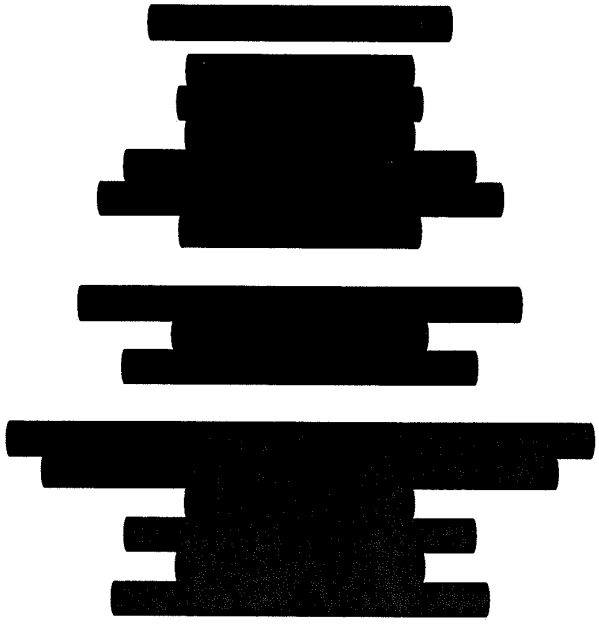
Protocol Number: CA209214
IND Number: 122,840
Ex-US Non-IND
EUDRACT Number 2014-001750-42
Date: 17-Jul-2014
Revised Date: 13-Nov-2017

Clinical Protocol CA209214

A Phase 3, Randomized, Open-Label Study of Nivolumab Combined with Ipilimumab Versus Sunitinib Monotherapy in Subjects with Previously Untreated, Advanced or Metastatic Renal Cell Carcinoma

(**CheckMate 214, CHECK**point pathway and nivolu**MA**b clinical **T**rial **E**valuation 214)

Revised Protocol Number: 03
Incorporates amendment(s): 14 and Administrative Letter 01



Motzer RJ, *et al. J Immunother Cancer* 2021; 8:e000891. doi: 10.1136/jitc-2020-000891

Replace all previous version(s) of the protocol with this revised protocol and please provide a copy of this revised protocol to all study personnel under your supervision, and archive the previous versions.

DOCUMENT HISTORY

Document	Date of Issue	Summary of Change
Revised Protocol 03	13-Nov-2017	Incorporates Amendment 14
Amendment 14	13-Nov-2017	Protocol amendment is being implemented to provide modifications to the protocol based on recommendations of the study's independent Data Monitoring Committee (DMC) after their review of the planned interim analysis of overall survival (OS), which met the pre-specified boundary for statistical significance for the co-primary endpoint of OS. As a result of the DMC assessment, this protocol amendment is being implemented to provide a mechanism for eligible subjects randomized to sunitinib treatment (Arm B) to receive nivolumab combined with ipilimumab therapy in a crossover extension phase. This amendment also provides the options for Arm A subjects to: 1) switch to a flat dose of nivolumab at 240 mg every 2 weeks if they are currently receiving nivolumab 3 mg/kg every 2 weeks and 2) discontinue treatment after 2 years even in the absence of disease progression or unacceptable toxicity. Protocol amendment also indicates that the interim analysis results should be considered the final primary analysis results of the protocol.
Administrative Letter 01	13-Feb-2017	Updated Medical Monitor and removed Study Director
Revised Protocol 02	04-Aug-2016	Incorporates Amendment 13
Amendment 13	04-Aug-2016	Added Objective Response Rate (ORR) as an additional co-Primary Endpoint. Included required updates based on Version 15 of the Nivolumab Investigator Brochure. Added language that allows for collection of additional survival data outside the original protocol specified visit windows. Added Study Director.
Revised Protocol 01	05-Nov-2014	Incorporates Amendment 04
Amendment 04	05-Nov-2014	Added an additional secondary objective related to incidence of AEs. Updated the IMDC prognostic factor for corrected calcium criteria. Added additional LFT testing for Arm A subjects. Incorporated minor changes to correct and/or maintain consistency throughout the protocol.
Original Protocol	17-Jul-2014	Not applicable

SYNOPSIS

Clinical Protocol CA209214

Protocol Title: A Phase 3, Randomized, Open-Label Study of Nivolumab Combined with Ipilimumab Versus Sunitinib Monotherapy in Subjects with Previously Untreated, Advanced or Metastatic Renal Cell Carcinoma

(CheckMate 214, CHECKpoint pathway and nivolu**MA**b clinical **T**rial **E**valuation 214**)**

Investigational Product(s), Dose and Mode of Administration, Duration of Treatment with Investigational Product(s):

- Nivolumab administered IV over 60 minutes at 3 mg/kg combined with ipilimumab administered IV over 30 minutes at 1 mg/kg every 3 weeks for 4 doses followed by nivolumab administered IV over 60 minutes at 3 mg/kg every 2 weeks or sunitinib 50 mg po Day 1 - 28 of each 42 day cycle until disease progression, unacceptable toxicity or other reasons specified in the protocol. Under Amendment 14, subjects receiving nivolumab at 3mg/kg every 2 weeks will have the option to switch to intravenous nivolumab dosing over 60 minutes at 240 mg every 2 weeks until disease progression, unacceptable toxicity or other discontinuation criteria specified in the protocol.

Study Phase: 3

Research Hypothesis: Treatment with nivolumab combined with ipilimumab will improve Objective Response Rate (ORR), Progression Free Survival (PFS), and Overall Survival (OS) compared to sunitinib monotherapy in subjects with previously untreated metastatic renal cell carcinoma (mRCC).

Objectives:

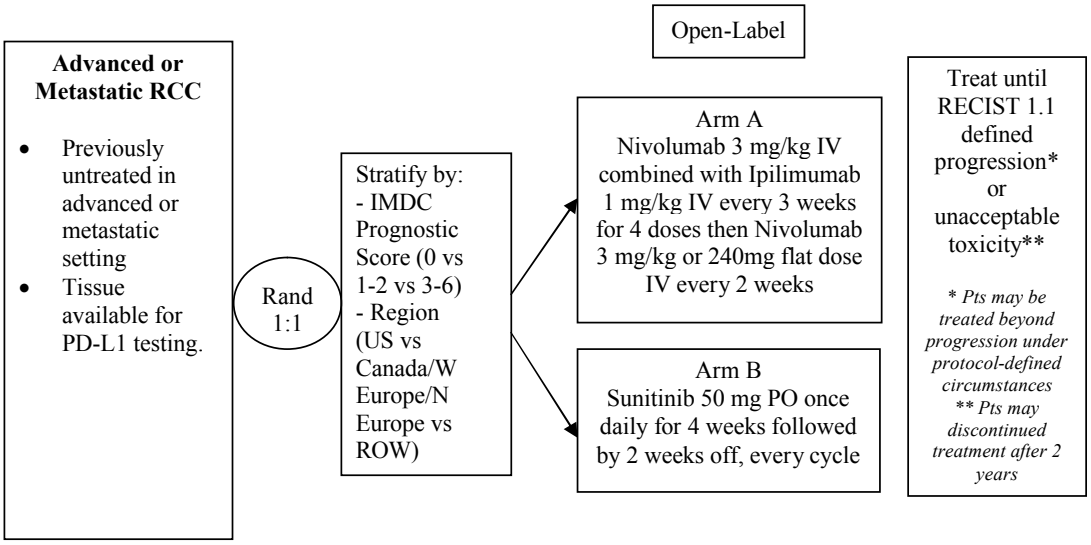
Primary Objectives

- To describe the ORR of nivolumab combined with ipilimumab and sunitinib monotherapy in intermediate and poor risk subjects with previously untreated mRCC based on IRRC assessments
- To compare the PFS of nivolumab combined with ipilimumab to sunitinib monotherapy in intermediate and poor-risk subjects with previously untreated mRCC, based on Independent Radiation Review Committee (IRRC) assessments
- To compare the OS of nivolumab combined with ipilimumab to sunitinib monotherapy in intermediate and poor-risk subjects with previously untreated mRCC

Key Secondary Objectives

- To compare the PFS of nivolumab combined with ipilimumab to sunitinib monotherapy in any-risk subjects with previously untreated mRCC, based on IRRC assessments
- To compare the OS of nivolumab combined with ipilimumab to sunitinib monotherapy in any-risk subjects with previously untreated mRCC
- To estimate the objective response rate (ORR) of nivolumab combined with ipilimumab and sunitinib monotherapy in subjects with previously untreated mRCC (any-risk), based on IRRC assessments
- To estimate the incidence of AEs of nivolumab combined with ipilimumab and sunitinib monotherapy in all treated subjects with previously untreated mRCC

Figure -1: Study Design:



Study Population:

Key Inclusion Criteria:

- Histological confirmation of RCC with a clear-cell component
- Advanced (not amenable to curative surgery or radiation therapy) or metastatic (AJCC Stage IV) RCC
- No prior systemic therapy for RCC with the following exception:
  a) One prior adjuvant or neoadjuvant therapy for completely resectable RCC if such therapy did not include an agent that targets VEGF or VEGF receptors and if recurrence occurred at least 6 months after the last dose of adjuvant or neoadjuvant therapy.
- Karnofsky Performance Status (KPS) of at least 70%
- Measurable disease as per RECIST 1.1
- Tumor tissue (formalin-fixed paraffin-embedded (FFPE) archival or recent acquisition) must be received by the central vendor (block or unstained slides) in order to randomize a subject to study treatment. (Note: Fine Needle Aspiration [FNA] and bone metastases samples are not acceptable for submission).
- Patients with favorable, intermediate and poor risk categories will be eligible for the study. Patients must be categorized according to favorable versus intermediate/poor risk status at registration.

  To be eligible for the Intermediate and Poor-Risk cohort, at least one of the following prognostic factors as per International Metastatic RCC Database Consortium (IMDC) must be present:

  a) KPS equal to 70
  b) Less than 1 year from diagnosis to randomization
  c) Hemoglobin less than the LLN
  d) Corrected calcium concentration greater than 10 mg/dL
  e) Absolute neutrophil count greater than the ULN
  f) Platelet count greater than the ULN

  If none of the above factors are present, subjects are only eligible for the favorable-risk cohort. The favorable-risk cohort may close to enrollment earlier than the intermediate- or poor-risk cohort.

Key Exclusion Criteria:

- Any history of or current CNS metastases. Baseline imaging of the brain is required within 28 days prior to randomization.
- Prior systemic treatment with VEGF or VEGF receptor targeted therapy (including, but not limited to, sunitinib, pazopanib, axitinib, tivozanib, and bevacizumab).
- Prior treatment with an anti-PD-1, anti-PD-L1, anti-PD-L2, anti-CD137, or anti-CTLA-4 antibody, or any other antibody or drug specifically targeting T-cell co-stimulation or checkpoint pathways.
- Any active or recent history of a known or suspected autoimmune disease or recent history of a syndrome that required systemic corticosteroids (> 10 mg daily prednisone equivalent) or immunosuppressive medications except for syndromes which would not be expected to recur in the absence of an external trigger. Subjects with vitiligo or type I diabetes mellitus or residual hypothyroidism due to autoimmune thyroiditis only requiring hormone replacement are permitted to enroll.
- Any condition requiring systemic treatment with corticosteroids (> 10 mg daily prednisone equivalents) or other immunosuppressive medications within 14 days prior to first dose of study drug. Inhaled steroids and adrenal replacement steroid doses > 10 mg daily prednisone equivalents are permitted in the absence of active autoimmune disease.
- Uncontrolled adrenal insufficiency.
- Ongoing symptomatic cardiac dysrhythmias, uncontrolled atrial fibrillation, or prolongation of the Fridericia corrected QT (QTcF) interval defined as > 450 msec for males and > 470 msec for females, where $QTcF = QT / \sqrt[3]{RR}$
- Poorly controlled hypertension (defined as systolic blood pressure (SBP) of ≥ 150 mmHg or diastolic blood pressure (DBP) of ≥ 90 mmHg), despite antihypertensive therapy
- History of any of the following cardiovascular conditions within 12 months of enrollment: cardiac angioplasty or stenting, myocardial infarction, unstable angina, coronary artery by-pass graft surgery, symptomatic peripheral vascular disease, class III or IV congestive heart failure, as defined by the New York Heart Association
- History of cerebrovascular accident including transient ischemic attack within the past 12 months
- History of deep vein thrombosis (DVT) unless adequately treated with low molecular weight heparin.
- History of pulmonary embolism within the past 6 months unless stable, asymptomatic, and treated with low molecular weight heparin for at least 6 weeks.
- History of abdominal fistula, gastrointestinal perforation, or intra-abdominal abscess within the past 6 months.
- Serious, non-healing wound or ulcer.
- Evidence of active bleeding or bleeding susceptibility; or medically significant hemorrhage within prior 30 days.
- Any requirement for anti-coagulation, except for low molecular weight heparin.
- Prior malignancy active within the previous 3 years except for locally curable cancers that have been apparently cured, such as basal or squamous cell skin cancer, superficial bladder cancer, or carcinoma in situ of the prostate, cervix, or breast.
- Known history of testing positive for human immunodeficiency virus (HIV) or known acquired immunodeficiency syndrome (AIDS).
- Any positive test for hepatitis B or hepatitis C virus indicating acute or chronic infection.
- Known medical condition (eg, a condition associated with diarrhea or acute diverticulitis) that, in the investigator's opinion, would increase the risk associated with study participation or study drug administration or interfere with the interpretation of safety results.
- Major surgery (eg, nephrectomy) less than 28 days prior to the first dose of study drug.
- Anti-cancer therapy less than 28 days prior to the first dose of study drug or palliative, focal radiation therapy less than 14 days prior to the first dose of study drug.

- Presence of any toxicities attributed to prior anti-cancer therapy, other than alopecia, that have not resolved to Grade 1 (NCI CTCAE v4) or baseline before administration of study drug.
- Receiving concomitant CYP3A4 inducers or strong CYP3A4 inhibitors (See Appendix 4).
- Impairment of gastrointestinal function or gastrointestinal disease that may significantly alter the absorption of sunitinib (eg, malabsorptive disorder, ulcerative disease, uncontrolled nausea, vomiting, diarrhea, or small bowel resection).
- Left ventricular ejection fraction (LVEF) less than the LLN as assessed by echocardiography or multigated acquisition (MUGA) scan.
- Any of the following laboratory test findings:
  a) WBC < 2,000/mm^3
  b) Neutrophils < 1,500/mm^3
  c) Platelets < 100,000/mm^3
  d) AST or ALT > 3 x ULN (> 5 x ULN if liver metastases are present)
  e) Total Bilirubin > 1.5 x ULN (except subjects with Gilbert Syndrome, who can have total bilirubin < 3.0 mg/dL)
  f) Serum creatinine > 1.5 x upper limit of normal (ULN) or creatinine clearance < 40 mL/min (measured or calculated by Cockroft-Gault formula)

Amendment 14 Update:

Specific eligibility criteria for subjects in the poor or intermediate cohorts originally randomized to the sunitinib Arm B and now entering the nivolumab combined with ipilimumab crossover extension phase are included in the protocol in Section 3.1.1.

Study Drug: includes both Investigational [Medicinal] Products (IP/IMP) and Non-investigational [Medicinal] Products (Non-IP/Non-IMP) as listed:

Study Drug for CA209214		
Medication	**Potency**	**IP/Non-IP**
Nivolumab	3 mg/kg	IP
Ipilimumab	1 mg/kg	IP
Sunitinib	50 mg	IP

Study Assessments: Objective Response Rate, Overall Survival, and Progression Free Survival are the co-primary endpoints of the study. Subjects will be assessed for response by CT or MRI beginning at 12 weeks (± 1 week) after randomization and continuing every 6 weeks (± 1 week) for the first 13 months and then every 12 weeks (± 1 week) until progression or treatment discontinuation, whichever occurs later. Overall survival is defined as the time from randomization to the date of death.

Amendment 14 Update:

The schedule of assessments for subjects in the poor or intermediate cohorts originally randomized to the sunitinib Arm B and now entering the nivolumab combined with ipilimumab crossover extension phase are included in the protocol in Section 5.1.

Statistical Considerations:

Sample Size: The sample size of the study accounts for the three co-primary efficacy endpoints: ORR, based on IRRC assessments, PFS, based on IRRC assessments and OS, evaluated in intermediate and poor-risk subjects with previously untreated mRCC. The overall alpha for this study is 0.05, which is split with 0.001 to evaluate ORR, 0.009 to evaluate PFS and 0.04 to evaluate OS.

ORR will be analyzed initially on a descriptive basis and will occupy an administrative adjustment of alpha of 0.001. PFS will be evaluated for treatment effect at an alpha of 0.009 (two-sided, penalized 0.001 from a 0.01 allocation) with at least 90% power; no interim analysis of PFS is planned. OS will be evaluated for treatment effect at an alpha level of 0.04 (two-sided) with 90% power, accounting for two formal interim analyses to assess efficacy.

It is estimated that approximately 1070 previously untreated mRCC subjects will be randomized in a 1:1 ratio. Among them, approximately 820 subjects (76.6%) with intermediate/poor risk subjects and approximately 250 (23.4%) subjects with favorable risk as per IMDC (IMDC prognostic score = 0) will be randomized. Assuming a fixed accrual rate of 57 subjects per month (40 intermediate/poor risk subjects per month), it will take approximately 20.5 months to randomize 1070 subjects (820 intermediate/poor risk subjects).

Sample size justification for ORR estimate

One of the primary objectives of the study is to describe the ORR (based on IRRC assessment) of nivolumab combined with ipilimumab to sunitinib monotherapy in intermediate and poor-risk subjects with previously untreated mRCC.

The primary analysis of ORR in the intermediate and poor-risk randomized subjects will be performed when these patients have an approximate 6 month minimum follow-up from the completion of enrollment. This will allow sufficient follow-up for ORR to have a stable estimate, adequate safety follow-up as well as information on duration of response in this population.

The maximum width of the exact two-sided 95% confidence interval (CI) is 9.9% when the ORR is expected to be in the 20% to 50% range. Table 8.1-1 summarizes the 95% exact CI when observed ORRs are 20% to 50%, respectively.

For example if at least 123 responders are observed among the 410 nivolumab and ipilimumab combination intermediate/poor risk randomized subjects (ie, ORR ≥ 30%) then the lower bound of the 95% CI is above 25.6%.

Sample size justification for PFS comparison

One of the primary objectives of the study is to compare the progression-free survival (based on IRRC assessments) of nivolumab combined with ipilimumab to sunitinib monotherapy in intermediate and poor-risk subjects with previously untreated mRCC. The number of events and power for this study were calculated assuming an exponential distribution for PFS in each arm.

For this comparison of PFS, it will be required to observe at least 591 PFS events among the randomized intermediate/poor risk subjects in the two respective treatment arms for a two-sided experiment-wise $\alpha = 0.009$ log-rank test, to show a statistically significant difference in PFS between the treatment arms with at least 90% power when the true hazard ratio of the experimental arm to control arm is 0.73. The HR of 0.73 is equivalent to demonstrating a 37.8% improvement in median PFS, assuming a median PFS of 9 months in the sunitinib monotherapy arm (weighted median estimate assuming a median PFS of 11 months in intermediate risk subjects and a median PFS of 4 months in poor risk subjects) and a median PFS of 12.4 months in the experimental treatment arm.

Under the assumptions for accrual and PFS distribution stated above, it will take approximately 31 months from FPFV to observe the required number of PFS events for the final PFS analysis (20.5 months for accrual and 10.5 months for minimum follow up). It is projected that an observed HR of 0.807 or less corresponding to a 2.1 month or greater improvement in median PFS (9 vs 11.1 months) for this comparison, would result in a statistically significant improvement in the final analysis of PFS.

Sample size justification for OS comparison:

One of the primary objectives of the study is to compare the overall survival of nivolumab combined with ipilimumab to sunitinib monotherapy in intermediate and poor-risk subjects with previously untreated mRCC. The number of events and power of this study were calculated assuming an exponential distribution for OS in each arm.

Approximately 639 events (ie, deaths), observed among the randomized intermediate/poor risk subjects, provides 90% power to detect a hazard ratio (HR) of 0.766 with an overall type 1 error of 0.04 (two-sided). The HR of 0.766 corresponds to a 30.6% increase in the median OS, assuming a median OS of 20 months for sunitinib monotherapy (weighted median estimate assuming a median OS of 26 months in intermediate risk subjects and a median OS of 8 months in poor risk subjects) and 23.5 months for experimental treatment arms respectively. It is projected that an observed hazard ratio of 0.846 or less, which corresponds to a 3.6 months or greater improvement in median OS (20 mo vs. 23.6 mo), would result in a statistically significant improvement in OS for the experimental arm at the final OS analysis

Two formal interim analyses of OS are planned for this study. The first interim analysis is planned at the time of final PFS analysis and it is expected to observe 370 events (58% of the targeted OS events for final analysis) and the second after observing 479 events (75% of targeted OS events needed for final analysis). The stopping boundaries at interim and final analyses will be derived based on the number of deaths using O'Brien and Fleming α spending function.

Under the assumptions stated above on accrual and OS distribution, it will approximately take 61 months from FPFV to observe the required number of OS events for the final OS analysis (20.5 months for accrual and 40.5 months for minimum follow up).

Endpoints:

Co-Primary Endpoints

Objective Response Rate, Progression-free Survival and Overall Survival are the co-primary endpoints.

Objective Response Rate

Objective response rate is defined as the proportion of randomized subjects who achieve a best response of complete response (CR) or partial response (PR) using the RECIST1.1 criteria based on IRRC assessment. BOR is defined as the best response designation, as determined by the IRC, recorded between the date of randomization and the date of objectively documented progression per RECIST 1.1 or the date of subsequent therapy (including tumor-directed radiotherapy and tumor-directed surgery), whichever occurs first. For subjects without documented progression or subsequent therapy, all available response designations will contribute to the BOR assessment. As described in Section 5.4, confirmation of response is required. Duration of response (DOR) is defined as the time between the date of first documented response (CR or PR) to the date of the first documented progression as determined by the IRC (per RECIST 1.1), 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 (Section 8.3.1.2). Time to Objective Response (TTR) is defined as the time from randomization to the date of the first confirmed documented response (CR or PR), as assessed by the IRC. DOR and TTR will be evaluated for responders (confirmed CR or PR) only.

Primary Definition of Progression-free Survival

The primary definition PFS is specified as the time between the date of randomization and the first date of documented progression, based on IRRC assessments (as per RECIST 1.1 criteria), or death due to any cause, whichever occurs first. Subjects who die without a reported progression will be considered to have progressed on the date of their death. The following censoring rules will be applied for the primary definition of PFS.

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

Secondary Definition of Progression-free Survival

The secondary definition of PFS is defined as the time between the date of randomization and the first date of documented progression, based on IRRC assessments (as per RECIST 1.1 criteria), or death due to any cause,

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

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

Finally, PFS based on investigator assessments will also be analyzed applying both the primary and the secondary definitions.

More detail on PFS will be provided in a separate Statistical Analysis Plan.

Overall Survival

Overall survival is defined as the time from randomization to the date of death from any cause. For subjects that are alive, their survival time will be censored at the date of last contact ("last known alive date"). Overall survival will be censored for subjects at the date of randomization if they were randomized but had no follow-up.

Secondary Endpoints

AE Incidence Rate

Adverse events incident rate is defined as the proportion subjects with any grade adverse events among subjects treated in each treatment arm. Events reported from the first dose and up to and including 100 days following the last dose of study treatment could be included in estimating this incidence rate.

Analyses: One of the primary objectives of the study is to describe the objective response rate per IRRC in the two treatment arms among intermediate and poor risk subjects. The ORR analysis will occupy a 0.001 administrative allocation of alpha.

The number and percentage of subjects in each category of best overall response per IRRC (complete response [CR], partial response [PR], stable disease [SD], progressive disease [PD], or unable to determine [UD]) according to the IRRC will be presented, by treatment group. An estimate of the response rate and an associated exact two-sided 95% CI (Clopper and Pearson[36]) will be presented, by treatment group.

Sensitivity analysis based on investigator-determined ORR may also be performed. DOR and TTR will also be evaluated. Descriptive analysis of the response in the investigator's choice group (ie, subjects treated with investigator's choice among ORR population) will also be provided.

At the time of the formal ORR analysis, no PFS or OS analysis will be conducted because of the immaturity of those specific endpoints. A reduced analysis will be defined in the data presentation plan.

One of the primary objectives of the study is to compare the progression-free survival (based on IRRC assessments) of nivolumab combined with ipilimumab to sunitinib monotherapy in intermediate and poor-risk subjects with previously untreated mRCC. A two-sided stratified 0.009 log-rank test will be used to do a formal comparison of PFS.

A stratified log-rank test will be used to compare the PFS of subjects randomized to nivolumab combined with ipilimumab to that of subjects randomized to sunitinib. Median PFS will be estimated via the Kaplan-Meier product limit method. Two-sided 99.1% CI for the median PFS will be computed for each randomized arm. Kaplan-Meier plots of PFS will be presented. Hazard ratios (HR) and corresponding two-sided (1-adjusted α)% confidence intervals (CI) will be estimated using a Cox proportional hazards model, with treatment arm as a single covariate, stratified by the stratification factors, corresponding to each comparison of PFS.

The totality of PFS results will be presented in a single graphical display that includes Kaplan-Meier curves for the two treatment arms, the log-rank p-values for the formal comparison, the HRs and corresponding CIs, and the two median estimates and corresponding CIs.

OS will be compared between the treatment arms using a two sided, $\alpha = 0.04$ level log-rank test (adjusted for interim analyses), stratified using the same factor as in PFS. A similar analysis as in PFS will be conducted for OS.

Hazard ratios (HR) and corresponding two-sided 96% confidence intervals (CI) will be estimated using a Cox proportional hazards model, with treatment arm as a single covariate, stratified by the stratification factors, corresponding to the comparison of OS.

Amendment 14 Update:

Optional Switch to Nivolumab 240mg Flat Dosing and Optional Discontinuation after 2 Years of Study Treatment (Arm A)

Arm A subjects receiving treatment with nivolumab at the time of Amendment 14 will continue to be monitored as specified in the protocol. They may continue to receive treatment with nivolumab at the same dose or switch to nivolumab at a flat dose of 240mg given every 2 weeks (see Section 1.4.10.4 for the rationale for the nivolumab 240mg flat dose).

Arm A subjects who have completed at least 2 years of treatment have the option to discontinue study treatment at the discretion of the subject and/or investigator (see Section 1.4.10.5 for the rationale for two year maximum treatment duration). For subjects who have not been assessed to have radiographic progression at the time of study treatment discontinuation, tumor assessments must continue to be performed until disease progression is documented.

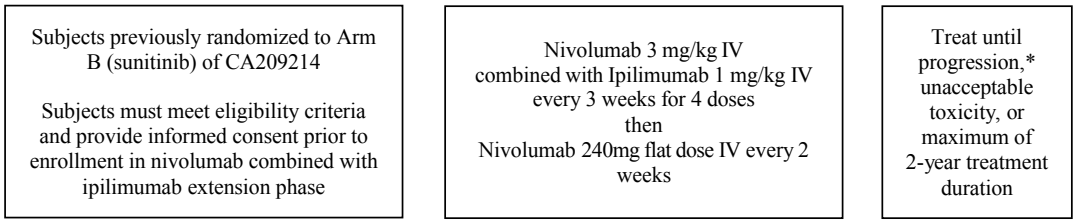
Nivolumab combined with Ipilimumab Crossover Extension Phase (Arm B)

With this amendment, all subjects in the poor or intermediate cohorts randomized to the sunitinib treatment (Arm B) who meet eligibility criteria may enter the nivolumab combined with ipilimumab crossover extension phase, according to the schema below. These subjects will be eligible to enter the crossover arm and receive BMS supplied study drug for a maximum of 2 years but no longer than up to 12 months after the approval of investigational product by the responsible health authority or until the investigational product becomes commercially available within the country, whichever occurs sooner. BMS reserves the right to terminate access to study drug if any of the following occur: a) the marketing application is rejected by the responsible health authority; b) the study is terminated due to safety concerns; c) the subject can obtain medication from a government sponsored or private health program; or d) therapeutic alternatives become available in the local market. These subjects will follow the assessment schedules outlined in Tables 5.1-5 and 5.1-6 of the protocol.

Subjects treated with sunitinib who have ended study treatment will be able to receive treatment with nivolumab combined with ipilimumab via the crossover extension phase of the study, assuming eligibility criteria are met (including a 14-day washout period for prior systemic anti-cancer therapy). Details are provided in Section 3.1.1.

Subjects currently receiving treatment with sunitinib may continue to be treated and monitored as specified in the protocol as long as they are continuing to derive benefit from sunitinib in the judgment of the investigator. These subjects may receive nivolumab combined with ipilimumab once they are discontinued from sunitinib therapy, assuming basic eligibility criteria are met (including a 14-day washout period from the last dose of sunitinib).

Nivolumab combined with Ipilimumab Crossover Extension Phase (schema for those previously randomized to sunitinib):



*Treatment beyond investigator-assessed RECIST 1.1-defined progression may be considered for subjects meeting criteria according to Section 4.5.7. Treatment beyond progression for subjects in the nivolumab combined with ipilimumab crossover extension phase should be discussed with the BMS Medical Monitor prior to subjects receiving additional study drug. Criteria for discontinuation of treatment beyond progression are described in Section 4.5.7. Subjects who discontinue study therapy for reasons other than progression should continue to be scanned until disease progression is documented.

TABLE OF CONTENTS

1 INTRODUCTION AND STUDY RATIONALE

1.1 Study Rationale

CA209214 (CheckMate 214, CHECKpoint pathway and nivoluMAb clinical Trial Evaluation) is a Phase 3, randomized, open-label study of nivolumab (BMS-936558) combined with ipilimumab vs sunitinib monotherapy in subjects with previously untreated, advanced or metastatic renal cell carcinoma (mRCC). In the Phase 1 setting, nivolumab combined with ipilimumab has demonstrated substantially greater clinical activity, as measured by objective response rate (ORR), than either agent alone. Given the durability of responses associated with immunotherapies, nivolumab combined with ipilimumab is hypothesized to lead to greater clinical benefit, as measured by objective response rate (ORR), progression-free survival (PFS), or overall survival (OS), compared to sunitinib, a widely used standard-of-care agent in this patient population. This study will allow for direct comparison of ORR, PFS and OS between arms. No approved drug has demonstrated an improvement in ORR, PFS or OS vs sunitinib in the Phase 3 setting. If nivolumab combined with ipilimumab has an acceptable safety profile and is shown to improve ORR, PFS, or OS, vs sunitinib, this study may support the approval of nivolumab combined with ipilimumab in subjects with previously untreated, advanced or metastatic RCC.

1.2 Research Hypothesis

Treatment with nivolumab combined with ipilimumab will improve ORR, PFS, or OS, compared to sunitinib monotherapy in subjects with previously untreated mRCC.

1.3 Objectives(s)

1.3.1 Primary Objectives

- To describe the ORR of nivolumab combined with ipilimumab and sunitinib monotherapy in intermediate and poor-risk subjects with previously untreated mRCC, based on IRRC assessments
- To compare the PFS of nivolumab combined with ipilimumab to sunitinib monotherapy in intermediate and poor-risk subjects with previously untreated mRCC, based on IRRC assessments
- To compare the OS of nivolumab combined with ipilimumab to sunitinib monotherapy in intermediate and poor-risk subjects with previously untreated mRCC

1.3.2 Secondary Objectives

- To compare the PFS of nivolumab combined with ipilimumab to sunitinib monotherapy in any-risk subjects with previously untreated mRCC, based on IRRC assessments
- To compare the OS of nivolumab combined with ipilimumab to sunitinib monotherapy in any-risk subjects with previously untreated mRCC
- To estimate the objective response rate (ORR) of nivolumab combined with ipilimumab and sunitinib monotherapy in subjects with previously untreated mRCC (any-risk), based on IRRC assessments

- To estimate the incidence of AEs of nivolumab combined with ipilimumab and sunitinib monotherapy in all treated subjects with previously untreated mRCC

1.3.3 Exploratory Objectives

- To assess the overall safety and tolerability of nivolumab combined with ipilimumab vs sunitinib monotherapy
- To estimate the ORR and PFS based on IRRC assessments and OS of nivolumab combined with ipilimumab vs sunitinib monotherapy in favorable risk subjects with previously untreated mRCC
- To characterize the pharmacokinetics (PK) of nivolumab and ipilimumab when co-administered
- To monitor immunogenicity of nivolumab and ipilimumab administered as combination therapy
    - To explore potential predictive biomarkers of clinical response to nivolumab-ipilimumab combination by analyzing tumor specimens and blood samples for proteins and genes involved in regulating immune responses (eg, PD-1, PD-L1PD-L2, CXCL10)
    - To assess the effects of single nucleotide polymorphisms (SNPs) in select genes (eg, PD-1, PD-L1, PD-L2, CTLA-4) on clinical endpoints and/or on the occurrence of adverse events
    - To explore associations between baseline measures of Myeloid Derived Suppressor Cells (MDSCs) and clinical outcomes
- To evaluate health related quality of life (HRQoL) as assessed by the Functional Assessment of Cancer Therapy-General (FACT-G)
- To assess disease related symptoms in each arm based on the NCCN Functional Assessment of Cancer Therapy- Kidney Symptom Index (FKSI-19)
- To assess changes in global health status in each treatment arm based on EuroQol's EQ-5D
- To assess healthcare resource utilization in each treatment arm

1.4 Product Development Background

1.4.1 Cancer Immunotherapy

Cancer immunotherapy rests on the premise that tumors can be recognized as foreign rather than as self and can be effectively attacked by an activated immune system. An effective immune response in this setting is thought to rely on immune surveillance of tumor antigens expressed on cancer cells that ultimately results in an adaptive immune response and cancer cell death. This functions by aborting the emergence of tumors as they arise and/or causing tumor shrinkage

where it is present. Meanwhile, tumor progression may depend upon acquisition of traits that allow cancer cells to evade immune surveillance and an effective immune response.[1] This evasion may occur by exploiting any of the checkpoints that control the regulatory immune response, including display of antigens and control of co-stimulatory pathways that affect the proliferation of cells involved in immunity. Current immunotherapy efforts attempt to break the apparent tolerance of the immune system to tumor cells and antigens by either introducing cancer antigens by therapeutic vaccination or by modulating regulatory checkpoints of the immune system - either directly by stimulation of immune cells by antibodies directed to receptors on T and B cells or indirectly by cytokine manipulation. T-cell stimulation is a complex process involving the integration of numerous positive, as well as negative, costimulatory signals in addition to antigen recognition by the T-cell receptor (TCR).[2] Collectively, these signals govern the balance between T-cell activation and tolerance to antigens.[1]

1.4.2 Programmed Death Receptor-1 (PD-1) and Nivolumab

Programmed death receptor-1 (PD-1, CD279), a 55 kD type I transmembrane protein, is a member of the CD28 family of T-cell costimulary receptors that also includes CD28, CTLA-4, ICOS, and BTLA.[2] PD-1 contains an intracellular membrane proximal immunoreceptor tyrosine inhibitory motif (ITIM) and a membrane distal immunoreceptor tyrosine-based switch motif (ITSM). Two ligands specific for PD-1 have been identified: PD-L1 (B7-H1/CD274) and PD-L2 (B7-DC/CD273). PD-L1 and PD-L2 have been shown to down-regulate T-cell activation upon binding to PD-1 in both murine and human systems.[3,4] PD-1 delivers a negative signal by the recruitment of a protein tyrosine phosphatase SHP-2 to the phosphorylated tyrosine residue in the ITSM in its cytoplasmic region.[5,6] PD-1 is primarily expressed on activated T cells, B cells and myeloid cells.[7]

Nivolumab is a fully human, IgG4 (kappa) isotype, mAb that binds PD-1. Blockade of the PD-1 pathway by nivolumab was studied using the mixed lymphocyte reaction (MLR). PD-1 blockade resulted in a reproducible enhancement of both proliferation and IFN-γ release in the MLR.[8] The effect of nivolumab on antigen-specific recall response was investigated using a CMV-restimulation assay with human peripheral blood mononuclear cells (PBMCs), and was evaluated by ELISA.

1.4.3 Cytotoxic T-Lymphocyte Antigen 4 (CTLA-4) and Ipilimumab

CTLA-4, an activation-induced T-cell surface molecule, is a member of the CD28:B7 immunoglobulin superfamily that competes with CD28 for B7. CTLA-4-mediated signals are inhibitory and turn off T cell-dependent immune responses.[9,10]

Ipilimumab is a fully human monoclonal IgG1κ that binds to the CTLA-4 antigen expressed on a subset of T cells from human and nonhuman primates. The proposed mechanism of action for ipilimumab is interference of the interaction of CTLA-4 with B7 molecules on APCs, with subsequent blockade of the inhibitory modulation of T-cell activation promoted by the CTLA-4/B7 interaction.

1.4.4 Preclinical Summary of Nivolumab Combined with Ipilimumab

Preclinical data indicate that the combination of PD-1 and CTLA-4 receptor blockade may improve antitumor activity. In vitro combinations of nivolumab plus ipilimumab increase IFN-γ production 2- to 7-fold over either agent alone in a mixed lymphocyte reaction. Increased antitumor activity of the combination was also observed in 3 of 5 syngeneic murine cancer models. In a murine melanoma vaccine model, blockade with either CTLA-4 or PD-1 antibodies increased the proportion of CTLA-4 and PD-1-expressing CD4/CD8 tumor infiltrating T effector cells, and dual blockade increased tumor infiltration of T effector cells and decreased intratumoral T regulatory cells, as compared to either agent alone[11].

Preclinically, a 4-week toxicity study of nivolumab in combination with ipilimumab conducted in cynomolgus monkeys demonstrated that the combination of nivolumab and ipilimumab resulted in dose-dependent gastrointestinal (GI) toxicity. Histologic findings included inflammatory changes in the large intestine, which increased in incidence and severity in a dose-dependent manner. GI toxicity/colitis was not observed in cynomolgus monkeys administered nivolumab alone, but was observed in monkeys receiving ipilimumab. Nivolumab in combination with ipilimumab was also associated with lymphoid hypocellularity of the cortex and/or medulla of the thymus and with acinar cell degranulation in the pancreas. Additional findings included interstitial mononuclear cell infiltrates in the kidneys, portal mononuclear cell infiltrates in the liver and myeloid hypercellularity in the bone marrow. Nivolumab in combination with ipilimumab at the high-dose level (ie, 50 mg/kg and 10 mg/kg, respectively) was associated with the death of 1 animal, attributed to acute gastric dilatation without histopathological evidence of colitis upon pathology evaluation of the GI tract.

1.4.5 Renal Cell Carcinoma: Background and Standard Treatments

Renal cell carcinoma (RCC) accounts for ~ 3% of all cancers in the US. This translates to 58,000 new cases a year with 13,000 associated deaths.[12] Metastatic disease is found in 30% of subjects at diagnosis. Close to 90-95% of metastatic disease is of the clear-cell histology.[13]

Multiple scoring systems are available to characterize prognosis in treatment-naive RCC. Two of the most commonly used are the Memorial Sloan-Kettering Cancer Center (MSKCC) prognostic scoring system and the International Metastatic RCC Database Consortium (IMDC) prognostic scoring system.[14,15] Each of these systems categorizes patients as favorable, intermediate, or poor-risk based on how many adverse prognostic factors are present (0: favorable-risk, 1-2: intermediate risk, 3 or more: poor-risk). The six parameters of importance for IDMC prognostic score classification are Karnofsky Performance Status (KPS), time from diagnosis to treatment, hemoglobin value, corrected calcium concentration, absolute neutrophil count, and platelet count. The five parameters included in the MSKCC prognostic score are KPS, nephrectomy status, hemoglobin value, LDH, and corrected calcium concentration. Time from diagnosis to treatment is often used in place of nephrectomy status. With each system, total number of adverse prognostic factors present has been shown to correlate with overall survival. Approximately 25% of patients are in the favorable-risk group, 50% are in the intermediate-risk group, and 25% are in the poor-risk group (mOS: ~ 9 months). In an analysis of 1028 patients

scored using the IMDC system, median OS for favorable, intermediate, and poor-risk patients is 43.2 months, 22.5 months, and 7.8 months, respectively.[16]

Until recently, the cytokines IL-2 and IFNα were the only active treatments for advanced or metastatic RCC. However, due to each of these agent's limited clinical benefit and substantial toxicity profile, newer targeted agents have largely replaced cytokines in the treatment of advanced or metastatic renal cell carcinoma.[17,18,19,20] The recognition of the importance of hypoxia inducible factor alpha (HIFα) signaling in the pathogenesis of clear-cell RCC has led to widespread study of two classes of targeted therapies, anti-angiogenic agents and mTOR inhibitors.[21] Targeting of angiogenesis is rational because constitutive HIFα activation leads to the upregulation or activation of several proteins including vascular endothelial growth factor (VEGF), which can subsequently lead to tumor proliferation and neovasculature formation. Targeting of the mTOR pathway is important because activation of the upstream PI3K/Akt/mTOR signaling pathway is one method by which constitutive HIFα activation or upregulation occurs. There are 7 agents for the treatment of RCC in the US and EU: 5 that target angiogenesis (ie, the VEGF-receptor tyrosine kinase inhibitors sorafenib, sunitinib, pazopanib, axitinib, and the VEGF-binding monoclonal antibody bevacizumab) and 2 that target the mTOR pathway (ie, everolimus and temsirolimus). Among these approved agents, none has demonstrated a statistically significant improvement in OS except for temsirolimus poor-risk patients. According to NCCN guidelines, sunitinib, temsirolimus (poor-risk only), bevacizumab plus interferon and pazopanib are Category 1 recommendations for first-line therapy of mRCC.[22] According to ESMO guidelines, sunitinib, bevacizumab plus interferon, and pazopanib are all standard treatment options for favorable and intermediate-risk patients, but sunitinib is the only one also considered an alternative to temsirolimus for the treatment of poor-risk patients.[23]

1.4.6 Sunitinib in Renal Cell Carcinoma

Sunitinib is a VEGF receptor TKI that is approved and recommended for the treatment of mRCC across prognostic groups.[22,23] In a randomized Phase 3 trial of sunitinib vs IFNα in treatment-naive subjects, mPFS and mOS were greater in the sunitinib group than in the IFNα group (mPFS: 11 mo vs 5 mo, HR = 0.539; p = < .001); mOS: 26.4 mo vs 21.8 mo, HR = 0.821; p = .051).[24] The ORR was also greater in the sunitinib group (47%) than in the IFNα group (12%). More recently, sunitinib was compared to pazopanib in a treatment-naive subjects in the Phase 3 COMPARZ study.[25] In this non-inferiority study, sunitinib and pazopanib demonstrated similar mPFS (8.4 mo for pazopanib vs 9.5 mo for sunitinib, HR = 1.05) and mOS (28.4 mo for pazopanib vs 29.3 mo for sunitinib, HR = 0.91, p = 0.28). The ORR of pazopanib and sunitinib was 31% and 24%, respectively. The most common (≥ 20)adverse reactions include fatigue, asthenia, fever, diarrhea, nausea, mucositis/stomatitis, vomiting, dyspepsia, abdominal pain, constipation, hypertension, peripheral edema, rash, hand-foot syndrome, skin discoloration, dry skin, hair color changes, altered taste, headache, back pain, arthralgia, extremity pain, cough, dyspnea, anorexia, and bleeding.[26] Other important adverse reactions

include hepatotoxicity, QT prolongation (including Torsades de Pointes), osteonecrosis of the jaw, tumor lysis syndrome, and thyroid dysfunction.

