Official Title of Study:

An Open-label Phase 2 Multi-cohort Trial of Nivolumab in Advanced or Metastatic Malignancies

NCT Number: NCT02832167

Document Date (Date in which document was last revised): Jun 24, 2018

STATISTICAL ANALYSIS PLAN
FOR CLINICAL STUDY REPORT

AN OPEN-LABEL PHASE 2 MULTI-COHORT TRIAL OF NIVOLUMAB IN ADVANCED OR METASTATIC MALIGNANCIES

PROTOCOL CA209627

VERSION # 2.0

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

[redacted]

[redacted]

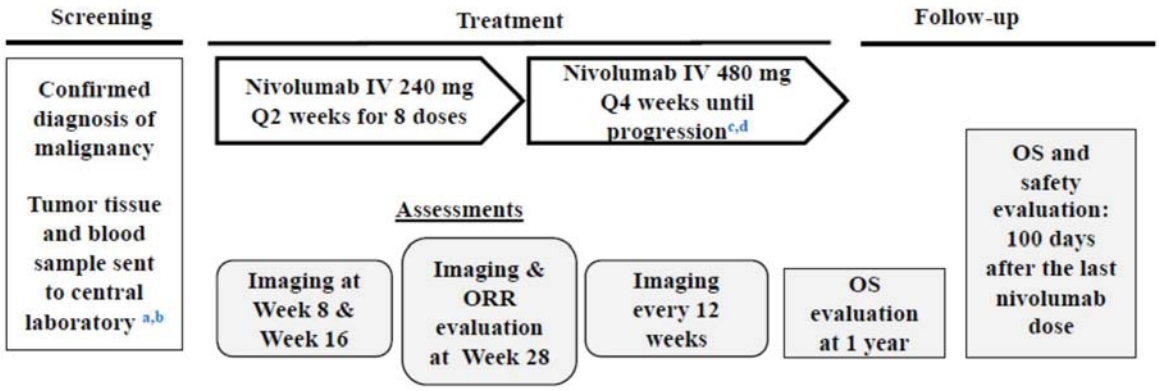
2 STUDY DESCRIPTION

2.1 Study Design

This is an open-label, multicenter, phase 2 study of nivolumab monotherapy in adult (≥ 18 years) subjects with advanced malignancies not previously evaluated with nivolumab or other I/O agent. Subjects must have received previous standard-of-care therapies (for primary therapy or post primary therapy).

The study design schematic is presented in Figure 1.

Figure 1: Study Design Schematic



a. Tumor tissue will be evaluated for sufficient tumor that meet the minimum quality requirements, received by a central laboratorybefore first treatment

b. Blood sample will be collected at screening or before first nivolumab administration

c. Subjects will be treated until confirmed progression or unacceptable toxicity, or 24 months of treatment

d. Time interval between dose 8 and dose 9 is 2 weeks; time interval between dose 9 and dose 10 is 4 weeks

Enrollment of up to 350 subjects is expected to require approximately 24 months. Interim monitoring and final analyses are planned. Subjects will receive study drug for the maximum treatment duration of 24 months.

The start of the trial is defined as first visit for first subject screened. End of trial is defined as the last scheduled procedure shown in the Time & Events Schedule for the last subject. Study completion is defined as the final date on which data for the primary endpoint was or is expected to be collected, if this is not the same.

2.2 Treatment Assignment

After informed consent has been obtained and the subject's eligibility is established, the subject will be enrolled and a number will be assigned through an interactive voice response system (IVRS). Every subject that signs the informed consent form must be assigned a subject number in IVRS. Once enrolled in IVRS, subjects that have met all eligibility criteria will be ready for treatment and drug vial assignment through the IVRS.

2.3 Blinding and Unblinding

Not applicable.

2.4 Protocol Amendments

This SAP incorporates the following protocol Amendment:

Amendment	Date of Issue	Summary of Major Changes
Revised Protocol 1	18-Jan-2018	• Changed primary end point from Clinical Benefit Rate at week 16 (CBR16) to Objective Response Rate (ORR). • Statistical analysis is modified to include a pause rule and clarification of final analysis of each bin. • Adverse Events were updated to include Immune-Mediated Adverse Events (IMAEs) • Vital signs were added to follow-up assessments • 28 days screening window removed • Adenocarcinoma of the small bowel and Adrenocortical carcinoma were added to inclusion criteria • Additional tumor types were added to the exclusion criteria • Treatment duration for 24 months
Revised Protocol 2	31-May-2018	• Incorporated change as per Administrative Letter 03 • Updated Medical Monitor information • Added minor changes to statistical section • Added language that Oncology TA was responsible for the design and conduct of the study and decisions regarding the protocol will be made with consultation with the Steering Committee • Added language for early stopping considerations

2.5 Data Monitoring and Other External Committees

There will be no independent Data Monitoring Committee (DMC) in this study. A Steering Committee comprised of trial investigators, trial statisticians, and BMS physicians will be created to design the trial and govern the trial's conduct, scope, and execution. The oncology therapeutic area of BMS has primary responsibility for overall design and conduct of the study. Decisions regarding the study protocol will be made by the sponsor upon consultation with the steering committee.

