reviewing the DMC recommendations, for providing guidance regarding the continuation or termination of the study, and for determining whether amendments to the protocol or changes to the study conduct are required.

An independent statistician external to BMS will perform the interim analysis of OS and will provide it to the DMC. If OS achieves the above pre-specified boundary for declaring statistical significance, the DMC chair will notify BMS following procedures specified in Section 6.1.2 of the DMC Charter. Additional details of the DMC responsibilities and procedures will be specified in the DMC Charter.

2.6 Blinded Independent Radiology Central Review

Tumor assessments for each participant should be submitted to the radiology vendor as they will be performed on an ongoing basis. At the time of investigator-assessed initial radiographic progression per RECIST 1.1 in any given participant, the site must request the Independent Central Review from the third party radiology vendor for confirmation of progression. The blinded, independent radiologists will review all available tumor assessments for that given participant and determine if RECIST 1.1 criteria for progression have been met. The independent assessment of whether or not the given participant met RECIST 1.1 criteria for progression will be provided to the site. Participants whose disease progression is not confirmed centrally will be required to continue tumor assessments (if clinically feasible) according to the protocol-specified schedule. Subsequent tumor assessments must be submitted to the third party radiology vendor for subsequent review and may be discontinued when the investigator and independent radiologists both assess the participant to have met RECIST 1.1 criteria for progression.

BICR will review tumor images in all treated participants to determine RECIST 1.1 response for analyses of ORR or PFS. Details of the BICR procedure will be specified in the imaging manual.

3 OBJECTIVES

3.1 Primary

To compare overall survival (OS) of nivolumab + ipilimumab combined with chemotherapy vs chemotherapy in participants with histologically confirmed stage IV NSCLC.

3.2 Secondary

- To compare progression free survival (PFS), based on BICR assessment, of nivolumab + ipilimumab combined with chemotherapy vs chemotherapy in participants with histologically confirmed stage IV NSCLC.
- To compare objective response rate (ORR), based on BICR assessment, of nivolumab + ipilimumab combined with chemotherapy vs chemotherapy in participants with histologically confirmed stage IV NSCLC. Further characterization of the response will include time to ORR and duration of response.

- To evaluate ORR, PFS, and OS in participants with histologically confirmed stage IV NSCLC treated with nivolumab + ipilimumab combined with chemotherapy vs chemotherapy with different PD-L1 expression levels.
- To evaluate tumor mutation burden as a potential predictive biomarker of efficacy (such as ORR, PFS, and OS) of nivolumab + ipilimumab in combination with chemotherapy using DNA derived from tumor and blood specimens.

[REDACTED]

4 ENDPOINTS

4.1 Efficacy Endpoints

4.1.1 Overall Survival

The primary objective will be measured by the primary endpoint of Overall survival (OS). OS is defined as the time from randomization to the date of death from any cause. For subjects that are alive, their survival time will be censored at the date of last contact date (or “last known alive date”). Overall survival will be censored at the date of randomization for subjects who were randomized but had no follow-up.

OS rate at time T is defined as the probability that a subject is alive at time T following randomization. OS rates at fixed time points (e.g. 6 months, depending on minimum follow-up) are defined as the probability that a subject is alive at time T following randomization.

Survival follow-up will be conducted every 3 months after subject’s off-treatment date.

4.1.2 Progression-Free Survival

Two definitions are used for analysis of Progression-Free Survival (PFS). The primary definition accounts for subsequent therapy by censoring at the last evaluable tumor assessment on or prior to the date of subsequent therapy. The secondary definition is irrespective of subsequent therapy and does not account for subsequent therapy.

Clinical deterioration in the absence of unequivocal evidence of progression (per RECIST v1.1 criteria) is not considered progression for purposes of determining PFS.

PFS rate at time T is defined as the probability that a subject has not progressed and is alive at time T following randomization. PFS rates at fixed time points (e.g. 6 months, depending on the minimum follow-up) are defined as the probability that a subject has not progressed and is alive at time T following randomization.

The first on-study tumor assessment is scheduled to be conducted at 6 weeks (± 1 week) following first dose date. Subsequent tumor assessments are scheduled every 6 weeks (± 1 week) up to Week 48, then every 12 weeks (± 1 week) until BICR assessed disease progression.

4.1.2.1 Primary Definition of Progression-Free Survival (Accounting for Subsequent Therapy)

The primary definition of PFS (PFS truncated at subsequent therapy) is defined as the time between the date of randomization and the date of first documented tumor progression, based on BICR assessments (per RECIST v1.1 criteria), or death due to any cause, whichever occurs first.

Subjects who die without a reported progression will be considered to have progressed on the date of their death. The following censoring rules will be applied for the primary definition of PFS:

- Subjects who did not progress or die will be censored on the date of their last evaluable tumor assessment.
- Subjects who did not have any on study tumor assessments and did not die will be censored on their date of randomization.
- Subjects who receive subsequent anti-cancer therapy prior to documented progression will be censored at the date of the last evaluable tumor assessment conducted on or prior to the date of initiation of the subsequent anti-cancer therapy.
- Subjects who did not have a documented progression and received subsequent anti-cancer therapy will be censored at the date of the last evaluable tumor assessment conducted on or prior to the initiation of the subsequent anti-cancer therapy.

Censoring rules for the primary definition of PFS (PFS truncated at subsequent therapy) are presented as follows and in [Table 4.1.2.1-1](#).

Figure 4.1.2.1-1: PFS Primary Definition

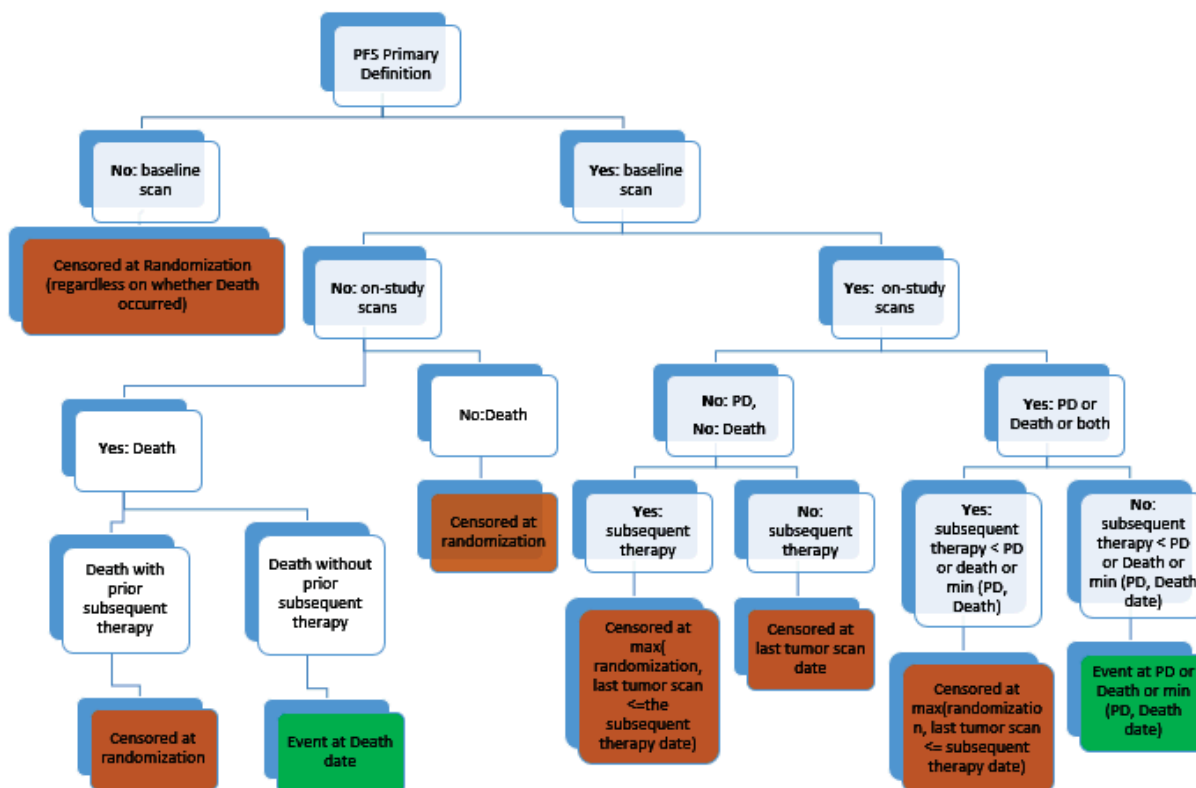


Table 4.1.2.1-1: Censoring Scheme used in Primary Definition of PFS

Situation	Date of Progression or Censoring	Outcome
No baseline tumor assessments*	Date of randomization	Censored
No on study tumor assessments and no death*	Date of randomization	Censored
Subsequent anti-cancer therapy started without death or progression per RECIST v1.1 reported prior or on the same day	Date of last evaluable tumor assessment prior to or on the date of initiation of the subsequent anti-cancer therapy	Censored
Documented progression per RECIST v1.1 and no new anti-cancer started before	Date of the first documented progression per RECIST v1.1 (excludes clinical progression)	Progressed
No progression and no death, and no new anti-cancer therapy started	Date of last evaluable tumor assessment	Censored
Death without progression per RECIST v1.1 and no new anti-cancer started before	Date of death	Progressed

* Tumor assessments and death if any, occurring after start of subsequent anti-cancer therapy are not considered.

4.1.2.2 Secondary Definition of Progression Free Survival (Irrespective of Subsequent Therapy)

The secondary definition of PFS (ITT definition) is defined as the time between the date of randomization and the date of first documented tumor progression, based on BICR assessments (per RECIST v1.1 criteria), or death due to any cause, whichever occurs first.

