Official Title of Study:

An Open-Label Exploratory Phase 2/3 Study of Nivolumab with Standard of Care Therapy vs Standard of Care Therapy for First-Line Treatment of Metastatic Colorectal Cancer (CheckMate 9X8: CHECKpoint pathway and nivoluMab clinical Trial Evaluation 9X8)

NCT Number: NCT03414983

Document Date (Date in which document was last revised): March 18, 2021

STATISTICAL ANALYSIS PLAN
FOR CLINICAL STUDY REPORT

AN OPEN-LABEL EXPLORATORY PHASE 2/3 STUDY OF NIVOLUMAB WITH STANDARD OF CARE THERAPY VS STANDARD OF CARE THERAPY FOR FIRST-LINE TREATMENT OF METASTATIC COLORECTAL CANCER

(CHECKMATE 9X8: CHECKPOINT PATHWAY AND NIVOLUMAB CLINICAL TRIAL EVALUATION 9X8)

PROTOCOL CA2099X8

VERSION # 4.0

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 BACKGROUND [redacted]

CA2099X8 (CHECKpoint pathway and nivoluMab clinical trial evaluation 9X8) is a Phase 2/3, randomized 2:1, open-label, multi-center trial to evaluate nivolumab (BMS-936558) in combination with standard of care (SOC) chemotherapy (mFOLFOX6) with bevacizumab for the treatment of first-line metastatic colorectal cancer (mCRC)[1]. [redacted]

Research Hypothesis:

Nivolumab plus SOC chemotherapy with bevacizumab (Nivo + SOC) will demonstrate superior efficacy (defined as progression-free survival [PFS]) compared to SOC chemotherapy with bevacizumab alone in participants with previously untreated mCRC. [redacted]

Schedule of Analyses:

In this study, one interim analysis (IA) and one final analysis (FA) are planned.

[redacted]

This document describes statistical analyses and outputs for CA2099X8 final analysis. At the time of this SAP update, one interim analysis was performed per previous version of SAP. Final analysis will be performed per this SAP [redacted]

2 STUDY DESCRIPTION

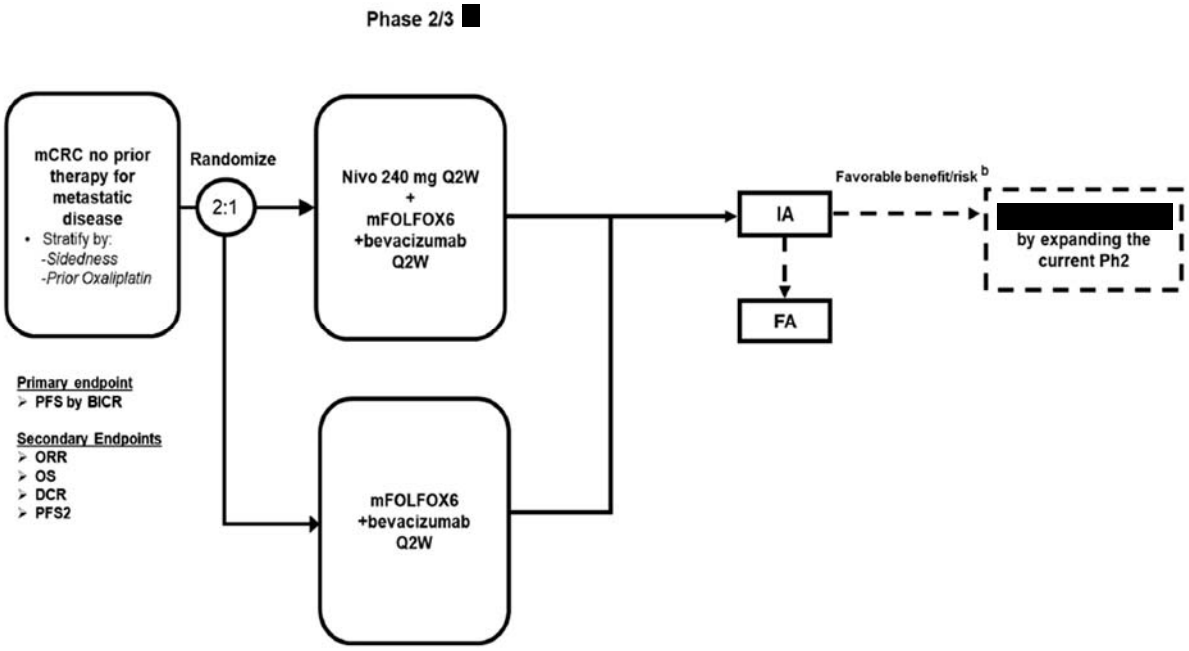
2.1 Study Design

This is a Phase 2/3, randomized, open-label study of nivolumab plus SOC, or SOC alone, in the first-line treatment of participants with metastatic CRC.

[redacted]

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

Figure 2.1-1: Study Design Schematic



DCR, disease control rate; FA; final analysis; IA, interim analysis; ORR, objective response rate; OS, overall survival; PFS, progression-free survival.

[redacted]

[b] Depending on the results at the time of the IA, different possible decisions may be made as indicated by dashed arrow from IA. If there is not enough evidence of favorable benefit/risk at IA, the Phase 2 study will continue to compete with 180 participants unless unacceptable safety concern arises.

One IA is planned in this Phase 2 study [redacted].

Enrollment and randomization will continue during the planned IA. The IA will only be based on the planned number of randomized participants. The study team/statistician will remain blinded to the additional participants who are not included in the IA. As such, when the study expends to a phase III cohort, these additional randomized participants will be included in the phase III cohort.

[redacted]

If the study does not expand at IA, enrollment will continue as planned to approximately180 randomized participants.

2.2 Treatment Assignment

CA2099X8 is a randomized, open-label study. After the participant's initial eligibility is established and informed consent has been obtained, the participant must be enrolled into the study by using an Interactive Response Technology (IRT) to obtain the participant number. Every participant that signs the informed consent form must be assigned a participant number using IRT.

All participants will be centrally randomized using the IRT.

Participants meeting all eligibility criteria will be randomized in a 2:1 ratio to nivolumab + SOC or to SOC arm. Randomization will be stratified by tumor sidedness (Left, Right, Transverse, and Unknown) and prior oxaliplatin-based adjuvant chemotherapy (Yes, No). The exact procedures for using the IRT will be detailed in the IRT manual.

2.3 Blinding and Unblinding

The trial is open label however no analyses or summaries generated by randomized treatment arms or treated arms for all randomized participants will be produced during the conduct of the study by the Sponsor.

At time of planned interim analysis, only the planned number of randomized participants will be unblinded to the study team. Decision of expanding to Phase III part of the study or continuing enrollment of Phase II part of the study will based on the results from these unblinded Participants.