3 OBJECTIVES

3.1 Primary

To evaluate investigator-assessed Objective Response Rate (ORR) of nivolumab monotherapy in advanced or metastatic malignancies.

3.2 Secondary

- To assess duration of investigator-assessed clinical response (duration of response)
- To assess time to response (TTR)
- To assess clinical benefit rate (CBR)
- To assess overall survival (OS) at 1 year.
- To assess safety of nivolumab in malignancies in this trial.
- To correlate clinical response and OS to programmed death-ligand 1 (PD-L1) expression.
- To correlate clinical response and OS to MisMatch Repair (MMR) alterations.

[redacted]

4 ENDPOINTS

4.1 Primary Endpoint(s)

The primary endpoint of this trial is investigator-assessed ORR of nivolumab monotherapy. ORR is defined as the number of participants with a best overall response of confirmed CR or PR divided by the number of all treated participants. Best overall response is defined as the best response designation, as determined by investigator, recorded between the date of first dose and the date of objectively documented progression per tumor-specific response criteria or the date of subsequent therapy, whichever occurs first. For participants without documented progression or subsequent therapy, all available response designations will contribute to the BOR determination. For subjects who continue treatment beyond progression, the BOR will be determined based on response assessments up to the time of initial RECIST 1.1 progression.

Tumor assessments will occur at Week 8, Week 16, Week 28 and then every 12 weeks until disease progression. BOR for each patient will be derived as one of the following categories:

- Complete response (CR): At least two objective statuses of CR documented with a minimum of 4 weeks apart before progression.
- Partial response (PR): At least one objective status of PR at one visit and then at the subsequent (not need to be second one, any visits after the first 'PR') visit with a CR or PR, and two objective statuses documented with a minimum of 4 weeks apart before progression.
- Stable disease (SD): At least one objective status of stable documented at least 6 weeks after first dosing date and before progression but not qualifying as CR, PR, Or for unconfirmed CR or PR documented at least 6 weeks after first dosing date and before progression.
- Progressive Disease (PD): Objective status of progression and not qualifying as CR, PR or SD.
- Not evaluable (NE).

4.2 Secondary Endpoint(s)

4.2.1 Duration of Response

DOR is defined as the time from first confirmed response (CR or PR) to the date of the first documented tumor progression as determined by investigator or death due to any cause, whichever occurs first. Participants who start subsequent therapy without a prior reported progression will be censored at the last evaluable tumor assessments prior to initiation of the subsequent anticancer therapy. Participants who die without a reported prior progression will be considered to have progressed on the date of their death. Participants who neither progress nor die will be censored on the date of their last evaluable tumor assessment.

4.2.2 TTR

TTR is defined as the time from first dosing date to the date of the first confirmed response, as assessed by investigator.

4.2.3 CBR

CBR is defined as the number of participants with a best overall response of confirmed CR or PR, or stable disease divided by the number of all treated subjects.

4.2.4 Overall Survival

OS is defined as time from the date of first dosing to the date of death. Subjects who did not die by the end of the study will be censored at the last known date alive. OS rate at 1 year is measured as the survival rate at 1 year from Kaplan-Meier curve of OS.

4.2.5 Safety and Tolerability

Safety and tolerability will be measured by the incidence of deaths adverse events (AEs) , serious adverse events (SAEs), adverse events leading to discontinuation, adverse events leading to dose modification, select adverse events (select AEs) leading to dose delay, and specific laboratory abnormalities (worst grade) in each tumor group. Toxicities will be graded using the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) version 4.0.

[redacted]

5 SAMPLE SIZE AND POWER

The maximum overall sample size across all tumor groups is 350 subjects. In addition, the maximum permitted sample size per group is 25 subjects, although a group with high posterior probability of success may be permitted a larger sample size at Sponsor discretion. It is expected that accrual may vary across groups, with some groups enrolling a higher number of subjects than other groups within the trial. Simulation studies were conducted to evaluate the performance of the analysis under various assumptions for the distribution of true underlying ORR across the tumor types. Operating characteristics including power and type I error were assessed.

When the treatment effects are similar across all tumor types, estimation efficiencies, due to borrowing, result in strong trial performance. When all groups are in truth effective, the individual groups generally exhibit power between 86.8% and 98.4%, under the scenario of odds ratio improvement of 3. When all groups are ineffective, the individual groups have between a 2.1% and 9.5% chance of mistakenly declaring group success. The group specific type I errors are the metric of interest when evaluating the false positive rate for this type of design. Full presentation and discussion of the simulation parameters, corresponding simulation results, and example trials are included in the Statistical Analysis and Modeling Report provided in the Appendices.

