**STATISTICAL ANALYSIS PLAN**

# An Open-label Phase 1b/2 Study of Binimetinib Administered in Combination with Nivolumab or Nivolumab Plus Ipilimumab in Patients with Previously Treated Microsatellite-stable (MSS) Metastatic Colorectal Cancer with RAS Mutation

| | |
|---|---|
| STUDY DRUG: | Binimetinib + Nivolumab + Ipilimumab |
| PROTOCOL NUMBER: | C4211004 (ARRAY-162-202) |
| VERSION: | 1.0 |
| DATE: | 09 October 2020 |

**Prepared by:**

PPD
PPD
PPD

**CONFIDENTIAL**

**The information contained in this document is regarded as confidential and may not be disclosed to another party without the permission of the author unless such disclosure is required by law or regulations. Persons to whom the information is disclosed must be informed that the information is confidential and may not be further disclosed by them.**

# Approval Signatures

| | |
|---|---|
| Date | Prepared by: PPD<br>PPD<br>PPD |
| Date | Reviewed by: PPD<br>PPD Biostatistics<br>PPD |
| Date | Approved by: PPD<br>PPD Biostatistics<br>Pfizer Inc. |

# TABLE OF CONTENTS

## LIST OF TABLES



## LIST OF FIGURES

## List of Abbreviations

| | |
|---|---|
| AE | Adverse event |
| ALP | Alkaline phosphatase |
| ALT | Alanine aminotransferase |
| aPTT | Activated partial thromboplastin time |
| AST | Aspartate aminotransferase |
| ATC | Anatomical therapeutic classification |
| BID | Twice daily |
| BMI | Body mass index |
| BOR | Best overall response |
| BNP | Brain natriuretic peptide |
| BUN | Blood urea nitrogen |
| C1D1 | Cycle 1 Day 1 |
| CA 19-9 | Cancer antigen 19-9 |
| CEA | Carcinoembryonic antigen |
| CK | Creatine kinase |
| Cmin | Minimum plasma concentration |
| CR | Complete response |
| CRC | Colorectal cancer |
| CRP | C-reactive protein |
| CSR | Clinical study report |
| CTCAE | Common terminology criteria for adverse events |
| CTLA-4 | Cytotoxic T-cell lymphoma-4 antigen |
| CV | Coefficient of variation |
| DCR | Disease control rate |
| DDS | Dose-determining set |
| DLT | Dose limiting toxicity |
| DOR | Duration of response |
| ECG | Electrocardiogram |
| ECHO | Echocardiogram |
| eCRF | Electronic case report form |
| ECOG | Eastern Cooperative Oncology Group |
| FAS | Full analysis set |
| FSH | Follicle-stimulating hormone |
| HIV | Human immunodeficiency virus |

| | |
|---|---|
| INR | International normalized ratio |
| IHC | Immunohistochemistry |
| CCI | CCI |
| LDH | Lactate dehydrogenase |
| LH | Luteinizing hormone |
| LLOQ | Lower Limit of quantification |
| LVEF | Left ventricular ejection fraction |
| mCRC | Metastatic colorectal cancer |
| MedDRA | Medical Dictionary for Regulatory Activities |
| MEK | Mitogen/extracellular signal regulated kinase |
| MSS | Microsatellite stable |
| MTD | Maximum tolerated dose |
| mTPI | Modified toxicity probability interval |
| MUGA | Multi-gated acquisition |
| NCI | National cancer institute |
| ORR | Objective response rate (overall response rate) |
| OS | Overall survival |
| PAS | Pharmacokinetic Analysis Set |
| PD | Progressive disease |
| PD-1 | Programmed death-1 |
| PD-L1 | Programmed death-ligand 1 |
| PD-L2 | Programmed cell death 1 ligand 2 |
| PFS | Progression-free survival |
| pH | Hydrogen ion concentration |
| PK | Pharmacokinetic |
| PR | Partial response |
| PT | Prothrombin time |
| Q4W | Every four weeks |
| Q8W | Every eight weeks |
| QTc | Corrected QT interval |
| QTcF | QT interval corrected for heart rate using Fredericia's formula |
| Racc_cmin | Accumulation ratio based on Cmin |
| RECIST | Response Evaluation Criteria In Solid Tumors |
| RP2D | Recommended phase 2 dose |
| SAE | Serious adverse event |
| SAP | Statistical Analysis Plan |
| SD | Stable disease |

| | |
|---|---|
| T3 | Triiodothyronine |
| T4 | Thyroxine |
| TBL | Total bilirubin |
| TSH | Thyroid-stimulating hormone |
| WBC | White blood cell(s) |
| WHO DRL | World Health Organization Drug Reference Listing |
| WHO ATC | World Health Organization Anatomical Therapeutic Chemical |

# 1. INTRODUCTION

This document provides the detailed statistical methodology for the analysis of data from study ARRAY-162-202 that will be presented in the Clinical Study Report (CSR). All changes to the planned analysis required before or after database lock will be made through an amendment or an addendum, respectively. Note that obvious corrections will be made at the time of analysis to address minor formatting or spelling mistakes without the need to amend these modules. The table, listing, and figure shells of the statistical analysis plan (SAP) can be found in a separate SAP shell document.

All changes to the planned analyses described in this document will be made prior to the database lock.

The analyses described herein are based on protocol version 4.0, dated 02MAR2018.

# 2. STUDY OBJECTIVES AND DESIGN

## 2.1. Study Objectives and Endpoints

The study objectives and corresponding endpoints are shown in the table below.

