STATISTICAL ANALYSIS PLAN

A RANDOMIZED PHASE 3 OPEN LABEL STUDY OF NIVOLUMAB VERSUS BEVACIZUMAB AND A SAFETY STUDY OF NIVOLUMAB OR NIVOLUMAB IN COMBINATION WITH IPILIMUMAB IN ADULT SUBJECTS WITH RECURRENT GLIOBLASTOMA (GBM)

PROTOCOL CA209143

[redacted]

TABLE OF CONTENTS

STATISTICAL ANALYSIS PLAN 1
TABLE OF CONTENTS 2
LIST OF TABLES 6
LIST OF FIGURES 7
1 BACKGROUND AND RATIONALE 8
2 STUDY DESCRIPTION 9
2.1 Study Design 9
2.2 Treatment Assignment 16
2.3 Blinding and Unblinding 17
2.4 Protocol Amendments 17
2.5 Data Monitoring Committee 19
3 OBJECTIVES 19
3.1 Primary 19
3.2 Secondary 20
3.3 Exploratory 20
4 ENDPOINTS 21
4.1 Primary 21
4.1.1 Cohorts 1, 1b, [redacted] *21*
4.1.2 Cohort 2 ..... *21*
4.2 Secondary 21
4.2.1 Cohort 2 ..... *21*
4.2.1.1 Overall Survival Rate at 12 Months (OS(12)) *21*
4.2.1.2 Progression Free Survival *22*
4.2.1.3 Objective Response Rate(ORR) *24*
[redacted]
4.2.2.1 Overall survival *25*
4.3 Exploratory 25
4.3.1 OS for Cohort 1 and 1b *25*
4.3.2 PFS for Cohort 1, 1b, [redacted] *25*
4.3.3 Duration of Response (DOR) *25*
4.3.4 Time to Response(TTR) *25*
4.3.5 Magnitude of Tumor burden *25*

[redacted]
4.3.8 PD-L1 Expression 25
[redacted]
5 SAMPLE SIZE AND POWER 30
6 STUDY PERIODS, TREATMENT REGIMENS AND POPULATIONS FOR ANALYSES 33
6.1 Study Periods 33
6.1.1 Baseline Period 33
6.1.2 Post Baseline Period 33
6.2 Treatment Regimens 34
6.3 Populations for Analyses 34
7 STATISTICAL ANALYSES 35
7.1 General Methods 35
7.2 Study Conduct 36
7.2.1 Accrual 36
7.2.2 Relevant Protocol Deviations 36
7.3 Study Population 37
7.3.1 Subject Disposition 37
7.3.2 Demographics and Baseline Characteristics 38
7.3.3 Medical History 38
7.3.4 Prior Therapy 38
7.3.5 Baseline Examinations 39
7.3.6 Discrepancies Between IVRS and CRF Stratification Factors (Cohort 2 only) 39
7.4 Extent of Exposure 39
7.4.1 Administration of Study Therapy 39
7.4.2 Modifications of Study Therapy 41

7.4.2.1 Dose delays *41*
7.4.2.2 Infusion Interruptions and Rate Changes *41*
7.4.2.3 Radiotherapy and Temozolomide *42*
7.4.2.4 Dose Reductions/Escalation *42*
7.4.3 Concomitant Medications *42*
7.5 Efficacy 42
7.5.1 Overall Survival for Cohort 2 ..... *42*
7.5.1.1 Primary Analysis *42*
7.5.1.2 OS sensitivity Analyses *43*
7.5.1.3 Consistency of Treatment Effect on OS in Subsets *44*
7.5.1.4 Subject Follow-Up *44*
7.5.1.5 Subsequent Therapy *44*
7.5.2 Interim Analysis for Cohort 2 ..... *45*
7.5.3 Overall Survival Rate at 12 Month for Cohort 2 ..... *45*
7.5.3.1 Primary Analysis *45*
7.5.4 Progression Free Survival for Cohort 2 ..... *45*
7.5.4.1 Primary Analysis *45*
7.5.4.2 Sensitivity Analysis for Cohort 2 ..... *46*
7.5.5 Objective Response Rate for Cohort 2 ..... *46*
7.5.5.1 Primary Analysis *47*
7.5.5.2 Further Characterization of ORR *47*
7.5.6 Other Efficacy Analyses for Cohort 2 ..... *47*
7.5.7 Efficacy Analyses for Cohorts 1, 1b, [redacted] *48*
7.6 Safety 48
7.6.1 Deaths *48*
7.6.2 Serious Adverse Events *48*
7.6.3 Adverse Events Leading to Discontinuation of Study Therapy *48*
7.6.4 Adverse Events Leading to Dose Delay of Study Therapy *48*
7.6.5 Adverse Events *48*
7.6.6 Select Adverse Events *48*
7.6.7 Immune-Mediated Adverse Events *49*
7.6.7.1 Incidence of IMAEs *50*
7.6.7.2 Management of IMAEs *51*

7.6.7.3 Time to Onset and Resolution of IMAEs *51*
7.6.7.4 Re-Challenge *51*
7.6.8 Immune Modulating Medication *51*
7.6.9 Multiple Events *52*
7.6.10 Clinical Laboratory Evaluations *52*
7.6.10.1 Hematology *52*
7.6.10.2 Serum Chemistry *52*
[redacted]
7.6.12 Vital Signs and Pulse Oximetry *52*
7.7 Pregnancy *52*
[redacted]
7.9 Biomarkers *52*
7.9.1 PD-L1 Expression *52*
7.9.2 Analysis Methods *53*
7.9.2.1 Distribution of PD-L1 Expression *53*
7.9.2.2 Association Between PD-L1 Status and Efficacy Measures *54*
7.9.2.3 Potential Predictive Relationship of PD-L1 Status for Efficacy Measures *55*
7.9.2.4 Association of Select AEs and PD-L1 Expression *55*
7.10 Neurologic Functioning Analyses *55*
[redacted]
8 CONVENTIONS *57*
9 CONTENT OF REPORTS *59*
10 DOCUMENT HISTORY *60*
11 REFERENCES *63*

LIST OF TABLES

Table 2.4-1: Protocol Amendments....17

Table 4.2.1.2-1: RANO Criteria for Response Assessment Incorporating MRI and Clinical Factors[a] 22

Table 4.2.1.2-2: Censoring Scheme used in Primary Analysis of PFS 24

[redacted]

Table 4.3.12.1-1: Time Windows for EORTC-QLQ-C30 and EORTC-QLQ-BN-20 Assessments 29

Table 5-2: Study Design - Original versus Revised 32

Table 7.4.1-1: Study Therapy Parameter Definitions....40

Table 7.5.4.2-1: Censoring Scheme 1 for Sensitivity Analysis of PFS....46

Table 7.6.7-1: Immune-Mediated Adverse Events....49

[redacted]

Table 10-1: Document History....60

LIST OF FIGURES

Figure 2.1-1: Study Design Schematic 11

1 BACKGROUND AND RATIONALE

This study is designed to include multiple Phase 1 cohorts to establish safety of nivolumab and the combination of nivolumab plus ipilimumab across different lines of Glioblastoma (GBM), and a Phase 3 cohort to compare the safety and efficacy of nivolumab versus bevacizumab in subjects with first recurrence of GBM. In particular, the objectives of this study are to 1) evaluate the safety and tolerability of nivolumab and nivolumab in combination with ipilimumab in subjects diagnosed with recurrent glioblastoma (GBM) in a safety lead-in group (Cohorts 1 and 1b); [redacted] 3) evaluate the safety, tolerability, and efficacy of nivolumab versus standard of care treatment (bevacizumab) in subjects diagnosed with recurrent GBM in a randomized trial (Cohort 2). It is expected that nivolumab will improve overall survival (OS) as compared with bevacizumab in subjects with recurrent GBM previously treated with first line radiotherapy and temozolomide.

This is the first study examining monoclonal antibodies targeting immune checkpoint inhibitors in subjects with different lines of GBM. To ensure that nivolumab monotherapy and nivolumab in combination with ipilimumab is tolerable, a safety lead-in (Cohorts 1and 1b) evaluating the tolerability of both treatment regimens was initiated prior to advancing a treatment arm into Cohort 2. Cohort 1 (selected US sites only) examined in a randomized fashion, the safety and tolerability of nivolumab monotherapy dosed at 3 mg/kg every 2 weeks (Arm N) or nivolumab 1 mg/kg + ipilimumab 3 mg/kg every 3 weeks for four doses, then nivolumab 3mg/kg every 2 weeks thereafter (Arm N1+I3). Cohort 1b (selected US sites only) examined, in a non-randomized fashion, the safety and tolerability of (nivolumab 3 mg/kg + ipilimumab 1 mg/kg every 3 weeks for four doses, then nivolumab 3mg/kg every 2 weeks thereafter (Arm N3+I1)). (Note: Cohort 1b was initiated after enrollment in Cohort 1 was completed via protocol amendment 03).

[redacted]

[redacted]

[redacted]

Treatment with nivolumab monotherapy in Cohort 1 met the protocol prespecified safety and tolerability profile (less than one-third of subjects required permanent discontinuation due to treatment related adverse events prior to receiving 4 doses of treatment) to advance to Cohort 2, the randomized portion of the study to compare nivolumab monotherapy versus bevacizumab. Since evaluation of a second dosing regimen for the combination therapy is on-going in Cohorts 1 and 1b, Cohort 2 was designed to evaluate nivolumab monotherapy versus bevacizumab. Evaluation of the efficacy of nivolumab + ipilimumab combination therapy will be evaluated separately, once there is sufficient data from Cohort 1 and Cohort 1b to select the appropriate dose of nivolumab + ipilimumab to advance into a randomized study.

Nivolumab dosing in Cohort 2 is based upon preliminary clinical experience in subjects with recurrent GBM from Cohort 1 and prior experience in the treatment of other tumor types (eg, melanoma). Subjects in the nivolumab arm will receive 3 mg/kg every 2 weeks, and subjects in the bevacizumab arm will be dosed at 10 mg/kg every 2 weeks. Subjects will continue to receive study medication until confirmed tumor progression, unacceptable toxicity, or other discontinuation criteria as described in protocol (section 3.5), whichever occurs first and then will enter a follow-up phase to gather information on overall survival.

This study will also provide data regarding survival rate at 12 months, progression free survival, objective response rate, [redacted].

Research Hypothesis:

Treatment with nivolumab monotherapy will improve overall survival (OS) as compared with bevacizumab in subjects with first recurrence of glioblastoma (GBM) treated with prior radiotherapy and temozolomide (Cohort 2).

This study also includes four Phase 1 safety lead-in cohorts at selected US sites to evaluate the safety and tolerability of nivolumab monotherapy [redacted] and nivolumab in combination with ipilimumab (Cohorts 1 and 1b). Enrollment to Cohort 2 (efficacy) will begin after the safety evaluation of nivolumab monotherapy of Cohort 1.