1.4.7 Nivolumab in Renal Cell Carcinoma

Nivolumab monotherapy has been studied in subjects with renal cell carcinoma in several BMS-sponsored studies, with the largest amount of data coming from two studies in subjects with mRCC: CA209009 and CA209010. In CA209010, 168 subjects who received at least one prior-anti-angiogenic therapy were randomized to receive nivolumab 0.3 mg/kg (n = 60), 2 mg/kg (n = 54), and 10 mg/kg (n = 54).[27] Median PFS was 2.7 mo, 4.0 mo, and 4.2 mo at 0.3, 2, and 10 mg/kg respectively. The ORR ranged from 20-22% across dose levels. Median OS was 18.2 mo at 0.3 mg/kg, but was not yet reached at the two highest dose levels. CA209009 enrolled a similar population to CA209010, but also included 23 subjects with treatment-naive RCC. Among treatment-naive subjects, all of whom received nivolumab 10 mg/kg every 3 weeks, the ORR was 13% (3/23).

CA209010 includes the largest safety database for nivolumab monotherapy in mRCC. All treated subjects (n = 167) were included in the safety analyses. Drug-related AEs of any grade occurred in 74.6%, 66.7%, and 77.8% of subjects treated at 0.3 mg/kg, 2 mg/kg, and 10 mg/kg respectively. The most common (≥ 10% in any group) drug-related AEs included fatigue, dry skin, rash, pruritis, arthralgia, nausea, diarrhea, decreased appetite, dry mouth, and hypersensitivity. Grade 3 drug-related AEs occurred in 5.1%, 16.7%, and 13% of subjects treated at 0.3 mg/kg, 2 mg/kg, and 10 mg/kg, respectively. Related Grade 3 events in at least 2 patients across dose levels included nausea, AST/ALT increased, and anemia. No drug-related Grade 4 or Grade 5 events occurred. No dose-toxicity relationship was identified except for hypersensitivity/infusion reaction which occurred most frequently in the 10 mg/kg treatment group.

1.4.8 Ipilimumab in Renal Cell Carcinoma

Ipilimumab monotherapy for the treatment of mRCC was studied in the Phase 2 clinical trial MDX010-11.[28] Two sequential cohorts were studied, each with a loading dose of 3 mg/kg followed by 3 doses of either 1 mg/kg (group 3-1; n = 21) or 3 mg/kg (group 3-3; n = 40). Subjects with stable disease or partial or complete response were allowed additional treatment. In group 3-1 (n = 21), one subject (5%) had a PR.[29] In group 3-3 (n = 40), 5 subjects (12.5 %) had a PR. Among 14 treatment-naive subjects in group 3-3, 3 (21%) had a PR.

In the ipilimumab monotherapy Phase 2 clinical trial MDX010-11, the major toxicities were colitis (all Grade 3 & 4; 14% in group 3-1, 33% in group3-3) and hypophisitis (1 grade 3/4, 1 grade 1/2 in group 3-3; none in group 3-1). Most reported AEs were Grade 1/2 (57% in group 3-1, 35% in group3-3) or Grade 3 (38% in group 3-1, 48 % in group 3-3).[30] Most reported AEs were Grade 1/2 (57% in group 3-1, 35% in group3-3) or Grade 3 (38% in group 3-1, 48 % in group 3-3). There were 6 subjects (15%) with Grade 4 AEs in group 3-3. The most common treatment-related AEs in group 3-1 (total 81%) and group 3-3 (total 93%) were diarrhea (38% and 40%, respectively) and fatigue (33% and 38%, respectively). Most AEs were

manageable with appropriate treatment, including high dose corticosteroids and hormone replacement.

1.4.9 Nivolumab Combined with Ipilimumab in Renal Cell Carcinoma

The combination of nivolumab with ipilimumab is currently being studied in the Phase 1 study CA209016.[31] Subjects with mRCC (favorable/intermediate MSKCC score; Karnofsky performance status ≥ 80%; untreated or any number of prior therapies) were randomized to receive nivolumab 3 mg/kg + ipilimumab 1 mg/kg (arm N3 + I1) or nivolumab 1 mg/kg + ipilimumab 3 mg/kg (arm N1 + I3) IV Q3W for 4 doses then nivolumab 3 mg/kg IV Q2W until progression/toxicity. The primary objective was to assess safety/tolerability; secondary objective was to assess antitumor activity. Subjects were randomized to N3 + I1 (n = 21) and N1 + I3 (n = 23). Most pts (n = 34; 77%) had prior systemic therapy (N3 + I1: 16; N1 + I3: 18). The confirmed ORR was 43% (N3 + I1) and 48% (N1 + I3) (Table 1.4.9-1). Duration of response (DOR) was 4.1+ to 42.1+ wks (7 of 9 responses ongoing) in N3 + I1, and 12.1+ to 35.1+ wks (9 of 11 responses ongoing) in N1 + I3. Best response of stable disease (SD) was seen in 5 (24%) pts (N3 + I1) and 8 (35%) pts (N1 + I3). Median PFS was 36.6 wks (N3 + I1) and 38.3 wks (N1 + I3); these data are still immature, with 11 of 21 events reported for N3 + I1 and 10 of 23 events reported for N1 + I3.

Table 1.4.9-1: Antitumor Activity

	N3 + I1 (n = 21)	**N1 + I3 (n = 23)**
Confirmed ORR, n (%) (95% CI)	9 (43) (21.8, 66.0)	11 (48) (26.8, 69.4)
Median duration of response, weeks (range)	31.1 (4.1+ - 42.1+)	Not reached (12.1+ - 35.1+)
Ongoing responses, % (n/N)	78 (7/9)	82 (9/11)
Best objective response, n (%) Complete response Partial response Stable disease Progressive disease Unable to determine	 0 9 (43) 5 (24) 5 (24) 1 (5)	 1 (4) 10 (43) 8 (35) 3 (13) 1 (4)
24 week PFS, % (95% CI)	65 (40, 82)	64 (41, 80)

The safety of nivolumab combined with ipilimumab was assessed in the Phase 1 study CA209016. Treatment-related adverse events (AEs) were seen in 39/44 pts (89%), including 16/21 (76.2%) in N3 + I1 and 23/23 (100%) in N1 + I3. Across the N3 + I1 and N1 + I3 arms, the most common (≥ 20%) treatment related AEs of any grade were fatigue (61%), diarrhea (32%), nausea (30%), rash (27%), pruritis (25%), ALT increased (23%), AST increased (20%), hypothyroidism (20%), and lipase increased (20%). Grade 3–4 related AEs occurred in 19 pts (29%), including 6/21 (29%) at N3 + I1 and 14/23 (61%) at N1 + I3. The most common (≥ 5%)

drug-related Grade 3-4 events were lipase increased (21%), ALT increased (14%), AST increased (7%), diarrhea (9%), fatigue (5%), amylase increased (5%), colitis (5%), lymphocyte count decreased (5%). No grade 3–4 pneumonitis was seen. No treatment-related deaths were reported.

Treatment-related AEs (including Grade 3-4), treatment-related AEs leading to discontinuation, and treatment-related SAEs all occurred more commonly in subjects in the N1 + I3 arm than in the N3 + I1 arm (Table 1.4.9-2).

Table 1.4.9-2: Safety by Dose

	N3 + I1 (n = 21)	N1 + I3 (n = 23)
Treatment-related AEs	76.2%	100%
Treatment-related Grade 3-4 AEs	28.6%	60.9%
Treatment-related SAEs	9.5%	26.1%
Treatment-related AEs leading to discontinuation	9.5%	26.1%

In summary, a similar robust level of clinical activity was observed with N3 + I1 and N1 + I3, but the N3 + I1 arm exhibited a more favorable safety profile.

1.4.10 Rationale for Study Design

1.4.10.1 Rationale for nivolumab combined with ipilimumab and choice of N3 + I1 dosing regimen

Data from CA209016 demonstrate a level of clinical activity, as measured by ORR, for the combination of nivolumab combined with ipilimumab that is substantially greater than that of either nivolumab or ipilimumab monotherapy in mRCC. These immunotherapy-induced responses are expected to be more durable than those induced by sunitinib monotherapy, and are therefore likely to translate into improvements in PFS, OS, or both vs sunitinib. The dosing regimen including nivolumab 3 mg/kg combined with ipilimumab 1 mg/kg was chosen because it exhibits similar clinical activity nivolumab 1 mg/kg combined with ipilimumab 3 mg/kg along with a more favorable safety profile.

1.4.10.2 Rationale for 2-arm design

The study will include 2 arms:

- Arm A: Nivolumab 3 mg/kg IV combined with Ipilimumab 1 mg/kg IV every 3 weeks for 4 doses then Nivolumab 3 mg/kg or 240mg flat dose IV every 2 weeks
- Arm B: Sunitinib 50 mg PO once daily for 4 weeks followed by 2 weeks off, continuously

The ORR for each of these regimens is reported to be approximately 40%. For nivolumab combined with ipilimumab, an ORR of 45% was reported in 44 subjects with mRCC treated in CA209016, including both treatment-naive and pre-treated patients, who received either nivolumab 3 mg/kg combined with ipilimumab 1 mg/kg or nivolumab 1 mg/kg combined with ipilimumab 3 mg/kg (ORR 43% for nivolumab 3 mg/kg + ipilimumab 1 mg/kg [n = 21], ORR 48% for nivolumab 1 mg/kg + ipilimumab 3 mg/kg [n = 23]). For sunitinib, the ORR for treatment-naive mRCC has been reported to be as high as 47%.[24]

This study does not include additional arms to demonstrate the contribution of nivolumab or ipilimumab monotherapy because the ORR for nivolumab and ipilimumab monotherapy is ~ 20% or less when each of these agents is used as monotherapy. In nivolumab study CA209009, the ORR among 23 subjects with treatment-naive mRCC who received nivolumab 10 mg/kg every 3 weeks was 13%. This dosing regimen, although different from the nivolumab 3 mg/kg dosing regimen used in the current study, provides a higher Cmin than would be expected with nivolumab 3 mg/kg every 2 week monotherapy. Therefore, the clinical activity is expected to be similar, as no dose-response relationship is reported for nivolumab for doses ranging from 0.3 to 10 mg/kg every 3 weeks.[27] In a phase 1 study of ipilimumab in mRCC, 14 treatment-naive subjects with mRCC received ipilimumab 3 mg/kg monotherapy. The ORR among these 14 subjects was 21%. In subjects who received ipilimumab 1 mg/kg monotherapy after a single 3 mg/kg dose, the ORR was 5%.[29]

Given that the ORR with either nivolumab or ipilimumab as monotherapy is ~ 20% or less, both agents in combination are expected to be required in order to achieve an ORR comparable to or better than that reported for sunitinib monotherapy.

1.4.10.3 Rationale for Choice of Primary Endpoints in Intermediate and Poor-Risk Subjects

The study will include co-primary endpoints of PFS and OS, to be evaluated in subjects with intermediate or poor prognosis according to IMDC prognostic criteria. Either OS or PFS, as long as there is no detriment in OS, have been successful endpoints to support drug approvals in mRCC. Sorafenib, sunitinib, bevacizumab, axitinib, and everolimus were each approved based on improvement in PFS. Temsirolimus was approved based on improvement in OS in a poor-risk population. Although tivozanib demonstrated a statistically significant improvement in PFS, this did not lead to regulatory approval, as a detriment in OS was noted in subjects on the tivozanib arm. Overall Response Rate (ORR) is included because of the improvement in ORR rates seen in CA209-016, with the intent to describe the ORR of nivolumab combined with ipilimumab and sunitinib monotherapy.

The evaluation of primary endpoints is limited to intermediate and poor-risk subjects, which comprise approximately 75% of the total treatment-naive mRCC population. Inclusion of this large subset of subjects in the primary endpoints of the study will allow for potential meaningful differences in efficacy to be detected earlier than if favorable-risk patients are included. This may allow the nivolumab ± ipilimumab combination to be made available to larger numbers of mRCC patients in a more timely fashion. Of note, favorable-risk patients will be included in the

study, with safety and efficacy endpoints including this population included among the secondary analyses.

1.4.10.4 Rationale for Nivolumab 240mg Flat Dose

Under Amendment 14, participants in Arm A who are receiving nivolumab 3 mg/kg every 2 weeks will have the option to switch to nivolumab 240 mg every 2 weeks. In addition, nivolumab 240 mg every 2 weeks will be the maintenance dosing used in the crossover extension phase.

The nivolumab dose of 240 mg every 2 weeks (Q2W) was selected based on clinical data and modeling and simulation approaches using population PK (PPK) and exposure-response analyses of data from studies in multiple tumor types (melanoma, non-small-cell lung cancer [NSCLC], and renal cell carcinoma [RCC]) where body weight normalized dosing (mg/kg) has been used.

PPK analyses have shown that the PK of nivolumab is linear with proportional exposure over a dose range of 0.1 to 10 mg/kg, and no differences in PK across ethnicities and tumor types were observed. Nivolumab clearance and volume of distribution were found to increase as the body weight increases, but less than the proportional with increasing weight, indicating that mg/kg dosing represents an over-adjustment for the effect of body weight on nivolumab PK. The PPK model previously developed using data from NSCLC subjects has recently been updated, using data from 1544 subjects from 7 studies investigating nivolumab in the treatment of melanoma, NSCLC, and RCC. In this dataset, the median (minimum - maximum) weight was 77 kg (35 kg - 160 kg) and thus, an approximately equivalent dose of 3 mg/kg for an 80 kg subject, nivolumab 240 mg Q2W was selected for future studies. To predict relevant summary exposures of nivolumab 240 mg Q2W, the PPK model was used to simulate nivolumab 3 mg/kg Q2W and 240 mg Q2W. In the simulations, the simulated patient populations consisted of 1000 subjects per treatment arm randomly sampled from aforementioned pooled database of cancer subjects.

Because no differences in PK were noted across ethnicities and tumor types, these simulated melanoma and NSCLC data will be applicable to subjects with other tumor types. The simulated measure of exposure of interest, time-averaged concentrations (Cavgss) for 240 mg Q2W are predicted to be similar for all subjects in reference to 80 kg subjects receiving 3 mg/kg Q2W.

Nivolumab is safe and well tolerated up to 10 mg/kg Q2W dose level. Adverse events have been broadly consistent across tumor types following monotherapy and have not demonstrated clear dose-response or exposure-response relationships. Additionally, the simulated median and 95th prediction interval of nivolumab summary exposures across body weight range (35 - 160 kg) are predicted to be maintained below the corresponding observed highest exposure experienced in nivolumab ie, 95th percentile following nivolumab 10 mg/kg Q2W from clinical study CA209003. Thus, while subjects in the lower body weight ranges would have greater exposures than 80 kg subjects, the exposures are predicted to be within the range of observed exposures at doses (up to 10 mg/kg Q2W) used in the nivolumab clinical program, and are not considered to put subjects at increased risk. For subjects with greater body weights, the simulated ranges of exposures are also not expected to affect efficacy, because the exposures predicted following administration of a 240 mg Q2W are on the flat part of the exposure-response curves for

previously investigated tumors, melanoma and NSCLC. Given the similarity of nivolumab PK across tumor types and the similar exposures predicted following administration of 240 mg flat dose compared to 3 mg/kg, it is expected that the safety and efficacy profile of 240 mg nivolumab will be similar to that of 3 mg/kg nivolumab. The USPI for nivolumab has been updated to allow 240 mg flat dose IV every 2 weeks for RCC and several other tumors.

1.4.10.5 Rationale for Two Year Maximum Duration of Treatment

The optimal duration of immunotherapy is an important question and continues to be investigated. Clinical trials across different tumors types in the nivolumab and ipilimumab development program indicate that most of the responses occur early, with a median time to response of 2-4 months, and emerging data suggests that benefit can be maintained in the absence of continued treatment. A recent analysis in a melanoma study suggests the majority of patients who discontinue nivolumab and/or ipilimumab for toxicity maintain disease control in the absence of further treatment.[32] Furthermore, a limited duration of ipilimumab, including only 4 induction doses, resulted in long term survival in patients with metastatic melanoma, with a sustained plateau in survival starting around 2 years after the start of treatment.[33]

Accumulating data suggest that 2 years of PD-1 checkpoint inhibitor treatment may be sufficient for long term benefit. CA209003, a dose-escalation cohort expansion trial evaluating the safety and clinical activity of nivolumab in patients with previously treated advanced solid tumors (including 129 subjects with NSCLC), specified a maximum treatment duration of 2 years. Among 16 subjects with non-small cell lung cancer (NSCLC) who discontinued nivolumab after completing 2 years of treatment, 12 subjects were alive >5 years and remained progression-free without any subsequent therapy (2) In the CA209003 NSCLC cohort, the overall survival (OS) curve begins to plateau after 2 years, with an OS rate of 25% at 2 years and 18% at 3 years.[34] These survival outcomes are similar to phase 3 studies in previously treated NSCLC, in which nivolumab treatment was continued until progression or unacceptable toxicity (2 year OS rates of 23% and 29%, and 3 year OS rates of 16%-18% for squamous and non-squamous NSCLC respectively).[35]

Similar results have been reported in clinical studies of pembrolizumab, another PD-1 inhibitor. Keynote-010 was a randomized phase 3 trial of pembrolizumab (at either 2 mg/kg or 10 mg/kg every 3 weeks) versus docetaxel in subjects with previously treated, PD-L1-positive, advanced NSCLC which specified a maximum treatment duration of 2 years for pembrolizumab. OS was significantly longer with both pembrolizumab 2 mg/kg (HR 0.72, $p = 0.00017$) and pembrolizumab 10 mg/kg (HR 0.60, $p < 0.00001$) compared to docetaxel, with an OS plateau developing beyond 2 years in both pembrolizumab arms. Among 690 patients who received pembrolizumab, 47 patients completed 2 years of pembrolizumab and stopped treatment. Most were able to maintain their response, including those with stable disease, with only 2 patients (4%) having confirmed progression after stopping at 2 years.[36]

Keynote-006 was a randomized phase 3 study of pembrolizumab versus ipilimumab in patients with advanced melanoma, which also specified a maximum 2 year duration of pembrolizumab treatment. 104 (19%) of 556 patients randomized to pembrolizumab completed 2 years of

treatment. With a median follow-up of 9 months after completion of pembrolizumab, the estimated risk of progression or death was 9% in these patients.[37]

Taken together, these data suggest that treatment beyond 2 years is unlikely to confer additional clinically meaningful benefit and that the risk of progression after discontinuing treatment at 2 years is low.

In contrast, a shorter duration of nivolumab of only 1 year was associated with increased risk of progression in previously treated patients with NSCLC, suggesting that treatment beyond 1 year is likely needed. In CA209153, patients with previously treated advanced NSCLC who completed 1 year of nivolumab therapy were randomized to either continue or stop treatment, with the option of retreatment upon progression. Among 163 patients still on treatment at 1 year and without progression, those who were randomized to continue nivolumab had significant improvement in progression-free survival (PFS) compared to those who were randomized to stop treatment, with median PFS (post-randomization) not reached vs 10.3 months, respectively; HR=0.42 (95% CI, 0.25 to 0.71).[38] With a median follow-up of 14.9 months post-randomization, there also was a trend for patients on continued treatment to live longer (OS HR = 0.63 [95% CI: 0.33, 1.20]). Of note, the PFS curves in both groups plateau approximately 1 year after randomization (i.e., 2 years after treatment initiation), suggesting that there may be minimal benefit in extending treatment beyond a total of 2 years.

Collectively, these data suggest that there is minimal if any benefit derived from continuing I-O treatment beyond two years in advanced tumors. However, even though immunotherapy is well tolerated, patients will be at risk for additional toxicity with longer term treatment. Therefore, in this study, treatment will be given for a maximum of 2 years from the start of study treatment for patients in the nivolumab combined with ipilimumab crossover extension phase. For patients on Arm A treatment, discontinuation of study treatment is allowed after 2 years of treatment, should the investigator feel it is an appropriate option for the subject.

1.4.10.6 Rationale for Crossover Extension Phase

The Data Monitoring Committee (DMC) for the CA209214 study convened on 06-Sep-2017 to review the first planned, Interim Analysis of the co-primary endpoint of overall survival (OS) in intermediate/poor risk subjects based on a database lock on 07-Aug-2017. The DMC confirmed that the pre-specified boundary for OS (adjusted significance boundary < 0.002) was surpassed and unanimously recommended that the study be stopped early by the Sponsor, Bristol-Myers Squibb.

Given the statistical significance of the OS co-primary endpoint, the results from the 07-Aug-2017 database lock represent the final analysis of CA209214. After a median follow-up of 25.2 months, the combination of nivolumab + ipilimumab demonstrated superior OS compared with sunitinib in intermediate/poor-risk subjects with advanced RCC, reducing the risk of death by 37%. The nivolumab + ipilimumab combination also demonstrated a significantly higher IRRC-assessed ORR compared to sunitinib (41.6% vs 26.5%), with objective responses that were deeper, including 9.4% of subjects achieving a complete response (CR), and more durable, with a median duration of response (DOR) not reached at the time of database lock. Clinically

meaningful improvement in PFS was also demonstrated with the nivolumab + ipilimumab combination vs sunitinib in intermediate/poor risk subjects. Exploratory analyses in favorable risk subjects showed more favorable efficacy in the sunitinib arm vs the nivolumab + ipilimumab arm.[39]

Based on these results, protocol Amendment 14 will provide a mechanism for eligible subjects randomized to sunitinib treatment (Arm B) to receive subsequent nivolumab combined with ipilimumab therapy for a maximum treatment period of 2 years as part of a crossover extension phase. Since the clinical benefit of nivolumab + ipilimumab combination in favorable risk subjects has not been clearly demonstrated, only subjects with intermediate/poor risk prior to randomization will be eligible for the crossover extension phase.

1.5 Overall Risk/Benefit Assessment

Patients with mRCC have multiple treatment options available to them, but none of the 7 available targeted agents have been able to demonstrate a significant improvement in overall survival when compared to each other. Median overall survival remains less than 4 years for treatment-naive patients with the most favorable prognosis, and is substantially shorter for patients who possess adverse prognostic factors. Therefore, new therapeutic options with the potential to provide greater survival across risk groups are needed. Nivolumab 3 mg/kg combined with ipilimumab 1 mg/kg has demonstrated substantial clinical activity, as measured by ORR, while still exhibiting an acceptable safety profile. These immunotherapy-induced responses are expected to be more durable than those induced by VEGF receptor therapy, and are therefore likely to translate into improvements in PFS and OS vs sunitinib.

2 ETHICAL CONSIDERATIONS

2.1 Good Clinical Practice

This study will be conducted in accordance with Good Clinical Practice (GCP), as defined by the International Conference on Harmonisation (ICH) and in accordance with the ethical principles underlying European Union Directive 2001/20/EC and the United States Code of Federal Regulations, Title 21, Part 50 (21CFR50).

The study will be conducted in compliance with the protocol. The protocol and any amendments and the subject informed consent will receive Institutional Review Board/Independent Ethics Committee (IRB/IEC) approval/favorable opinion prior to initiation of the study.

All potential serious breaches must be reported to BMS immediately. A serious breach is a breach of the conditions and principles of GCP in connection with the study or the protocol, which is likely to affect, to a significant degree, the safety or physical or mental integrity of the subjects of the study or the scientific value of the study.

Personnel involved in conducting this study will be qualified by education, training, and experience to perform their respective tasks.

This study will not use the services of study personnel where sanctions have been invoked or where there has been scientific misconduct or fraud (eg, loss of medical licensure, debarment).

2.2 Institutional Review Board/Independent Ethics Committee

Before study initiation, the investigator must have written and dated approval/favorable opinion from the IRB/IEC for the protocol, consent form, subject recruitment materials (eg, advertisements), and any other written information to be provided to subjects. The investigator or BMS should also provide the IRB/IEC with a copy of the Investigator Brochure or product labeling information to be provided to subjects and any updates.

The investigator or BMS should provide the IRB/IEC with reports, updates and other information (eg, expedited safety reports, amendments, and administrative letters) according to regulatory requirements or institution procedures.

2.3 Informed Consent

Investigators must ensure that subjects are clearly and fully informed about the purpose, potential risks, and other critical issues regarding clinical studies in which they volunteer to participate.

In situations where consent cannot be given to subjects, their legally acceptable representatives (as per country guidelines) are clearly and fully informed about the purpose, potential risks, and other critical issues regarding clinical studies in which the subject volunteers to participate.

BMS will provide the investigator with an appropriate (ie, Global or Local) sample informed consent form which will include all elements required by ICH, GCP and applicable regulatory requirements. The sample informed consent form will adhere to the ethical principles that have their origin in the Declaration of Helsinki.

Investigators must:

1) Provide a copy of the consent form and written information about the study in the language in which the subject is most proficient prior to clinical study participation. The language must be non-technical and easily understood.
2) Allow time necessary for subject or subject's legally acceptable representative to inquire about the details of the study.
3) Obtain an informed consent signed and personally dated by the subject or the subject's legally acceptable representative and by the person who conducted the informed consent discussion.
4) Obtain the IRB/IEC’s written approval/favorable opinion of the written informed consent form and any other information to be provided to the subjects, prior to the beginning of the study, and after any revisions are completed for new information.
5) If informed consent is initially given by a subject’s legally acceptable representative or legal guardian, and the subject subsequently becomes capable of making and communicating his or her informed consent during the study, consent must additionally be obtained from the subject.
6) Revise the informed consent whenever important new information becomes available that is relevant to the subject's consent. The investigator, or a person designated by the investigator, should fully inform the subject or the subject's legally acceptable representative or legal guardian, of all pertinent aspects of the study and of any new information relevant to the

subject's willingness to continue participation in the study. This communication should be documented.

The confidentiality of records that could identify subjects must be protected, respecting the privacy and confidentiality rules applicable to regulatory requirements, the subjects' signed ICF and, in the US, the subjects' signed HIPAA Authorization.

The consent form must also include a statement that BMS and regulatory authorities have direct access to subject records.

The rights, safety, and well-being of the study subjects are the most important considerations and should prevail over interests of science and society.

3 INVESTIGATIONAL PLAN

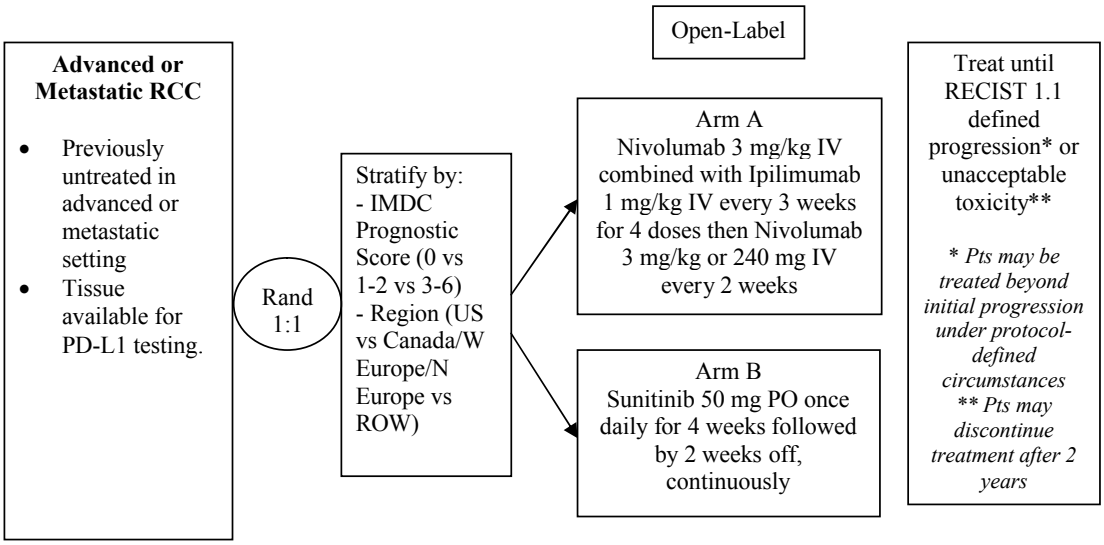
3.1 Study Design and Duration

This is a Phase 3, randomized, open-label study of nivolumab 3 mg/kg combined with ipilimumab 1 mg/kg every 3 weeks for 4 doses followed by nivolumab 3 mg/kg every 2 weeks vs sunitinib monotherapy using the approved dose and schedule (50 mg po once daily for 4 weeks followed by 2 weeks off, every cycle) in adult (≥ 18 years) subjects with previously untreated advanced or metastatic RCC. The study is expected to randomize approximately 820 subjects with intermediate or poor prognosis and up to approximately 250 subjects with favorable prognosis as per IMDC criteria. Tumor tissue, archival or recent acquisition, must be received by the central vendor in order to be randomized. Subjects must have advanced (not amenable to curative surgery or radiation) or metastatic (AJCC Stage IV) RCC, and must not have received prior systemic therapy for the treatment of advanced or metastatic RCC. Prior adjuvant or neoadjuvant therapy is allowed if such therapy did not include an agent that targets VEGF or VEGF receptors and was completed at least 6 weeks prior to randomization. Subjects will be randomized 1:1 and stratified by IMDC prognostic score (0 vs 1-2 vs 3-6) and region (US vs Canada/W Europe/N Europe vs Rest of World). Subjects will be randomized to Arm A (nivolumab 3 mg/kg IV combined with ipilimumab 1 mg/kg IV every 3 weeks for 4 doses, followed by nivolumab 3 mg/kg (*Amendment 14:* or 240mg) IV every 2 weeks) or Arm B (sunitinib using the approved dose and schedule of 50 mg PO once daily for 4 weeks followed by two weeks off, continuously). No dose increases or reductions will be allowed for nivolumab. Dose modifications for sunitinib will be allowed as per the approved product label. A maximum of 2 sunitinib dose reductions in 12.5 mg increments will be allowed. Subjects will be assessed for response (RECIST 1.1) by CT or MRI beginning 12 weeks (± 1 week) from randomization and continuing every 6 weeks (± 1 week) for the first 13 months and then every 12 weeks until progression or treatment discontinuation, whichever occurs later. Subjects will be allowed to continue study therapy after initial investigator-assessed RECIST 1.1-defined progression if assessed by the investigator to be deriving clinical benefit and tolerating study drug. Such subjects should discontinue study therapy when further progression is documented (see Section 4.5.7). The co-primary endpoints of this study are ORR and PFS in intermediate and poor-risk subjects, as assessed by an Independent Radiology Review Committee (IRRC) and OS

in intermediate and poor-risk subjects. However, the ORR analysis is intended to be descriptive and will occupy a 0.001 administrative adjustment of alpha. The analysis of ORR will occur after approximately 6 months of minimum follow-up. The final analysis of PFS will occur after 591 events. The final analysis of OS will occur after approximately 639 events (ie, deaths) have occurred. Interim analyses of OS will occur at the time of final PFS analysis and after at least 479 events (75% of targeted OS events needed for final analysis) have occurred. Key secondary endpoints include PFS, OS and ORR regardless of prognostic score.

The study design schematic is presented in Figure 3.1-1.

Figure 3.1-1: Study Design Schematic



Amendment 14 Update:

Optional Switch to Nivolumab 240mg Flat Dosing and Optional Discontinuation after 2 Years of Study Treatment (Arm A)

Arm A subjects receiving treatment with nivolumab at the time of Amendment 14 will continue to be monitored as specified in the protocol. They may continue to receive treatment with nivolumab at the same dose or switch to nivolumab at a flat dose of 240mg given every 2 weeks (see Section 1.4.10.4 for the rationale for the nivolumab 240mg flat dose).

Arm A subjects who have completed at least 2 years of treatment have the option to discontinue study treatment at the discretion of the subject and/or investigator (see Section 1.4.10.5 for the rationale for two year maximum treatment duration). For subjects who have not been assessed to have radiographic progression at the time of study treatment discontinuation, tumor assessments must continue to be performed until disease progression is documented.

Nivolumab combined with Ipilimumab Crossover Extension Phase (Arm B)

Based on the positive results of the final analysis of the study (see Section 1.4.10.6 for details), Protocol Amendment 14 will provide a mechanism for subjects randomized to Arm B who meet eligibility criteria specified in Section 3.1.1 to receive subsequent nivolumab combined with ipilimumab therapy as part of a crossover extension phase. In the crossover extension phase, subjects will receive nivolumab 3 mg/kg IV combined with ipilimumab 1 mg/k IV every 3 weeks for 4 doses, followed by nivolumab 240 mg IV every 2 weeks, until progression, unacceptable toxicity, or a maximum of 2 years from the first nivolumab combined with ipilimumab dose. Continuation of treatment in the crossover extension phase beyond investigator-assessed progression is permitted if the criteria specified in Section 4.5.7 have been met and provided that treatment does not extend beyond 2 years from the first nivolumab + ipilimumab dose given in the crossover extension phase. Additional procedures required during the crossover extension phase are specified in Table 5.1-6 and Table 5.1-5.

This study will consist of four phases: screening, treatment, follow-up, and nivolumab combined with ipilimumab crossover extension.

Screening Phase:

- Begins by establishing the subject's initial eligibility and signing of the informed consent form (ICF).
- Subject is enrolled using the Interactive Voice Response System (IVRS).
- Tumor tissue must be received at the Central Laboratory for biomarker analyses in order for the subject to be randomized. If an insufficient amount of tumor tissue is available prior to the start of the screening phase, subjects must consent to allow the acquisition of additional tumor tissue.

Treatment Phase:

- Begins with the randomization call to the IVRS. The subject will be randomly assigned to either the nivolumab combined with ipilimumab arm (Arm A) or the sunitinib arm (Arm B).
- A negative pregnancy test must be documented within 24 hours prior to the start of investigational product.
- PRO (Patient Reported Outcome) instruments must be completed after randomization, prior to the first dose of study therapy and according to the schedule in Table 5.1-2.
- Within 3 days from randomization the subject must receive the first dose of study medication (Day 1 of Cycle 1).
- On-study laboratory assessments (Cycle 2 and beyond) should be drawn within 72 hours prior to dosing.
- Adverse event assessments should be documented at each clinic visit.
- WOCBP must have a pregnancy test during week 1 and week 4 for cycles 1-2 and week 1 and week 5 starting from cycle 3. Table 5.1-2 and Table 5.1-3.

PK samples and immunogenicity samples will be collected according to the schedule in Table 5.5.1-1.

- Study drug dosing may be delayed for toxicity.
- Each cycle during the treatment phase is expected to last 6 weeks, unless there are delays.

For subjects on the nivolumab + ipilimumab arm (Arm A)

For the first 2 cycles;

- Nivolumab and ipilimumab are administered every 3 weeks for 4 doses

Starting cycle 3:

- Nivolumab is administered every 2 weeks

For subjects on the sunitinib arm (Arm B)

Starting cycle 1:

- Sunitinib is administered daily for 4 weeks, followed by 2 weeks off

For all subjects, regardless of arm:

- Treated subjects will be evaluated for response according to the RECIST 1.1 guidelines beginning 12 weeks (± 1 week) after randomization and continuing every 6 weeks (± 1 week) for the first 13 months, and then every 12 weeks (± 1 week) until disease progression or treatment discontinuation, whichever occurs later.
- This treatment phase ends when the subject is discontinued from study therapy.

Follow-Up Phase

- Begins when the decision to discontinue a subject from study therapy is made (no further treatment with study therapy).
- Two follow-up visits include collection of PK/immunogenicity samples.
- Subjects who discontinue treatment for reasons other than tumor progression will continue to have tumor assessments beginning 12 weeks (± 1week) after randomization and continuing every 6 weeks (± 1 week) for the first 13 months from randomization, and every 12 weeks (± 1 week) thereafter until documented tumor progression.
- Subjects will be followed for drug-related toxicities until these toxicities resolve, return to baseline or are deemed irreversible. All adverse events will be documented for a minimum of 100 days after last dose.
- After completion of the first two follow-up visits, subjects will be followed every 3 months for survival.

- PRO instruments will be completed according to the schedule in Table 5.1-7.

Nivolumab combined with Ipilimumab Crossover Extension Phase:

Screening

- Begins by establishing the subject's eligibility according to Section 3.1.1 and signing of the informed consent form (ICF) for the crossover extension phase.
- Subject is entered into the crossover extension phase using the Interactive Voice Response System (IVRS).

Treatment

- A negative pregnancy test must be documented within 24 hours prior to the start of nivolumab combined with ipilimumab.
- PRO (Patient Reported Outcome) instruments should and be completed according to the schedule in Table 5.1-5 and Table 5.1-6.
- On-study laboratory assessments (Cycle 2 and beyond) should be drawn within 72 hours prior to dosing.
- Adverse event assessments should be documented at each clinic visit.
- WOCBP must have a pregnancy test every 4 weeks, independent of study drug dosing.
- Study drug dosing may be delayed for toxicity.
- Each cycle during the treatment phase is expected to last 6 weeks, unless there are delays.
  - For the first 2 cycles;
    - Nivolumab and ipilimumab are administered every 3 weeks for 4 doses.
  - Starting Cycle 3:
    - Nivolumab is administered every 2 weeks.
- Treated subjects will be evaluated for response according to the RECIST 1.1 guidelines beginning 12 weeks (± 1 week) after randomization and continuing every 8 weeks (± 1 week) for the first 13 months, and then every 12 weeks (± 1 week) until disease progression or treatment discontinuation, whichever occurs later.
- The treatment phase ends when the subject is discontinued from study therapy.

Follow-Up

- Begins when the decision to discontinue a subject from study therapy is made (no further treatment with study therapy).
- Two follow-up visits will be conducted. Subjects will not provide PK or immunogenicity samples.
- Subjects who discontinue treatment for reasons other than tumor progression will continue to have tumor assessments beginning 12 weeks (± 1week) after randomization and continuing every 8 weeks (± 1 week) for the first 13 months from randomization, and every 12 weeks (± 1 week) thereafter until documented tumor progression.
- Subjects will be followed for drug-related toxicities until these toxicities resolve, return to baseline or are deemed irreversible. All adverse events will be documented for a

minimum of 100 days after last dose.

- After completion of the first two follow-up visits, subjects will be followed every 3 months for survival.
- PRO instruments will be completed according to the schedule in Table 5.1-7.

The total duration of the study from start of randomization to final analysis of OS is expected to be 61 months (20.5 months of accrual + 40.5 months of follow-up), assuming a fixed accrual rate of 57 subjects per month (including 40 IMDC poor/intermediate risk subjects per month). The enrollment will stop once approximately 820 intermediate/poor risk subjects have been randomized regardless of the number of favorable risk subjects randomized. Additional survival follow-up may continue for up to 5 years from the primary analysis of survival. The study will end once survival follow-up has concluded.

3.1.1 Nivolumab combined with Ipilimumab Crossover Extension Phase (Only for Subjects Originally Randomized to Arm B)

Inclusion Criteria for the Crossover Extension Phase

1) Signed Written Informed Consent

a) Subjects must have signed and dated an IRB/IEC approved written Informed Consent Form in accordance with regulatory and institutional guidelines. This must be obtained before the performance of any protocol-related procedures that are not part of normal subject care.

b) Subjects must be willing and able to comply with scheduled visits, treatment schedule, laboratory tests and other requirements of the study.

2) Target Population

a) Subjects previously randomized to sunitinib treatment (Arm B) who were classified as either intermediate or poor risk per IMDC prognostic score prior to randomization.

b) Prior anti-cancer therapy, including sunitinib and palliative radiotherapy, must have been completed at least 14 days prior to first dose of nivolumab combined with ipilimumab.

c) All toxicities attributed to prior anti-cancer therapy other than alopecia and fatigue must have resolved to Grade 1 or baseline prior to first dose of nivolumab combined with ipilimumab.

d) KPS of at least 70% (See Appendix 2)

e) Laboratory values must meet the following criteria and should be obtained within 14 days prior to first dose of nivolumab combined with ipilimumab:

i) WBC $\geq 2000/\mu L$

ii) Neutrophils $\geq 1500/\mu L$

iii) Platelets $\geq 100 \times 10^3/\mu L$

iv) Hemoglobin ≥ 9.0 g/dL

v) Serum creatinine ≤ 1.5 x ULN or creatinine clearance (CrCl) ≥ 40 mL/minute

(using Cockcroft/Gault formula):

$$\text{Female CrCl} = \frac{(140 - \text{age in years}) \times \text{weight in kg} \times 0.85}{72 \times \text{serum creatinine in mg/dL}}$$

$$\text{Male CrCl} = \frac{(140 - \text{age in years}) \times \text{weight in kg} \times 1.00}{72 \times \text{serum creatinine in mg/dL}}$$

vi) AST/ALT ≤ 3.0 x ULN (≤5 x ULN for subjects with liver metastases)
vii) Total Bilirubin ≤ 1.5 x ULN (except subjects with Gilbert Syndrome, who can have total bilirubin < 3.0 mg/dL.

Exclusion Criteria for the Crossover Extension Phase

1) Medical History and Concurrent Diseases

a) Subjects with active, known or suspected autoimmune disease. Subjects with type I diabetes mellitus, hypothyroidism only requiring hormone replacement, skin disorders (such as vitiligo, psoriasis) not requiring systemic treatment, or conditions not expected to recur in the absence of an external trigger are permitted to enroll.

b) Subjects with a condition requiring systemic treatment with either corticosteroids (> 10 mg daily prednisone equivalent) or other immunosuppressive medications within 14 days prior to the first dose of nivolumab combined with ipilimumab. Corticosteroids with minimal systemic absorption (for example, topical, inhalational, or as specified in Section 3.4.3) and adrenal replacement steroid doses > 10 mg daily prednisone or equivalent are permitted in the absence of active autoimmune disease.

c) Subjects must have recovered from the effects of major surgery or significant traumatic injury at least 28 days prior to the first dose of nivolumab combined with ipilimumab.

d) For Arm B subjects receiving sunitinib treatment or in the follow-up phase who have not received any subsequent systemic therapy at the time of Amendment 14: treatment with any subsequent systemic anticancer therapy.

e) For Arm B subjects in the follow-up phase who are receiving or have received any subsequent systemic anticancer therapy at the time of Amendment 14: treatment with any additional line of subsequent systemic anticancer therapy beyond the one being given or last received at the time of Amendment 14

f) Prior treatment with an anti-PD-1,anti-PD-L1, anti-PD-L2, anti-CD137, or anti-CTLA-4 antibody, or any other antibody or drug specifically targeting T-cell co-stimulation or checkpoint pathways.

g) Uncontrolled adrenal insufficiency.

2) Physical and Laboratory Test Findings

a) Known history of testing positive for human immunodeficiency virus (HIV) or known acquired immunodeficiency syndrome (AIDS).

b) Known positive test for Hepatitis B virus or Hepatitis C virus indicating acute or chronic infection.

3.2 Post Study Access to Therapy

At the conclusion of the study, subjects who continue to demonstrate clinical benefit and subjects who have no exceeded 2 years of study treatment in the crossover extension phase will be eligible to receive BMS supplied study drug. Study drug will be provided via an extension of the study, a rollover study requiring approval by responsible health authority and ethics committee or through another mechanism at the discretion of BMS. BMS reserves the right to terminate access to BMS supplied study drug if any of the following occur: a) the marketing application is rejected by responsible health authority; b) the study is terminated due to safety concerns; c) the subject can obtain medication from a government sponsored or private health program; or d) therapeutic alternatives become available in the local market.