In addition, the Blinded Independent Review Committee (BICR) will perform the central imaging review without knowledge of the treatment group assignment. At time of planned interim analysis, BICR data from the planned number of randomized participants will be transferred to the Sponsor and included for interim analysis.

2.4 Protocol Amendments

Global amendments incorporated in the protocol are described in Table 2.4-1.

Table 2.4-1: Protocol Amendments

Document	Date of Issue	Summary of Change
Protocol Amendment 07	13-Nov-2020	[redacted]
Administrative Letter 05	28-Sep-2020	
Revised Protocol 06	02-Jun-2020	Major change: Clarifies study treatment duration for standard of care (SOC)

Table 2.4-1: Protocol Amendments

Document	Date of Issue	Summary of Change
		SOC may continue, after the maximum treatment duration of 24 months for nivolumab, until progression, unacceptable toxicity, withdrawal of consent, or the end of the study, whichever comes first.
Revised Protocol 05	19-Dec-2019	Removed interim analysis 2 (IA2).
Revised Protocol 04	07-Dec-2018	Clarified the infusion days for fluorouracil, updated Appendix 3, and aligned Appendix 4 with contraceptive guidance for nivolumab and other components of study treatment.
Revised Protocol 03	13-Sep-2018	Incorporates Administrative Letter 03, and reinstates exclusion criteria 2j.
Administrative Letter 03	11-Sep-2018	Adding Appendix 11 back into the document
Revised Protocol 02	06-Aug-2018	Incorporates Administrative Letter 02 and updates to be consistent with SmPC and label information. Revises prohibited and/or restricted treatments; laboratory tests, assessments, and other analyses; dose modification criteria; urinalysis; [redacted] sample collection requirements. [redacted] clarifies eligibility criteria. [redacted] Minor changes and corrections including revisions to reflect the most recent language for BMS studies.
Administrative Letter 02	23-Apr-2018	[redacted]
Revised Protocol 01	02-Feb-2018	Incorporates Administrative Letter 01, revisions to laboratory tests exclusion criteria, revisions to dose modification criteria, urinalysis [redacted] sample collection requirements. [redacted] Addition of a definition of DILI for participants with elevated liver function tests at baseline. Updated rationale for the 2-year treatment duration, in addition of a number of other minor changes and corrections.
Administrative Letter 01	07-Dec-2017	[redacted]
Original Protocol	13-Oct-2017	Not applicable

[redacted]

2.6 Blinded Independent Review Committee (BICR)

A BICR will be utilized in this study for determination of BICR-assessed endpoints. Details of BICR procedures will be specified in the BICR Charter.

3 OBJECTIVES

3.1 Primary

- To compare the efficacy of nivolumab plus standard of care (SOC) chemotherapy plus bevacizumab (Nivo + SOC) with SOC chemotherapy plus bevacizumab in participants with mCRC

3.2 Secondary

- To evaluate preliminary efficacy in all randomized participants
- To characterize the safety and tolerability of nivolumab in combination with standard therapy with bevacizumab

4 ENDPOINTS

4.1 Primary Endpoint

The primary endpoint is progression-free survival (PFS) that is assessed by BICR in All Randomized participants.

PFS is defined as the time from randomization to the date of the first documented progression, as determined by BICR (per RECIST 1.1), or death due to any cause, whichever occurs first.

- Participants who die without a reported prior progression per BICR (and die without start of subsequent therapy) will be considered to have progressed on the date of death.
- Participants who did not have documented progression per BICR and who did not die, will be censored at the date of the last tumor assessment on or prior to initiation of the subsequent anti-cancer therapy, if any.
- Participants who did not have baseline tumor assessment will be censored on the date they were randomized

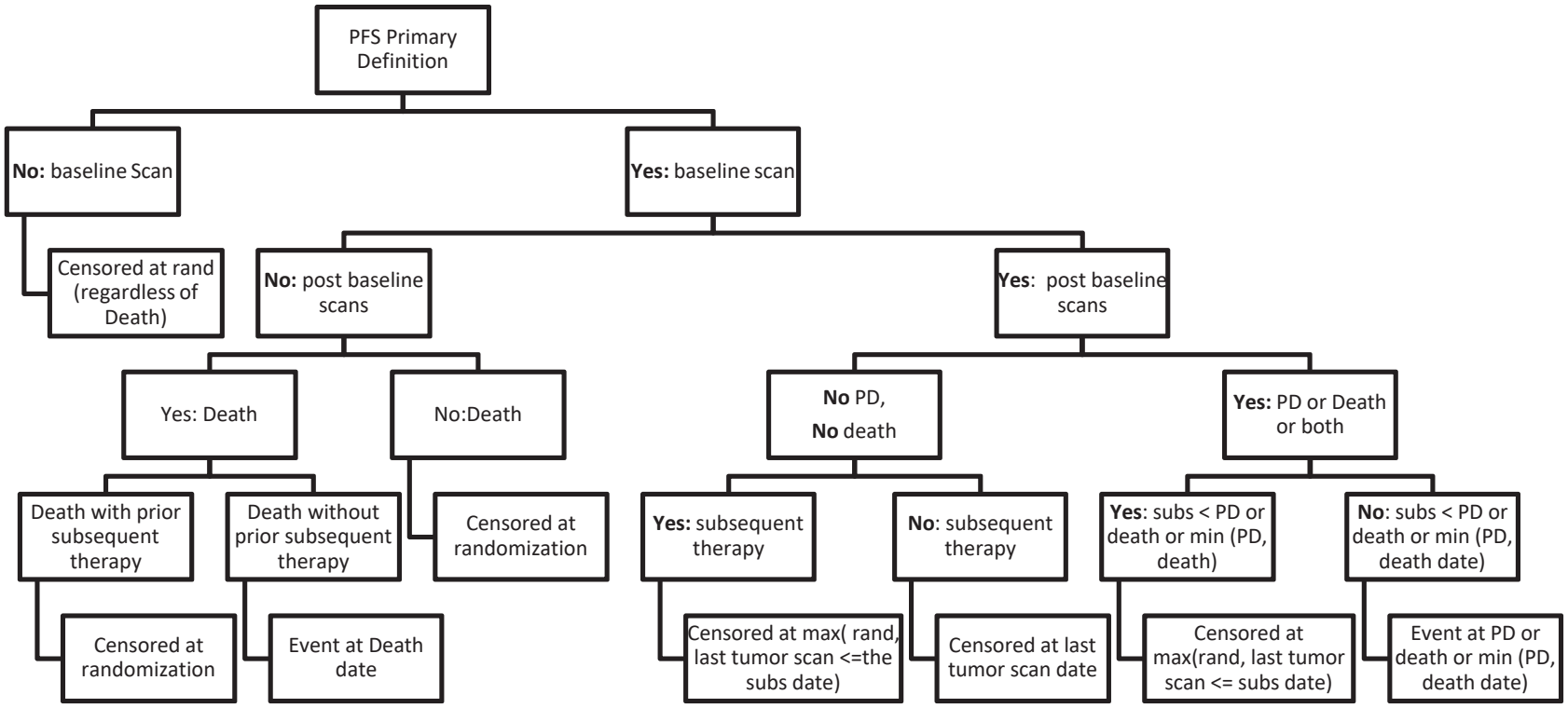
- Participants who did not have any post baseline tumor assessments and did not die (or died after initiation of the subsequent anti-cancer therapy) will be censored at the randomization date.
- Participants who started any subsequent anti-cancer therapy without a prior reported progression per BICR will be censored at the last tumor assessment on or prior to initiation of the subsequent anti-cancer therapy.
- Baseline tumor assessment is defined as tumor scans prior to or on randomization date

Censoring rules for the primary analysis of PFS are summarized in Table 4.1-1. Further explanation for various censoring scenarios are presented in Figure 4.1-1.