Subjects who die without a reported progression will be considered to have progressed on the date of their death. The following censoring rules will be applied for the secondary definition of PFS:

- Subjects who did not progress or die will be censored on the date of their last evaluable tumor assessment.
- Subjects who did not have any on study tumor assessments and did not die will be censored on their date of randomization.

Censoring rules for the secondary definition of PFS (ITT definition) are presented as follows and in Table 4.1.2.2-1.

Figure 4.1.2.2-1: PFS Secondary Definition

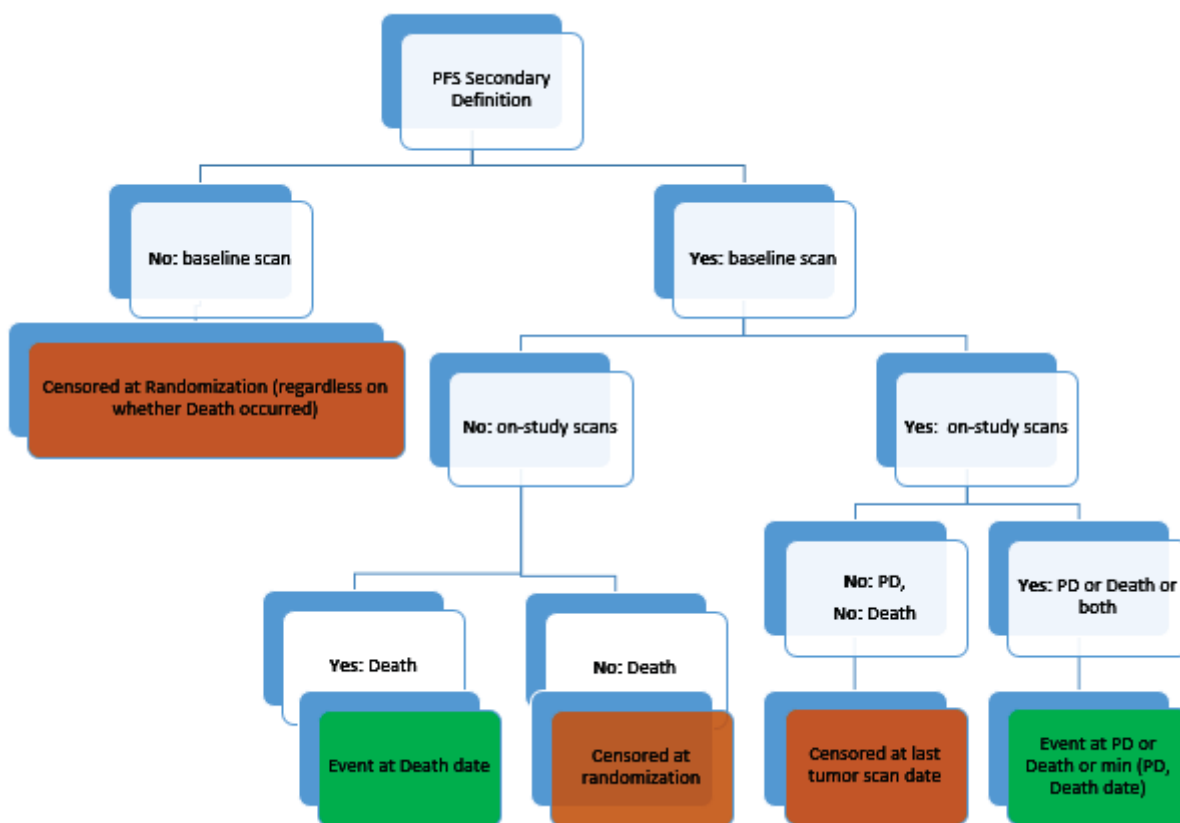


Table 4.1.2.2-1: Censoring Scheme for Secondary definition of PFS

Situation	Date of Progression of Censoring	Outcome
No baseline tumor assessment	Date of randomization	Censored
No on-study tumor assessments and no death	Date of randomization	Censored
Documented progression per RECIST v1.1	Date of first documented progression per RECIST v1.1 criteria (excludes clinical progression)	Progressed
No progression and no death	Date of last evaluable tumor assessment	Censored
Death without progression per RECIST v1.1	Date of death	Progressed

4.1.3 Objective Response Rate

Objective Response Rate (ORR) is defined as the number of randomized subjects who achieve a best response of confirmed complete response (CR) or confirmed partial response (PR) based on BICR assessments (using RECIST v1.1 criteria) divided by the number of all randomized subjects. Best Overall Response (BOR) is defined as the best response, as determined by the BICR, recorded between the date of randomization and the date of objectively documented progression per RECIST v1.1 criteria or the date of initiation of palliative local therapy or the date of subsequent therapy, whichever occurs first. For subjects without documented progression or palliative local therapy or subsequent therapy, all available response designations will contribute to the BOR determination. Confirmation of response is required at least 4 weeks after the initial response. Further characterization of the response will include time to response and depth of response (maximum tumor shrinkage in target lesions).

4.1.3.1 Time to Response

Time to Response (TTR) is defined as the time from randomization to the date of the first confirmed documented response (CR or PR), as assessed by the BICR. TTR will be evaluated for responders (confirmed CR or PR) only.

4.1.3.2 Duration of Response

Duration of Response (DOR) is defined as the time between the date of first confirmed documented response (CR or PR) to the date of the first documented tumor progression as determined by the BICR (per RECIST v1.1 criteria), or death due to any cause, whichever occurs first. Subjects who start subsequent therapy (including palliative local therapy) without a prior reported progression will be censored at the last evaluable tumor assessments prior to initiation of the subsequent anticancer therapy (including palliative local therapy). Subjects who die without a reported prior progression will be considered to have progressed on the date of their death. Subjects who neither

progress nor die, DOR will be censored on the date of their last evaluable tumor assessment. DOR will be evaluated for responders (confirmed CR or PR) only.

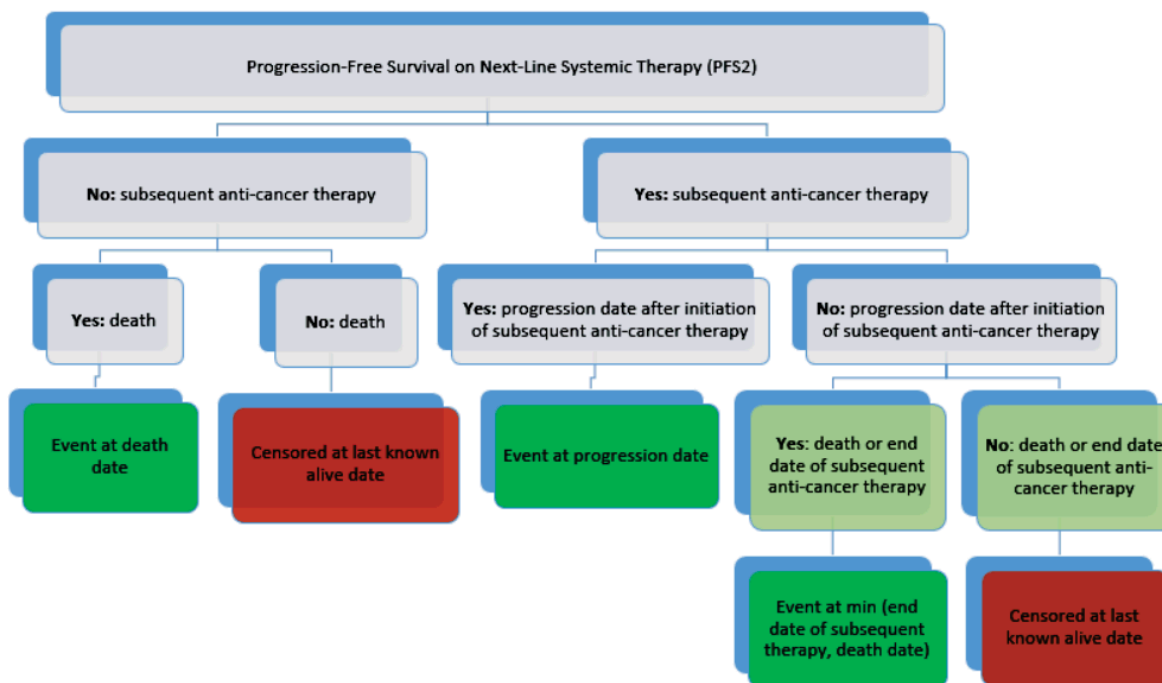
4.1.4 PFS2

PFS on next-line therapy (PFS2) is defined as the time from randomization to objectively documented progression after the next line of therapy, per investigator assessment, or to death from any cause, whichever occurs first. Subjects who were alive and without progression after the next line of therapy will be censored at last known alive date.

The following censoring rules will be applied for PFS2:

- Subjects who did not receive subsequent anti-cancer therapy (i.e. second-line therapy):
 - Subjects who died, the death date is the event date;
 - Else the subject’s PFS2 is censored at his last known alive date.
- Subjects who received subsequent anti-cancer therapy (i.e. second-line therapy):
 - Subjects who had a disease progression after the start of subsequent anti-cancer therapy, this disease progression date is the event date;
 - Else if a subject died or discontinued subsequent anti-cancer therapy, the date of min (death, discontinuation of subsequent anti-cancer therapy) is the event date;
 - Else the subject’s PFS2 is censored at his last known alive date.

Figure 4.1.4-1: PFS2 Definition



4.2 Safety endpoints

The assessment of safety will be based on the incidence of adverse events (AEs), serious adverse events (SAEs), adverse events leading to discontinuation, adverse events leading to dose modification, select adverse events (select AEs) for EU/ROW Submissions, immune-mediated AEs (IMAEs) for US Submission, other events of special interest (OEOSI), and deaths. The use of immune modulating concomitant medication will be also summarized. In addition clinical laboratory tests, and immunogenicity (i.e. development of anti-drug antibody) will be analyzed.