Schedule of Analyses for Cohort 2:

OS is the primary endpoint for cohort 2 in this study. The final OS analysis is planned to be performed when at least 300 events have been observed.

An independent Data Monitoring Committee (DMC) will monitor the data for safety and have access to periodic interim safety and efficacy reports to allow for a benefit/risk assessment. Details are specified in the DMC Charter[1].

2 STUDY DESCRIPTION

2.1 Study Design

This is a randomized, open-label, multicenter, phase III study of nivolumab monotherapy versus bevacizumab and a safety study of nivolumab and nivolumab in combination with ipilimumab

therapy in adult (≥ 18 years) subjects with a first recurrence of glioblastoma (GBM) after treatment with radiotherapy and temozolomide. Since the use of nivolumab and ipilimumab has not previously been studied in subjects with GBM, a safety lead-in (Cohort 1) consisting of approximately 10 subjects in each of two treatment arms will be enrolled and randomized to treatment with either Arm N (nivolumab monotherapy) or Arm N1+I3 (nivolumab combined with ipilimumab). Tolerability and safety of a treatment arm will be determined after all subjects in the arm have completed four doses or have discontinued dosing prior to completing four doses. Subjects will continue treatment until confirmed progression or study discontinuation for any other reason. In order to better characterize preliminary safety and tolerability for the combination therapy, an additional, non-randomized safety cohort (Cohort 1b) to evaluate an alternate dosing regimen for the combination therapy (nivolumab 3mg/kg + ipilimumab 1 mg/kg every 3 weeks for 4 cycles followed by nivolumab 3 mg/kg thereafter) was initiated based on a site specific amendment to the protocol for selected US sites. Information from subjects treated with combination therapy in Cohort 1 and Cohort 1b will provide a basis for dose selection for future studies in subjects with GBM.

[redacted]

Enrollment in Cohort 2 was planned to begin after safety and tolerability of the safety lead-in had been evaluated and determination regarding which treatments arms (Arm N and/or Arm N+I) will be studied. Treatment with nivolumab monotherapy in Cohort 1 met the protocol prespecified safety and tolerability profile (< one-third of subjects required permanent discontinuation due to treatment related adverse events prior to receiving 4 doses of treatment) to advance to Cohort 2, the randomized portion of the study to compare nivolumab monotherapy versus bevacizumab. The randomized portion of this study (Cohort 2) will be limited to nivolumab monotherapy versus bevacizumab. Evaluation of the efficacy of nivolumab + ipilimumab combination therapy will be evaluated separately, once there is sufficient data from Cohort 1 and Cohort 1b to select the appropriate dose to advance into a randomized study.

All subjects in Cohort 1, 1b, [redacted] and 2 will be followed for safety and tolerability, tumor progression, and overall survival. Tumor progression or response endpoints will be assessed using the Radiologic Assessment in Neuro-Oncology (RANO) criteria. Treatment with study

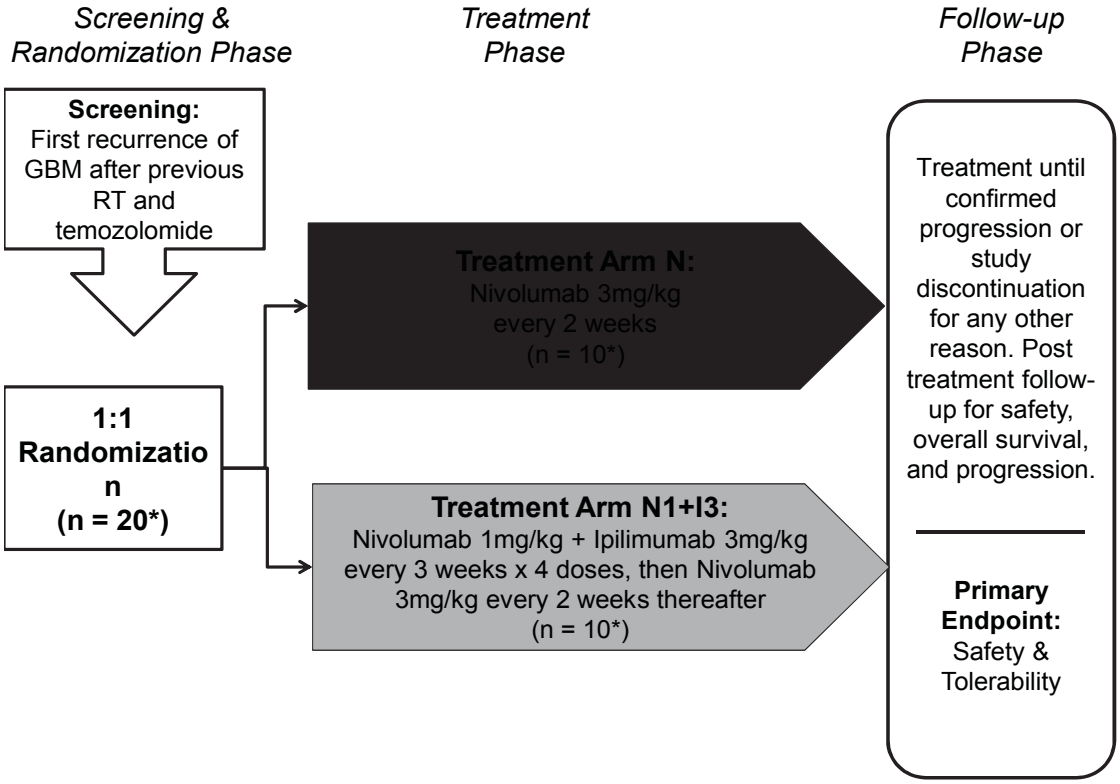
medication will continue until confirmed progression, unacceptable toxicity or other discontinuation criteria as described in protocol (section 3.5). A Data Safety Monitoring Committee will meet regularly during the study to ensure that subject safety is carefully monitored

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

Figure 2.1-1: Study Design Schematic

Cohort 1 – Safety Lead-In for Recurrent GBM:

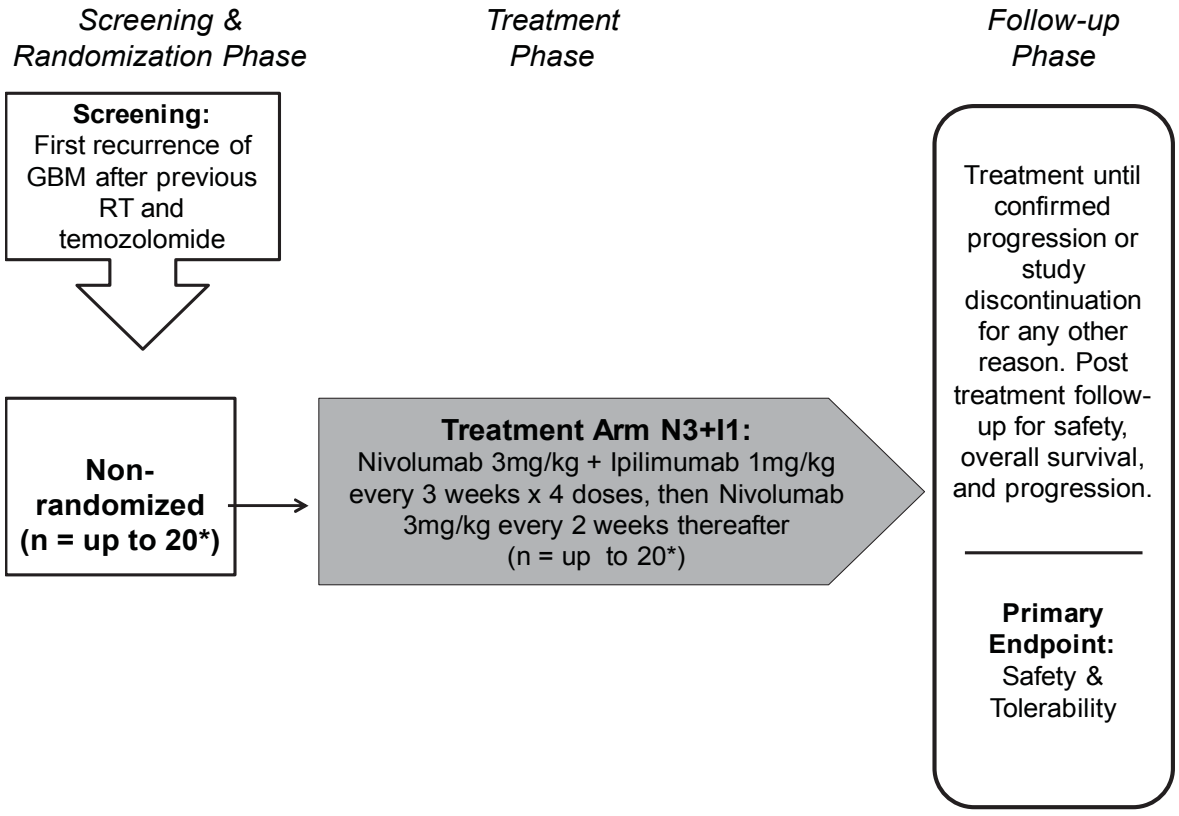
- Primary endpoint: Safety and tolerability of Nivolumab and N1 + I3 in subjects with recurrent GBM
- Demonstration of acceptable tolerability and safety before advancing to Cohort 2



*Cohort 1 will randomize 20 subjects in a 1:1 fashion to assess the tolerability and safety of the investigational agents and suitability of advancing the treatment arms to Cohort 2, the randomized efficacy phase of the study. The tolerability and safety assessment of each arm will occur after the last randomized subject has completed four doses or has discontinued dosing prior to completing four doses.

