Statistical Analysis Plan for

A DOUBLE-BLIND, RANDOMIZED, TWO ARM PHASE 2 STUDY OF NIVOLUMAB IN COMBINATION WITH IPILIMUMAB VERSUS NIVOLUMAB IN COMBINATION WITH IPILIMUMAB PLACEBO IN RECURRENT OR METASTATIC SQUAMOUS CELL CARCINOMA OF THE HEAD AND NECK (SCCHN)

NCT02823574

30-May-2018

STATISTICAL ANALYSIS PLAN
FOR CLINICAL STUDY REPORT

A DOUBLE-BLIND, RANDOMIZED, TWO ARM PHASE 2 STUDY OF NIVOLUMAB IN COMBINATION WITH IPILIMUMAB VERSUS NIVOLUMAB IN COMBINATION WITH IPILIMUMAB PLACEBO IN RECURRENT OR METASTATIC SQUAMOUS CELL CARCINOMA OF THE HEAD AND NECK (SCCHN)

PROTOCOL(S) CA209714

VERSION # 3.0

TABLE OF CONTENTS

[redacted]

LIST OF TABLES

LIST OF FIGURES

[redacted]

2 STUDY DESCRIPTION

2.1 Study Design

This is a randomized (2:1), double blinded, Phase 2 trial in subjects ≥ 18 years old with untreated metastatic SCCHN or recurrent SCCHN that is not amenable to curative therapy, evaluating nivolumab in combination with ipilimumab, vs. nivolumab in combination with ipilimumab placebo as a first line treatment.

Approximately 396 subjects will be randomized to the two treatment arms in a 2:1 ratio and stratified by platinum refractory subgroup (yes vs no), PD-L1 status (expressing vs non-expressing/non-evaluable/indeterminate) and HPV p-16 status (oropharyngeal Cancer HPV

Positive (including HPV Positive with unknown primary site) vs oropharyngeal Cancer HPV Negative /non-oropharyngeal Cancer).

Subjects will be treated with one of the following:

Arm A: Nivo/Ipi Combo

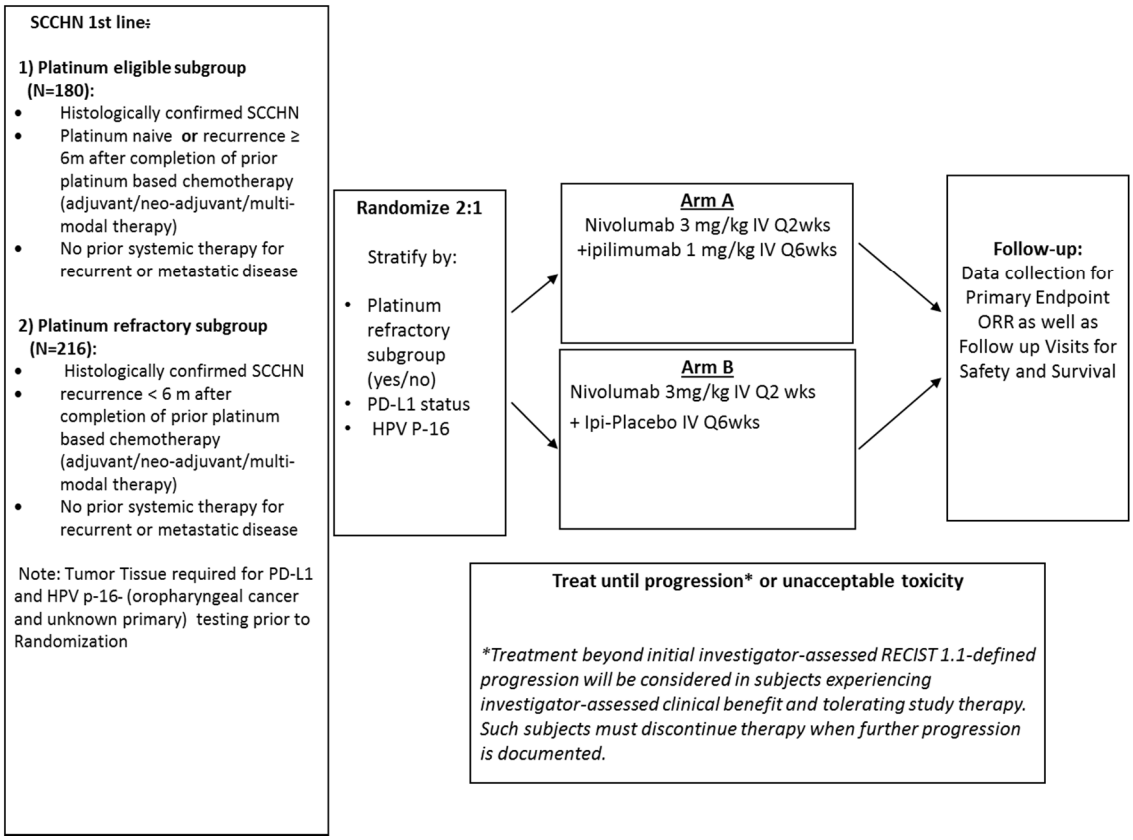
- Nivolumab 3 mg/kg IV will be administered every 2 weeks
- Ipilimumab 1 mg/kg IV will be administered every 6 weeks following the administration of nivolumab.

Arm B: Nivo and Ipilimumab-placebo

- Nivolumab 3 mg/kg IV will be administered every 2 weeks
- Ipilimumab Placebo IV will be administered every 6 weeks following the administration of nivolumab.

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

Figure 2.1-1: Study Design Schema



Tumor progression or response endpoints will be assessed using Response Evaluation Criteria In Solid Tumors (RECIST 1.1) criteria. Treatment with study medication will continue until RECIST 1.1 defined progression, unacceptable toxicity, or withdrawal of consent. Dose reductions will not be allowed for nivolumab or ipilimumab or ipilimumab-placebo. Treatment beyond initial

investigator-assessed progression (either clinical or radiographical) is permitted for nivolumab and ipilimumab or ipilimumab-placebo if the subject has an investigator-assessed clinical benefit and is tolerating study drug. The primary analysis of ORR for each of the platinum eligible and platinum refractory subgroups will be conducted after the last randomized subject in that subgroup has been followed for at least 6 months. Survival follow-up may continue for up to 5 years from the time of this analysis. The study will end once survival follow-up has concluded.

2.2 Treatment Assignment

After the subject's initial eligibility is established and informed consent has been obtained, the subject must be enrolled into the study by calling an interactive voice response system (IVRS) to obtain the subject number.

Once enrolled in the IVRS, subjects that have met all eligibility criteria will be ready to be randomized through the IVRS in a 2:1 ratio, with twice as many subjects receiving nivolumab and ipilimumab as nivolumab and ipilimumab-placebo:

- Arm A: Nivo 3 mg/kg Q2W + Ipi 1 mg/kg Q6W
- Arm B: Nivo 3 mg/kg Q2W + Ipi Placebo Q6W

The stratification factors are:

- Platinum refractory subgroup (yes vs no).
- PD-L1 status (expressing vs non-expressing or non-evaluable or indeterminate). PD-L1 status is set to expressing when tumor cell PD-L1 expression >=1%. (up to 20% of randomized subjects can be included into the study as non-evaluable or indeterminate).
- HPV p-16 status (oropharyngeal Cancer HPV Positive (including unknow primary with HPV p16 positive) vs oropharyngeal Cancer HPV Negative /non-oropharyngeal Cancer).

The randomization procedures will be carried out via permuted blocks within each stratum. The exact procedures for using the IVRS will be detailed in the IVRS manual.

2.3 Blinding and Unblinding

The Sponsor, subjects, investigator and site staff will be blinded to the study drug administered. Each investigative site must assign an unblinded pharmacist/designee, and an unblinded site monitor will be assigned by the sponsor to provide oversight of drug supply and other unblinded study documentation.

Before breaking the blind of an individual subject's treatment during the blinded portion of the study, the investigator should determine that the unblinded information is necessary, ie, that it will alter the subject's immediate management. The Principal Investigator should only call for emergency unblinding during the blinded portion of the study AFTER the decision to discontinue the subject has been made.

The Sponsor will remain blinded until the time of the planned interim analysis of the primary endpoint.

For this study, the method of unblinding is through the Interactive Response Technology (IRT).