## Table 1. Objectives and Related Endpoints

| Objective | Endpoint |
|---|---|
| **Primary** | |
| **Phase 1b**:<br>• Determine the MTD and RP2D of binimetinib administered in combination with nivolumab<br>• Determine the MTD and RP2D of binimetinib administered in combination with nivolumab plus ipilimumab | • Incidence of dose-limiting toxicities (DLTs) resulting from binimetinib in combination with nivolumab<br>• Incidence of DLTs resulting from binimetinib in combination with nivolumab plus ipilimumab |
| **Phase 2:** Assess the preliminary antitumor activity of the treatment combinations based on Response Evaluation Criteria in Solid Tumors (RECIST) version 1.1 | ORR per RECIST v1.1 |
| **Secondary** | |
| **Both phases**:<br>• Assess the activity of the treatment combinations based on RECIST version 1.1<br>• Characterize the safety profile of the treatment combinations<br>• Characterize the PK of binimetinib in both treatment combinations | • ORR per RECIST v1.1 (Phase Ib only)<br>• DOR per RECIST v1.1<br>• Rate of CR per RECIST v1.1<br>• Incidence and severity of AEs graded according to the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) v4.03, and changes in clinical laboratory parameters<br>• Sparse plasma concentrations for binimetinib |
| **Exploratory** | |
| **Both phases:**<br>• Obtain preliminary estimates of PFS and OS | • PFS per RECIST v1.1<br>• OS |
| CCI [redacted] | |

Abbreviations: AE = adverse events; CR = compete response; DLT = dose-limiting toxicity; DOR = duration of response; MTD = maximum tolerated dose; ORR = overall response rate; OS = overall survival; PFS = progression-free survival; PK = pharmacokinetic; R2PD = recommended Phase 2 dose; RECIST = Response Evaluation Criteria In Solid Tumors;

## 2.2. Study Design

This is a multicenter, open-label Phase 1b/2 study to determine the maximum tolerated dose (MTD) and recommended Phase 2 dose (RP2D) and schedule of binimetinib, and to assess the safety, efficacy, and pharmacokinetics (PK) of binimetinib administered in combination with nivolumab or nivolumab plus ipilimumab in patients with previously treated microsatellite stable (MSS) metastatic colorectal cancer (mCRC) with documented *RAS* mutation. The study will include a dose-finding period in Phase 1b followed by a randomized Phase 2 period (Figure 1). Patients will be assigned (Phase 1b) or randomized (Phase 2) to one of the following arms:

**Phase 1b (Dose Determination):**

Arm 1A - Binimetinib plus nivolumab (Doublet)

Arm 1B - Binimetinib with nivolumab plus ipilimumab (Triplet)

**Phase 2 (Efficacy Assessment):**

Arm 2A - Binimetinib plus nivolumab (Doublet)

Arm 2B - Binimetinib with nivolumab plus ipilimumab (Triplet)

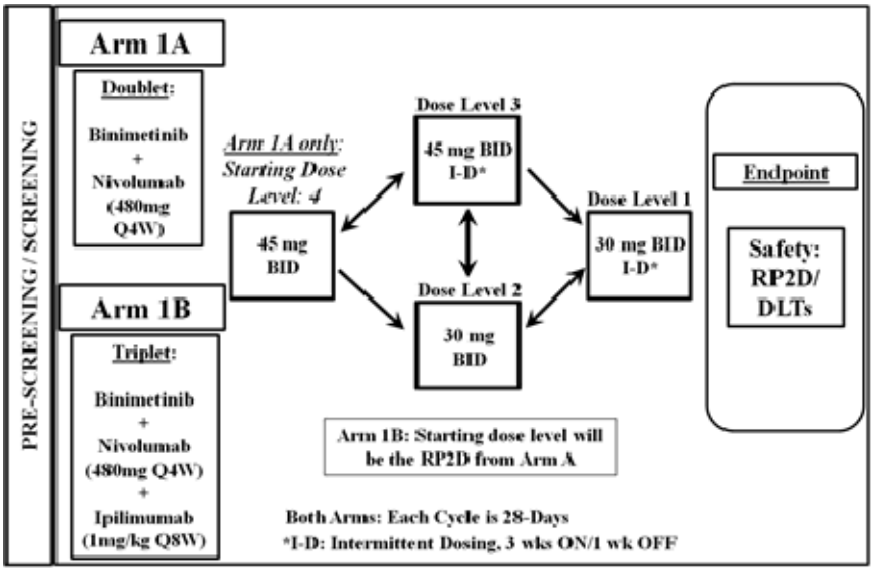
The aim of Phase 1b (n~42) is to determine the MTD and RP2D and schedule of binimetinib in combination with nivolumab (Doublet) and the MTD and RP2D and schedule of binimetinib in combination with nivolumab + ipilimumab (Triplet).

The aim of Phase 2 (n~48) is to assess the clinical efficacy of the doublet and triplet and will further characterize the safety of the drug combinations.

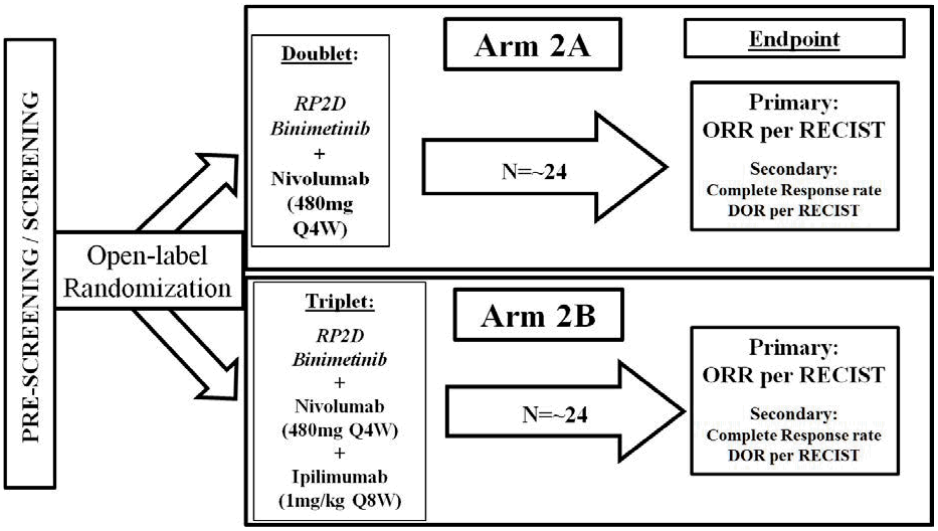
Patients will be treated during 28-day cycles and will continue until progression of disease, unacceptable toxicity develops, withdrawal of informal consent, initiation of subsequent anticancer therapy, the patient is lost to follow-up, or death.