Cohort 1b: Safety evaluation of combination of Nivolumab 3mg/kg +ipilimumab 1mg/kg for recurrent GBM

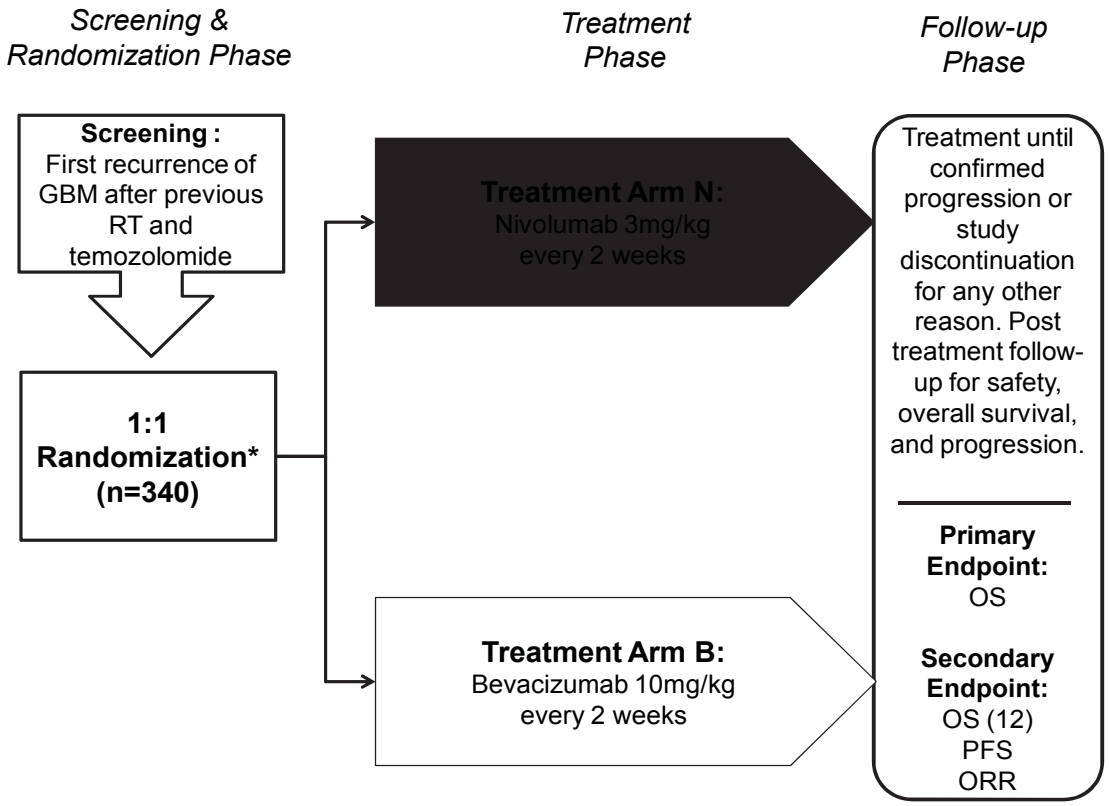
- Primary endpoint: Safety and tolerability of N3+ I1 combination arm in recurrent GBM



*Cohort 1b will be non-randomized with up to 20 subjects treated at selected US sites only.

Cohort 2 – Randomized, 2-Arm Efficacy and Safety Study for Recurrent GBM:

- **<u>Primary endpoint:</u> Overall Survival; <u>Secondary endpoints:</u> OS(12), PFS, and ORR**
- **1:1 Randomization**



*This study will consist of three phases: screening, treatment, and follow up. After confirmed progression or study discontinuation for any other reason, study treatment will be discontinued and subjects will enter the post treatment follow-up phase to assess safety, progression and overall survival. Randomization will be stratified by presence of a measurable lesion at baseline (yes/no).

2.2 Treatment Assignment

The subject number will be assigned through an interactive voice response system (IVRS) once the subject has signed the informed consent form and is registered. Every subject that signs the informed consent form must be assigned a subject number in IVRS. Specific instructions for using IVRS will be provided to the investigational site in a separate document.

The investigator or designee will register the subject for enrollment by following the enrollment procedures established by BMS. The following information is required for enrollment:

- Date that informed consent was obtained
- Date of birth
- Gender at birth

Once enrolled in IVRS, enrolled subjects that have met all eligibility criteria will be ready to be randomized through the IVRS. The following information is required for subject randomization:

- Subject number
- Date of birth
- Presence of a measurable lesion at baseline (yes/no) (Cohort 2 only)

Subjects meeting all eligibility criteria will be randomized in a 1:1 ratio to Arm N or Arm N1+I3 for Cohort 1. Subjects meeting all eligibility criteria will be randomized in a 1:1 ratio to Arm N or Arm B, stratified by presence of a measurable lesion at baseline (yes/no) for Cohort 2. Subjects enrolled to Cohort 1b will be assigned to Arm N3+I1.

[redacted]

2.3 Blinding and Unblinding

Not applicable. This is an open-label study.

2.4 Protocol Amendments

This SAP (version 4.0) corresponds to the revised protocol 4d (US only) titled "A Randomized Phase 3 Open Label Study of Nivolumab versus Bevacizumab and Multiple Phase 1 Safety Cohorts of Nivolumab or Nivolumab in Combination with Ipilimumab Across Different Lines of Glioblastoma" and revised protocol 6 titled "A Randomized Phase 3 Open Label Study of Nivolumab versus Bevacizumab and a Safety Study of Nivolumab or Nivolumab in Combination with Ipilimumab in Adult Subjects with Recurrent Glioblastoma (GBM)".

It incorporates the following amendments:

Table 2.4-1: Protocol Amendments

Amendment	Date of Issue	Summary of Major Changes
Revised Protocol 01 (Incorporate Amendment 01)	10-Dec-2013	Adds 12-lead ECG at screening and as clinically indicated while on treatment. Clarifies guidance regarding corticosteroid use during the study. Includes guidance on the use of PPI or H2 blockers for patients on chronic steroids. Excludes patients with a history of gastrointestinal diverticulitis Corrects inconsistencies in the statistical considerations section.
Revised Protocol 01 (Incorporate Amendment 02)	18-Feb-2014	Add "HIV" to list of laboratory parameters to be tested at screening (this is site specific amendment for Germany only)
Revised Protocol 01a	14-May-2014	Adds an additional combination treatment arm of nivolumab 3mg/kg in combination with ipilimumab

Table 2.4-1: Protocol Amendments

Amendment	Date of Issue	Summary of Major Changes
(Incorporate Amendment 03)		1 mg/kg every 3 weeks x 4 doses followed by nivolumab 3 mg/kg every 2 weeks thereafter. The additional combination arm is identified as Cohort 1b and will treat up to 20 subjects..
Revised Protocol 02 (Incorporate Amendment 04)	15-Jul-2014	Clarify the study design and rationale for the initiation of the efficacy study (Cohort 2). Treatment with nivolumab monotherapy in Cohort 1 met the protocol prespecified safety and tolerability profile to advance to a randomized trial versus bevacizumab in Cohort 2. Information regarding the combination dosing of nivolumab plus ipilimumab from Cohort 1 and 1b is still being evaluated Combination dosing will not be included in Cohort 2 The study design, rationale for dose selection, study hypothesis, study objectives, and statistical considerations were updated accordingly Provides guidance on treatment after a confirmed complete response Update inclusion and exclusion criteria [redacted] Clarifies continuation of treatment post-resection Clarifies analysis of best overall response (BOR)
Revised Protocol 3 (Includes amendment 05)	15-Aug-2014	[redacted] • Added clarification for allowing resection for assessment of progression/pseudoprogression • Added central neuropathic review of tumor samples obtained after biopsy or resection in subjects for whom determination of progression versus pseudoprogression cannot be determined • Changed confirmation of progression to be 12 weeks rather than 8 weeks after initial radiologic progression for subjects who meet criteria for continuation of study treatment • Added table to summarize assessment of best overall response (BOR) • Combined site specific amendment for Cohort 1b into global protocol
Revised Protocol 03a (Includes Amendment 06)	04-Feb-2015	• Updated title of protocol • [redacted]
Revised Protocol 4a	27-Feb-2015	• Increase sample size of Cohort 2 by 120 subjects

Table 2.4-1: Protocol Amendments

Amendment	Date of Issue	Summary of Major Changes
(Includes amendment 07)		to 340 • Change from 2 planned interim analyses to a single interim analysis at 80% of the total number of events • Added clarification to allow collection of OS data outside the protocol specified windows
[redacted]	[redacted]	[redacted]
Revised Protocol 4c (US specific) and 5 (Global) (Includes amendment 09)	23-Feb-2016	• To modify the timing of interim (IA) and final (FA) analyses based on external data suggesting delayed separation of overall survival (OS) Kaplan-Meier (KM) curve and a long -lasting plateau at the end of OS KM curve.
Revised protocol 4d (US specific) and 6 (Global)		• Cancel interim analysis and proceed directly to final analysis (at least 300 OS events).

2.5 Data Monitoring Committee

An independent Data Monitoring Committee (DMC) has been established to provide oversight of safety and efficacy considerations, study conduct, and risk-benefit ratio. Following review, the DMC will recommend continuation, modification, or discontinuation of this study based on reported safety and efficacy data. Details of DMC responsibilities and procedures are specified in the DMC charter[1]. Representatives of the Sponsor will serve only as coordinators of the committee, without having full member responsibilities or privileges. In addition, the Sponsor will independently review safety data in a blinded manner during the cohort 2 portion of this trial to ensure that any safety issues are identified and addressed. Since the safety lead-in portion of trial will help characterize the safety and tolerability of nivolumab and nivolumab plus ipilimumab, safety review of Cohorts 1, 1b, [redacted] will be in unblinded manner during the study.

3 OBJECTIVES

3.1 Primary

Cohort 1, 1b, [redacted] (Safety Lead-in):

- To evaluate the safety and tolerability of nivolumab and nivolumab in combination with ipilimumab in subjects with recurrent GBM (Cohorts 1 and 1b)

[redacted]

Cohort 2:

- To compare the overall survival (OS) of nivolumab versus bevacizumab in subjects with recurrent GBM

3.2 Secondary

[redacted]

Cohort 2 only:

- To compare the overall survival rate at 12 months (OS(12)) of nivolumab versus bevacizumab
- To compare progression free survival (PFS) of nivolumab versus bevacizumab
- To compare the objective response rate (ORR) of nivolumab versus bevacizumab

3.3 Exploratory

These are exploratory endpoints for all cohorts unless otherwise specified.

- To evaluate progression free survival (PFS) and objective response rate (ORR) of nivolumab (Cohort 1, [redacted]) and nivolumab in combination with ipilimumab (Cohort 1and 1b)
- To evaluate the safety of nivolumab and bevacizumab (Cohort 2, only)

[redacted]

[redacted]

- To describe steroid sparing effects in glioblastoma subjects treated with nivolumab, nivolumab plus ipilimumab, or bevacizumab.

[redacted]

4 ENDPOINTS

4.1 Primary

4.1.1 Cohorts 1, 1b, [redacted]

The primary objective (to assess the safety and tolerability of nivolumab monotherapy or nivolumab combined with ipilimumab in subjects with recurrent or newly diagnosed GBM) will be primarily assessed by the rate of treatment-related AEs leading to drug discontinuations prior to complete four dose of treatment. In addition, safety and tolerability will be analyzed through the incidence of adverse events, serious adverse events, and specific laboratory abnormalities (worst grade) in each treatment arm. Toxicities will be graded using the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) version 4.0.

4.1.2 Cohort 2

The primary endpoint of Cohort 2 is overall survival (OS). OS for cohort 2 is defined as the time between the date of randomization and the date of death due to any cause. For subjects without documentation of death, OS will be censored at the last known alive date. OS will be followed based on the study assessment schedule while subjects are on study drug and every 3 months via in-person or phone contact during the survival follow-up phase of the study.