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 (e.g., laboratory tests and vital signs) will be defined as evaluations with a date on or prior to the day of first dose of study treatment

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

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

6.2 Treatment Regimens

Treatment period, Phase 1: Subjects will receive treatment with nivolumab 240 mg as a 30 minute infusion on Day 1 of a treatment cycle every 2 weeks (14 days) for 8 doses until progression, unacceptable toxicity, withdrawal of consent, the study ends, or until Phase 2 dosing begins, whichever occurs first.

Treatment period, Phase 2: Beginning with Dose 9, subjects will receive nivolumab 480 mg as a 30 minute infusion every 4 weeks (± 3 days) until progression, unacceptable toxicity, withdrawal of consent, or the study ends, whichever occurs first. Subjects will receive study drug for the maximum treatment duration of 24 months.

6.3 Populations for Analyses

The following subject populations will be considered in this trial:

- **All enrolled subjects:** All subjects who signed an informed consent form and were registered into the IVRS.
- **All treated subjects:** All subjects who received at least 1 dose of nivolumab.
- **Trial update analysis subjects:** All subjects who received at least 1 dose of nivolumab and 28 weeks have elapsed after the date of 1st dose
- **PK subjects:** All treated subjects with available serum time-concentration data.

- **Biomarker subjects:** All treated subjects with available biomarker data.

7 STATISTICAL ANALYSES

7.1 General Methods

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

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

7.2 Study Conduct

Unless otherwise specified, the study conduct data will be presented on all treated subjects by tumor group.

7.2.1 Accrual

The following will be presented on the enrolled population

- Number of subjects accrued by country and investigational site
- Number of subjects accrued by month.

A by-subject listing of accrual by tumor type will be produced.

7.2.2 Relevant Protocol Deviations

The following programmable deviations from inclusion and exclusion criteria will be considered as relevant protocol deviations and be summarized on treated subjects. 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.

At Entrance:

- Subjects with baseline ECOG performance status > 1
- Subjects without measurable disease at baseline by CT or MRI per RECIST 1.1 criteria
- Subjects with screening laboratory values not meeting criteria listed in Target Population of Section 3.3.1 Inclusion Criteria of the study protocol
- Life expectancy < 16 weeks

- Subjects with target disease exclusions (listed in Target Disease Exclusions of Section 3.3.2 Exclusion Criteria of the study protocol)

On-study:

- Subjects receiving prohibited anti-cancer therapy (systemic therapy, radiotherapy, surgery) while on study therapy.
- Assessments not performed per protocol
- Study drug not administered as per protocol

A by subject listing of relevant protocol deviation will also be produced.

7.3 Study Population

Unless otherwise specified, the study population data will be presented on all treated subjects by tumor group.

7.3.1 Subject Disposition

Summary tables reflecting the number of subjects who are enrolled, who are treated and who completed the study will be presented by tumor type. The number of subjects who do not complete treatment phase part 1, treatment phase part 2 and follow-up, both overall and according to reasons for non-completion, will be summarized by tumor type.

A by subject listing of disposition will be provided.

7.3.2 Demographics and Baseline Characteristics

Descriptive statistics of the following baseline characteristics will be summarized by tumor group.

- Age
- Gender
- Race
- Ethnicity
- Smoking history: Never/Current/Former/Unknown
- Current Disease Diagnosis
  - Time from diagnosis to study treatment first dose
  - Stage: I/II/III/IV-Non-Metastatic/IV-Metastatic

7.3.3 Physical Measurements

Descriptive statistics of the following physical measurements will be summarized by tumor group.

- Baseline weight (kg)
- Body weight height (cm)
- Body mass index (BMI) (kg/m^2)

- Baseline Eastern Cooperative Oncology Group (ECOG) performance status

Measurements will also be summarized by tumor group and study visit.

7.3.4 Medical History

General medical history will be coded and grouped into Preferred Terms (PT) by System Organ Class (SOC), using Medical Dictionary for Regulatory Activities (MedDRA) version 18.1 or higher. Listings and summaries will be based on the SOCs and PTs. Summaries will be provided by tumor group.

A by subject listing of disease specific medical history will be provided.

7.3.5 Prior Therapy

Descriptive statistics of prior therapies will be summarized for all treated subjects by tumor group.

- Prior Surgery Related to Cancer
  - Subjects with surgery (yes/no)
  - Time from surgery to study treatment first dose
- Prior Radiotherapy
  - Subjects with radiotherapy (yes/no)
  - Time from therapy stop to study treatment first dose
- Prior Systemic Cancer Therapy
  - Subjects with systemic cancer therapy (yes/no)
  - Setting of regimen (adjuvant therapy/metastatic disease/neo-adjuvant therapy)
  - Best response to regimen (CR/PR/SD/PD/unable to determine/not applicable)
  - Time from therapy stop to study treatment first dose

A by subject listing of the above prior therapies will be provided.