2.4 Protocol Amendments

This SAP (version3.0) incorporates the following amendments:

Table 2.4-1: Protocol Amendments

Amendment	Date of Issue	Summary of Major Changes
Revised Protocol 01 (Incorporate Amendment 01)	19-Jul-2016	This amendment is updating the Protocol to reduce some of the burden on subjects and investigators, ensure consistency with the most current version of the nivolumab Investigator Brochure (v15) and the nivolumab label, and provide clarification of definitions and minor corrections. The major changes are as follows: • Change the title and contact information of the medical monitor. • Change on the biomarkers collection schedule • PK and IMG Follow up visit samples no longer required to be collected. • Updated Algorithms for Renal, Pulmonary, Hepatic and Skin to match with updated Nivolumab IB v15. (includes correction identified in nivo IB v15 erratum)
Revised Protocol 02 (Incorporate Amendment 05)	27-Jun-2017	1.Increase in the size of the platinum eligible population in order to provide greater estimation precision. 2. Alignment of protocol with responses to regulatory authorities. 3. Minor changes to eligible criteria and study processes. 4. Clarification of outstanding issues and corretion of typographical errors.
Revised Protocol 03	14-Nov-2017	Correct mistakes introduced by Revised Protocol 02 and add a potential interim analysis.
Revised Protocol 04	16-May-2018	1. Incorporate Administrative letter (change in study personnel) [redacted] 3. Clarify timing of the planned interim analysis and final analysis and update statistical analysis plan 4. Other minor edits and clarifications.

2.5 Data monitoring Committee

A Data Monitoring Committee (DMC) will be utilized to provide general oversight and safety considerations for this study. The DMC will provide advice to the Sponsor regarding actions the committee deems necessary for the continuing protection of subjects treated in this study. The DMC will be charged with assessing such actions in light of an acceptable risk/benefit/ profile for nivolumab in combination with placebo, and nivolumab in combination with ipilimumab. The DMC will act in an advisory capacity to BMS and will monitor subject safety data for the study approximately every 6 months for the duration of the trial.

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.

3 OBJECTIVES

3.1 Primary

To compare the ORR and assess DOR of the treatment of nivolumab in combination with ipilimumab vs. nivolumab in combination with ipilimumab placebo, as determined by a blinded independent central review (BICR) using RECIST 1.1 criteria, for first line treatment of recurrent or metastatic SCCHN in the platinum refractory setting.

3.2 Secondary

- To estimate the ORR and assess DOR of the treatment of nivolumab in combination with ipilimumab vs nivolumab in combination with ipilimumab placebo, as determined by BICR using RECIST 1.1 criteria, for first line treatment of recurrent or metastatic SCCHN in the platinum eligible setting
- To assess progression-free survival (PFS), as determined by BICR, of nivolumab in combination with ipilimumab vs. nivolumab in combination with ipilimumab placebo for first line treatment of recurrent or metastatic SCCHN in the platinum eligible and platinum refractory settings, separately and overall
- To assess overall survival (OS), of nivolumab in combination with ipilimumab vs. nivolumab in combination with ipilimiumab placebo for first line treatment of recurrent or metastatic SCCHN in the platinum eligible and platinum refractory settings, separately and overall
- To assess efficacy (ORR, DOR, PFS and OS) by PD-L1 expression of nivolumab in combination with ipilumumab compared to nivolumab in combination with ipilimumab placebo for first line treatment of recurrent or metastatic SCCHN in the platinum eligible and platinum refractory settings, separately and overall
- To assess efficacy (ORR, DOR, PFS and OS) by HPV p-16 status of nivolumab in combination with ipilumumab compared to nivolumab in combination with ipilimumab placebo for first line treatment of recurrent or metastatic SCCHN in the platinum eligible and platinum refractory settings, separately and overall

- To evaluate tumor mutation burden, as a potential predictive biomarker of efficacy (such as ORR, DOR, PFS and OS) of nivolumab in combination with ipilimumab or ipilimumab placebo for first-line treatment of recurrent or metastatic SCCHN in the platinum eligible and platinum refractory settings, separately and overall.

[redacted]

4 ENDPOINTS

4.1 Primary Endpoint

4.1.1 Objective Response Rate by BICR and Duration of Objective Response in Platinum Refractory Subgroup

Objective response rate by BICR and Duration of Objective Response (DOR) in the platinum refractory subgroup are the primary endpoints of this study. ORR is defined as the number of randomized subjects with a best overall response (BOR) of a complete response (CR) or partial

response (PR), assessed by BICR per RECIST 1.1, divided by the number of randomized subjects for each treatment group. Note that to achieve a best response of CR or PR, confirmation is required.

Best Overall Response

BOR is defined as the best response designation, as determined by BICR, recorded between the date of randomization and the date of progression, as assessed by BICR per RECIST 1.1, or the date of subsequent anticancer therapy (including tumor-directed radiotherapy and tumor-directed surgery), whichever occurs first. For subjects without evidence of RECIST 1.1 progression or subsequent anticancer therapy, all available response designations will contribute to the BOR assessment. For subjects who continue treatment beyond progression, the BOR will be determined based on response designations up to the time of initial RECIST 1.1 progression.

Tumor assessments should occur every 6 weeks (± 1 week) for the first 48 weeks, then every 12 weeks (± 1 week) until disease progression or subsequent therapy (whichever occurs later).

Duration of Objective Response

Duration of Objective Response (DOR) is defined as the time between the date of first confirmed response (CR or PR) to the date of the first documented progression as determined by BICR (per RECIST 1.1), or death due to any cause, whichever occurs first. Subjects who neither progress nor die will be censored on the date of their last evaluable tumor assessment. Subjects who started any subsequent anti-cancer therapy without a prior reported progression will be censored at the last evaluable tumor assessment prior to or on initiation of the subsequent anti-cancer therapy. DOR will be evaluated for responders (i.e. subjects with confirmed CR or PR) only.

Time to Objective Response

Time to Objective Response (TTR) is a supporting endpoint for ORR. TTR is defined as the time from randomization to the date of the first confirmed response (CR or PR), as assessed by BICR. TTR will be evaluated for responders (i.e. subjects with a BOR of confirmed CR or PR) only.

4.2 Secondary Endpoints

4.2.1 Objective Response Rate by BICR and Duration of Objective Response in Platinum Eligible Subgroup

The secondary endpoints of ORR and DOR by BICR in platinum eligible subgroup are defined similarly as described for the primary endpoint.

4.2.2 Progression-free Survival by BICR

4.2.2.1 Primary Definition of Progression-free Survival by BICR

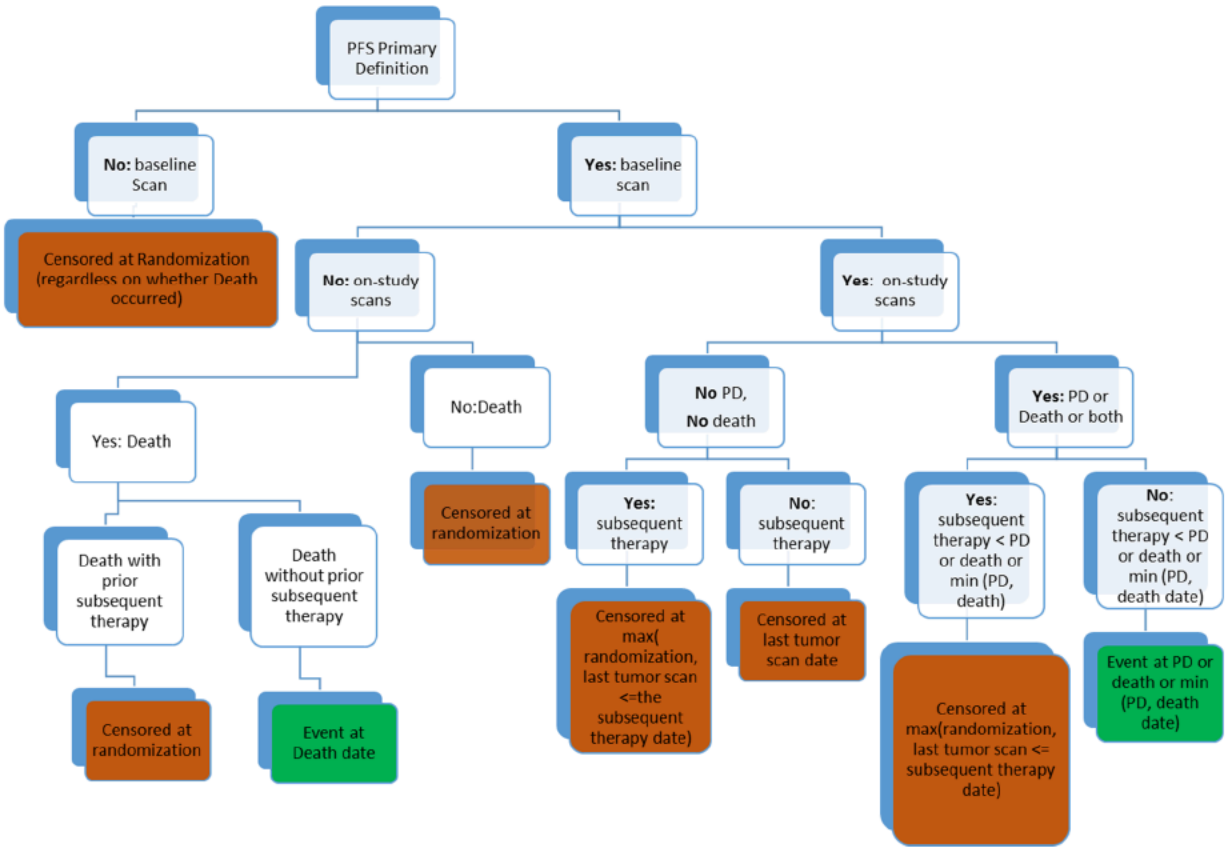
The BICR-assessed PFS is defined as the time from randomization to the date of first documented disease progression, as assessed by the BICR using RECIST 1.1 criteria, or death due to any cause, whichever occurs first. Subjects who died without a reported progression will be considered to have progressed on the date of their death. 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 the date they were randomized. Subjects who started any subsequent anti-cancer therapy, including tumor-directed radiotherapy and tumor-directed surgery, without a prior reported progression will be censored on the date of their last evaluable tumor assessment prior to/on the initiation of the subsequent anti-cancer therapy.