[REDACTED]

4.3.2.1 PD-L1 Expression

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 immunohistochemistry (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.

Subjects with missing PD-L1 expression are subjects with no tumor tissue sample available for evaluation.

PD-L1 expression will be collected in the IRT 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.

4.3.2.2 Tumor Mutation Burden (TMB)

Tumor Tissue TMB

Tumor tissue will be evaluated by targeted and/or whole exome sequencing for potential association with clinical outcomes. Gene expression signatures such as but not limited to those

associated with inflammatory processes and/or immune related signaling may also be evaluated by targeted and/or whole transcriptome RNA sequencing for potential association with clinical outcomes.

Blood TMB using Plasma Cell Free DNA (cfDNA)

Cell free DNA (cfDNA) are small fragments of DNA that are shed from tumor and from non-malignant cells and can be found circulating in the peripheral bloodstream. cfDNA derived from malignant cells can be isolated and analyzed from blood-derived plasma for genetic features including, but not limited to somatic mutations using targeted sequencing-based methods (eg, Next Generation sequencing). To complement the planned genomics analyses outlined for tumor assessments, plasma will be collected at baseline and on-treatment to isolate cfDNA baseline and on-treatment changes in mutational burden (either global, or individual genes) will be evaluated for association with treatment outcomes.

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

5 SAMPLE SIZE AND POWER

The sample size is based on the comparison of the primary endpoint of OS between nivolumab and ipilimumab plus chemotherapy and chemotherapy alone, with a two-sided overall alpha of 0.05. The number of events was estimated assuming an exponential distribution of OS for the chemotherapy arm and a piecewise exponential distribution with a 3 month delayed treatment effect for the nivolumab + ipilimumab + chemotherapy arm²³⁴⁵.

Approximately 700 participants will be randomized to treatment arm and control arm in a 1:1 ratio. Approximately 402 events (ie, deaths), observed among the 700 randomized participants, provides 81% power to detect an average hazard ratio (HR) of 0.75 with a type 1 error of 0.05 (two-sided). The average HR of 0.75 resulted from an assumed targeted hazard ratio of 1 for the initial 3 months from randomization and a targeted hazard ratio of 0.68 for the time beyond 3 months from randomization and corresponds to a 33% increase in the median OS, assuming a median OS of 13.93 months for chemotherapy alone and 18.57 months for nivolumab and ipilimumab plus chemotherapy respectively.

There is one planned interim analysis of OS for superiority at approximately 80% of total events, ie, 322. The stopping boundaries at the interim and final analyses will be based on the actual number of OS events at the time of the analysis using Lan-DeMets alpha spending function with O'Brien-Fleming boundaries. It is estimated that it will take approximately 24/29 months from the randomization of the first participant to observe the required number of events for the interim and the final OS analysis, with approximately 15 months of accrual to reach 700 randomized participants, assuming a piecewise accrual rate (5 for the first 2 months, 23 for the next 3 months, 47 for the next 3 months, and 70/month starting from the eighth month).

Comparison of PFS per BICR is one of the key secondary objectives. 596 events provides approximately 94% power to detect a hazard ratio (HR) of 0.75 with a type 1 error of 0.05 (two-sided). The HR of 0.75 corresponds to a 33% increase in the median PFS per BICR, assuming a median PFS per BICR of 6 months for chemotherapy alone and 8 months for nivolumab and ipilimumab plus chemotherapy respectively. One interim analysis of PFS per BICR is planned at the time of the interim analysis of the primary analysis of OS. This formal interim comparison of PFS per BICR would be conducted in a hierarchical manner, i.e. if the formal comparison of OS is statistically significant then the formal comparison of PFS per BICR would be conducted (see Section 7.5). The stopping boundaries at the interim and final analyses will be based on the actual number of PFS per BICR events at the time of the analysis using Lan-DeMets alpha spending function with O'Brien-Fleming boundaries. At the final analysis time of OS, 596 PFS per BICR events are expected. Approximately 524 out of 596 (88%) PFS per BICR are expected at the time of the interim analysis. If the interim analysis is performed exactly at 524 events, then PFS per BICR would be considered statistically significant if the p-value is ≤ 0.034 , providing approximately 88% power. The type 1 error to be used for the final analysis of PFS per BICR would then be 0.041.

This study includes a sub-study to allow enrollment of patients from China (site-specific protocol Amendment 02). Data from these additional subjects will be reported separately. Subjects from

China randomized on or before the end of global study accrual will be included in the population used for the primary analysis clinical study report. The required number of deaths for the interim and final OS analyses are based on the global study population.

6 STUDY PERIODS, TREATMENT REGIMENS AND POPULATIONS FOR ANALYSES

6.1 Study Periods

- Baseline period:
 - Baseline evaluations or pre-treatment events will be defined as evaluations or events that occur before the date and time of the first dose of study treatment. Evaluations (laboratory tests, pulse oximetry, vital signs, and biomarkers) 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.
 - 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, pulse oximetry, vital signs, and biomarkers) 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 observations on or prior to the first dose of study treatment then the latest non missing observation on or before first dose date (and time if collected) will be used as the baseline in the analyses. If multiple observations exist on the latest collection date (and time if collected), the record with the latest data entry date and time will be used. If multiple observations exist on the latest collection date (and time if collected) and data entry date and time, then the first observation is used as baseline, unless otherwise specified.
 - ◆ For PD-L1, non-missing is identified as those with quantifiable test result. After applying the rule above, if there are no records with a quantifiable test result, then select those with indeterminate result (“INDETERMINATE”). If there are no records with indeterminate test result, then select those with unavailable result (“NOT EVALUABLE”). If there are no records with unavailable test result, then select those which are not reported or not available result (all other records).
 - ◆ For Anti-Drug Antibody (ADA), the record related to the most recent assessment among those records where date (and time if collected) of Nivolumab/Ipilimumab immunoglobulin (IMG) assessment is less than the date (and time if collected) of the first Nivolumab/Ipilimumab dose date.
- Post baseline period:
 - On-treatment AEs will be defined as AEs with an onset date and time on or after the date and time of the first dose of study treatment (or with an onset date on or after the day of first dose of study treatment if time is not collected or is missing). For subjects who are off

study treatment, AEs will be included if event occurred within a safety window of 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 (such as laboratory tests, pulse oximetry, vital signs, etc.) will be defined as evaluations taken after the day (and time, if collected and not missing) of first dose of study treatment. For subjects who are off study treatment, evaluations should be within a safety window of 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 subjects who are off study treatment

6.2 Treatment Regimens

Treatment group “as randomized” corresponds to the treatment group assigned by the Interactive Response Technology (IRT) system.

The treatment group “as treated” will be same as the treatment group “as randomized” by IRT unless a subject received the incorrect study treatment for the entire period of treatment, in which case the subject’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

- Global study population: all subjects enrolled during the global accrual window (from first patient first consent date to last patient outside of China’s consent date). Any patient from China enrolled during the global accrual window will be included.
- Enrolled subjects: All subjects who signed the informed consent form and obtained a subject number during the global enrollment period.
- Randomized subjects: All subjects from the global study population who were randomized through the IRT.
- Treated subjects: All randomized subjects from the global study population who received at least one dose of any study treatment.
- Response evaluable subjects: All randomized subjects from the global study population with baseline and at least one on-study tumor assessment. This population is used for analysis of tumor burden changes over the time.
- Tumor Tissue TMB evaluable subjects: All randomized subjects from the global study population with baseline evaluable tumor tissue TMB (non-missing numeric).
- Blood TMB evaluable subjects: All randomized subjects from the global study population with baseline evaluable blood TMB (non-missing numeric).
- China Cohort: Subjects enrolled from China Mainland.

The primary population of analysis of analysis is the global study population.

Unless otherwise specified, the safety analyses will include all treated subjects, global study population.

Unless otherwise specified, the efficacy analyses will include all randomized subjects, global study population.

7 STATISTICAL ANALYSES

7.1 General Methods

Unless otherwise noted, discrete variables will be tabulated by the frequency and proportion of subjects falling into each category, grouped by treatment. Percentages given in these tables will be rounded to the first decimal and, therefore, may not always sum to 100%. Percentages less than 0.1 will be indicated as '< 0.1'. Continuous variables will be summarized by treatment group using the mean, standard deviation, median, minimum, and maximum values.

Time-to-event variables (e.g. time-to resolution) will be analyzed using the Kaplan-Meier technique. When specified, the median will be reported along with 95% CI using Brookmeyer and Crowley method¹⁰ (using log-log transformation for constructing the confidence intervals¹¹).

The conventions to be used for imputing missing and partial dates for analyses requiring dates are described in Section 8.

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 SOC 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 algorithms¹² 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

- $(\text{Date of last dose of study treatment} - \text{date of first dose of study treatment} + 31 \text{ days (or 101 days, depending on the analysis)}) / 365.25$, 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.
- $(\text{Last known alive date} - \text{date of first dose of study treatment} + 1) / 365.25$, for subjects who are still on-treatment or who are off study treatment and were followed less than 30 days (or 100 days depending on the analysis) after last dose of study treatment.

7.1.1.1 Select Adverse Events (EU/ROW Submissions)

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 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.

[REDACTED]

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.



7.2 Study Conduct

The following programmable deviations will be considered as relevant protocol deviations.

Eligibility:

- Subjects with misclassified PD-L1 stratification level (IRT vs clinical database)
- Subjects who do not have measurable disease at baseline per investigator

On-study:

- Subjects receiving concurrent anti-cancer therapy
- Subjects treated differently than as randomized (subjects who received the wrong treatment for the entire treatment period, excluding the never treated)

The Relevant Protocol Deviations will be summarized and listed based on all randomized subjects, by treatment group and overall.

Enrollment by country and site, and enrollment by month will be summarized and listed for all enrolled subjects.