7.3.6 Baseline Examinations

Subjects with abnormal baseline physical examination will be tabulated by examination criteria and also by tumor type.

7.4 Extent of Exposure

Unless otherwise specified, the exposure data will be presented on treated subjects by tumor group.

7.4.1 Administration of Study Therapy

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

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

- Cumulative dose
- Duration of treatment

A by-subject listing of extent of exposure: treatment phase, number of doses, date of first and last dose, cumulative dose, relative dose intensity, duration of treatment, and reason for discontinuation.

Table 7.4-1: Administration of Study Therapy: Definition of Parameters

	nivolumab
Dosing schedule per protocol	**Phase 1: 240 mg Q2 weeks for 8 doses. Phase 2: 480 mg Q4 weeks until progression**
Cumulative Dose	*Cum dose (mg) is sum of the doses administered to a subject during the treatment period*
Relative dose intensity (%)	*Phase 1: [Cum dose (mg)/((Last dose date - Start dose date + 14) 240 / 14)] x 100* *Phase 2: [Cum dose (mg)/((Last dose date - Start dose date + 28) x 480 / 28)] x 100*
Duration of treatment	*Last dose date - Start dose date +1*

7.4.2 Modifications of Study Therapy

7.4.2.1 Dose Delays

Subjects should begin study treatment within 3 calendar days of treatment assignment. If required for clinic scheduling, subjects may be treated up to 3 days before or after the scheduled date, ie, at intervals of not less than 12 days from the previous dose during Q2W cycles. For Q4W dosing cycles, participants may be dosed within a ±3 day window. A dose given more than 3 days after the intended dose date will be considered a delay. Subsequent treatments should be based on the actual date of administration of the previous dose of drug. A dose will be considered as actually delayed if the delay is exceeding 3 days after the intended dose date (i.e., greater than or equal to 4 days from scheduled dosing date). Length of dose delay is defined as (latter dose date – previous actual dose date - 14) for phase 1 and (latter dose date – previous actual dose date - 28) for phase 2. Dose delay may happen more than once. The phase 2 is beginning with dose 9 (2 weeks following the 8th dose). Dose delays will be divided into following categories: 4 - < 8 days, 8 - < 15 days, 15 - < 43, > 42 days. Reason for dose delay will be retrieved from CRF dosing pages.

The following parameters will be summarized:

- Number of dose delayed per subject, Length of Delay, and Reason for Dose Delay
- Number of subjects with at least one dose delayed along with reason for dose delay

7.4.2.2 Dose Modifications

There will be no dose escalations or reductions of nivolumab allowed.

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

The following will be summarized:

- Number of subjects with at least one infusion with IV rate reduced along with the reason of the rate reduction
- Number of subjects with at least one dose infusion interrupted along with the reason for the interruptions and number of infusions interrupted per subject

A by subject listing of study drug administered will be provided. A batch listing number will be also provided.

[redacted]

7.5 Efficacy

7.5.1 Overall Efficacy Analysis

Unless otherwise specified, all efficacy analyses will be performed for all treated subjects.

7.5.2 Primary Analysis

7.5.2.1 Statistical Hypotheses

Let Yi be the response indicator for the i^{th} subject. Define $\pi g = Pr(Yi = 1 \mid gi = g)$ as the underlying probability of response for group g for the experimental treatment and Rg as the assumed probability of response for group g within the (historical) control population. Transformation to the logit scale is applied for modeling purposes. Let θg be the mean log odds treatment effect, i.e.

$$\theta_g = log\left(\pi_g/(1-\pi_g)\right) - log\left(R_g/(1-R_g)\right).$$

Thus, θg is the logistic regression coefficient for the treatment within tumor group g. The primary analysis is a set of group specific tests of $\theta g > 0$ (equivalently, $\pi g > Rg$) meaning that the treatment is better than the assumed control rate within that group. Thus, the following hypotheses are tested for each tumor group g:

$$H0g : \theta g \leq 0,$$

$$H1g : \theta g > 0.$$

The assumed control ORRs (R_g) vary by group and for each tumor group are provided in Table 7.5-1.