Table 4.1-1: Censoring Scheme Used in Primary Analysis of PFS

Situation	Date of Progression or Censoring	Outcome
No baseline tumor assessment	Date of randomization	Censored
No post baseline tumor assessments and no death	Date of randomization	Censored
Documented progression	Date of the first documented tumor progression	Progressed
No progression and no death	Date of last tumor assessment	Censored
Subsequent anti-cancer therapies	Date of last tumor assessment prior to initiation of the subsequent therapy	Censored
Death without progression and without initiation of subsequent anti-cancer therapy	Date of death	Progressed

Figure 4.1-1: Graphical Representation of the Primary PFS Definition



rand = randomization, subs = subsequent therapy

4.2 Secondary Endpoints

4.2.1 Progression-Free Survival (PFS)

Progression-free survival by investigator assessment in All Randomized participants is defined the same way as for the primary endpoint (per BICR), except that only tumor assessments by investigator will be taken into account.

4.2.2 Objective Response Rate (ORR)

ORR as assessed by BICR per RECIST 1.1 in All Randomized participants is defined as the number of participants with a best overall response (BOR) of confirmed CR or PR divided by the number of randomized participants in a population of interest. BOR is defined as the best response designation as determined by BICR per RECIST 1.1, recorded between the date of randomization and the date of objectively documented progression per RECIST 1.1 or the date of subsequent anti-cancer therapy (including tumor-directed radiotherapy and tumor-directed surgery), whichever occurs first. For participants without documented progression or subsequent anti-cancer therapy, all available response designations will contribute to the BOR determination.

For a BOR of CR or PR, the initial response assessment must be confirmed by a consecutive assessment no less than 28 days later. In the case of stable disease (SD), measurements must have met the SD criteria at least 42 days after randomization.

ORR as assessed by investigator in All Randomized participants is defined similarly as the endpoint as assessed by BICR, except that only tumor assessments by investigator will be taken into account.

In this study, the case report form (CRF) page of BOR will be filled under the following conditions when investigators review the tumor scans: 1) a participant has disease progression; 2) a participant has complete response; 3) a participant is determined to be not evaluable or off study (for example, a participant died before first on-study assessment). Therefore, at the database lock of interim analysis, participants who are on treatment or who have response of PR or SD may not qualify to complete the BOR CRF page. As such, the following imputation algorithm will be used to determine the BOR for participants who have not been ready to complete BOR page in order to conduct ORR analysis per investigator assessment.

- If a participant has PR date (from confirmed PR) reported in the CRF page of investigator best overall response date (RECIST 1.1 response criteria) and the BOR page is not filled yet, the participant's investigator BOR will be imputed as PR. This reflects the situation where a PR has been determined for the participant but the PR may further become CR at later time points, and thus only the PR date is recorded but the BOR page is not filled.
- If a participant does not have a BOR page but has at least one on-study tumor assessment that occurs beyond 42 days after randomization, the participant's BOR will be imputed as SD.
- Otherwise, the participant's investigator BOR will remain missing and not be imputed. This reflects a situation where a participant recently randomizes to the study and has not yet reached the timing of first scheduled tumor scan.

4.2.3 Duration of Response (DoR)

Duration of Response (as assessed by BICR and as assessed by investigator) is defined as the time between the date of first confirmed response (CR or PR) to the date of the first documented progression (per RECIST 1.1) or death due to any cause, whichever occurs first. The rules of censoring are the same as PFS. DoR will be evaluated only for participants with confirmed CR or PR.

4.2.4 Disease Control Rate (DCR)

DCR is defined as the proportion of participants whose BOR is confirmed CR or confirmed PR or SD among All Randomized participants of a population of interest based on BICR-assessed BOR or investigator-assessed BOR.

4.2.5 Time to Response (TTR)

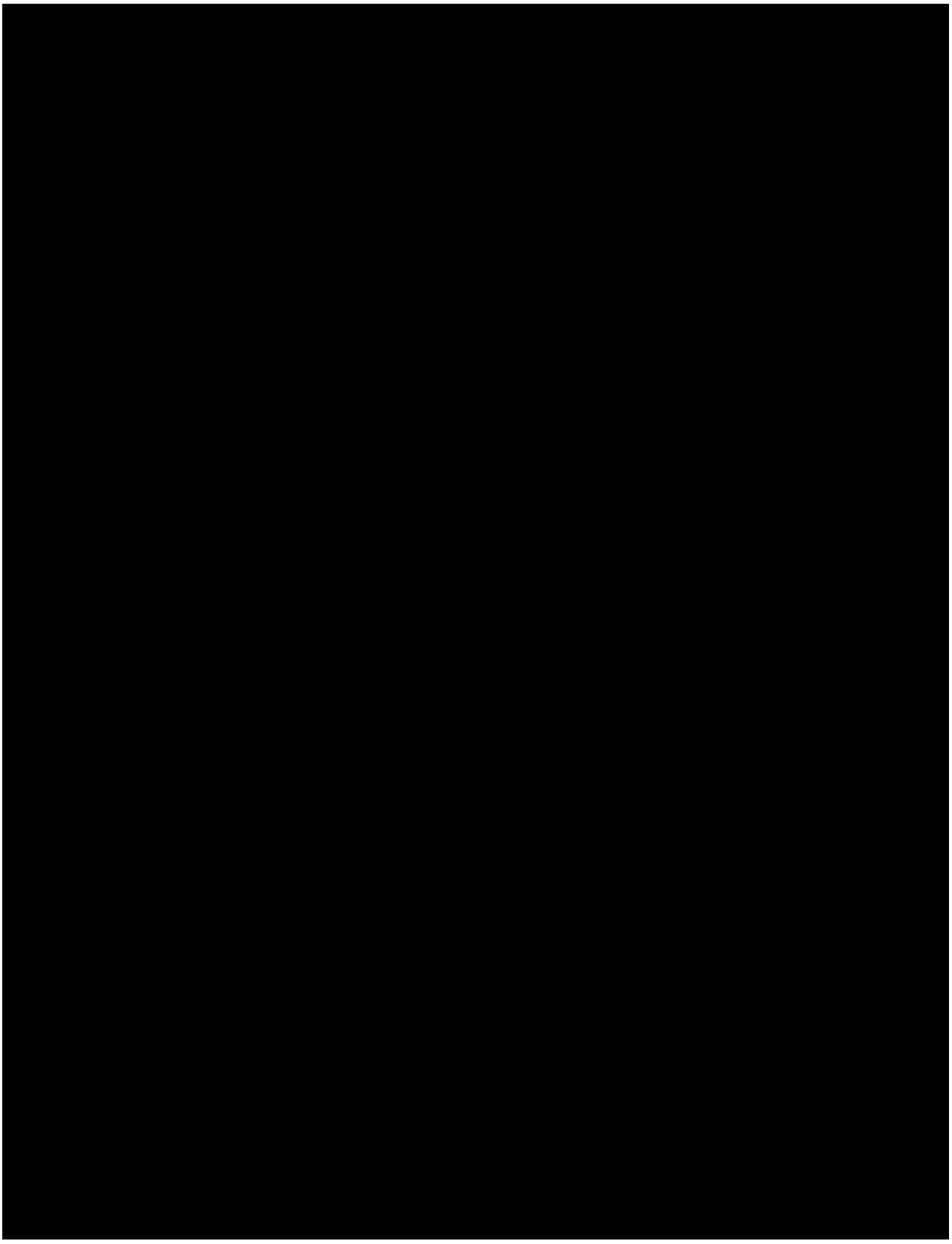
TTR is defined as the time from the randomization date to the date of the first confirmed CR or PR (as assessed by BICR and as assessed by investigator). TTR is derived for responders only.