A by-subject listing of batch numbers for all treated subjects will be provided.

7.3 Study Population

Analyses in this section will be tabulated for all randomized subjects by treatment group as randomized, unless otherwise specified.

7.3.1 *Subject Disposition*

The total number of subjects enrolled (randomized or not randomized) will be presented along with the reason for not being randomized. This analysis will be performed on the all enrolled subjects population only.

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

Number of subjects who discontinued study treatment along with corresponding reason will be tabulated by treatment group as treated. Reason for discontinuation will be derived from subject status CRF page. This analysis will be performed only on the all treated subjects population.

A by-subject listing for all treated subjects will be provided showing the subject's off treatment date and whether the subject continue in the treatment period/study along with the reason for going off treatment period/study. A by-subject listing for all enrolled subjects will also be provided,

showing whether the subject was randomized/treated along with the reason for not being randomized/treated.

7.3.2 Demographics and Other Baseline Disease Characteristics

The following demographic and baseline disease characteristics will be summarized and listed by treatment group as randomized:

- Age (continuous)
- Age categorization (< 65, ≥ 65 and < 75, ≥ 75 and < 85, ≥ 85, ≥ 75, ≥ 65)
- Sex (Male vs. Female)
- Race (White, Black or African American, Asian, American Indian, or Alaska Native, Native Hawaiian or Other Pacific Islander, Other, Asian Indian, Chinese, Japanese, Asian Other)
- Ethnicity (Hispanic/Latino and Not Hispanic/Latino)
- Country by geographic region
- ECOG performance status (0, 1, >1)
- Baseline weight (kg) (continuous)
- Tumor histology (Squamous vs. Non-squamous)
- Prior surgery related to cancer (Yes vs. No)
- Prior radiotherapy (Yes vs. No)
- Time from initial disease diagnosis to randomization (< 1 year, ≥ 1 year, 1 - < 2 years, 2 - < 3 years, 3 - < 4 years, 4 - < 5 years, ≥ 5 years)
- Disease stage (Stage IV vs. Recurrent to Metastatic Disease)
- Smoking status (Current/Former, Never Smoked, Unknown)
- Region (Europe, North America, Rest of the World, Asia)
- Baseline PD-L1+ status based on a 1% cut off (≥ 1% vs. < 1% or indeterminate)
- PD-L1 expression subgroups (<1%, ≥1%, 1-49%, ≥50%)
- CNS Metastasis (Yes vs. No)
- Liver Metastasis (Yes vs. No)
- Bone Metastasis (Yes vs. No)
- Sites of diseases (all lesions)
- Number of disease sites per subject (all lesions)
- Tumor burden: sum of the diameters of target lesions at baseline
- Pre-treatment tumor assessments (per investigator)

Summary table (cross-tabulation) by treatment group for stratification factors (PD-L1 status, histology, sex) will be provided to show any discrepancies between what was reported through IRT vs. CRF data or clinical database at baseline. This summary will be performed based on all randomized subjects.

A listing of randomization scheme presenting randomized treatment group and as treated treatment group will be provided for all randomized subjects.

7.3.3 **Medical History**

A by-subject listing of general medical history for all randomized subjects will be provided.

7.3.4 **Prior Therapy Agents**

Prior cancer therapy will be summarized by treatment group and overall.

Prior systemic cancer therapy will be summarized by treatment group and overall and listed by subject.

Prior radiotherapy and prior surgery related to cancer will be listed by subject.

7.3.5 **Physical Examinations**

Subjects with abnormal baseline physical examination will be listed by subject.

7.3.6 **Baseline Physical Measurements**

Baseline physical measurements will be listed by subject.

7.4 **Extent of Exposure**

Listings will include all available exposure data. Analyses will be performed by treatment group “as treated” in all treated subjects, unless otherwise specified.

7.4.1 **Administration of Study Therapy**

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

- Number of doses received
- Cumulative dose
- Relative dose intensity (%) using the following categories: < 50%; 50 - < 70%; 70 - < 90%; 90 - < 110%; ≥ 110%

Duration of study therapy will be summarized (descriptive statistics) by treatment group. Duration of combination treatment is the last dosing date of any drug component minus the first dosing date of any drug component plus one during the treatment phase.

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

Table 7.4.1-1: Administration of study therapy - Nivolumab and Ipilimumab

	nivolumab	Ipilimumab
Dosing schedule per protocol	360 mg every 3 weeks	1 mg/kg every 6 weeks
Dose	<i>Dose (mg/kg) is defined as Total Dose administered (mg). Dose administered in</i>	<i>Dose (mg/kg) is defined as Total Dose administered (mg)/Most recent weight (kg). Dose</i>

Table 7.4.1-1: Administration of study therapy - Nivolumab and Ipilimumab

	nivolumab	Ipilimumab
	mg at each dosing date is collected on the CRF.	administered in mg at each dosing date and weight are collected on the CRF.
Cumulative Dose	<i>Cum dose (mg) is sum of the doses (mg) administered to a subject during the treatment period.</i>	<i>Cum dose (mg/kg) is sum of the doses (mg/kg) administered to a subject during the treatment period.</i>
Relative dose intensity (%)	$[Cum\ dose\ (mg) / ((Last\ Nivolumab\ dose\ date - Nivolumab\ Start\ dose\ date + 21) \times 360 / 21)] \times 100$	$[Cum\ dose\ (mg/kg) / ((Last\ dose\ date - Start\ dose\ date + 42) \times 1 / 42)] \times 100$
Duration of study therapy	<i>Last Nivolumab dose date - Nivolumab Start dose date +1</i> <i>(Start dose date of the combination is the first dose date among the first dose dates of all drugs of the combination and last dose date of the combination is the last date among the last dose dates of all drugs of the combination)</i>	<i>Last Ipilimumab dose date - Ipilimumab Start dose date +1</i>

Table 7.4.1-2: Administration of study therapy - Carboplatin/Paclitaxel

	Carboplatin	Paclitaxel
Dosing schedule per protocol	AUC 6 every 3 weeks	200 mg/m ² every 3 weeks
Dose	<i>Dose (AUC) is defined as Total Dose administered (mg)/(creatinine clearance + 25). Dose administered in mg at each dosing date is collected on the CRF and creatinine clearance derived from the CRF data.</i>	<i>Dose (mg/kg) is defined as Total Dose administered (mg)/Most recent BSA (m²). Dose administered in mg at each dosing date is collected on the CRF and BSA is derived from most recent weight and baseline height also collected on the CRF.</i>
Cumulative Dose	<i>Cum Dose (AUC) is the sum of the doses (AUC) administered to a subject.</i>	<i>Cum dose (mg/m²) is sum of the doses (mg/m²) administered to a subject during the treatment period.</i>
Relative dose intensity (%)	$Cum\ dose\ (AUC) / ((Last\ dose\ date\ of\ Carbo - Start\ dose\ date\ of\ Carbo + 21) \times 6 / 21) \times 100$	$[Cum\ dose\ (mg/m^2) / ((Last\ paclitaxel\ dose\ date - paclitaxel\ Start\ dose\ date + 21) \times 200 / 21)] \times 100$
Duration of study therapy	Last dose date - Start dose date + 1 <i>(Start dose date of the combination is the first dose date between the first dose dates of the two drugs of the combination and</i>	Last dose date - Start dose date + 1

last dose date of the combination is the last date between the last dose dates of the two drugs of the combination)

Table 7.4.1-3: Administration of study therapy - Carboplatin and Pemetrexed

	Carboplatin	Pemetrexed
Dosing schedule per protocol	AUC 6 every 3 weeks	500 mg/m ² every 3 weeks
Dose	<i>Dose (AUC)</i> is defined as Total Dose administered (mg)/(creatinine clearance + 25). Dose administered in mg at each dosing date is collected on the CRF and creatinine clearance derived from the CRF data.	<i>Dose (mg/kg)</i> is defined as Total Dose administered (mg)/Most recent BSA (m ²). Dose administered in mg at each dosing date is collected on the CRF and BSA is derived from most recent weight and baseline height also collected on the CRF.
Cumulative Dose	<i>Cum Dose (AUC)</i> is the sum of the doses (AUC) administered to a subject.	<i>Cum dose (mg/m²)</i> is sum of the doses (mg/m ²) administered to a subject during the treatment period.
Relative dose intensity (%)	$\frac{\text{Cum dose (AUC)}}{(\text{Last dose date of Carbo} - \text{Start dose date of Carbo} + 21) \times 6 / 21} \times 100$	$\left[\frac{\text{Cum dose (mg/m}^2\text{)}}{(\text{Last pemetrexed dose date} - \text{pemetrexed Start dose date} + 21) \times 500 / 21} \right] \times 100$
Duration of study therapy	Last dose date - Start dose date + 1 <i>(Start dose date of the combination is the first dose date between the first dose dates of the two drugs of the combination and last dose date of the combination is the last date between the last dose dates of the two drugs of the combination)</i>	Last dose date - Start dose date + 1

Table 7.4.1-4: Administration of study therapy - Cisplatin/Pemetrexed

	Cisplatin	Pemetrexed
Dosing schedule per protocol	75 mg/m ² every 3 weeks	500 mg/m ² every 3 weeks
Dose	<i>Dose (mg/kg)</i> is defined as Total Dose administered (mg)/Most recent BSA (m ²). Dose administered in mg at each dosing date is collected on the CRF and BSA is derived from most recent weight and baseline height also collected on the CRF.	<i>Dose (mg/kg)</i> is defined as Total Dose administered (mg)/Most recent BSA (m ²). Dose administered in mg at each dosing date is collected on the CRF and BSA is derived from most recent