Table 7.5-1: Historical Control ORRs

Group Index g	Tumor Type	Assumed Control ORR (R_g)
1	Anal Cancer (squamous cell histology)	15%
2	Biliary Tract Cancer (includes nonresectable disease), intrahepatic or extrahepatic cholangiocarcinoma, gallbladder cancer, and ampullary carcinoma	10%
3	Carcinoid after SSAs (Ki67 less than 20%)	7.5%
4	Squamous Cell Cancer of the Cervix (exocervix), Squamous Cell Cancer of the Vagina	20%
5	Endometrial Cancer (after primary treatment including RT). Subjects with tumors > 10% estrogen receptor positive pathology in the primary tumor are excluded.	20%
6	Non-squamous Cell Cancer of the Head and Neck (including Cancer of the Salivary Gland, Adenoid Cystic Carcinoma)	10%
7	Histiocytoses (including Erdheim Chester Disease [macrophage disorder], Langerhans Cell Histiocytosis [LCH, dendritic cell disorder])	15%
8	Lynch Syndrome Associated Cancers (excluding HNPCC)	12.5%
9	Medullary Thyroid Cancer (after TKI [vandetanib or cabozantinib])	20%
10	Merkel Cell Carcinoma (includes unresectable disease)	15%
11	Mesothelioma	10%
12	Nasopharyngeal Carcinoma	17.5%
13	Neuroendocrine Tumors (poorly differentiated, Ki67 > 20%)	10%
14	Neuroendocrine Tumors (well to moderately differentiated) after SSAs or everolimus, including insulinomas	15%
15	Non-Lung Small Cell Carcinoma (includes Small Cell Carcinoma of the ovary of pulmonary or hypercalcemic type)	10%
16	Penile Cancer	15%
17	Rare Women's Cancers: Clear Cell (> 50% clear cell by pathology)	5%

Group Index g	Tumor Type	Assumed Control ORR (R_g)
18	Soft-Tissue Sarcoma: Liposarcoma, Leiomyosarcoma, Malignant Peripheral Nerve Sheath Tumor, NF-1	10%
19	Testicular Cancer (chemotherapy resistant disease or relapsed within 2 years of primary therapy)	10%
20	Thymic Carcinoma or Invasive Thymoma	10%
21	Thyroid Cancer (papillary or follicular), after failing RAI and approved kinase inhibition [lenvatinib]	15%
22	Thyroid Cancer: anaplastic first line. In BRAF V600e positive subjects, investigators may administer nivolumab after vemurafenib.	5%
23	Uterine Sarcoma (excluding endometrial stromal sarcoma)	15%
24	Vulvar Cancer (post vulvectomy, cisplatin and radiotherapy)	15%
25	Adenoid Cystic Carcinoma	10%

7.5.2.2 Final Efficacy Analyses

The primary endpoint of this trial is ORR. ORR is defined as the number of participants with a best overall response of confirmed CR or PR divided by the number of all treated participants. Best overall response is defined as the best response designation, as determined by investigator, recorded between the date of first dose and the date of objectively documented progression per tumor-specific response criteria or the date of subsequent therapy, whichever occurs first. For participants without documented progression or subsequent therapy, all available response designations will contribute to the BOR determination. For subjects who continue treatment beyond progression, the BOR will be determined based on response assessments up to the time of initial RECIST 1.1 progression. Tumor assessments will occur at Week 8, Week 16, week28 and then every 12 weeks until disease progression. Once a group is stopped for futility, the group will not be re-opened for further enrollment.

A final analysis for the study will occur at least 28 weeks after the last subject of the study's first treatment. Separate final analyses will be conducted for each cohort. The final analysis for each group declares success if the Bayesian posterior probability that the treatment effect in that group (treatment compared to historical control) is above 80%. Formally:

$$Pr(\pi_g > R_g) > 0.80$$

7.5.2.3 Trial Update Efficacy Analyses

Trial update analyses will be conducted at most every 4 weeks, and includes accruing ORR data information for all subjects into the modeling framework. At each analysis, the groups will be evaluated for early stopping for futility or success by comparing posterior quantities for the ORR to pre-specified early stopping criteria based on the hierarchical model. The design will limit the maximum number of subjects per group to 25 subjects. Available resources will be used more effectively by enrolling subjects in groups that offer greater promise of efficacy. The trial update analysis subject population will be used for efficacy analyses using the hierarchical model. The all treated subjects population will be used for supporting efficacy displays.

Early Futility: If there a sufficiently low probability (20%) that the ORR in a group exceeds the historical rate R_g by at least 10%, then the group will stop enrollment early for futility. Formally, enrollment will stop early for futility if:

$$Pr(\pi g > Rg+0.10) < 0.20$$

A group is only eligible for early stopping for futility once a minimum of 5 subjects have had the opportunity to complete 28 weeks of follow-up for ORR, i.e. up to 28 weeks after the date of first dose, in that group.

Early Success: If there a sufficiently high probability (95%) that the ORR in a group exceeds the historical rate Rg, then the group may stop enrollment early for success. Formally, enrollment may stop early for success if:

$$Pr(\pi g > Rg) > 0.95$$

A group is only eligible for early stopping for success once a minimum of 8 subjects have had the opportunity to complete 28 weeks of follow-up for ORR, i.e. up to 28 weeks after the date of first dose, in that group.