3.3 Study Population

For entry into the study, the following criteria MUST be met.

3.3.1 Inclusion Criteria

1. Signed Written Informed Consent

a) Subjects must have signed and dated an IRB/IEC approved written informed consent form in accordance with regulatory and institutional guidelines. This must be obtained before the performance of any protocol related procedures that are not part of normal subject care.

b) Subjects must be willing and able to comply with scheduled visits, treatment schedule, laboratory testing, and other requirements of the study.

2. Target Population

a) Histological confirmation of RCC with a clear-cell component.

b) Advanced (not amenable to curative surgery or radiation therapy) or metastatic (AJCC Stage IV) RCC

c) No prior systemic therapy for RCC with the following exception:

i) One prior adjuvant or neoadjuvant therapy for completely resectable RCC if such therapy did not include an agent that targets VEGF or VEGF receptors and if recurrence occurred at least 6 months after the last dose of adjuvant or neoadjuvant therapy.

d) KPS of at least 70% (See Appendix 2)

e) Measurable disease as per RECIST v1.1 (See Appendix 3)

f) Tumor tissue (FFPE archival or recent acquisition) must be received by the central vendor (block or unstained slides) in order to randomize a subject to study treatment. (Note: Fine Needle Aspiration [FNA] and bone metastases samples are not acceptable for submission).

g) Patients with favorable, intermediate and poor risk categories will be eligible for the study. Patients must be categorized according to favorable versus intermediate/poor risk status at registration.

To be eligible for the intermediate or poor-risk cohort, at least one of the following prognostic factors as per the International Metastatic RCC Database Consortium (IMDC) criteria must be present:

i) KPS equal to 70%

ii) Less than 1 year from diagnosis to randomization

iii) Hemoglobin less than the lower limit of normal (LLN)

iv) Corrected calcium concentration greater than 10 mg/dL (Appendix 1)

v) Absolute neutrophil count greater than the ULN

vi) Platelet count greater than the ULN

If none of the above factors are present, subjects are only eligible for the favorable-risk cohort. The favorable-risk cohort may close to enrollment earlier than the intermediate- or poor-risk cohort.

3. Age and Reproductive Status

a) Males and Females, ≥ 18 years of age

b) Women of childbearing potential (WOCBP) must have a negative serum or urine pregnancy test (minimum sensitivity 25 IU/L or equivalent units of HCG) within 24 hours prior to the start of study drug.

c) Women must not be breastfeeding

d) Women of childbearing potential (WOCBP) must agree to follow instructions for method(s) of contraception for a period of 30 days (duration of ovulatory cycle) plus the time required for the investigational drug to undergo approximately five half-lives. The terminal half-life of the active metabolite of sunitinib is up to 110 hours.

i) WOCBP randomized to receive nivolumab + ipilimumab should use an adequate method to avoid pregnancy for 23 weeks (30 days plus the time required for nivolumab to undergo approximately five half-lives) after the last dose of investigational drug.

ii) WOCBP randomized to receive sunitinib should use an adequate method to avoid pregnancy for 8 weeks (30 days plus the time required for the active metabolite of sunitinib to undergo five half-lives)

e) Males who are sexually active with WOCBP must agree to follow instructions for method(s) of contraception for a period of 90 days (duration of sperm turnover) plus the

time required for the investigational drug to undergo approximately five half-lives. The terminal half-life of the active metabolite of sunitinib is up to 110 hours.

i) Males randomized to receive nivolumab combined with ipilimumab who are sexually active with WOCBP must continue contraception for 31 weeks (90 days plus the time required for nivolumab to undergo approximately five half-lives) after the last dose of investigational drug.

ii) Males randomized to receive sunitinib who are sexually active with WOCBP must continue contraception for 16 weeks (90 days plus the time required for the active metabolite of sunitinib to undergo five half-lives) after the last dose of investigational drug.

f) Azoospermic males and WOCBP who are continuously not heterosexually active are exempt from contraceptive requirements. However they must still undergo pregnancy testing as described in this section.

Investigators shall counsel WOCBP and male subjects who are sexually active with WOCBP on the importance of pregnancy prevention and the implications of an unexpected pregnancy Investigators shall advise WOCBP and male subjects who are sexually active with WOCBP on the use of highly effective methods of contraception. Highly effective methods of contraception have a failure rate of < 1% when used consistently and correctly.

At a minimum, subjects must agree to the use of one highly effective method as listed below:

HIGHLY EFFECTIVE METHODS OF CONTRACEPTION

- Hormonal methods of contraception including combined oral contraceptive pills, vaginal ring, injectables, implants and intrauterine devices (IUDs) such as Mirena® by WOCBP subject or male subject's WOCBP partner Female partners of male subjects participating in the study may use hormone based contraceptives as one of the acceptable methods of contraception since they will not be receiving study drug
- Nonhormonal IUDs, such as ParaGard®
- Tubal ligation
- Vasectomy
- Complete Abstinence*

*Complete abstinence is defined as complete avoidance of heterosexual intercourse and is an acceptable form of contraception for all study drugs. Subjects who choose complete abstinence are not required to use a second method of contraception, but female subjects must continue to have pregnancy tests. Acceptable alternate methods of highly effective contraception must be discussed in the event that the subject chooses to forego complete abstinence.

LESS EFFECTIVE (UNACCETPABLE) METHODS OF CONTRACEPTION

- Diaphragm with spermicide

- Cervical cap with spermicide
- Vaginal sponge
- Male Condom with or without spermicide
- Progestin only pills by WOCBP subject or male subject's WOCBP partner
- Female Condom*
- Periodic abstinence (calendar, symptothermal, post-ovulation methods)
- Withdrawal (coitus interruptus)
- Spermicide only
- Lactation amenorrhea method (LAM)

* A male and female condom must not be used together.

3.3.2 Exclusion Criteria

1. Target Disease Exceptions

a) Any history of or current CNS metastases. Baseline imaging of the brain by MRI (preferred) or CT scan is required within 28 days prior to randomization.

2. Medical History and Concurrent Diseases

a) Prior systemic treatment with VEGF or VEGF receptor targeted therapy (including, but not limited to, sunitinib, pazopanib, axitinib, tivozanib, and bevacizumab).

b) Prior treatment with an anti-PD-1, anti-PD-L1, anti-PD-L2, anti-CD137, or anti-CTLA-4 antibody, or any other antibody or drug specifically targeting T-cell co-stimulation or checkpoint pathways.

c) Any active or recent history of a known or suspected autoimmune disease or recent history of a syndrome that required systemic corticosteroids (> 10 mg daily prednisone equivalent) or immunosuppressive medications except for syndromes which would not be expected to recur in the absence of an external trigger. Subjects with vitiligo or type I diabetes mellitus or residual hypothyroidism due to autoimmune thyroiditis only requiring hormone replacement are permitted to enroll.

d) Any condition requiring systemic treatment with corticosteroids (> 10 mg daily prednisone equivalents) or other immunosuppressive medications within 14 days prior to first dose of study drug. Inhaled steroids and adrenal replacement steroid doses > 10 mg daily prednisone equivalents are permitted in the absence of active autoimmune disease.

e) Uncontrolled adrenal insufficiency.

f) Ongoing symptomatic cardiac dysrhythmias, uncontrolled atrial fibrillation, or prolongation of the Fridericia corrected QT (QTcF) interval defined as > 450 msec for males and > 470 msec for females, where $QTcF = QT / \sqrt[3]{RR}$

g) Poorly controlled hypertension (defined as systolic blood pressure (SBP) of ≥ 150 mmHg or diastolic blood pressure (DBP) of ≥ 90 mmHg), despite antihypertensive therapy.

h) History of any of the following cardiovascular conditions within 12 months of enrollment: cardiac angioplasty or stenting, myocardial infarction, unstable angina, coronary artery by-pass graft surgery, symptomatic peripheral vascular disease, class III or IV congestive heart failure, as defined by the New York Heart Association.

i) History of cerebrovascular accident including transient ischemic attack within the past 12 months.

j) History of deep vein thrombosis (DVT) unless adequately treated with low molecular weight heparin

k) History of pulmonary embolism within the past 6 months unless stable, asymptomatic, and treated with low molecular weight heparin for at least 6 weeks.

l) History of abdominal fistula, gastrointestinal perforation, or intra-abdominal abscess within the past 6 months.

m) Serious, non-healing wound or ulcer.

n) Evidence of active bleeding or bleeding susceptibility; or medically significant hemorrhage within prior 30 days.

o) Any requirement for anti-coagulation, except for low molecular weight heparin.

p) Prior malignancy active within the previous 3 years except for locally curable cancers that have been apparently cured, such as basal or squamous cell skin cancer, superficial bladder cancer, or carcinoma in situ of the prostate, cervix, or breast.

q) Known history of testing positive for human immunodeficiency virus (HIV) or known acquired immunodeficiency syndrome (AIDS).

r) Any positive test for hepatitis B or hepatitis C virus indicating acute or chronic infection.

s) Known medical condition (eg, a condition associated with diarrhea or acute diverticulitis) that, in the investigator's opinion, would increase the risk associated with study participation or study drug administration or interfere with the interpretation of safety results.

t) Major surgery (eg, nephrectomy) less than 28 days prior to the first dose of study drug.

u) Anti-cancer therapy less than 28 days prior to the first dose of study drug or palliative, focal radiation therapy less than 14 days prior to the first dose of study drug.

v) Presence of any toxicities attributed to prior anti-cancer therapy, other than alopecia, that have not resolved to Grade 1 (NCI CTCAE v4) or baseline before administration of study drug.

w) Receiving concomitant CYP3A4 inducers or strong CYP3A4 inhibitors (See Appendix 4).

x) Impairment of gastrointestinal function or gastrointestinal disease that may significantly alter the absorption of sunitinib (eg, malabsorptive disorder, ulcerative disease, uncontrolled nausea, vomiting, diarrhea, or small bowel resection).

3. Physical and Laboratory Test Findings

a) Left ventricular ejection fraction (LVEF) less than the LLN as assessed by echocardiography or multigated acquisition (MUGA) scan.

b) Any of the following laboratory test findings:

i) WBC < 2,000/mm^3

ii) Neutrophils < 1,500/mm^3

iii) Platelets < 100,000/mm^3

iv) AST or ALT > 3 x ULN (> 5 x ULN if liver metastases are present)

v) Total Bilirubin > 1.5 x ULN (except subjects with Gilbert Syndrome, who can have total bilirubin < 3.0 mg/dL)

vi) Serum creatinine > 1.5 x upper limit of normal (ULN) or creatinine clearance < 40 mL/min (measured or calculated by Cockroft-Gault formula):

$$\text{Female CrCl} = \frac{(140 - \text{age in years}) \times \text{weight in kg} \times 0.85}{72 \times \text{serum creatinine in mg/dL}}$$

$$\text{Male CrCl} = \frac{(140 - \text{age in years}) \times \text{weight in kg} \times 1.00}{72 \times \text{serum creatinine in mg/dL}}$$

4. Allergies and Adverse Drug Reaction

a) History of severe hypersensitivity reaction to any monoclonal antibody.

5. Other Exclusion Criteria

a) Prisoners or subjects who are involuntarily incarcerated

b) Subjects who are compulsorily detained for treatment of either a psychiatric or physical (eg, infectious disease) illness

Eligibility criteria for this study have been carefully considered to ensure the safety of the study subjects and that the results of the study can be used. It is imperative that subjects fully meet all eligibility criteria.

Subject Re-enrollment: This study permits the re-enrollment of a subject that has discontinued the study as a pre-treatment failure (ie, subject has not been randomized / has not been treated). If re-enrolled, the subject must be re-consented.

3.3.3 Women of Childbearing Potential

A Women of childbearing potential (WOCBP) is defined as any female who has experienced menarche and who has not undergone surgical sterilization (hysterectomy or bilateral oophorectomy) and is not postmenopausal. Menopause is defined as 12 months of amenorrhea in a woman over age 45 years in the absence of other biological or physiological causes. In

addition, females under the age of 55 years must have a serum follicle stimulating hormone, (FSH) level > 40mIU/mL to confirm menopause.

3.4 Concomitant Treatments

3.4.1 Prohibited and/or Restricted Treatments

The following medications are prohibited during the study:

- Immunosuppressive agents (except to treat a drug-related adverse event)
- Systemic corticosteroids > 10 mg daily prednisone equivalent (except as stated in Section 3.4.3 below or to treat a drug-related adverse event).
- Any concurrent antineoplastic therapy (ie, chemotherapy, hormonal therapy, immunotherapy, radiation therapy except for palliative radiation therapy (Section 3.4.2), surgical resection except for palliative surgical resection (Section 3.4.2), or standard or investigational agents for treatment of cancer).

Supportive care for disease-related symptoms may be offered to all subjects on the trial.

Note: Initiation of CYP3A4 inducers and inhibitors (Appendix 4) is not prohibited after dosing has begun, however Arm B subjects should follow sunitinib dose modification recommendations (Section 4.5.3.2).

3.4.2 Other Restrictions and Precautions

Palliative (limited-field) radiation therapy and palliative surgical resection are permitted, if the following criteria are met:

- The subject will be considered to have progressed at the time of palliative therapy and must meet criteria to continue with treatment beyond progression (Section 4.5.7)
- The case is discussed with the BMS Medical Monitor or Study Director.

3.4.3 Permitted Therapy

Subjects are permitted to use topical, ocular, intra-articular, intranasal, and inhalational corticosteroids (with minimal systemic absorption). Physiologic replacement doses of systemic corticosteroids are permitted, even if > 10 mg/day prednisone equivalents. A brief course of corticosteroids for prophylaxis (eg, contrast dye allergy) or for treatment of non-autoimmune conditions (eg, delayed-type hypersensitivity reaction caused by contact allergen) is permitted.

Concomitant medications are recorded at baseline and throughout the treatment phase of the study in the appropriate section of the CRF. All medications (prescriptions or over the counter medications) continued at the start of the study or started during the study and different from the study drug must be documented in the concomitant therapy section of the CRF.

3.5 Discontinuation of Subjects following any Treatment with Study Drug

Subjects MUST discontinue investigational product (and non-investigational product at the discretion of the investigator) for any of the following reasons:

- Subject's request to stop study treatment
- Any clinical adverse event (AE), laboratory abnormality or intercurrent illness which, in the opinion of the investigator, indicates that continued participation in the study is not in the best interest of the subject
- Protocol defined disease progression (subjects may be permitted to continue treatment beyond initial disease progression see Section 4.5.7)
- Termination of the study by Bristol-Myers Squibb (BMS)
- Loss of ability to freely provide consent through imprisonment or involuntarily incarceration for treatment of either a psychiatric or physical (eg, infectious disease) illness
- Additional protocol specified reasons for discontinuation (see Section 4.5.5)

In the case of pregnancy, the investigator must immediately notify the BMS Medical Monitor/designee of this event. In most cases, the study drug will be permanently discontinued in an appropriate manner. If the investigator determines a possible favorable benefit/risk ratio that warrants continuation of study drug, a discussion between the investigator and the BMS Medical Monitor/designee must occur.

All subjects who discontinue study drug should comply with protocol specified follow-up procedures as outlined in Section 5. The only exception to this requirement is when a subject withdraws consent for all study procedures including post-treatment study follow-up or loses the ability to consent freely (ie, is imprisoned or involuntarily incarcerated for the treatment of either a psychiatric or physical illness).

If study drug is discontinued prior to the subject's completion of the study, the reason for the discontinuation must be documented in the subject's medical records and entered on the appropriate case report form (CRF) page.

3.6 Post Study Drug Study Follow up

In this study, overall survival is a key endpoint of the study. Post study follow-up is of critical importance and is essential to preserving subject safety and the integrity of the study. Subjects who discontinue study drug must continue to be followed for collection of outcome and/or survival follow-up data as required and in line with Section 5 until death or the conclusion of the study.

3.6.1 Withdrawal of Consent

Subjects who request to discontinue study drug will remain in the study and must continue to be followed for protocol specified follow-up procedures. The only exception to this is when a subject specifically withdraws consent for any further contact with him/her or persons previously authorized by subject to provide this information. Subjects should notify the investigator of the

decision to withdraw consent from future follow-up **in writing**, whenever possible. The withdrawal of consent should be explained in detail in the medical records by the investigator, as to whether the withdrawal is from further treatment with study drug only or also from study procedures and/or post treatment study follow-up, and entered on the appropriate CRF page. In the event that vital status (whether the subject is alive or dead) is being measured, publicly available information should be used to determine vital status only as appropriately directed in accordance with local law.

3.6.2 Lost to Follow-Up

All reasonable efforts must be made to locate subjects to determine and report their ongoing status. This includes follow-up with persons authorized by the subject as noted above. Lost to follow-up is defined by the inability to reach the subject after a minimum of three documented phone calls, faxes, or emails as well as lack of response by subject to one registered mail letter. All attempts should be documented in the subject's medical records. If it is determined that the subject has died, the site will use permissible local methods to obtain the date and cause of death.

If investigator's use of third-party representative to assist in the follow-up portion of the study has been included in the subject's informed consent, then the investigator may use a Sponsor-retained third-party representative to assist site staff with obtaining subject's contact information or other public vital status data necessary to complete the follow-up portion of the study. The site staff and representative will consult publicly available sources, such as public health registries and databases, in order to obtain updated contact information. If after all attempts, the subject remains lost to follow-up, then the last known alive date as determined by the investigator should be reported and documented in the subject's medical records.

4 STUDY DRUG

Study drug includes both Investigational [Medicinal] Product (IP/IMP) and Non-investigational [Medicinal] Product (Non-IP/Non-IMP) and can consist of the following:

- All products, active or placebo, being tested or used as a comparator in a clinical trial.
- Study required premedication, and
- Other drugs administered as part of the study that are critical to claims of efficacy (eg, background therapy, rescue medications)
- Diagnostic agents: (such as glucose for glucose challenge) given as part of the protocol requirements must also be included in the dosing data collection.

Table 4-1: Study Drugs for CA209214

Product Description / Class and Dosage Form	Potency	IP/ Non-IMP	Blinded or Open Label	Packaging/ Appearance	Storage Conditions (per label)
BMS-936558-01 (Nivolumab) Solution for Injection[a]	100 mg (10 mg/mL)	IP	Open Label	10 mL per vial (5 or 10 vials/carton)	Store at 2° - 8 °C. Protect from light and freezing.
Ipilimumab Solution for Injection	200 mg (5mg/mL)	IP	Open Label	40 mL per vial (4 vials/carton)	Store at 2° - 8 °C. Protect from light and freezing.
Sunitinib Malate Capsule[b]	12.5 mg	IP	Open Label	28 capsules per wallet card or 30 capsules per bottle	Store at 15° - 25 °C.

[a] May be labeled as "BMS-936558-01" or "Nivolumab"

[b] Sunitinib may be obtained by the investigational sites in certain countries as local commercial product (which may be available as a different potency/package size than listed above) if local regulations allow and agreed to by BMS.

Premedications or medications used to treat infusion-related reactions should be sourced by the investigative sites if available and permitted by local regulations. Solutions used as diluent (ie, 0.9% Sodium Chloride Injection or 5% Dextrose Injection) should also be sourced by investigative sites if available and permitted by local regulations.

4.1 Investigational Product

An investigational product, also known as investigational medicinal product in some regions, is defined a pharmaceutical form of an active substance or placebo being tested or used as a reference in a clinical study, including products already with a marketing authorization but used or assembled (formulated or packaged) differently than the authorized form, or used for an unauthorized indication, or when used to gain further information about the authorized form.

The investigational product should be stored in a secure area according to local regulations. It is the responsibility of the investigator to ensure that investigational product is only dispensed to study subjects. The investigational product must be dispensed only from official study sites by authorized personnel according to local regulations.

4.2 Non-investigational Product

Other medications used as support or escape medication for preventative, diagnostic, or therapeutic reasons, as components of the standard of care for a given diagnosis, may be considered as non-investigational products.

In this protocol, non-investigational product(s) is/are: medications used to treat nivolumab infusion-related reactions (eg, steroids, anti-emetics); these non-investigational products should be sourced by the investigator sites if available and permitted by local regulations.

4.3 Storage and Dispensing

The product storage manager should ensure that the study drug is stored in accordance with the environmental conditions (temperature, light, and humidity) as determined by BMS. If concerns regarding the quality or appearance of the study drug arise, the study drug should not be dispensed and contact BMS immediately.

Study drug not supplied by BMS will be stored in accordance with the package insert.

Investigational product documentation (whether supplied by BMS or not) must be maintained that includes all processes required to ensure drug is accurately administered. This includes documentation of drug storage, administration and, as applicable, storage temperatures, reconstitution, and use of required processes (eg, required diluents, administration sets).

Infusion-related supplies (eg, IV bags, in-line filters, 0.9% NaCl solution) will not be supplied by the sponsor and should be purchased locally if permitted by local regulations.

For nivolumab and ipilimumab, please refer to the current version of the Investigator Brochures and/or pharmacy reference sheets for complete storage, handling, dispensing, and infusion information.

Nivolumab is to be administered as an approximately 60-minute IV infusion. At the end of the infusion, flush the line with a sufficient quantity of normal saline. Ipilimumab is to be administered as an approximately 30-minute IV infusion. At the end of the infusion, flush the line with a sufficient quantity of normal saline or 5% dextrose solution. When both study drugs are to be administered on the same day, separate infusion bags and filters must be used for each infusion. Nivolumab is to be administered first. The nivolumab infusion must be promptly

followed by a saline flush to clear the line of nivolumab before starting the ipilimumab infusion. The second infusion will always be ipilimumab, and will start at least 30 minutes after completion of the nivolumab infusion.

For sunitinib, please refer to the appropriate SmPC or package insert and/or pharmacy reference sheets for complete storage, handling, and administration information.

4.4 Method of Assigning Subject Identification

CA209214 is a randomized, open-label study. After the subject's initial eligibility is established and informed consent has been obtained, the subject must be enrolled into the study by calling an interactive voice response system (IVRS) to obtain the subject number. 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 and the required tumor tissue has been received by the central laboratory will be ready to be randomized through the IVRS. The following information is required for subject randomization:

- Subject number
- Date of birth
- KPS less than 80 (ie, KPS equal to 70)? Yes/No
- Less than 1 year from initial diagnosis of RCC (eg, nephrectomy or first diagnostic biopsy) to randomization? Yes/No
- Hemoglobin less than the LLN? Yes/No
- Corrected calcium greater than 10 mg/dL? Yes/No (Appendix 1)
- Absolute neutrophil count greater than the ULN? Yes/No
- Platelet count greater than the ULN? Yes/No

Subjects meeting all eligibility criteria will be randomized in a 1:1 ratio to Arm A (nivolumab combined with ipilimumab) or Arm B (sunitinib), stratified by the following factors:

- IMDC Prognostic Score (Total Number of IMDC Adverse Prognostic Factors Present)
  - 0
  - 1-2
  - 3-6

- Region
  - US
  - Canada/W Europe/N Europe
  - Rest of World

No more than 820 intermediate/poor risk subjects and 250 favorable risk subjects will be randomized in this study. These restrictions will be implemented via the IVRS system

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

Amendment 14 Update:

IVRS will be amended to allow all subjects in the poor or intermediate cohorts previously randomized to Arm B (sunitinib) to receive treatment with nivolumab combined with ipilimumab. The IVRS will assign the nivolumab combined with ipilimumab treatment for all subjects eligible for the crossover extension phase. Procedural information will be provided in a separate document.

Subjects currently randomized to Arm B (sunitinib) may also continue obtaining treatment, as previously done so through the IVRS, as long as they are continuing to derive benefit from sunitinib in the judgement of the investigator.

4.5 Selection and Timing of Dose for Each Subject

The dosing schedule is detailed in Table 4.5-1 and Table 4.5-2.

Table 4.5-1: Dosing Schedule for Cycle 1 and Cycle 2

1 Cycle = 6 weeks						
	Week 1	**Week 2**	**Week 3**	**Week 4**	**Week 5**	**Week 6**
Arm A (Nivolumab 3 mg/kg + Ipilimumab 1 mg/kg)	Day 1 3 mg/kg Nivolumab + 1 mg/kg Ipilimumab			Day 1 3 mg/kg Nivolumab + 1 mg/kg Ipilimumab		
Arm B (Sunitinib)	Sunitinib 50 mg PO once daily x 4 weeks				2 weeks off	

Table 4.5-2: Dosing Schedule Cycle 3 and Beyond

1 Cycle = 6 weeks						
	Week 1	**Week 2**	**Week 3**	**Week 4**	**Week 5**	**Week 6**
Arm A (Nivolumab 3 mg/kg + Ipilimumab 1 mg/kg)	Day 1 3 mg/kg or 240mg flat dose Nivolumab		Day 1 3 mg/kg or 240mg flat dose Nivolumab		Day 1 3 mg/kg or 240mg flat dose Nivolumab	
Arm B (Sunitinib)	Sunitinib 50 mg PO once daily x 4 weeks				2 weeks off	

The first dose is to be administered within 3 days following randomization, except as noted in Table 5.1-2.

For Arm A (Nivolumab combined with Ipilimumab):

When nivolumab and ipilimumab are to be administered on the same day, separate infusion bags and filters must be used for each infusion. Nivolumab is to be administered first. The second infusion will always be ipilimumab, and will at least 30 minutes after completion of the nivolumab infusion.

Ipilimumab may be diluted in 0.9% Sodium Chloride Solution or 5% Dextrose solution. Nivolumab may be diluted in 0.9% Sodium Chloride Solution.

The dosing calculations for the 3mg/kg dosing should be based on the body weight, assessed at either the day of dosing, the start of each cycle, or the last recorded weight and if the subject's weight on the day of dosing differs by > 10% from the weight used to calculate the dose, then the dose must be recalculated. All doses should be rounded to the nearest milligram. There will be no dose modifications allowed.

During cycles 1 and 2, subjects may be dosed no less than 19 days from the previous dose of drug.

Starting from cycle 3, subjects may be dosed no less than 12 days from the previous dose of drug.

Treatment compliance will be monitored by drug accountability as well as the subject's medical record and eCRF.

Amendment 14 Update:

Subjects receiving nivolumab every 2 weeks at 3mg/kg as part of Arm A will have the option to switch to nivolumab at 240mg every 2 weeks as a 60 minutes IV infusion.

Subjects who enter the nivolumab combined with ipilimumab crossover extension phase from Arm B will receive a flat dose of 240mg every 2 weeks IV infusion upon completion of the combination phase with ipilimumab.

4.5.1 Antiemetic Premedications

Antiemetic premedications should not be routinely administered prior to dosing of drugs. See section 4.5.6 for premedication recommendations following a nivolumab or ipilimumab-related infusion reaction.

4.5.2 Dose Delay Criteria

Dose delay criteria apply for all drug-related adverse events (regardless of whether or not the event is attributed to nivolumab, ipilimumab, or both). All study drugs must be delayed until treatment can resume (see Section 4.5.4).

4.5.2.1 Dose Delay Criteria for Arm A (Nivolumab combined with Ipilimumab)

Nivolumab and ipilimumab administration should be delayed for the following:

- Any Grade ≥ 2 non-skin, drug-related adverse event, with the following exceptions:
  - Grade 2 drug-related fatigue or laboratory abnormalities do not require a treatment delay
- Any Grade 3 skin, drug-related adverse event
- Any Grade 3 drug-related laboratory abnormality, with the following exceptions for AST, ALT, or total bilirubin:
  - If a subject has a baseline AST, ALT, or total bilirubin that is within normal limits, delay dosing for drug-related Grade ≥ 2 toxicity
  - If a subject has baseline AST, ALT, or total bilirubin within the Grade 1 toxicity range, delay dosing for drug-related Grade ≥ 3 toxicity
- Grade 3 amylase or lipase abnormalities that are not associated with symptoms or clinical manifestations of pancreatitis do not require a dose delay. It is recommended to consult with the BMS Medical Monitor or Study Director for Grade 3 amylase or lipase abnormalities.
- Any adverse event, laboratory abnormality, or intercurrent illness which, in the judgment of the investigator, warrants delaying the dose of study medication.

During cycles 1 and 2, both nivolumab and ipilimumab must be delayed at the same time.

Because of the potential for clinically meaningful nivolumab or ipilimumab-related AEs requiring early recognition and prompt intervention, management algorithms have been developed for suspected pulmonary toxicity, GI, hepatotoxicity, endocrinopathy, skin toxicity, neurological toxicity and renal toxicity.

In order to standardize the management across of subjects on Arm A, for the overlapping adverse event management algorithms present in both the BMS-936558 (nivolumab) and ipilimumab IB (**GI, hepatic, and endocrine** algorithms), the recommendations are to follow the BMS-936558 (nivolumab) IB adverse event algorithms as opposed to the ipilimumab IB algorithms.

The algorithms recommended for utilization in CA209214 are included in Appendix 5.

4.5.2.2 Dose Delay Criteria for Arm B (Sunitinib)

Sunitinib dose delays should be based on instructions in the approved product label and should be considered for any severe or intolerable drug-related adverse events.

Within a cycle, missed doses of sunitinib should be skipped. Subjects should never be dosed during the 2-week off period of each 6-week cycle, even if treatment delays occurred earlier in the cycle and therapy is ready to be resumed. If treatment is delayed past the end of the 6-week cycle, the start of the next cycle should be delayed until treatment with sunitinib resumes.

Prior to resuming therapy after a dose delay, refer to Section 4.5.3.2 for dose reduction recommendations and Section 4.5.5.2 for discontinuation criteria.

For this protocol, the following sunitinib dose delay recommendations should be followed:

- Related Grade 1 and 2 toxicities do not require dose delay, with the following exceptions:
  - For related Grade 2 hemorrhage, delay until the toxicity resolves to ≤ Grade 1 or baseline.
  - Arterial thrombosis of any grade requires discontinuation (Section 4.5.5.2).
- For related Grade 3 non-hematological toxicity, delay until the toxicity resolves to ≤ Grade 1 or baseline, with the following exceptions:
  - Recurrent Grade 3 drug-related hemorrhage after dose reduction requires discontinuation (Section 4.5.5.2).
  - Related AST or ALT > 8 x ULN requires discontinuation (Section 4.5.5.2).
  - Concurrent AST or ALT > 3 x ULN and total bilirubin > 2 x ULN requires discontinuation (Section 4.5.5.2).
- For related Grade 3 hematologic toxicity, delay until the toxicity resolves to ≤ Grade 2 or baseline.
- For related Grade 4 non-hematological toxicity, reduce the dose by 12.5 mg, with the following exceptions:
  - Related Grade 4 symptomatic venous thrombosis requires discontinuation (Section 4.5.5.2).
  - Related Grade 4 cardiac disorder requires discontinuation (Section 4.5.5.2).
  - Amylase or lipase elevations do not require dose reduction if not accompanied by other evidence of pancreatitis.
- For related Grade 4 hematological toxicity, delay until the toxicity resolves to ≤ Grade 1 or baseline.

4.5.3 Dose Modifications

4.5.3.1 Dose Modifications for Arm A (Nivolumab combined with Ipilimumab)

Dose reductions or dose escalations of nivolumab or ipilimumab are not permitted, including for subjects enrolled in the crossover extension phase.

4.5.3.2 Dose Modifications for Arm B (Sunitinib)

Sunitinib Dose Reductions

Sunitinib dose reductions are permitted as per the approved product label for safety reasons or when a concomitant strong CYP3A4 inhibitor is needed (Appendix 4). Selection of an alternative concomitant medication with minimal or no enzyme inhibition potential is recommended whenever possible.

Dose reductions should occur in 12.5 mg decrements. No more than 2 dose reductions are allowed. If more than 2 dose reductions are necessary (ie, reduction to less than 25 mg daily), the subject must be permanently discontinued (Section 4.5.5).

At the time a dose reduction is considered, also refer to Section 4.5.2.2 for dose delays recommendations and Section 4.5.5.2 for discontinuation criteria.

For this protocol, the following sunitinib dose reduction recommendations should be followed:

- Related Grade 1 and 2 toxicities do not require dose reduction, with the following exception:
  - Arterial thrombosis of any grade requires discontinuation (Section 4.5.5.2).
- For related Grade 3 non-hematological toxicity, reduce the dose by 12.5 mg at the discretion of the investigator, with the following exceptions:
  - Recurrent Grade 3 drug-related hemorrhage after dose reduction requires discontinuation (Section 4.5.5.2).
  - Related AST or ALT > 8 x ULN requires discontinuation (Section 4.5.5.2).
  - Concurrent AST or ALT > 3 x ULN and total bilirubin > 2 x ULN requires discontinuation (Section 4.5.5.2).
- For related Grade 3 hematologic toxicity, dose reduction is not required except for recurrent Grade 3 neutropenia or thrombocytopenia.
- For related Grade 4 non-hematological toxicity, reduce the dose by 12.5 mg, with the following exceptions:
  - Related Grade 4 hemorrhage requires discontinuation.
  - Related Grade 4 symptomatic venous thrombosis requires discontinuation (Section 4.5.5.2).
  - Related Grade 4 cardiac disorder requires discontinuation (Section 4.5.5.2).
  - Amylase or lipase elevations do not require dose reduction if not accompanied by other evidence of pancreatitis.
- For related Grade 4 hematological toxicity, reduce the dose by 12.5 mg.

Sunitinib Dose Escalations

Sunitinib dose escalations are permitted as per the approved product label when a concomitant CYP3A4 inducer is needed (Appendix 4). Selection of an alternative concomitant medication with minimal or no enzyme induction potential is recommended whenever possible.

4.5.4 Criteria to Resume Treatment

4.5.4.1 Criteria to Resume Treatment on Arm A (Nivolumab combined with Ipilimumab)

Missed doses of nivolumab and/or ipilimumab should be administered as soon as the subject meets criteria to resume treatment. If a dose has been missed, the subject should not wait until the next scheduled dosing date.

Subjects may resume treatment with study drug when the drug-related AE(s) resolve to Grade ≤ 1 or baseline value, with the following exceptions:

- Subjects may resume treatment in the presence of Grade 2 fatigue.
- Subjects who have not experienced a Grade 3 drug-related skin AE may resume treatment in the presence of Grade 2 skin toxicity.
- Subjects with baseline Grade 1 AST/ALT or total bilirubin who require dose delays for reasons other than a 2-grade shift in AST/ALT or total bilirubin may resume treatment in the presence of Grade 2 AST/ALT OR total bilirubin.
- Subjects with combined Grade 2 AST/ALT AND total bilirubin values meeting discontinuation parameters (Section 4.5.5) should have treatment permanently discontinued.
- Drug-related pulmonary toxicity, diarrhea, or colitis, must have resolved to baseline before treatment is resumed.
- Drug-related endocrinopathies adequately controlled with only physiologic hormone replacement may resume treatment.

If treatment is delayed > 6 weeks, the subject must be permanently discontinued from study therapy, except as specified in Section 4.5.5.

During cycles 1 and 2, both nivolumab and ipilimumab must be resumed on the same day. All four doses of nivolumab combined with ipilimumab must be given prior to beginning nivolumab monotherapy (cycle 3 and beyond).

If the subject is unable to resume both nivolumab and ipilimumab, permanent discontinuation is required (Section 4.5.5).

4.5.4.2 Criteria to Resume Treatment on Arm B (Sunitinib)

Within a cycle, missed doses of sunitinib should be skipped and not replaced. Subjects should never be dosed during the 2-week off period of each 6-week cycle, even if treatment delays occurred earlier in the cycle and therapy is ready to be resumed. If treatment is delayed past the end of the 6-week cycle, the start of the next cycle should be delayed until treatment with sunitinib resumes.

If treatment is delayed > 6 weeks for any reason, the subject must be permanently discontinued from study therapy, except in cases where permission to resume treatment is granted by the BMS Medical Monitor or Study Director.

Criteria to resume treatment are dependent on the reason for delay and are included in Section 4.5.2.2.

4.5.5 Discontinuation Criteria

4.5.5.1 Discontinuation Criteria for Arm A (Nivolumab combined with Ipilimumab)

Treatment with nivolumab and ipilimumab should be permanently discontinued for any of the following:

- Any Grade 2 drug-related uveitis or eye pain or blurred vision that does not respond to topical therapy and does not improve to Grade 1 severity within the re-treatment period OR requires systemic treatment
- Any Grade 3 non-skin, drug-related adverse event lasting > 7 days, with the following exceptions for drug-related laboratory abnormalities, uveitis, pneumonitis, bronchospasm, diarrhea, colitis, neurologic toxicity, hypersensitivity reactions, and infusion reactions:
  - Grade 3 drug-related uveitis, pneumonitis, bronchospasm, diarrhea, colitis, neurologic toxicity, hypersensitivity reaction, or infusion reaction of any duration requires discontinuation
  - Grade 3 drug-related laboratory abnormalities do not require treatment discontinuation except:
    - Grade 3 drug-related thrombocytopenia > 7 days or associated with bleeding requires discontinuation
    - Any drug-related liver function test (LFT) abnormality that meets the following criteria require discontinuation:
      - AST or ALT > 8 x ULN
      - Total bilirubin > 5 x ULN
      - Concurrent AST or ALT > 3 x ULN and total bilirubin > 2 x ULN
- Any Grade 4 drug-related adverse event or laboratory abnormality, except for the following events which do not require discontinuation:
  - Grade 4 amylase or lipase abnormalities that are not associated with symptoms or clinical manifestations of pancreatitis. It is recommended to consult with the BMS Medical Monitor or Study Director for Grade 4 amylase or lipase abnormalities.
  - Isolated Grade 4 electrolyte imbalances/abnormalities that are not associated with clinical sequelae and are corrected with supplementation/appropriate management within 72 hours of their onset
- Any dosing interruption lasting > 6 weeks unless the BMS Medical Monitor or Study Director is consulted and agrees with the rationale for resuming therapy after a delay > 6 weeks. Note that tumor assessments should continue as per protocol even if dosing is interrupted.
- Any adverse event, laboratory abnormality, or intercurrent illness which, in the judgment of the Investigator, presents a substantial clinical risk to the subject with continued nivolumab or ipilimumab dosing

During cycles 1 and 2, both nivolumab and ipilimumab must be discontinued at the same time.

4.5.5.2 Discontinuation Criteria for Arm B (Sunitinib)

Treatment with sunitinib should be permanently discontinued for any of the following:

- Any requirement for more than 2 sunitinib dose reductions.
- Any Grade drug-related arterial thrombosis.
- Grade 4 drug-related hemorrhage or recurrent Grade 3 drug-related hemorrhage after dose reduction.
- Grade 4 drug-related symptomatic venous thrombosis.
- Grade 4 drug-related cardiac toxicity.
- Two or more symptomatic episodes of hypertension despite modification of antihypertensive medication(s) and reduction of sunitinib dose.
- Any drug-related liver function test (LFT) abnormality that meets the following criteria require discontinuation:
    1. AST or ALT > 8 x ULN.
    2. Concurrent AST or ALT > 3 x ULN and total bilirubin >2 x ULN.
- Any dosing interruption lasting > 6 weeks unless the BMS Medical Monitor or Study Director is consulted and agrees with the rationale for resuming therapy after a delay > 6 weeks. Note that tumor assessments should continue as per protocol even if dosing is interrupted.
- Any adverse event, laboratory abnormality, or intercurrent illness which, in the judgment of the Investigator, presents a substantial clinical risk to the subject with continued sunitinib dosing.

4.5.6 Treatment of Nivolumab or Ipilimumab-Related Infusion Reactions

Since nivolumab and ipilimumab contain only human immunoglobulin protein sequences, each is unlikely to be immunogenic and induce infusion or hypersensitivity reactions. However, if such a reaction were to occur, it might manifest with fever, chills, rigors, headache, rash, pruritis, arthralgias, hypo- or hypertension, bronchospasm, or other symptoms. All Grade 3 or 4 infusion reactions should be reported within 24 hours to the BMS Medical Monitor or Study Director and reported as an SAE if criteria are met. Infusion reactions should be graded according to NCI CTCAE (version 4.0) guidelines.

Treatment recommendations are provided below and may be modified based on local treatment standards and guidelines as appropriate:

For Grade 1 symptoms: (Mild reaction; infusion interruption not indicated; intervention not indicated)

Remain at bedside and monitor subject until recovery from symptoms. The following prophylactic premedications are recommended for future infusions: diphenhydramine 50 mg

(or equivalent) and/or paracetamol 325 to 1000 mg (acetaminophen) at least 30 minutes before additional nivolumab administrations.

For Grade 2 symptoms: (Moderate reaction requires therapy or infusion interruption but responds promptly to symptomatic treatment [eg, antihistamines, non-steroidal anti-inflammatory drugs, narcotics, corticosteroids, bronchodilators, IV fluids]; prophylactic medications indicated for ≤ 24 hours).

Stop the nivolumab or ipilimumab infusion, begin an IV infusion of normal saline, and treat the subject with diphenhydramine 50 mg IV (or equivalent) and/or paracetamol 325 to 1000 mg (acetaminophen); remain at bedside and monitor subject until resolution of symptoms. Corticosteroid or bronchodilator therapy may also be administered as appropriate. If the infusion is interrupted, then restart the infusion at 50% of the original infusion rate when symptoms resolve; if no further complications ensue after 30 minutes, the rate may be increased to 100% of the original infusion rate. Monitor subject closely. If symptoms recur then no further nivolumab or ipilimumab will be administered at that visit. Administer diphenhydramine 50 mg IV, and remain at bedside and monitor the subject until resolution of symptoms. The amount of study drug infused must be recorded on the electronic case report form (eCRF). The following prophylactic premedications are recommended for future infusions: diphenhydramine 50 mg (or equivalent) and/or paracetamol 325 to 1000 mg (acetaminophen) should be administered at least 30 minutes before additional nivolumab or ipilimumab administrations. If necessary, corticosteroids (recommended dose: up to 25 mg of IV hydrocortisone or equivalent) may be used.

For Grade 3 or Grade 4 symptoms: (Severe reaction, Grade 3: prolonged [ie, not rapidly responsive to symptomatic medication and/or brief interruption of infusion]; recurrence of symptoms following initial improvement; hospitalization indicated for other clinical sequelae [eg, renal impairment, pulmonary infiltrates]). Grade 4: (life-threatening; pressor or ventilatory support indicated).

Immediately discontinue infusion of nivolumab or ipilimumab. Begin an IV infusion of normal saline, and treat the subject as follows. Recommend bronchodilators, epinephrine 0.2 to 1 mg of a 1:1,000 solution for subcutaneous administration or 0.1 to 0.25 mg of a 1:10,000 solution injected slowly for IV administration, and/or diphenhydramine 50 mg IV with methylprednisolone 100 mg IV (or equivalent), as needed. Subject should be monitored until the investigator is comfortable that the symptoms will not recur. Nivolumab or ipilimumab will be permanently discontinued. Investigators should follow their institutional guidelines for the treatment of anaphylaxis. Remain at bedside and monitor subject until recovery from symptoms. In the case of late-occurring hypersensitivity symptoms (eg, appearance of a localized or generalized pruritis within 1 week after treatment), symptomatic treatment may be given (eg, oral antihistamine, or corticosteroids).