- For subjects without any baseline TA, they will be censored at randomization date
- For subjects with baseline TA but without any on study TA, if no death information is available, subjects will be censored at randomization date; if death information is available, subjects will be either an event at death date if no subsequent therapy is in the record or censored at randomization date if subsequent therapy exists.
- For subjects with baseline and on-study TA,
  a) Subjects who did not have disease progression per BICR assessment or die will be censored on the date of last tumor assessment on study if there's no subsequent therapy or last tumor assessment conducted on/prior to the subsequent therapy if there's one.
  b) Subjects who did have disease progression per BICR assessment or die will be an event at disease progression per BICR assessment or death date if there's no subsequent therapy before the event or censored at the last tumor assessment date conducted on/prior to the subsequent therapy date.

Further explanation for various censoring scenarios for the primary definition of PFS are presented in Figure 4.2.2.1-1.

Figure 4.2.2.1-1: Graphic display of PFS Primary Definition



4.2.2.2 Secondary Definition of Progression-free Survival by BICR

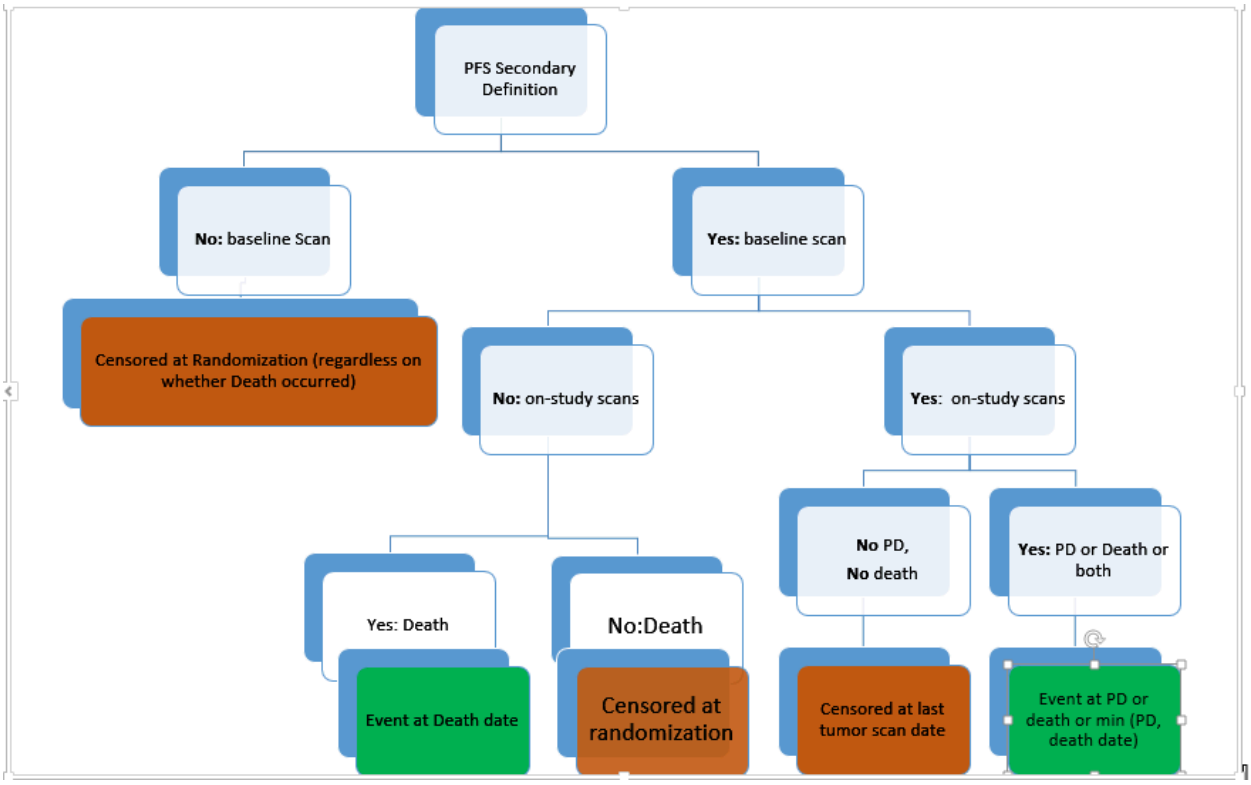
The secondary definition of PFS (ITT definition) is the time between the date of randomization and the first date of documented progression, as determined by BICR (as per RECIST 1.1 criteria), or death due to any cause, whichever occurs first. Subjects who died without a reported progression will be considered to have progressed on the date of their death. A subject's PFS time will not be censored for subsequent anti-cancer therapy prior to a progression event.

The following censoring rules will be applied for the secondary definition of PFS.

- For subjects without any baseline TA, they will be censored at randomization date
- For subjects with baseline TA but without any on study TA, if no death information is available, subjects will be censored at randomization date; if death information is available, subjects will be an event at death date.
- For subjects with baseline and on-study TA,
  a) Subjects who did not have disease progression per BICR assessment or die will be censored on the date of last tumor assessment on study.
  b) Subjects who did have disease progression per BICR assessment or die will be an event at disease progression per BICR assessment or death date.

Further explanation for various censoring scenarios for the secondary definition of PFS are presented in Figure 4.2.2.2-1.

Figure 4.2.2.2-1: Graphic Display of PFS Secondary Definition



4.2.2.3 Progression-free survival rate

The progression free survival rate at time *T* is defined as the probability that a subject has not progressed and is alive at time *T* following randomization.

Tumor assessments should occur every 6 weeks (± 1 week) for the first 48 weeks, then every 12 weeks (± 1 week) until disease progression or subsequent therapy (whichever occurs later).

4.2.3 Overall Survival

OS is defined as the time between the date of randomization and the date of death. For subjects without documentation of death, OS will be censored on the last date the subject was known to be alive. The overall survival rate at time *T* is defined as the probability that a subject is alive at time *T* following randomization.

[redacted]

[redacted]

[redacted]

[redacted]

5 SAMPLE SIZE AND POWER

The primary objective is to compare the ORR of nivolumab in combination with ipilimumab vs nivolumab in combination with ipilimumab placebo, as determined by BICR using Response Evaluation Criteria In Solid Tumors (RECIST 1.1) criteria in platinum refractory subjects with recurrent or metastatic SCCHN. The alpha level for the ORR is adjusted for one planned interim analysis using Lan-DeMets alpha spending function with O'Brien-Fleming boundaries, which will be determined by the fraction of subjects included in the interim analysis. Given that the interim analysis for the primary endpoint is expected to be performed when 70% of platinum refractory subjects have reached 6 months follow up after randomization, the alpha level is expected to be 0.015 for the interim analysis and 0.045 for the final analysis.