A summary of the hierarchical model analysis will be provided. For each group the following measures will be reported; the total number of evaluable subjects, the protocol-specified historical rate (Rg), the observed ORR, the estimated mean ORR and 95% credible interval (based on the posterior 2.5% and 97.5% percentiles), the posterior probability that ORR exceeds its corresponding Rg, and the posterior probability that ORR exceeds its corresponding Rg by at least 10%. Enrollment decisions and final analysis decisions based on the protocol-specified decision rules will be provided.

A summary of BOR (CR, PR, SD, progression) will be provided for each tumor group in order to support the model-based analyses for ORR.

7.5.2.4 Pause Rule

The Bayesian hierarchical model incorporates information from all cohorts into each analysis, however the primary driver of the analysis for each cohort is the data from that specific cohort. Thus, it is important that enough data is available in each cohort where possible. To minimize the possibility of excessive enrollment imbalance between the groups, enrollment may be paused in cohort at any time where 12 or more patients have enrolled, but have not had, yet, an opportunity to complete their 28 week evaluation visit (dropouts are not included in this calculation if they have no post-baseline tumor assessment). A pause will be lifted when a trial update occurs and the following conditions apply:

- At most 3 subjects have not had an opportunity to complete 28 weeks of follow-up or have discontinued due to disease progression.
- Neither early success or early futility criteria has been met for a group.

In addition, based on considerations such as emerging external data, cohorts may be paused by Sponsor decision in consultation with Steering Committee.

7.5.3 Secondary Efficacy Analyses

7.5.3.1 Duration of Response

DOR is defined as the time from first documented response (CR or PR) to the date of the first documented tumor progression as determined by investigator or death due to any cause, whichever occurs first. Participants who start subsequent therapy without a prior reported progression will be censored at the last evaluable tumor assessments prior to initiation of the subsequent anticancer therapy. Participants who die without a reported prior progression will be considered to have progressed on the date of their death. Participants who neither progress nor die will be censored on the date of their last evaluable tumor assessment. Duration of response only applies to subjects, whose best overall response is CR or PR according to RECIST 1.1. It is defined as the time from first confirmed CR or PR until the first documented sign of disease progression (according to RECIST 1.1 criteria) or death. The following censoring rules will be applied:

Table 7.5-2: Censoring Rules for the Duration of Response Analysis

Situation	Date of Progression or Censoring	Outcome
No baseline tumor assessments and no death	First dosing date	Censored
No on study tumor assessments and no death	First dosing date	Censored
New anticancer treatment started without a prior reported progression per RECIST 1.1 or death	Date of last evaluable tumor assessment prior to or on the date of initiation of the subsequent anti-cancer therapy	Censored
Progression per RECIST 1.1 documented at scheduled or unscheduled visit and no new anticancer treatment started before	Date of the first documented tumor progression	Progressed
Subject progression free (per RECIST 1.1) and no new anticancer treatment started	Date of last tumor assessment	Censored

Death without prior progression per RECIST 1.1 and no new anticancer treatment started	Date of death	Progressed/death

The duration of response will be analyzed using Kaplan-Meier methods. Rates for Probability of Duration of Response lasting beyond 6, 9, 12 months will be computed from the Kaplan Meier estimate and corresponding 95% confidence interval will be derived based on Greenwood formula for variance derivation and on log-log transformation applied on the survivor function.

7.5.3.2 Time to Response

Time to response (TTR) is defined as the time from first dosing date to the date of the first confirmed response, as assessed by investigator. Summary statistics of TTR will be provided for participants who achieve PR or CR.

7.5.3.3 Clinical Benefit Rate

Clinical benefit rate (CBR) is defined as the number of participants with a best overall response of confirmed CR or PR, or stable disease divided by the number of all treated subjects. CBR will be provided along with the corresponding 2-sided 95% confidence intervals using the method based on the Binomial distribution.

7.5.3.4 Overall Survival

The OS rate at 1 year will be calculated by tumor group. OS is defined as time from the date of enrollment to the date of death. Subjects who did not die by the end of the study will be censored at the last known date alive.

Time to event distribution will be estimated using Kaplan Meier techniques. This will be done for OS. Median OS along with 95% CI will be constructed based on a log-log transformed CI for the survivor function. Rates at 1-year will be derived from the Kaplan Meier estimate and corresponding confidence interval will be derived based on Greenwood formula for variance derivation and on log-log transformation applied on the survivor function.

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

- On-study: (on-treatment and not progressed, on-treatment progressed, in follow-up)
- Off-study: (lost to follow-up, withdraw consent, etc.)

7.6 Safety

Unless otherwise specified, the safety data will be presented by tumor group.

Whenever appropriate, analysis of all safety data will be performed by following the BMS nivolumab Core Safety SAP (version 4.0)[1] and the comprehensive specifications for the

development of data displays for tables, listings and graphs should follow the BMS nivolumab Core Safety Data Presentation Plan (DPP)[2].