Table 7.4.1-4: Administration of study therapy - Cisplatin/Pemetrexed

	Cisplatin	Pemetrexed
		weight and baseline height also collected on the CRF.
Cumulative Dose	<i>Cum dose (mg/m²) is sum of the doses (mg/m²) administered to a subject during the treatment period.</i>	<i>Cum dose (mg/m²) is sum of the doses (mg/m²) administered to a subject during the treatment period.</i>
Relative dose intensity (%)	$\left[\frac{\text{Cum dose (mg/m}^2\text{)}}{(\text{Last pemetrexed dose date - pemetrexed Start dose date} + 21) \times 75 / 21} \right] \times 100$	$\left[\frac{\text{Cum dose (mg/m}^2\text{)}}{(\text{Last pemetrexed dose date - pemetrexed Start dose date} + 21) \times 500 / 21} \right] \times 100$
Duration of study therapy	Last dose date - Start dose date + 1 <i>(Start dose date of the combination is the first dose date between the first dose dates of the two drugs of the combination and last dose date of the combination is the last date between the last dose dates of the two drugs of the combination)</i>	Last dose date - Start dose date + 1

Creatinine clearance will be calculated using the Crockroft-Gault formula, defined as:

$$CrCL \text{ (ml/min)} = \frac{(140 - \text{age(in years)}) * \text{weight(in kg)}}{72 * \text{serum creatinine (in mg/dL)}}$$

for males, and

$$CrCL \text{ (ml/min)} = \frac{(140 - \text{age(in years)}) * \text{weight(in kg)}}{72 * \text{serum creatinine (in mg/dL)}} * 0.85$$

for females. The most recent weight will be used. If the computed creatinine clearance is more than 125 ml/min, then the creatinine clearance value should be capped at 125 ml/min for dose exposure computations.

7.4.2 Modifications of Study Therapy

7.4.2.1 Dose Delays

Each study medication (nivolumab, ipilimumab, and chemotherapy drugs) 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, ipilimumab, and chemotherapy drugs. 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 subjects with at least one dose delayed, the number of dose delays per subject, the reason for dose delay and the length of dose delay.

7.4.2.2 Infusion Interruptions and Rate Changes

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

The following parameters will be summarized by treatment group:

- Number of subjects with at least one dose infusion interruption, the reason for interruption, and the number of infusion interruptions per subject.
- Number of subjects with at least one IV infusion rate reduction, the reason for reduction and the number of infusion with IV rate reduction per subject.

7.4.2.3 Dose Escalations

Dose escalations (within subject) are not permitted for either nivolumab or ipilimumab.

7.4.2.4 Dose Reductions

Dose reductions (within subject) are not permitted for either nivolumab or ipilimumab.

Dose reductions of chemotherapy may be modified for toxicity. Dose levels of chemotherapy are defined in the protocol as follows:

Table 7.4.2.4-1: Dose Modifications of Chemotherapeutic Agents

Dose Level	Carboplatin	Pemetrexed	Paclitaxel	Cisplatin
Starting dose	AUC 6 or AUC 5	500 mg/m ²	200 mg/m ²	75 mg/m ²
First dose reduction	AUC 5 (if starting dose is AUC 6) or AUC 4 (if starting dose is AUC 5)	375 mg/m ²	150 mg/m ²	56 mg/m ²
Second dose reduction	AUC 4 (if starting dose is AUC 6) or AUC 3 (if starting dose is AUC 5)	250 mg/m ²	100 mg/m ²	38 mg/m ²
Third dose reduction	Discontinue	Discontinue	Discontinue	Discontinue

Source: Table 7.4.2.2-1 in CA2099LA Protocol

For any cycle, it will be defined as a dose reduction if the observed dose level (based on calculated administered dose) is below protocol specified dose level. Dose ranges for dose levels of platinum doublet chemotherapy are defined in [Table 7.4.2.4-2](#).

The reason for dose reduction as reported by the investigator will be tabulated for all instances of dose reductions based on the CRF page. A category “Not Reported” will be defined for all reductions with no reason reported by the investigator.

Table 7.4.2.4-2: Calculated Dose Ranges and Related Dose Levels

Dose Level	Carboplatin (AUC)	Pemetrexed (mg/m ²)	Paclitaxel (mg/m ²)	Cisplatin (mg/m ²)
Level 0	With pemetrexed: ≥ 5.5 (start of AUC 6) or ≥ 4.5 (start of AUC 5) With Paclitaxel: ≥ 5.5	≥ 437.5	With Carboplatin: ≥ 175	≥ 65.5
Level -1	With pemetrexed: < 5.5 and ≥ 4.5 (start of AUC 6) or < 4.5 and ≥ 3.5 (start of AUC 5) With Paclitaxel: < 5.5 and ≥ 4.5	< 437.5 and ≥ 312.5	With Carboplatin: < 175 and ≥ 125	< 65.5 and ≥ 47
Level -2	With pemetrexed: < 4.5 (start of AUC 6) or < 3.5 (start of AUC 5) With Paclitaxel: ≥ 5.5	< 312.5	With Carboplatin: < 125	< 47

The following will be summarized:

- Number and percentage of subjects with at least one dose reduction, the reason of dose reduction, and the number of dose reductions per subject.

7.4.2.5 Partial Discontinuation of Ipilimumab

Subjects treated with nivolumab and ipilimumab may discontinue ipilimumab and continue to receive nivolumab (ie, partial discontinuation). Continuation of ipilimumab after discontinuation of nivolumab is not allowed on study.

The following will be summarized for subjects receiving the combination of nivolumab and ipilimumab:

- Number and percentage of subjects who had partial discontinuation of ipilimumab
- Reason for partial discontinuation

Reason for partial discontinuation will be retrieved from the dosing CRF pages.

7.4.3 Concomitant Medications

Concomitant medications, defined as medications other than study medications which are taken at any time on-treatment (i.e. on or after the first day of study therapy and within 100 days following the last dose of study therapy), will be coded using the UMC WHO Drug Global Dictionary.

The following summary table will be provided:

- Concomitant medications (subjects with any concomitant medication, subjects by medication class and generic term)

Prior medications, defined as non-study medications with a start date before consent date, and current medications, defined as non-study medications with a start date before the first date of study medication and stop date after consent date, will be coded using the UMC WHO Drug Global Dictionary.

The following summary table will be provided:

- Prior/current medications (subjects with any prior/current medication, subjects by medication class and generic term)

By-subject listings will accompany the tables.

7.4.3.1 Immune modulating medication

Immune modulating concomitant medications are medications entered on an immune modulating medication form or available from the most current pre-defined list of immune modulating medications. The list of anatomic class, therapeutic class and generic name used for the selection at the time of the database lock will be provided.

The percentage of subjects who received immune modulating concomitant medication for

- management of adverse event
- premedication
- other use
- any use
- management of drug-related select adverse event (any grade, grade 3-5) by select AE category/subcategory (EU/ROW Submissions)
- management of IMAEs (any grade, grade 3-5) by IMAE category (US Submission)

will be reported separately for each treatment group (percentages of treated subjects by medication class and generic term).

For each category/subcategory of drug-related select AEs (any grade, grade 3-5) and IMAEs (any grade, grade 3-5), the following will be reported for each treatment group:

- The total immune modulating medication treatment duration (excluding overlaps), duration of high dose of corticosteroid, initial dose of corticosteroid, and tapering duration (summary statistics)

Duration represents the total duration the subject received the concomitant medication of interest. If the subject took the medication periodically, then DURATION in the summation of all use. Initial dose represents the dose of the concomitant medication of interest received at the start of the event. In the case multiple medications started on the same date, the highest equivalent dose is chosen and converted to mg/kg by dividing by the subject's recent weight.

These analyses, except the ones related to IMAEs will be conducted using the 30-day safety window. The analyses related to IMAEs will be conducted using the 100-day safety window.

7.4.3.2 Subsequent Cancer Therapy

Number and percentage of subjects receiving subsequent cancer therapies will be summarized for all randomized subjects. Categories include:

- Subsequent systemic therapy
 - Immunotherapy including commercial Nivolumab (anti-PD1 agents, anti-PD-L1 agents, anti-CTLA-4 agents and others) by drug name
 - Other anti-cancer agents excluding all immunotherapy (approved and investigational) by drug name
 - Palliative local therapy (including on-treatment)
- Subsequent surgery (limited to: tumor resection, curative; tumor resection, palliative; incisional biopsy; excisional biopsy towards censoring for progression free survival per primary definition) for treatment of tumors
- Subsequent radiotherapy for treatment of tumors
- Any combination of the above

A by-subject listing of subsequent cancer therapy will also be produced for all randomized subjects.

7.5 Efficacy

Principal analyses of progression free survival (PFS) and objective response rate (ORR) will be based on the Blinded Independent Central Review (BICR) evaluation, unless noted otherwise.

Analyses in this section will be tabulated for all randomized subjects by treatment group as randomized, unless otherwise specified.

Unless stated otherwise, whenever a stratified analysis is specified, the following stratifications factors (recorded at randomization as per IRT) will be used:

- Histology (Squamous vs. Non-squamous)
- Sex (Male vs. Female)
- PD-L1 Level ($\geq 1\%$ vs. $< 1\%$ or not quantifiable)

For assessing the secondary objectives of this study, a hierarchical testing procedure¹³ will be used so that the overall experiment-wise Type I error rate is two-sided 0.05.

A separate group sequential spending function will be used to adjust for the overall Type I error at the interim and final analysis for each of the formal comparisons. The actual level of significance which each objective is assessed will be determined based on the individual group sequential spending function and the amount of information spent at the time of the analysis. The significance level for ORR per BICR will be pre-specified as $\alpha=0.025$ for the interim analysis as an equal Bonferroni splitting of the two-sided $\alpha=0.05$.