Approximately 396 subjects (216 platinum refractory subjects and 180 platinum eligible subjects) will be randomized to either nivolumab plus ipilimumab or nivolumab in combination with ipilimumab placebo in a 2:1 ratio.

A sample size of 216 randomized platinum refractory subjects (144 and 72 respectively) will provide 84% power for testing the odds ratio of nivolumab plus ipilimumab over Nivolumab in combination with placebo, with a 0.050 two-sided significance level, assuming ORR of 35% and 15% (odds ratio of 3.051) in the nivolumab plus ipilimumab and nivolumab in combination with ipilimumab placebo treatment group respectively (odds ratio of proportions test using EAST v6).

Only the primary endpoint will be tested, none of the secondary endpoints.

The first secondary objective is to estimate the ORR of the treatment of nivolumab in combination with ipilimumab vs nivolumab in combination with ipilimumab placebo, as determined by BICR using RECIST 1.1 criteria for first line treatment of recurrent or metastatic SCCHN in the platinum eligible setting.

Approximately 180 subjects (120 and 60 respectively) will be randomized in the platinum eligible subgroup. For a sample size of 120 subjects randomized to nivolumab plus ipilimumab, the maximum width of the exact two-sided 95% confidence interval (CI) is 18.6% when the ORR is expected to be in the 10% to 55% range. The table below summarizes the 95% exact CI when observed ORRs are between 10% and 55% respectively.

Table 5-1: Observed ORR with Exact 95% CI in Platinum Eligible Subjects randomized to nivolumab plus ipilimumab (N=120).

Table 5-1: Observed ORR with Exact 95% CI in Platinum Eligible Subjects randomized to nivolumab plus ipilimumab (N=120)

Observed ORR	95% Exact CI
10%	(5.3%, 16.8%)
20%	(13.3%, 28.3%)
30%	(22.0%, 39.0%)
40%	(31.2%, 49.3%)

Observed ORR	95% Exact CI
50%	(40.7%, 59.3%)
55%	(45.7%, 64.1%)

About 60 subjects will be randomized to nivolumab plus ipilimumab placebo arm. The following table summarizes the 95% exact CI when observed ORRs are between 10% and 55% respectively.

Table 5-2: Observed ORR with Exact 95% CI in Platinum Eligible Subjects randomized to nivolumab plus ipilimumab placebo (N=60)

Observed ORR	95% Exact CI
10%	(3.8%, 20.5%)
20%	(10.8%, 32.3%)
30%	(18.8%, 43.2%)
40%	(27.6%, 53.5%)
50%	(36.8%, 63.2%)
55%	(41.6%, 67.9%)

6 STUDY PERIODS, TREATMENT REGIMENS AND POPULATIONS FOR ANALYSES

6.1 Study Periods

6.1.1 Baseline Period

Baseline evaluations or events will be defined as evaluations or events that occur before the date and time of the first dose of study treatment.

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 and vital signs) will be defined as evaluations with a date on or prior to the day of first dose of study treatment

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

6.1.2 Post Baseline Period

On-treatment AEs will be defined as AEs with an onset date and time on or after the date and time of the first dose of study treatment (or with an onset date on or after the day of first dose of study

treatment if time is not collected or is missing). For subjects who are off study treatment, AEs will be counted as on-treatment if the event occurred within 30 days (or 100 days depending on analysis) (see Core Safety SAP), of the last dose of study treatment. No "subtracting rule" will be applied when an AE occurs both pre-treatment and post-treatment with the same preferred term and grade.

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

6.2 Treatment Regimens

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

Subjects will be randomized, in a 2:1 ratio, to one of the following treatment arms:

- Arm A: Nivo 3 mg/kg Q2W + Ipi 1 mg/kg Q6W
- Arm B: Nivo 3 mg/kg Q2W + Ipi Placebo Q6W

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

6.3 Populations for Analyses

- Enrolled subjects: All subjects who signed an informed consent form and were registered into the IVRS. This is the dataset for pre-treatment disposition.
- Randomized subjects: All enrolled subjects who were randomized to any treatment arm in the study. This is the primary dataset for analyses of study conduct, study population, and efficacy.
- Treated subjects: All randomized subjects who received at least one dose of study drug (nivolumab or ipilimumab). This is the primary dataset for analyses of exposure and safety.
- Response evaluable subjects: randomized subjects whose change in the sum of diameters of target lesions was assessed (i.e., target lesion measurements were made at baseline and at least one on-study visit.)
- PK subjects: All subjects treated with Nivolomab or ipilimumab with available serum time-concentration data.
- Biomarker subjects:
    - All tested PD-L1 subjects: All subjects who had a tumor biopsy specimen available for assessment of PD-L1 expression. [Note: This population is not just limited to randomized subjects since some screen failure subjects may have tumor biopsy specimens for PD-L1 testing]
    - All randomized subjects with quantifiable PD-L1 expression at baseline: See definitions of baseline and quantifiable PD-L1 expression in Section 4.3.6.
    - All randomized subjects with quantifiable TMB at baseline.

- PRO Subjects: All randomized subjects with available Patient Reported Outcomes (PRO) data.
- Immunogenicity (ADA evaluable) subjects: Treated subjects with baseline and at least one post-baseline pre-infusion immunogencity assessment. See Core/Integrated Statistical Analysis Plan for Immunogenicity

All analyses will be performed using the treatment arm as randomized (intent to treat), with the exception of dosing and safety, for which the treatment arm as received will be used.

7 STATISTICAL ANALYSES

7.1 General Methods

Unless otherwise noted, the bulleted titles in the following subsections describe tabulations of discrete variables, by the frequency and proportion of subjects falling into each category, grouped by treatment (with total). Percentages given in these tables will be rounded and, therefore, may not always sum to 100%. Continuous variables will be summarized by treatment group (with total) using descriptive statistics; ie, number of non-missing observations (n), mean, standard deviation (STD), median, minimum, maximum and quartiles. If a missing category is not being presented in the data display, only those subjects with non-missing values for the parameter being assessed are included in the percentage calculation.

Time to event distributions (i.e. progression free survival, overall survival, time to response, and duration of response) will be estimated using the Kaplan-Meier product limit method. When appropriate, estimates of the median time to event will be presented, along with associated 95% CIs. The confidence interval for the median will be obtained based on the log{-log} transformation of the survival function.[5] Kaplan-Meier estimates of rates at fixed time points (e.g. OS at 12 months) will be presented along with 95% confidence intervals for those rates. The confidence interval for the rate at time t_0, $S(t_0)$, will be obtained by first computing the interval for $\log(-\log(S(t_0))$ and then back transforming.[6] (The methods specified above for computing the confidence intervals for median time to event and for survival rates are the default methods in SAS.[7])

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

Confidence intervals for binomial proportions will be derived using the Clopper-Pearson method.[8]

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

7.2 Study Conduct

7.2.1 Accrual

The accrual pattern will be summarized per country, investigational site, and per month for all enrolled and randomized subjects. Randomization date, first dosing date, country, investigational site will be presented in a by subject listing of accrual.

Furthermore, the accrual pattern will be summarized by the stratification factors for all randomized subjects alone.

7.2.2 Relevant Protocol Deviations

The following programmable deviations will be considered as relevant protocol deviations and summarized by treatment group and total in all randomized subjects, overall and for platinum refractory and platinum eligible subgroups separately. Non-programmable relevant eligibility and on-treatment protocol deviations, as well as significant (both programmable and non-programmable) eligibility and on-treatment protocol deviations will be reported through ClinSIGHT listings.

Eligibility:

- Subjects without measurable disease at baseline.
- Subject with baseline ECOG PS > 1
- Subjects who received prior treatment with an anti-PD-1, anti-PD-L1, anti-CD137, anti-CTLA-4 antibody
- Any prior treatment with any non-platinum containing systemic therapy regimen for SCCHN (adjuvant, neoadjuvant, or multi-modal treatment)

On-study:

- Subjects receiving concurrent anti-cancer therapy (defined as chemotherapy, hormonal immunotherapy, radiation therapy directed to target lesions, surgery (surgical resection of tumor only), standard or investigational agents for treatment of SCCHN).
- Subject treated differently as randomized (subjects who received the wrong treatment, excluding the never treated).

Listings will also be provided.