4.2 Secondary

4.2.1 Cohort 2

4.2.1.1 Overall Survival Rate at 12 Months (OS(12))

OS(12) is the percentage of subjects who are still alive at 12 months after randomization. It is estimated using Kaplan-Meier (KM) product-limit method.

4.2.1.2 Progression Free Survival

PFS for cohort 2 is defined as the time from randomization to the date of the first documented tumor progression (per RANO, Table 4.2.1.2-1) or death due to any cause. Subjects who die without a reported prior progression will be considered to have progressed on the date of death. Subjects who did not have documented disease progression or die will be censored at the date of the last assessment for disease progression. Subjects who did not have any on study tumor assessment and did not have disease progression or die will be censored at the randomization date. Subjects who started any subsequent anti-cancer therapy without a prior reported progression will be censored at the last assessment for disease progression prior to initiation of the subsequent anti-cancer therapy which excludes the surgical resection for differentiating radiology progression from pseudoprogression (diagnostic surgical resection, protocol section 4.3.7). Subject who had diagnostic surgical resection without a prior documented progression and tumor pathology does not confirm disease progression will be censored at the last tumor assessment prior to initiation of the diagnostic surgical resection. First tumor assessment will occur at the end of week 6 (+/- 1 week), end week 12 (+/- 1 week) and then every 8 weeks (+/- 1 week) thereafter until disease progression is documented.

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.

Censoring rules for the primary analysis of PFS are presented in Table 4.2.1.2-2. Alternate censoring rules for sensitivity analyses are specified in Section 7.5.4.2.

Table 4.2.1.2-1: RANO Criteria for Response Assessment Incorporating MRI and Clinical Factors[a]

Response	Criteria
Complete response	Requires all of the following: complete disappearance of all enhancing measurable and nonmeasurable disease sustained for at least 4 weeks; no new lesions; stable or improved nonenhancing (T2/FLAIR) lesions; patients must be off corticosteroids (or on physiologic replacement doses only); and stable or improved clinically. Note: Patients with nonmeasurable disease only cannot have a complete response; the best response possible is stable disease.
Partial response	Requires all of the following: ≥ 50% decrease compared with baseline in the sum of products of perpendicular diameters of all measurable enhancing lesions sustained for at least 4 weeks; no progression of nonmeasurable disease; no new lesions; stable or improved nonenhancing (T2/FLAIR) lesions on same or lower dose of corticosteroids compared with baseline scan; the corticosteroid dose at the time of the scan evaluation should be no greater than the dose at time of baseline scan; and stable or improved clinically. Note: Patients with nonmeasurable disease only cannot have a partial response; the best response possible is stable disease.
Stable disease	Requires all of the following: does not qualify for complete response, partial response, or progression; stable nonenhancing (T2/FLAIR) lesions on same or lower dose of corticosteroids compared with baseline scan. In the event that the corticosteroid dose was increased for new symptoms and signs without confirmation of disease progression on neuroimaging, and subsequent follow-up imaging shows that this increase in corticosteroids was required because of disease progression, the last scan

Table 4.2.1.2-1: RANO Criteria for Response Assessment Incorporating MRI and Clinical Factors[a]

Response	Criteria
	considered to show stable disease will be the scan obtained when the corticosteroid dose was equivalent to the baseline dose.
Progression	Defined by any of the following: ≥ 25% increase in sum of the products of perpendicular diameters of enhancing lesions compared with the smallest tumor measurement obtained either at baseline (if no decrease) or best response, on stable or increasing doses of corticosteroids[*]; significant increase in T2/FLAIR nonenhancing lesions on stable or increasing doses of corticosteroids compared with baseline scan or best response after initiation of therapy[*] not cause by comorbid events (eg, radiation therapy, demyelination, ischemic injury, infection, seizures, postoperative changes, or other treatment effects); any new lesion; clear clinical deterioration not attributable to other causes apart from the tumor (eg, seizures, medication adverse effects, complications of therapy, cerebrovascular events, infection, and so on) or changes in corticosteroid dose; failure to return for evaluation as a result of death or deteriorating condition; or clear progression of nonmeasurable disease.

[a] From: Wen PY, Macdonald DR, Reardon DA, Cloughesy TF, Sorenson AG, Galanis E, et al. Updated Response Assessment Criteria for High-Grade Gliomas: Response Assessment in Neuro-Oncology Working Group; J Clin Oncol. 2010 Apr 10;28(11):1963-72.

NOTE: Radiologic interpretation guidelines, definitions and tumor measurement instructions will be provided separately. All measurable and nonmeasurable lesions must be assessed using the same techniques as at baseline.

Abbreviations: MRI, magnetic resonance imaging; FLAIR, fluid-attenuated inversion recovery

[*] Stable doses of corticosteroids include subjects not on corticosteroids

Table 4.2.1.2-2: Censoring Scheme used in Primary Analysis of PFS

Situation	Date of Progression or Censoring	Outcome
No baseline tumor assessments and no disease progression/death	Date of Randomization	Censored
No on study tumor assessments and no disease progression/death	Date of Randomization	Censored
Documented progression per investigator assessment (including clinical progression)	Date of the first documented progression per RANO	Progressed
Diagnostic surgical resection without documented progression prior or on the same day	Date of last tumor assessment prior to initiation of the diagnostic surgical resection	Censored
Diagnostic surgical resection with documented progression (per CRF) prior to surgical resection	Date of progression (per CRF)	Progressed
No progression and no death	Date of last assessment for disease progression	Censored
Subsequent therapies, including new anticancer therapy, tumor-directed radiotherapy, or tumor-directed surgery received, without progression reported prior or on the same day	Date of last assessment for disease progression prior to initiation of the subsequent therapy	Censored
Death after initiation of subsequent cancer therapy	Date of last assessment for disease progression prior to initiation of the subsequent therapy	Censored
Death without progression	Date of death	Progressed

4.2.1.3 Objective Response Rate(ORR)

ORR is defined as the number of subjects whose best overall (BOR) response is a *confirmed* complete response (CR) or *confirmed* partial response (PR) divided by response-evaluable subjects (section 6.3). BOR is defined as the best response designation, recorded between the date of randomization and the date of objectively documented progression per RANO criteria, the date of subsequent therapy or the date of pathology results from diagnostic surgical resection, whichever occurs first. For subjects without documented progression, subsequent therapy, or diagnostic surgical resection, all available response designations will contribute to the BOR assessments. BOR will be determined by investigator reported response based on RANO criteria.

For purposes of this protocol analysis, if a response evaluable subject discontinues from the study or receives a subsequent therapy prior to the 6 week tumor assessment; this subject will be counted in the denominator as a non-responder.

[redacted]

4.3 Exploratory

4.3.1 OS for Cohort 1 and 1b

OS for cohorts 1 and 1b is same as OS for cohort 2 (see Section 4.1.2. for details)

4.3.2 PFS for Cohort 1, 1b, [redacted]

PFS for cohorts 1, 1b, [redacted] is defined as the time from first dose to the date of the first documented tumor progression or death due to any cause. Censoring rules for the analysis of PFS for cohorts 1, 1b, [redacted] are same as censoring rules for cohort 2 and are presented in Table 4.2.1.2-2

4.3.3 Duration of Response (DOR)

DOR is defined as the time between the date of first documented response (CR or PR) to the date of first documented tumor progression (per RANO criteria), or death due to any cause, whichever occurs first. For subjects who neither progress nor die, the duration of objective response will be censored at the same time they will be censored for the primary definition of PFS (Table 4.2.1.2-2). DOR will be evaluated for responders (ie, subjects with confirmed CR or PR) only.

4.3.4 Time to Response(TTR)

TTR is defined as the time from randomization for cohort 2 and first dose for cohorts 1, 1b, [redacted] to the date of the first documented CR or PR per investigator reported RANO criteria. TTR will be evaluated for responders only.

4.3.5 Magnitude of Tumor burden

The magnitude of tumor burden is defined as the percent decrease in tumor size from baseline to nadir, observed up until the date of progression, per RANO criteria, the date of subsequent anticancer therapy (including tumor-directed radiotherapy and tumor-directed surgery including diagnostic surgical resection), or death, whichever occurs first.

[redacted]

4.3.8 PD-L1 Expression

PD-L1 expression missing: Subjects without an available tumor biopsy specimen for PD-L1 evaluation will be considered as PD-L1 expression missing.

For subjects with an available tumor biopsy specimen(s), the following will be considered:

- PD-L1 expression is defined as the percent of tumor cell membrane staining in a minimum of 100 evaluable tumor cells per validated Dako PD-L1 IHC assay unless otherwise specified. This is referred as quantifiable PD-L1 expression. If the PD-L1 staining could not be quantified, it is further classifies as:
- Indeterminate: Tumor cell membrane staining hampered for reasons attributed to the biology of the tumor biopsy specimen and not because of improper sample preparation or handling
- Not evaluable: Tumor biopsy specimen was not optimally collected or prepared (e.g. PD-L1 expression is neither quantifiable nor indeterminate)
- Baseline PD-L1 expression: If more than one tumor biopsy specimen is available, baseline PD-L1 expression will be determined from the most recently collected specimen (prior to first dose of study treatment) with a quantifiable result. If all specimens for a given subject are either indeterminate or not evaluable, then the PD-L1 expression will be considered indeterminate as long as at least one specimen is indeterminate. Otherwise, PD-L1 expression will be considered not evaluable.

PD-L1 status is a dichotomized variable using a X% cut-off for quantifiable PD-L1 expression:

- PD-L1 ≥ X %: ≥ X % PD-L1 expression
- PD-L1 < X %: < X % PD-L1 expression

Where X% denotes the PD-L1 expression cut-off of 1% and 5%. Additional cut off values may also be explored. Values above and below the cut off are referred to as PD-L1 positive and negative respectively

[redacted]

- bladder control: (Q20-1)/3 * 100

[redacted]

5 SAMPLE SIZE AND POWER

This study comprises of [redacted] safety cohorts (cohorts 1, 1b, [redacted]) and a randomized phase 3 randomized, open-label, study of nivolumab monotherapy versus bevacizumab (cohort 2) in adult (≥ 18 years) subjects with GBM.

Cohort 1 consists of two arms to evaluate the safety and tolerability profile of nivolumab (N) and nivolumab in combination with ipilimumab (N1+I3, nivolumab 1 mg/kg + ipilimumab 3 mg/kg) in subjects with recurrent GBM. For this purpose, approximately 20 subjects were randomized in a 1:1 ratio to arm N or arm N1 + I3. Cohort 1b was added to characterize an alternative dosing regimen for the combination therapy of nivolumab and ipilimumab (N3+ I1, nivolumab 3 mg/kg + ipilimumab 1 mg/kg) in subjects with recurrent GBM. Cohort 1b included up to 20 treated subjects. [redacted]

Cohort 2 is a randomized two arm study with primary objective to compare the overall survival (OS) of nivolumab (N) versus bevacizumab (B).