4.5.7 Treatment Beyond Disease Progression

Accumulating evidence[40] indicates a minority of subjects treated with immunotherapy may derive clinical benefit despite initial evidence of PD.

Subjects, regardless of study arm, will be permitted to continue treatment beyond initial investigator assessed progression as long as they meet the following criteria:

- Investigator-assessed clinical benefit
  and
- Subject is tolerating study drug.

The assessment of clinical benefit should take into account whether the subject is clinically deteriorating and unlikely to receive further benefit from continued treatment.

All decisions to continue treatment beyond initial progression should be discussed with the BMS Medical Monitor or Study Director and documented in the study records.

Subjects must be re-consented with an ICF addendum to continue treatment.

Subjects should discontinue study therapy upon evidence of further progression, defined as an additional 10% or greater increase in tumor burden volume from time of initial progression (including all target lesions and new measurable lesions).

New lesions are considered measurable at the time of initial progression if the longest diameter is at least 10 mm (except for pathological lymph nodes, which must have a short axis of at least 15 mm). Any new lesion considered non-measurable at the time of initial progression may become measurable and therefore included in the tumor burden measurement if the longest diameter increases to at least 10 mm (except for pathological lymph nodes, which must have an increase in short axis to at least 15 mm).

Amendment 14 Update:

Subjects enrolled in the nivolumab combined with ipilimumab crossover extension phase will be permitted to continue study therapy beyond initial investigator-assessed RECIST 1.1-defined progression, as defined in this section, up to a maximum of 2 years from the first dose of nivolumab combined with ipilimumab.

4.5.8 Immunotherapy Adverse Event Management

Because of the potential for clinically meaningful nivolumab or ipilimumab-related AEs requiring early recognition and prompt intervention, management algorithms have been developed for suspected pulmonary toxicity, GI toxicity, hepatotoxicity, endocrinopathy, skin toxicity, neurological toxicity, and renal toxicity.

These adverse event management algorithms are included in Appendix 5.

4.6 Blinding/Unblinding

Not applicable.

4.7 Treatment Compliance

Trained medical personnel will administer nivolumab and ipilimumab and dispense other study medication.

Treatment compliance will be monitored by drug accountability, as well as by recording administration of all medications in the CRF. The date and time of start and end of infusion and the exact amount given at each infusion will be recorded. Any missed doses or interruptions in sunitinib administration will be recorded. In case the treatment has to be interrupted during an infusion and the dosing is not resumed, the medical personnel should evaluate the percentage of dose received by the patient and document it in the patient record.

Any reason for non-compliance should also be documented.

4.8 Destruction of Study Drug

For this study, study drugs (those supplied by BMS or sourced by the investigator) such as partially used study drug containers, vials and syringes may be destroyed on site.

Any unused study drugs can only be destroyed after being inspected and reconciled by the responsible Study Monitor unless study drug containers must be immediately destroyed as required for safety, or to meet local regulations (eg, cytotoxics or biologics).

On-site destruction is allowed provided the following minimal standards are met:

- On-site disposal practices must not expose humans to risks from the drug.
- On-site disposal practices and procedures are in agreement with applicable laws and regulations, including any special requirements for controlled or hazardous substances.
- Written procedures for on-site disposal are available and followed. The procedures must be filed with the site's SOPs and a copy provided to BMS upon request.
- Records are maintained that allow for traceability of each container, including the date disposed of, quantity disposed, and identification of the person disposing the containers. The method of disposal, ie, incinerator, licensed sanitary landfill, or licensed waste disposal vendor must be documented.
- Accountability and disposal records are complete, up-to-date, and available for the Monitor to review throughout the clinical trial period.

If conditions for destruction cannot be met the responsible Study Monitor will make arrangements for return of study drug.

It is the investigator's responsibility to arrange for disposal of all empty containers, provided that procedures for proper disposal have been established according to applicable federal, state, local, and institutional guidelines and procedures, and provided that appropriate records of disposal are kept.

4.9 Return of Study Drug

If study drug will not be destroyed upon completion or termination of the study, all unused and/or partially used study drug that was supplied by BMS must be returned to BMS. The return of study drug will be arranged by the responsible Study Monitor.

It is the investigator's responsibility to arrange for disposal of all empty containers, provided that procedures for proper disposal have been established according to applicable federal, state, local,

and institutional guidelines and procedures, and provided that appropriate records of disposal are kept.

Arrangements for the return of study drug will be made by the responsible Study Monitor.

4.10 Retained Samples for Bioavailability / Bioequivalence

Not applicable.

Motzer RJ, *et al. J Immunother Cancer* 2021; 8:e000891. doi: 10.1136/jitc-2020-000891

5 STUDY ASSESSMENTS AND PROCEDURES

5.1 Flow Chart/Time and Events Schedule

Table 5.1-1: Screening Assessments (CA209214)

Procedure	Screening Visit	Notes
Eligibility Assessments		
Informed Consent	X	
Inclusion/Exclusion Criteria	X	All inclusion/exclusion criteria should be assessed prior to randomization
Medical History	X	
Tumor Tissue Samples	X	Sufficient tumor tissue, archival or recent acquisition, (block or minimum of 10 slides; obtained from core biopsy, punch biopsy, excisional biopsy or surgical specimen) received at Central Laboratory.
Safety Assessments		
Physical Examination	X	
Vital Signs and oxygen saturation	X	Including BP, HR, & temperature. Obtain at the screening visit and within 72 hours prior to first dose
Physical Measurements (including performance status)	X	Height and weight and Karnofsky Performance Status.
ECG[a]	X	Within 28 days prior to randomization
Cardiac Ejection Fraction	X	Within 28 days prior to randomization
Assessment of Signs and Symptoms	X	Within 14 days prior to randomization
Concomitant Medication Collection	X	Within 14 days prior to randomization
Laboratory Tests	X	CBC w/differential, Chemistry panel including: LDH, AST, ALT, ALP, T.Bili, BUN or serum urea level, creatinine, Ca, albumin, Mg, Na, K, Cl, glucose, amylase, lipase, TSH, Free T4, Free T3, Hep B/C(HBV sAG, HCV antibody or HCV RNA), within 14 days prior to randomization
Pregnancy Test (WOCBP Only)	X	

Table 5.1-1: Screening Assessments (CA209214)		
Procedure	**Screening Visit**	**Notes**
Efficacy Assessment		
Screening/Baseline Tumor Assessments	X	CT/MRI of the chest, abdomen, pelvis and all known sites of disease and MRI (preferred) or CT scan of the Brain within 28 days prior to randomization

[a] Fridericia corrected Qt required

Table 5.1-2: On-study Assessments Cycles 1 and 2 (CA209214)

Procedure	Cycle 1 and Cycle 2[a] (Cycle = 6 weeks)						Notes
	Day 1 Week 1	Day 1 Week 2	Day 1 Week 3	Day 1 Week 4	Day 1 Week 5	Day 1 Week 6	
Safety Assessments							
Targeted Physical Examination	X			X			To be performed only as clinically indicated within 72 hours prior to dosing.
Vital Signs and Oxygen Saturation	X			X			Including BP, HR, temperature and oxygen saturation by pulse oximetry. Pulse oximetry at rest and after exertion prior to dosing.
Physical Measurements (including performance status)	X			X			Weight and KPS within 72 hours prior to dosing
Adverse Events Assessment	Continuously						
Review of Concomitant Medications	X			X			
Laboratory Tests	X			X			Within 72 hrs prior to dosing to include CBC w/ differential, AST, ALT, ALP, T.Bili, BUN or serum urea level, creatinine, Ca, Mg, Na, K, Cl, LDH, glucose, amylase, lipase, TSH (with reflexive Free T4 and Free T3). Note: C1W1D1 labs do not need to be repeated if they were performed within 14 days of dosing.
Pregnancy Test (WOCBP Only)	X			X			Within 24 hours prior to administration of first dose of study drug and thereafter Q 4weekly independent of study drug dosing. Serum or Urine
Immunogenicity blood sample	See Table 5.5.1-1 for details regarding specific sample timing						
Pharmacokinetic Samples							
PK samples	See Table 5.5.1-1 for details regarding specific sample timing						

Table 5.1-2: On-study Assessments Cycles 1 and 2 (CA209214)

Procedure	Cycle 1 and Cycle 2[a] (Cycle = 6 weeks)						Notes
	Day 1 Week 1	Day 1 Week 2	Day 1 Week 3	Day 1 Week 4	Day 1 Week 5	Day 1 Week 6	
Exploratory Biomarker Testing							
Exploratory Serum Biomarkers	Y			Y			To be collected pre-dose; Y= only for Cycle 1
Peripheral Blood RNA	Y			Y			To be collected pre-dose; Y= only for Cycle 1
Whole Blood Sample for Genotyping	Y						EDTA Tubes for DNA. Must be obtained prior to dosing. Y= only for Cycle 1.
Myeloid Derived Suppressor Cells (MDSCs)	Y						Cyto-Chex BCT for MDSC stabilization. Must be obtained prior to dosing. Y = only for cycle 1.
Peripheral Blood Mononuclear Cells (PBMCs)	X			Y			Please refer to lab manual for instructions. Y = only for cycle 1 3 Collections Total - Cycle 1 Week1, Cycle 1 Week4, Cycle 2 Week 1
Efficacy Assessments							
Tumor Assessments[b]	FIRST tumor assessment should be performed at 12 weeks (± 1 wk) following randomization. SUBSEQUENT tumor assessments should occur every 6 weeks (1± wk) until disease progression. CT/MRI of the chest, abdomen, pelvis and all known sites of disease. Use same imaging method as was used at screening/baseline.						
Clinical Drug Supplies							
IVRS-Randomize	Y						Y=Only for cycle 1
Administer Nivolumab and Ipilimumab (Arm A)	X			X			First dose to be administered within 3 days following randomization. Subsequent doses may be administered within 3 days after the scheduled date if necessary.

Table 5.1-2: On-study Assessments Cycles 1 and 2 (CA209214)

Procedure	Cycle 1 and Cycle 2[a] (Cycle = 6 weeks)						Notes
	Day 1 Week 1	Day 1 Week 2	Day 1 Week 3	Day 1 Week 4	Day 1 Week 5	Day 1 Week 6	
Dispense Sunitinib (Arm B)	X						First dose to be administered within 3 days following randomization.[c]
Outcomes Research Assessments							
FACT-G	X			X			D1W1 should be completed after randomization but prior to any study-related procedures.
EQ-5D	X			X			D1W1 should be completed after randomization but prior to any study-related procedures.
FKSI-19	X			X			D1W1 should be completed after randomization but prior to any study-related procedures.
Healthcare Resource Utilization	X			X			

[a] If a dose is delayed, the procedures scheduled for that same timepoint should also be delayed to coincide with when that timepoint's dosing actually occurs.

[b] Tumor assessments will be performed at weeks 12, 18, 24, 30 etc. (± 1 week) from randomization, these time points are independent of dosing.

[c] C1D1 dose should occur within 3 days of treatment assignment. Subjects requiring preapproval from insurance or other sources to obtain sunitinib may delay start of treatment an additional 5 days (total of 8 days from treatment assignment in IVRS) to obtain approvals. Treatment should start as soon as possible after approval received. If additional time is required, notify site manager. For these subjects, CBC and/or Chemistry should be repeated ≤ 72 hours prior to the initial dose if more than 5 days has elapsed since testing performed or if the investigator notes any changes in the subject's condition that warrant repeat testing.

Table 5.1-3: On-study Assessments Cycle 3 and Beyond (CA209214)[a]

Procedure	Cycle 3 and Beyond (Cycle = 6 weeks)						Notes
	Day 1 Week 1	Day 1 Week 2	Day 1 Week 3	Day 1 Week 4	Day 1 Week 5	Day 1 Week 6	
Safety Assessments							
Targeted Physical Examination	X		X: Arm A only		X		To be performed only if clinically indicated within 72 hours prior to dosing
Vital Signs and Oxygen Saturation	X		X: Arm A only		X		Including BP, HR, temperature and oxygen saturation by pulse oximetry. Pulse oximetry at rest and after exertion prior to dosing.
Physical Measurements (including performance status)	X		X: Arm A only		X		Weight and KPS within 72 hours prior to dosing.
Adverse Events Assessment	Continuously						
Review of Concomitant Medications	X		X: Arm A only		X		
Laboratory Tests	X		X: Arm A only AST, ALT, ALP, and T. Bili only		X		Within 72 hrs prior to re-dosing to include CBC w/ differential, AST, ALT, ALP, T.Bili, BUN or serum urea level, creatinine. Ca, Mg, Na, K, Cl, LDH, glucose, amylase, lipase, TSH (with reflexive Free T4 and Free T3)
Pregnancy Test (WOCBP Only)	X				X		Within 24 hours prior to administration of study drug. Serum or Urine
Immunogenicity blood sample	See Table 5.5.1-1 for details regarding specific sample timing						

Table 5.1-3: On-study Assessments Cycle 3 and Beyond (CA209214)[a]

Procedure	Cycle 3 and Beyond (Cycle = 6 weeks)						Notes
	Day 1 Week 1	Day 1 Week 2	Day 1 Week 3	Day 1 Week 4	Day 1 Week 5	Day 1 Week 6	
Pharmacokinetic Samples							
PK samples	See Table 5.5.1-1 for details regarding specific sample timing						
Efficacy Assessments							
Tumor Assessments	FIRST tumor assessment should first be performed at 12 weeks (± 1 wk) following randomization. SUBSEQUENT tumor assessments should occur every 6 weeks (1± wk) up to first 13 months, then every 12 weeks until disease progression. CT/MRI of chest, abdomen, pelvis and all known sites of disease. Use same imaging method as was used at screening/baseline.						
Clinical Drug Supplies							
Administer Nivolumab (Arm A)[b,c]	X		X		X		Subsequent doses may be administered within 3 days after the scheduled date if necessary. See Section 4.5
Dispense Sunitinib (Arm B)	X						
Outcomes Research Assessments							
FACT-G	X				Y		Prior to any study-related procedures. Y=only during 1st 6 months
EQ-5D	X				Y		Prior to any study-related procedures. Y=only during 1st 6 months
FKSI-19	X				Y		Prior to any study-related procedures. Y=only during 1st 6 months
Healthcare Resource Utilization	X				Y		Y=only during 1st 6 months

[a] If a dose is delayed, the procedures scheduled for that same timepoint should also be delayed to coincide with when that timepoint's dosing actually occurs.

[b] Subjects can receive nivolumab at 3mg/kg or 240mg flat dose.

[c] Arm A subjects who have completed at least 2 years of treatment have the option to discontinue study treatment at the discretion of the subject and/or investigator.

Table 5.1-4: Screening Assessments (CA209214) for Subjects Previously Randomized to Arm B Entering Nivolumab Combined with Ipilimumab Crossover Extension Phase

Procedure	Screening Visit	Notes
Eligibility Assessments		
Informed Consent	X	
Inclusion/Exclusion Criteria	X	Assessed prior to calling IVRS and registering the subject for crossover extension phase
Medical History	X	
Safety Assessments		
Physical Examination	X	
Vital Signs and Oxygen Saturation	X	Including BP, HR, & temperature. Obtain vital signs within 72 hours prior to first dose of nivolumab combined with ipilimumab.
Physical Measurements (including performance status)	X	Height and weight and Karnofsky Performance Status
ECG[a]	X	Within 28 days prior to first dose of nivolumab combined with ipilimumab
Assessment of Signs and Symptoms	X	After obtaining Informed Consent, assess all signs and symptoms within 14 days prior to the first dose of nivolumab combined with ipilimumab
Concomitant Medication Collection	X	Within 14 days prior to first dose of nivolumab combined with ipilimumab
Laboratory Tests	X	Labs performed locally within 14 days prior to first dose of nivolumab combined with ipilimumab (unless otherwise specified): CBC w/Differential, Chemistry Panel including: LDH, AST, ALT, ALP, T.Bili, BUN or Serum Urea Level, Creatinine, Ca, Albumin, Mg, Na, K, Cl, Glucose, Amylase, Lipase, TSH, Free T4, Free T3
Pregnancy Test (WOCBP Only)	X	Performed within 24 hours prior to first dose of nivolumab combined with ipilimumab (serum or urine)

Table 5.1-4: Screening Assessments (CA209214) for Subjects Previously Randomized to Arm B Entering Nivolumab Combined with Ipilimumab Crossover Extension Phase

Procedure	Screening Visit	Notes
Efficacy Assessments		
Radiographic Tumor Assessment (Chest, Abdomen, Pelvis)	X	Should be performed within 28 days prior to first dose of nivolumab combined with ipilimumab. MRI of brain (with contract unless contraindicated) is required in subjects with known history of brain metastases. Additional site of known or suspected disease (including CNS) should be imaged prior to first dose of nivolumab combined with ipilimumab.
Other		
Patient-Reported Outcome Measurements	X	Prior to first dose of nivolumab combined with ipilimumab: FKSI-DRS and EQ-5D

[a] Fridericia corrected Qt required

Table 5.1-5: On-Study Assessments Cycles 1 and 2 (CA209214) for Subjects Previously Randomized to Arm B Entering Nivolumab Combined with Ipilimumab Crossover Extension Phase

Procedure	Cycle 1 and Cycle 2[a] (Cycle = 6 Weeks)						Notes
	Day 1 Week 1	Day1 Week 2	Day 1 Week 3	Day 1 Week 4	Day 1 Week 5	Day 1 Week 6	
Safety Assessments							
Targeted Physical Examination	X			X			To be performed only as clinically indicated within 72 hours prior to dosing
Vital Signs and Oxygen Saturation	X			X			Including BP, HR, temperature and oxygen saturation by pulse oximetry. Pulse oximetry at rest and after exertion prior to dosing.
Physical Measurements (including performance status)	X			X			Weight and KPS within 72 hours prior to dosing
Adverse Events Assessment	Continuously						
Review of Concomitant Medications	X			X			
Laboratory Tests	X			X			Within 72 hours prior to dosing to include CBC w/Differential, LFTs, BUN or Serum Urea, Creatinine, AST, ALT, ALP, T.Bili, Ca, Mg, Na, K, Cl, Glucose, Amylase, Lipase, TSH (with Reflective Free T4 and Free T3). ***Note:*** C1W1D1 labs do not need to be repeated if they were performed within 14 days of dosing.
Pregnancy Test (WOCBP Only)	X			X			Serum or Urine - within 24 hours prior to administration of first dose of nivolumab combined with ipilimumab, thereafter Q4 weekly independent of study drug dosing.

Table 5.1-5: On-Study Assessments Cycles 1 and 2 (CA209214) for Subjects Previously Randomized to Arm B Entering Nivolumab Combined with Ipilimumab Crossover Extension Phase

Procedure	Cycle 1 and Cycle 2[a] (Cycle = 6 Weeks)						Notes
Efficacy Assessments							
Tumor Assessments[b]	FIRST tumor assessment should be performed at 12 weeks (± 1 wk) following first dose of nivolumab combined with ipilimumab. SUBSEQUENT tumor assessments should occur every 8 weeks (1± wk) until disease progression. CT/MRI of the chest, abdomen, pelvis and all known sites of disease. Use same imaging method as was used at screening/baseline.						
Clinical Drug Supplies							
IVRS - Register Crossover Extension Phase	Y						Y = Only for Cycle 1
Administer Study Treatment	X			X			Record study drug infusion start and stop times for nivolumab and ipilimumab. Doses of nivolumab combined with ipilimumab should not be given less than 19 days from the previous dose.
Outcomes Research Assessments							
EQ-5D	X			X			Prior to any study-related procedures
FKSI-19	X			X			Prior to any study-related procedures
Healthcare Resource Utilization	X			X			Except Cycle 1 (See Section 5.7.1)

[a] If a dose is delayed, the procedures scheduled for that same timepoint should also be delayed to coincide with when that timepoint's dosing actually occurs.

[b] Tumor assessments will be performed at Weeks 12, 20, 28, 36 etc. (± 1 week) from first dose of nivolumab combined with ipilimumab; these time points are independent of dosing.

Table 5.1-6: On-Study Assessments Cycle 3 and Beyond (CA209214) for Subjects Previously Randomized to Arm B Entering Nivolumab Combined with Ipilimumab Crossover Extension Phase

Procedure	Cycle 3 and Beyond[a] (Cycle = 6 Weeks)						Notes
	Day 1 Week 1	Day 1 Week 2	Day 1 Week 3	Day 1 Week 4	Day 1 Week 5	Day 1 Week 6	
Safety Assessments							
Targeted Physical Examination	X		X		X		To be performed only if clinically indicated within 72 hours prior to dosing
Vital Signs and Oxygen Saturation	X		X		X		Including BP, HR, temperature and oxygen saturation by pulse oximetry. Pulse oximetry at rest and after exertion prior to dosing
Physical Measurements (including performance status)	X		X		X		Weight and KPS within 72 hours prior to dosing
Adverse Events Assessment	Continuously						
Review of Concomitant Medications	X		X		X		
Laboratory Tests	X		Y		X		**X:** Within 72 hours prior to dosing to include CBC w/Differential, LFTs, BUN or Serum Urea, Creatinine, AST, ALT, ALP, T.Bili, Ca, Mg, Na, K, Cl, Glucose, Amylase, Lipase, TSH (with Reflexive Free T4 and Free T3) **Y:** AST, ALT, ALP, and T.Bili only
Pregnancy Test (WOCBP Only)	X				X		Serum or Urine - within 24 hours prior to administration of nivolumab

Table 5.1-6: On-Study Assessments Cycle 3 and Beyond (CA209214) for Subjects Previously Randomized to Arm B Entering Nivolumab Combined with Ipilimumab Crossover Extension Phase

Procedure	Cycle 3 and Beyond[a] (Cycle = 6 Weeks)						Notes
Efficacy Assessments							
Tumor Assessments[b]	FIRST tumor assessment should be performed at 12 weeks (± 1 wk) following first dose of nivolumab combined with ipilimumab. SUBSEQUENT tumor assessments should occur every 8 weeks (1± wk) until disease progression. CT/MRI of the chest, abdomen, pelvis and all known sites of disease. Use same imaging method as was used at screening/baseline.						
Clinical Drug Supplies							
Administer Study Treatment	X		X		X		Subjects will receive nivolumab 240mg flat dose. Record study drug infusion start and stop times. Doses of nivolumab should not be given less than 12 days from the previous dose. The first dose of nivolumab monotherapy should not be given less than 19 days from the previous nivolumab combined with ipilimumab dose.
Outcomes Research Assessments							
EQ-5D	X				Y		Prior to any study-related procedures. Y = Only during the first 6 months
FKSI-19	X				Y		Prior to any study-related procedures. Y = Only during the first 6 months
Healthcare Resource Utilization	X				Y		Y = Only during the first 6 months

[a] If a dose is delayed, the procedures scheduled for that same timepoint should also be delayed to coincide with when that timepoint's dosing actually occurs.

[b] Tumor assessments will be performed at Weeks 12, 20, 28, 36 etc. (± 1 week) from first dose of nivolumab combined with ipilimumab; these time points are independent of dosing.

Table 5.1-7: Follow-up Assessments (CA209214) - *All Subjects*

Procedure	Follow-Up[a], Visits 1 and 2	Survival[b], Follow-up Visits	Notes
Safety Assessments			
Targeted Physical Examination	X		To assess for potential late emergent study drug related issues
Adverse Events Assessment	X	X	
Laboratory Tests	X		On site/local CBC w/differential, AST, ALT, ALP, T.Bili, BUN or serum urea level, creatinine and TSH for X01, repeat at X02 (or XX02 for crossover extension phase) if study drug related toxicity persists.
Pregnancy Test (WOCBP Only)	X		Serum or urine
Review of Concomitant Medication	X		
Immunogenicity blood sample	X		Refer to Table 5.5.1-1 for details regarding specific sample timing; Only for subjects ORIGINALLY randomized to Arm A
Outcomes Research Assessments			
FACT-G	X		Follow-up visits 1 and 2 only.
EQ-5D	X	Y	X=entered by patient, Y=Instrument to be entered by site every 3 months for the first 12 months then every 6 months thereafter.
FKSI-19	X		Follow-up visits 1 and 2 only.
Healthcare Resource Utilization	X		Follow-up visits 1 and 2 only
Survival Status			
Subject Status	X	X	Every 3 months, may be accomplished by visit or phone contact, to include subsequent anti-cancer therapy

Table 5.1-7: Follow-up Assessments (CA209214) - *All Subjects*

Procedure	Follow-Up[a], Visits 1 and 2	Survival[b], Follow-up Visits	Notes
Efficacy Assessments			
Tumor Assessments	Only for subjects without RECIST 1.1 defined progression on study therapy. For All Subjects **Not** in the Nivolumab Combined with Ipilimumab Crossover Extension Phase: FIRST tumor assessment should first be performed at 12 weeks (± 1 wk) following randomization SUBSEQUENT tumor assessments should occur every 6 weeks (± 1 wk) thereafter for the first 13 months, then every 12 wks (± 1 wk) until disease progression CT/MRI of the chest, abdomen, pelvis and all known sites of disease. Use same imaging method as was used at screening/baseline. For All Subjects in the Nivolumab Combined with Ipilimumab Crossover Extension Phase: FIRST tumor assessment should first be performed at 12 weeks (± 1 wk) following randomization SUBSEQUENT tumor assessments should occur every 8 weeks (± 1 wk) thereafter for the first 13 months, then every 12 wks (± 1 wk) until disease progression CT/MRI of the chest, abdomen, pelvis and all known sites of disease. Use same imaging method as was used at screening/baseline.		
Pharmacokinetic Samples			
PK samples	X		See Table 5.5.1-1 for schedule of assessments; Only for subjects ORIGINALLY randomized to Arm A

[a] Follow-up visit 1 (FU1) = 30 days from the last dose +/- 7 days or coincide with the date of discontinuation (+/- 7 days) if date of discontinuation is greater than 37 days after last dose, Follow-up visit 2 (FU2) = 84 days (+/- 7 days) from follow-up visit 1; subjects will follow the same timeframe and procedures for the Nivolumab combined with Ipilimumab Crossover Extension phase, with the exception of the PK sample collection.

[b] Survival visits = every 3 months from FU2 +/- 14 days; subjects will follow the same timeframe and procedures for the Nivolumab combined with Ipilimumab Crossover Extension phase.

5.1.1 Retesting During Screening or Lead-in Period

Retesting of laboratory parameters and/or other assessments within any single Screening or Lead-in period will be permitted (in addition to any parameters that require a confirmatory value).

Any new result will override the previous result (ie, the most current result prior to Randomization) and is the value by which study inclusion will be assessed, as it represents the subject's most current, clinical state.

5.2 Study Materials

- NCI CTCAE version 4.0
- Nivolumab Investigator Brochure
- Ipilimumab Investigator Brochure
- Pharmacy Binder
- Laboratory manuals for collection and handling of blood (including PK, biomarker and immunogenicity) and tissue specimens
- Site manual for operation of interactive voice response system, including enrollment/randomization worksheets
- Manual for entry of local laboratory data
- Pregnancy Surveillance Forms
- CA209214 Imaging Manual

5.3 Safety Assessments

At baseline, a medical history will be obtained to capture relevant underlying conditions. The baseline examinations should include weight, height, Karnofsky Performance Status, BP, HR, and temperature. Baseline signs and symptoms are those that are assessed within 14 days prior to randomization. Concomitant medications will be collected from within 14 days prior to randomization through the study treatment period (see Section 5.1).

Baseline local laboratory assessments should be done within 14 days prior to randomization to include: CBC w/differential, LFTs (ALT, AST, total bilirubin, alkaline phosphatase), BUN or serum urea level, creatinine, Ca, albumin, Mg, Na, K, Cl, LDH, glucose, amylase, lipase, TSH, Free T4, Free T3, and Hep B and C testing (HBV sAg, HCV Ab or HCV RNA) (see Table 5.1-1).

Pregnancy testing for WOCBP (done locally) must be performed within 24 hours prior to the initial administration of study drug at baseline and then 4 weekly during the treatment period and at the safety follow-up visits.

Subjects will be evaluated for safety if they have received any study drug. Toxicity assessments will be continuous during the treatment phase. During the safety follow-up phase Table 5.1-7, toxicity assessments should be done in person. Once subjects reach the survival follow-up phase

either in person or documented telephone calls to assess the subject's status are acceptable. Adverse events and laboratory values will be graded according to the NCI-CTCAE version 4.0.

On-study weight and Karnofsky Performance status should be assessed on Day 1 of Weeks 1, and 4 during cycles 1 and 2 and Day 1 of Weeks 1, 3 (Arm A only) and 5 starting from Cycle 3 and vital signs should be assessed at each on-study visit. Vital signs should also be taken as per institutional standard of care prior to, during and after dosing. Oxygen saturation by pulse oximetry at rest and after exertion should be assessed at each on-study visit prior to dosing. The start and stop times of the nivolumab and ipilimumab infusions should be documented. Physical examinations are to be performed as clinically indicated. If there are any new or worsening clinically significant changes since the last exam, report changes on the appropriate non-serious or serious adverse event page.

On study local laboratory assessments should be done within 72 hours of dosing to include; CBC w/differential, LFTs (ALT, AST, total bilirubin, alkaline phosphatase), BUN or serum urea level, creatinine, Ca, Mg, Na, K, Cl, LDH, glucose, amylase, lipase, TSH with reflexive Free T4, Free T3 on Day 1 of Weeks 1 and 4 for Cycles 1 and 2 and on Day 1 of Weeks 1 and 5 starting from Cycle 3. Additional measures including non-study required laboratory tests should be performed as clinically indicated or to comply with local regulations. Laboratory toxicities (eg, suspected drug inducted liver enzyme elevations) will be monitored during the follow-up phase via on site/local labs until all study drug related toxicities resolve, return to baseline or are deemed irreversible.

In addition, for subjects on the nivolumab combined with ipilimumab arm (Arm A), LFTs should also be assessed prior to the second dose of each cycle. The results of these labs should be reviewed prior to dosing.

Oxygen saturation by pulse oximetry should be obtained prior to each dosing and at any time a subject has any new or worsening respiratory symptoms. A reading at rest and on exertion should be obtained at each time point. The extent of the exertion should be based on the judgment of the investigator, but should remain consistent for each individual subject throughout the study. If the patient's subject's status changes, the investigator can alter the extent of exertion based on their medical judgment. If a subject shows changes on pulse oximetry or other pulmonary related signs (hypoxia, fever) or symptoms (eg, dyspnea, cough, fever) consistent with possible pulmonary adverse events, the patient subject should be immediately evaluated to rule out pulmonary toxicity. An algorithm for the management of suspected pulmonary toxicity can be found in the nivolumab Investigator's Brochure and in Appendix 5.

Some of the previously referred to assessments may not be captured as data in the eCRF. They are intended to be used as safety monitoring by the treating physician. Additional testing or assessments may be performed as clinically necessary or where required by institutional or local regulations.

BMS may request that survival data be collected on all randomized subjects outside of the protocol defined window (Table 5.1-7). At the time of this request, each subject will be contacted to determine their survival status unless the subject has withdrawn consent for all contact.

Amendment 14 Update:

For subjects moving into the nivolumab combined with ipilimumab crossover extension phase, please refer to Table 5.1-4, Table 5.1-5, and Table 5.1-6 for the schedule of screening and on-study assessments.

5.3.1 Imaging Assessment for the Study

Any incidental findings of potential clinical relevance that are not directly associated with the objectives of the protocol should be evaluated and handled by the Study Investigator as per standard medical/clinical judgment.

Scans will be submitted to an Imaging Corelab by an IRCC.

5.3.1.1 CT/MRI

Both contrast-enhanced Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) scans acquired on dedicated CT/MRI equipment are adequate imaging modalities for this study.

CT with contrast (or MRI) of the chest, abdomen, pelvis and all other known sites of disease are to be performed for tumor assessments at baseline (Table 5.1-1), at 12 weeks (±1 wk) after randomization and then every 6 weeks (± 1 wk) as per Table 5.1-2 and Table 5.1-3, until disease progression or treatment is discontinued (whichever occurs later).

CT scans should be acquired with 5 mm slices with no intervening gap (contiguous). Should a subject have a contraindication for CT IV contrast, a non-contrast CT of the chest and a contrast enhanced MRI of the abdomen and pelvis may be obtained. MRIs should be acquired with slice thickness of ≤ 5 mm with no gap (contiguous).

Every attempt should be made to image each subject using an identical acquisition protocol on the same scanner for all imaging time.

Note: Use of CT component of a PET/CT scanner:

Combined modality scanning such as with FDG-PET/CT is increasingly used in clinical care, and is a modality/technology that is in rapid evolution; therefore, the recommendations outlined here may change rather quickly with time. At present, low dose or attenuation correction CT portions of a combined FDG-PET/CT are of limited use in anatomically based efficacy assessments and it is therefore suggested that they should not be substituted for dedicated diagnostic contrast enhanced CT scans for anatomically based RECIST 1.1 measurements. However, if a site can document that the CT performed as part of a FDG-PET/CT is of identical diagnostic quality to a diagnostic CT (with IV and oral contrast) then the CT portion of the FDG-PET/CT can be used for RECIST 1.1 measurements. Note, however, that the FDG-PET portion of the CT introduces additional data which may bias an investigator if it is not routinely or serially performed.

5.3.1.2 MRI/CT Brain

MRI (preferred) or CT of brain is required at screening in order to rule out active metastatic disease (subjects with a history of brain metastasis are excluded from the study). MRI or CT

brain scans during on-study treatment and follow up periods are required only if clinically indicated for new signs and symptoms that suggest central nervous system (CNS) involvement.

5.4 Efficacy Assessments

Study evaluations will take place in accordance with the flow charts in Section 5.1. Baseline assessments should be performed within 28 days prior to randomization utilizing CT or MRI. In addition to chest, abdomen, pelvis, and brain, all known sites of disease should be assessed at baseline. Subjects who cannot receive IV contrast at the start of study should be imaged by MRI of abdomen/pelvis with IV contrast and CT of chest without contrast. Subsequent assessments should include chest, abdomen, and pelvis, and all known sites of disease and should use the same imaging method as was used at baseline. Patients initially imaged with CT of chest, abdomen, and pelvis with IV contrast who can no longer receive contrast can be monitored by CT of chest, abdomen, and pelvis without IV contrast. Subjects will be evaluated for tumor response beginning 12 weeks (± 1 wk) from randomization and continuing every 6 weeks (± 1 wk) for the first 13 months from randomization and every 12 weeks (± 1 wk) thereafter, until disease progression is documented or treatment is discontinued (whichever occurs later). Tumor assessments for ongoing study treatment decisions will be completed by the investigator using RECIST (Response Evaluation Criteria in Solid Tumors) 1.1 criteria.

Amendment 14 Update:

Subjects will be evaluated for tumor response beginning 12 weeks (±1 week) from the first dose of nivolumab combined with ipilimumab (start of the crossover extension phase) and continuing every 8 weeks (±1 week) for the first 13 months from the first dose and every 12 weeks (±1 week) thereafter, until disease progression is documented or treatment is discontinued (whichever occurs later).

5.5 Pharmacokinetic and Immunogenicity Assessments

Samples for PK and immunogenicity assessments will be collected for all subjects originally randomized to Arm A as described in Table 5.5.1-1. All timepoints are relative to the start of nivolumab drug administration. All on-treatment timepoints are intended to align with days on which study drug is administered. If it is known that a dose is going to be delayed, then the predose sample should be collected just prior to the delayed dose. However, if a predose sample is collected but the dose is subsequently delayed, an additional predose sample should not be collected. Further details of sample collection, processing, and shipment will be provided in the laboratory procedures manual.

5.5.1 Pharmacokinetics and Immunogenicity Collection and Processing

A detailed schedule of PK and immunogenicity evaluations is provided in Table 5.5.1-1 PK samples will be analyzed for nivolumab and ipilimumab by a validated immunoassay. Immunogenicity samples will be analyzed for anti-nivolumab and anti-ipilimumab antibodies by a validated immunoassay; samples may also be analyzed for neutralizing antibodies by a validated functional cell-based method. Serum samples may be analyzed by an exploratory method that measures nivolumab and ipilimumab anti-drug antibodies for technology exploration

purposes; exploratory results will not be reported. Serum samples designated for PK or biomarker assessments may also be used for immunogenicity analysis if required (eg, insufficient volume for complete immunogenicity assessment or to follow up on suspected immunogenicity related AE).

Table 5.5.1-1: Pharmacokinetic (PK) and Immunogenicity Sample Collections for Subjects Originally Randomized to Arm A (CA209214)

Study Day[a]	Event (Relative To Nivolumab Dosing) Hour	Time (Relative To Nivolumab Dosing) Hour: Min	Pharmacokinetic Blood Sample for Nivolumab	Immunogenicity Blood Sample for Nivolumab	Pharmacokinetic Blood Sample for Ipilimumab	Immunogenicity Blood Sample for Ipilimumab
C1W1D1	Predose[b]	00:00	X	X	X	X
C1W1D1	EOI[c]	02:00	X		X	
C1W4D1	Predose[b]	00:00	X	X	X	X
C2W1D1	Predose[b]	00:00	X	X	X	X
C2W1D1	EOI[c]	02:00	X		X	
C3W1D1	Predose[b]	00:00	X	X	X	X
W1D1 of every 2nd cycle after C3W1D1 up to C9W1D1	Predose[b]	00:00	X	X	X	X
W1D1 of every 3rd cycle after C9W1D1 until end of study treatment	Predose[b]	00:00	X	X		
First 2 Follow-up visits (approximately up to 100 days from the discontinuation of study drug)	N/A	N/A	X	X	X[d]	X[d]

[a] If a subject permanently discontinues study drug treatment during the sampling period, they will move to sampling at the follow up visits

[b] Predose: All pre-dose samples for nivolumab and ipilimumab should be taken prior to the start of nivolumab infusion.

[c] EOI: End of Infusion. EOI samples for both nivolumab and ipilimumab should be collected after the end of the ipilimumab infusion

[d] Ipilimumab follow-up samples are only to be collected if a subject discontinues treatment prior to C6D1

5.6 Biomarker Assessments

A variety of factors that may impact the immunomodulatory properties and efficacy of nivolumab-ipilimumab combination will be investigated in tumor tissue and in peripheral blood specimens taken from all randomized subjects prior to or during treatment as outlined in Table 5.1-1 and Table 5.1-2. Data from these investigations will be evaluated for pharmacodynamic effects, where applicable, and for associations with response, survival endpoints, and/or safety (adverse events). Comparative analyses between the two treatment arms will be used to identify biomarkers with predictive versus prognostic value. Several analyses will be completed and are described briefly below. Complete instructions on the collection, processing, handling, and shipment of all samples will be provided in a separate lab procedure manual.

5.6.1 Tumor Tissue Specimens

Archival or current formalin-fixed, paraffin-embedded tumor tissue must be sent to the central vendor/laboratory prior to a subject being randomized. A reference laboratory will receive the samples for IHC-based analyses aimed at quantifying the expression of proteins involved in PD-1 signaling such as PD-1, PD-L1, and PD-L2. Additional IHC analyses may be completed to determine the relative abundance of other protein markers associated with tumor-infiltrating immune cells (eg, CD4, CD8) and/or with RCC disease progression. The abundance of each protein monitored (or combinations of proteins) will be correlated with clinical endpoints. FFPE tissue may also be evaluated by fluorescent in-situ hybridization (FISH), genetic mutation detection methods, and/or by qRT-PCR as part of additional exploratory analyses seeking biomarker associations with clinical endpoints.

5.6.2 Peripheral Blood Specimens

5.6.2.1 Serum-Soluble Proteins

To understand the prevalence of circulating proteins and the impact they may have on clinical activity, the protein concentrations of a panel of cytokines, chemokines, and other relevant immunomodulatory or RCC-associated serum-soluble factors (eg, soluble PD-L1; VEGF) will be investigated prior to and on-treatment as outlined in Table 5.1-2.

5.6.2.2 Serum miRNA

MicroRNAs are broadly expressed, small RNAs that regulate the abundance of mRNA transcripts and their translation into protein. Global miRNA expression profiling has become increasingly common in cancer research, and miRNA signatures that are correlated to stage of disease or to clinical outcomes are now available for a variety of cancer types. Expression profiling of miRNA may also be useful in identifying molecular markers for the prediction of drug-responses and for prospective stratification. Intriguingly, miRNAs are stable in serum and may represent miRNAs over-expressed in tumors and/or reflect immune system activity. Serum taken from subjects randomized to each treatment arm may be analyzed for miRNA content by microarray and/or by similar methodologies (eg quantitative RT-PCR). The resulting miRNA expression profiles will be evaluated for associations with response and survival data. Ultimately, this approach may lead to the identification of unique miRNA signatures associated with clinical benefit.

5.6.2.3 Single Nucleotide Polymorphisms (Genotyping)

Whole blood will be collected from all subjects prior to first dose to generate genomic DNA for single nucleotide polymorphism (SNP) analyses. These analyses will focus on SNPs within genes associated with PD-1 and other immunoregulatory signaling pathways to determine if natural variation within those genes is associated with clinical benefit and/or with adverse events.

5.6.2.4 Peripheral Blood RNA

Transcriptional profiling of tumor[41] and of whole blood (unpublished) has been completed as part of clinical studies of iplimumab, leading to the identification of genes apparently associated with benefit from treatment. To determine if these or other genes are associated with clinical benefit from nivolumab-ipilimumab combination, gene expression profiling of whole blood RNA obtained pre-treatment may be completed. Gene expression may be assessed by microarray, qRT-PCR, and/or similar methodologies, with emphasis on genes with relevant immune function.

5.6.2.5 Myeloid Derived Suppressor Cells (MDSCs)

Myeloid derived suppressor cells are an immune cell population capable of suppressing T cell activation and proliferation. Preliminary data suggest that low pre-treatment MDSC levels in peripheral blood may be associated with better overall survival in melanoma patients treated with the immunotherapeutic agent ipilimumab. MDSCs will be quantified at pre-treatment to assess associations with outcomes.

5.6.2.6 Peripheral Blood Mononuclear Cells (PBMCs)

Peripheral blood mononuclear cells (PBMCs) in whole blood will be obtained at baseline and on treatment. Cells may be assessed by flow cytometry or by ELIspot. Nucleic acid may be prepared from cells to assess T cell rearrangements by next generation sequencing. Lymphocyte subsets to be assayed by flow cytometry may include, but are not limited to CD8+ and CD4+ T-cell subsets (activated; effector/memory; regulatory), NK cells, and populations of those cells as defined by the expression of activation, exhaustion, or signaling markers such as ICOS, HLA-DR, PD-1, CTLA-4, and/or intracellular IFNγ.

5.7 Outcomes Research Assessments

Patient reported outcomes will be captured through the use of three validated self reported questionnaires: the Functional Assessment of Cancer Therapy-General (FACT-G), the NCCN Functional Assessment of Cancer Therapy- Kidney Symptom Index (FKSI-19) and the EuroQoL Group's EQ-5D.

The FACT-G is a 27-item questionnaire that measures general cancer health related quality of life. It is one of the most widely used HrQoL cancer specific scales and has been validated in numerous types of cancer patients, across cultures, and in many languages. The scale is a compilation of general questions divided into four primary HrQoL dimensions: Physical Well-Being, Social/Family Well-Being, Emotional Well-Being, and Functional Well-Being. Summary scores can be calculated for each domain in addition to a single overall summary score.