4.2.6 Overall Survival (OS)

OS is defined as the time between the date of randomization and the date of death. For participants without documentation of death, OS will be censored on the last date the participant was known to be alive.

4.2.7 Safety Endpoints

Safety and tolerability objective will be measured by the incidence of adverse events (AEs), serious adverse events (SAEs), adverse events leading to discontinuation, adverse events leading to dose delays, immune mediated adverse events, select 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. See details in the Core Safety SAP.



[redacted]

[redacted]

[redacted]

[redacted]

6 STUDY PERIODS, TREATMENT REGIMENS AND POPULATIONS FOR ANALYSES

6.1 Study Periods

Refer to Core Safety SAP[13] for definitions of baseline period and post baseline period for AEs and evaluations of laboratory tests, and vital signs.

6.2 Treatment Regimens

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

- Arm A in the IRT system: Experimental arm: Nivo 240mg + mFOLFOX6 + Bevacizumab all Q2W
- Arm B in the IRT system: Control arm: mFOLFOX6 + Bevacizumab both Q2W

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

6.3 Populations for Analyses

All analyses will be performed using the treatment group as randomized (intent-to-treat) with the exception of extent of exposure (dosing) and safety, for which the treatment group as treated will be used.

When appropriate, the following definitions of different types of participants will be applied to different populations of interest defined as below:

- **Enrolled:** Participants who signed an informed consent form and were registered into the IRT system.
- **Randomized**: All participants who were randomized to any treatment group in the study. This is the primary population for analyses of efficacy.
- **Treated:** All participants who received at least one dose of study treatment (nivolumab or SOC). This is the primary population for analyses of exposure and safety.
- **Response Evaluable:** Randomized participants with 1) a BOR of CR, PR, SD, or PD; 2) target lesion(s) assessed at baseline and 3) at least one post-baseline time point with all baseline target lesion(s) assessed

[redacted]

[redacted]

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 participants falling into each category, grouped by treatment (with total, as needed). 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, as needed) using the mean, standard deviation, median, minimum and maximum values.

7.1.1 General Methods: Efficacy

Stratification Factors

For all stratified analyses, stratification factors will be based on the randomization factors. Specifically, stratification factors used for analyses are as follows:

- for All Randomized Participants
  - tumor sidedness (left, right, transverse, and unknown)
  - prior oxaliplatin-based adjuvant chemotherapy (yes vs. no)

Unless otherwise specified, stratification factors are used as entered in the IRT.

General Methods for Time-to-event Variables

Time to event distribution (e.g. OS, PFS, DoR) will be estimated using Kaplan-Meier (KM) method.

Median along with 95% confidence interval (CI) will be constructed based on a log-log transformed CI for the survivor function S(t)[14],[15]. Rates at fixed timepoints (e.g. OS at 6 months) will be derived from the KM estimate and corresponding CI will be derived based on Greenwood formula[16] for variance derivation and on log-log transformation applied on the survivor function S(t)[17].

Unless otherwise specified, the stratified log-rank test (for primary endpoint of PFS) will be performed to test the comparison of time to event distributions between two treatment groups. The stratified hazard ratio (HR) between two treatment groups along with CI will be obtained by fitting a stratified Cox model (using randomization stratification factors specified above) with the treatment group variable as unique covariate.

General Methods for binomial proportion variables

Confidence intervals for binomial proportions will be derived using the Clopper-Pearson method[18]. The difference in rates (e.g., ORR) between the two treatment groups along with their two-sided 95% CI will be estimated using Cochran-Mantel-Haenszel (CMH) method of weighting[19], adjusting for the stratification factors.

7.2 Study Conduct

7.2.1 Accrual

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

7.2.2 Relevant Protocol Deviations

The relevant Protocol Deviations will be summarized for all randomized participants, by treatment group and overall. The following programmable deviations will be considered as relevant protocol deviations. Non-programmable relevant eligibility and on-treatment protocol deviations, as well as significant (both programmable and nonprogrammable) eligibility and on-treatment protocol deviations will be reported via other approach.

At Entrance

- Participants who have received prior chemotherapy for metastatic colorectal cancer
  - For participants who have received adjuvant or neo-adjuvant chemotherapy, there is no more than 6 months gap (≤ 6 months) between the completion of that therapy and diagnosis of recurrent or metastatic disease.
- Participants with baseline ECOG performance status > 1.
- Participants without any measurable disease per investigator by RECIST 1.1 criteria at baseline.

On-study

- Participants treated differently as randomized (participants who received the wrong treatment consistently, excluding the never treated).

A participant listing will also be produced.

7.3 Study Population

7.3.1 Subject Disposition

The total number of participants enrolled (randomized or not randomized) will be presented along with the reason for not being randomized.

Number of participants randomized but not treated along with the reason will be tabulated by treatment group as randomized. This analysis will be performed on all randomized participants.

Number of participants who discontinued study treatment along with corresponding reason will be tabulated by treatment group as treated. This analysis will be performed on all treated participants.