7.6.1 Deaths

Deaths will be summarized by tumor group:

- All deaths, reasons for death

Deaths within 30 days and 100 days of last dose received, reasons for deathBy-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 tumor group:

- Overall summary of SAEs by worst CTC grade (grade 1, 2, 3, 4, 5, unknown) presented by SOC/PT using the 30-day and repeated using the 100-daysafety window
- Overall summary of SAEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT using the 30-day and repeated using the 100-day safety window and for AEs with extended follow-up. This table will be restricted to events with incidence greater or equal to 1% in any tumor group.
- Overall summary of drug-related SAEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT using the 30-day and repeated using the 100-day safety window and AEs with extended follow-up

By-subject SAE listing will be provided for the All Enrolled Subjects population.

7.6.3 Adverse Events Leading to Discontinuation of Study Therapy

AEs leading to discontinuation will be summarized by tumor group:

- Overall summary of AEs leading to discontinuation by worst CTC grade (grade 1, 2, 3, 4, 5, unknown) presented by SOC/PT
- 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 (grade 1, 2, 3, 4, 5, unknown) 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 analysis will be conducted using the 30-day and repeated using the 100-day safety window.

By-subject AEs leading to discontinuation listing will be provided.

7.6.4 Adverse Events Leading to Dose Modification

AEs leading to dose delay/interruptions will be summarized for each tumor group:

- Overall summary of AEs leading to dose delay/reduced by worst CTC grade (grade 1, 2, 3, 4, 5, unknown) presented by SOC/PT

The analysis will be conducted using the 30-day and repeated using the 100-day safety window.

By-subject AEs leading to dose delay/reduced listing will be provided.

7.6.5 Adverse Events

Adverse events will be summarized by tumor group:

- Overall summary of any AEs by worst CTC grade (grade 1, 2, 3, 4, 5, unknown) presented by SOC/PT using the 30-day and repeated using the 100-day safety window
- Overall summary of any AEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT using the 30-day and repeated using the 100-day safety window and for AEs with extended follow-up. This table will be restricted to events with an incidence greater or equal to 5% in any tumor group.
- Overall summary of drug-related AEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT using the 30-day and repeated using the 100-day safety window and for AEs with extended follow-up
- Overall summary of any non-serious AEs presented by SOC/PT using the 30-day and repeated using the 100-day safety window. This table will be restricted to events with an incidence greater or equal to 5% in any tumor group.Summary of late emergent drug-related AEs by worst CTC grade presented by SOC/PT for AEs reported beyond 100 days after last dose of study therapy

By-subject AE listing will be provided.

7.6.6 Select Adverse Events

Unless otherwise specified, analyses will be performed by select AE category.

7.6.6.1 Incidence of Select AE

Select AEs (see Table 11-3 of CA209 Core Safety SAP) will be summarized by tumor group for each category/subcategory:

- Overall summary of any select AEs by worst CTC grade presented by Category or Subcategory/PT (any grade, grade 3-4, grade 5)
- Overall summary of drug-related select AEs by worst CTC grade presented by Category or Subcategory / PT (any grade, grade 3-4, grade 5)
- Overall summary of any serious select AEs by worst CTC grade presented by Category or Subcategory /PT (any grade, grade 3-4, grade 5)

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

The analyses will be conducted using the 30-day and repeated using the 100-day safety window and/or extended follow-up. Refer to CA209 Core Safety DPP for details,

By-subject select AE listing will be provided.

7.6.6.2 Time to Onset/Resolution of Select AE

Time-to onset, time-to resolution and time-to resolution where immune modulating medication was initiated will be analyzed for each specific category/subcategory:

- Summary of time to onset of drug-related select AEs by category for treated subjects who experienced at least one drug- related select adverse event from the category
- Summary of time to resolution of drug-related select AEs by category for treated subjects who experienced at least one drug- related select adverse event from the category
- Summary of time to resolution of drug-related select AEs where immune modulating medication was initiated by category for treated subjects who experienced at least one drug-related select adverse event from the category

7.6.7 Immune Modulating Medication

The percentage of subjects who received immune modulating concomitant medication for

- management of adverse event
- premedication
- other use

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

For each category/subcategory of select AEs (see Table 11-3 of CA209 Core Safety SAP), the following will be reported for each tumor group:

- Percentage of subjects who received immune modulating concomitant medication for management of any select AE in the category among subjects who experienced at least one select adverse event in the category/subcategory.
- The total medication treatment duration, duration of high dose of corticosteroid and tapering duration (summary statistics)

These analyses will be performed on any select AEs, drug-related select AEs, grade 3-5 select AEs and drug-related select AEs.

Total duration of all immune modulating medications (excluding overlaps) given for select AEs management will be reported.