The NCCN FKSI-19 is a 19-item scale that measures tumor specific HrQoL in kidney cancer patients. The FKSI-19 uses five Likert-type response categories that range from "not at all" to "very much". Patients are asked to circle the response category that best characterizes their response over the last 7 days on 19 items that include symptoms such as lack of energy, fatigue, appetite, coughing, shortness of breath, pain, nausea and ability to work.

General health status will be measured using the EQ-5D. The EQ-5D is a standardized instrument for use as a measure of self-reported general health status. The EQ-5D comprises 5 dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety) and a visual analog rating scale (VAS). The utility data generated from the EQ-5D is recommended for and commonly used in cost effectiveness analysis.

5.7.1 Healthcare Resource Utilization

Healthcare resource utilization data (eg, hospitalizations, non-protocol specified medical visits, etc) will be collected for all randomized subjects. Specifically, healthcare resource utilization is evaluated based on the number of medical care encounters such as hospital admissions and their duration, outpatient visits, diagnostic tests and procedures, concomitant medications and reasons for the encounters.

Resource utilization questions will be asked during the study and at the two follow up visits as outlined in Table 5.1-2, Table 5.1-3, Table 5.1-5, Table 5.1-6, and Table 5.1-7.

5.8 Other Assessments

5.8.1 Immunogenicity Assessments

Blood samples for immunogenicity analyses of nivolumab and/or ipilimumab will be collected according to the schedule given in Table 5.5.1-1. Samples collected from subjects in each treatment arm will be evaluated for development of Anti-Drug Antibody (ADA) for nivolumab and/or ipilimumab by validated immunoassays. Samples may also be analyzed for neutralizing ADA response to nivolumab and/or ipilimumab. (Neutralizing ADA testing conditioned upon assay availability.) For subjects treated in the nivolumab combined with ipilimumab crossover extension phase, bloods samples for immunogenicity will NOT be collected in the crossover extension phase.

5.9 Results of Central Assessments

Not Applicable.

6 ADVERSE EVENTS

An ***Adverse Event (AE)*** is defined as any new untoward medical occurrence or worsening of a preexisting medical condition in a clinical investigation subject administered study drug and that does not necessarily have a causal relationship with this treatment. An AE can therefore be any unfavorable and unintended sign (such as an abnormal laboratory finding), symptom, or disease temporally associated with the use of study drug, whether or not considered related to the study drug.

The causal relationship to study drug is determined by a physician and should be used to assess all adverse events (AE). The causal relationship can be one of the following:

Related: There is a reasonable causal relationship between study drug administration and the AE.

Not related: There is not a reasonable causal relationship between study drug administration and the AE.

The term "reasonable causal relationship" means there is evidence to suggest a causal relationship.

Adverse events can be spontaneously reported or elicited during open-ended questioning, examination, or evaluation of a subject. (In order to prevent reporting bias, subjects should not be questioned regarding the specific occurrence of one or more AEs.)

6.1 Serious Adverse Events

A S***erious Adverse Event (SAE)*** is any untoward medical occurrence that at any dose:

- results in death
- is life-threatening (defined as an event in which the subject was at risk of death at the time of the event; it does not refer to an event which hypothetically might have caused death if it were more severe)
- requires inpatient hospitalization or causes prolongation of existing hospitalization (see **NOTE** below)
- results in persistent or significant disability/incapacity
- is a congenital anomaly/birth defect
- is an important medical event (defined as a medical event(s) that may not be immediately life-threatening or result in death or hospitalization but, based upon appropriate medical and scientific judgment, may jeopardize the subject or may require intervention [eg, medical, surgical] to prevent one of the other serious outcomes listed in the definition above.) Examples of such events include, but are not limited to, intensive treatment in an emergency room or at home for allergic bronchospasm; blood dyscrasias or convulsions that do not result in hospitalization.) Potential drug induced liver injury (DILI) is also considered an important medical event. (See Section 6.6 for the definition of potential DILI.)

Suspected transmission of an infectious agent (eg, pathogenic or nonpathogenic) via the study drug is an SAE.

Although pregnancy, overdose, cancer, and potential drug induced liver injury (DILI) are not always serious by regulatory definition, these events must be handled as SAEs. (See Section 6.1.1 for reporting pregnancies).

Any component of a study endpoint that is considered related to study therapy (eg, death is an endpoint, if death occurred due to anaphylaxis, anaphylaxis must be reported) should be reported as SAE (see Section 6.1.1 for reporting details).

NOTE:

The following hospitalizations are not considered SAEs in BMS clinical studies:

- a visit to the emergency room or other hospital department < 24 hours, that does not result in admission (unless considered an important medical or life-threatening event)
- elective surgery, planned prior to signing consent
- admissions as per protocol for a planned medical/surgical procedure
- routine health assessment requiring admission for baseline/trending of health status (eg, routine colonoscopy)
- medical/surgical admission other than to remedy ill health and planned prior to entry into the study. Appropriate documentation is required in these cases
- admission encountered for another life circumstance that carries no bearing on health status and requires no medical/surgical intervention (eg, lack of housing, economic inadequacy, caregiver respite, family circumstances, administrative reason).
- Admission for administration of anticancer therapy in the absence of any other SAEs (applies to oncology protocols)

6.1.1 Serious Adverse Event Collection and Reporting

Sections 5.6.1 and 5.6.2 in the Investigator Brochure (IB) represent the Reference Safety Information to determine expectedness of serious adverse events for expedited reporting. Following the subject's written consent to participate in the study, all SAEs, whether related or not related to study drug, must be collected, including those thought to be associated with protocol-specified procedures. **All SAEs must be collected that occur during the screening period and within 100 days of discontinuation of dosing**. If applicable, SAEs must be collected that relate to any later protocol-specified procedure (eg, a follow-up skin biopsy).

The investigator should report any SAE that occurs after these time periods and that is believed to be related to study drug or protocol-specified procedure.

An SAE report should be completed for any event where doubt exists regarding its seriousness.

If the investigator believes that an SAE is not related to study drug, but is potentially related to the conditions of the study (such as withdrawal of previous therapy or a complication of a study procedure), the relationship should be specified in the narrative section of the SAE Report Form.

SAEs, whether related or not related to study drug, and pregnancies must be reported to BMS (or designee) within 24 hours. SAEs must be recorded on the SAE Report Form; pregnancies on a Pregnancy Surveillance Form (electronic or paper forms). The preferred method for SAE data reporting collection is through the eCRF. The paper SAE/pregnancy surveillance forms are only intended as a back-up option when the eCRF system is not functioning. In this case, the paper forms are to be transmitted via email or confirmed facsimile (fax) transmission to:

SAE Email Address: Refer to Contact Information list.

SAE Facsimile Number: Refer to Contact Information list.

For studies capturing SAEs through electronic data capture (EDC), electronic submission is the required method for reporting. The paper forms should be used and submitted immediately, only in the event the electronic system is unavailable for transmission. When paper forms are used, the original paper forms are to remain on site.

SAE Telephone Contact (required for SAE and pregnancy reporting): Refer to Contact Information list.

If only limited information is initially available, follow-up reports are required. (Note: Follow-up SAE reports should include the same investigator term(s) initially reported.)

If an ongoing SAE changes in its intensity or relationship to study drug or if new information becomes available, a follow-up SAE report should be sent within 24 hours to the BMS (or designee) using the same procedure used for transmitting the initial SAE report.

All SAEs should be followed to resolution or stabilization.

6.2 Nonserious Adverse Events

A ***nonserious adverse event*** is an AE not classified as serious.

6.2.1 Nonserious Adverse Event Collection and Reporting

The collection of nonserious AE information should begin at initiation of study drug and continue until 100 days from the last dose of study drug. Nonserious AE information should also be collected from the start of a placebo lead-in period or other observational period intended to establish a baseline status for the subjects.

Nonserious AEs should be followed to resolution or stabilization, or reported as SAEs if they become serious (see Section 6.1.1). Follow-up is also required for nonserious AEs that cause interruption or discontinuation of study drug and for those present at the end of study treatment as appropriate. All identified nonserious AEs must be recorded and described on the nonserious AE page of the CRF (paper or electronic).

Completion of supplemental CRFs may be requested for AEs and/or laboratory abnormalities that are reported/identified during the course of the study.

6.3 Laboratory Test Result Abnormalities

The following laboratory test result abnormalities should be captured on the nonserious AE CRF page or SAE Report Form (paper or electronic) as appropriate:

- Any laboratory test result that is clinically significant or meets the definition of an SAE
- Any laboratory test result abnormality that required the subject to have study drug discontinued or interrupted
- Any laboratory test result abnormality that required the subject to receive specific corrective therapy.

It is expected that wherever possible, the clinical rather than laboratory term would be used by the reporting investigator (eg, anemia versus low hemoglobin value).

6.4 Pregnancy

If, following initiation of the study drug, it is subsequently discovered that a study subject is pregnant or may have been pregnant at the time of study exposure, including during approximately half lives after product administration, the investigator must immediately notify the BMS Medical Monitor/designee of this event and complete and forward a Pregnancy Surveillance Form to BMS Designee within 24 hours and in accordance with SAE reporting procedures described in Section 6.1.1.

In most cases, the study drug will be permanently discontinued in an appropriate manner (eg, dose tapering if necessary for subject safety).

In the rare event that the benefit of continuing study drug is thought to outweigh the risk, after consultation with BMS, the pregnant subject may continue study drug after a thorough discussion of benefits and risk with the subject

Protocol-required procedures for study discontinuation and follow-up must be performed on the subject unless contraindicated by pregnancy (eg, x-ray studies). Other appropriate pregnancy follow-up procedures should be considered if indicated.

The investigator must immediately notify the BMS (or designee) Medical Monitor or Study Director of this event and complete and forward a Pregnancy Surveillance Form to BMS (or designee) within 24 hours and in accordance with SAE reporting procedures described in Section 6.1.1.

Follow-up information regarding the course of the pregnancy, including perinatal and neonatal outcome and, where applicable, offspring information must be reported on the Pregnancy Surveillance Form.

Any pregnancy that occurs in a female partner of a male study participant should be reported to BMS. Information on this pregnancy will be collected on the Pregnancy Surveillance Form.

6.5 Overdose

All occurrences of overdose must be reported as SAEs (see Section 6.1.1 for reporting details).

6.6 Potential Drug Induced Liver Injury (DILI)

Wherever possible, timely confirmation of initial liver-related laboratory abnormalities should occur prior to the reporting of a potential DILI event. All occurrences of potential DILIs, meeting the defined criteria, must be reported as SAEs (see Section 6.1.1 for reporting details).

Potential drug induced liver injury is defined as:

ALT or AST elevation > 3 times upper limit of normal (ULN)

AND

Total bilirubin > 2 times ULN, without initial findings of cholestasis (elevated serum alkaline phosphatase),

AND

No other immediately apparent possible causes of AT elevation and hyperbilirubinemia, including, but not limited to, viral hepatitis, pre-existing chronic or acute liver disease, or the administration of other drug(s) known to be hepatotoxic.

6.7 Other Safety Considerations

Any significant worsening noted during interim or final physical examinations, electrocardiogram, x-ray filming, any other potential safety assessment required or not required by protocol should also be recorded as a nonserious or serious AE, as appropriate, and reported accordingly.

7 DATA MONITORING COMMITTEE AND OTHER EXTERNAL COMMITTEES

To provide independent oversight of safety, efficacy, and study conduct, a data monitoring committee (DMC) will be instituted. The DMC will meet regularly to ensure that subject safety is carefully monitored. The DMC will convene additional ad hoc meetings if necessary. Following each meeting, the DMC will recommend continuation, modification, or discontinuation of the study based on observed toxicities. The DMC will also review the interim analysis results and inform BMS whether stopping criteria for superiority are met at that time. A separate DMC charter will describe the activities of this committee in more detail.

IRRC assessments will be utilized in this study for determination for PFS and ORR endpoints. The IRRC will review all available tumor assessment scans for all randomized subjects. Details of IRRC responsibilities and procedures will be specified in the IRRC charter.

8 STATISTICAL CONSIDERATIONS

8.1 Sample Size Determination

The sample size of the study accounts for the three co-primary efficacy endpoints: ORR based on IRRC assessments, PFS based on IRRC assessments and OS, evaluated in intermediate and poor-risk subjects with previously untreated mRCC. The overall alpha for this study is 0.05, which is split with 0.001 to evaluate ORR, 0.009 to evaluate PFS and 0.04 to evaluate OS.

ORR will be analyzed initially on a descriptive basis and will occupy an administrative adjustment of alpha of 0.001. PFS will be evaluated for treatment effect at an alpha of 0.009 (two-sided, penalized 0.001 from a 0.01 allocation) with at least 90% power; no interim analysis of PFS is planned. OS will be evaluated for treatment effect at an alpha level of 0.04 (two-sided) with 90% power, accounting for two formal interim analyses to assess efficacy.

It is estimated that approximately 1070 previously untreated mRCC subjects will be randomized in a 1:1 ratio. Among them, 820 subjects (76.6%) with intermediate/poor risk subjects and approximately 250 (23.4%) subjects with favorable risk as per IMDC (IMDC prognostic score = 0) will be randomized. Assuming a fixed accrual rate of 57 subjects per month (40 intermediate/poor risk subjects per month), it will take 20.5 months to randomize 1070 subjects (820 intermediate/poor risk subjects).

Assuming a 21% screen failure rate, it is estimated that approximately 1355 subjects will be enrolled in order to have 820 intermediate/poor-risk subjects randomized. The primary analysis is based on intermediate/poor risk subjects as per IMDC prognostic score and the number of PFS/OS events observed among them. The enrollment will stop once approximately 820 intermediate/poor risk subjects have been randomized regardless of the number of favorable risk subjects.

Sample size justification for ORR estimate

One of the primary objectives of the study is to describe the ORR (based on IRRC assessment) of nivolumab combined with ipilimumab to sunitinib monotherapy in intermediate and poor-risk subjects with previously untreated mRCC.

The primary analysis of ORR in the intermediate and poor-risk randomized subjects will be performed when these patients have an approximate 6 month minimum follow-up from the completion of enrollment. This will allow sufficient follow-up for ORR to have a stable estimate, adequate safety follow-up as well as information on duration of response in this population.

The maximum width of the exact two-sided 95% confidence interval (CI) is 9.9% when the ORR is expected to be in the 20% to 50% range. Table 8.1-1 summarizes the 95% exact CI when observed ORRs are 20% to 50%, respectively.

Table 8.1-1: Observed ORR with exact 95% CI

Observed ORR	95% Exact CI
20%	(16.2% - 24.2%)
25%	(21.0% - 29.6%)
30%	(25.6% - 34.7%)
35%	(30.5% - 40.0%)
40%	(35.2% - 44.9%)
45%	(40.2% - 50.1%)
50%	(45.1% - 54.9%)

For example, if at least 123 responders are observed among the 410 nivolumab and ipilimumab combination intermediate/poor risk randomized subjects (ie, ORR ≥ 30%) then the lower bound of the 95% CI is above 25.6%.

Sample size justification for PFS comparison

One of the primary objectives of the study is to compare the progression-free survival (based on IRRC assessment) of nivolumab combined with ipilimumab to sunitinib monotherapy in intermediate and poor-risk subjects with previously untreated mRCC. The number of events and power for this study were calculated assuming an exponential distribution for PFS in each arm.

For this comparison of PFS, it will be required to observe at least 591 PFS events among the randomized intermediate/poor risk subjects in the two respective treatment arms for a two-sided experiment-wise $\alpha = 0.009$ log-rank test, to show a statistically significant difference in PFS between the treatment arms with at least 90% power when the true hazard ratio of the experimental arm to control arm is 0.73. The HR of 0.73 is equivalent to demonstrating a 37.8% improvement in median PFS, assuming a median PFS of 9 months (weighted median estimate assuming a median PFS of 11 months in intermediate risk subjects and a median PFS of 4 months in poor risk subjects)[42] in the sunitinib monotherapy arm and a median PFS of 12.4 months in the experimental treatment arm.

Under the assumptions for accrual and PFS distribution stated above, it will take approximately 31 months from FPFV to observe the required number of PFS events for the final PFS analysis (20.5 months for accrual and 10.5 months for minimum follow up). It is projected that an observed HR of 0.807 or less corresponding to a 2.1 month or greater improvement in median PFS (9 vs 11.1 months) for this comparison, would result in a statistically significant improvement in the final analysis of PFS.

Sample size justification for OS comparison:

One of the primary objectives of the study is to compare the overall survival of nivolumab combined with ipilimumab to sunitinib monotherapy in intermediate and poor-risk subjects with previously untreated mRCC. The number of events and power of this study were calculated assuming an exponential distribution for OS in each arm.

Approximately 639 events (ie, deaths), observed among the randomized intermediate/poor risk subjects, provides 90% power to detect a hazard ratio (HR) of 0.766 with an overall type 1 error of 0.04 (two-sided). The HR of 0.766 corresponds to a 30.6% increase in the median OS, assuming a median OS of 20 months for sunitinib monotherapy (weighted median estimate assuming a median OS of 26 months in intermediate risk subjects and a median OS of 8 months in poor risk subjects)[43] and 26.1 months for experimental treatment arms respectively. It is projected that an observed hazard ratio of 0.846 or less, which corresponds to a 3.6 months or greater improvement in median OS (20 mo vs. 23.6 mo), would result in a statistically significant improvement in OS for the experimental arm at the final OS analysis.

Two formal interim analyses of OS are planned for this study. The first interim analysis is planned at the time of final PFS analysis and it is expected to observe 370 events (58% of the targeted OS events for final analysis) and the second after observing 479 events (75% of targeted OS events needed for final analysis). The stopping boundaries at interim and final analyses will be derived based on the number of deaths using O'Brien and Fleming α spending function.

Under the assumptions stated above on accrual and OS distribution, it will approximately take 61 months from FPFV to observe the required number of OS events for the final OS analysis (20.5 months for accrual and 40.5 months for minimum follow up).

In summary, it is expected to take:

- Approximately 20.5 months to complete accrual

- Approximately 27 months from FPFV to obtain a minimum follow-up of 6 months on the intermediate and poor risk randomized subjects for the descriptive analysis of ORR (PFS and OS will not be analyzed until sufficient events have occurred)
- Approximately 31 months from FPFV to obtain the required number of PFS events (ie, at least 591 events among the 820 intermediate and poor risk randomized subjects) and deaths for the first formal interim analysis of OS (ie, approximately 370 deaths among the same population)
- Approximately 40 months from FPFV to obtain the required deaths for the second formal interim analysis of OS (ie, 479 deaths among the intermediate and poor risk randomized subjects)
- Approximately 61 months from FPFV to obtain the required deaths for the final analysis of OS (ie, 639 deaths among the intermediate and poor risk randomized subjects).

Table 8.1-2 summarizes sample size design parameters and schedule of primary endpoint analyses planned in this study.

Table 8.1-2: Summary of sample size parameters and schedule of analyses

Co-Primary Endpoints	ORR	PFS	OS
Primary analysis population	Intermediate/poor risk subjects (IMDC score ≥ 1)		
Accrual rate per month	40		
Power	N/A	90%	90%
Alpha	Administrative 0.001	0.009 2-sided	0.04 2-sided (0.0045 at IA1, 0.0131 at IA2 , 0.0354 at FA)
Hypothesized Median Control vs. exp (months)	25% vs 40%	9 vs. 12.4	20 vs. 26.1
Hypothesized Hazard ratio	N/A	0.726	0.766
Critical Hazard ratio (Observed hazard ratio at which a statistically significant difference would be observed) / Difference in median (months) Corresponding to a minimal clinically significant effect size	N/A	0.807 / 2.1	0.846/ 3.6
Critical HR at interim analysis-1(IA1) /effect size	N/A	N/A	0.74/ 6.9
Expected number of event for IA1 (percentage of target events)	N/A	N/A	370(58%)
Timing of IA1 from FPFV l(months)	N/A	N/A	31
Critical HR at interim analysis-2(IA2) /effect size	N/A	N/A	0.8/ 5.1
Target number of event for IA2 (percentage of target events)	N/A	N/A	479 (75%)
Timing of IA2 from FPFV l(months)	N/A	N/A	40

Table 8.1-2: Summary of sample size parameters and schedule of analyses

Co-Primary Endpoints	ORR	PFS	OS
Accrual Duration (months)	20.5	20.5	20.5
Timing of final analysis(FA) from FPFV (months)	27	31	61
Sample size[a]	820	820	820
Target number of events(Event Goal)	N/A	591	639

[a] East version 5.4 was used for sample size / power computation

8.2 Populations for Analyses

- **All enrolled subjects**: All subjects who signed an informed consent form and were registered into the IVRS.
- **All randomized subjects (any risk subjects)**: All subjects who were randomized to any treatment arm in the study. This population is considered as the secondary efficacy analysis population. Analysis of demography, protocol deviations, baseline characteristics, secondary efficacy analysis and outcome research analysis will be performed for this population.
- **Intermediate/poor risk subjects**: All randomized subjects with baseline IMDC prognostic score ≥ 1 at the time randomization. This is the primary efficacy analysis population. Analysis of demography, protocol deviations, baseline characteristics and primary efficacy analysis will be performed for this population.
- **All treated subjects**: All subjects who received any dose of study therapy. This is the primary dataset for drug exposure and safety analysis.
- **Favorable risk subjects**: All randomized subjects with baseline IMDC prognostic score = 0 at the time randomization. This population would be used for conducting exploratory analysis of efficacy endpoints.
- **PK subjects**: All subjects with available serum time-concentration data from randomized subjects dosed with nivolumab.
- **Immunogenicity subjects**: All subjects with available data from randomized subjects dosed with nivolumab.

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

8.3 Endpoints

The primary objectives of this study to describe ORR and is to compare PFS (based on IRRC assessment) and OS of nivolumab combined with ipilimumab to sunitinib monotherapy in intermediate and poor-risk subjects with previously untreated mRCC. This is measured by the three co-primary endpoints defined in Section 8.3.1.

The first secondary objective of this study is to compare PFS (based on IRRC assessment) in the two treatment arms in the all randomized population. This would be measured by the same definitions of PFS, as specified in sections 8.3.1.1 and 8.3.1.3 respectively, in the all randomized population.

The second secondary objective of this study is to compare OS in the two treatment arms in the all randomized population. This would be measured by the same definition of OS, as specified in section 8.3.1.4, in the all randomized population.

The third secondary objective of this study is assessing ORR in the two treatment arms in all randomized population. This would be measured by the definition of ORR as specified in section 8.3.2.1, in intermediate/poor risk subjects and all randomized population respectively.

8.3.1 Co-Primary Endpoints

Object response rate, progression free survival, and overall survival are the co-primary endpoints.

8.3.1.1 Objective Response Rate

Objective response rate is defined as the proportion of randomized subjects who achieve a best response of complete response (CR) or partial response (PR) using the RECIST1.1 criteria based on IRRC assessment. BOR is defined as the best response designation, as determined by the IRC, recorded between the date of randomization and the date of objectively documented progression per RECIST 1.1 or the date of subsequent therapy (including tumor-directed radiotherapy and tumor-directed surgery), whichever occurs first. For subjects without documented progression or subsequent therapy, all available response designations will contribute to the BOR assessment. As described in Section 5.4, confirmation of response is required. Duration of response (DOR) is defined as the time between the date of first documented response (CR or PR) to the date of the first documented progression as determined by the IRC (per RECIST 1.1), 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 (Section 8.3.1.3). Time to Objective Response (TTR) is defined as the time from randomization to the date of the first confirmed documented response (CR or PR), as assessed by the IRC. DOR and TTR will be evaluated for responders (confirmed CR or PR) only.

8.3.1.2 Primary Definition of Progression-Free Survival

The primary definition PFS is specified as the time between the date of randomization and the first date of documented progression, based on IRRC assessment (as per RECIST 1.1 criteria), or death due to any cause, whichever occurs first. Subjects who die without a reported progression will be considered to have progressed on the date of their death. The following censoring rules will be applied for the primary definition of PFS.

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

- Subjects who receive subsequent systemic anti-cancer therapy prior to documented progression will be censored at the date of the last tumor assessment prior to the initiation of the new therapy.

8.3.1.3 Secondary Definition of Progression-Free Survival

The secondary definition of PFS is defined as the time between the date of randomization and the first date of documented progression, based on IRRC assessment (as per RECIST 1.1 criteria), or death due to any cause, whichever occurs first. Subjects who die without a reported progression will be considered to have progressed on the date of their death. The following censoring rules will be applied for the primary definition of PFS.

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

8.3.1.4 Overall Survival

Overall survival is defined as the time from randomization to the date of death from any cause. For subjects that are alive, their survival time will be censored at the date of last contact ("last known alive date"). Overall survival will be censored for subjects at the date of randomization if they were randomized but had no follow-up.

8.3.2 Secondary Endpoint(s)

8.3.2.1 Adverse Event Incidence Rate

Adverse events incident rate is defined as the proportion subjects with any grade adverse events among subjects treated in each treatment arm. Events reported from the first dose and up to and including 100 days following the last dose of study treatment could be included in estimating this incidence rate.

8.3.3 Exploratory Endpoint(s)

Safety and tolerability objective will be measured by the incidence of adverse events, serious adverse events, deaths and laboratory abnormalities. Toxicities will be graded using the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) version 4.0.

PFS (based on IRRC assessment) and OS will be estimated in the two treatment arms among subjects with favorable risk as per IMDC prognostic criteria.

Nivolumab PK will be characterized using nivolumab serum concentration, which will also be used to determine exposure in each subject. Samples will be collected to characterize pharmacokinetics of nivolumab and these exposures will be used to explore exposure-safety and exposure-efficacy relationships.

HRQoL will be assessed by FACT-G and FKSI-19. Global health status will be assessed by EQ-5D instrument.

Disease-related symptom progression rate is defined as the proportion of subjects who have disease-related symptom progression as measured by the FKSI-19. The minimum important change in the FKSI-19 used to define symptom progression is approximately a change of two points and that definition has been used for this mRCC symptom scale in other mRCC trials. Disease-related symptom progression is defined as a decrease of two points in the FKSI-19 relative to the subject's baseline FKSI-19 score without returning to above that point during the remainder of the study.

8.4 Analyses

8.4.1 Demographics and Baseline Characteristics

Demographics and baseline characteristics will be summarized by treatment arm as randomized using descriptive statistics for intermediate/poor risk subjects as well as all randomized subjects.

8.4.2 Efficacy Analyses

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

The key secondary objectives will be tested in the following hierarchical order, after conducting the primary objective analyses:

- PFS based on IRRC assessment among all randomized subjects
- OS among all randomized subjects

The formal testing of PFS based on IRRC assessment, at a two sided 0.009 significance level, among all randomized subjects will take place if PFS based on IRRC assessment among intermediate/poor risk subjects is statistically significant. Likewise, the testing of OS, at a two sided 0.04 significance level, among all randomized subjects will take place only if OS intermediate/poor risk subjects are statistically significant. The detail of the testing procedure will be specified in the statistical analysis plan.

8.4.2.1 Objective Response Rate: Co-primary Endpoint

One of the primary objectives of the study is to describe the objective response rate per IRRC in the two treatment arms among intermediate and poor risk subjects. The ORR analysis will occupy a 0.001 administrative allocation of alpha.

The number and percentage of subjects in each category of best overall response per IRRC (complete response [CR], partial response [PR]], stable disease [SD], progressive disease [PD], or unable to determine [UD]) according to the IRRC will be presented, by treatment group. An estimate of the response rate and an associated exact two-sided 95% CI (Clopper and Pearson[44]) will be presented, by treatment group.

Sensitivity analysis based on investigator-determined ORR may also be performed. DOR and TTR will also be evaluated. Descriptive analysis of the response in the investigator's choice group (ie, subjects treated with investigator's choice among ORR population) will also be provided.

At the time of the formal ORR analysis, no PFS or OS analysis will be conducted because of the immaturity of those specific endpoints. A reduced analysis will be defined in the data presentation plan.

8.4.2.2 Progression-Free Survival Analysis: Co-primary Endpoint

One of the primary objectives of the study is to compare the progression-free survival (based on IRRC assessment) of nivolumab combined with ipilimumab to sunitinib monotherapy in intermediate and poor-risk subjects with previously untreated mRCC. All analyses specified in this section will be conducted for PFS derived as per primary and secondary definitions.

The primary formal comparison of PFS will be conducted using a two-sided 0.009 stratified log-rank test, with IMDC Prognostic Score (1-2 vs 3-6) and Region (US vs Canada/W Europe/N Europe vs ROW) as stratification factors among intermediate/poor risk subjects.

Median PFS will be estimated via the Kaplan-Meier product limit method. Two-sided 95% CI for the median PFS will be computed for each randomized arm. Kaplan-Meier plots of PFS will be presented. Hazard ratios (HR) and corresponding two-sided 99.1% confidence intervals (CI) will be estimated using a Cox proportional hazards model, with treatment arm as a single covariate, stratified by the stratification factors, corresponding to the comparison of PFS.

The totality of PFS results will be presented in a single graphical display that includes Kaplan-Meier curves for the two treatment arms, the log-rank p-values for the formal comparison, the HRs and corresponding CI, and the median PFS estimates and corresponding CIs.

The following supportive analyses of PFS will also be conducted:

A stratified multivariate Cox regression model will be used in order to estimate the treatment effect after adjustment for possible imbalances in known or potential prognostic factors. The covariates included in this model will be specified in the statistical analysis plan.

PFS using an un-stratified log rank test. The hazard ratio associated with treatment will be presented along with the associated two-sided 95% CIs.

8.4.2.3 Overall Survival Analysis: Co-primary Endpoint

One of the primary objectives of the study is to compare the overall survival of nivolumab combined with ipilimumab to sunitinib monotherapy in intermediate and poor-risk subjects with previously untreated mRCC. OS will be compared between the treatment arms using a two sided, $\alpha = 0.04$ level log-rank test (adjusted for interim analyses), stratified using the same factor as in PFS. A similar analysis as in PFS will be conducted for OS. Hazard ratios (HR) and corresponding two-sided 96% confidence intervals (CI) will be estimated using a Cox

proportional hazards model, with treatment arm as a single covariate, stratified by the stratification factors, corresponding to the comparison of OS.

8.4.2.4 Progression-Free Survival Analysis: Secondary Objective

One of the secondary objectives of the study is to compare the progression-free survival (based on IRRC assessment) of nivolumab combined with ipilimumab to sunitinib monotherapy in all randomized with previously untreated mRCC. All analyses specified in this section will be conducted for PFS derived as per primary and secondary definitions.

A formal comparison of PFS in all randomized subjects will be conducted using a two-sided 0.009 stratified log-rank test, with IMDC Prognostic Score (0 vs 1-2 vs 3-6) and Region (US vs Canada/W Europe/N Europe vs ROW) as stratification factors among all randomized subjects, only if the PFS comparison among intermediate/poor risk subjects towards the primary objective assessment is found to be statistically significant. Analyses of PFS among all randomized subjects will be similar to those conducted towards the assessment of the primary PFS objective.

8.4.2.5 Overall Survival Analysis: Secondary Objective

One of the secondary objectives of the study is to compare the overall survival of nivolumab combined with ipilimumab to sunitinib monotherapy in all randomized with previously untreated mRCC.

A formal comparison of OS in all randomized subjects will be conducted using a two-sided 0.04 stratified log-rank test, with IMDC Prognostic Score (0 vs 1-2 vs 3-6) and Region (US vs Canada/W Europe/N Europe vs ROW) as stratification factors among all randomized subjects, only if the OS comparison among intermediate/poor risk subjects towards the primary objective assessment is found to be statistically significant. Analyses of OS among all randomized subjects will be similar to those conducted towards the assessment of the primary OS objective.

8.4.2.6 Objective Response Rate Analysis: Secondary Objective

One of the secondary objectives of the study is to estimate the objective response rate in the two treatment arms among all randomized subjects.

Estimates of response rate, along with its exact two-sided 95% CI by Clopper-Pearson method, will be computed within each treatment arm. A two sided, 95% CI for difference of response rate between the treatment arms will also be computed.

8.4.3 Safety Analyses

The safety analysis will be performed in all treated subjects. Descriptive statistics of safety will be presented using the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) version 4.0 by treatment arm. All AEs, drug-related AEs, SAEs and drug-related SAEs will be tabulated using the worst grade per NCI CTCAE v 4.0 criteria by system organ class and preferred term. On-study lab parameters including hematology, coagulation, chemistry, liver function and renal function will be summarized using worse grade per NCI CCAE v 4.0 criteria.

8.4.3.1 AE Incidence Rate Analysis: Secondary Objective

One of the secondary objectives of the study is to estimate the adverse event incidence rate in the two treatment arms among treated subjects.

Estimates of AE incidence rate, along with its exact two-sided 95% CI by Clopper-Pearson method, will be computed within each treatment arm.

8.4.4 Pharmacokinetic Analyses

The nivolumab and ipilimumab concentration data obtained in this study may be combined with data from other studies in the clinical development programs to develop or refine a population PK model. These models may be used to evaluate the effects of intrinsic and extrinsic covariates on the PK of nivolumab and ipilimumab and to determine measures of individual exposure. In addition, model determined exposures may be used for exposure-response analyses.

8.4.5 Biomarker Analyses

8.4.5.1 Pharmacodynamic Analyses

To assess pharmacodynamic effects in serum, blood RNA, or peripheral cells obtained from subjects on each treatment arm, summary statistics for biomarkers of immunoregulatory activity (eg, IFN-inducible proteins, miRNAs, gene expression, immune cells) and their corresponding changes (or percent changes) from baseline will be tabulated by planned study visit. In addition, the time course of biomarker outcomes will be investigated graphically. If there is indication of a meaningful pattern across time, further analysis may be completed to characterize the relationship. Possible associations between changes in biomarker measures of interest and exposure to study drug will be explored graphically.

8.4.5.2 Pharmacogenomic and Exploratory Analyses

Potential relationships between biomarker data and efficacy or safety endpoints will be investigated as part of an analysis plan aimed at identifying baseline biomarkers that may be used to prospectively identify patients likely (or not likely) to respond to nivolumab and to identify subjects who may be predisposed to having adverse reactions to treatment. These exploratory predictive biomarker analyses will be completed with biomarkers measured in blood and in tumor samples and will focus primarily, as outlined in the exploratory objectives, on SNPs in select genes associated with immunity or on the expression of selected proteins in tumor specimens, such as PD-1, PD-L1, and PD-L2. Similar analyses will be completed with data regarding serum-soluble factors, blood RNA and/or immune cell types.

Associations between biomarkers and efficacy measures will be analyzed on all subjects treated with at least one dose of study medication and with corresponding efficacy and biomarker measurements. Efficacy measures will include response, PFS, and OS. Demographic and case-history factors will be examined to determine whether stratification or adjustments should be made within the subsequent statistical analyses, and if necessary, the appropriate stratification or adjustment will be made.

Biomarkers will be summarized graphically as they relate to efficacy and safety endpoints, as applicable. Summary statistics will be tabulated. SNP allele frequencies will be summarized. The relationships between binary measures (eg, response) and candidate biomarkers will be investigated using logistic regression. Associations will be summarized in terms of point and interval estimates of hazard ratios, odds ratios, or other statistics, as appropriate for the analyses completed. Models to predict clinical activity based on combinations of biomarkers may also be investigated.

Additional post hoc statistical analyses not specified in the protocol, such as alternative modeling approaches may be completed. All analyses described in this section are based on the availability of the data

8.4.6 Outcomes Research Analyses

Descriptive summary statistics of Quality of life (QoL) assessments will be presented at baseline and each on-study time point, unless otherwise specified. Mean changes from baseline for each of the three scales will be calculated for each subject at each on-study time point. In addition, subject compliance will be described per time point by the proportion of subjects who filled out the QoL assessments over the numbers of subject known to be alive and eligible for assessment at these time points.

8.4.7 Other Analyses

8.4.7.1 Immunogenicity Analyses

Immunogenicity may be reported for ADA positive status (such as persistent positive, other positive, only last sample positive, baseline positive) and ADA negative status, relative to baseline. In addition, presence of neutralizing antibody may be reported, if applicable. Effect of immunogenicity on safety/efficacy and biomarkers and PK may be explored.

8.4.7.2 Resource Utilization

Resource Utilization analysis will be described in a separate statistical analysis plan.

8.5 Interim Analyses

Two interim analyses of OS are planned. First interim analysis is scheduled at the time of final PFS analysis and it is expected after observing 370 deaths (approx 58% of the targeted OS events) and the second interim analysis is scheduled after 478 deaths (approx 75% of total deaths) have been observed among intermediate/poor risk subjects based on above accrual rate and the exponential distribution in each arm. These formal comparisons of OS will allow for early stopping for superiority, and the boundaries for declaring superiority will be derived based on the actual number of deaths using Lan-DeMets spending function with O'Brien and Fleming type of boundary in EAST v5.4. If the first interim analysis is performed exactly at 370 deaths , the boundary in terms of statistical significance for declaring superiority would be 0.0045 (HR=0.744 , 6.9 months improvement in median OS) and if the second interim analysis is performed at exactly 479 deaths, the boundary in terms of statistical significance at the interim analysis for declaring superiority would be 0.0131 (or 0.8 with regard to HR boundary, which corresponds to 5.1 months improvement in median OS under the assumed control arm hazard

function). The boundary for declaring superiority in terms of statistical significance for the final analysis after 639 events would be 0.0354. An independent statistician external to BMS will perform interim analysis.

In addition to the formal planned interim analysis for OS, the DMC will have access to periodic un-blinded interim reports of efficacy and safety to allow a risk/benefit assessment. Details will be included in the DMC charter.

9 STUDY MANAGEMENT

9.1 Compliance

9.1.1 Compliance with the Protocol and Protocol Revisions

The study shall be conducted as described in this approved protocol. All revisions to the protocol must be discussed with, and be prepared by, BMS. The investigator should not implement any deviation or change to the protocol without prior review and documented approval/favorable opinion from the IRB/IEC and Regulatory Authority(ies), if required by local regulations, of an amendment, except where necessary to eliminate an immediate hazard(s) to study subjects.

If a deviation or change to a protocol is implemented to eliminate an immediate hazard(s) prior to obtaining IRB/IEC approval/favorable opinion, as soon as possible the deviation or change will be submitted to:

- IRB/IEC for review and approval/favorable opinion
- BMS
- Regulatory Authority(ies), if required by local regulations

Documentation of approval signed by the chairperson or designee of the IRB(s)/IEC(s) must be sent to BMS.

If an amendment substantially alters the study design or increases the potential risk to the subject: (1) the consent form must be revised and submitted to the IRB(s)/IEC(s) for review and approval/favorable opinion; (2) the revised form must be used to obtain consent from subjects currently enrolled in the study if they are affected by the amendment; and (3) the new form must be used to obtain consent from new subjects prior to enrollment.

If the revision is done via an administrative letter, investigators must inform their IRB(s)/IEC(s).

9.1.2 Monitoring

BMS representatives will review data centrally to identify potential issues to determine a schedule of on-site visits for targeted review of study records.

Representatives of BMS must be allowed to visit all study site locations periodically to assess the data quality and study integrity. On site they will review study records and directly compare them with source documents, discuss the conduct of the study with the investigator, and verify that the facilities remain acceptable. Certain CRF pages and/or electronic files may serve as the source documents.

In addition, the study may be evaluated by BMS internal auditors and government inspectors who must be allowed access to CRFs, source documents, other study files, and study facilities. BMS audit reports will be kept confidential.

The investigator must notify BMS promptly of any inspections scheduled by regulatory authorities, and promptly forward copies of inspection reports to BMS.

9.1.2.1 Source Documentation

The Investigator is responsible for ensuring that the source data are accurate, legible, contemporaneous, original and attributable, whether the data are hand-written on paper or entered electronically. If source data are created (first entered), modified, maintained, archived, retrieved, or transmitted electronically via computerized systems (and/or any other kind of electronic devices) as part of regulated clinical trial activities, such systems must be compliant with all applicable laws and regulations governing use of electronic records and/or electronic signatures. Such systems may include, but are not limited to, electronic medical/health records (EMRs/EHRs), adverse event tracking/reporting, protocol required assessments, and/or drug accountability records).

When paper records from such systems are used in place of electronic format to perform regulated activities, such paper records should be certified copies. A certified copy consists of a copy of original information that has been verified, as indicated by a dated signature, as an exact copy having all of the same attributes and information as the original.

9.1.3 Investigational Site Training

Bristol-Myers Squibb will provide quality investigational staff training prior to study initiation. Training topics will include but are not limited to: GCP, AE reporting, study details and procedure, electronic CRFs, study documentation, informed consent, and enrollment of WOCBP.

9.2 Records

9.2.1 Records Retention

The investigator must retain all study records and source documents for the maximum period required by applicable regulations and guidelines, or institution procedures, or for the period specified by BMS, whichever is longer. The investigator must contact BMS prior to destroying any records associated with the study.

BMS will notify the investigator when the study records are no longer needed.

If the investigator withdraws from the study (eg, relocation, retirement), the records shall be transferred to a mutually agreed upon designee (eg, another investigator, IRB). Notice of such transfer will be given in writing to BMS.

9.2.2 Study Drug Records

It is the responsibility of the investigator to ensure that a current disposition record of investigational product (inventoried and dispensed) is maintained at the study site. Records or logs must comply with applicable regulations and guidelines and should include:

- amount received and placed in storage area
- amount currently in storage area
- label identification number or batch number
- amount dispensed to and returned by each subject, including unique subject identifiers
- amount transferred to another area/site for dispensing or storage
- nonstudy disposition (eg, lost, wasted)
- amount destroyed at study site, if applicable
- amount returned to BMS
- retain samples for bioavailability/bioequivalence, if applicable
- dates and initials of person responsible for Investigational Product dispensing/accountability, as per the Delegation of Authority Form.

BMS will provide forms to facilitate inventory control if the investigational site does not have an established system that meets these requirements.

9.2.3 Case Report Forms

An investigator is required to prepare and maintain adequate and accurate case histories designed to record all observations and other data pertinent to the investigation on each individual treated or entered as a control in the investigation. Data that are derived from source documents and reported on the CRF must be consistent with the source documents or the discrepancies must be explained. Additional clinical information may be collected and analyzed in an effort to enhance understanding of product safety. CRFs may be requested for AEs and/or laboratory abnormalities that are reported or identified during the course of the study.

For sites using the BMS electronic data capture tool, electronic CRFs will be prepared for all data collection fields except for fields specific to SAEs and pregnancy, which will be reported on the paper or electronic SAE form and Pregnancy Surveillance form, respectively. Spaces may be left blank only in those circumstances permitted by study-specific CRF completion guidelines provided by BMS.

The confidentiality of records that could identify subjects must be protected, respecting the privacy and confidentiality rules in accordance with the applicable regulatory requirement(s).

The investigator will maintain a signature sheet to document signatures and initials of all persons authorized to make entries and/or corrections on CRFs.

The completed CRF, including any paper or electronic SAE/pregnancy CRFs, must be promptly reviewed, signed, and dated by the investigator or qualified physician who is a subinvestigator and who is delegated this task on the Delegation of Authority Form .For electronic CRFs, review and approval/signature is completed electronically through the BMS electronic data capture tool. The investigator must retain a copy of the CRFs including records of the changes and corrections.