A participant listing for all treated participants will be provided showing the participant's randomization date, first and last dosing date, off treatment date and reason for going off treatment.

A participant listing for all enrolled participants will also be provided, showing the participant's race, gender, age, consent date and reason for not being randomized.

7.3.2 Demographics and Other Baseline Characteristics

Descriptive statistics will be summarized for the following baseline characteristics for all randomized participants by treatment as randomized. All baseline presentations will identify participants with missing measurements. Listings will also be provided.

7.3.2.1 Demographics

- Age (descriptive statistics)
- Age categories (<65, ≥ 65 and <75, ≥ 75 years)
- Gender (male, female)
- Race/Ethnicity
- Region (US/Canada vs. Europe vs. Rest of World)

- Baseline ECOG Performance Status (0, 1, >1, Not reported)
- Weight (descriptive statistics)

7.3.2.2 CRC disease Characteristics

- Tumor sidedness (Left, Right, Transverse, and Unknown) (*stratification factor*)
- Prior oxaliplatin-based chemotherapy (Yes, No) (*stratification factor*)
- Time from initial disease Diagnosis to Randomization (< 3month, 3month-< 6months, 6months -< 1 year, 1-< 2 years, 2 - < 3 years, 3 -< 4 years, 4 - < 5 years, ≥ 5 years)
- Time from Initial Metastatic Disease Diagnosis to Randomization (< 3month, 3month-< 6months, 6months -< 1 year, 1-< 2 years, 2 - < 3 years, 3 -< 4 years, 4 - < 5 years, ≥ 5 years)
- Disease diagnosis: initial diagnosis (Stage, classification, Location, cell type), initial metastatic (stage, location) and study entry (stage)
- TNM classification at initial diagnosis:
  - Tumor (Tx, T0, T1, T2, T3, T4a, T4b)
  - Nodes (Nx, N0, N1a, N1b, N1c, N2a, N2b)
  - Metastasis (M0, M1a, M1b, M1c)

[redacted]

- All lesions (Investigator and BICR Tumor Assessments at Baseline): organ of lesion,
- Target Lesions (Investigator and BICR Tumor Assessments at Baseline): Presence of target lesions, organ of target lesion, summary of diameters of target lesions

7.3.2.3 Other Baseline Characteristics

- Smoking Status
- Alcohol use

7.3.3 Medical History

- General medical history will be listed by participant and pretreatment events will be tabulated.

7.3.4 Prior Cancer Therapy

- Prior anti-cancer therapy will be summarized:
- Prior surgery related to current cancer (yes or no).
- Prior radiotherapy (yes or no).

- Prior systemic therapy (setting, Time from completion of prior adjuvant/neo-adjuvant therapy to randomization(< 6 months, 6 - < 12month, ≥ 12 months), etc)
- Prior/current non-study medication classified by anatomic and therapeutic classes.

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

7.3.5 Baseline Examinations

Participants with abnormal baseline physical examination will be tabulated by examination criteria and by treatment as randomized for all randomized participants.

7.3.5.1 Discrepancies Between IRT and CRF Stratification Factors

A summary table (cross-tabulation) by treatment group as randomized for stratification factor will be provided to show any discrepancies between what was reported through IVRS and CRF data.

7.4 Extent of Exposure

Analyses in this section will be performed in all treated participants by treatment group as treated. Listing will include all available exposure data.

7.4.1 Administration of Study Therapy

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

- Relative dose intensity (%) using the following categories: < 50%; 50 - < 70%; 70 - < 90%; 90 - < 110%; ≥ 110%.
- Number of doses received (summary statistics)
- Cumulative dose
- Duration of treatment: duration of treatment will be presented using a Kaplan-Meier curve whereby the last dose date will be the event date for participants who discontinued study therapy. Participants who are still on study therapy will be censored on their last dose date. Median duration of treatment and associated 95% CI will be provided.
- A by-participant listing of dosing of study medication (record of study medication, infusion details, and dose change) and a listing of batch number will be also provided.
- Time from randomization to first dose date (<= 3 days, 4 - 5 days, 6 - 7 days, 8 - 14 days, 15 - 21 days, >21 days)

The selection and timing of dose for each participant is as follows (see details in Protocol Section 7.1):

Table 7.4.1-1: Selection and Timing of Dose

Study Treatment	Unit dose strength(s)/Dosage level(s)	Dosage formulation Frequency of Administration	Route of Administration
Oxaliplatin	85 mg/m^2	Q2W	IV
Leucovorin	400 mg/m^2 or 350 mg/m^2 as per local standards	Q2W	IV
Fluorouracil	400 mg/m^2	Q2W	IV
Fluorouracil	1200 mg/m^2 Continuous infusion Day #1 and Day #2 OR 2400 mg/m^2 Continuous infusion over 46 hours Day #1 through Day #2 as per local standards	Q2W	IV
Bevacizumab	5 mg/kg	Q2W	IV
Nivolumab	240 mg	Q2W	IV

IV, intravenous.

Table 7.4.1-2: Administration of Study Therapy - Definition of Parameters

Parameter	Definition per study medication
	Nivolumab
Dosing schedule per protocol	240mg every 2 weeks
Dose	Dose (mg) is defined as the total drug administrated at each dosing date as collected on the CRF
Cumulative Dose	The sum of all doses (mg) administered to a participant during the treatment period
Relative dose intensity (%)	Cum dose (mg) /[(Last dose date - Start dose date + 14) x 240 /14] x 100
Duration of treatment	Last dose date - Start dose date +1
	Bevacizumab
Dosing schedule per protocol	5 mg/kg every 2 weeks
Dose	Dose (mg/kg) is defined as total dose administered at each dosing date (mg) divided by the most recent weight (kg). Dose administered and weight are collected from the CRF.
Cumulative Dose	Cum Dose (mg/kg) is the sum of the doses administered to a participant during the treatment period.
Relative dose intensity (%)	Cum dose (mg/kg) /[(Last dose date - Start dose date +14) x 5/14] x 100
Duration of treatment	Last dose date - Start dose date +1

Table 7.4.1-2: Administration of Study Therapy - Definition of Parameters

Parameter	Definition per study medication
	Oxaliplatin/ Leucovorin/ Fluorouracil
Dosing schedule per protocol	Every 2 weeks
Dose[a]	Dose (mg/m2) is defined as the total dose administered (mg) / most recent BSA
Cumulative Dose	The sum of all doses (mg/m2) administered to a participant during the treatment period
Relative dose intensity (%)	Cumulative dose (mg/m2) / [Last dose date –Start dose date + 14) x ds / 14] ds is dose level for Oxaliplatin or Leucovorin or Fluorouracil, respectively as in Table 7.4.1-1
Duration of treatment	Last dose date - Start dose date +1

[a] * Dose administered in mg at each dosing date and BSA (computed using most recent weight and baseline height) are collected on the CRF.