**Figure 1. Study Design Schematic**

**Phase 1b**




**Phase 2**




### 2.2.1. Phase 1b

Phase 1b of the study will consist of dose-finding cohorts in Arm 1A and Arm 1B. Enrollment of patients in the dose-finding period and all dose-escalation and de-escalation decisions will be guided by the modified toxicity probability interval (mTPI-2) design (Guo et al. 2017).[5] The mTPI-2 design is a model-based approach that is guided by a prespecified decision matrix that recommends escalating, reducing, or maintaining the same dose or stopping dose escalation based on the number of patients with dose limiting toxicities (DLTs) observed in the dose level under evaluation. Doses of binimetinib, nivolumab, and ipilimumab to be explored in the dose-finding period are presented in protocol Table 2.

There will be a period of at least 7 days between C1D1 of the first patient and C1D1 of the second patient enrolled during Phase 1b (Arms 1A and 1B).

Arm 1A (Doublet):

The first patient enrolled will receive 480 mg nivolumab every 4 weeks (Q4W) in combination with the Starting Dose Level of 45 mg twice daily (BID) binimetinib. Subsequent patients will receive 480 mg nivolumab Q4W in combination with binimetinib at a dose and schedule that will be determined based on the cumulative toxicities observed in current and prior patients.

Arm 1B (Triplet):

The first patient enrolled will receive 480 mg nivolumab Q4W with 1 mg/kg ipilimumab every 8 weeks (Q8W) in combination with binimetinib. The binimetinib Starting Dose Level will be assigned once the RP2D from Arm 1A has been determined. Subsequent patients will receive 480 mg nivolumab Q4W with 1 mg/kg ipilimumab Q8W in combination with binimetinib, at a dose and schedule that will be determined based on the cumulative toxicities observed in current and prior patients.

The Starting Dose Level of Arm 1B will be the RP2D of binimetinib from Arm 1A.

The target DLT probability for Cycle 1 is 30%, with an equivalence interval, ie, an acceptable interval, of 25% to 35%. Dosing in Phase 1b will continue within an arm until 9 patients have been treated at one dose level with a recommendation to stay (or a recommendation to escalate if there is not a higher dose level or if the tolerability of the next dose level is unacceptable), or the maximum sample size within a Phase 1b arm ($n$=21) has been reached. If further safety evaluations are required, additional patients may be added to a cohort.

If the first 2 patients in a previously untested dose level experience a DLT, the next cohort will be opened at the next lower dose level/schedule or an intermediate dose level. However, if the first 2 patients in a new cohort at a previously tested dose combination/schedule experience a DLT, further enrollment to that cohort will stop and the Sponsor's Medical Monitor and the Investigators will meet to discuss all available data for that dose level/schedule and determine if additional patients may be enrolled in the dose cohort or if a new cohort of patients will be enrolled at a lower dose level.

Toxicities will only be considered DLTs if they occur in Cycle 1 although overall safety including later cycles will be considered for dose escalations and for each RP2D determination.

Patients are required to complete Cycle 1 (≥75% of the planned cumulative dose of binimetinib) to be considered evaluable for MTD determination unless discontinuation occurred due to a DLT.

The dose-finding period of the study will evaluate the safety and tolerability of the combinations of nivolumab and binimetinib with or without ipilimumab for one cycle (the first 28 days of treatment in Phase 1b).

### 2.2.2. Phase 2

Phase 2 of the study will consist of two arms to investigate the safety and clinical activity of the RP2Ds established from Phase 1b, Arm 2A (nivolumab in combination with binimetinib) and Arm 2B (nivolumab plus ipilimumab in combination with binimetinib).

- Arm 2A: Patients randomized to the nivolumab and binimetinib arm will receive nivolumab administered as 480 mg Q4W as a 30-minute (±5 minutes) IV infusion (not including time required to flush the IV lines) on Day 1 of each treatment cycle and the recommended dose of binimetinib in Arm 1A of Phase 1b.

- Arm 2B: Patients randomized to the nivolumab, ipilimumab, and binimetinib arm will receive nivolumab administered as 480 mg Q4W as a 30-minute (±5 minutes) IV infusion (not including time required to flush the IV lines) on Day 1 of each treatment cycle, ipilimumab administered as 1 mg/kg Q8W as a 30-minute (±5 minutes) IV infusion (not including time required to flush the IV lines), and the recommended dose of binimetinib determined in Arm 1B of Phase 1b.

Dosing calculations for ipilimumab should be based on the body weight assessed at baseline. It is not necessary to re-calculate subsequent doses if the patient weight is within 10% of the weight used to calculate the initial dose.

The primary analysis will take place after all patients who receive at least one dose of any study drug have had the opportunity for at least two post-baseline tumor assessments (or have discontinued tumor assessments beforehand).

## 2.3. Sample Size Justification

For the dose-finding period of this study (Phase 1b), the primary objective is to determine the RP2D of binimetinib in combination with nivolumab with or without ipilimumab. The total number of patients enrolled in the dose-finding period will depend on the number of dose levels tested and the number of patients treated in each cohort before the RP2D has been determined.

Approximately 90 patients are planned. A maximum of approximately 42 patients will be enrolled in Phase 1b and a minimum of approximately 48 patients will be enrolled in Phase 2.

Ji and Wang (2013)[6] recommended a sample size of $n = k \times (d+1)$ for Phase I studies that implement the mTPI design, where $k$ is the cohort size and $d$ is the number of dose levels. In the Phase I period of the study, patients are evaluated in group of 3 to 6 at each dose level. With 4 possible dose levels, a maximum of approximately 21 patients are planned to be treated in each arm, with a target of 9 patients at the MTD.