Each individual electronically signing electronic CRFs must meet BMS training requirements and must only access the BMS electronic data capture tool using the unique user account provided by BMS. User accounts are not to be shared or reassigned to other individuals.

9.3 Clinical Study Report and Publications

A Signatory Investigator must be selected to sign the clinical study report.

For this protocol, the Signatory Investigator will be selected as appropriate based on the following criteria:

- Subject recruitment (eg, among the top quartile of enrollers)
- Involvement in trial design
- Other criteria (as determined by the study team)

The data collected during this study are confidential and proprietary to BMS. Any publications or abstracts arising from this study require approval by BMS prior to publication or presentation and must adhere to BMS's publication requirements as set forth in the approved clinical trial agreement (CTA). All draft publications, including abstracts or detailed summaries of any proposed presentations, must be submitted to BMS at the earliest practicable time for review, but at any event not less than 30 days before submission or presentation unless otherwise set forth in the CTA. BMS shall have the right to delete any confidential or proprietary information contained in any proposed presentation or abstract and may delay publication for up to 60 days for purposes of filing a patent application.

10 GLOSSARY OF TERMS

Term	Definition
Complete Abstinence	If one form of contraception is required, Complete Abstinence is defined as complete avoidance of heterosexual intercourse and is an acceptable form of contraception for all study drugs. Female subjects must continue to have pregnancy tests. Acceptable alternate methods of highly effective contraception must be discussed in the event that the subject chooses to forego complete abstinence. If two forms of contraception is required, Complete abstinence is defined as complete avoidance of heterosexual intercourse and is an acceptable form of contraception for all study drugs. Subjects who choose complete abstinence are not required to use a second method of contraception, but female subjects must continue to have pregnancy tests. Acceptable alternate methods of highly effective contraception must be discussed in the event that the subject chooses to forego complete abstinence. **Expanded definition** Complete abstinence as defined as complete avoidance of heterosexual intercourse is an acceptable form of contraception for all study drugs. This also means that abstinence is the preferred and usual lifestyle of the patient. This does not mean periodic abstinence (eg, calendar, ovulation, symptothermal, profession of abstinence for entry into a clinical

Term	Definition
	trial, post-ovulation methods) and withdrawal, which are not acceptable methods of contraception. Subjects who choose complete abstinence are not required to use a second method of contraception, but female subjects must continue to have pregnancy tests. Acceptable alternate methods of highly effective contraception must be discussed in the event that the subject chooses to forego complete abstinence

11 LIST OF ABBREVIATIONS

Abbreviation	Term
ADA	Anti-drug antibody
AE	Adverse event
AJCC	American Joint Committee on Cancer
ALT	Alanine transaminase
AST	Aspartate transaminase
BMS	Bristol-Myers Squibb
BOR	Best overall response
BUN	Blood urea nitrogen
CMV	Cytomegalovirus
CR	Complete response
CRC	Colorectal cancer
CrCl	Creatinine clearance
CRF	Case report form
CT	Computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
DCF	Data clarification form
DILI	Drug-induced liver injury
DLT	Dose-limiting toxicity
DMC	Data monitoring committee
ECG	Electrocardiogram
eCRF	Electronic case report form
EDC	Electronic data capture
ELISA	Enzyme-linked immunosorbent assay
EOI	End of infusion
FFPE	Formalin-fixed paraffin-embedded
FKSI	Functional Assessment of Cancer Therapy-Kidney Symptom Index
FPFV	First Patient First Visit
FSH	Follicle-stimulating hormone
FU	Follow-up

Abbreviation	Term
GCP	Good clinical practices
GMP	Good manufacturing practices
HCV	Hepatitis C virus
HBV	Hepatitis B virus
HDL	High-density lipoprotein
HIFα	Hypoxia inducible factor α
HIPAA	Health Information Portability and Accountability Act
HRT	Hormone replacement therapy
HRU	Health Resource Utilization
ICF	Informed consent form
ICH	International Conference on Harmonisation
IHC	Immunohistochemistry
IMDC	International Metastatic RCC Database Consortium
ITIM	Immunoreceptor tyrosine inhibitory motif
ITSM	Immunoreceptor tyrosine-based switch motif
IV	Intravenous
IFN	Interferon
IRB/IEC	Institutional review board/independent ethics committee
IRRC	Independent Radiology Review Committee
IVRS	Interactive voice response system
KPS	Karnofsky Performance Score
LDL	Low-density lipoprotein
LFT	Liver function test
mAb	Monoclonal antibody
MedDRA	Medical Dictionary for Regulatory Activities
MEL	Metastatic melanoma
miRNA	Micro-ribonucleic acid
MLR	Mixed lymphocyte reaction
MRI	Magnetic resonance imaging
MSKCC	Memorial Sloan-Kettering Cancer Center

Abbreviation	Term
MTD	Maximum-tolerated dose
mTOR	Mammalian target of rapamycin
NCI	National Cancer Institute
NSCLC	Non-small cell lung cancer
ORR	Objective response rate
OS	Overall survival
PBMC	Peripheral blood mononuclear cell
PD	Progressive disease
PD-1	Programmed death-1
PD-L1	Programmed death-ligand 1
PD-L2	Programmed death-ligand 2
PFS	Progression-free survival
PK	Pharmacokinetics
PO	Per os (by mouth)
PR	Partial response
PRO	Patient-reported outcome
RCC	Renal cell carcinoma
RECIST	Response Evaluation Criteria in Solid Tumors
RNA	Ribonucleic acid
RT-PCR	Reverse transcriptase - polymerase chain reaction
SAE	Serious adverse event
sAg	Surface antigen
SD	Stable disease
SNP	Single nucleotide polymorphism
SOP	Standard operating procedures
TCR	T-cell receptor
TKI	Tyrosine kinase inhibitor
TSH	Thyroid stimulating hormone
ULN	Upper limit of normal
US	United States

Abbreviation	Term
VEGF	Vascular endothelial growth factor
VEGFr	Vascular endothelial growth factor receptor
WOCBP	Women of child bearing potential

12 REFERENCES

REFERENCES FOR THE USE OF CONDOMS WITH SPERMICIDE.

Kestelman P. et. al., Efficacy of the Simultaneous Use of Condoms and Spermicides Family Planning Perspectives.Vol 23 (5); October 1991.

Gabbay MB, Thomas J, Gibbs A, Hold P. A Randomized Crossover Trial of The Impact of Additional Spermicide on Condom Failure Rates. Sex Transm Dis 2008; 35: 862-8.

LITERATURE REFERENCES

[1] Greenwald RJ, Freeman GJ, Sharpe AH. The B7 family revisited. *Annu Rev Immunol* 2004; 23:515-48.

[2] Freeman GJ, Long AJ, Iwai Y, et al. Engagement of the PD-1 immunoinhibitory receptor by a novel B7 family member leads to negative regulation of lymphocyte activation. *J Exp Med* 2000; 192(7):1027-34.

[3] Latchman Y, Wood CR, Chernova T, et al. PD-L2 is a second ligand for PD-1 and inhibits T cell activation. *Nat Immunol* 2001; 2(3):261-8.

[4] Carter LL, Fouser LA, Jussif J, et al. PD-1:PD-L inhibitory pathway affects both CD4+ and CD8+ T cells and is overcome by IL-2. *Eur J Immunol* 2002; 32(3):634-43.

[5] Chemnitz JM, Parry RV, Nichols KE, June CH, Riley JL. SHP -1 and SHP-2 associate with immunoreceptor tyrosine-based switch motif of programmed death 1 upon primary human T cell stimulation, but only receptor ligation prevents T cell activation. *J Immunol* 2004; 173(2):945-54.

[6] Sheppard KA, Fitz LJ, Lee JM, et al. PD-1 inhibits T-cell receptor induced phosphorylation of the ZAP70/CD3zeta signalosome and downstream signaling to PKC-theta. *FEBS Letters* 2004 Sep 10; 574(1-3):37-41.

[7] Nishimura H, Honjo T. PD-1: an inhibitory immunoreceptor involved in peripheral tolerance. *Trends Immunol*. 2001;22: 265-8.

[8] Effect of MDX-1106 on CD4+ T cells during an allogeneic mixed lymphocyte reaction (MLR) (Medarex Study No. MDX-1106-026-R). Bristol-Myers Squibb Pharmaceutical Research Institute; 2006. Document Control No. 930046581.

[9] Alegre ML, Noel PJ, Eisfelder BJ, et al. Regulation of surface and intracellular expression of CTLA4 on mouse T cells. J Immunol 1996; 157:4762-70.

[10] Postow MA, Harding J, Wolchok JD. Targeting immune checkpoints: Releasing the restraints on anti-tumor immunity for patients with melanoma. Cancer J 2012; 18:153-159.

[11] Curran MA, Montalvo W, Yagita H, et al. PD-1 and CTLA-4 combination blockade expands infiltrating T cells and reduces regulatory T and myeloid cells within B16 melanoma tumors. PNAS 2010; 107: 4275-80.

[12] Cancer Facts and Figures 2010. American Cancer Society. Atlanta: American Cancer Society; 2010.

[13] Cancer of the Kidney: Section 3. W. Marston Linehan, Brian I. Rini, James C. Yang, in Cancer : Principles & Practice of Oncology / edited by Vincent T. DeVita, Samuel Hellman, Steven A. Rosenberg. Philadelphia: Lipincott-Raven; 8th edition.

[14] Motzer RJ, Mazumdar M, Bacik J, et al. Survival and prognostic stratification of 670 patients with advanced renal cell carcinoma. J clin oncol. 1999; 17: 2530-2540.

[15] Heng DYC, Wanling X, Regan M, et al. Prognostic factors for overall survival in patients with metastatic renal cell carcinoma treated with vascular endothelial growth factor-targeted agents: Results from a large, multicenter study. J Clin Oncol. 2009; 27:5794-5799.

[16] Heng DYC, Wanling X, Regan MM, et al. External validation and comparison with other models of the International Metastatic Renal-Cell Carcinoma Database Consortium prognostic model: A population-based study. Lancet Oncol. 2013;14:141-148.

[17] Aldesleukin (IL-2) US Prescribing Information, April 2010.

[18] Fyfe G, Fisher RI, Rosenberg SA, et al. Results of treatment of 255 subjects with metastatic renal cell carcinoma who received high-dose recombinant interleukin-2 therapy. J Clin. Oncol. 1995; 139: 688-696.

[19] Coppin, C., Porzsolt F., Autenrieth M., Kumpf J., Coldman A., Wilt T., Immunotherapy for advanced renal cell cancer (Review). The Cochrane Library. 2008, Issue 4.

[20] Interferon alfa-2b, recombinant, Prescribing Information, August 2009.

[21] Mulders, P. Vascular endothelial growth factor and mTOR pathways in renal cell carcinoma: differences and synergies of two targeted mechanisms. BJU Int 2009; 104(11):1585-1589.

[22] NCCN Clinical Practice Guidelines in Oncology; Kidney Cancer. Version 2.2014. Accessed at http://www.nccn.org/professionals/physician_gls/pdf/kidney.pdf on 2-April-2014.

[23] Escudier B, Eisen T, et al. Renal cell carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2012; 23 Suppl 7 vii65-71.

[24] Motzer RJ, Hutson TE, Tomczak P, et al. Overall survival and updated results for sunitinib compared with interferon alfa in patients with metastatic renal cell carcinoma. J Clin Oncol 2009; 27: 3584-90.

[25] Motzer RJ, Hutson TE, Cella D, et al. Pazopanib versus sunitinib in metastatic renal-cell carcinoma. N Engl J Med 2013; 369:722-731.

[26] Sutent (sunitinib malate) US Prescribing Information. April 2014.

[27] Final Clinical Study Report for CA209010: A randomized, blinded, Phase 2 dose-ranging study of nivolumab (MDX-1106, BMS-936558) in subjects with progressive advanced/metastatic clear-cell renal cell carcinoma (RCC) who have received prior anti-angiogenic therapy. Bristol-Myers Squibb Company; 2013. Document Control No. 930074316.

[28] A Phase 2 Open-label Study of Single Agent Ipilimumab for the Treatment of IL-2 Refractory or IL-2 Ineligible Patients with Stage IV Renal Cancer; NCT00057889; Final Clinical Study Report 930025954 v1.0 (28-January-2008)

[29] Yang JC, Hughes M, Kammula U et al. Ipilimumab (Anti-CTLA4 Antibody) causes regression of metastatic renal cell cancer associated with enteritis and hypophysitis. J Immunother 2007; 30(8): 825-830

[30] A Phase 2 Open-label Study of Single Agent Ipilimumab for the Treatment of IL-2 Refractory or IL-2 Ineligible Patients with Stage IV Renal Cancer; NCT00057889; Final Clinical Study Report 930025954 v1.0 (28-January-2008)

[31] CA209016, Data on file.

[32] Schadendorf D, et al. Pooled Analysis of Long-Term Survival Data from Phase II and Phase III Trials of Ipilimumab in Unresectable or Metastatic Melanoma. J Clin Oncol 2015; 33: 1889-94.

[33] Schadendorf D, et al. Efficacy and quality of life outcomes in patients with advanced melanoma (MEL) who discontinued treatment with nivolumab (NIVO) plus ipilimumab (IPI) due to toxicity in a phase III trial (CheckMate 067). 2016 Annual Meeting of the European Association of Dermato-Oncology; August 31-September 3, 2016; Vienna, Austria.

[34] Brahmer J, et al. Five-year follow-up from the CA209-003 study of nivolumab in previously treated advanced non-small cell lung cancer: clinical characteristics of long-term survivors. Oral presentation presented at: American Association for Cancer Research (AACR) Annual Meeting; April 1-5, 2017; Washington, DC, USA.

[35] Felip E et al., Three-year follow-up from Checkmate 017/057: Nivolumab versus docetaxel in patients with previously treated advanced non-small lung cancer (NSCLC). Poster discussion presentation at the European Society of Medical Oncology Annual Meeting. 2017 Sep 8-12; Madrid, Spain. Poster 1301PD.

[36] Herbst RS et al. KEYNOTE-010: Durable clinical benefit in patients with previously treated, PD-L1-expressing NSCLC who completed pembrolizumab. Poster presentation at the World Conference on Lung Cancer 2016 Dec 4-7; Vienna, Austria.

[37] Robert, C et al. Long-term outcomes in patients with ipilimumab-naive advanced melanoma in the phase 3 KEYNOTE-006 study who completed pembrolizumab treatment; Journal of Clinical Oncology 2017 35:15_suppl, 9504-9504.

[38] Spigel DR, et al. Randomized results of continuous vs 1-year fixed-duration in patients with advanced non-small cell lung cancer. Oral presentation at the European Society of Medical Oncology Annual Meeting. 2017 Sep 8-12; Madrid, Spain.

[39] Escudier, B, et al. CheckMate 214: Efficacy and safety of nivolumab + ipilimumab (N+I) v sunitinib (S) for treatment-naïve advanced or metastatic renal cell carcinoma (mRCC), including IMDC risk and PD-L1 expression subgroups. Oral presentation at the European Society of Medical Oncology Annual Meeting. 2017 Sep 8-12; Madrid, Spain.

[40] Wolchok JD, Hoos A, O'Day S, et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. Clin Cancer Res. 2009; 15: 7412-20.

[41] Ji RR, Chasalow SD, Wang et al. An immune-active tumor microenvironment favors clinical response to ipilimumab. Cancer Immunol Immunother. 2012; 61:1019-31

[42] Motzer R.J, Hutson T.E., Tomczak P., Michaelson M.D. et al. Sunitinib versus interferon alfa in metastatic renal-cell carcinoma N Engl J Med, 356 (2) (2007), pp. 115–124

[43] Motzer R.J, Hutson T.E., McCann L., Deen K. et al. Overall Survival in Renal-Cell Carcinoma with Pazopanib versus Sunitinib N Engl J Med, 370 (18) (2014), pp. 1769-1770

[44] Clopper, C.; Pearson, E. S. (1934). "The use of confidence or fiducial limits illustrated in the case of the binomial". *Biometrika* **26**: 404–413.

APPENDIX 1 INTERNATIONAL METASTATIC RCC DATABASE CONSORTIUM (IMDC) PROGNOSTIC CRITERIA

Adverse Prognostic Factors
Clinical
KPS < 80% Time from diagnosis to treatment < 1 year
Laboratory
Hemoglobin < LLN Corrected calcium > ULN Absolute neutrophil count > ULN Platelet count > ULN

LLN = Lower limit of normal
ULN = Upper limit of normal

Corrected calcium (mg/dL) = measured total Ca (mg/dL) + 0.8 (4.0 - serum albumin [g/dL]), where 4.0 represents the average albumin level in g/dL.

Corrected calcium (mmol/L) = measured total Ca (mmol/L) + 0.02 (40 - serum albumin [g/L]), where 40 represents the average albumin level in g/L

Risk Group Based on Number of Adverse Prognostic Factors	
Number of Adverse Prognostic Factors Present	**Risk Group**
0	Favorable
1-2	Intermediate
3-6	Poor

Reference: Heng D, Xie W, Regan M, et al. Prognostic factors for overall survival in patients with metastatic renal cell carcinoma treated with vascular endothelial growth factor-targeted agents: results from a large, multicenter study. J Clin Oncol 2009; 27(34):5794-5799.

APPENDIX 2 PERFORMANCE STATUS SCALES

STATUS	SCALES		STATUS
	KARNOFSKY	ZUBROD-ECOG-WHO	
Normal, no complaints	100	0	Normal activity
Able to carry on normal activities Minor signs or symptoms of disease	90	0	Symptoms, but fully ambulatory
Normal activity with effort	80	1	
Cares for self. Unable to carry on normal activity or to do active work	70	1	Symptomatic, but in bed < 50% of the day.
Requires occasional assistance, but able to care for most of his needs	60	2	
Requires considerable assistance and frequent medical care	50	2	Needs to be in bed > 50% of the day, but not bedridden
Disabled. Requires special care and assistance	40	3	
Severely disabled. Hospitalization indicated though death non imminent	30	3	Unable to get out of bed
Very sick. Hospitalization necessary. Active supportive treatment necessary	20	4	
Moribund	10	4	
Dead	0	5	Dead

APPENDIX 3 RECIST 1.1 GUIDELINES

1 EVALUATION OF LESIONS

At baseline, tumor lesions/lymph nodes will be categorized measurable or non-measurable as follows:

1.1 Measurable

Tumor lesions: Must be accurately measured in at least one dimension (longest diameter in the plane of measurement is to be recorded) with a minimum size of:

1. 10 mm by CT scan (CT scan slice thickness no greater than 5 mm)
2. 10 mm caliper measurement by clinical exam (lesions which cannot be accurately measured with calipers should be recorded as non-measurable)
3. 20 mm by chest x-ray

Malignant lymph nodes: To be considered pathologically enlarged and measurable, a lymph node must be ≥ 15 mm in short axis when assessed by CT scan (CT scan slice thickness recommended to be no greater than 5 mm).

Lymph nodes merit special mention since they are normal anatomical structures which may be visible by imaging even if not involved by tumor. Pathological nodes which are defined as measurable and may be identified as target lesions must meet the criterion of a short axis of ≥ 15 mm by CT scan. Only the short axis of these nodes will contribute to the baseline sum. The short axis of the node is the diameter normally used by radiologists to judge if a node is involved by solid tumor. Nodal size is normally reported as two dimensions in the plane in which the image is obtained (for CT scan this is almost always the axial plane; for MRI the plane of acquisition may be axial, saggital or coronal). The smaller of these measures is the short axis. For example, an abdominal node which is reported as being 20 mm x 30 mm has a short axis of 20 mm and qualifies as a malignant, measurable node. In this example, 20 mm should be recorded as the node measurement. All other pathological nodes (those with short axis ≥ 10 mm but < 15 mm) should be considered non-target lesions. Nodes that have a short axis < 10 mm are considered non-pathological and should not be recorded or followed.

1.2 Non-Measurable

All other lesions are considered non-measurable, including small lesions (longest diameter < 10mm or pathological lymph nodes with ≥ 10 to < 15 mm short axis) as well as truly non-measurable lesions. Lesions considered truly non-measurable include: leptomeningeal disease, ascites, pleural or pericardial effusion, inflammatory breast disease, lymphangitic involvement of skin or lung, abdominal masses/abdominal organomegaly identified by physical exam that is not measurable by reproducible imaging techniques.

2 BASELINE DOCUMENTATION OF 'TARGET' AND 'NON-TARGET' LESIONS

When more than one measurable lesion is present at baseline all lesions up to a maximum of five lesions total (and a maximum of two lesions per organ) representative of all involved organs should be identified as target lesions and will be recorded and measured at baseline (this means in instances where patients have only one or two organ sites involved a maximum of two and four lesions respectively will be recorded).

Target lesions should be selected on the basis of their size (lesions with the longest diameter), be representative of all involved organs, but in addition should be those that lend themselves to reproducible repeated measurements. It may be the case that, on occasion, the largest lesion does not lend itself to reproducible measurement in which circumstance the next largest lesion which can be measured reproducibly should be selected.

A sum of the diameters (longest for non-nodal lesions, short axis for nodal lesions) for all target lesions will be calculated and reported as the baseline sum diameters. If lymph nodes are to be included in the sum, then as noted above, only the short axis is added into the sum. The baseline sum diameters will be used as reference to further characterize any objective tumor regression in the measurable dimension of the disease.

All other lesions (or sites of disease) including pathological lymph nodes should be identified as non-target lesions and should also be recorded at baseline. Measurements are not required and these lesions should be followed as 'present', 'absent', or in rare cases 'unequivocal progression' (more details to follow). In addition, it is possible to record multiple nontarget lesions involving the same organ as a single item on the case record form (eg, 'multiple enlarged pelvic lymph nodes' or 'multiple liver metastases').

3 RESPONSE CRITERIA

3.1 Evaluation of Target Lesions

Complete Response (CR): Disappearance of all target lesions. Any pathological lymph nodes (whether target or non-target) must have reduction in short axis to < 10 mm.

Partial Response (PR): At least a 30% decrease in the sum of diameters of target lesions, taking as reference the baseline sum diameters.

Progressive Disease (PD): At least a 20% increase in the sum of diameters of target lesions, taking as reference the smallest sum on study (this includes the baseline sum if that is the smallest on study). In addition to the relative increase of 20%, the sum must also demonstrate an absolute increase of at least 5 mm. (Note: the appearance of one or more new lesions is also considered progression).

Stable Disease (SD): Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD, taking as reference the smallest sum diameters while on study.

3.1.1 Special Notes on the Assessment of Target Lesions

3.1.1.1 Lymph nodes

Lymph nodes identified as target lesions should always have the actual short axis measurement recorded (measured in the same anatomical plane as the baseline examination), even if the nodes regress to below 10 mm on study. This means that when lymph nodes are included as target lesions, the 'sum' of lesions may not be zero even if complete response criteria are met, since a normal lymph node is defined as having a short axis of < 10 mm. Case report forms or other data collection methods may therefore be designed to have target nodal lesions recorded in a separate section where, in order to qualify for CR, each node must achieve a short axis < 10 mm. For PR, SD and PD, the actual short axis measurement of the nodes is to be included in the sum of target lesions.

3.1.1.2 Target lesions that become 'too small to measure'

While on study, all lesions (nodal and non-nodal) recorded at baseline should have their actual measurements recorded at each subsequent evaluation, even when very small (eg, 2 mm). However, sometimes lesions or lymph nodes which are recorded as target lesions at baseline become so faint on CT scan that the radiologist may not feel comfortable assigning an exact measure and may report them as being 'too small to measure'. When this occurs it is important that a value be recorded on the case report form. If it is the opinion of the radiologist that the lesion has likely disappeared, the measurement should be recorded as 0 mm. If the lesion is believed to be present and is faintly seen but too small to measure, a default value of 5 mm should be assigned (Note: It is less likely that this rule will be used for lymph nodes since they usually have a definable size when normal and are frequently surrounded by fat such as in the retroperitoneum; however, if a lymph node is believed to be present and is faintly seen but too small to measure, a default value of 5 mm should be assigned in this circumstance as well). This default value is derived from the 5 mm CT slice thickness (but should not be changed with varying CT slice thickness). The measurement of these lesions is potentially non-reproducible, therefore providing this default value will prevent false responses or progressions based upon measurement error. To reiterate, however, if the radiologist is able to provide an actual measure, that should be recorded, even if it is below 5 mm.

3.1.1.3 Lesions that split or coalesce on treatment

When non-nodal lesions 'fragment', the longest diameters of the fragmented portions should be added together to calculate the target lesion sum. Similarly, as lesions coalesce, a plane between them may be maintained that would aid in obtaining maximal diameter measurements of each individual lesion. If the lesions have truly coalesced such that they are no longer separable, the vector of the longest diameter in this instance should be the maximal longest diameter for the 'coalesced lesion'.

3.2 Evaluation of Non-Target Lesions

This section provides the definitions of the criteria used to determine the tumor response for the group of non-target lesions. While some non-target lesions may actually be measurable, they

need not be measured and instead should be assessed only qualitatively at the time points specified in the protocol.

Complete Response (CR): Disappearance of all non-target lesions and normalization of tumor marker level. All lymph nodes must be non-pathological in size (< 10mm short axis).

Non-CR/Non-PD: Persistence of one or more non-target lesion(s) and/or maintenance of tumor marker level above the normal limits.

Progressive Disease (PD): Unequivocal progression (see comments below) of existing non-target lesions. (Note: the appearance of one or more new lesions is also considered progression).

3.2.1 Special Notes on Assessment of Progression of Non-Target Disease

The concept of progression of non-target disease requires additional explanation as follows:

3.2.1.1 When the patient also has measurable disease

In this setting, to achieve 'unequivocal progression' on the basis of the non-target disease, there must be an overall level of substantial worsening in non-target disease such that, even in presence of SD or PR in target disease, the overall tumor burden has increased sufficiently to merit discontinuation of therapy (see examples in Appendix 2 and further details below). A modest 'increase' in the size of one or more non-target lesions is usually not sufficient to qualify for unequivocal progression status. The designation of overall progression solely on the basis of change in non-target disease in the face of SD or PR of target disease will therefore be extremely rare.

3.2.1.2 When the patient has only non-measurable disease

This circumstance arises in some trials when it is not a criterion of study entry to have measurable disease. The same general concepts apply here as noted above, however, in this instance there is no measurable disease assessment to factor into the interpretation of an increase in non-measurable disease burden. Because worsening in non-target disease cannot be easily quantified (by definition: if all lesions are truly non-measurable) a useful test that can be applied when assessing patients for unequivocal progression is to consider if the increase in overall disease burden based on the change in non-measurable disease is comparable in magnitude to the increase that would be required to declare PD for measurable disease: ie, an increase in tumor burden representing an additional 73% increase in 'volume' (which is equivalent to a 20% increase diameter in a measurable lesion). Examples include an increase in a pleural effusion from 'trace' to 'large', an increase in lymphangitic disease from localized to widespread, or may be described in protocols as 'sufficient to require a change in therapy'. If 'unequivocal progression' is seen, the patient should be considered to have had overall PD at that point. While it would be ideal to have objective criteria to apply to non-measurable disease, the very nature of that disease makes it impossible to do so; therefore the increase must be substantial.

Motzer RJ, *et al. J Immunother Cancer* 2021; 8:e000891. doi: 10.1136/jitc-2020-000891

3.2.2 New Lesions

The appearance of new malignant lesions denotes disease progression; therefore, some comments on detection of new lesions are important. There are no specific criteria for the identification of new radiographic lesions; however, the finding of a new lesion should be unequivocal: ie, not attributable to differences in scanning technique, change in imaging modality or findings thought to represent something other than tumor (for example, some 'new' bone lesions may be simply healing or flare of pre-existing lesions). This is particularly important when the patient's baseline lesions show partial or complete response. For example, necrosis of a liver lesion may be reported on a CT scan report as a 'new' cystic lesion, which it is not.

A lesion identified on a follow-up study in an anatomical location that was not scanned at baseline is considered a new lesion and will indicate disease progression. An example of this is the patient who has visceral disease at baseline and while on study has a CT or MRI brain ordered which reveals metastases. The patient's brain metastases are considered to be evidence of PD even if he/she did not have brain imaging at baseline.

If a new lesion is equivocal, for example because of its small size, continued therapy and follow-up evaluation will clarify if it represents truly new disease. If repeat scans confirm there is definitely a new lesion, then progression should be declared using the date of the initial scan. While FDG-PET response assessments need additional study, it is sometimes reasonable to incorporate the use of FDG-PET scanning to complement CT scanning in assessment of progression (particularly possible 'new' disease). New lesions on the basis of FDG-PET imaging can be identified according to the following algorithm:

1. Negative FDG-PET at baseline, with a positive FDG-PET at follow-up is a sign of PD based on a new lesion.
2. No FDG-PET at baseline and a positive FDG-PET at follow-up: If the positive FDG-PET at follow-up corresponds to a new site of disease confirmed by CT, this is PD. If the positive FDG-PET at follow-up is not confirmed as a new site of disease on CT, additional follow-up CT scans are needed to determine if there is truly progression occurring at that site (if so, the date of PD will be the date of the initial abnormal FDG-PET scan). If the positive FDG-PET at follow-up corresponds to a pre-existing site of disease on CT that is not progressing on the basis of the anatomic images, this is not PD.

3.3 Response Assessment

3.3.1 Evaluation of Best Overall Response

The best overall response is the best response recorded from the start of the study treatment until the end of treatment taking into account any requirement for confirmation. The patient's best overall response assignment will depend on the findings of both target and non-target disease and will also take into consideration the appearance of new lesions. Furthermore, depending on the nature of the study and the protocol requirements, it may also require confirmatory measurement.

Motzer RJ, *et al. J Immunother Cancer* 2021; 8:e000891. doi: 10.1136/jitc-2020-000891

3.3.2 Time Point Response

It is assumed that at each protocol specified time point, a response assessment occurs. Table 3.3.2-1 provides a summary of the overall response status calculation at each time point for patients who have measurable disease at baseline. When patients have non-measurable (therefore non-target) disease only, Table 3.3.2-2 is to be used.

Table 3.3.2-1: Time Point Response: Patients With Target (+/- Non-Target) Disease

Target Lesions	Non-Target Lesions	New Lesions	Overall Response
CR	CR	No	CR
CR	Non-CR/non-PD	No	PR
CR	Not evaluated	No	PR
PR	Non-PD or not all evaluated	No	PR
SD	Non-PD or not all evaluated	No	SD
Not all evaluated	Non-PD	No	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD

CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease and NE = inevaluable

Table 3.3.2-2: Time Point Response: Patients with Non-target Disease Only

Non-Target Lesions	New Lesions	Overall Response
CR	No	CR
Non-CR/non-PD	No	Non-CR/non-PD[a]
Not all evaluated	No	NE
Unequivocal PD	Yes or No	PD
Any	Yes	PD

CR = complete response, PD = progressive disease and NE = inevaluable

[a] Non-CR/non-PD is preferred over SD for non-target disease since SD is increasingly used as endpoint for assessment of efficacy in some trials so to assign this category when no lesions can be measured is not advised.

3.3.3 Best Overall Response

Best response determination of complete or partial response requires confirmation: Complete or partial responses may be claimed only if the criteria for each are met at a subsequent time point

of ≥ 4 weeks later. In this circumstance, the best overall response can be interpreted as in Table 3.3.3-1.

Special note on response assessment: When nodal disease is included in the sum of target lesions and the nodes decrease to 'normal' size (< 10 mm), they may still have a measurement reported on scans. This measurement should be recorded even though the nodes are normal in order not to overstate progression should it be based on increase in size of the nodes. As noted earlier, this means that patients with CR may not have a total sum of 'zero' on the case report form (CRF).

Table 3.3.3-1: Best Overall Response (Confirmation of CR&PR Required)

Overall Response First Time Point	Overall Response Subsequent Time Point	BEST Overall Response
CR	CR	CR
CR	PR	SD, PD OR PR[a]
CR	SD	SD provided minimum criteria for SD duration[b] met, otherwise, PD
CR	PD	SD provided minimum criteria for SD duration[b] met, otherwise, PD
CR	NE	SD provided minimum criteria for SD duration[b] met, otherwise, NE
PR	CR	PR
PR	PR	PR
PR	SD	SD
PR	PD	SD provided minimum criteria for SD duration[b] met, otherwise, PD
PR	NE	SD provided minimum criteria for SD duration [b]met, otherwise, NE
NE	NE	NE

CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease, and NE = inevaluable

[a] If a CR is truly met at first time point, then any disease seen at a subsequent time point, even disease meeting PR criteria relative to baseline, makes the disease PD at that point (since disease must have reappeared after CR). Best response would depend on whether minimum duration for SD was met. However, sometimes 'CR' may be claimed when subsequent scans suggest small lesions were likely still present and in fact the patient had PR, not CR at the first time point. Under these circumstances, the original CR should be changed to PR and the best response is PR.

[b] Minimum criteria for SD duration is 6 weeks.

3.3.4 Confirmation Scans

Verification of Response: To be assigned a status of CR or PR, changes in tumor measurements must be confirmed by consecutive repeat assessments that should be performed no less than 28 days after the criteria for response are first met. For this study, the next scheduled tumor assessment can meet this requirement.

Verification of Progression: Progression of disease should be verified in cases where progression is equivocal. If repeat scans confirm PD, then progression should be declared using the date of the initial scan. If repeat scans do not confirm PD, then the subject is considered to not have progressive disease.

APPENDIX 4 INDUCERS AND STRONG INHIBITORS OF CYP3A4

Inducers and Strong Inhibitors of CYP3A4
CYP3A4 Inducers
Phenytoin
Carbamazepine
Rifampin
Rifabutin
Rifapentine
Phenobarbital
Dexamethasone
Strong CYP3A4 Inhibitors
Ketoconazole
Itraconazole
Voriconazole
Clarithromycin
Erythromycin
Telithromycin
Nefazodone
Saquinavir
Ritonavir
Atazanavir
Indinavir
Nelfinavir

Notes:

The above list is not exhaustive.

Grapefruit, grapefruit juice and other foods that are known to inhibit CYP3A4 activity should be avoided during treatment.

St. John's Wort (*Hypericum perforatum*) is known to be an inducer of CYP3A4 and should be avoided during treatment.

APPENDIX 5 MANAGEMENT ALGORITHMS FOR IMMUNO-ONCOLOGY AGENTS

These general guidelines constitute guidance to the Investigator and may be supplemented by discussions with the Medical Monitor representing the Sponsor. The guidance applies to all immuno-oncology agents and regimens.

A general principle is that differential diagnoses should be diligently evaluated according to standard medical practice. Non-inflammatory etiologies should be considered and appropriately treated.

Corticosteroids are a primary therapy for immuno-oncology drug-related adverse events. The oral equivalent of the recommended IV doses may be considered for ambulatory patients with low-grade toxicity. The lower bioavailability of oral corticosteroids should be taken into account when switching to the equivalent dose of oral corticosteroids.

Consultation with a medical or surgical specialist, especially prior to an invasive diagnostic or therapeutic procedure, is recommended.

The frequency and severity of the related adverse events covered by these algorithms will depend on the immuno-oncology agent or regimen being used.

GI Adverse Event Management Algorithm

Rule out non-inflammatory causes. If non-inflammatory cause is identified, treat accordingly and continue I-O therapy. Opiates/narcotics may mask symptoms of perforation. Infliximab should not be used in cases of perforation or sepsis.

Grade of Diarrhea/ Colitis (NCI CTCAE v4)	Management	Follow-up
Grade 1 Diarrhea:< 4 stools/day over baseline; Colitis: asymptomatic	• Continue I-O therapy per protocol • Symptomatic treatment	• Close monitoring for worsening symptoms. • Educate patient to report worsening immediately If worsens: • Treat as Grade 2 or 3/4
Grade 2 Diarrhea: 4-6 stools per day over baseline; IV fluids indicated <24 hrs; not interfering with ADL Colitis: abdominal pain; blood in stool	• Delay I-O therapy per protocol • Symptomatic treatment	If improves to grade 1: • Resume I-O therapy per protocol If persists > 5-7 days or recur: • 0.5-1.0 mg/kg/day methylprednisolone or oral equivalent • When symptoms improve to grade 1, taper steroids over at least 1 month, consider prophylactic antibiotics for opportunistic infections, and resume I-O therapy per protocol. If worsens or persists > 3-5 days with oral steroids: • Treat as grade 3/4
Grade 3-4 Diarrhea (G3): ≥7 stools per day over baseline; incontinence; IV fluids ≥24 hrs; interfering with ADL Colitis (G3): severe abdominal pain, medical intervention indicated, peritoneal signs G4: life-threatening, perforation	• Discontinue I-O therapy per protocol • 1.0 to 2.0 mg/kg/day methylprednisolone IV or IV equivalent • Add prophylactic antibiotics for opportunistic infections • Consider lower endoscopy	If improves: • Continue steroids until grade 1, then taper over at least 1 month If persists > 3-5 days, or recurs after improvement: • Add infliximab 5 mg/kg (if no contraindication). Note: Infliximab should not be used in cases of perforation or sepsis

Patients on IV steroids may be switched to an equivalent dose of oral corticosteroids (e.g. prednisone) at start of tapering or earlier, once sustained clinical improvement is observed. Lower bioavailability of oral corticosteroids should be taken into account when switching to the equivalent dose of oral corticosteroids.

Updated 05-Jul-2016

Renal Adverse Event Management Algorithm

Rule out non-inflammatory causes. If non-inflammatory cause, treat accordingly and continue I-O therapy

Grade of Creatinine Elevation (NCI CTCAE v4)	Management	Follow-up
Grade 1 Creatinine > ULN and > than baseline but ≤ 1.5x baseline	• Continue I-O therapy per protocol • Monitor creatinine weekly	**If returns to baseline:** •Resume routine creatinine monitoring per protocol **If worsens:** •Treat as Grade 2 or 3/4
Grade 2-3 Creatinine > 1.5x baseline to ≤ 6x ULN	• Delay I-O therapy per protocol • Monitor creatinine every 2-3 days • 0.5 to 1.0 mg/kg/day methylprednisolone IV or oral equivalent • Consider renal biopsy with nephrology consult	**If returns to Grade 1:** •Taper steroids over at least 1 month, consider prophylactic antibiotics for opportunistic infections, and resume I-O therapy and routine creatinine monitoring per protocol **If elevations persist > 7 days or worsen:** •Treat as Grade 4
Grade 4 Creatinine > 6x ULN	• Discontinue I-O therapy per protocol • Monitor creatinine daily • 1.0-2.0 mg/kg/day methylprednisolone IV or IV equivalent • Consult nephrologist • Consider renal biopsy	**If returns to Grade 1 :** Taper steroids over at least 1 month and add prophylactic antibiotics for opportunistic infections

Patients on IV steroids may be switched to an equivalent dose of oral corticosteroids (e.g. prednisone) at start of tapering or earlier, once sustained clinical improvement is observed. Lower bioavailability of oral corticosteroids should be taken into account when switching to the equivalent dose of oral corticosteroids.

Updated 05-Jul-2016

Pulmonary Adverse Event Management Algorithm

Rule out non-inflammatory causes. If non-inflammatory cause, treat accordingly and continue I-O therapy. Evaluate with imaging and pulmonary consultation.

Grade of Pneumonitis (NCI CTCAE v4)	Management	Follow-up
Grade 1 Radiographic changes only	• Consider delay of I-O therapy • Monitor for symptoms every 2-3 days • Consider Pulmonary and ID consults	• Re-image at least every 3 weeks If worsens: • Treat as Grade 2 or 3-4
Grade 2 Mild to moderate new symptoms	• Delay I-O therapy per protocol • Pulmonary and ID consults • Monitor symptoms daily, consider hospitalization • 1.0 mg/kg/day methyl-prednisolone IV or oral equivalent • Consider bronchoscopy , lung biopsy	• Re-image every 1-3 days If improves: • When symptoms return to near baseline, taper steroids over at least 1 month and then resume I-O therapy per protocol and consider prophylactic antibiotics If not improving after 2 weeks or worsening: • Treat as Grade 3-4
Grade 3-4 Severe new symptoms; New/worsening hypoxia; Life-threatening	• Discontinue I-O therapy per protocol • Hospitalize • Pulmonary and ID consults • 2-4 mg/kg/day methylprednisolone IV or IV equivalent • Add prophylactic antibiotics for opportunistic infections • Consider bronchoscopy, lung biopsy	If improves to baseline: • Taper steroids over at least 6 weeks If not improving after 48 hours or worsening: • Add additional immunosuppression

Patients on IV steroids may be switched to an equivalent dose of oral corticosteroids (e.g. prednisone) at start of tapering or earlier, once sustained clinical improvement is observed. Lower bioavailability of oral corticosteroids should be taken into account when switching to the equivalent dose of oral corticosteroids.

Updated 05-Jul-2016

Hepatic Adverse Event Management Algorithm

Rule out non-inflammatory causes. If non-inflammatory cause, treat accordingly and continue I-O therapy. Consider imaging for obstruction.