The hierarchical ordering of the key secondary endpoints is as follows:

- 1) Progression Free Survival per BICR
- 2) Objective Response Rate per BICR

The statistical testing will be carried out using the following sequential procedure:

- 1) The primary endpoint of OS will be tested first.
- 2) If the p-value of OS is not statistically significant either at the interim or the final analysis, then no further statistical testing regarding the secondary endpoints will be conducted. However, estimates, along with their 95% CI, will be provided for those (i.e., medians and HR, rates and odds ratios) at the time of the final analysis.
- 3) If the p-value of OS is statistically significant either at the interim or the final analysis, PFS per BICR will be tested. If the p-value of PFS per BICR is statistically significant, ORR per BICR will be tested and the p-value will be provided. If the p-value of PFS is not statistically significant, then no statistical testing of ORR per BICR will be conducted. For ORR per BICR, estimates (rates and odds ratios) and associated 95% CI will be provided, regardless of the outcome of PFS per BICR testing.

Confidence intervals (CI) for primary and secondary endpoint analyses included in hierarchy will be based on nominal significance level adjusted for primary endpoints and interim analyses to preserve overall type one error rate.

Alpha (α) for the CI will be the same as nominal significance level for hypothesis testing. CIs for other endpoints will be at the two-sided 95% level. All p-values reported will be two-sided. P-values will be rounded to the fourth decimal place. Point estimates and confidence bounds for efficacy variables will be rounded to the second decimal place.

A by-subject listing of efficacy results will be presented including treatment group, treatment duration, BICR progression date, overall survival, death date, etc.

7.5.1 Analysis of Overall Survival

The primary objective of the study is to compare the overall survival between treatment groups in all randomized subjects.

Overall survival will be compared between the treatment groups (nivolumab and ipilimumab plus chemotherapy versus chemotherapy alone) at the interim and final analyses, using stratified log-rank test. The stratification factors will be histology (squamous vs. non-squamous), sex (male vs. female), and PD-L1 level ($\geq 1\%$ vs. $< 1\%$ or not quantifiable). A Lan DeMets α -spending function with O'Brien and Fleming type of boundary will be employed to determine the nominal significance levels for the interim and final analyses. The stratified hazard ratio between the treatment groups will be presented along with $100 \times (1 - \alpha)\%$ CI (adjusted for interim). In addition, two-sided p-value will also be reported for the analysis of OS.

OS will be estimated using the KM techniques. A two-sided 95% CI for median OS in each treatment group will be computed via the log-log transformation method. OS rates at fixed time points (e.g. 6 months, depending on the minimum follow-up) will be presented along with their

associated 95% CIs. These estimates will be derived from the Kaplan Meier estimate and corresponding CIs will be derived based on Greenwood formula for variance derivation and on log-log transformation applied on the survivor function.

The status of subjects who are censored in the OS KM analysis will be tabulated for each treatment group using the following categories:

- On-study (on-treatment, in follow-up)
- Off-study (lost to follow-up, withdraw consent, never treated)

A by-subject listing will be presented including treatment group, first and last dose date, whether the subject died, and if censored, the reason, event/censored date and OS duration.

7.5.2 Supportive Analyses of Overall Survival

The following OS sensitivity analyses will be conducted in all randomized subjects:

- 1) Delayed effect of immunotherapy interventions may cause a late separation in the OS KM curves and non-proportional hazards. OS will be compared between treatment groups via two-sided 0.05 stratified weighted log-rank test among subjects. The two-sided stratified weighted log-rank p-value will be reported using G ($\rho = 0$, $\gamma = 1$) weights, in the terminology of Fleming and Harrington¹⁴.

The Fleming Harrington test can be unstable, so it is possible, though uncommon, that the p-value for this trial will not be estimable.

The estimate of the OS hazard ratio in the period before and following 3 months will be calculated using a stratified time-dependent Cox model with effects for treatment and period-by-treatment interaction. In this model, period is a binary variable indicating pre- vs. post-3 months. Ties will be handled using the exact method. A two-sided 95% CI for the hazard ratio will also be presented.

- 2) Stratified log-rank test in the All Treated Subjects population, using arm as treated. This analysis will be performed only if the proportion of randomized but never treated subjects exceeds 10%.
- 3) A multivariate Cox regression model will be used in order to estimate the treatment effect after adjustment for possible imbalances in known or potential prognostic factors. The factors used in the randomization, which, by definition, will be balanced across treatment groups, will still be included in the model as stratification factors. However, all additional factors will be incorporated as covariates. The additional factors, which are all measured at baseline, will include:
 - a) ECOG performance status (0, ≥ 1)
 - b) Baseline histology (squamous, non-squamous) (only for randomized subjects)
 - c) Sex (male, female)
 - d) PD-L1 subgroups ($\geq 1\%$ vs $<1\%$)

The level of the covariate normally associated with the worst prognosis will be coded as the reference level. The hazard ratio associated with treatment and with each of the baseline covariates will be presented along with associated 95% CIs.

- 4) OS using stratification factors as obtained from the baseline CRF pages (instead of IRT). The hazard ratio associated with treatment will be presented along with the associated two-sided 95% CIs. This analysis will be performed only if at least one stratification variable/factor at randomization (as per IRT) and baseline are not concordant for at least 10% of the randomized subjects.
- 5) OS using an un-stratified log rank test. The hazard ratio associated with treatment will be presented along with the associated two-sided 95% CIs.
- 6) OS for subjects with no relevant protocol deviations. This analysis will be conducted only if there are more than 10% subjects with relevant protocol deviations. The hazard ratio associated with treatment will be presented along with the associated two-sided 95% CIs.
- 7) The method of Gail and Simon¹⁵ will be used to test for a qualitative interaction between treatment and strata. This test will be conducted at $\alpha=0.10$ level. The p-value reported from this specific analysis is for descriptive purposes alone.
- 8) To examine the assumption of proportional hazards in the Cox regression model, in addition to treatment, a time-dependent variable defined by treatment by time interaction will be added into the model. A two-sided Wald Chi-square p-value of less than 0.1 may indicate a potential nonconstant treatment effect. In that case, additional exploratory analyses may be performed.

7.5.3 **Subset Analyses of Overall Survival**

The influence of baseline and demographic characteristics on the treatment effect among all randomized subjects will be explored via exploratory subset analyses. The median OS based on KM product-limit method along with two-sided 95% CIs will be produced for the following subgroups:

- Age categorization (<65, ≥65 and <75, ≥75 and <85, ≥85, ≥75, and ≥65)
- Sex (male vs. female)
- Race (white, black, asian, other)
- Region (Europe, North America, Rest of World, Asia)
- ECOG performance (0, ≥1)
- Baseline histology (squamous, non-squamous)
- Smoking status (current/former, never smoked, unknown)
- PD-L1 subgroups (<1%, ≥1%, 1-49%, ≥50%)
- Disease stage (stage IV, recurrent to metastatic disease)
- CNS Metastasis (yes vs. no)
- Liver Metastasis (yes vs. no)
- Bone Metastasis (yes vs. no)
- Tumor Tissue TMB Evaluable (≥10 Mut/MB, < 10 Mut/MB, Overall)
- Tumor Tissue TMB Not Evaluable
- Blood TMB Evaluable (≥16 Mut/MB, < 16 Mut/MB, ≥20 Mut/MB and < 20 Mut/MB, Overall)
- Blood TMB Not Evaluable

A forest plot of the OS hazard ratios (along with the 95% CIs) will be produced for each level of the subgroups listed above. The analysis comparing treatment (i.e., Hazard Ratio) will be conducted if the number of subjects in the subgroup category is more than 10.

7.5.4 Interim Analysis of Overall Survival

An independent statistician external to BMS will perform the analysis. In addition to the formal planned interim analysis for OS, the Data Monitoring Committee (DMC) will have access to periodic un-blinded interim reports of efficacy and safety to allow a risk/benefit assessment. Details are included in the DMC charter.

One interim analysis of OS is planned. The interim analysis is scheduled when 322 events (approximately 80% of the targeted OS events) have been observed among randomized subjects. This formal comparison of OS will allow for early stopping for superiority, and the boundaries for declaring superiority will be derived based on the actual number of events using Lan-DeMets spending function with O'Brien and Fleming type of boundary. If the interim analysis is performed exactly at 322 events, the boundary in terms of statistical significance for declaring superiority would be 0.024 (HR boundary of 0.78). The boundary for declaring superiority in terms of statistical significance for the final analysis after 402 events would be 0.042 (HR boundary of 0.82).

The DMC will review the safety and efficacy data from the informal interim analyses and will determine if the study should continue with or without changes or if accrual should be stopped. Subject enrollment will continue while waiting for the DMC's decisions.

The chair of the DMC and the sponsor can call an unscheduled review of the safety data.

At the time of the formal interim analysis for superiority of OS, the DMC may recommend continuing or stopping the trial. If the trial continues beyond the formal interim analysis, the nominal critical point for the final OS analysis will be determined using the recalculated information fraction at the time of the interim analysis, as described above. The final OS hazard ratio and corresponding confidence interval will be reported whereby the confidence interval will be adjusted accordingly (i.e. using the recalculated nominal α level at the final analysis).

If the trial is stopped for superiority of OS at the interim, the p-value from the interim stratified log-rank test will be considered the final primary analysis result.

7.5.5 Analysis of Progression-Free Survival

One of the secondary objectives of the study is to compare the progression-free survival (as determined by BICR) between treatment groups in all randomized subjects.

Hierarchical testing of PFS (as determined by BICR) will be performed upon demonstration of superiority in OS at OS at interim or final analyses for all randomized subjects (see Section 7.5). P-values will not be presented for PFS if the primary endpoint OS was not statistically significant. In this case, descriptive analysis and treatment difference estimations for PFS will still be presented.

PFS will be compared between the treatment groups via stratified log-rank test among all randomized subjects. The stratification factors will be histology (squamous vs. non-squamous), sex (male vs. female), and PD-L1 level ($\geq 1\%$ vs. $< 1\%$ or not quantifiable).