In the Phase 2 period of the study, a minimum of approximately 24 patients will be treated in each arm. Randomization into Phase 2 will continue until at least 33 patients have received the doublet and the triplet, respectively, at the RP2D (Phase 1b and Phase 2 combined).

The primary analysis of the response rate and efficacy endpoints will include all Phase 2 patients plus those treated at the same dose/schedule in Phase 1b; thus, it is expected that a minimum of 33 patients in each arm will be available for the primary efficacy evaluation.

In the doublet arm, a response rate of 5% or less will be considered unacceptably low. Hence, in, the null hypothesis that the response rate is at most 5% will be tested against the alternative hypothesis that the response rate is greater than 5%. The targeted response rate is 20%. Based on a test of one proportion using an arcsine transformation, with these assumptions, and a two-tailed alpha = 0.10 significance level, 33 patients will provide approximately 86% power.

In the triplet arm, a response rate of 10% or less will be considered unacceptably low. Hence, the null hypothesis that the response rate is at most 10% will be tested against the alternative hypothesis that the response rate is greater than 10%. The targeted response rate is 30%. With these assumptions, and a two-tailed alpha = 0.10 significance level, 33 patients will provide approximately 91% power.

The study is not adequately powered to compare the response rate between treatment arms; these comparisons will be descriptive only.

## 2.4. Analysis Sets

The number of patients in each analysis set will be summarized by treatment group, as well as listed by patient.

### 2.4.1. Full Analysis Set

The Full Analysis Set (FAS) will consist of all patients who receive at least one dose of any study drug in Phase Ib and all patients randomized to study treatment in Phase 2. According to the intention-to-treat principle, patients will be analyzed according to the treatment they have been assigned during randomization.

### 2.4.2. Safety Set

The Safety Set (SS) will consist of all patients who receive at least one dose of any study drug.

### 2.4.3. Dose-Determining Set

The Dose-determining Set (DDS) includes all patients from Phase 1b who either experience a DLT or receive at least 75% of the planned binimetinib dose intensity during the first cycle of treatment.

Patients in the DDS who do not experience DLT during the first cycle will be considered to have sufficient safety evaluations if they have been observed for ≥28 days following the first dose and are considered by both the Sponsor and Investigators to have enough safety data to conclude that a DLT did not occur.

### 2.4.4. Pharmacokinetic Analysis Set

The pharmacokinetic analysis set (PAS) consists of all patients who receive at least one dose of binimetinib and had at least one evaluable post-baseline binimetinib concentration measurement.

## 2.5. Treatment Assignment and Treatment Groups

### 2.5.1. Duration of Exposure

For binimetinib: continuous dosing (BID)

- **Duration of exposure** (days) = date of last (non-zero) dose of study drug – date of first dose of study drug + 1.

For binimetinib: intermittent dosing (3 weeks on/1 week off, BID)

- **Duration of exposure** (days) = date of last (non-zero) dose of study drug – date of first dose of study drug + 1 – (number of completed cycles * 7).

Number of completed cycles will be determined by taking only the integer part of [(date of last (non-zero) dose of study drug – date of first (non-zero) dose of study drug + 1) / 28].

For nivolumab: every four weeks (Q4W)

- **Duration of exposure** (days) = date of last (non-zero) dose of study drug – date of first dose of study drug + 28.

For ipilimumab: every eight weeks (Q8W)

- **Duration of exposure** (days) = date of last (non-zero) dose of study drug – date of first dose of study drug + 56.

If last known date of study drug exposure is later than the report cutoff date, it will be replaced by the report cutoff date.

For patients who did not take any drug, the duration of drug exposure is by definition equal to zero.

Duration of exposure will be converted from days to months by dividing the number of days by 30.4375 and will be summarized in months including descriptive statistics (n, mean, median, minimum, maximum) and a frequency distribution. In addition, duration of exposure along with confirmed tumor response per RECIST criteria will be presented graphically.

### 2.5.2. Cumulative Dose

**Planned cumulative dose** (mg) = sum of all protocol specified doses across each planned day of dosing.

**Actual cumulative dose** (mg) = sum of all actual doses taken during the dosing period.

For patients who did not take any drug the actual cumulative dose is by definition equal to zero.

### 2.5.3. Dose Intensity and Relative Dose Intensity

For binimetinib: continuous dosing (BID) or intermittent dosing (3 weeks on/1 week off, BID)

- **Planned dose intensity** (mg/day) = planned cumulative dose (mg)/duration of exposure (days).
- **Actual dose intensity** (mg/day) = actual cumulative dose (mg)/duration of exposure (days).
- **Relative dose intensity** = 100*[actual dose intensity (mg/day)/planned dose intensity (mg/day)].

For nivolumab: every four weeks (Q4W)

- **Planned dose intensity** (mg/4 weeks) = planned cumulative dose (mg)/[number of planned doses].
- **Actual dose intensity** (mg/4 weeks) = actual cumulative dose (mg)/[number of planned doses].
- **Relative dose intensity** = 100*[actual dose intensity (mg/4 weeks)/planned dose intensity (mg/4 weeks)].

For ipilimumab: every eight weeks (Q8W)

- **Planned dose intensity** (mg/8 weeks) = planned cumulative dose (mg)/[number of planned doses].
- **Actual dose intensity** (mg/8 weeks) = actual cumulative dose (mg)/[number of planned doses].
- **Relative dose intensity** = 100*[actual dose intensity (mg/8 weeks)/planned dose intensity (mg/8 weeks)].

A summary of exposure, including duration, cumulative dose, actual dose intensity, and relative dose intensity (including categories <50%, 50%-<75%, 75%-<90%, 90%-<110%, and ≥110%, if applicable), will be presented for each study drug. Duration of exposure, cumulative dose, actual dose intensity, and relative dose intensity will also be listed for each patient by treatment group.