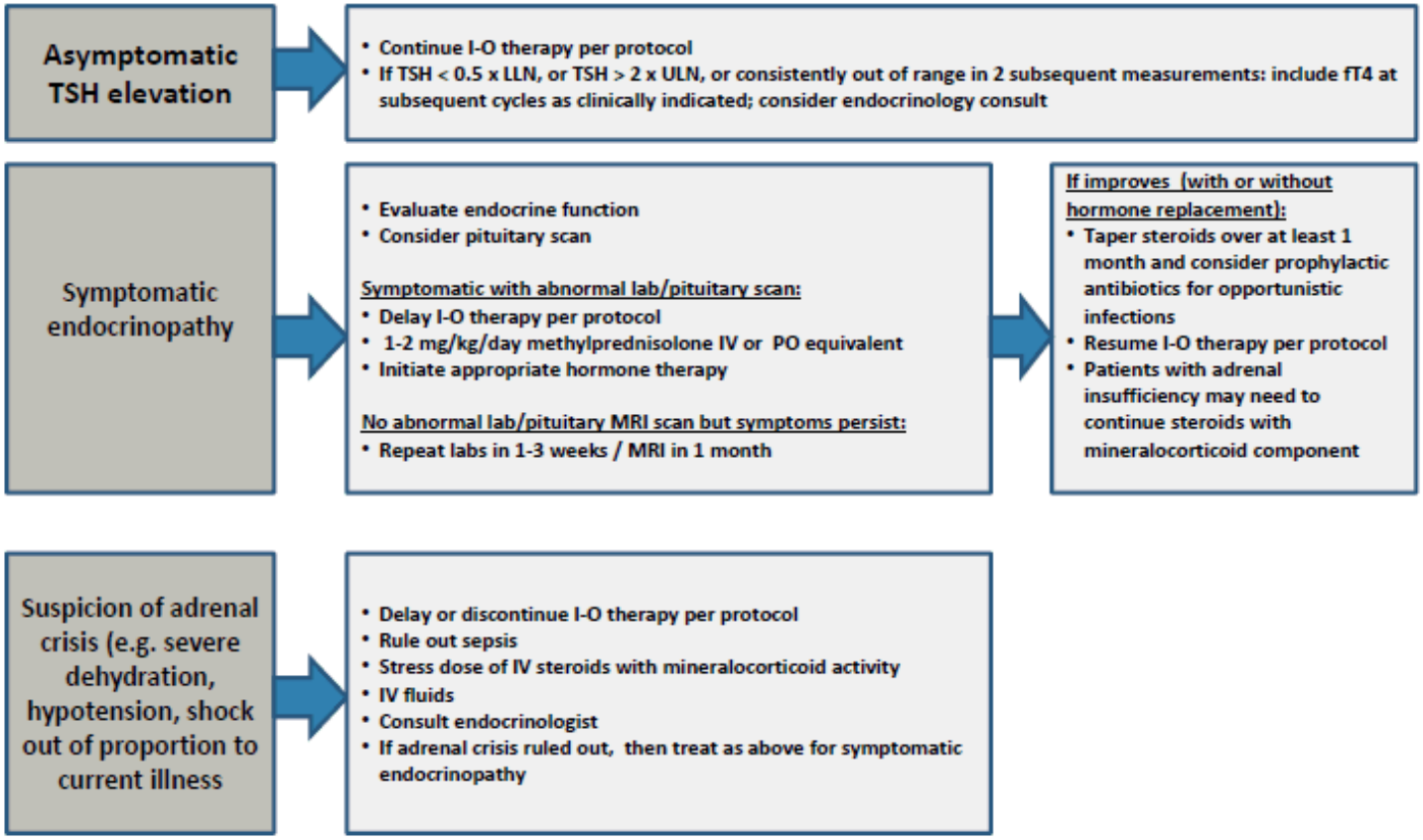
Grade of Liver Test Elevation (NCI CTCAE v4)	Management	Follow-up
Grade 1 AST or ALT > ULN to 3.0 x ULN and/or T. bili > ULN to 1.5 x ULN	• Continue I-O therapy per protocol	• Continue LFT monitoring per protocol If worsens: • Treat as Grade 2 or 3-4
Grade 2 AST or ALT > 3.0 to ≤ 5 x ULN and/or T. bili > 1.5 to ≤ 3 x ULN	• Delay I-O therapy per protocol • Increase frequency of monitoring to every 3 days	If returns to baseline: •Resume routine monitoring, resume I-O therapy per protocol If elevations persist > 5-7 days or worsen : • 0.5-1 mg/kg/day methylprednisolone or oral equivalent and when LFT returns to grade 1 or baseline, taper steroids over at least 1 month, consider prophylactic antibiotics for opportunistic infections, and resume I-O therapy per protocol
Grade 3-4 AST or ALT > 5 x ULN or T.bili >3 x ULN	• Discontinue I-O therapy* • Increase frequency of monitoring to every 1-2 days • 1.0 to 2.0 mg/kg/day methylprednisolone IV or IV equivalent** • Add prophylactic antibiotics for opportunistic infections • Consult gastroenterologist	If returns to grade 2: • Taper steroids over at least 1 month If does not improve in >3-5 days, worsens or rebounds: • Add mycophenolate mofetil 1 g BID • If no response within an additional 3-5 days, consider other immunosuppressants per local guidelines

Patients on IV steroids may be switched to an equivalent dose of oral corticosteroids (e.g. prednisone) at start of tapering or earlier, once sustained clinical improvement is observed. Lower bioavailability of oral corticosteroids should be taken into account when switching to the equivalent dose of oral corticosteroids.
*I-O therapy may be delayed rather than discontinued if AST/ALT ≤ 8 x ULN or T.bili ≤ 5 x ULN.
**The recommended starting dose for grade 4 hepatitis is 2 mg/kg/day methylprednisolone IV.

Updated 05-Jul-2016

Endocrinopathy Management Algorithm

Rule out non-inflammatory causes. If non-inflammatory cause, treat accordingly and continue I-O therapy. Consider visual field testing, endocrinology consultation, and imaging.

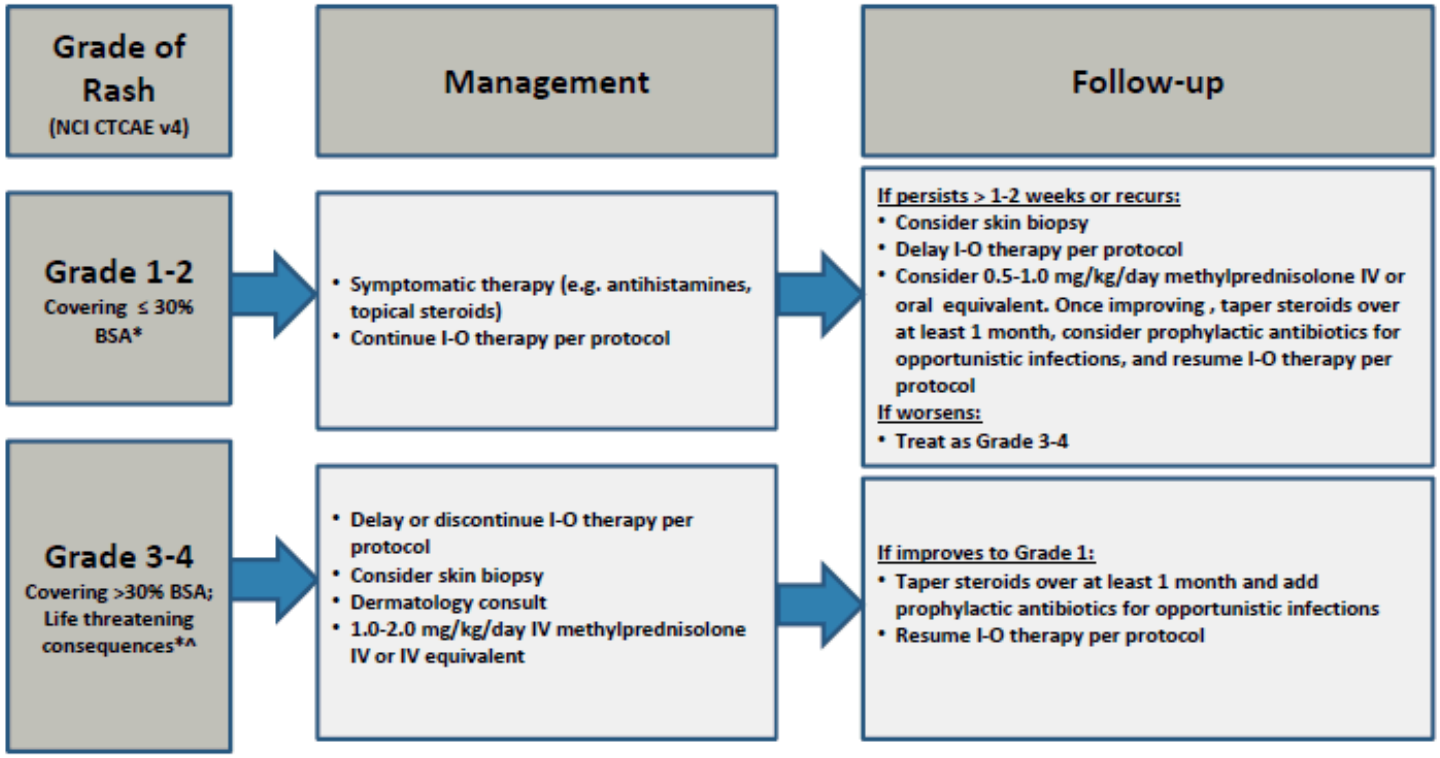


Patients on IV steroids may be switched to an equivalent dose of oral corticosteroids (e.g. prednisone) at start of tapering or earlier, once sustained clinical improvement is observed. Lower bioavailability of oral corticosteroids should be taken into account when switching to the equivalent dose of oral corticosteroids.

Updated 05-Jul-2016

Skin Adverse Event Management Algorithm

Rule out non-inflammatory causes. If non-inflammatory cause, treat accordingly and continue I-O therapy.

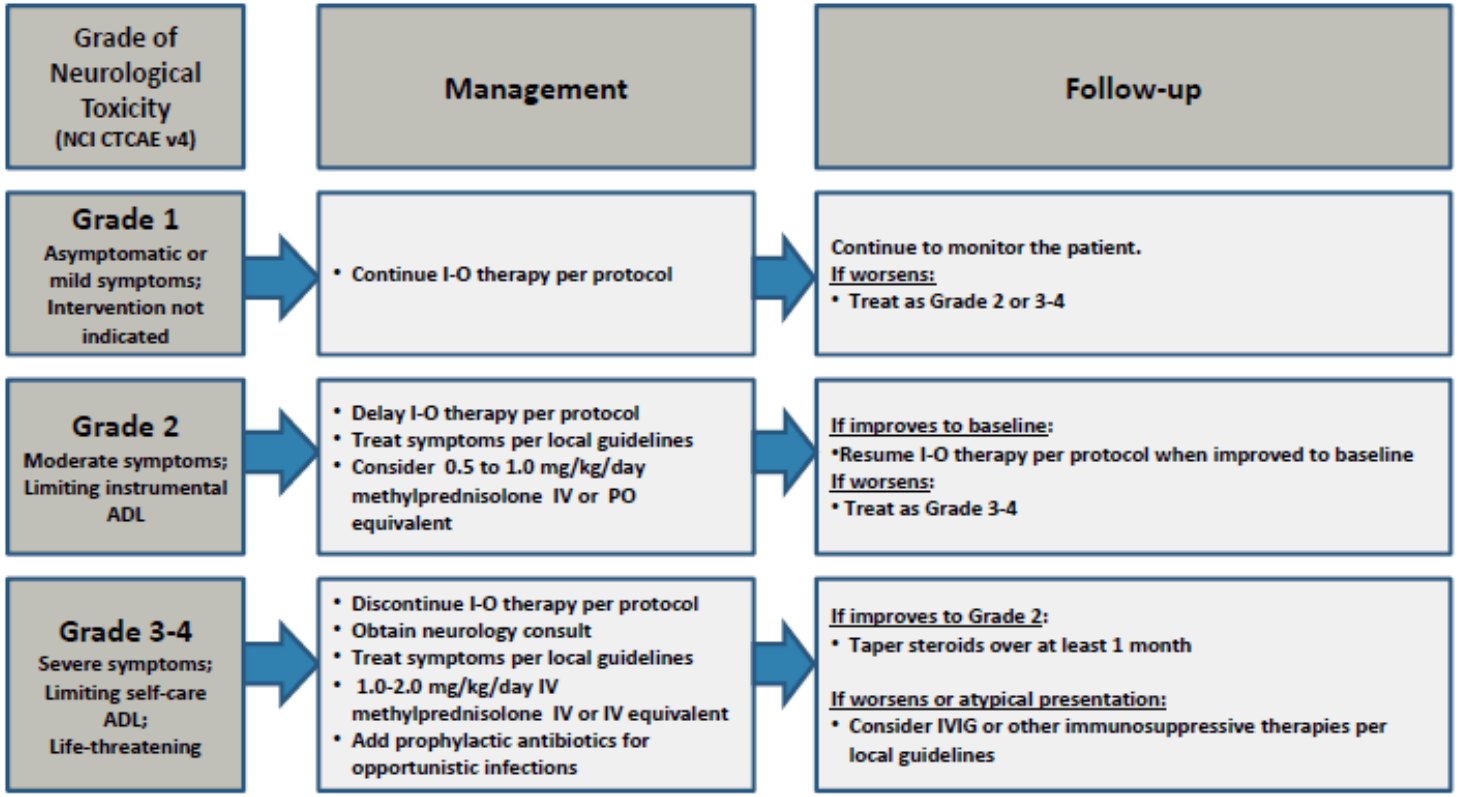


Patients on IV steroids may be switched to an equivalent dose of oral corticosteroids (e.g. prednisone) at start of tapering or earlier, once sustained clinical improvement is observed. Lower bioavailability of oral corticosteroids should be taken into account when switching to the equivalent dose of oral corticosteroids.
*Refer to NCI CTCAE v4 for term-specific grading criteria.
^If SJS/TEN is suspected, withhold I-O therapy and refer patient for specialized care for assessment and treatment. If SJS or TEN is diagnosed, permanently discontinue I-O therapy.

Updated 05-Jul-2016

Neurological Adverse Event Management Algorithm

Rule out non-inflammatory causes. If non-inflammatory cause, treat accordingly and continue I-O therapy.



Patients on IV steroids may be switched to an equivalent dose of oral corticosteroids (e.g. prednisone) at start of tapering or earlier, once sustained clinical improvement is observed. Lower bioavailability of oral corticosteroids should be taken into account when switching to the equivalent dose of oral corticosteroids.

Updated 05-Jul-2016

STATISTICAL ANALYSIS PLAN

FOR CLINICAL STUDY REPORT

A PHASE 3, RANDOMIZED, OPEN-LABEL STUDY OF NIVOLUMAB COMBINED WITH IPILIMUMAB VERSUS SUNITINIB MONOTHERAPY IN SUBJECTS WITH PREVIOUSLY UNTREATED, ADVANCED OR METASTATIC RENAL CELL CARCINOMA

PROTOCOL(S) CA209214

VERSION # 4.0

REVISION HISTORY

Revision	Date	Revised By	Changes Made -- Reasons for the Change
1.0	9/16/2014	Prabhu Bhagavatheeswaran	Original issue
2.0	2/4/2016	Brent McHenry	Added weighted log-rank (Section 7.5.2) and PD-L1 analyses (Section 7.7). General updates for typos, inconsistencies, etc.
3.0	8/17/2016	Brent McHenry	Incorporate protocol amendment (revprot02) to include ORR as co-primary endpoint.
3.1	3/09/2017	Brent McHenry	Update PFS primary/secondary definition to be consistent with updated program standards.
4.0	6/14/2017	Brent McHenry	Due to slowing PFS events rate per IRRC, the timing of the final PFS lock was advanced.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 BACKGROUND AND RATIONALE

CA209214 (CheckMate 214, CHECKpoint pathway and nivoluMAb clinical Trial Evaluation) is a Phase 3, randomized, open-label study of nivolumab (BMS-936558) combined with ipilimumab vs sunitinib monotherapy in subjects with previously untreated, advanced or metastatic renal cell carcinoma (RCC). In the Phase 1 setting, nivolumab combined with ipilimumab has demonstrated substantially improved activity compared to either agent alone. Given the durability of responses associated with immunotherapies, nivolumab combined with ipilimumab is hypothesized to lead to greater clinical benefit, as measured by objective response rate (ORR), progression-free survival (PFS) or overall survival (OS), than sunitinib, a widely used standard-of-care agent in this patient population. This study will allow for direct comparison of PFS and OS between arms. No approved drug has demonstrated an improvement in ORR, PFS or OS vs sunitinib in the Phase 3 setting. If nivolumab combined with ipilimumab has an acceptable safety profile and is shown to improve ORR, PFS, OS, or both vs sunitinib, this study may support the approval of nivolumab combined with ipilimumab in subjects with previously untreated, advanced or metastatic RCC.

Research Hypothesis:

Treatment with nivolumab combined with ipilimumab will improve ORR, PFS or OS compared to sunitinib monotherapy in subjects with previously untreated, advanced or metastatic RCC.

Schedule of Analyses:

This study will be monitored by an independent Data Monitoring Committee (DMC).Details are specified in the DMC charter[1]

Formal analysis of ORR (as per IRRC) in the intermediate/poor-risk randomized subjects will be performed when these patients have at least 6 months of minimum follow-up from the completion of enrollment approximately 27 months from FPFV. At the time of the formal ORR analysis, no analysis of PFS or OS will be conducted.

Formal final analysis of PFS (as per IRRC) is scheduled after observing at least 591 PFS events among the randomized intermediate/poor risk subjects in the two respective treatment arms. Under the assumptions stated above on accrual and PFS distribution, it will approximately take 31 months from FPFV to observe the required number of PFS events.

Two interim analyses of OS are planned. The first interim analysis is planned at the time of final PFS analysis and it is expected to observe 370 events (58% of the targeted OS events for final analysis) among the randomized intermediate/poor risk subjects, which is expected at around 31 months from FPFV. The second interim analysis after observing 479 events (75% of targeted OS events needed for final analysis) among the randomized intermediate/poor risk subjects, which is expected at around 40 months from FPFV. It will approximately take 61 months from FPFV to observe the required number of 639 OS events for the final OS analysis.

Secondary endpoints will be analyzed at the time of the final analysis of co-primary endpoints based on a hierarchical testing strategy. In the event that the interim analysis for superiority of overall survival is positive, final (CSR) analyses will be performed prior to achieving 639 deaths; additional details can be found in section 7.5.13.

2 STUDY DESCRIPTION

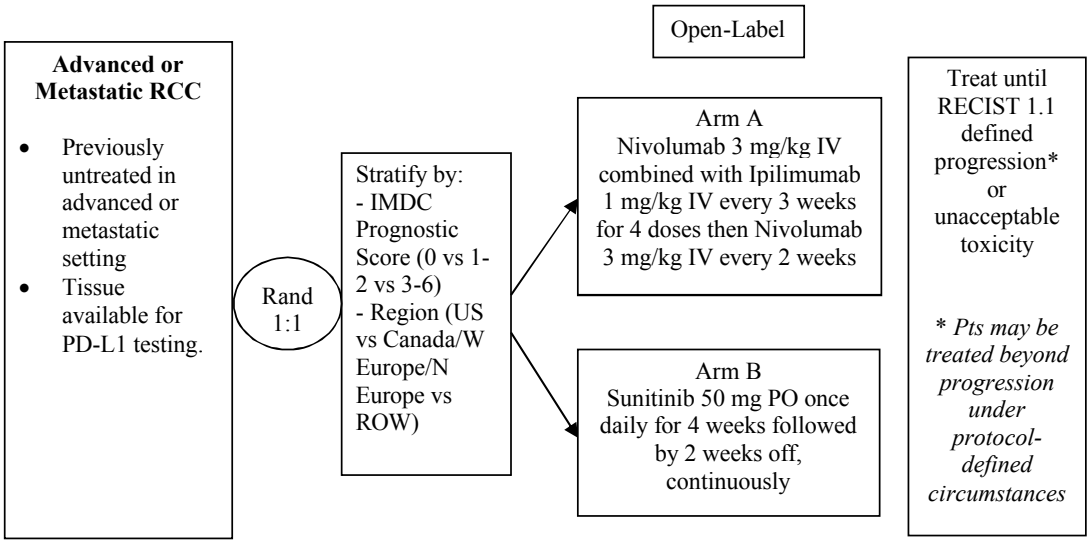
2.1 Study Design

This is a Phase 3, randomized, open-label study of nivolumab 3 mg/kg combined with ipilimumab 1 mg/kg every 3 weeks for 4 doses followed by nivolumab 3 mg/kg every 2 weeks vs sunitinib monotherapy using the approved dose and schedule (50 mg po once daily for 4 weeks followed by 2 weeks off, continuously) in adult (≥ 18 years) subjects with previously untreated advanced or metastatic RCC. The study is expected to randomize approximately 820 subjects with intermediate or poor prognosis and up to approximately 250 subjects with favorable prognosis as per IMDC criteria. Tumor tissue from an unresectable or metastatic site of disease must be received by the central vendor in order to be randomized. Subjects must have advanced (not amenable to curative surgery or radiation) or metastatic (AJCC Stage IV) RCC, and must not have received prior systemic therapy for the treatment of advanced or metastatic RCC. Prior adjuvant or neoadjuvant therapy is allowed if such therapy did not include an agent that targets VEGF or VEGF receptors and was completed at least 6 weeks prior to randomization. Subjects will be randomized 1:1 and stratified by IMDC prognostic score (0 vs 1-2 vs 3-6) and region (US vs Canada/W Europe/N Europe vs Rest of World). Subjects will be randomized to Arm A (nivolumab 3 mg/kg IV combined with ipilimumab 1 mg/kg IV every 3 weeks for 4 doses, followed by nivolumab 3 mg/kg IV every 2 weeks) or Arm B (sunitinib using the approved dose and schedule of 50 mg PO once daily for 4 weeks followed by two weeks off, continuously).

Subjects will be assessed for response (RECIST 1.1) by CT or MRI beginning 12 weeks (± 1 week) from randomization and continuing every 6 weeks (± 1 week) for the first 13 months and then every 12 weeks until progression or treatment discontinuation, whichever occurs later. Subjects will be allowed to continue study therapy after initial investigator-assessed RECIST 1.1-defined progression if assessed by the investigator to be deriving clinical benefit and tolerating study drug. Such subjects should discontinue study therapy when further progression is documented (see Section 4.3.6). The co-primary endpoints of this study are ORR and PFS in intermediate and poor-risk subjects, as assessed by an independent review committee (IRRC) and OS in intermediate and poor-risk subjects. ORR will be analyzed initially on a descriptive basis and will occupy an administrative adjustment of alpha of 0.001.

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

Figure 2.1-1: Study Design Schematic



2.2 Treatment Assignment

CA209214 is a randomized, open label study. After the subject's initial eligibility is established and informed consent has been obtained, the subject must be enrolled into the study by calling an interactive voice response system (IVRS) to obtain the subject number. 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
- KPS less than 80 (ie, equal to 70)? Yes/No
- Less than 1 year from diagnosis to randomization? Yes/No
- Hemoglobin less than the LLN? Yes/No
- Corrected calcium greater than the ULN? Yes/No
- Absolute neutrophil count greater than the ULN? Yes/No
- Platelet count greater than the ULN? Yes/No

Subjects meeting all eligibility criteria will be randomized in a 1:1 ratio to Arm A (nivolumab combined with ipilimumab) or Arm B (sunitinib), stratified by the following factors:

- IMDC Prognostic Score (Total Number of IMDC Adverse Prognostic Factors Present)
  - 0
  - 1-2
  - 3-6

- Region
  - US
  - Canada/W Europe/N Europe
  - ROW

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

2.3 Blinding and Unblinding

This is an open label study.

2.4 Protocol Amendments

Not applicable.

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. 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 conduct of this trial to ensure that any safety issues are identified and addressed.

The DMC will conduct the first review of the safety data after at least 20 subjects are treated and followed for at least 1 month. The DMC will conduct its second review of the safety data after at least 50 subjects are treated and followed for at least 1 month. The DMC will conduct its third review of the safety data focusing on the initial approximately 12 Japanese subjects treated and followed for at least 1 month. The DMC will then review safety and the available efficacy data pertaining to co-primary endpoints to evaluate safety in the context of benefit, every six months thereafter.

The DMC will also review the formal analysis of ORR (per IRRC) scheduled at around 27 months (when all patients have at least 6 months of follow-up) from FPFV. Details of the formal ORR analyses can be found in section 7.5.13.

The DMC will also review the formal final analysis of PFS (as per IRRC) and first interim analysis of superiority of OS scheduled at around 31 months (approximately 591 PFS events and 370 OS events) from FPFV. A second interim analysis of overall survival will be at around 40 months (approximately 479 OS events) from FPFV. Details of the interim analyses can be found in section 7.5.13.

2.6 Independent Radiological Review Committee

An independent Radiological Review Committee (IRRC) has been established to provide an independent imaging review of images obtained in subjects participating in this study. Details of IRRC responsibilities and processes may be found in the IRRC Charter. The IRRC determined PFS and ORR endpoints will be utilized as a part of primary and secondary efficacy analyses.

3 OBJECTIVES

3.1 Primary

- To describe the ORR of nivolumab combined with ipilimumab and sunitinib monotherapy in intermediate and poor-risk subjects with previously untreated mRCC, as assessed by IRRC.
- To compare the PFS of nivolumab combined with ipilimumab to sunitinib monotherapy in intermediate and poor-risk subjects with previously untreated mRCC, as assessed by IRRC.
- To compare the OS of nivolumab combined with ipilimumab to sunitinib monotherapy in intermediate and poor-risk subjects with previously untreated mRCC.

3.2 Secondary

- To compare the PFS of nivolumab combined with ipilimumab to sunitinib monotherapy in all randomized subjects with previously untreated mRCC, as assessed by IRRC.
- To compare the OS of nivolumab combined with ipilimumab to sunitinib monotherapy in all randomized subjects with previously untreated mRCC.
- To estimate the objective response rate (ORR) of nivolumab combined with ipilimumab to sunitinib monotherapy in subjects with previously untreated mRCC (any-risk), as assessed by IRRC.

3.3 Exploratory

- To assess the overall safety and tolerability of nivolumab combined with ipilimumab vs sunitinib monotherapy.
- To estimate the PFS based on IRRC assessments and OS of nivolumab combined with ipilimumab vs sunitinib monotherapy in favorable risk subjects with previously untreated mRCC.
- To characterize the pharmacokinetics (PK) of nivolumab and ipilimumab when co-administered.
- To evaluate immunogenicity of nivolumab and ipilimumab administered as combination therapy.

- [redacted]
- To evaluate health related quality of life (HRQoL) as assessed by the Functional Assessment of Cancer Therapy-General (FACT-G).
- To assess disease related symptoms in each arm based on the NCCN Functional Assessment of Cancer Therapy- Kidney Symptom Index (FKSI-19).
- To assess changes in global health status in each treatment arm based on EuroQol's EQ-5D.
- To assess healthcare resource utilization in each treatment arm.

4 ENDPOINTS

The primary objectives of this study are to describe ORR (as assessed by an IRRC) and to compare PFS (as assessed by an IRRC) and OS of nivolumab combined with ipilimumab to sunitinib monotherapy in intermediate and poor-risk subjects with previously untreated mRCC. This is measured by the three co-primary endpoints defined in section 4.1.

The first secondary objective of this study is to compare PFS (as assessed by an IRRC) in the two treatment arms in the all randomized population. This would be measured by the same definitions of PFS, as specified in sections 4.1.1 and 4.1.3 respectively, in the all randomized population.

The second secondary objective of this study is to compare OS in the two treatment arms in the all randomized population. This would be measured by the same definition of OS, as specified in section 4.1.4, in the all randomized population.

The third secondary objective of this study is assessing ORR in the two treatment arms in all randomized population. This would be measured by the definition of ORR as specified in section 4.2.1, in the all randomized population.

4.1 Co-Primary Endpoints

Objective response rate, overall survival and progression-free survival are the co-primary endpoints.

4.1.1 Objective Response Rate

Objective response rate is defined as the proportion of randomized subjects who achieve a best response of complete response (CR) or partial response (PR) using the RECIST 1.1 criteria based on IRRC assessment. BOR is defined as the best response designation, as determined by the IRRC, recorded between the date of randomization and the date of objectively documented progression per RECIST 1.1 or the date of subsequent therapy (including tumor-directed radiotherapy and

tumor-directed surgery), whichever occurs first. For subjects without documented progression or subsequent therapy, all available response designations will contribute to the BOR assessment. Confirmation of response is required at least 4 weeks after the initial response. Duration of response (DOR) is defined as the time between the date of first documented response (CR or PR) to the date of the first documented progression as determined by the IRRC (per RECIST 1.1), 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. Time to Objective Response (TTR) is defined as the time from randomization to the date of the first confirmed documented response (CR or PR), as assessed by the IRRC. DOR and TTR will be evaluated for responders (confirmed CR or PR) only.

4.1.2 Primary Definition of Progression-free Survival

The primary definition of PFS (PFS truncated at subsequent therapy) is specified as the time between the date of randomization and the first date of documented progression, as determined by the IRRC (as per RECIST 1.1 criteria), or death due to any cause, whichever occurs first.

Subsequent therapy includes anticancer therapy, tumor directed radiotherapy, or tumor directed surgery as shown in Table 4.1.2-1. Subjects who die without a reported progression will be considered to have progressed on the date of their death. The following censoring rules will be applied for the primary definition of PFS.

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

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.

The first on-study tumor assessment is scheduled to be conducted at 12 weeks (± 1 week) following randomization. Subsequent tumor assessments are scheduled every 6 weeks (± 1 week) up to 13 months, then every 12 weeks until disease progression.

Censoring rules for the primary definition of PFS (PFS truncated at subsequent therapy) are presented as follows and in Table 4.1.2-1.

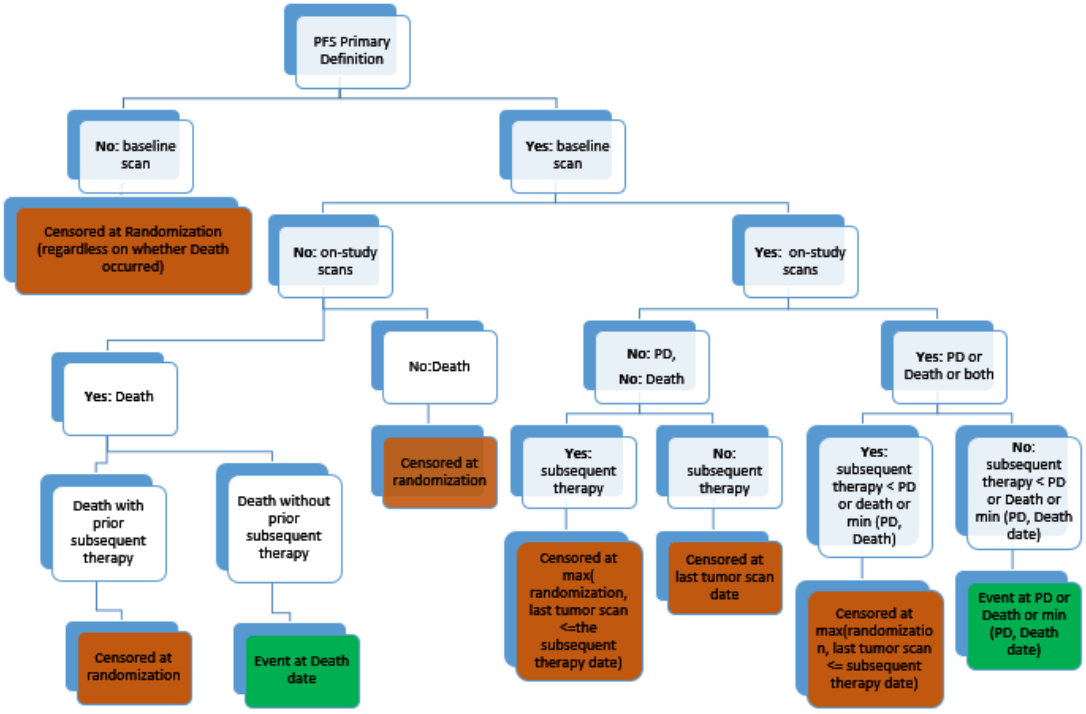


Table 4.1.2-1: Censoring Scheme used in Primary Definition of PFS

Situation	Date of Progression or Censoring	Outcome
No baseline tumor assessments	Date of randomization	Censored
No on study tumor assessments and no death	Date of randomization	Censored
Documented progression	Date of the first documented progression per RECIST 1.1 (excludes clinical progression)	Progressed
No progression and no death	Date of last evaluable tumor assessment	Censored
New anticancer therapy, tumor-directed radiotherapy, or tumor-directed surgery received without progression reported prior or on the same day	Date of last evaluable tumor assessment prior to initiation of the subsequent therapy	Censored
Death without progression	Date of death	Progressed

4.1.3 Secondary Definition of Progression-free Survival

The secondary definition of PFS (ITT definition) is defined as the time between the date of randomization and the first date of documented progression, as determined by the IRRC (as per RECIST 1.1 criteria), or death due to any cause, whichever occurs first. Subjects who die without

a reported progression will be considered to have progressed on the date of their death. The following censoring rules will be applied for the secondary definition of PFS.

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

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.

The first on-study tumor assessment is scheduled to be conducted at 12 weeks (± 1 week) following randomization. Subsequent tumor assessments are scheduled every 6 weeks (± 1 week) up to 13 months, then every 12 weeks until disease progression.

Censoring rules for the secondary definition of PFS are presented as follows and in Table 4.1.3-1.

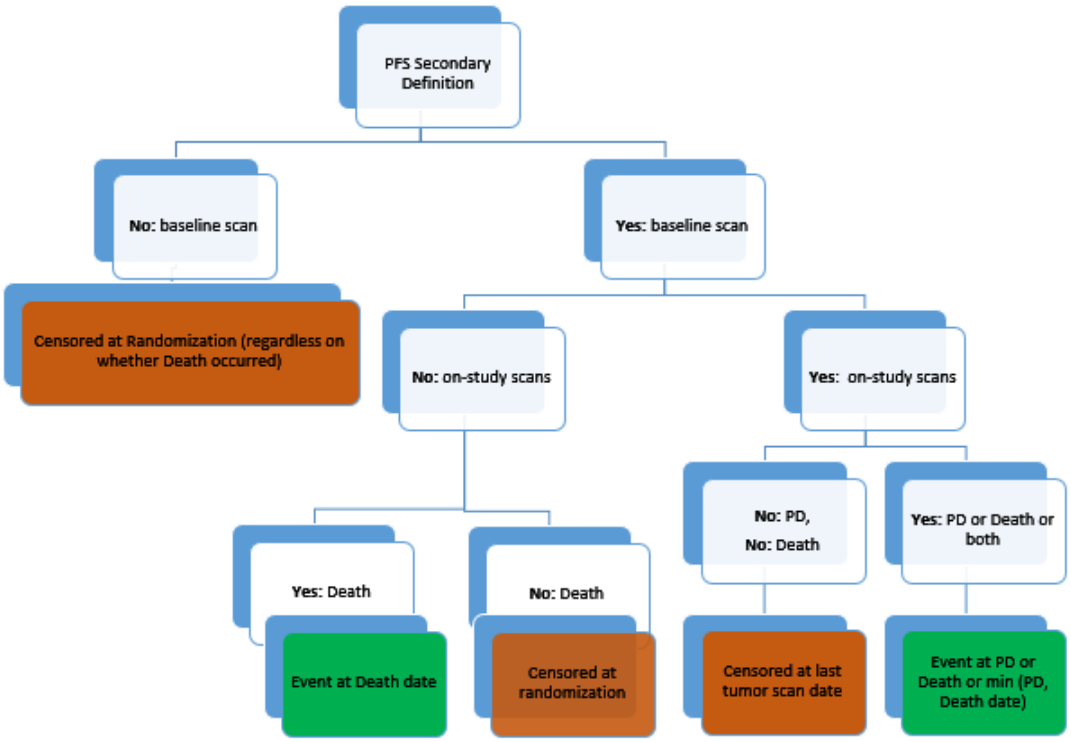


Table 4.1.3-1: Censoring Scheme for Secondary definition of PFS

Situation	Date of Progression of Censoring	Outcome
No baseline tumor assessment	Date of randomization	Censored
No on-study tumor assessments and no death	Date of randomization	Censored
Documented progression	Date of first documented progression per RECIST 1.1 (excludes clinical progression)	Progressed
No progression and no death	Date of last evaluable tumor assessment	Censored
Death without progression	Date of death	Progressed

4.1.4 Overall Survival

Overall survival is defined as the time from randomization to the date of death from any cause. For subjects that are alive, their survival time will be censored at the date of last contact ("last known alive date"). Overall survival will be censored for subjects at the date of randomization if they were randomized but had no follow-up.

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

4.2 Secondary Endpoint

4.2.1 Objective Response Rate

Objective response rate is defined as the proportion of randomized subjects who achieve a best response of complete response (CR) or partial response (PR) using the RECIST 1.1 criteria as per IRRC assessment.

The ORR (as per IRRC assessment) is defined as the number of subjects whose *confirmed* best objective (BOR) response is a complete response (CR) or partial response (PR) divided by the number of subjects in the population of interest. The BOR is defined as the best response designation, as determined by the IRRC per RECIST 1.1. For subjects without document progression or subsequent therapy, all available response designations will contribute to the BOR determination. Subsequent therapy includes anticancer therapy, tumor directed radiotherapy, or tumor directed surgery. The BOR will be determined based on response designations up to the date of last evaluable tumor assessment prior to initiation of the subsequent therapy. For subjects who continue treatment beyond progression, the BOR will be determined based on response designations up to the time of initial RECIST 1.1 progression.

The first on-study tumor assessment is scheduled to be conducted at 12 weeks (± 1 week) following randomization. Subsequent tumor assessments are scheduled every 6 weeks (± 1 week) up to 13 months, then every 12 weeks until disease progression.

4.2.1.1 Further Characterization of ORR

4.2.1.1.1 Duration of Objective Response

Duration of Objective Response (DOR) is defined as the time between the date of first documented response (CR or PR) to the date of the first documented progression as determined by the IRRC (per RECIST 1.1), 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 4.1.2-1). DOR will be evaluated for responders (i.e. subjects with confirmed CR or PR) only.

4.2.1.1.2 Time to Objective Response

Time to Objective Response (TTR) is defined as the time from randomization to the date of the first confirmed documented response (CR or PR), as assessed by the IRRC. TTR will be evaluated for responders among the population of interest (i.e. subjects with a BOR of CR or PR).

4.3 Exploratory Endpoint(s)

4.3.1 Safety

Safety and tolerability will be measured by the incidence of deaths, adverse events, serious adverse events, adverse events leading to discontinuation, adverse events leading to dose delay, select adverse events and specific laboratory abnormalities (worst grade) in each treatment group. Toxicities will be graded using the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) version 4.0. See details in the Core Safety SAP[2].

4.3.2 Pharmacokinetics

Pharmacokinetics will be measured by tshe serum concentration of nivolumab and/or ipilimumab. Samples will be collected to characterize pharmacokinetics of and to explore exposure-safety and exposure-efficacy relationships.

4.3.3 Immunogenicity

Refer to Core Safety SAP.

4.3.4 Biomarkers

[redacted]

4.3.5 FACT-G and FKSI-19 assessment

The FACT-G is a 27-item questionnaire that measures general cancer health related quality of life. It is one of the most widely used HrQoL cancer specific scales and has been validated in numerous types of cancer patients, across cultures, and in many languages. The scale is a compilation of general questions divided into four primary HrQoL dimensions: Physical Well-Being, Social/Family Well-Being, Emotional Well-Being, and Functional Well-Being. Summary scores can be calculated for each domain in addition to a single overall summary score.

The NCCN FKSI-19 is a 19-item scale that measures tumor specific HrQoL in kidney cancer patients. The FKSI-19 uses five Likert-type response categories that range from "not at all" to "very much." Patients are asked to circle the response category that best characterizes their response over the last 7 days on 19 items that include symptoms such as lack of energy, fatigue, appetite, coughing, and shortness of breath, pain, nausea and ability to work.

4.3.6 EuroQoL EQ-5D

Subjects' overall health status will be assessed using the EuroQol Group's self-reported health status measure (EQ-5D-3L)[3]. EQ-5D essentially has 2 components; the EQ-5D descriptive system and the EQ visual analogue scale (EQ VAS).

The EQ-5D descriptive system comprises the following 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension has 3 levels: no problems, some problems and severe problems. Once the data have been collected and a database created, a scoring function can be used to assign a value (i.e., EQ-5D index score) to self-reported health states from a set of population-based preference weights.

The EQ VAS records the subject's self-rated health state on a 100-point vertical, visual analogue scale (0 = worst imaginable health state; 100 = best imaginable health state)[4].

5 SAMPLE SIZE AND POWER

The sample size of the study accounts for the three co-primary efficacy endpoints: ORR and PFS as per IRRC and OS evaluated in intermediate and poor-risk subjects with previously untreated mRCC. The overall alpha for this study is 0.05, which is split with 0.001 to evaluate ORR, 0.009 to evaluate PFS and 0.04 to evaluate OS.

ORR will be analyzed initially on a descriptive basis and will occupy an administrative adjustment of alpha of 0.001. PFS will be evaluated for treatment effect at an alpha of 0.009 (two-sided, penalized 0.001 from a 0.01 allocation), with at least 90% power; no interim analysis of PFS is planned. OS will be evaluated for treatment effect at an alpha level of 0.04 (two-sided) with 90% power, accounting for two formal interim analyses to assess efficacy.

It is estimated that approximately 1070 previously untreated mRCC subjects will be randomized in a 1:1 ratio. Among them, 820 subjects (76.6%) with intermediate/poor risk and approximately 250 (23.4%) subjects with favorable risk as per IMDC (IMDC prognostic score = 0) will be randomized. Assuming a fixed accrual rate of 69 subjects per month (53 intermediate/poor risk

subjects per month), it will take 16 months to randomize 1070 subjects (820 intermediate/poor risk subjects).

Assuming a 21% screen failure rate, it is estimated that approximately 1355 subjects will be enrolled in order to have 820 intermediate/poor-risk subjects randomized. The primary analysis is based on intermediate/poor risk subjects as per IMDC prognostic score and the number of PFS/OS events observed among them. The enrollment will stop once approximately 820 intermediate/poor risk subjects have been randomized.

Sample size justification for ORR estimate

One of the primary objectives of the study is to describe the ORR (based on IRRC assessment) of nivolumab combined with ipilimumab to sunitinib monotherapy in intermediate and poor-risk subjects with previously untreated mRCC.

The primary analysis of ORR in the intermediate and poor-risk randomized subjects will be performed when these patients have at least 6 months of minimum follow-up from the completion of enrollment. This will allow sufficient follow-up for ORR to have a stable estimate, adequate safety follow-up as well as information on duration of response in this population.

The maximum width of the exact two-sided 95% confidence interval (CI) is 9.9% when the ORR is expected to be in the 20% to 50% range. Table 5-1 summarizes the 95% exact CI when observed ORRs are 20% to 50%, respectively.

Table 5-1: Observed ORR with exact 95% CI

Observed ORR	95% Exact CI
20%	(16.2% - 24.2%)
25%	(21.0% - 29.6%)
30%	(25.6% - 34.7%)
35%	(30.5% - 40.0%)
40%	(35.2% - 44.9%)
45%	(40.2% - 50.1%)
50%	(45.1% - 54.9%)

For example if at least 123 responders are observed among the 410 nivolumab and ipilimumab combination intermediate/poor risk randomized subjects (i.e. ORR ≥ 30%) then the lower bound of the 95% CI is above 25.6%.

Sample size justification for PFS comparison

One of the primary objectives of the study is to compare the progression-free survival (as determined by IRRC) of nivolumab combined with ipilimumab to sunitinib monotherapy in intermediate and poor-risk subjects with previously untreated mRCC. The number of events and power for this study were calculated assuming an exponential distribution for PFS in each arm.

For this comparison of PFS, it will be required to observe approximately 465 PFS events among the randomized intermediate/poor risk subjects in the two respective treatment arms for a two-sided experiment-wise $\alpha = 0.009$ log-rank test, to show a statistically significant difference in PFS between the treatment arms with approximately 80% power when the true hazard ratio of the experimental arm to control arm is 0.73. The HR of 0.73 is equivalent to demonstrating a 37.8% improvement in median PFS, assuming a median PFS of 9 months in the sunitinib montherapy arm (weighted median estimate assuming a median PFS of 11 months in intermediate risk subjects and median PFS of 4 months in poor risk subjects)[5] and a median PFS of 12.4 months in the experimental treatment arm.

Under the assumptions for accrual and PFS distribution stated above, it will take approximately 35 months from FPFV to observe the required number of PFS events for the final PFS analysis (16 months for accrual and 19 months for minimum follow up). It is projected that an observed HR of 0.785 or less corresponding to a 2.5 month or greater improvement in median PFS (9 vs 11.5 months) for this comparison, would result in a statistically significant improvement in the final analysis of PFS.

Update to Timing of Final PFS Analysis

The number of events for the primary endpoint, PFS per IRRC accounting for subsequent therapy, was observed to be lower than originally assumed per protocol. At 28 months after FPFV there were approximately 72% of the 591 target PFS events per IRRC among intermediate/poor risk subjects, with only 5-10 events occurring monthly over the previous 6 months. This event rate is expected to continue and as a result the target number of events is unlikely to occur even in the next 1-2 years. Thus the timing of the final PFS analysis was advanced with lower power than originally planned as summarized in Table 5-2.