The primary definition of PFS adjusting for subsequent anticancer therapy will be used in this analysis. The two-sided log-rank p-value will be reported.

The estimate of the PFS hazard ratio between treatment groups will be calculated using a stratified Cox proportional hazards model, with treatment as the sole covariate. Ties will be handled using the exact method. The $100*(1-\alpha)\%$ CI (adjusted for multiplicity, see Section 7.5) will be provided. For descriptive purposes, a two-sided 95% CI for the hazard ratio will also be presented.

The PFS function for each treatment group will be estimated using the KM product limit method and will be displayed graphically. A two-sided 95% CI for median PFS in each treatment group will be computed via the log-log transformation method. PFS rates at fixed time points (e.g. 6 months, depending on the minimum follow-up) will be presented along with their associated 95% CIs. These estimates will be derived from the Kaplan Meier estimate and corresponding CIs will be derived based on Greenwood¹⁶ formula for variance derivation and on log-log transformation applied on the survivor function¹⁷.

Analyses of PFS will also be conducted based on the secondary definition of PFS. These analyses will be the same as those specified above.

The source of PFS event (progression or death) will be summarized by treatment group. The status of subjects who are censored (as per primary definition of PFS) in the PFS KM analysis will be tabulated for each treatment group including the following categories:

- On-study (on-treatment, in follow-up)
- Off-study (lost to follow-up, withdraw consent, never treated)
- No baseline tumor assessment
- No on-study tumor assessment and no death
- Received subsequent anticancer therapy

A by-subject listing will be presented including treatment group, PFS duration under the primary definition, PFS duration on the ITT definition, whether the subject was censored under the primary definition, and if censored, the reason, and whether the subject was censored under the ITT definition, and if censored, the reason.

A by-subject listing of lesion evaluations per BICR will be presented.

7.5.6 Supportive Analyses of Progression-Free Survival

Unless otherwise specified, the following sensitivity analyses will be conducted using both the primary and the secondary definition of PFS in all randomized subjects:

- 1) PFS using stratification factors as obtained from the baseline CRF pages (instead of IRT). The hazard ratio associated with treatment will be presented along with the associated two-sided 95% CIs. This analysis will be performed only if at least one stratification variable/factor at

randomization (as per IRT) and baseline are not concordant for at least 10% of the randomized subjects.

- 2) PFS using the investigator's assessment. The hazard ratio associated with treatment and median PFS will be presented along with the associated two-sided 95% CIs.

A cross tabulation of PFS assessment by BICR versus PFS assessment by investigator will be presented, by treatment group. Concordance Rate of event will be computed as the frequency with which investigator and BICR agree on classification of a subject as event vs censored as a proportion of the total number of randomized subjects assessed by both the investigator and BICR.

- 3) PFS per BICR account for missing tumor assessment prior to PFS event (progression or death). This analysis will be performed only if at least 20% of events have missing prior tumor assessment within the primary efficacy analysis population. It will apply the following restriction to the primary definition: if the elapsed time between the PFS event and the last on-study assessment immediately prior to the event (or randomization date if no on-study scan) is two or more missed visits (more than 12 weeks + 10 days), the subject will be censored at his/her last tumor assessment prior to the PFS event (or randomization date if no on-study scan).

A by-subject listing of PFS assessment per BICR and investigator will be presented.

7.5.7 Subset Analyses of Progression-Free Survival

The influence of baseline and demographic characteristics on the treatment effect among all randomized subjects will be explored via exploratory subset analyses. The median PFS based on KM product-limit method along with two-sided 95% CIs will be produced for the following subgroups:

- Age categorization (<65, ≥65 and <75, ≥75 and <85, ≥85, ≥75, and ≥65)
- Sex (male vs. female)
- Race (white, black, asian, other)
- Region (Europe, North America, Rest of World, Asia)
- ECOG performance (0, ≥1)
- Baseline histology (squamous, non-squamous)
- Smoking status (current/former, never smoked, unknown)
- PD-L1 subgroups (<1%, ≥1%, 1-49%, ≥50%)
- Disease stage (stage IV, recurrent to metastatic disease)
- CNS Metastasis (yes vs. no)
- Liver Metastasis (yes vs. no)
- Bone Metastasis (yes vs. no)
- Tumor Tissue TMB Evaluable (≥10 Mut/MB, < 10 Mut/MB, Overall)
- Tumor Tissue TMB Not Evaluable
- Blood TMB Evaluable (≥16 Mut/MB, < 16 Mut/MB, ≥20 Mut/MB and < 20 Mut/MB, Overall)
- Blood TMB Not Evaluable

A forest plot of the PFS hazard ratios (along with the 95% CIs) will be produced for each level of the subgroups listed above. The analysis comparing treatment (i.e., Hazard Ratio) will be conducted if the number of subjects in the subgroup category is more than 10.

7.5.8 Current Status of PFS and OS Follow-up

The extent of follow-up for survival, defined as the time between randomization date and last known alive date (for subjects who are alive) or death date (for subjects who died), will be summarized descriptively (median, min, max, etc.) in months for all randomized subjects.

The currentness of follow-up for survival, defined as the time between last OS contact (i.e., last known alive date or death date) and cutoff date (defined by last patient last visit date), will be summarized in months for all randomized subjects. Subjects who died and subjects with last known alive date on or after data cut-off date will have zero value for currentness of follow-up.

Minimum follow-up for OS for all randomized subjects, defined as the time from cutoff date to last subject's randomization date, will be summarized in months for all randomized subjects.

Time from last evaluable tumor assessment to cutoff date in months will be summarized by treatment group and overall for all randomized subjects. Subjects who have a PFS event will be considered as current for this analysis. The secondary definition of PFS will be used for this summary.

In addition, time to treatment discontinuation will be summarized and presented by treatment group using a Kaplan-Meier curve whereby the last dose date will be the event date for those subjects who are off study therapy. Median duration of study therapy and associated 95% CI will be provided. Subjects who are still on study therapy will be censored on their last dose date.

A by-subject listing will also be produced to accompany the subject time from last evaluable tumor assessment.

7.5.9 Analysis of Objective Response Rate

One of the secondary objectives of the study is to estimate the ORR per BICR in the treatment groups among all randomized subjects.

Hierarchical testing of ORR will be performed upon demonstration of superiority in OS and superiority in PFS per BICR at OS interim or final analyses for all randomized subjects (see Section 7.5).

The number and percentage of subjects in each category of BOR per BICR (complete response [CR], partial response [PR], stable disease [SD], progressive disease [PD], or unable to determine [UTD]) will be presented, by treatment group. Estimates of response rate, along with its exact two-sided 95% CI by Clopper and Pearson¹⁸ will be presented, by treatment group. An estimate of the difference in response rates between the treatment groups along with the corresponding two sided CI will also be computed using the following Cochran-Mantel-Haenszel (CMH) method of weighting, adjusting for stratification factors¹⁹. The formula is

$$\hat{\theta} = \frac{\sum_i w_i \hat{\theta}}{\sum_i w_i} \sim N \left[\theta, \frac{\sum_i w_i^2 \left[\frac{p_{ix}(1-p_{ix})}{n_{ix}-1} + \frac{p_{iy}(1-p_{iy})}{n_{iy}-1} \right]}{(\sum_i w_i)^2} \right]$$

where $\hat{\theta} = p_{ix} - p_{iy}$ is the difference in rates in the *i*th stratum, $w_i = \frac{n_{ix}n_{iy}}{n_{ix}+n_{iy}}$, and n_{ix} and n_{iy} are the number of subjects randomized to treatments x and y, respectively, in the *i*th stratum. Stratification factors will be PD-L1 level, histology, and sex, as entered into the IRT.

A two sided 95% CI for odds ratio of response between the treatment groups will also be computed.

Similar analyses will be repeated based on the investigator's assessment of ORR. A cross tabulation of BICR best response versus the investigator best response will be presented, by treatment group and by response categories. Concordance Rate of Responders will be computed as the frequency with which investigator and BICR agree on classification of a subject as responder vs. non responder/UTD as a proportion of the total number of randomized subjects assessed by both the investigator and BICR.

The following subject-level graphics will also be provided:

- For the responders only, time courses of the following events of interest will be graphically displayed: tumor response, progression, last dose received, and death.
- For response evaluable subjects (randomized subjects with baseline and at least one on-study tumor assessment),
 - A bar plot showing the best % reduction from baseline in sum of diameter of target lesions based on BICR assessment for each subject will be produced (excluding assessments after PD and assessments after start of subsequent anti-cancer therapy).
 - A plot of individual time course of tumor burden change per BICR assessment will be produced.

A by-subject listing of best overall response will be presented including treatment group, best overall response per BICR and dates of CR/PR/progression.

A by-subject listing of per time point tumor response per BICR will be presented.

7.5.10 Subset Analyses of Objective Response

The influence of baseline and demographic characteristics on the treatment effect will be explored via exploratory subset analysis. BICR assessment of ORR will be summarized for the following subgroups:

- Age categorization (<65, ≥65 and <75, ≥75 and <85, ≥85, ≥75, and ≥65)
- Sex (male vs. female)
- Race (white, black, asian, other)
- Region (Europe, North America, Rest of World, Asia)
- ECOG performance (0, ≥1)
- Baseline histology (squamous, non-squamous)

- Smoking status (current/former, never smoked, unknown)
- PD-L1 subgroups (<1%, ≥1%, 1-49%, ≥50%)
- Disease stage (stage IV, recurrent to metastatic disease)
- CNS Metastasis (yes vs. no)
- Liver Metastasis (yes vs. no)
- Bone Metastasis (yes vs. no)
- Tumor Tissue TMB Evaluable (≥10 Mut/MB, < 10 Mut/MB, Overall)
- Tumor Tissue TMB Not Evaluable
- Blood TMB Evaluable (≥16 Mut/MB, < 16 Mut/MB, ≥20 Mut/MB and < 20 Mut/MB, Overall)
- Blood TMB Not Evaluable

A forest plot of treatment effect on ORR per BICR in the above subgroups will be produced. The un-weighted differences in ORR between the two treatment groups and corresponding 95% two-sided CI using the method of Newcombe^{20,21} will be provided. The analysis comparing treatment (i.e., ORR difference) will be conducted if the number of subjects in the subgroup category is more than 10.