7.4.2 Modifications of Study Therapy

7.4.2.1 Dose delays

Treatment may be delayed for up to a maximum of 6 weeks from the last dose without approval from BMS clinical trial physician. A dose will be considered as actually delayed if the delay is exceeding 3 days (i.e., greater than or equal to 4 days from scheduled dosing date). Length of delay is defined as (duration of interval-from prior dose in days - 14). Dose delays will be divided into following categories: 4 - < 8 days, 8 - < 15 days, 15 - < 42 days, ≥ 42 days. Reason for dose delay will be retrieved from CRF dosing pages.

The following parameters will be summarized for Nivolumab, Bevacizumab, Oxaliplatin, Leucovorin, and Fluorouracil:

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

7.4.2.2 Infusion Interruptions and Rate Changes

Each study therapy infusion may be interrupted. This information will be retrieved from CRF dosing pages. The following parameters will be summarized for Nivolumab, Bevacizumab, Oxaliplatin, Leucovorin, and Fluorouracil:

- Number of participants with at least one dose infusion interrupted along with reason (per interruption) for the interruptions and number of infusions interrupted per participant
- Number of participants with at least one IV infusion rate reduction, number of IV infusion rate reduction per participant, and the reason for reduction

7.4.2.3 Dose Reductions

Per protocol, there will be no dose escalations or reductions of nivolumab. However, for bevacizumab and chemotherapy, dose reduction is permitted (as specified in the protocol). This information will be retrieved from CRF dosing pages.

For participants treated with chemotherapy, the following will be summarized:

- Number of participants with at least one dose reduction along with the reason of the dose reduction.

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 (participants with any concomitant medication, participants by medication class and generic term).

A by-participant listing will accompany the table.

7.4.4 Subsequent Therapy

- Number and percentage of participants receiving subsequent therapies will be summarized. Categories include:
- Immunotherapy including commercial Nivolumab (anti-PD1 agents, anti-PD-L1 agents, anti-CTLA-4 agents and others) by drug name
- Other anti-cancer agents excluding all immunotherapy (approved and investigational) by drug name
- Surgery for treatment of tumors
- Radiotherapy for treatment of tumors

A participant listing of follow-up therapy will also be produced for participants who had any subsequent therapy.

7.5 Efficacy

Unless otherwise specified, analyses in this section will be performed in all randomized participants by treatment group as randomized.

7.5.1 Primary Efficacy Endpoint

7.5.1.1 Primary Analysis

The distribution of PFS as assessed by BICR will be compared between treatment groups using a two-sided log rank test, stratified by the stratification factors as specified by

- Tumor sidedness (left, right, transverse, and unknown)
- Prior oxaliplatin-based adjuvant chemotherapy (yes vs. no)

A two-sided significance level of 0.2 will be used.

HR with its associated two-sided 80% and 95% CIs will be estimated via a stratified Cox model (using the above randomization stratification factors) with treatment group as the only covariate in the model.

PFS curves for each treatment group will be estimated and plotted using the KM product-limit method. Median PFS along with two-sided 95% CI will be constructed based on a log-log transformed CI for the survival function[14,15]. In addition, PFS rates at a specified time point (e.g., 6, 9, 12, 15, 18, 24 months), and the corresponding two-sided 95% CIs using the log-log transformation, will be computed. Minimum follow-up (see definition in Section 7.5.2.4) must be longer than a time period to generate a rate.

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

1) Censored on randomization date
2) Censored on date of last disease assessment on-study

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

7.5.1.2 Sensitivity Analyses

Progression-Free Survival

- *Analysis using a 2-sided, un-stratified log-rank tes*t, an un-stratified Cox proportional hazards model with treatment as the single covariate
- *Analysis for participants with no relevant deviation.* This analysis will be conducted only if there are more than 10% participants with relevant protocol deviations.

The above analysis will use the same significance level as the corresponding primary analysis. Estimate of the HR, its two-sided 80% and 95% CI and descriptive p-value will be presented.

Sensitivity analyses for PFS to investigate alternative censoring schemes

- *Stratified Analysis (using stratification factors per IRT) accounting for assessment on/after subsequent therapy:* PFS will be defined similarly to the primary definition except that events (progression or death) and disease assessments that occurred on or after subsequent anti-cancer

therapy will be considered (no time point truncation) (see censoring scheme in Table 7.5.1.2-1).

The above analysis will use the same significance level, stratified by stratification factors as the corresponding primary analysis. Estimate of the HR, its two-sided 80% and 95% CI and p-value will be presented.

Table 7.5.1.2-1: Censoring Scheme - PFS Sensitivity Analysis

Situation	Date of Progression or Censoring	Outcome
No baseline tumor assessment	Date of randomization	Censored
No post baseline tumor assessments and no death	Date of randomization	Censored
Documented progression per BICR assessment according to RECIST 1.1	Date of the first documented tumor progression	Progressed
No progression and no death	Date of last tumor assessment	Censored
Subsequent anti-cancer therapies started without a prior reported progression	Date of subsequent anti-cancer therapy not considered	None
Death without progression and without initiation of subsequent anti-cancer therapy	Date of death	Progressed
Death without progression after initiation of subsequent anti-cancer therapy	Date of death	Progressed

7.5.2 Secondary Efficacy Endpoints

7.5.2.1 Objective Response Rate and Related Endpoints

ORR (as assessed by BICR and investigator) will be estimated in each treatment group along with the exact 95% CI using Clopper-Pearson method. An estimate of the difference in ORRs and corresponding 95% CI will be calculated using CMH methodology and adjusted by the stratification factors as specified in Section 7.1.1. The stratified odds ratios (Mantel-Haenszel estimator) between the treatments will be provided along with the corresponding two-sided 95% CI.

Best Overall Response (as assessed by BICR and by investigator) will be summarized by response category for each treatment group.

Duration of Response (as assessed by BICR and by investigator) in each treatment group will be estimated using KM product-limit method for participants who achieve confirmed PR or CR. Median values along with two-sided 95% CI will be calculated.

Time to objective response (as assessed by BICR and by investigator): summary statistics will be provided for each treatment group for participants who achieve PR or CR.

Disease control rate: an estimate and corresponding two-sided 95% exact CI will be provided using Clopper-Pearson method.

The following participant-level graphics will also be provided:

- For the responders only, time courses of the following events of interest will be graphically displayed: first tumor response, tumor progression, last dose received, and death (a.k.a. swimming plot).
- For response evaluable participants, a waterfall plot showing the best reduction from baseline in target lesion based on will be produced (excluding assessments after PD and assessments after start of subsequent anti-cancer therapy).
- For response evaluable participants, spider plot of tumor target lesion percent change from baseline based on will be produced.