### 2.5.4. Dose Modifications

**Dose reduction:**

Dose reductions are not permitted for nivolumab and ipilimumab. For binimetinib, a dose reduction is defined as a decrease in dose from the protocol planned dose and a decrease from the previous non-zero dose, even if this decrease has been directly preceded by an interruption. For example, for binimetinib, in the sequence of total daily dose 90 mg – 0 mg – 60 mg, the 60 mg dose will be counted as a reduction.

If a patient moves from a higher than protocol planned dose down to the planned dose then this is not be counted as a reduction, however if they move directly from a higher than planned dose down to a lower than protocol planned dose or the planned dose on a less frequent regimen, then this is counted as a reduction.

If the dose on the first dosing record is lower than protocol planned dose this is also counted as a reduction.

**Dose interruption:**

A dose interruption is defined as an actual dose equal to zero (where the planned dose is not zero), between the first and the last non-zero doses. Dosing records with zero dose as the last dosing record will not be counted as an interruption.

Frequency counts and percentages of patients who have binimetinib dose reductions or any study drug interruptions, and the corresponding reasons, will be provided. The number of dose interruptions per patient, and the duration of dose interruptions (days) will also be summarized for each study drug.

## 2.6. Key Definitions

### 2.6.1. Study Day

The study day will be derived. Cycle 1 Day 1 (C1D1) will coincide with the first day of treatment with any study drug.

For assessments/events occurring on or after the start date of study treatment, study day will be calculated as:

Study day (days) = event date – start date of study treatment + 1.

Therefore, the first day of study treatment is study day 1.

For all assessment/events occurring prior to the start of the study treatment, study day will be negative and will be calculated as follows. There is no day 0 for study day.

Study day (days) = event date – start date of study treatment.

### 2.6.2. Baseline

Baseline is the last available and valid assessment performed before the first administration of study treatment, unless otherwise stated under the related assessment section of the protocol and/or the SAP. Baseline can be the day before first treatment administration or the same day as first treatment administration if a pre-dose assessment/value is available (eg, electrocardiogram (ECG), PK samples, samples for biomarkers).

If time is recorded for the first treatment dose and for a specific assessment performed the day of first dose, this assessment will be considered as baseline only if it is actually performed before the first dose, as checked using both times.

If time is not recorded, a specific assessment performed the day of first dose administration will be considered as baseline if, according to protocol, it should be performed before the first dose.

Patients with no data on a particular parameter before the first treatment administration will have a missing baseline for this parameter.

Computation of baseline for ECG, biomarker, and other endpoints are described in each specific section.

For safety comparison against baseline (eg, for laboratory parameters, vital signs, weight, etc.), baseline is considered as the last available assessment or value collected prior to start of treatment. If no time is provided for an assessment on C1D1, it will be treated as pre-dose assessment.

The ECG baseline will be the average of triplicate ECG measurements prior to the start of treatment on C1D1.

### 2.6.3. On-treatment Period

On-treatment period is defined as the time from the first dose date of any study drug through minimum of (last dose of study treatment + 30 days, start day of subsequent anticancer therapy - 1 day).

### 2.6.4. Last Contact Date

Last contact date will be derived for patients not known to have died on or before the analysis cut-off date. Imputed dates will not be considered for the determination of last contact date. Only dates associated with patient visits or actual assessment of the patient will be used in the derivation. Dates associated with a technical operation unrelated to patient status (eg, the date a blood sample was processed) will not be used. Assessment dates after

the cut-off date will not be applied to derive the last contact date. Last contact date will only be derived using the latest complete date among the following:

- Study drug start and end dates with non-missing dose (doses of 0 are allowed);
- RECIST assessment date with evaluation marked as done;
- Laboratory/PK collection date with sample collection marked as done;
- Vital sign and ECG assessment date with non-missing parameter value;
- Performance status assessment date with non-missing performance status;
- Start/end date of adverse events with non-missing verbatim term;
- Start/end date of antineoplastic therapies administered after study treatment; discontinuation with non-missing medication/procedure term;
- Cardiac imaging assessment date marked as done;
- Dermatolic exam date with non-missing result;
- Physcial exam date with non-missing result of normal or abnormal;
- Ophthalmic exam date marked as done;
- Biomarker collection date with sample collected marked as done;
- Date of contact for most recent post-treatment survival assessment.

## 2.7. Imputation Rules for Partial or Missing Dates

For computation of time intervals (eg, elapsed time between initial diagnosis and first recurrence/relapse), the time interval should be set to missing when the imputation rule leads to a negative value.

For patients who discontinue the study with ongoing events, the discontinuation date will be used as the completion date of the event.

### 2.7.1. Adverse Event (AE) Date Imputation

Missing and partial date for AE will be handled according to rules specified below.

There will be no attempt to impute the following:

- Completely missing AE start dates.
- AE start dates missing the year.

For partial AE start date, the date imputation will be based on the temporal relation between the partial date and start of treatment date.

For partial AE end date or completely missing end date (AE is ongoing), the date imputation will be based on the temporal relation between the partial date, and the last contact date.

Table 2 provides examples of the different considered imputations for AE start date.

Table 3 provides examples of the different considered imputations for AE end date.

AEs that are completely missing a start date will be considered as treatment emergent if the end date of the AE is either missing or occurs after the treatment start date.