Results from Ipilimumab and Nivolumab studies have demonstrated long term survival benefit in patients treated with immuno-oncology drugs observed as a long lasting plateau towards the end of survival curve. Data from these studies have also suggested delayed effect observed as late separation of survival curves between experimental and treatment arms. Both long-term survival benefit and delayed onset of benefit may be attributed to immuno-oncology drugs based on their mechanism of action. Recent data from a different immunotherapy compound also supports this observation in recurrent GBM population.[6]

Based on observations from these studies, following key assumptions are considered appropriate for this study design:

- 4-month delayed separation of OS curves between nivolumab and bevacizumab
- Bevacizumab Arm: OS follows exponential distribution with median of 9.2 months
- Nivolumab Arm: 10% "cure rate" (long-term survival) and exponential distribution with a median of 13.7 months after first four months for "non-cured" nivolumab subjects.

Final analysis of OS will be conducted when at least 300 deaths have been observed among 369 randomized subjects. There will be no interim analysis for efficacy.

Given the observed accrual of 369 randomized subjects and survival assumptions stated above, simulations were conducted using piecewise exponential model for treatment group which closely approximates the assumed cure rate model for Nivolumab. Based on the assumptions for survival model, following hazard rate for Nivo/hazard ratio is used for simulation:

Months	0-<4	4-<8	8-<12	12-<16	16-<20	20-<24	24-<28	28-<32	32-<36	36+
Hazard rate for Nivolumab	.075	0.045	0.043	0.042	0.041	0.039	0.037	0.035	0.034	0.032
Hazard Ratio*	1.00	0.591	0.576	0.557	0.538	0.518	0.488	0.465	0.446	0.43

*Based on hazard rate for bevacizumab arm = 0.07534, which corresponds to median of 9.2 months under exponential distribution.

Based on simulations, power at FA will be approximately 92% based on this model.

In case target number of OS events is not met for FA, 25 months after last patient is randomized, FA will be conducted. Follow-up of 25 months is considered ample for mature OS curve. Type I error rate will be 5%.

All simulations were conducted using East v6.3.1.

Three hundred and sixty-nine (369) subjects have been randomized to the two arms (nivolumab vs bevacizumab) in a 1:1 ratio stratified by presence of a measureable lesion at baseline (yes/no) in 9 months. Details for comparison between previous and revised design are included in Table 5-1.

Table 5-1: Study Design - Original versus Revised

	Previous Design [redacted]	Revised Design
N randomized / Randomization	**n=369 (actual)/ 1:1 (nivo: bevacizumab)**	**n=369 (actual)/ 1:1 (nivo: bevacizumab)**
Target OS effect	**Bevacizumab: mOS = 9.2 months (follows exponential distribution)** **Nivolumab: Non-proportional hazards model based on following key assumptions:** **4 months of delayed separation of curves** **10% 'cure rate' for nivolumab** **13.7 months of mOS for 'non-cured'** **Power (based on simulation): ~90%**	**Bevacizumab: mOS = 9.2 months (follows exponential distribution)** **Nivolumab: Non-proportional hazards model based on following key assumptions:** **4 months of delayed separation of curves** **10% 'cure rate' for nivolumab** **13.7 months of mOS for 'non-cured'** **Power (based on simulation): ~92%**
Timing of OS IA (from first subject randomized)	**≥ 90%(270) OS events (27 months, 18 months after last patient randomized)**	**No interim analysis**
Final OS (from first subject randomized)	**300 OS events (34 months, 25 months after last patient randomized)**	**300 OS events (34 months, 25 months after last patient randomized)**

The overall survival at 12 months (OS(12)), progression free survival (PFS), and objective response rate (ORR) are secondary endpoints of Cohort 2 and will be tested using a hierarchical testing procedure in order to control a family-wise type I error of 0.05. Study will have more than 90% power to detect HR of 0.67 or lower for PFS (assuming a median PFS of 4.2 months for arm B) and a difference of 20% or more for ORR at two-sided significance level of 0.05.

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 (or date of randomization in case of no 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.

If more than one tumor biopsy specimen is available, baseline PD-L1 expression will be determined from the most recently collected specimen (prior to first dose of study treatment) with a measurable result. If all specimens for a given subject are either indeterminate or unknown, then the PD-L1 expression will be considered indeterminate as long as at least one specimen is indeterminate. Otherwise, PD-L1 expression will be considered unknown.

6.1.2 Post Baseline Period

On-treatment AEs will be defined as AEs with an onset date-time on or after the date-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[2)] of the last dose of study treatment.

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. 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.

6.2 Treatment Regimens

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

Cohort 1:

- Arm N3: nivolumab 3mg/kg
- Arm N1+I3: nivolumab 1mg/kg + ipilimumab 3mg/kg

- Cohort 1b:
  - Arm N3+I1: nivolumab 3mg/kg + ipilimumab 1mg/kg

Cohort 2:

- Arm N: nivolumab 3mg/kg
- Arm B: bevacizumab 10mg/kg

[redacted]

[redacted]

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

All subjects will be treated with Arm N3+I1 (nivolumab 3mg/kg + ipilimumab 1mg/kg) in cohort 1b.

6.3 Populations for Analyses

- Enrolled Subjects: All subjects who signed an informed consent form and were registered into the IVRS.
- Randomized Subjects: All enrolled subjects who were randomized to any treatment arm in the study. This is the primary dataset for analyses of baseline demographics and efficacy analyses of Cohort 2.
- Treated subjects in Cohorts 1, 1b, [redacted]: All subjects in Cohorts 1, 1b, [redacted] who received at least one dose of study drug. This is the dataset for baseline demographics, efficacy, and safety evaluation of Cohorts 1, 1b, [redacted]
- Treated subjects in Cohort 2: All randomized subjects in Cohort 2 who received at least one dose of study drug. This is the dataset for safety evaluation of Cohort 2.

- Response-Evaluable Subjects: All randomized subjects (for Cohort 2) or all treated subjects (for Cohorts 1, 1b, [redacted]) with measurable disease at a baseline tumor assessment (Source: CRF).
- PK subjects: All randomized subjects (for Cohort 2) or all treated subjects (for Cohorts 1, 1b, 1c and 1d) with available serum time-concentration data.

[redacted]

- All tested PD-L1 subjects: All subjects who had a tumor biopsy specimen available for assessment of PD-L1 expression.
  - Cohort 2
    - All randomized PD-L1 subjects: All tested PD-L1 subjects who are randomized to the study in and with baseline PD-L1 expression.
    - All evaluable PD-L1 subjects: All randomized PD-L1 subjects with quantifiable PD-L1 expression.
    - All treated PD-L1 subjects: All randomized PD-L1 subjects who received at least one dose of study treatment.
  - Cohorts 1, 1b, [redacted]
    - All evaluable PD-L1 subjects: All treated PD-L1 subjects with quantifiable PD-L1 expression.
    - All treated PD-L1 subjects: All treated PD-L1 subjects who received at least one dose of study treatment.
- Biomarker subjects: All randomized subjects (for Cohort 2) or all treated subjects (for Cohorts 1, 1b, [redacted]) with available biomarker data

[redacted]

7 STATISTICAL ANALYSES

7.1 General Methods

Unless otherwise noted, the 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 the mean, standard deviation, median, minimum and maximum values. 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 (ie, progression free survival, overall survival, time to response and duration of response) will be estimated using Kaplan Meier techniques. When appropriate, the median along with the corresponding log-log transformed 95% CI will be estimated. Rates at fixed timepoints (eg, OS at 12 months) will be derived from the Kaplan Meier estimate along with their corresponding log-log transformed 95% confidence intervals[7]. Confidence intervals for binomial proportions will be derived using the Clopper-Pearson method[8].

For cohort 2, unless otherwise specified, the stratified log-rank test will be performed to test the comparison between time to event distributions (OS and PFS). Stratification factors will be the presence of a measureable lesion at baseline (yes/no) as entered into the IVRS.

For cohort 2, unless otherwise specified, the stratified hazard ratio between 2 groups along with CI will be obtained by fitting a stratified Cox model with the group variable as unique covariate. Stratification factor will be the presence of a measureable lesion at baseline (yes/no) as entered into the IVRS.

For cohort 2, difference in ORR between the two arms and corresponding 95% CI will be estimated using Cochran-Mantel-Haenszel (CMH) method of weighting, adjusting for the presence of measurable lesion at baseline as entered into the IVRS

P-values other than those provided for the OS primary analysis and the hierarchical analysis of key secondary endpoints (Section 4.2) are for descriptive purpose only and not adjusted for multiplicity.

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 by cohorts. Randomization date, first dosing date, country, investigational site will be presented in a by cohort and subject listing of accrual.

7.2.2 Relevant Protocol Deviations

The following programmable deviations will be considered as relevant protocol deviations and summarized by cohort, treatment group and overall within each cohort in all randomized subjects for cohort 2 and all treated subjects for cohorts 1, 1b, [redacted]. 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 without histologically confirmed diagnosis of World Health Organization Grade IV malignant glioma (glioblastoma or gliosarcoma)

- Subjects without previous first line treatment with at least radiotherapy and temozolomide (for Cohorts 1, 1b and 2 only)
- Subjects without documented first recurrence of GBM by diagnostic biopsy or contrast enhanced magnetic resonance imaging (MRI) per RANO criteria(for Cohorts 1, 1b and 2 only)
- Subjects without Karnofsky performance status of 70 or higher

On-study:

- Subjects receiving anti-cancer therapy (chemotherapy, hormonal therapy, immunotherapy, standard or investigational agents for treatment of cancer) other than study therapy while on study therapy.
- Subjects treated differently than as assigned (subjects who received the wrong treatment, excluding the never treated) (for Cohort and 2 only).

Listings will also be provided.

7.3 Study Population

Unless otherwise specified, the following analyses will be presented by cohort and treatment group as "treated" for cohorts 1, 1b, and as "randomized" for cohort 2. Analysis population for cohorts1, 1b, will consist of the treated population and for cohort 2 will consist of the randomized population.

7.3.1 Subject Disposition

The total number of subjects enrolled (randomized or not randomized) will be presented by cohort along with the reason for not being randomized for cohorts 1 and 2.

Number of subjects who discontinued study treatment along with corresponding reason will be tabulated by cohort and treatment group as treated. Reason for discontinuation will be derived from subject status CRF page.

Number of subjects randomized but not treated along with the reason will be tabulated by cohort and treatment group as randomized for cohorts 1 and 2. Number of subjects enrolled but not treated along with the reason will be tabulated for cohort 1b.