7.5.2.2 Sensitivity Analyses for ORR

As sensitivity analysis, a summary of BICR-assessed ORR based on response-evaluable participants will be presented. In addition, for interim analyses, a summary of both BICR-assessed and investigator-provided ORR will be provided on response-evaluable participants among the interim analysis participants.

7.5.2.3 PFS and OS

PFS (investigator assessments per RECIST v1.1) and OS will be summarized descriptively for randomized participants using the KM product-limit method. Median values of PFS and OS, along with two-sided 95% CI (based on the log-log transformation), will also be calculated.

KM curve of PFS and OS will be generated. PFS and OS rates at specific time points (eg. 6 month) will be estimated using KM estimates. PFS and OS rates at additional time points (e.g.12, 18 and 24 months) may also be estimated depending on whether minimum follow-up will be longer than time point to generate the rate. Associated two-sided 95% CIs will be calculated using the Greenwood's formula[14,15].

Status of Censored Participants

The status of participants who are censored in the PFS (investigator assessment) Kaplan-Meier analysis will be tabulated using the same categories as PFS by BICR in Section 7.5.1.1.

The status of participants who are censored in the OS Kaplan-Meier analysis will be tabulated using following categories:

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

Participant listings of PFS and OS will be produced.

7.5.2.4 Participants Follow-up

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

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

The currentness of follow-up for OS is defined as the time between last OS contact (i.e., last known date alive or death date) and clinical cut-off date. Participants who died and participants with last known alive date after data cut-off date will be considered as current for this analysis.

The currentness of follow-up for PFS is defined as the time between last tumor assessment date (regardless of initiation of subsequent therapy) and clinical cut-off date. Participants who have a PFS event (regardless of initiation of subsequent therapy) and participants with last tumor assessment date on or after clinical cut-off will be considered as current for this analysis.

The currentness of follow-up for OS and PFS will be summarized by treatment group, and will be categorized into the following categories: 0 day, 1day - 3 months, 3-6 months, 6-9 months, 9-12 months, and > 12 months.

7.5.2.5 Consistency of Treatment Effect in Subsets

To assess consistency of treatment effects in different subsets, a "forest" plot of PFS/OS unstratified HR (and 95% CI) will be produced for the following subgroups.

- Age category (< 65, ≥ 65 and < 75, ≥ 75)
- Gender (male, female)
- Race (white, black, Asian, and other)
- Region(Us/Canada, Europe, Asia, Rest of world)
- ECOG (0, 1, >1 if exist)
- Tumor sidedness (Left, Right, Transverse, and Unknown) (*stratification factor--IRT*)
- Disease stage at initial diagnosis (Stage I, Stage II, Stage III, Stage IV)
- Disease site at initial diagnosis (Colon vs. Rectum)
- Disease classification at initial diagnosis (Gx, G1, G2, G3, G4)
- Disease cell type at initial diagnosis (adenocarcinoma, cribriform comedo-type adenocarcinoma, medullary, micropapillary, colloid carcinoma, serrated adenocarcinoma, signet ring cell)
- Time from Initial Disease Diagnosis to Randomization (< 1 year, 1 - < 3 years, ≥ 3 years)
- Time from Initial Metastatic Disease Diagnosis to Randomization (< 3month, 3month-<6monther, >=6month)
- Metastatic disease location(liver, lung, etc)
- Number of organs with metastases at baseline (≤ 1 vs. ≥ 2)

[redacted]

[redacted]

- Prior curative surgery (Yes, No)
- Prior radiotherapy (Yes, No)
- Prior adjuvant/neoadjuvant chemotherapy (Yes, No)
- Prior oxilaplatin use (Yes, No) (*stratification factor-IRT*)

If subset category has less than 10 participants per treatment group, HR will not be computed/displayed. Number of events and median PFS/OS along with 95% CI will be displayed for each treatment group.

7.5.2.6 ORR Subgroups

To assess consistency of ORR, ORR per BICR will be summarized for the subgroups [redacted] as defined in the Section 7.5.2.5. BICR-assessed ORR will be computed for each treatment as randomized along with the exact 95% CI using the Clopper-Pearson method.

The ORR subgroup analyses will also be presented for response-evaluable participants from the interim analysis participants for interim analyses.

[redacted]

7.6 Safety

In general, safety presentation will use the all treated population (presented as treated) and will follow the Core Safety SAP[13].

7.6.1 Deaths

See Core Safety SAP

7.6.2 Serious Adverse Events

See Core Safety SAP

7.6.3 Adverse Events Leading to Discontinuation of Study Therapy

See Core Safety SAP

7.6.4 Adverse Events Leading to Dose Modification

See Core Safety SAP

7.6.5 Adverse Events

See Core Safety SAP

7.6.6 Select Adverse Events

See Core Safety SAP

7.6.7 Immune Modulating Medication

See Core Safety SAP

7.6.8 Multiple Events

See Core Safety SAP

7.6.9 Other Events of Special Interest

See Core Safety SAP

7.6.10 Immune Mediated Adverse Events

See Core Safety SAP

7.6.11 Clinical Laboratory Evaluations

See Core Safety SAP

7.6.12 Vital Signs

See Core Safety SAP

[redacted]

7.6.14 Pregnancy

See Core Safety SAP

7.6.15 Adverse Events by Subgroup

See Core Safety SAP

[redacted]

[redacted]

[redacted]

[redacted]

7.11 Interim Analysis

No multiplicity control of type I error rate for planned IA (and no formal statistical test will be performed). Strong type I error control will apply to the phase III expansion cohort, if decision is made to expand to Phase III part of the study.

The following analyses will be performed for the IA (if applicable).

Study Conduct and Study Population

- relevant protocol deviations
- Demographic and baseline disease characteristics
- Disposition: frequency table of participants who discontinued treatment and reasons for discontinuation
- Disposition: frequency table of participants who were randomized but not treated and reasons for not treated
- Baseline Physical Measurements Summary
- Prior cancer therapy summary
- discrepancy in stratification factor between IRT and CRF

Extent of exposure

- Time from randomization to first dose of study therapy
- Summary of administration of study therapy
- Summary of dose delays for all study therapies
- Summary of infusion interruptions or rate reduction for all study therapies
- Summary of dose reductions for bevacizumab and chemotherapy
-

Efficacy Analyses:

- Summary of ORR (per BICR and investigator)
- Summary of ORR (per BICR and investigator) within subgroups (as defined in Section 7.5.2.5) by treatment groups as randomized.
- Summary of ORR (per BICR and investigator) on all response-evaluable participants
- Summary of ORR (per BICR and investigator) with subgroups on all response-evaluable participants
- TTR and DOR (per BICR and investigator)
- Spider plot
- Waterfall plot
- Swimmer plot

- Kaplan Meier plot of PFS (per BICR and investigator).
- Kaplan Meier plot of OS
- Summary of PFS (per BICR and investigator) and OS using the KM product-limit method. Median values of PFS and OS, along with two-sided 95% CI (based on the log-log transformation), PFS and OS rate at specific timepoint