7.3 Study Population

Unless otherwise specified, analyses will be performed for all randomized subjects by treatment group as randomized, overall and for platinum refractory and platinum eligible subgroups separately.

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 will be tabulated by treatment group as randomized.

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

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

7.3.2 Demographics and Other Baseline Characteristics

The following baseline characteristics (CRF source/clinical database) will be summarized by treatment arm as randomized:

- Age (descriptive statistics)
- Age categorization (<65 vs. ≥65)
- Age categorization (< 65, 65- <75, ≥ 75)
- Gender (Male vs. Female)
- Race (White, Black , Asian, Other)
- Ethnicity (Hispanic/Latino, Not Hispanic/Latino, Unknown)*

  * Required for US subjects only
- Country
- Region (North America vs. EU vs.ROW)
- Baseline ECOG performance status(0,1)
- Platinum-refractory (yes/no)
- HPV-16 status (Oropharyngeal HPV positive, HPV positive with unknown primary site, Oropharyngeal HPV negative, other site with HPV positive, other site with HPV negative, other site with no HPV result available)
- PD-L1 (expressing (>=1%), non-expressing (<1%), non-evaluable, indeterminate)
- Subjects with CNS metastasis at Baseline (Yes, No) - per radiographic tumor screening assessment
- Smoking status (former/current, never, unknown)
- Alcohol status (former, current, never, unknown)
- Tumor mutation burden (TMB)

Baseline Stratification Factors as entered into the IVRS will tabulated for all randomized subjects:

- PD-L1 status (expressing vs non-expressing/non-evaluable/indeterminate)
- HPV p-16 status (oropharyngeal HPV p16 positive (including unknow primary with HPV p16 positive) vs oropharyngeal HPV p16 negative or nonoropharyngeal)
- Platinum refractory subgroup (yes vs no).

7.3.3 Baseline Disease Characteristics

The following baseline disease characteristics will be summarized.

- Disease stage at study entry (locally recurrent, locally recurrent and metastatic, metastatic)
- Disease stage at initial diagnosis (stage I, II, III, IVA, IVB, IVC)
- TMN Classification at initial diagnosis
- Time from Initial Disease Diagnosis to Randomization (< 1 year, 1 - < 2 year, 2 - < 3 year, 3 - < 4 year, 4 - < 5 year, ≥ 5 year).
- Sites of diseases (all lesions) per BICR
- Number of disease sites per subject (all lesions) per BICR
- Number of target lesions, non-target lesions and disease sites at baseline as per BICR
- Tumor burden: sum of the diameters of target lesions at baseline per BICR
- Sites of diseases (all lesions) per investigator
- Subsites of diseases per investigator

7.3.4 Prior Therapy

Prior systemic anti-center therapy

Prior systemic anti-center therapy (which will be identified from the CRF page "Prior Systemic Therapy") will be categorized using the BMS WHO drug dictionary and will be summarized overall and by treatment group. The latest version of the drug dictionary at the time of the analysis will be used. The number and percentage of subjects with any prior systemic anti-cancer therapy and the number of subjects receiving each therapy (generic term) will be presented.

- Prior systemic cancer therapy summary classified by setting, therapeutic class and generic name
- Best response to most recent prior systemic therapy regimen (CR, PR, SD, PD, Unable to determine, Not applicable, Not reported)
- Time from completion of most recent prior systemic therapy regimen to randomization (< 3 months, 3 -< 6 months, 6 -< 12 months, 12 -< 24 months, >= 24 months)*

In addition, frequencies and percents will be provided for the following:

- Prior surgery related to cancer (yes, no, not reported) and reason for prior surgery
- Prior radiotherapy (yes, no, not reported) , reason for prior RT, total dose received
- Time from completion of most recent radiotherapy to randomization (< 3 months, 3 - 6 months, > 6 months)

*The date of the last dose of platinum or the date of progression may be incomplete. If only the day is missing, it will be imputed as the 1st of the month. If either the month or the year is missing, the date will be considered missing.

Other prior therapy:

- Prior/current non-study medication classified by anatomic and therapeutic classes.

Agents and medication will be reported using the generic name. A listing by subject will also be provided.

7.3.5 Baseline Safety Laboratory Tests

Baseline safety laboratory evaluations will be the latest available samples taken on or before Cycle 1, Day 1. Baseline safety laboratory values will be presented by CTC severity grade, by treatment group. These analyses will be restricted to treated subjects. Separate tables will be generated for hematology and Serum Chemistry.

CTC grades will be derived as part of the analysis data set programming using version 4.0.

7.3.6 Medical history

General medical history will be tabulated and also listed by subject.

7.3.7 Baseline Examinations

Subjects with abnormal baseline physical examination will be tabulated by examination criteria and by treatment arm.

7.3.8 Baseline Physical Examination

Summary of baseline height and weight will be tabulated and presented.

7.3.9 Discrepancies Between IVRS and CRF stratification factors

Summary tables (cross-tabulations) by treatment arm for each baseline stratification factor will be provided to show any discrepancies between what was reported through IVRS vs. CRF data/clinical database.

- PD-L1 status (expressing vs non-expressing/non-evaluable/indeterminate)
- HPV p-16 status (oropharyngeal HPV p16 positive (including unknow primary with HPV p16 positive) vs oropharyngeal HPV p16 negative or nonoropharyngeal)
- Platinum refractory subgroup (yes vs no).

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, overall and for platinum refractory and platinum eligible subgroups separately, unless otherwise specified.

7.4.1 Administration of Study Therapy

The following parameters will be summarized (descriptive statistics) by treatment group, using all treated population.

- Time from randomization to first dose of study therapy (≤ 3, 4 - 5, 6-7, 8-14, 15 - 21, > 21 to 28, >28 days)
- Duration of treatment (in month) will be presented 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 treatment and associated 95% CI will be provided. Subjects who are still on study therapy will be censored on their last dose date.

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

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

Table 7.4.1-1 below summarizes the key parameters used for analyses of dosing.

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

	Nivolumab	Ipilimumab
Dosing schedule per protocol	3 mg/kg every 2 weeks	1 mg/kg every 6 weeks
Dose	Dose (mg/kg) is defined as the minimum of [vial strength (mg/mL) *total volume infused (mL)] /most recent weight (kg) and nominal dose (mg/kg) * total volume infused (mL) /total volume prepared (mL)	Dose (mg/kg) is defined as the minimum of [vial strength (mg/mL) *total volume infused (mL)] /most recent weight (kg) and nominal dose (mg/kg) * total volume infused (mL) /total volume prepared (mL)
Cumulative Dose	Cum dose (mg/kg) is sum of the doses (mg/kg) administered to a subject during the treatment period.	Cum dose (mg/kg) is sum of the doses (mg/kg) administered to a subject during the treatment period.
Relative dose intensity (%)	Cum dose (mg/kg) / [(Last Nivo dose date - Start Nivo dose date + 14) x 3 / 14] x 100	Cum dose (mg/kg) / [(Last Ipi dose date - Start Ipi dose date + 42) x 1 / 42] x 100
Duration of treatment	Last dose date - Start dose date +1	Last dose date - Start dose date +1

Volume infused, volume prepared, and weight are collected on the CRF

7.4.2 Modifications of Study Therapy

7.4.2.1 Dose delays

For nivolumab, treatment may be delayed for up to a maximum of 6 weeks from the last dose. Subjects receiving nivolumab may be dosed no less than 12 days from the previous dose.

Each study medication infusion may be delayed independently. 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) from previous dose for any given study medication. The length of delay for nivolumab is defined as (duration of previous cycle in days - 14). The length of delay for ipilimumab or ipilimumab-placebo is defined as (duration of previous cycle in days - 42). Dose delays for each study medication will be divided into following categories: on-time, 4 - 7 days, 8 - 14 days, 15 - 42, > 42 days. Reason for dose delay will be retrieved from the respective CRF

dosing pages for each study medication. If different reasons for delay are recorded (in the same cycle), both the reasons will be reported.

The following parameters will be summarized, by drug:

- Number of subjects with at least one dose delayed, number of doses delayed per subject, length of dose delay and Reason for dose delay

7.4.2.2 Infusion Interruptions and Rate Changes

The following parameters will be summarized from the CRF pages for all treatment drugs:

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

7.4.2.3 Dose Reductions/Escalation

Dose reduction/escalation is not permitted for nivolumab or ipilimumab or ipilimumab-placebo.

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 WHO Drug Dictionary.

The following summary tables will be provided:

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

A by-subject listing will accompany the table.