A subject listing for all randomized subjects in cohorts 1, and 2 will be provided showing the subject's randomization date, first and last dosing date, off study date and reason for going off-study. A subject listing for all enrolled subjects in cohort 1b, will be provided showing the subject's enrollment date, first and last dosing date, off study date and reason for going off-study. A subject listing for subjects not randomized in

cohorts 1, [redacted] and 2 will also be provided, showing the subject's race, gender, age, consent date and reason for not being randomized.

7.3.2 Demographics and Baseline Characteristics

The following baseline characteristics will be summarized by cohort and treatment group. All baseline presentations will identify subjects with missing measurements. Listings will also be provided.

- Age (descriptive statistics)
- Age category I (<65, ≥ 65)
- Age category II (<65, ≥ 65- <75, ≥ 75)
- Gender (male, female)
- Race (white, black, asian, other)
- Region (US/Canada, Europe, Rest of the World)
- Presence of a measureable lesion at baseline (Yes, No) (Cohort 2 only. source: IVRS)
- Baseline Karnofsky performance status (100%, 90%, 80%, ...)
- Baseline pathology (Glioblastoma, Gliosarcoma)
- Weight (descriptive statistics)
- Gene Promoter Methylation (Methylated, Unmethylated, Unknown, Not reported)
- Smoking Status (Current/former, Never, Unknown, Not Reported)
- Baseline Corticosteroid Use (< 4mg/day, ≥ 4 mg/day,-No) (Based on average corticosteroid use 5 days prior to start of dosing (randomization date for subjects not treated) in dexamethasone equivalents)
- Time from Initial Disease Diagnosis to Recurrent Disease Diagnosis (<6 months,6 - < 1year, 1 - <2 years, 2 - <3 years, 3 - <4 years and >= 4 years)
- All lesions (Investigator Tumor Assessments at Baseline): sites of disease, number of disease sites per subject.
- Target lesions (Investigator Tumor Assessments at Baseline): Presence of target lesions, site of target lesion, sum of longest diameter of target lesion.
- Pre-treatment events: summarized by worst CTC grade presented by SOC/PT

7.3.3 Medical History

General medical history will be listed by subject.

7.3.4 Prior Therapy

The following will be summarized by cohort and treatment group.

- Prior systemic cancer therapy (yes/no)
- Prior agent received (generic name)
- Prior surgery related to cancer (yes/no)
- Prior radiotherapy (yes/no)

Agents and medication will be reported using the generic name. A listing by subject will also be provided including prior/current non study medication.

7.3.5 Baseline Examinations

Subjects with abnormal baseline physical exam results will be tabulated by examination criteria (eg, neck, cardiovascular, lungs, etc), by cohort and by treatment group.

7.3.6 Discrepancies Between IVRS and CRF Stratification Factors (Cohort 2 only)

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

- Presence of a measureable lesion at baseline (Yes, No)

7.4 Extent of Exposure

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

7.4.1 Administration of Study Therapy

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

- Time from randomization and to first dose of study therapy (0 to 3 days, > 3 to 7, > 7 to 14, > 14 to 21, > 21 to 28, > 28) (cohort 2)

The following parameters will be summarized (descriptive statistics) by cohort, study therapy (nivolumab, ipilimumab and bevacizumab) and treatment group:

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

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

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

The key parameters used to calculate dosing data are shown below.

Table 7.4.1-1: Study Therapy Parameter Definitions

Monotherapy of Nivolumab and Bevacizumab		
	Nivolumab	**Bevacizumab**
Dosing schedule per protocol	3 mg/kg every 2 weeks	10 mg/kg every 2 weeks
Dose	*Dose (mg/kg)* is defined as Total Dose administered (mg)/Most recent weight (kg). Dose administered in mg at each dosing date and weight are collected on the CRF.	*Dose (mg/kg)* is defined as Total Dose administered (mg)/Most recent weight (kg). Dose administered in mg at each dosing date and weight are collected on the CRF
Cumulative Dose	*Cum dose (mg/kg) is* sum of the doses (mg/kg) administered to a subject.	*Cum dose (mg/kg) is* sum of the doses (mg/kg) administered to a subject.
Relative dose intensity (%)	Cum dose (mg/kg)/[(Last dose date - Start dose date + 14) x 3/14] x 100	Cum dose (mg/kg)/[(Last dose date - Start dose date + 14) x 10/14] x 100
Duration of treatment	Last dose date - Start dose date +1	Last dose date - Start dose date +1

Nivolumab combined with Ipilimumab		
	Nivolumab	**Ipilimumab**
Dosing Schedule per Protocol	1 mg/kg every 3 weeks for 4 doses followed by 3 mg/kg every 2 weeks	3 mg/kg every 3 weeks for 4 doses
Dose	*Dose (mg/kg)* is defined as Total Dose administered (mg)/Most recent weight (kg). Dose administered in mg at each dosing date and weight are collected on the CRF	*Dose (mg/kg)* is defined as Total Dose administered (mg)/Most recent weight (kg). Dose administered in mg at each dosing date and weight are collected on the CRF
Cumulative Dose	Cum Dose (mg/kg) is the sum of the doses administered to a subject.	Cum Dose (mg/kg) is the sum of the doses administered to a subject.
Cycle Duration$_{(i)}$ (wk)	(Dose date$_{(i+1)}$ - Dose date$_{(i)}$)/7	N/A
Cycle Intensity$_{(i)}$ (mg/kg/wk)	Dose$_{(i)}$/Cycle Duration$_{(i)}$	N/A
Relative Cycle Intensity $_{(i)}$ (%)	(Cycle Intensity$_{(i)}$/intended dose per week) * 100	N/A
Relative Dose Intensity (%)	Sum of all Relative Cycle Intensities divided by N	Cum dose /[(Last dose date - Start dose date + 21) x 3/21] x 100
Duration of Treatment	*Last dose date - Start dose date +1*	*Last dose date - Start dose date +1*

7.4.2 Modifications of Study Therapy

7.4.2.1 Dose delays

Nivolumab, Ipilimumab treatment may be delayed for up to a maximum of 6 weeks from the last dose. A dose will be considered as actually delayed if the delay is exceeding 3 days (ie, greater than or equal to 4 days from scheduled dosing date) for nivolumab, ipilimumab and bevacizumab. All studies drugs must be delayed until treatment can resume. Length of delay for Nivolumab is defined as (duration of previous cycle in days - 14) for Arm N3, Arm B, Arm N1+I3/N3+I1(week 13 and following), Arm N3+RT+TMZ, ARM N3+RT and (duration of previous cycle in days - 21) for Arm N1+I3/N3+I1(week 1 - 12)(for Nivolumab and Ipilimumab). Dose delays will be divided into following categories: 4 - < 8 days, 8 - < 15 days, 15 - < 42, ≥ 42 days. Reason for dose delay/omission will be retrieved from CRF dosing pages.

The following parameters will be summarized by cohort and treatment group.

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

7.4.2.2 Infusion Interruptions and Rate Changes

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

The following parameters will be summarized by cohort and treatment group:

- 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

[redacted]

7.4.2.4 Dose Reductions/Escalation

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

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 by cohort and treatment group 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, the primary population will consist in the all randomized subjects in cohort 2 and the analysis will be performed by treatment group as “randomized”.

7.5.1 Overall Survival for Cohort 2

7.5.1.1 Primary Analysis

The OS curves for each treatment group will be estimated using the Kaplan-Meier (KM) product-limit method. Median OS and the corresponding two-sided 95% confidence intervals using the log-log transformation will be computed.

The distribution of OS will be compared in two randomized arms via a two-sided 5% stratified log-rank test.

In addition, a stratified Cox proportional hazards regression model will be used to estimate hazard ratio (see Section 7.1) between treatment groups along with the 95 % CI.

Survival rates at 6, 9, 12, and 18 months will be estimated using KM estimates on the OS curve for each randomized arm. Minimum follow-up must be longer than the time point to generate the rate. Associated two-sided 95% CIs will be calculated.

The status of subjects who are censored in the OS Kaplan-Meier 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, ...).

To examine the assumption of proportional hazards in the Cox regression model, in addition to treatment, a time-dependent variable defined by treatment by time interaction will be added into the model. A two-sided Wald Chi-Square p-values of less than 0.1 will indicate a potential nonconstant treatment effect. In that case, additional exploratory analyses may be performed.

7.5.1.2 OS sensitivity Analyses

The following OS sensitivity analyses will be performed:

1) OS will be compared between treatment groups using a two-sided 5% unstratified log-rank test.
2) OS will be compared between the treatment groups using the strata as determined at baseline (CRF source). This analysis will be performed only if stratification variable at IVRS and at baseline disagree for at least 10% of the randomized subjects.
3) OS will be compared between the treatment groups using a two-sided 5% stratified log-rank test in All treated subject population, using arm, as randomized. This analysis will be performed only if the proportion of randomized but never treated subjects exceeds 5%.
4) A multivariate Cox regression model will be used to estimate the treatment effect after including the following covariates measured at baseline: Backward selection method will be used to eliminate non-significant covariates at level 0.15.
   a) Age (continuous covariate)
   b) Steroid Use (Yes, No)
   c) Performance Status (Karnofsky scale) (≤ 80, > 80)
   d) Time from GBM diagnosis to recurrence(continuous covariate)

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

Estimate of the hazard ratio, its two-sided 95% CI.

7.5.1.3 Consistency of Treatment Effect on OS in Subsets

To assess consistency of treatment effects in different subsets, a "forest" plot of the OS hazard ratio (and 95% CI) will be produced for the following variables, but not limited to:

- Baseline measurable lesion (yes vs. no) (source: CRF)
- Region (US/Canada vs. W. Europe vs. Rest of World)
- Age categorization (< 65, ≥ 65 - < 75, ≥ 75, ≥ 65)
- Age categorization (< 50, ≥ 50 - <65)
- Gender (Male vs. Female)
- Race (White, African American, Asian, Other)
- Smoking status (yes vs. no, unknown)
- Baseline Performance Status (Karnofsky scale) (≤ 80, > 80)
- Baseline Pathology (Glioblastoma vs. Gliosarcoma)
- Prior I/O Vaccine(Yes/no)
- MGMT status at baseline (Methylated, unmethylated, Unknown)
- Prior Corticosteroid use (No, Yes)
- Time from GBM diagnosis to recurrence (≤ 12 months, >12 months)

If a subgroup category has less than 10 subjects in a treatment group, then HR will not be reported for that subgroup.

7.5.1.4 Subject Follow-Up

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) for cohort 2. For cohorts 1, 1b, [redacted] it will be defined as the time between dose start date and last known date alive (for subjects who are alive) or death date (for subjects who died). It will be summarized descriptively (median, min, max) for all randomized subjects.