Safety Analyses:

- Death summary
- Summary of Any Adverse Events by worst CTC Grade
- Summary of Drug-Related Adverse Events by worst CTC Grade
- Summary of Serious Adverse Events by worst CTC Grade
- Summary of Drug-Related Serious Adverse Events by worst CTC Grade
- Summary of Adverse Events leading to discontinuation by worst CTC Grade
- Summary of Drug-Related Adverse Events leading to discontinuation by worst CTC Grade
- Summary of Adverse Events Leading to dose delay and dose reduction
- Summary of Any Select Adverse Events by worst CTC Grade
- Summary of Drug-Related Select Adverse Events by worst CTC Grade and by Category
- summary of Immune-Mediated Adverse Event
- Summary of Laboratory Parameter Changes from Baseline Grade (US unit only)
- Summary of laboratory parameter worst grade (US unit only)
- Summary of Abnormal Liver Test
- Summary of abnormal Thyroid Test

[redacted]

8 CONTENT OF REPORTS

8.1 Interim Analysis

Please refer to Section 7.11.

[redacted]

9 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[20]. Missing and partial Non-Study Medication Domain

dates will be imputed using the derivation algorithm described in BMS Non-Study Medication Domain Requirements Specification[21].

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

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

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 case, 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 will 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 response, and time to response) will be calculated as follows:

Duration = (Last date - first date + 1)

[redacted]

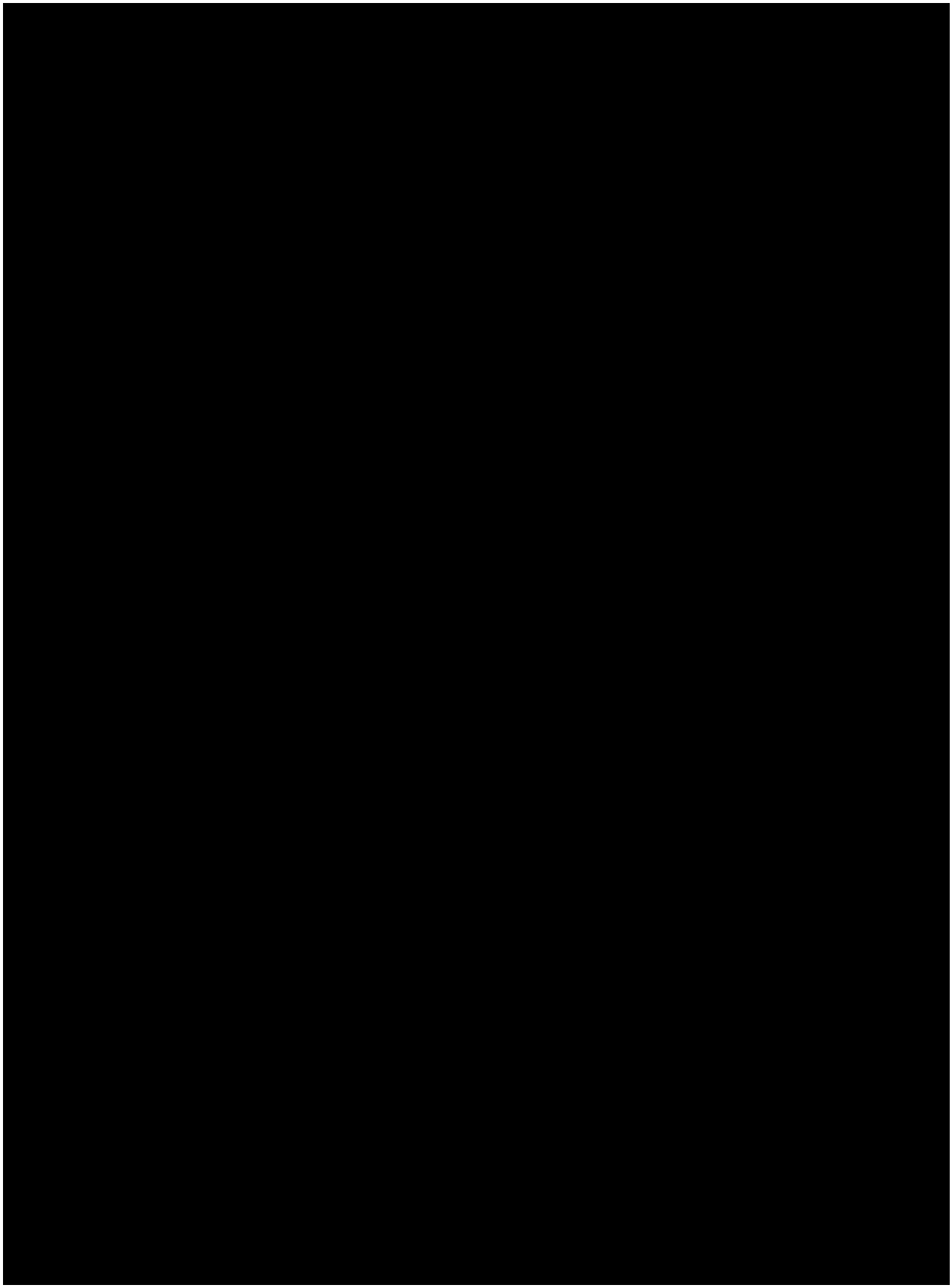
10 CONTENT OF REPORTS

All analyses described in the 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, figures and listings

11 DOCUMENT HISTORY

Table 11-1: Document History

Version Number	Author(s)	Description
1.0	[redacted]	Initial version 07-Nov-2018
2.0		Removed IA2 related text (per protocol revision 05); updated protocol amendment section.
2.1		Updated PFS2 censoring rule (section 4.3.3) to be consistent with program level PFS2 derivation
3.0		[redacted] • Improved resolution of Figure 2.1-1. • Updated protocol amendment section. [redacted] • Added Table 5.1-1. • Minor clarifications and updates in section 7.2.2, 7.3.5.1, 7.4.2.1 and 7.5.1.1.
4.0		[redacted] • Added Section 8 to clarify analyses conducted prior to this revision.



Bristol Myers Squibb™

Signatures

Document: CA2099X8 CSR Statistical Analysis Plan Version 4

Event	User Name	User ID	Reason	Date (ET)	[redacted]
Task Completed (Approve (eSignature)): I approve this document.	[redacted]	[redacted]	[redacted]	12-Mar-2021 14:25:34	[redacted]
Task Completed (Approve (eSignature)): I approve this document.	[redacted]	[redacted]	[redacted]	12-Mar-2021 14:44:08	[redacted]
Task Completed (Approve (eSignature)): I approve this document.	[redacted]	[redacted]	[redacted]	15-Mar-2021 06:26:35	[redacted]