7.5 Efficacy

Unless otherwise specified, efficacy analyses will be performed for all randomized subjects by treatment group as randomized, overall and for platinum refractory and platinum eligible subgroups separately.

CIs for efficacy endpoints will be at the two-sided 95% level, except for the primary endpoint ORR for platinum-refractory subjects where CI is anticipated to be provided at 98.5% level for the interim and 95.5% level for the final analysis (see Section 7.5.1). The p-values presented in the clinical study report will be rounded to the fourth decimal place. Point estimates and confidence intervals for efficacy variables will be rounded to the second decimal place. P-value will only be provided for the comparison of ORR between nivolumab in combination with ipilimumab vs nivolumab in combination with ipilimumab placebo in platinum refractory randomized subjects.

7.5.1 Primary Endpoint ORR in Platinum Refractory Subgroup

The primary endpoint of the study is ORR in platinum refractory subgroup, as determined by a blinded independent central review (BICR) using RECIST 1.1 criteria.

7.5.1.1 Primary Analysis of ORR

The comparison of ORR between nivolumab in combination with ipilimumab vs nivolumab in combination with ipilimumab placebo in platinum refractory randomized subjects will be carried out using a two-sided Cochran-Mantel-Haenszel (CMH) test stratified by the stratification factors PD-L1 status (expressing vs non-expressing/non-evaluable/indeterminate) and HPV p-16 status (oropharyngeal HPV p16 positive (including unknow primary with HPV p16 positive) vs oropharyngeal HPV p16 negative or nonoropharyngeal) as recorded in the IVRS. The significance level will be adjusted for an interim analysis which will be performed when around 70% of platinum-refractory subjects have been followed for at least 6 months after randomization. Using Lan-DeMets alpha spending function with O'Brien-Fleming boundaries, the significance level for this comparison will be determined by the fraction of subjects included in the interim analysis. If the analysis is performed when exactly 70% of platinum refractory subjects have reached 6 months follow up after randomization, the alpha level is 0.015 for the interim and 0.045 for the final. An associated odds ratio (nivolumab/ipilimumab vs. nivolumab/ipilimumab placebo) and corresponding Ci will be calculated. ORR will be summarized by a binomial response rate and its corresponding two-sided exact CI using Clopper-Pearson method for each treatment group. An estimate of the difference in ORRs between arms (nivolumab/ipilimumab vs. nivolumab/ipilimumab placebo) and corresponding CI will be calculated using Cochran-Mantel-Haenszel (CMH) methodology and adjusted by the stratification factors as recorded in the IVRS.[9]

The formula is

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

where $\hat{\theta} = p_{ix} - p_{iy}$ is the difference in rates in the ith 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 ith stratum.

BOR will be summarized by response category for each treatment group. The number and percentage of subjects in each category of best overall response (BOR) per BICR (complete response [CR], partial response [PR]], stable disease [SD], progressive disease [PD], or unable to determine [UD]) will be presented, by treatment group.

7.5.1.2 Sensitivity Analyses of ORR

The following sensitivity analyses will be conducted using the primary definition of ORR in the randomized subjects:

- If one stratification variable at IVRS and at baseline (CRF and clinical database) disagrees for at least 10% of the randomized subjects, similar analysis of ORR as primary analysis will be performed using the strata as determined at baseline.
- ORR per BICR for subjects with no relevant deviation. This analysis will be conducted only if there are more than 5% subjects with relevant protocol deviations.
- ORR per BICR in treated subjects using treatment group "as treated" if more than 10% randomized subjects in any treatment group were never treated.
- ORR per BICR analysis for response evaluable subjects (randomized subjects with baseline and at least one on-study tumor assessment), if more than 10% randomized subjects in any treatment group were not response evaluable.
- ORR per BICR where all response designations contribute to the BOR determination, regardless of start of subsequent anti-cancer therapy.

To assess concordance between BICR and investigator assessments, BOR will be cross-tabulated by treatment group and assessment type (Investigator vs. BICR). Concordance Rate of Responders will be computed as the frequency with which Investigator and BICR agree on classification of a subject as responder/non responder as a proportion of the total number of subjects assessed within each treatment group.

7.5.1.3 Subset Analyses of ORR

The influence of baseline and demographic characteristics on the treatment effect will be explored via exploratory subset analyses for the following factors using the primary definition of ORR per BICR.

- Age categorization (<65 vs. 65-<75 vs ≥75)
- Gender (Male vs. Female)
- Race(White, Black, Asian, Other)
- Region (North America vs. EU vs.ROW)
- ECOG performance status(0, 1)
- Disease stage at study entry (locally recurrent vs locally recurrent and metastatic vs metastatic)
- HPV p-16 status (oropharyngeal Cancer HPV Positive (including unknown primary with HPV Positive) vs oropharyngeal Cancer HPV Negative /non-oropharyngeal Cancer) (from IVRS)
- HPV p-16 status (oropharyngeal Cancer HPV Positive (including unknown primary with HPV Positive) vs oropharyngeal Cancer HPV Negative /non-oropharyngeal Cancer) (based on CRF data)
- Oropharyngeal Cancer HPV p-16 status (oropharyngeal Cancer HPV Positive vs oropharyngeal Cancer HPV Negative) (based on CRF data)
- PD-L1 status (expressing vs non-expressing/non-evaluable/indeterminate) (from IVRS)

- PD-L1 status (expressing, non-expressing, non-evaluable/indeterminate) (based on clinical database)
- Tumor associated immune cell PD-L1
- PD-L1 combined positive score (CPS)
- Tumor mutational burden
- Gene expression profilePlatinum-refractory (yes/no) (from IVRS)
- Platinum-refractory (yes/no) (based on CRF data)
- Tobacco use (Former/Current, Never, Unknown)
- Alcohol use (Former, Current, Never, Unknown)
- Prior surgery (Yes, No)
- Prior radiotherapy (Yes, No)
- Site of primary tumor (Oral Cavity, Larynx, Oropharynx, Hypopharynx, Unknown P16 primary location, Other)
- Best response to the most recent regimen (CR/PR, SD, PD, Unknown or Not reported)
- Disease stage at initial diagnosis (stage I, II, III, IVA, IVB, IVC)
- Time from initial disease diagnosis to randomization (< 1 year, 1 - < 2 year, 2 - < 3 year, 3 - < 4 year, 4 - < 5 year, ≥ 5 year)

BICR-determined ORR along with the exact 95% CI using Clopper-Pearson method will be displayed for each treatment group. If a subset category has less than 10 subjects per treatment group, ORR will not be computed/displayed.

A forest plot of the unweighted ORR response rate differences (nivolumab/ipilimumab vs. nivolumab/ipilimumab placebo) and corresponding 95% CIs using the method of Newcombe will be produced for each level of the subgroups listed above.

7.5.1.4 Duration of Objective Response

Duration of response (DOR) curves in each treatment group will be estimated using KM product-limit method for subjects with a BOR of PR or CR per BICR assessment. Median DOR, corresponding two-sided 95% CI, and range will be calculated. Proportion of subjects with duration of response at least 6 months will be estimated with corresponding two-sided 95% CI. Summary statistics will be computed based on a log-log transformed CI for the survivor function.

7.5.1.5 Time to Objective Response

Summary statistics of time to objective response (TTR) will be provided for each treatment group for subjects who achieve PR or CR.

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

[redacted]

[redacted]

7.5.2 Secondary Endpoint ORR in Platinum Eligible Subgroup

The secondary endpoint of ORR in platinum eligible subgroup will be summarized by a binomial response rate and its corresponding two-sided 95% exact CI using Clopper-Pearson method for each treatment group. ORR will also be estimated in baseline subgroups as specified in Section 7.5.1.3. DOR and TTR will be analyzed similarly as described in Sections 7.5.1.4 and 7.5.1.5. Analyses described in Sections 7.5.1.6 and 7.5.1.7 will also be performed similarly for platinum eligible subgroup.

7.5.3 Progression Free Survival

7.5.3.1 PFS primary analysis

The PFS function based on BICR assessment for each treatment arm will be estimated using Kaplan Meier technique and will be displayed graphically in the platinum refractory and platinum eligible subgroups, separately. The primary definition of PFS will be used in this analysis (Section 4.2.2). Median Survival time along with 95% CI will be constructed based on a log-log transformed CI for the survivor function. The hazard ratio and the 95% confidence interval will be estimated in a stratified Cox proportional hazards model using the randomized arm as a single covariate, stratified by PD-L1 status and HPV p-16 status, in the platinum refractory and platinum eligible subgroups, separately. PFS rates at 3, 6, 12 and 24 months, with 95% CIs will be estimated using Kaplan-Meier methodology. Minimum follow-up must be greater than or equal to the time-point to generate the rate.