The currentness of follow-up, defined as the time between last OS contact (ie, last known date alive or death date) and data cut-off date, will be summarized by treatment group. Subjects who died and subjects with a Last Known Date Alive on or after LPLV will have a value of '0' for currentness of follow-up. The currentness of follow-up will be categorized into the following categories: 0 days, 1-30 days, 31-60 days, 61-90 days, 91-120 days, 120-150 days, 151 or more days.

7.5.1.5 Subsequent Therapy

Subsequent therapy will be summarized by treatment group and listed.

- Subsequent Therapy
  - Systemic anti-cancer therapy by drug name
  - Surgery

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

7.5.2 Interim Analysis for Cohort 2

There will be no interim analysis for primary endpoint of OS for cohort 2.

7.5.3 Overall Survival Rate at 12 Month for Cohort 2

7.5.3.1 Primary Analysis

The OS(12) for each treatment group will be estimated using the Kaplan-Meier (KM) product-limit method and the corresponding two-sided 95% confidence intervals using log-log transformation will be computed. OS(12) will be compared between the two randomized arms via a two-sided (5% alpha level), Z test with variance estimation based on Greenwood formula[9] for variance derivation using log(-log) transformation. This comparison will only be tested if the OS comparison is positive.

7.5.4 Progression Free Survival for Cohort 2

7.5.4.1 Primary Analysis

The PFS curves for each treatment group will be estimated using the Kaplan-Meier (KM) product-limit method. Median PFS and the corresponding two-sided 95% confidence intervals using the log-log transformation will be computed.

PFS will be compared between the two randomized arms using two-sided (5% alpha level) stratified log-rank test. This comparison will only be tested if both of OS and OS(12) comparisons are positive

In addition, a stratified Cox proportional hazards regression model will be used to estimate hazard ratio (see Section 7.1) between treatment groups along with the 95% CI.

PFS rates at 6, 9, 12, and 18 months will be estimated using KM estimates on the PFS curve for each treatment group. Minimum follow-up must be longer than the time point to generate the rate. The associated two-sided 95% CI will also be calculated.

The source of progression (death vs. progression) will be summarized by treatment group.

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

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

7.5.4.2 Sensitivity Analysis for Cohort 2

Sensitivity analyses of PFS will also be performed using the following modification:

1) PFS accounting for assessment after subsequent therapy (PFS ITT). It will be defined similar to the primary definition except that events (progression or death) and tumor assessments that occurred after subsequent anticancer therapy or diagnostic surgical resection will be taken into account (see censoring scheme 1 for sensitivity analysis in Table 7.5.4.2-1).

Table 7.5.4.2-1: Censoring Scheme 1 for Sensitivity Analysis of PFS

Situation	Date of Progression of Censoring	Outcome
No baseline tumor assessment and no disease progression/death	Date of randomization	Censored
No on-study tumor assessments and no disease progression/death	Date of randomization	Censored
Documented progression per investigator assessment (including clinical progression)	Date of first documented progression per RANO	Progressed
No progression and no death	Date of last assessment for disease progression	Censored
New anticancer treatment started without a prior reported progression	Date of subsequent anti-cancer therapy not considered	None
Diagnostic surgical resection without documented progression prior or on the same day and without confirmed disease progression from tumor pathology	Date of diagnostic surgical resection not considered	None
Death after initiation of subsequent cancer therapy	Date of death	Progressed
Death without progression	Date of death	Progressed

2) PFS not considering clinical progression as event: It will be defined same as the primary definition except not considering clinical progression as progression event. Subjects with no baseline or on-study tumor assessment and no death will be censored on the date of randomization.
3) PFS (ITT) not counting clinical progression as event: It will be defined same as PFS ITT defined above except not considering clinical progression as progression event. Subjects with no baseline or on-study tumor assessment and no death will be censored on the date of randomization.

7.5.5 Objective Response Rate for Cohort 2

The population will consist in the response-evaluable subjects.

7.5.5.1 Primary Analysis

ORR estimates and corresponding 95% CIs will be provided by treatment group. BOR will be tabulated for each treatment group.

Comparison of ORR between the treatment groups will be conducted using stratified Cochran-Mantel-Haenszel (CMH) test (two-sided, 5% alpha level). This test will only be performed if all of the OS, OS(12), and PFS comparisons are positive.

An estimate of the difference in ORR between arms and a corresponding two sided, 95% CI for will also be computed using the following Cochran-Mantel-Haenszel (CMH) method of weighting, adjusting for the stratification factor.[10] 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.

7.5.5.2 Further Characterization of ORR

DOR will be summarized for subjects who achieve confirmed PR or CR using the Kaplan-Meier (KM) product-limit method for each treatment group. Median DOR, corresponding two-sided 95% CIs and range will also be calculated. In addition, the percentage of responders still in response at different time points (3, 6, 9 and 12 months) will be presented based on the KM plot. Minimum follow-up must be longer than the timepoint to generate the rate.

The magnitude of reduction in tumor burden will be summarized descriptively.

7.5.6 Other Efficacy Analyses for Cohort 2

The following subject-level graphics will also be provided by treatment group as randomized:

- For all responders, time courses of the following events of interest will be graphically displayed: tumor response, progression, last dose received, and death.
- For response evaluable subjects, a waterfall plot showing the best reduction in target lesion based will be produced.
- For response-evaluable subjects, a plot of individual subjects' percent change in target lesion tumor burden from baseline.

7.5.7 Efficacy Analyses for Cohorts 1, 1b, [redacted]

The following analyses maybe performed if sufficient data from cohorts 1, 1b, [redacted] are available. Unless otherwise specified, the population will consist in the all treated subjects in cohorts 1, 1b,[redacted]. The analysis will be performed by treatment group as "treated".

OS curves for each treatment group will be estimated using the KM product limit method. Median OS and corresponding two-sided, 95% confidence intervals will be computed.

PFS curves for each treatment group will be estimated using the KM product-limit method. Median PFS and corresponding two-sided, 95% confidence intervals will be computed.

Estimates of ORR and corresponding 95% CIs will be provided by treatment group. BOR will be tabulated for each treatment group. This analysis will be performed only on the response-evaluable subject population.

DOR will be summarized for subjects who achieve confirmed PR or CR using the Kaplan-Meier (KM) product-limit method for each treatment group. Median DOR, corresponding two-sided 95% CIs and range will also be calculated. This analysis will be performed only on subjects with PR or CR.

Durability of stable disease will be described.

For response-evaluable subjects, a plot of individual subjects' percent change in target lesion tumor burden from baseline.

7.6 Safety

For all safety related analyses, refer to the Core Safety SAP[2]. Safety will be summarized for: all treated subjects, by cohort and treatment.

7.6.1 Deaths

See Core Safety SAP[2].

7.6.2 Serious Adverse Events

See Core Safety SAP[2].

7.6.3 Adverse Events Leading to Discontinuation of Study Therapy

See Core Safety SAP[2].

7.6.4 Adverse Events Leading to Dose Delay of Study Therapy

See Core Safety SAP[2].

7.6.5 Adverse Events

See Core Safety SAP[2].

7.6.6 Select Adverse Events

See Core Safety SAP[2].

7.6.7 Immune-Mediated Adverse Events

- The immune-mediated adverse events (IMAE) consist of a list of preferred terms grouped by specific category (eg, pneumonitis, diarrhea/colitis, see Core Safety SAP for select adverse events). The IMAE categories and terms are defined by the Sponsor and the list that is the most current at the time of analysis will be used. This list will be based on the most current version of MedDRA at the time when integrated database will be built. The final list will be included as appendix in CSR.

Table 7.6.7-1: Immune-Mediated Adverse Events

Class Name of IMAE	PT TERM	Order in the output
PNEUMONITIS	PNEUMONITIS Interstitial lung disease	1
DIARRHEA/COLITIS	COLITIS DIARRHEA ENTEROCOLITIS	2
HEPATITIS	HEPATOTOXICITY HEPATITIS HEPATITIS ACUTE AUTOIMMUNE HEPATITIS ASPARTATE AMINOTRANSFERASE INCREASED ALANINE AMINOTRANSFERASE INCREASED BLOOD BILIRUBIN INCREASED HYPERBILIRUBINAEMIA BLOOD ALKALINE PHOSPHATASE INCREASED	3
ADRENAL INSUFFICIENCY	ADRENAL INSUFFICIENCY	4
HYPOTHYROIDISM/THYROIDITIS	HYPOTHYROIDISM THYROIDITIS ACUTE AUTOIMMUNE THYROIDITIS THYROIDITIS	5
HYPOTHYROIDISM	HYPOTHYROIDISM	6
THYROIDITIS	AUTOIMMUNE THYROIDITIS THYROIDITIS THYROIDITIS ACUTE	7
HYPERTHYROIDISM	HYPERTHYROIDISM	12
HYPOPHYSITIS	HYPOPHYSITIS	13
DIABETES MELLITUS	DIABETES MELLITUS DIABETIC KETOACIDOSIS	8
NEPHRITIS AND RENAL	NEPHRITIS	9

Table 7.6.7-1: Immune-Mediated Adverse Events

Class Name of IMAE	PT TERM	Order in the output
DYSFUNCTION	NEPHRITIS ALLERGIC	
	TUBULOINTERSTITIAL NEPHRITIS	
	RENAL FAILURE ACUTE	
	RENAL FAILURE	
	BLOOD CREATININE INCREASED	
RASH	RASH	10
	Rash maculo-papular	
HYPERSENSITIVITY	HYPERSENSITIVITY	11
	INFUSION RELATED REACTION	
OTHER	Myasthenic Syndrome	14
	Myasthenia Gravis	
	Myasthenia Gravis Crisis	
	Myasthenic Syndrome	
	Demyelination Event	
	Demyelination	
	Guillain-Barre Syndrome	
	Guillain-Barre Syndrome	
	Miller Fisher Syndrome	
	Pancreatitis Event	
	Autoimmune Pancreatitis	
	Pancreatitis	
	Pancreatitis Acute	
	Pancreatitis Necrotising	
	Uveitis Event	
	Chorioretinitis	
	Cyclitis	
	Intermediate uveitis	
	Iridocyclitis	
	Iritis	
	Uveitis	

- All summaries in the following subsections for IMAEs based on extended follow-up (100 day window after the last dose).