7.5.11 Time to Tumor Response and Duration of Response

Duration of response (DOR) and time to response (TTR) will also be evaluated for subjects who achieved confirmed PR or CR. The DOR for each treatment group will be estimated using the Kaplan-Meier (KM) product limit method and will be displayed graphically. A table will be produced presenting number of events, number of subjects involved, medians, and 95% CIs for the medians. Median values of DOR, along with two-sided 95% CI in each treatment group will be computed based on a log-log transformation method.

The status of subjects who are censored in the DOR KM analysis will be tabulated for each treatment group including the following categories:

- Ongoing follow-up (current [last scan within adequate window vs cutoff date], not current)
- Off-study (lost to follow-up, withdraw consent, never treated)
- Received subsequent anticancer therapy.

TTR, which does not involve censoring, will be summarized by treatment group in all responders using descriptive statistics.

Cumulative Response Rates will be tabulated for Week 6, Month 3, 6, 8, and 12, and overall response rate will be provided.

A by-subject listing will be presented including treatment group, best response, time to response, duration of response, whether the subject was censored for duration of response, and, if so, the reason.

7.5.12 PFS2

One of the exploratory objectives of the study is to compare PFS2 between treatment groups in all randomized subjects.

PFS2 will be analyzed similarly to PFS:

- Median values based on KM method, along with two-sided 95% CI using Brookmeyer and Crowley method will be calculated. The estimate of standard error will be calculated using the Greenwood formula;
- PFS2 will be graphically displayed along with the median and 95% CI.

A by-subject listing of PFS and PFS2 will be provided.

7.6 Safety

Analyses in this section will be tabulated for all treated subjects 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/ROW Submissions)

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.

7.6.7 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 or ipilimumab 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
- 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 Laboratory Parameters

The analysis population for each laboratory test is restricted to treated subjects who underwent that laboratory test. Laboratory tests (in addition to the tests specified below) with CTC criteria collected in the specific studies may also be included in the summaries.

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
- ALP > 1.5 x ULN
- Concurrent (within 1 day) ALT or AST > 3 x ULN and total bilirubin > 1.5 x ULN
- Concurrent (within 30 days) ALT or AST > 3 x ULN and total bilirubin > 1.5 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 \leq 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 \geq 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 \geq 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 \leq 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 and pulse oximetry (i.e. % oxygen saturation) collected on the CRF will be provided in separate listings.

7.6.12 Physical Measurements

Physical measurements will be listed by subject.

7.6.13 Non-Protocol Medical Procedures

Non-protocol medical procedures will be listed by subject.

[REDACTED]

Clinical implications

[REDACTED]

[REDACTED]

[REDACTED]

7.6.15 Pregnancy

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

7.6.16 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 - < 85 vs. ≥ 85 vs. ≥ 75 vs. ≥ 65)
- Region (Europe, North America, Rest of World, Asia)

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

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

7.8.1 Distribution of PD-L1 Expression

Descriptive statistics of PD-L1 expression:

- Listing of all PD-L1 IHC data, all randomized subjects.
- Summary of tumor specimen acquisition and characteristics, all randomized subjects.
- Summary statistics of PD-L1 expression in all randomized subjects with quantifiable PD-L1 expression.
- Summary of BOR and ORR by baseline PD-L1 expression in all randomized subjects.
- Cumulative distribution plot of baseline PD-L1 expression versus population percentile in all randomized subjects with quantifiable PD-L1 expression.
- Box plots of PD-L1 expression versus Response Status (BICR assessment) in all randomized subjects with quantifiable PD-L1 expression.
- Waterfall plot of Individual PD-L1 expression in all randomized subjects with quantifiable PD-L1 expression.

7.8.2 Tumor Mutation Burden (TMB)

7.8.2.1 Tumor Tissue TMB

The analyses are based on tumor tissue TMB evaluable subjects, defined as subjects with tissue TMB data available. It is known that not all subjects will provide tumor tissue TMB data, due to factors such as available remaining tissue and inherent failure rates of the TMB process.

The descriptive analyses of tumor tissue TMB at baseline will be conducted:

- Listing of all tumor tissue TMB data.
- Summary of tumor specimen characteristics
- Cumulative distribution plot of TMB at baseline versus population percentile in all subjects with evaluable tumor tissue TMB.

In addition, the joint distribution of PD-L1 and tumor tissue TMB among both PD-L1 and tumor tissue TMB-evaluable subjects will be examined.

The pre-specified tumor tissue TMB thresholds (≥ 10 Mut/MB, < 10 Mut/MB) will be used to categorize subjects in subgroups. The following descriptive analyses of efficacy endpoints (OS, PFS per BICR, and ORR) among tumor tissue TMB-evaluable subjects will be presented separately by treatment group.

A summary of OS/PFS per BICR by tumor tissue TMB subgroups (high, low) will be presented according to TMB cutpoints (10 Mut/MB). OS/PFS function for each tumor tissue TMB subgroup will be estimated using the KM product limit method for each treatment arm and will be displayed graphically. A two-sided 95% CI for median OS/PFS in each arm will be computed via log-log transformation method. Hazard ratios of OS/PFS and corresponding two-sided 95% CIs will be estimated using an unstratified Cox proportional hazards model, with treatment group as a single covariate.

The comparison of ORR per BICR by tumor tissue TMB subgroups (high, low) will be conducted. An estimate of the difference in ORRs and corresponding 95% CI will be calculated using unstratified CMH methodology.

The number and percentage of subjects in each category of BOR per BICR will be presented by tumor tissue TMB subgroups, by treatment group. A summary of DOR (as determined per BICR) by tumor tissue TMB subgroups will also be presented. A two-sided 95% CI for median DOR in each arm will be computed via the log-log transformation method.

7.8.2.2 Blood TMB

The analyses are based on blood TMB evaluable subjects, defined as subjects with blood TMB data available. It is known that not all subjects will provide blood TMB data, due to inherent failure rates of the TMB process.

The descriptive analyses of blood TMB at baseline will be conducted:

- Listing of all blood TMB data.
- Cumulative distribution plot of TMB at baseline versus population percentile in all subjects with evaluable blood TMB.

In addition, the joint distribution of PD-L1 and blood TMB among both PD-L1 and blood TMB-evaluable subjects will be examined.

The pre-specified blood TMB thresholds (≥ 16 Mut/MB and < 16 Mut/MB; ≥ 20 Mut/MB and < 20 Mut/MB) will be used to categorize subjects in subgroups. The following descriptive analyses of efficacy endpoints (OS, PFS per BICR, and ORR) among blood TMB-evaluable subjects will be presented separately by treatment group.

A summary of OS/PFS per BICR by blood TMB subgroups (high, low) will be presented according to TMB cutpoints (16 Mut/MB and 20 Mut/MB). OS/PFS function for each tumor tissue TMB subgroup will be estimated using the KM product limit method for each treatment arm and will be displayed graphically. A two-sided 95% CI for median OS/PFS in each arm will be computed via log-log transformation method. Hazard ratios of OS/PFS and corresponding two-sided 95% CIs will be estimated using an unstratified Cox proportional hazards model, with treatment group as a single covariate.

The comparison of ORR per BICR by blood TMB subgroups (high, low) will be conducted. An estimate of the difference in ORRs and corresponding 95% CI will be calculated using unstratified CMH methodology.

The number and percentage of subjects in each category of BOR per BICR will be presented by blood TMB subgroups, by treatment group. A summary of DOR (as determined per BICR) by blood TMB subgroups will also be presented. A two-sided 95% CI for median DOR in each arm will be computed via the log-log transformation method.

7.8.2.3 Correlation between Tissue TMB and Blood TMB

As an exploratory analysis, the correlation between baseline tumor tissue TMB and blood TMB will be assessed using Spearman correlation based on rank order. The 95% CI for Spearman correlation will be provided based on Fisher transformation.

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[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 Specification²²
- For missing and partial adverse event resolution dates, imputation will be performed as follows:
 - If only the day of the month is missing, the last day of the month will be used to replace the missing day. If the imputed date is after the death date or the last known alive date, then the latest known alive date or death date is considered as the resolution date.
 - If the day and month are missing or a date is completely missing, it will be considered as missing.
- Missing and partial non-study medication domain dates will be imputed using the derivation algorithm described in 4.1.3 of BMS Non-Study Medication Domain Requirements Specification²³.
- Missing and partial radiotherapy and surgery dates will be imputed using algorithm described in [APPENDIX 2](#).
- 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 alive date 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 alive date.
 - 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 progression after start of study therapy, 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. In case of the date of death is present and complete, the imputed progression date will be compared to the date of death. The minimum of the imputed progression date and date of death will be considered as the date of progression.
 - If the day and month are missing or a date is completely missing, it will be considered as missing.
- For date of progression to prior therapies, 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.

- For other partial/missing dates, the following conventions were 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:

$$1 \text{ month} = 30.4375 \text{ days and } 1 \text{ year} = 365.25 \text{ days.}$$

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

$$\text{Duration} = (\text{Last date} - \text{first date} + 1)$$

Last known alive date will be defined based on all appropriate dates collected on the CRF.

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

The analyses described in this SAP will be included in the study-specific CSR(s). Refer to the Data Presentation Plan for mock-ups of all tables and listings.

10 DOCUMENT HISTORY

Table 10-1: Document History

Version Number	Author(s)	Description
1	Stephanie Meadows-Shropshire	Initial version.