Sample size justification for OS comparison:

One of the primary objectives of the study is to compare the overall survival of nivolumab combined with ipilimumab to sunitinib monotherapy in intermediate and poor-risk subjects with previously untreated mRCC. The number of events and power of this study were calculated assuming an exponential distribution for OS in each arm.

Approximately 639 events (ie, deaths), observed among the randomized intermediate/poor risk subjects, provides 90% power to detect a hazard ratio (HR) of 0.766 with an overall type 1 error of 0.04 (two-sided). The HR of 0.766 corresponds to a 30.6% increase in the median OS, assuming a median OS of 20 months for sunitinib monotherapy (weighted median estimate assuming a median OS of 26 months in intermediate risk subjects and a median of 8 months in poor risk subjects)[6] and 26.1 months for experimental treatment arms, respectively. It is projected that an observed hazard ratio of 0.846 or less, which corresponds to a 3.6 months or greater improvement in median OS (20 mo vs. 23.6 mo), would result in a statistically significant improvement in OS for the experimental arm at the final OS analysis.

Two formal interim analyses of OS are planned for this study. The first interim analysis is planned at the time of final PFS analysis and it is expected to observe 330 events (52% of the targeted OS

events for final analysis) and the second after observing 479 events (75% of targeted OS events needed for final analysis). The stopping boundaries at the interim and final analyses will be derived based on the number of deaths using O'Brien and Fleming α–spending function.

Under the assumptions stated above on accrual and OS distribution, it will approximately take 65 months from FPFV to observe the required number of OS events for the final OS analysis (16 months for accrual and 49 months for minimum follow up).

In summary, it is expected to take:

- Approximately 16 months to complete accrual
- Approximately 22 months from FPFV to obtain at least a minimum follow-up of 6 months on the intermediate and poor risk randomized subjects for the descriptive analysis of ORR
- Approximately 35 months from FPFV to obtain the approximate number of PFS events (i.e. approximately 465 events among the 820 intermediate and poor risk randomized subjects) and deaths for the first formal interim analysis of OS (i.e. approximately 330 deaths among the same population)
- Approximately 46 months from FPFV to obtain the required deaths for the second formal interim analysis of OS (i.e. 479 deaths among the intermediate and poor risk randomized subjects)
- Approximately 65 months from FPFV to obtain the required deaths for the final analysis of OS (i.e. 639 deaths among the intermediate and poor risk randomized subjects).

Table 5-2 summarizes sample size design parameters and schedule of primary endpoint analyses planned in this study.

Table 5-2: Summary of sample size parameters and schedule of analyse

Co-Primary Endpoints	**ORR**	**PFS**	**OS**
Primary analysis population	Intermediate/poor risk subjects (IMDC score ≥ 1)		
Accrual rate per month	53[b]		
Power	N/A	~80%	90%
Alpha	Administrative 0.001	0.009 2-sided	0.04 2-sided (0.0024 at IA1, 0.0137 at IA2 , 0.0354 at FA)
Hypothesized Median Control vs. exp (months)	25% vs 40%	9 vs. 12.4	20 vs. 26.1
Hypothesized Hazard ratio	N/A	0.726	0.766
Critical Hazard ratio (Observed hazard ratio at which a statistically significant difference would be observed) / Difference in median (months) Corresponding to a minimal clinically significant effect size	N/A	0.785 / 2.5	0.846/ 3.6
Critical HR at interim analysis-1(IA1) /effect size	N/A	N/A	0.72/ 7.8

Table 5-2: Summary of sample size parameters and schedule of analyse

Co-Primary Endpoints	ORR	PFS	OS
Expected number of event for IA1 (percentage of target events)	N/A	N/A	330 (52%)
Timing of IA1 from FPFV l(months)	N/A	N/A	35
Critical HR at interim analysis-2(IA2) /effect size	N/A	N/A	0.8 / 5.1
Target number of event for IA2 (percentage of target events)	N/A	N/A	479 (75%)
Timing of IA2 from FPFV (months)	N/A	N/A	46
Accrual Duration (months)	16	16	16
Timing of final analysis (FA) from FPFV (months)	22	35	65
Sample size[a]	820	820	820
Target number of events (Event Goal)	N/A	465	639

[a] East version 5.4 was used for sample size / power computation.

[b] Accrual rate adjusted to reflect observed accrual.

6 STUDY PERIODS, TREATMENT REGIMENS AND POPULATIONS FOR ANALYSES

6.1 Study Periods

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, for all treated subjects. For subjects who are randomized but not treated, baseline evaluation or events will be defined as those that occur before the date and time of randomization.
- In cases where the time (onset time of event or evaluation time and dosing time) is missing or not collected, the following definitions will apply:
  - Pre-treatment AEs will be defined as AEs with an onset date prior to but not including the day of the first dose of study treatment.
  - Baseline evaluations (laboratory tests, 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.

Post baseline period:

- On-treatment AEs will be defined as AEs with an onset date-time on or after the datetime of the first dose of study treatment (or with an onset date on or after the day of first dose of study treatment if time is not collected or is missing). For subjects who are off study treatment, AEs will be counted as on-treatment if the event occurred within 100 days of the last dose of study treatment. No "subtracting rule" will be applied when an AE occurs both pre-treatment and post-treatment with the same preferred term and grade.
- On-treatment evaluations (laboratory tests, pulse oximetry and vital signs) will be defined as evaluations taken after the day (and time, if collected and not missing) of first dose of study treatment. For subjects who are off study treatment, evaluations should be within 100 days of the last dose of study treatment.
- Late emergent drug-related AEs will be defined as drug-related AEs with an onset date greater than 100 days after the last dose of study treatment in subjects who are off study treatment.

6.2 Treatment Regimens

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

- Arm A: Experimental arm: nivolumab + ipilimumab
- Arm B: Control arm: sunitinib

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.

6.3 Populations for Analyses

- **All enrolled subjects**: All subjects who signed an informed consent form and were registered into the IVRS
- **All randomized subjects**: All subjects who were randomized to any treatment arm in the study. This population is considered as the secondary efficacy analysis population. Analysis of demography, protocol deviations, baseline characteristics, secondary efficacy analysis and outcome research analysis will be performed for this population.
- **Intermediate/poor risk subjects**: All randomized subjects with baseline IMDC prognostic score ≥ 1 at the time of randomization (IVRS). This is the primary efficacy analysis population. Analysis of demography, protocol deviations, baseline characteristics and primary efficacy analysis will be performed for this population.
- **All treated subjects**: All subjects who received any dose of study therapy. This is the primary dataset for drug exposure and safety analysis.

- **All treated intermediate/poor risk subjects**: All intermediate/poor risk subjects who received any dose of study therapy. **Favorable risk subjects**: All randomized subjects with baseline IMDC prognostic score = 0 at the time of randomization (IVRS). This population would be used for conducting exploratory analysis of efficacy endpoints.
- **PK subjects**: All subjects with available serum time-concentration data from randomized subjects dosed with nivolumab.
- **Immunogenicity subjects**: All subjects with available data from randomized subjects dosed with nivolumab.
- **PD-L1 treated subjects:** All subjects with a PD-L1 assessment at baseline who received any dose of study therapy.

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

7 STATISTICAL ANALYSES

7.1 General Methods

Unless otherwise noted, the bulleted titles in the following subsections describe tabulations of discrete variables, by the frequency and proportion of subjects falling into each category, grouped by treatment (with total). Percentages given in these tables will be rounded and, therefore, may not always sum to 100%. Continuous variables will be summarized by treatment group (with total) using 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 (i.e. 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 95% CI will be estimated using log-log transformation. Rates at fixed time points (e.g. 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]

The unweighted difference in ORRs between the two treatment arms and corresponding asymptotic 95% CI will be estimated using a Newcombe method.[9]

Unless otherwise specified, the stratified log-rank test will be performed to test the comparison between time to event distributions (PFS and OS). 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 a unique covariate.

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

7.2 Study Conduct

7.2.1 Accrual

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

Furthermore, the accrual pattern will be summarized per month by the stratification factors.

7.2.2 Relevant Protocol Deviations

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

Eligibility:

- Subjects with baseline KPS < 70%
- Subjects who received prior systemic anti-cancer treatment in the metastatic setting
- Subjects without histologically confirmed RCC with a clear-cell component, documented advanced or metastatic (AJCC Stage IV) RCC

Eligibility (only for intermediate/poor risk subjects):

- Subjects with a baseline IMDC prognostic score < 1

On-study:

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

Listings will also be provided.

7.3 Study Population

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

7.3.1 Subject Disposition

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

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

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

A subject listing for all randomized subjects 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 subjects not randomized will also be provided, showing the subject's race, gender, age, consent date and reason for not being randomized.

7.3.2 Demographics and Other Baseline Disease Characteristics

The following baseline characteristics will be summarized by treatment arm as randomized:

- Age
- Age categorization (< 65, ≥ 65 and < 75, ≥ 75 and < 85, ≥ 85, ≥ 75, ≥ 65)
- Gender (Male vs. Female)
- Race (White, Black or African American, Asian, Other)
- Ethnicity (Hispanic/Latino and Not Hispanic/Latino)
- Karnofsky performance status(70, 80, 90, 100)
- Baseline IMDC prognostic score (0, 1-2, ≥ 3) (source: CRF)
- Prior nephrectomy
- Prior radiotherapy
- Time from initial disease diagnosis to randomization (<1 year, ≥1 year)
- LDH level (≤ 1.5 x ULN, >1.5 x ULN)
- Hemoglobin (<LLN, ≥ LLN)
- Corrected Calcium (≤ 10 mg/dl, >10mg/dl)
- Alkaline phosphatase (< ULN, ≥ ULN)
- Region (per IVRS)
- Baseline PD-L1+ status based on a 1% cut off
- Baseline PD-L1+ status based on a 5% cut off
- Baseline PD-L1+ status based on a 10% cut off
- Sites of diseases (all lesions)
- Number of disease sites per subject (all lesions)
- Tumor burden: sum of the diameters of target lesions at baseline
- Pre-treatment events: summarized by worst CTC grade presented by SOC/PT

7.3.3 Medical history

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

7.3.4 Prior therapy agents

- Prior adjuvant or neo-adjuvant therapy agents for localized or locally advanced RCC will be summarized.

7.3.5 Baseline examinations

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

7.3.6 Baseline Physical Examination

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

7.3.7 Discrepancies Between IVRS and CRF stratification factors

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

- Baseline IMDC prognostic score (0 vs. 1-2 vs. ≥ 3)

7.4 Extent of Exposure

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

7.4.1 Administration of Study Therapy

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

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

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

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

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.

Table 7.4.1-1: Study Therapy Parameter Definitions

	Nivolumab	Ipilimumab	Sunitinib
Dosing Schedule per Protocol	3 mg/kg every 3 weeks for 4 doses followed by 3 mg/kg every 2 weeks	1 mg/kg every 3 weeks for 4 doses	50 mg PO once daily for 4 weeks followed by 2 weeks off.
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.	mg
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.	mg sum of the doses administered to a subject
Cycle Duration$_{(i)}$ (wk)	(Dose date$_{(i+1)}$ - Dose date$_{(i)}$)/7	N/A	N/A
Cycle Intensity$_{(i)}$ (mg/kg/wk)	Dose$_{(i)}$/Cycle Duration$_{(i)}$	N/A	N/A
Relative Cycle Intensity $_{(i)}$ (%)	(Cycle Intensity$_{(i)}$/intended dose per week)(i) * 100	N/A	N/A
Relative Dose Intensity (%)	Sum of all Relative Cycle Intensities divided by N	Cum dose /[(Last dose date - Start dose date + 21) x 1/21] x 100	See below
Duration of Treatment	*Last dose date - Start dose date +1*	*Last dose date - Start dose date +1*	*Last dose date - Start dose date +15*

Volume infused, volume prepared, and weight are collected on the CRF. Nominal nivolumab dose collected in IVRS.

i = 1, 2,...,N, where N = number of infusions. Cycle Duration $_{(N)}$ = 3 weeks for nominal 3 mg/kg nivolumab doses in combination period and 2 weeks for nominal 3 mg/kg in monotherapy period. Intended dose per week is 1 mg/kg for nominal 3 mg/kg nivolumab doses in combination period and 1.5 mg/kg for nominal 3 mg/kg nivolumab doses in monotherapy period.

Additional Parameters -Sunitinib treatment

Average daily dose (in mg/day) is defined as:

Sum of all Sunitinib doses in mg actually received / duration of treatment in days

Since Sunitinib treatment consists of 50 mg PO daily dose for 4 weeks followed by 2 weeks of washout period, the planned dose intensity of Sunitinib is 33.33 mg/day (50 mg x 28 days / 42 days)

Relative dose intensity for Sunitinib (%) is defined as: *(Average daily dose / 33.33) x 100*

7.4.2 Modifications of Study Therapy

7.4.2.1 Dose Delays

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

Dose Delays/Omission/interruption-Sunitinib treatment

Dose omission/delayed/interruption for subjects treated with Sunitinib will occur if the subject did not receive any dose during at least one day for a different reason than "Dosing Error" or "No Change"; provided that treatment was resumed afterwards. Reason for dose omission will be retrieved from CRF dosing pages. Dose delay/omission/interruption is considered the same since this is a continuous daily dosing.

It is worth noting that during the two week mandatory washout period, a daily dose of 0 mg will be entered in the CRF pages, with corresponding reason for dose modification recorded as "No Change".

The following parameters will be summarized by treatment group:

- Number of dose delays per subject, length of delay, and reason for delay

7.4.2.2 Infusion Interruptions and Rate Changes

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

The following parameters will be summarized by treatment group:

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

7.4.2.3 Dose Escalations

Dose escalations are not permitted for either nivolumab or ipilimumab.

7.4.2.4 Dose Reductions

Dose reductions are not permitted for either nivolumab or ipilimumab.

Dose reduction for subjects treated with Sunitinib is defined as at least one day with a non zero dose smaller than 50mg and smaller than previous non zero dose with a CRF reason different from "Dosing Error" or "No Change".

7.4.2.5 Dose Omissions

Dose omissions are not permitted for either nivolumab or ipilimumab.

7.4.3 Concomitant Medications

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

The following summary tables will be provided:

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

A by-subject listing will accompany the table.

7.5 Efficacy

The formal analysis of ORR in the intermediate and poor-risk randomized subjects will be performed when these patients have an at least 6 month minimum follow-up from the completion of enrollment. An administrative allocation of 0.001 alpha will be used. At the time the ORR analysis, a reduced set of the total analyses defined in this document will be performed (e.g. no PFS/OS analysis). The relevant analyses will be specifically defined in the data presentation plan.

Principal analyses of PFS and ORR will be based on the IRRC evaluation, unless noted otherwise. Unless stated otherwise, whenever a stratified analysis is specified using intermediate/poor risk subjects, the following stratifications factors (recorded at randomization as per IVRS) will be used:

- IMDC prognostic risk score (1-2 vs. 3-6)
- Region (US vs.Canada/W.Europe/N.Europe vs. ROW)

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

- IMDC prognostic risk score (0 vs.1-2 vs. 3-6)
- Region (US vs.Canada/W.Europe/N.Europe vs. ROW)

For assessing the secondary objectives of this study, a hierarchical testing procedure[10] will be used so that the overall experiment-wise Type I error rate is 0.05.

The key secondary objectives among all randomized subjects will be tested after conducting the primary objective analyses on intermediate/poor risk subjects.

The formal testing of ORR per IRRC using 95% exact CIs among all randomized subjects will take place if the 95% exact CI of ORR per IRRC across treatment groups among all randomized intermediate/poor risk subjects do not overlap. The analysis of ORR is descriptive so no p-values for rate differences will be reported.

Similarly, the formal testing of PFS as per IRRC, at a two sided 0.009 significance level, among all randomized subjects will take place if PFS per IRRC among intermediate/poor risk subjects is statistically significant. Likewise, the testing of OS, at a two sided 0.04 significance level, among all randomized subjects will take place only if OS intermediate/poor risk subjects is statistically significant.

All p-values reported will be two-sided. Confidence Interval for co-primary and secondary endpoint analyses included in hierarchy (PFS and OS) will be based on nominal significance level adjusted for co-primary endpoints and interim analyses to preserve overall type one error rate (See sections 7.5.1, 7.5.6, 7.5.7 and 7.5.8 for details). Alpha (α) for the confidence interval will be the same as nominal significance level for hypothesis testing. CIs for other endpoints will be at the two-sided 95% level. The p-values presented in the clinical study report will be rounded to the fourth decimal place. Point estimates and confidence bounds for efficacy variables will be rounded to the second decimal place.

7.5.1 Analysis of Objective Response Rate towards Primary Objective

One of the primary objectives of the study is to estimate the objective response rate per IRRC in the two treatment arms among intermediate and poor risk subjects. For the ORR per IRRC analysis both the final and interim CRF pages collect BOR following this algorithm:

- Apply final BOR if available (from level1 VA), otherwise use the most recent interim BOR (from level1 TM)

Estimates of response rate, along with its exact two-sided 95% CI by Clopper-Pearson method, will be computed within each treatment arm. A two sided 95% CI for difference of response rate between the treatment arms will also be computed.

Sensitivity analysis based on investigator-determined ORR may also be performed. DOR and TTR will also be evaluated. At the time of the formal ORR analysis, there will be no formal analysis of PFS and OS.

One of the secondary objectives of the study is to estimate the objective response rate in the two treatment arms among all randomized subjects.

The number and percentage of subjects in each category of best overall response per IRRC (complete response [CR], partial response [PR]], stable disease [SD], progressive disease [PD], or unable to determine [UD]) according to the IRRC will be presented, by treatment group. An estimate of the response rate and an associated exact two-sided 95% CI (Clopper and Pearson[11]) will be presented, by treatment group.

A 2-sided, 95% confidence interval for the difference of ORR between treatment arms will be computed for all randomized subjects by the method of DerSimonian and Laird[12], using a fixed-effects model (setting Δ^2 equal to zero), adjusting for the stratification factors. The weighted response rate difference and 95% CI can be determined using the following formula:

$$\hat{\theta} = \frac{\sum_{i=1}^{12} \hat{\theta_i} w_i}{\sum_{i=1}^{12} w_i} \sim N(\theta, 1/\sum_{i=1}^{12} w_i)$$

where $\hat{\theta_i}$ is the response rate difference of the i[th] stratum and $w_i = 1/var(\hat{\theta}_i)$.

Similar analyses will be repeated based on the investigator's assessment of ORR. A cross tabulation of IRRC best response versus the investigator best response will be presented, by treatment group.

7.5.2 Analysis of Progression-Free Survival towards Primary Objective

One of the primary objectives of the study is to compare the progression-free survival (as determined by IRRC) of nivolumab combined with ipilimumab to sunitinib monotherapy in intermediate and poor-risk subjects with previously untreated mRCC. All the analyses outlined in this section are specified for intermediate/poor risk subjects population.

The principal analysis of PFS (as determined by IRRC) will be to compare the two treatment arms via two sided 0.009 stratified log-rank test among intermediate/poor risk subjects. The primary definition of PFS will be used in this analysis. The two-sided log-rank p-value will be reported.

The estimate of the PFS hazard ratio, of nivolumab combined with ipilimumab to sunitinib monotherapy, will be calculated using a stratified Cox proportional hazards model, with treatment as the sole covariate. Ties will be handled using the Exact method. A two-sided, 99.1% CI for the hazard ratio will also be presented.

The PFS function for each treatment arm will be estimated using the KM product limit method and will be displayed graphically. A two-sided 95% CI for median PFS in each arm will be computed via the log-log transformation method. Estimates for one-year and two-year PFS rates will be presented along with their associated 95% CIs. Minimum follow-up must be greater than or equal to the time-point to generate the rate. These estimates will come from the KM curve and their standard errors (SEs) and associated CIs, will be computed using log-log transformed 95% confidence intervals[13].

The method of Gail and Simon[14] will be used to test for a qualitative interaction between treatment and strata, IMDC prognostic risk score (1-2 vs. 3-6) and Region (US vs. Canada/W.Europe/N.Europe vs. ROW). This test will be conducted at α= 0.10 level. The p-value reported from this specific analysis is for descriptive purposes only.

The proportional hazards assumption will be assessed via the following hazard rate model, which contains a time dependent covariate:

$$\lambda(t, z) = \lambda_i(t) e^{(\{b_1 + b_2 \times [\log(t)]\} \times Z)}, \qquad i = \{1 - 6\}$$

where i=1-6 corresponds to each of the six levels the stratum can take, and Z is the treatment indicator, which is equal to 1 for the combination arm and 0 for the control arm.. The null hypothesis, that the proportional hazards assumption is valid, i.e., that $b_2=0$, will be tested against the alternative hypothesis that $b_2 \neq 0$ using a Wald statistic at $\alpha= 0.10$ level. The p-value reported from this specific analysis is for descriptive purposes alone. A plot of smoothed scaled Schoenfeld residuals of the above model will be used to graphically illustrate the evolution of the hazard ratio over time.

The source of PFS event (progression or death) will be summarized by treatment group.

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

The status of subjects who are censored (as per primary definition of PFS) in the PFS KM analysis will be tabulated for each randomized treatment group using the following categories:

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

7.5.3 Supportive Analyses of Progression-Free Survival

The following sensitivity analyses will be conducted using both the primary and the ITT definitions of PFS in the intermediate/poor risk subjects:

1. Delayed effect of immunotherapy interventions may cause a late separation in the progression free survival KM curves and non-proportional hazards as was observed in the second line phase 3 mRCC study (CA209-025). The principal analysis of PFS (as determined by IRRC) will be to compare the two treatment arms via two sided 0.009 stratified weighted log-rank test among intermediate/poor risk subjects. The primary definition of PFS will be used in this analysis. The two-sided stratified weighted log-rank p-value will be reported using G (rho = 0, gamma = 1) weights, in the terminology of Fleming and Harrington[15].

   The Fleming Harrington test can be unstable, so it is possible, though uncommon, that the p-value for this trial will not be estimable. In this case, the primary analysis will default to using the two sided 0.009 stratified log-rank test among intermediate/poor risk subjects (as specified in section 7.5.2).

   The estimate of the PFS hazard ratio in the period following 6 months, of nivolumab combined with ipilimumab compared to sunitinib monotherapy, will be calculated using a stratified time-dependent Cox model with effects for treatment and period-by-treatment interaction. In this model, period is a binary variable indicating pre- vs post- 6 months. The second line phase 3 mRCC study (CA209-025) served as the basis for the 6 month delayed treatment effect in PFS. Ties will be handled using the exact method. A two-sided 99.1% CI for the hazard ratio will also be presented.

2. A multivariate Cox regression model will be used in order to estimate the treatment effect after adjustment for possible imbalances in known or potential prognostic factors. The factors used in the randomization, which, by definition, will be balanced across arms, will still be included in the model as stratification factors. However, all additional factors will be incorporated as covariates. The additional factors, which are all measured at baseline, will include:

    a. LDH (≤ 1.5 x ULN, > 1.5 x ULN)

    b. Previous Nephrectomy (Yes, No)

   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 99.1% CIs.

3. PFS using stratification factors as obtained from the baseline CRF pages (instead of IVRS). The hazard ratio associated with treatment will be presented along with the associated two-sided 99.1% CIs. This analysis will be performed only if at least one stratification variable/factor at randomization (as per IVRS) and baseline are not concordant for at least 10% of the randomized intermediate / poor risk subjects.
4. PFS using the investigator's assessment. The hazard ratio associated with treatment and median PFS will be presented along with the associated two-sided 99.1% CIs.
5. PFS using an un-stratified log rank test. The hazard ratio associated with treatment will be presented along with the associated two-sided 99.1% CIs.
6. PFS using an unstratified Cox proportional hazards model, adjusted, using as covariates only the two stratification factors used in randomization. The hazard ratio associated with treatment will be presented along with the associated two-sided 99.1% CIs.
7. PFS for subjects with no relevant deviation. This analysis will be conducted only if there are more than 10% subjects with relevant protocol deviations. The hazard ratio associated with treatment will be presented along with the associated two-sided 99.1% CIs.

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

7.5.4 Concordance Between IRRC and Investigator Assessments of Progression

For the purpose of assessing concordance between the IRRC and investigator tumor assessments among intermediate/poor risk subjects, progression status will be categorized as documented progression, death or censored. A cross tabulation between the IRRC and the investigator

progression status will be presented, by treatment group. The secondary definition of PFS (ITT) will be used for this analysis.

The number of subjects with the same timing of IRRC and investigator documented progression will be summarized. In addition, a frequency table of the time between the date of the IRRC documented progression and the date of the investigator documented progression (weeks) will be presented by treatment group, for subjects who had documented progression according to both the IRRC and the investigator at different time points. The time between dates will be defined as:

$$\frac{\left(\begin{array}{c} date\ of\ IRC\ documented\ progression- \\ date\ of\ investigator\ documented\ progression \end{array}\right)}{7}$$

The time between dates will be categorized as < -12 weeks, -12 weeks to < -8 weeks, -8 weeks to < -4 weeks, -4 weeks to < -2 weeks, -2 weeks to < 0 weeks, 0 weeks to < 2 weeks, 2 weeks to < 4 weeks, 4 weeks to < 8 weeks, 8 weeks to < 12 weeks, ≥ 12 weeks. Subjects who only progressed per investigator and had a death event per IRRC will also be summarized along with the difference in the timing of events. The time between dates will be categorized as < 10 weeks, 10 to < 20 weeks, ≥ 20 weeks.

A by subject listing of IRRC and investigator PFS status and the time between progression dates according to the IRRC and the investigator will be provided.

7.5.5 Subset Analyses of Progression-Free Survival

The influence of baseline and demographic characteristics on the treatment effect among intermediate/poor risk subjects will be explored via exploratory subset analyses for the following factors

- Age categorization (< 65 vs. ≥ 65 - < 75 vs. ≥ 75)
- Gender (Male vs. Female)
- Race
- Region (US vs.Canada/W.Europe/N.Europe vs. ROW)
- Karnofsky performance status(< 90 vs. ≥ 90)
- Baseline IMDC prognostic score (1-2, ≥ 3) (source: CRF)
- Prior adjuvant or neo-adjuvant therapy for localized or locally advanced RCC (Yes, No)
- Prior Nephrectomy (Yes, No)
- Prior Radiotherapy (Yes, No)
- Time from initial disease diagnosis to randomization (< 1 year, ≥ 1 year)
- LDH level (≤ 1.5 x ULN , > 1.5 x ULN)
- Hemoglobin (< LLN , ≥ LLN)
- Corrected Calcium (≤ 10 mg/dl, > 10mg/dl)
- Alkaline phosphatase (< ULN, ≥ ULN)
- Baseline PD-L1+ status based on a 1% cut off
- Baseline PD-L1+ status based on a 5% cut off

- Baseline PD-L1+ status based on a 10% cut off

A forest plot of the PFS hazard ratios (along with the 95% CIs) will be produced for each level of the subgroups listed above.

All the above mentioned analyses (except age, race, region, and gender) will be conducted if the number of subjects in each subgroup is more than 20. Estimates of median PFS would be computed for all the subsets.

7.5.6 Analysis of Overall Survival towards the Primary Objective

One of the primary objectives of the study is to compare the overall survival of nivolumab combined with ipilimumab to sunitinib monotherapy in intermediate and poor-risk subjects with previously untreated mRCC. All the analyses outlined in this section are specified for intermediate/poor risk subjects population.

Overall survival will be compared between the treatment arms at the interim and final analyses, using stratified log-rank test. The stratification factors will be those used in the analysis of PFS. An O'Brien and Fleming α-spending function will be employed to determine the nominal significance levels for the interim and final analyses. The stratified hazard ratio between the treatment groups will be presented along with 100*(1- α)% CI (adjusted for interim). In addition, two-sided p-value will also be reported for the primary analysis of OS.

All analyses performed for PFS (detailed in section 7.5.1) will be repeated for OS. Supportive analyses 1, 2, 4, 5 and 6 of PFS (detailed in section 7.5.3) as well as the subset analyses (detailed in section 7.5.5) will also be repeated for OS.

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

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

Estimates for 1, 2 and 3-year OS rates will be presented along with their associated 95% CIs. These analyses will be only performed if the minimum follow-up for OS has reached corresponding to that endpoint. These estimates and their standard errors (SEs) will be come from the KM curve for use in constructing CIs computed using log-log transformed 95% confidence intervals.

7.5.7 Analysis of Progression-Free Survival towards the Secondary Objective

One of the key secondary objectives of the study is to compare the progression-free survival (as determined by IRRC) of nivolumab combined with ipilimumab to sunitinib monotherapy in all randomized subjects with previously untreated mRCC.

If the formal comparison of PFS (as per IRRC) in the intermediate/poor risk subjects is found to be statistically significant, then PFS (as determined by IRRC) will be compared between the two treatment arms via two sided 0.009 stratified log-rank test among intermediate/poor risk subjects.

The primary definition of PFS will be used in this analysis. The two-sided log-rank p-value will be reported. Further analysis of PFS will include estimation of the hazard ratio and estimation of the PFS distribution in each treatment group.

The estimate of the PFS hazard ratio, of nivolumab combined with ipilimumab to sunitinib monotherapy, will be calculated using a stratified Cox proportional hazards model, with treatment as the sole covariate. Ties will be handled using the Exact method. A two-sided, 99.1% CI for the hazard ratio will also be presented.

The PFS function for each treatment arm will be estimated using the KM product limit method and will be displayed graphically. A two-sided 95% CI for median PFS in each arm will be computed via the log-log transformation method. Estimates for one-year and two-year PFS rates will be presented along with their associated 95% CIs. Minimum follow-up must be greater than or equal to the time-point to generate the rate. These estimates and their standard errors (SEs) will be come from the KM curve for use in constructing CIs computed using log-log transformed 95% confidence intervals.

The source of PFS event (progression or death) will be summarized by treatment group.

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

The status of subjects who are censored (as per primary definition of PFS) in the PFS KM analysis will be tabulated for each randomized treatment group using the following categories:

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

Received subsequent anticancer therapy

Supportive analyis (3) of PFS (detailed in section 7.5.3) as well as the subset analyses (detailed in section 7.5.5) will also be repeated for PFS among all randomized subjects, if the principal comparison of PFS among all randomized subjects is found to be statistically significant.

7.5.8 Analysis of Overall Survival towards the Secondary Objective

One of the key secondary objectives of the study is to compare the overall survival of nivolumab combined with ipilimumab to sunitinib monotherapy in all randomized subjects with previously untreated mRCC. The testing of OS, at a two sided 100*(1-α)% significance level, among all randomized subjects will take place only if OS intermediate/poor risk subjects is statistically significant.

Analyses of OS towards secondary objective will be similar to PFS analyses outlined in 7.5.7.

7.5.9 Current status of PFS and OS follow-up

Time from last tumor assessment to data cut-off in months will be summarized by treatment arm and overall for all randomized subjects. Subjects who have a PFS event will be considered as current for this analysis. The ITT definition of PFS will be used for this summary.

In addition Kaplan-Meier plots of time from randomization to post-baseline tumor assessment will be produced by treatment arm for the first twelve assessments.

Current status of OS follow-up will be summarized in months, by computing the time from "last known alive" date to data cut-off date. Subjects who have a death event will be considered as current for this analysis.

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

7.5.10 Analysis of Objective Response

One of the secondary objectives of the study is to estimate the objective response rate in the two treatment arms among intermediate/poor risk and all randomized subjects separately.

The number and percentage of subjects in each category of best overall response per IRRC (complete response [CR], partial response [PR]], stable disease [SD] (including Non-CR/Non-PD), progressive disease [PD], or unable to determine [UD]) according to the IRRC will be presented, by treatment group. An estimate of the response rate and an associated exact two-sided 95% CI (Clopper and Pearson[16]) will be presented, by treatment group.

A 2-sided, 95% confidence interval for the difference of ORR between treatment arms will be computed for all randomized subjects by the method of DerSimonian and Laird[17], using a fixed-effects model (setting Δ^2 equal to zero), adjusting for the stratification factors. The weighted response rate difference and 95% CI can be determined using the following formula:

$$\hat{\theta} = \frac{\sum_{i=1}^{12} \hat{\theta}_i w_i}{\sum_{i=1}^{12} w_i} \sim N(\theta, 1/\sum_{i=1}^{12} w_i)$$

where $\hat{\theta}_i$ is the response rate difference of the i[th] stratum and $w_i = 1/var(\hat{\theta}_i)$.

Similar analyses will be repeated based on the investigator's assessment of ORR. A cross tabulation of IRRC best response versus the investigator best response will be presented, by treatment group.

7.5.11 Subset Analyses of Objective Response

The influence of baseline and demographic characteristics on the treatment effect will be explored via exploratory subset analysis. The subsets will be the same as those analyzed for PFS and will be reported based on the IRRC assessment of ORR.

The un-weighted differences in ORR between the two treatment groups and corresponding 95% two-sided CI using the method of Newcombe will be provided.

7.5.12 Time to Tumor Response, Time in Response, and Duration of Response

The distributions of duration of response will be estimated, by arm, using the KM product limit method. The KM estimates will be presented graphically and tables will be produced presenting number of events, number of subjects involved, medians, and 95% CIs for the medians.

Time to tumor response, which does not involve censoring, will be summarized by treatment group, using descriptive statistics.

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

7.5.13 Interim Analysis of Overall Survival

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

Two interim analyses of OS are planned. The first and second interim analyses are scheduled at the time of final PFS analysis when 370 deaths (approximately 58% of the targeted OS events) are expected and after 479 deaths (approximately 75% of total deaths) have been observed, respectively, among intermediate/poor risk subjects based on above accrual rate and the exponential distribution in each arm. These formal comparisons of OS will allow for early stopping for superiority, and the boundaries for declaring superiority will be derived based on the actual number of deaths using Lan-DeMets spending function with O'Brien and Fleming type of boundary in EAST v5.4. If the first interim analysis is performed exactly at 370 deaths, the boundary in terms of statistical significance for declaring superiority would be 0.0045 (HR=0.74, 6.9 months improvement in median OS) and if the second interim analysis is performed at exactly 479 deaths, the boundary in terms of statistical significance at the interim analysis for declaring superiority would be 0.0131 (or 0.8 with regard to HR boundary, which corresponds to 5.1 months improvement in median OS under the assumed control arm hazard function). The boundary for declaring superiority in terms of statistical significance for the final analysis after 639 events would be 0.0354.

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

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

Implications of OS Interim Analysis

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

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

7.6 Safety

Safety summary tables will be generated for all treated subjects. Listings will include all available data.

7.6.1 All Adverse Events

See CORE Safety SAP[2]. In addition, summary tables will be presented for all treated intermediate/poor risk subjects.

7.6.2 Deaths

See CORE Safety SAP[2]. In addition, summary tables will be presented for all treated intermediate/poor risk subjects.

7.6.3 Serious Adverse Events

See CORE Safety SAP[2]. In addition, summary tables will be presented for all treated intermediate/poor risk subjects.

7.6.4 Adverse Events Leading to Discontinuation/Modification of Study Therapy

See CORE Safety SAP[2].

7.6.5 Multiple Events

See CORE Safety SAP[2].

7.6.6 Other Observations Related to Safety

See CORE Safety SAP[2].

7.6.7 Select Adverse Events

See CORE Safety SAP[2].

7.6.8 Immune-Mediated Adverse Events

See CORE Safety SAP[2].

7.6.9 Other Events of Special Interest

See CORE Safety SAP[2].

7.6.10 Clinical Laboratory Evaluations

7.6.10.1 Hematology

See CORE Safety SAP[2]. In addition, summary tables will be presented for all treated intermediate/poor risk subjects.

7.6.10.2 Serum Chemistry

Amylase and lipase will be summarized in addition to the serum chemistry parameters described in the CORE safety SAP[2].

7.6.11 Immunogenicity

See CORE Safety SAP[2].

7.6.12 Vital Signs and Physical Findings

See CORE Safety SAP[2].

7.6.13 Quality of Life

Fact-G, FSKI-19 and EQ-5D instruments will be summarized and listed.

7.7 Biomarker Analysis

[redacted]

7.7.1 PD-L1 Expression

Analyses 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. If there is an indication of a meaningful association, future work will evaluate PD-L1 expression as a predictive biomarker, including selection of an optimal PD-L1 expression cut-off to classify subjects as PD-L1 positive or PD-L1 negative. Cut-off selection and validation will be conducted across studies and reported outside of individual clinical study reports.

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

PD-L1 status is a categorical variable by X% cut off for quantifiable PD-L1 expression:

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

where X denotes the possible PD-L1 expression cut-off values of 1%, 5% and 10%. Additional cut off values may also be explored.

Analyses of PD-L1 will include:

- Examine the distribution of PD-L1 expression
- Assess potential association between PD-L1 status and efficacy measures
- Evaluate the potential predictive relationship of the PD-L1 status and efficacy measures
- Test performance statistics such as sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV)
- Assess potential association between PD-L1 status and overall AEs

7.7.1.1 Analyses

1) Descriptive statistics of PD-L1 expression and PD-L1 status, analyses will be based on all evaluable PD-L1 subjects if not otherwise specified:
   - Listing of all PD-L1 IHC data, all tested PD-L1 subjects.
   - Summary of tumor specimen acquisition and characteristics, all randomized subjects.
   - Summary statistics of PD-L1 expression by treatment groups of select subgroups and overall.
   - Box plot of PD-L1 expression by treatment group and overall.
   - Cumulative distribution plot of PD-L1 expression versus population percentile by treatment group and overall.
   - Waterfall plots of individual PD-L1 expression by treatment group and overall.
   - Frequency of PD L1 Status (X%), including indeterminate and unknown if over 5% of subjects fall in this category, by treatment group for select subgroups and overall, all PD-L1 tested subjects.
   - Frequency of PD-L1 immune cell expression categories (no tumor associated immune cells, no PD-L1 expression tumor associated immune cells, and rare, intermediate and numerous PD-L1-expressing tumor-associated immune cells) by treatment group, overall and by tumor PD-L1 expression status.

Note: Selected subgroups are identical to the subgroups used for OS subgroup analysis (Section 7.5.4)

2) Evaluation of associations between PD-L1 status and efficacy measures, analyses will be based all PD-L1 tested subjects 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 unknown or indeterminate subgroup

   Analyses for OS endpoint:

   For each of the subgroup:
   - OS curves for each treatment group will be estimated using the Kaplan-Meier product limit method. Two-sided, 95% confidence intervals for median OS will be constructed based on a log-log transformed CI for the survivor function S(t).
   - Forest plot of Hazard Ratios with 95% CIs

   Analyses for ORR (BOR):
   - Box plots of PD-L1 expression versus Response Status by treatment group

   For each of the subgroup:
   - Frequency and percentage of BOR will be summarized for each treatment group.

- ORR will be computed by treatment group along with exact 95% CIs using the Clopper-Pearson method.

Analyses for PFS endpoint:

For each of the subgroup:

- PFS curves for each treatment group will be estimated using the Kaplan-Meier product limit method. Two-sided, 95% confidence intervals for median PFS will be constructed based on a log-log transformed CI for the survivor function S(t).
- Forest plot of Hazard Ratios with 95% CIs

3) Evaluation of the potential predictive relationship of PD-L1 status for efficacy measures all PD-L1 evaluable subjects

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 treatment vs. control and its associated 95% CI for each of the PD-L1 status subgroup
- Hazard ratio PD-L1 positive vs. negative and its associated 95% CI within each treatment group.

Analyses for ORR endpoint:

A logistic regression model will be fitted for response (yes=CR or PR, No=SD or PD or unknown) with treatment, PD-L1 status and the treatment by PD-L1 status interaction. Although the study is not designed to have appropriate power to formally test the interaction of the model, an interaction test at significance level of 0.2 will warrant further exploration and the following statistics will be reported:

- Interaction p-value
- Odds ratio of treatment vs. control and its associated 95% CI will be reported for each of the PD-L1 status subgroup
- Odds ratio of PD-L1 positive vs. negative and its associated 95% CI will be reported for each treatment group

Analyses for PFS endpoint:

A Cox proportional hazards regression model will be fitted for PFS with treatment, PD-L1 status, and treatment by PD-L1 status interaction. Although the study is not designed to have appropriate power to formally test the interaction of the model, an interaction test at

significance level of 0.2 will warrant further exploration and the following statistics will be reported:

- Interaction p-value
- Hazard ratio of treatment vs. control and its associated 95% CI for each of the PD-L1 status subgroup
- Hazard ratio PD-L1 positive vs. negative and its associated 95% CI within each treatment group.

4) Test performance statistics for PD-L1 status vs. efficacy measures on all PD-L1 evaluable subjects:

   For each treatment group the following will be produced:

   - 2 by 2 contingency table of PD-L1 status by response status (yes=CR or PR; No=SD or PD or unknown).
   - Sensitivity, specificity, PPV and NPV will be reported along with the contingency table.

5) Association of all select AE and PD-L1 expression, all PD-L1 treated subjects

   Overall summary of any select AEs by worst CTC grade (any grade, grade 3-4, grade 5) presented by SOC/PT by treatment group for following subgroups will be provided

   - Each PD-L1 status subgroup
   - PD-L1 unknown or indeterminate subgroup.

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[18]. 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[19].

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

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

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

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

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

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

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

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

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

Duration = (Last date - first date + 1)

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

9 CONTENT OF REPORTS

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

10 REFERENCES

1 CA209-214 Data Monitoring Committee charter

2 Core Safety Statistical Analysis Plan for Nivolumab Program. BMS Document 930072588 3.0.

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

4 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.

5 Motzer R.J, Hutson T.E., Tomczak P., Michaelson M.D. et al. Sunitinib versus interferon alfa in metastatic renal-cell carcinoma N Engl J Med, 356 (2) (2007), pp. 115–124.

6 Motzer R.J, Hutson T.E., McCann L., Deen K. et al. Overall Survival in Renal-Cell Carcinoma with Pazopanib versus Sunitinib N Engl J Med, 370 (18) (2014), pp. 1769-1770.

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 Fagerland MW, Lydersen S, Laake P. Recommended confidence intervals for two independent binomial proportions. Stat Methods Med Res 2011. Published Online First 13 Oct 2011.

10 Glimm E, Maurer W, Bretz F. Hierarchical testing of multiple endpoints in group-sequential trials. Statistics in Medicine. 2010;29:219-228.

11 Clopper, C.; Pearson, E. S. (1934). "The use of confidence or fiducial limits illustrated in the case of the binomial". *Biometrika* **26**: 404–413.

12 DerSimonian R, Laird N. Meta-analysis in clinical trials. Controlled Clinical Trials, 1986; 7: 177-188

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

14 Gail, M. and Simon, R. “Testing for qualitative interactions between treatment effects and subject subsets”. *Biometrics*, 1985; 41:361-372.

15 Fleming TR, Harrington DP 1991: Counting processes and survival analysis. New York: John Wiley and Sons.

16 Clopper, C.; Pearson, E. S. (1934). "The use of confidence or fiducial limits illustrated in the case of the binomial". *Biometrika* **26**: 404–413.

17 DerSimonian R, Laird N. Meta-analysis in clinical trials. Controlled Clinical Trials, 1986; 7: 177-188

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

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