7.6.7.1 Incidence of IMAEs

Immune-Mediated AEs by worst CTC grade (both classifications, that are, "Any Grade, Grade 3-4, Grade 5" and "Worst CTC Grade", unless noted otherwise) will be summarized by Category/PT for each category:

- Summary of any endocrine IMAE
- Summary of IMAE where immune modulating medication was initiated
- Summary of IMAEs [only by Worst CTC Grade]
- Summary of endocrine serious IMAEs
- Summary of serious IMAEs where immune modulating medication was initiated
- Summary of any endocrine IMAEs leading to dose delay or reduction
- Summary of IMAEs leading to dose delay or reduction where immune modulating medication was initiated
- Summary of endocrine IMAEs leading to discontinuation
- Summary of IMAEs leading to discontinuation where immune modulating medication was initiated

7.6.7.2 Management of IMAEs

Following summaries will be presented

- Summary of duration of immune modulating concomitant medication for IMAE management
- Summary of immune modulating concomitant medication for IMAE management
- Summary of immune modulating concomitant medication for Grade 3 to 5 IMAEs management

7.6.7.3 Time to Onset and Resolution of IMAEs

Time to onset and resolution for IMAEs will be analyzed. Following summaries will be presented

- Summary of time to onset of endocrine IMAE
- Summary of time to onset of IMAEs where immune modulating medication was initiated
- Summary of time to resolution of endocrine IMAEs
- Summary of time to resolution of endocrine IMAEs where immune modulating medication was initiated

7.6.7.4 Re-Challenge

Subjects who had infusion of nivolumab after the onset of IMAE were considered re-challenged subjects. Following summary will be presented for re-challenged subjects.

- Summary of subjects who were re-challenged with nivolumab by IMAE category

Listings of IMAEs will also be provided

7.6.8 Immune Modulating Medication

See Core Safety SAP[2].

7.6.9 Multiple Events

See Core Safety SAP[2].

7.6.10 Clinical Laboratory Evaluations

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

7.6.10.1 Hematology

See Core Safety SAP[2].

7.6.10.2 Serum Chemistry

See Core Safety SAP[2].

[redacted]

7.6.12 Vital Signs and Pulse Oximetry

See Core Safety SAP[2].

7.7 Pregnancy

See Core Safety SAP[2].

[redacted]

7.9 Biomarkers

Analyses for PD-L1 are described below. Analyses of exploratory biomarkers other than PD-L1 expression will be documented in a separate statistical analysis plan and may be reported external to the clinical study report. Following analyses will be presented for cohort 2. Summaries will be reported for Cohorts 1, 1b, [redacted] if there is sufficient data.

7.9.1 PD-L1 Expression

Analyses of efficacy outcomes as a function of PD-L1 expression are descriptive in nature and intended to examine the distribution of PD-L1 expression and assess potential associations between PD-L1 expression and efficacy measures.

PD-L1 expression in this section will be defined based on validated Dako PD-L1 IHC assay.

The following PD-L1 expression subgroups will be considered.

- Each baseline quantifiable PD-L1 expression status subgroup:
  - PD-L1 ≥ X%
  - PD-L1 < X%

Baseline PD-L1 expression not evaluable or indeterminate subgroup.

Where X denotes PD-L1 expression cut-off, which will be specified at later point in time.

Analyses of PD-L1 will include:

1) Examine the distribution of PD-L1 expression
2) Assess potential association between PD-L1 status and efficacy measures

7.9.2 Analysis Methods

Analyses of PD-L1 will include:

- Examine the distribution of PD-L1 expression
- Assess potential associations between PD-L1 expression and efficacy measures
- Assess potential associations between PD-L1 expression using different cut offs and select AEs.

The following analyses will be performed:

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 by cohort and treatment group, all tested PD-L1 subjects.
- Summary of tumor specimen acquisition and characteristics by cohort and treatment group, all randomized subjects for cohort 2 and all treated subjects for cohorts1, 1b, [redacted]
- Cumulative distribution plot of PD-L1 expression at baseline versus population percentile by treatment group and overall, all evaluable PD-L1 subjects. This analysis will be performed only for cohort 2.
- 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. This analysis will be performed only for cohort 2.
- Box plot of PD-L1 expression by treatment group and overall, all randomized subjects in cohort 2

7.9.2.2 Association Between PD-L1 Status and Efficacy Measures

Analyses will be performed using all randomized PD-L1 subjects in Cohort 2 if not otherwise specified. Each analysis will be performed for the subgroups listed below if not otherwise specified.

- Each PD-L1 status subgroup
- PD-L1 indeterminate or not evaluable subgroup

Analyses for OS/PFS endpoint (for Cohort 2 only):

A separate Cox proportional hazards regression model will be fitted for OS and PFS with PD-L1+ expression level as a sole covariate among all subjects treated with Nivolumab monotherapy and Bevacizumab, 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 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. Hazard Ratios with 95% CIs will also be reported

Analyses for ORR endpoint (for Cohort 2 only):

A separate logistic regression model will be fitted for ORR with PD-L1+ expression level as a sole covariate for subjects treated with Nivolumab and Bevacizumab, 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 Nivolumab treated subjects:
  - A plot of the ROC curve
  - A plot of estimated true positive fraction and false positive fraction vs. PD-L1+ expression (X-axis)

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

Analyses will be performed using all evaluable PD-L1 subjects in Cohort 2 if not otherwise specified.

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 vs. bevacizumab and its associated 95% CI for each of the 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 of Select AEs and PD-L1 Expression

Select adverse events will be summarized by worst CTC Grade and baseline PD-L1 Status (cutoff at 5%) for each cohort and treatment group using all treated PD-L1 subjects and the PD-L1 indeterminate or not evaluable 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[11]. 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[12].

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 (eg, 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.

10 DOCUMENT HISTORY

Table 10-1: Document History

Version Number	Author(s)	Description
1.0	[redacted]	Initial version dated 11-Aug-2014
2.0	[redacted]	• Revised SAP according to protocol amendments 05, 06 and 07. • [redacted] • Increased the sample size for cohort 2 by 120 subjects to 340. • Changed from two planned interim analyses to single interim analysis at 80% (215) of the total number of events. • [redacted] • Added analyses for PFS and ORR endpoints based on IRRC assessment as sensitivity analyses. • Added sensitivity analyses for OS (section 7.5.1.2) and analysis to examine the assumption of proportional hazards • Added prior I/O vaccine to the OS subset analyses. • Updated Section 7.5.2 to make it consistent with the data being provided to DMC for benefit/risk assessment. • Added section 7.6.7 to describe analyses for immune-mediated adverse events • Updated Section 7.9 to include flexibility for using appropriate PD-L1 expression cut-off value for analyzing PD-L1 data. Also biomarker analyses for safety cohorts (1, 1b, [redacted]) will be presented if there is sufficient data. .
3.0	[redacted]	• Revised SAP as per protocol amendments 08 and 09. This included updating section 5 (sample size and power) to modify timing of IA and FA. [redacted] • PFS Endpoint: o Censoring for primary analyses for PFS was updated to censor at randomization date in case of no baseline or tumor assessment only if there is no progression and death date. o Sensitivity analyses were added to understand the impact of clinical progression. Two sensitivity analyses were added that do not consider clinical progression as a PFS event. o Analyses based on IRRC assessment was removed since independent review is not planned. o Table 4.2.1.2-2 was updated for clarification for the case

Table 10-1: Document History

Version Number	Author(s)	Description
		of death after subsequent therapy and progression before diagnostic resection. o Table 7.5.4.2-1 was updated to be consistent with table 4.2.1.2-2 and for clarity. • ORR Endpoint: o Analysis based on IRRC assessment was removed since independent review is no longer planned o Method to estimate difference in ORR is changed to using Cochran-Mantel-Haenszel (CMH) method of weighting to adjust for stratification factor. [redacted] • Clarified definition for all tested PD-L1 subjects for cohorts 1, 1b, [redacted] and 2. [redacted] • Relevant protocol deviation of "subjects without documented first recurrence of GBM within 21 days of randomization" has changed to "subjects without documented first recurrence of GBM" without any window since that was cleaning more meaningful. Other minor updates were made for clarification. On-study deviation of "subjects treated differently than as assigned " was updated to be defined only for cohort 2 , not cohort 1. • Demographics and baseline characteristics: o added baseline corticosteroid use summaries o Updated categories for "Time from Initial Disease Diagnosis to Recurrent Disease Diagnosis" and smoking status. • Extent of Exposure: [redacted] o Updated table 7.4.1-1 to be consistent with the way data is collected for Nivolumab and Ipilimumab • OS Sensitivity analyses: o Added multivariate analysis to estimate the treatment effect adjusting for prognostic factors. o Added following factors for assessing the consistency of treatment effects in different subsets for OS: MGMT status, prior corticosteroid use and time from GBM diagnosis to recurrence. Updated the categories for baseline performance status. • Clarification for the subsequent therapy summaries: removed the categories that were not clinically relevant and added that summaries will be provided by treatment group.

Table 10-1: Document History

Version Number	Author(s)	Description
		• PD-L1 Expression Analysis methods: o Added boxplots of PD_l1 expression by treatment group for cohort 2 o Added analyses understand relationship between for PFS and OS endpoint and PD-L1 expression, and ROC analysis with ORR for PD-L1 expression o Removed analysis for predictive relationship of PD-L1 status and efficacy measures. • Imputation algorithm for death dates has been updated to be consistent with other Nivolumab studies. • EQ-5D: Included summary for completion rates.
4.0	[redacted]	• Revised SAP as per protocol amendment 10. The main purpose of the amendment is to cancel the interim analysis. Revised section 5 (sample size and power) and section 7.5 to remove references to interim analysis.

11 REFERENCES

1 Data Monitoring Committee Charter. Study CA209143. BMS Document 930076684.1.

2 Core Safety Statistical Analysis Plan for CA209, Bristol-Myers Squibb

3 Aaronson N.K., Ahmedzai S., et al. The European Organisation for Research and Treatment of Cancer QLQ-30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst, 85: 365-376, 1993.

4 http://www.euroqol.org/about-eq-5d/publications/user-guide.html

5 The EuroQol Group: EuroQol: A new facility for the measurement of health-related quality of life—The EuroQol Group. Health Policy 16:199-208, 1990.

6 Reardon DA, Li G, Recht L, et al. ReACT: Overall survival from a randomized phase II study of rindopepimut (CDX-110) plus bevacizumab in relapsed glioblastoma. J Clin Oncol 33, 2015 (suppl; abstr 2009).

7 Kalbfleisch RL. The Statistical Analysis of Failure Time Data. Wiley-Interscience, 2002.

8 Clopper, CJ and Pearson, ES. The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika 26: 404-423, 1934.

9 Greenwood, M. The errors of sampling of the survivorship tables, Reports on Public Health and Statistical Subjects, 33, Appendix 1, HMSO, London, 1926

10 Statistical Methodology in the Pharmaceutical Sciences/ edited by Berry DA, Chapter 13, Categorical Data Analysis, p 415 and 417 ff., Marcel Dekker, 1990

11 Adverse Event Domain Requirements Specification. Bristol-Myers Squibb Co. PRI. Version 2.1. April 23, 2012.

12 Non-Study Medication Domain Requirements Specification. Bristol-Myers Squibb Co. PRI. Version 2.2 April 24, 2012.