**Table 2. Adverse Event Start Date Imputation Example Scenarios**

| Partial AE Start Date | Treatment Start Date | Temporal Relationship Compared to Treatment Start | Imputed Date |
|---|---|---|---|
| 12mmyyyy | 20OCT2001 | Uncertain | <blank> |
| ddmmm2000 | 20OCT2001 | Before | 01JUL2000 |
| ddmmm2002 | 20OCT2001 | After | 01JAN2002 |
| ddmmm2001 | 20OCT2001 | Uncertain | 21OCT2001 |
| ddSEP2001 | 20OCT2001 | Before | 15SEP2001 |
| ddOCT2001 | 20OCT2001 | Uncertain | 21OCT2001 |
| ddNOV2001 | 20OCT2001 | After | 01NOV2001 |

**Table 3. Adverse Event End Date Imputation Example Scenarios**

| Partial AE End Date | Last Contact Date | Ongoing | Imputed Date |
|---|---|---|---|
| Missing | 20OCT2001 | Yes | 20OCT2001 |
| ddmmm2000 | 20OCT2001 | No | 31DEC2000 |
| ddmmm2002 | 20OCT2001 | No | 31DEC2002 |
| ddmmm2001 | 20OCT2001 | No | 20OCT2001 |
| ddmmm2001 | 20OCT2001 | Yes | 31DEC2001 |
| ddSEP2001 | 20OCT2001 | No | 30SEP2001 |
| ddOCT2001 | 20OCT2001 | No | 20OCT2001 |
| ddOCT2001 | 20OCT2001 | Yes | 31OCT2001 |

### 2.7.2. Prior and Concomitant Medication and Non-drug Therapy Date Imputation

The imputation of the start date of prior and concomitant medication and non-drug therapy will follow the same conventions as for AE start date. The imputation of the end date of prior and concomitant medications will also follow the same conventions as for AE end date. However, imputation of non-drug therapy end date will be different, as the data collected on non-drug therapy does not include whether the therapy is ongoing. The imputation of non-drug therapy end date is as follows: completely missing end dates will not be imputed, end dates with only day missing will be imputed to the last day of the month, and end dates with both day and month missing will be imputed to December 31st.

### 2.7.3. Imputation of Initial Diagnosis of Cancer Date and Antineoplastic Therapy (Surgery, Radiotherapy, and Medications) Dates

#### 2.7.3.1. Date of Initial Diagnosis of Cancer

Incomplete dates of initial diagnosis of cancer will be imputed as follows:

- If only the day is missing, it will be imputed to the 15th day of the month.
- If both the day and month are missing and the year is prior to the year of the treatment start date, impute as July 1st.
- If both the day and month are missing and the year is the same as the year of the treatment start date, imputed as January 1st.
- If the date is completely missing, no imputation will be performed.

#### 2.7.3.2. Antineoplastic Therapies Date Imputation

**Prior therapies**

Start date:

In general, follow the same rules that are applied to the imputation of an AE/concomitant medication start date, except:

- Completely missing start dates are imputed as treatment start date - 1;
- If only day is missing, and month and year match that of the treatment start date, impute as treatment start date - 1;
- If both day and month are missing, and the year matches that of the treatment start date, then impute as treatment start date - 1.

End date:

- Imputed date = min(treatment start date - 1, last day of the month), if day is missing.
- Imputed date = min(treatment start date - 1, December 31st), if month and day are missing.
- Completely missing end dates will not be imputed.

Date of progression recorded on prior antineoplastic medication eCRF page

If date of progression is partial with day missing, then date of progression will be imputed as min(last day of the month, treatment start date - 1).

**Post therapies**

Start date:

- Imputed date = max(last date of study drug + 1, first day of the month), if day is missing.
- Imputed date = max(last date of study drug + 1, January 1st), if day and month are missing.
- Completely missing start dates will be imputed as last date of study drug + 1.

Last date of study drug is the date of treatment discontinuation as collected from the disposition page.

End date: No imputation.

### 2.7.4. Imputation of Last Date of Study Drug Administration

If the last date of study drug is completely missing and there is no end of treatment eCRF page and no death date, the patient should be considered to be on-going and use the cut-off date for the analysis as the last dosing date.

If the last date of study drug is completely or partially missing and there is either an end of treatment eCRF page or a death date, then imputed last dose date will be as follows:

- = 31DECYYYY, if **only Year** is available and Year < Year of min (EOT date, death date).
- = Last day of the month, if **both Year and Month** are available and **Year = Year** of min (EOT date, death date) and Month < the month of min (EOT date, death date).
- = Last day of the month, if **both Year and Month** are available and **Year < Year** of min (EOT date, death date).

- = min (EOT date, death date), for all other cases.

The imputed date will be compared with start date of study treatment:

- If the imputed date < start date of study drug, then last date of study drug is set to start date of study drug.

- Otherwise, use the imputed date.

# 3. DATA ANALYSIS METHODS

A CSR will be prepared based on all data collected from patients enrolled in Phase 1b and Phase 2. Data from participating centers in this study protocol will be combined, so that an adequate number of patients will be available for analysis. No center effect will be assessed. The data will be summarized with respect to demographic and baseline characteristics, efficacy observations and measurements, safety observations and measurements, and all relevant pharmacokinetic measurements using descriptive statistics (*n*, mean, standard deviation, median, minimum, and maximum) for quantitative data and contingency tables (frequencies and percentages) for qualitative data.

**Phase 1b dose de-escalation/escalation data**: Patients that are treated at the same dose level and schedule in each arm will be pooled into a single treatment group. All summaries, listings, figures and analyses will be performed by treatment group and overall (as applicable).

**Phase 2 data**: All summaries, listings, figures and analyses will be performed by study arm and overall (as applicable).

**Note**: Patients from the Phase 1b and the Phase 2 arms treated at the same dose level and schedule will be pooled in all summaries, except those on the DDS, which is only defined for Phase 1b patients.

## 3.1. Study Conduct

Unless specified otherwise, summaries described in this section will be based on the FAS. Summaries will be produced by treatment group and overall.

### 3.1.1. Demographics and Baseline Characteristics

Demographic, and other baseline data including age, gender, race, ethnicity, height, weight, body mass index (BMI), and Eastern Cooperative Oncology Group (ECOG) performance status, will be listed individually by patient and summarized by treatment group in Phase 1b and by study arm in Phase 2 using descriptive statistics (continuous data) or contingency tables (categorical data). BMI ($kg/m^2$) will be defined as: [weight (lb)*0.453592]/[(height (in)*0.0254)$^2$], which simplifies to 703 * weight (lb)/[height (in)$^2$]. In addition, age (<65, ≥65 years) and weight (<120, 120-165, ≥165 lb) categories will be summarized.