The source of PFS event as assessed by BICR (RECIST 1.1 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 randomized treatment group using the following categories:

- On-study (on-treatment, follow-up)
- Off-study (lost to follow-up, withdrawn consent, never treated)
- Received subsequent anticancer therapy

- No baseline tumor assessment

7.5.3.2 Sensitivity Analyses of Progression Free Survival

The following sensitivity analyses will be conducted using the primary definition of PFS in the randomized subjects:

- PFS per BICR using stratification factors as obtained from the baseline CRF pages and clinical database (instead of IVRS). The hazard ratio associated with treatment will be presented along with the associated two-sided 95% CIs. This analysis will be performed only if the stratification variable/factor at randomization (as per IVRS) and baseline are not concordant for at least 10% of randomized subjects.
- PFS per BICR using an un-stratified Cox model. The hazard ratio associated with treatment will be presented along with the associated two-sided 95% CIs.
- The primary PFS per BICR analysis will be repeated using secondary PFS definition which accounts for the tumor scans post subsequent therapies. The hazard ratio associated with treatment and median PFS will be presented along with the associated two-sided 95% CIs.

A by-subject listing for all randomized subjects will be presented including treatment arm, IVRS stratification level, PFS under the primary definition with censoring status and reason, PFS under the secondary definition with censoring status and reason.

7.5.3.3 Current status of Progression Free Survival follow-up

Time from last tumor assessment to data cut-off in months will be summarized by treatment arm. 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.

By-subject listings will also be produced to accompany the subject time from last tumor assessment table for all randomized subjects.

[redacted]

7.5.3.6 Subsequent Therapy

The following information pertaining to subsequent therapies will be summarized for all randomized subjects by treatment arm, as randomized:

Number and percentage of subjects receiving subsequent therapies including:

- Immunotherapy (anti-PD1 agents, anti-PD-L1 agents, anti-CTLA-4 agents and others) by drug name
- Other agents excluding Nivolumab (approved and investigational) by drug name
- Surgery
- Radiotherapy
- Any combination of the above

A subject listing of follow-up therapy will be produced for subjects who had any subsequent therapy for all randomized subjects.

7.5.4 Overall Survival

7.5.4.1 Primary Analysis of Overall Survival

OS distribution for each treatment arm will be estimated using Kaplan Meier technique and will be displayed graphically in the platinum refractory and platinum eligible subgroups, separately. Median Survival time in each arm along with 95% CI will be constructed based on a log-log transformed CI for the survivor function. The hazard ratio and the 95% confidence interval will be estimated in a stratified Cox proportional hazards model using the randomized arm as a single covariate, stratified by PD-L1 status and HPV p-16 status, in the platinum refractory and platinum eligible subgroups, separately. OS rates at 6, 9, 12, 18, 24, 36, 48 and 60 months with 95% CIs will be estimated using Kaplan-Meier methodology. These analyses will be performed only if the minimum follow-up for OS has been reached corresponding to that endpoint.

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

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

To examine the assumption of proportional hazards in the Cox regression model, in addition to treatment, a time-dependent variable, in the form of a 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 non-constant hazard ratio.

7.5.4.2 Overall Survival Sensitivity Analyses

The following OS sensitivity analyses will be performed.

- OS using stratification factors as obtained from the baseline CRF pages and clinical database (instead of IVRS). The hazard ratio associated with treatment will be presented along with the associated two-sided 95% CIs. This analysis will be performed only if the stratification

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

- OS using an unstratified Cox model. The hazard ratio associated with treatment will be presented along with the associated two-sided 95% CIs.

A by-subject listing will be presented including treatment arm, OS, whether the subject was censored, and if censored, the reason.

7.5.4.3 Subset Analyses of Overall Survival

The influence of baseline and demographic characteristics on the treatment effect among randomized subjects will be explored via exploratory subset analyses for the same factors as ORR subset analyses (Section 7.5.1.3).

Number of events and median OS along with 95% CI will be displayed for each treatment group. A forest plot of the OS unstratified hazard ratios (along with 95% CI) will be produced for each level of the subgroups listed above. If subset category has less than 10 subjects per treatment group, HR will not be computed/displayed.

7.5.4.4 Subject Follow-up for Overall Survival

The minimum follow-up will be reported. The minimum follow-up is defined as the time interval between the last patient's randomization date and the clinical cutoff date.

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

The currentness of follow-up, defined as the time between last OS contact (i.e., last known date alive or death date) and data cut-off date, will be summarized by treatment group. Subjects who died before data cut-off date will automatically have zero value for currentness of follow-up. For subjects with last known date alive after data cut-off date, they will have zero value for currentness of follow-up as well. The currentness of follow-up will be categorized into the following categories: 0 days, 1-3 months, 3-6 months, 6-9 months, 9-12 months and ≥ 12 months.

[redacted]

7.5.6 Interim Analysis

One interim analysis will be performed when around 70% of platinum-refractory subjects have been followed for at least 6 months after randomization. The comparison of ORR between nivolumab in combination with ipilimumab vs nivolumab in combination with ipilimumab placebo in platinum refractory randomized subjects will be carried out using a two-sided Cochran-Mantel-Haenszel (CMH) test stratified by the stratification factors as recorded in the IVRS. Using Lan-

DeMets alpha spending function with O'Brien-Fleming boundaries, the significance level for this comparison will be determined by the fraction of subjects included in the interim analysis. If the analysis is performed when exactly 70% of platinum refractory subjects have reached 6 months follow up after randomization, the alpha level is 0.015. Point estimate of ORR, and its corresponding two-sided 98.5% exact CI will be provided for each treatment group. The difference in ORRs between the two treatment groups and its corresponding 98.5% CI will be calculated using CMH methodology and adjusted by the stratification factors as recorded in the IVRS.

The following sensitivity analyses may be conducted using the primary definition of ORR in platinum-refractory subjects:

- If one stratification variable at IVRS and at baseline (CRF and clinical database) disagrees for at least 10% of the randomized subjects, similar analysis of ORR will be performed using the strata as determined at baseline.
- ORR per BICR for subjects with no relevant deviation. This analysis will be conducted only if there are more than 5% subjects with relevant protocol deviations.

The DOR distribution will be estimated using Kaplan-Meier technique. Median survival time along with 95% CI will be provided for each treatment group.

The PFS and OS distribution will also be estimated using Kaplan-Meier technique. Median survival time along with 95% CI will be provided for each treatment group. The hazard ratio and the 95% confidence interval will be estimated in a stratified Cox proportional hazards model using the randomized arm as a single covariate, stratified by PD-L1 status and HPV p-16 status.

Similar descriptive analyses may also be performed for the platinum-eligible subgroups, overall study population (both platinum-eligible and platinum refractory subgroups combined together) and by tumor cellPD-L1, tumor associated immune cell PD-L1, PD-L1 combined positive score (CPS), tumor mutation burden, gene expression profile,, HPV status, and selected other biomarkers status in each subgroup alone or combined.

7.6 Safety

For all safety related analyses, refer to the Core Safety SAP. Safety will be summarized for treated subjects, by arm, as treated. Unless otherwise specified, analyses will be performed for overall and for platinum refractory and platinum eligible subgroups separately.

7.6.1 Deaths

See Core Safety SAP.

7.6.2 Serious Adverse Events

See Core Safety SAP.

7.6.3 Adverse Events Leading to Discontinuation of Study Therapy

See Core Safety SAP.

7.6.4 Adverse Events Leading to Dose Delay of Study Therapy

See Core Safety SAP.

7.6.5 Adverse Events

See Core Safety SAP.

7.6.6 Select Adverse Events

See Core Safety SAP.

7.6.7 Immune Modulating Medication

See Core Safety SAP.

7.6.8 Multiple Events

See Core Safety SAP.

7.6.9 Clinical Laboratory Evaluations

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

7.6.9.1 Hematology

See Core Safety SAP.

7.6.9.2 Serum Chemistry

See Core Safety SAP.

7.6.10 Immunogenicity

See Core Safety SAP

7.6.11 Vital Signs and Pulse Oximetry

See Core Safety SAP.

7.7 Pregnancy

See Core Safety SAP.