The analysis will be conducted using the 30-day and repeated using the 100-day safety window, if applicable. Refer to CA209 Core Safety DPP for details,

7.6.8 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 the subjects treated 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 tumor group. This analysis will be limited to the rate of all AEs and all drug-related AEs. Refer to CA209 Core Safety DPP for details,

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 tumor groups.

7.6.9 Clinical Laboratory Evaluations

Clinical laboratory parameters (hematology, serum chemistry and electrolytes) will be evaluated. Laboratory tests will be graded using the NCI Common Terminology Criteria, version 4.0. Clinical laboratory data will be first analyzed using International System of Units (SI). Analyses will be repeated using US conventional units..

7.6.9.1 Hematology

The following will be summarized by tumor 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 300-day safety window and repeated using the 100-day safety window.

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

7.6.9.2 Serum Chemistry

The following will be summarized by tumor 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 300-day safety window and repeated using the 100-day safety window.

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

7.6.9.3 Electrolytes

The following will be summarized by tumor 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).

The analyses will be conducted using the 30-day safety window and repeated using the 100-day safety window.

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

7.6.9.4 Abnormal Thyroid Function Test

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

- TSH value > ULN and
  - with baseline TSH value ≤ ULN
  - at least one FT3/FT4 test value < LLN
- TSH < LLN and
  - with baseline TSH value ≥ LLN
  - at least one FT3/FT4 test value > ULN

The analyses will be conducted using the 30-day safety window and repeated using the 100-day safety window.

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

7.6.9.5 Abnormal Hepatic Function Test

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

- ALT or AST > 3 x ULN, > 5 x ULN, > 10 x ULN and > 20 x ULN
- Total bilirubin > 2 x ULN

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

The following scatterplots will be produced for the following hepatic laboratory parameters:

- Total bilirubin peak vs. AST peak
- Total bilirubin peak vs. ALT peak

On-treatment peak total bilirubin and on-treatment peak AST/ALT may or may not happen on the same day of liver testing.

The analyses will be conducted using the 30-day safety window and repeated using the 100-day safety window.

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

7.6.10 Vital Signs

Vital signs parameters (systolic blood pressure [mmHg], diastolic blood pressure [mmHg], heart rate [bpm], respiratory rate [breaths/min] and body temperature [°C]) will be listed. Summaries of vital signs parameters and respective changes from baseline will be provided for each vital signs parameter by tumor group and study visit.

Subjects with vital signs outside of a pre-specified range will also be listed.

7.6.11 ECOG Performance Status

Summaries of ECOG performance will be provided by tumor group and study visit. A by-subject listing of ECOG will be provided.

Classification of ECOG performance status is provided in Table 7.6-12.

Table 7.6-12: ECOG PERFORMANCE STATUS

ECOG PERFORMANCE STATUS	
Grade	ECOG
0	Fully active, able to carry on all pre-disease performance without restriction
1	Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g., light house work, office work
2	Ambulatory and capable of all self-care but unable to carry out any work activities. Up and about more than 50% of waking hours
3	Capable of only limited self-care, confined to bed or chair more than 50% of waking hours
4	Completely disabled. Cannot carry on any self-care. Totally confined to bed or chair

Table 7.6-12: ECOG PERFORMANCE STATUS

ECOG PERFORMANCE STATUS	
5	Dead

7.6.12 Pregnancy Testing

By-subject listing of pregnancy tests results will be provided for all treated female subjects.

[redacted]

8 CONVENTIONS

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

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

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

For date of 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 missing and partial adverse event onset dates, imputation will be performed using the Adverse Event Domain Requirements Specification[5].

For missing and partial adverse event resolution dates, imputation will be performed as follows (these conventions may change):

- 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 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 4.3.3 of BMS Non-Study Medication Domain Requirements Specification[6].

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)

8.1 Missing data in Efficacy endpoints

If measurements for one or more target lesions are missing for an evaluation and disease does not qualify as progression, the objective status for that evaluation is not evaluable.

For method used for 'target lesion', if the post-baseline method changed for same lesion for example, the baseline tumor measurement method is 'MRI', but post-baseline tumor measurement method changed to 'CT', then for that evaluation is not evaluable if not progressed.

If a target lesion (longest diameter for a non-nodal lesion, shortest axis for a nodal lesion) is documented as too small to measure without unequivocal complete disappearance of the lesion, a default value of 5mm will be assigned and which objective status will be assigned accordingly.

All statistical analyses will be carried out using SAS (Statistical Analysis System software, SAS Institute, North Carolina, USA) unless otherwise noted. The primary efficacy endpoint analyses and trial update analyses will be performed in custom code created in C++.

9 CONTENT OF REPORTS

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

10 DOCUMENT HISTORY

Table 10-1: Document History

Version	Author	Summary of Changes
1.0	[redacted]	Original version
2.0		

[redacted]