The summaries and listings for each part will be based on the assessments from the screening visit/baseline.

### 3.1.2. Subject Disposition

The following disposition categories will be summarized:

- Number (%) of patients who discontinued the study treatment;
- Primary reasons for end of treatment;
- Number (%) of patients who continue to be followed for study evaluation completion;
- Number (%) of patients who were no longer being followed for study evaluation completion;
- Primary reasons for study evaluation completion.

Patient disposition will also be listed.

Screen failure patients are those who were screened, but never started the study treatment for any reason. The data collected on these patients will not be included in any analyses. A listing of reasons for screen failure will be presented.

### 3.1.3. Protocol Deviation

All protocol deviations will be finalized before database lock. Protocol deviations will be tabulated and listed. The number and percentage of patients in the FAS with any protocol deviations will be tabulated by the deviation category. In addition, protocol deviations due to COVID-19 will be presented in a separate listing.

### 3.1.4. Medical History

Medical history will be summarized and listed. Medical history is coded using the latest Medical Dictionary for Regulatory Activities (MedDRA) terminology available at the time of reporting.

### 3.1.5. Disease History

The summary of disease history will include stage at study entry, location of primary tumor, status of tumor removal/resection prior to study entry, time since initial diagnosis to start of study treatment (months), and RAS mutation status, RAS alteration, and MSS assessment results. Disease history will also be listed for each patient by treatment group.

### 3.1.6. Prior Antineoplastic Therapy

Prior antineoplastic therapy will be summarized for three distinct subtypes (medication, radiotherapy, and surgery). The number (%) of patients who received, separately, any prior antineoplastic medication, radiotherapy, or surgery will be summarized. All prior antineoplastic therapies will be listed separately for medication, radiotherapy, and surgery.

- Prior antineoplastic medications will also be summarized by the total number of regimens (there can be more than one medication per regimen), setting at last medication, best response at last medication (defined as the best response during the last treatment regimen recorded), reason for discontinuation at last medication, time (in months) from end of last medication to start of treatment, and time (in months) from start of last medication to disease progression. The last medication is defined based on the last end date of all prior regimen components. Prior antineoplastic medications will also be summarized by Anatomical Therapeutic Chemical (ATC) class and preferred term.

- For prior antineoplastic radiotherapy, information about the last radiotherapy (based on end date) will be summarized: time (in months) between radiotherapy and start of study treatment, location, setting, method, and whether 30% or more bone marrow was radiated at last radiotherapy.

- For prior antineoplastic surgery, the time (months) between last surgery (non-biopsy procedure) and start of study treatment will be summarized.

Incomplete dates will be handled as described in Section 2.7.

### 3.1.7. Subsequent Antineoplastic Medications

Subsequent antineoplastic medications, defined as antineoplastic medications taken after discontinuation of study treatment, will be listed and tabulated by ATC class and preferred term.

### 3.1.8. Prior and Concomitant Therapy

Prior and concomitant therapies (defined as any medications, excluding study drug and prior antineoplastic treatments) administered in the study and recorded in the Prior and Concomitant Medications eCRF, will be summarized and listed using the Safety Set. These medications will be coded using the World Health Organization Drug Reference Listing (WHO DRL) dictionary that employs the WHO ATC classification system.

The number and percentage of patients with prior and concomitant medication will be presented as follows. A prior medication summary and listing will be produced for medications that were started and stopped prior to the start of study drug. A concomitant medications summary and listing will be produced for medications that started prior to study drug and continued during study, as well as those that started during the on-treatment period.

Incomplete dates will be handled as described in Section 2.7.

## 3.2. Analysis of the Primary Endpoints

### 3.2.1. Phase 1b

Estimation of the MTD and RP2D of binimetinib administered in combination with nivolumab, as well as of binimetinib administered in combination with nivolumab plus ipilimumab will be based on the incidence of DLTs in Cycle 1 for patients in the DDS. Dose limiting toxicities occurring during the first cycle of Phase 1b will be summarized. All Phase 1b DLTs will be listed by treatment group.

### 3.2.2. Phase 2

The preliminary antitumor activity of the treatment combinations will be based on the Objective Response Rate (ORR) per Response Evaluation Criteria in Solid Tumors (RECIST) version 1.1. The best overall response (BOR) as assessed by the Investigator per RECIST v1.1 will be determined for each patient and is described further below. The ORR will be calculated within each treatment group, where ORR is defined as the number of patients achieving an overall best response of complete response (CR) or partial response (PR) divided by the total number of patients in that treatment group. ORR will be calculated with an exact (Clopper-Pearson) two-sided 90% confidence interval (CI). In addition, the exact two-sided 95% CI of ORR will also be calculated. Disease Control Rate (DCR) is defined as the proportion of patients with a best overall response of CR, PR, or stable disease (SD). DCR with its corresponding exact (Clopper-Pearson) 95% CI will be calculated.

Both confirmed and unconfirmed ORR and DCR will be summarized, but the primary analysis will be based on confirmed responses. This analysis will be conducted using patients in the FAS.

The BOR is the best response recorded from the start of the treatment until disease progression/recurrence (taking as reference for progressive disease [PD] the smallest measurements recorded since the treatment started). However, any assessments taken more than 30 days after the last dose of study therapy will not be included in the BOR derivation. Moreover, if any alternative cancer therapy is taken while on study, any subsequent assessments will be excluded from the BOR determination.

Confirmed responses are those confirmed by a follow-up scan at the site, as there is no central adjudication of local results planned for this study. The confirmed BOR for each patient is determined from the sequence of overall (lesion) responses according to the following rules below:

- CR = at least two determinations of CR at least 4 weeks apart before progression, where confirmation required, or one determination of CR prior to progression, where confirmation not required.