[redacted]

[redacted]

7.9.2 Analyses

7.9.2.1 Distribution of PD-L1 Expression

Descriptive statistics of PD-L1 expression and PD-L1 status:

- Listing of all PD-L1 IHC data and treatment group, all PD-L1 tested subjects.
- Summary of tumor specimen acquisition and characteristics by treatment group for all randomized subjects.
- Cumulative distribution plot of PD-L1 expression at baseline versus population percentile by treatment group and overall, all evaluable PD-L1 subjects.
- Frequency of PD-L1 expression status by treatment group and overall, all randomized PD-L1 subjects, including indeterminate and not evaluable if over 5% of subjects in the population fall in this category.
- Box plot of PD-L1 expression by treatment group and overall, all randomized subjects

7.9.2.2 Association Between PD-L1 Status and Efficacy Measures

Analyses for ORR per BICR endpoint:

Evaluation of association between PD-L1+ tumor expression level and ORR in subjects treated with nivolumab/ipilimumab and nivolumab/ipilimumab-placebo

- A separate logistic regression model will be fitted for ORR with PD-L1+ expression level as a sole covariate for subjects treated with nivolumab/ipilimumab vs nivolumab/ipilimumab-placebo, respectively. An appropriate transformation of PD-L1+ expression may be considered depending on an assessment of fit of the model. The following summaries will be presented by treatment arm:
  - Odds ratios corresponding to a 1 and 10% PD-L1 expression change and its associated 95% CI, respectively
  - A plot of estimated response probability with 95% confidence band vs PD-L1+ expression (X-axis)
  - A plot of estimated response probability for each treatment arm in a single plot vs PD-L1+ expression (X-axis)
- Box plot of PD-L1 expression versus Response Status
- Receiver Operating Characteristics (ROC) analysis with ORR will be performed to help assess in-study (re-substitution) predictive accuracy of the logistic regression model and whether there is a clinically meaningful threshold of PD-L1 expression. The following summaries will be provided for treated subjects by treatment arm:
  - A plot of the ROC curve
  - A plot of estimated true positive fraction and false positive fraction vs. PD-L1+ expression (X-axis)

Analyses for PFS per BICR endpoint:

- A separate Cox proportional hazards regression model will be fitted for PFS with PD-L1+ expression level as a sole covariate among all subjects treated with nivolumab/ipilimumab and nivolumab/ipilimumab-placebo, respectively. An appropriate transformation of PD-L1+ expression may be considered depending on an assessment of fit of the model. The following summaries will be presented by treatment arm:
  - The hazard ratios corresponding to 1 and 10% PD-L1+ expression change along with its associated 95% CI, respectively
  - A plot of estimated $\log_e$(hazard) with 95% confidence band vs PD-L1+ expression(X-axis)
  - A plot of estimated $\log_e$(hazard) for each treatment arm in a single plot vs. PD-L1+ expression(X-axis)
- For each of the PD-L1 expression status subgroup:
  - PFS curves will be estimated using the Kaplan-Meier product limit method for each treatment group. Two-sided, 95% confidence intervals for median PFS will be computed by Brookmeyer and Crowley method.
  - Forest plot of Hazard Ratios with 95% CIs.

Analyses for OS endpoint:

- A separate Cox proportional hazards regression model will be fitted for OS with PD-L1+ expression level and among subjects treated with nivolumab/ipilimumab and nivolumab/ipilimumab-placebo, respectively. An appropriate transformation of PD-L1+ expression may be considered depending on an assessment of fit of the model. The following summaries will be presented by treatment arm:
  - Hazard ratios corresponding to a 1 and 10% PD-L1+ expression change and its associated 95% CI, respectively
  - A plot of estimated loge(hazard) with 95% confidence band vs PD-L1+ expression(X-axis)
  - A plot of estimated $\log_e$(hazard) for each treatment arm in a single plot vs. PD-L1+ expression(X-axis)
- For each of the PD-L1 expression status subgroup:
  - OS curves will be estimated using the Kaplan-Meier product limit method for each treatment group. Two-sided, 95% confidence intervals for median OS will be computed by Brookmeyer and Crowley method.
  - Forest plot of Hazard Ratios with 95% CIs.

7.9.2.3 Potential Predictive Relationship of PD-L1 Status for Efficacy Measures

Analyses will be performed using all evaluable PD-L1 subjects, if not otherwise specified.

Analyses for ORR per BICR:

A logistic regression model will be fitted for ORR with treatment, PD-L1 status, and treatment by PD-L1 status interaction. Although the study is not designed to have appropriate power to formally test the interaction of the model, an interaction test at significance level of 0.2 will warrant further exploration and the following statistics will be reported:

- Interaction p-value
- Odds ratio of nivolumab/ipilimumab vs nivolumab/ipilimumab-placebo and its associated 95% CI for each PD-L1 status subgroup
- Odds ratio of PD-L1 $\geq$ X% vs. $<$ X% and its associated 95% CI within each treatment group where X denotes the PD-L1 expression cut-off.

Analyses for PFS per BICR:

A Cox proportional hazards regression model will be fitted for PFS with treatment, PD-L1 status, and treatment by PD-L1 status interaction. Although the study is not designed to have appropriate power to formally test the interaction of the model, an interaction test at significance level of 0.2 will warrant further exploration and the following statistics will be reported:

- Interaction p-value
- Hazard ratio of nivolumab/ipilimumab vs nivolumab/ipilimumab-placebo and its associated 95% CI for each PD-L1 status subgroup
- Hazard ratio PD-L1 $\geq$ X% vs. $<$ X% and its associated 95% CI within each treatment group where X denotes the PD-L1 expression cut-off.

This analysis will be conducted using primary definition of PFS.

Analyses for OS endpoint:

A Cox proportional hazards regression model will be fitted for OS with treatment, PD-L1 status, and treatment by PD-L1 status interaction. Although the study is not designed to have appropriate

power to formally test the interaction of the model, an interaction test at significance level of 0.2 will warrant further exploration and the following statistics will be reported.

- Interaction p-value
- Hazard ratio of nivolumab/ipilimumab vs nivolumab/ipilimumab-placebo and its associated 95% CI for each PD-L1 status subgroup
- Hazard ratio PD-L1 ≥ X% vs. < X% and its associated 95% CI within each treatment group where X denotes the PD-L1 expression cut-off.

7.9.2.4 Association Between PD-L1 Status and select AE

Select AEs will be summarized by worst CTC Grade by treatment group for

- Each PD-L1 status subgroup by 1% cut-off (positive, negative)
- PD-L1 not evaluable or indeterminate subgroup

[redacted]

[redacted]

[redacted]

8 CONVENTIONS

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

For missing and partial adverse event onset dates, imputation will be performed using the Adverse Event Domain Requirements Specification[10]. Missing and partial Non-Study Medication Domain dates will be imputed using the derivation algorithm described in 4.3.3 of BMS Non-Study Medication Domain Requirements Specification[11].

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

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

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

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

*In cases where the date of death is present and complete, the imputed 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.

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

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

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

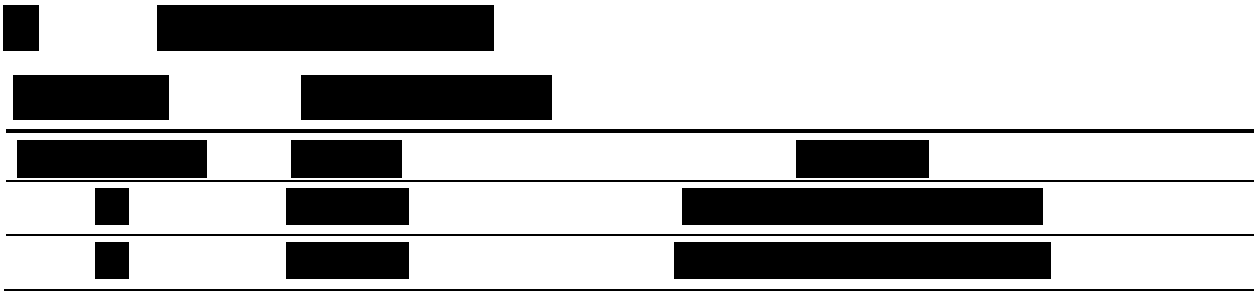
Duration (e.g. time from first diagnosis to first dosing date, duration of response, and time to response) will be calculated as follows:

Duration = (Last date - first date + 1)

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

9 CONTENT OF REPORTS

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



[redacted]