- PR = at least two determinations of PR or better at least 4 weeks apart before progression (and not qualifying for a CR), where confirmation required, or one determination of PR prior to progression, where confirmation not required.

- SD = an assessment of SD or better >6 weeks after start of treatment (and not qualifying for PD, PR or CR).

- PD = progression ≤12 weeks after start of treatment (and not qualifying for CR, PR, or SD).

- UNK (Unknown) = all other cases (ie, not qualifying for confirmed CR or PR and without SD after more than 6 weeks or early progression within the first 12 weeks).

Individual lesion measurements and overall response assessments will be listed by patient and assessment date.

A waterfall graph will be used to depict anti-tumor activity. This plot will display the best percentage change from baseline in the sum of diameters of target lesions for each patient. The term "best" means the largest shrinkage in tumor size, or if a patient only has assessments with tumor growth, take the assessment where the growth is minimal.

The best percentage change in target lesions from baseline will be derived across all the post-baseline assessments until documented disease progression, excluding assessments after start of subsequent antineoplastic therapy, as:

- Minimum of [(sum of target lesions at week XX – sum of target lesions at baseline)/sum of target lesions at baseline] × 100.

However, caution needs to be paid to the assessments, where an occurrence of a new lesion or worsening in non-target lesions (resulting in PD as an overall lesion response at given assessment) contradicts the measurements obtained on target lesions. These assessments will not be displayed as bars in the graph. If such a 'contradicting' assessment represents the only post-baseline assessment for a patient, then the patient will be represented by a special symbol (eg, *) in the waterfall graph.

The assessments with unknown overall response will be excluded. Patients without any valid assessments to calculate a percentage change from baseline value will be excluded from the graph. Only assessments performed before the start of any further antineoplastic therapies (ie, any additional secondary antineoplastic therapy or anti-cancer surgery) and not later than 30 days after the last administration of study drug will be included.

Patients will be ordered in the graph from left (worst change) to right (best change).

1. Bars above the horizontal axis (0%) representing tumor growth.

2. Bars under the horizontal axis (0%) representing tumor shrinkage.

3. Bars with value of 0 with * symbol representing patients with new lesions or non-target lesions that contradict the tumor measurements.

The total number of patients displayed in the graph (n) will be shown. The BOR (confirmed) per RECIST criteria will be shown above each of the displayed bars in the graph and the RAS mutation status for each patient will be displayed above the waterfall plot in a heat map. Symbols will be used to differentiate treatment groups. A horizontal threshold line at 20% and -30% will be shown.

In addition, a spider plot for each treatment arm will be created to plot the percentage change from baseline in the sum of diameters of target lesions by time since treatment start date (months) for each patient with at least one baseline and at least one post-baseline tumor assessment. A horizontal threshold line at 20% and -30% will be shown.

### 3.3. Analysis of the Secondary Endpoints

For all efficacy parameters, data will be listed, summarized, or analyzed by treatment group in Phase 1b and by study arm in Phase 2 using the FAS unless otherwise specified.

#### 3.3.1. Phase 1b

The antitumor activity of the treatment combinations will be based on the ORR per RECIST version 1.1. The ORR will be calculated for each treatment group. The ORR with exact (Clopper-Pearson)[2] 95% confidence intervals will be provided. Both confirmed and unconfirmed ORR will be summarized. This analysis will be conducted using the FAS.

#### 3.3.2. Both Phases

**Rate of CR**

Rate of CR is defined as the number of patients achieving an overall best response of CR per RECIST v1.1 divided by the total number of patients in that treatment group or treatment arm. Rate of CR will be calculated for each treatment group in Phase 1b and treatment arm in Phase 2. The rate of CR with exact (Clopper-Pearson) 95% confidence intervals will be provided. Both confirmed and unconfirmed rate of CR will be summarized. This analysis will be conducted using the FAS.

**Duration of Response**

Duration of response (DOR) per RECIST v1.1 will be calculated for confirmed “responders” only and is defined as the time between the date of first documented confirmed response (CR or PR) and the date of first documented progression or death due to any cause. If progression or death due to any cause has not occurred, the patient is censored at the date of last adequate tumor assessment other than unknown. If progression or death due to any cause has occurred after two or more missing assessments, or after new anticancer therapy is given then the patient is censored at the date of last adequate tumor assessment. In these cases, the date of last adequate tumor assessment is the date the last tumor assessment with overall lesion response of CR, PR, or SD which was made before an event or a censoring reason occurred.

An estimate of the DOR survival function will be constructed using the Kaplan-Meier (product-limit) method (Kaplan & Meier, 1958)[7] as implemented in PROC LIFETEST. The 25%, median, and 75% DOR (in weeks) will be summarized along with 95% confidence intervals as calculated from the PROC LIFETEST output (using method of [Brookmeyer & Crowley, 1982]).[1] Kaplan-Meier estimates with 95% confidence intervals at specific time points will be summarized as well. The confidence intervals are constructed using Greenwood's formula for the standard error of the Kaplan-Meier estimate. When the estimated survival function is close to zero or unity, symmetric intervals are inappropriate since they can lead to confidence limits that lie outside the interval [0, 1]. Any limit that is greater than unity will be replaced by 1.0. Any limit that is less than zero will be replaced by 0.0. A Kaplan-Meier plot of the survival function will be provided as well.

In addition, duration of response will be listed by patient and treatment group.

**Adverse Events and Changes in Clinical Laboratory Parameters**

The summary of incidence and severity of adverse events (AEs) graded according to the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE) v4.03 is described in Section 3.5.1. The summary of changes in clinical laboratory parameters is described in Section 3.5.5.

**Sparse plasma concentrations for binimetinib**

The summary of sparse plasma concentrations for binimetinib is described in Section 3.6.1.

## 3.4. Analysis of the Exploratory Endpoints

### 3.4.1. Progression-free Survival – per RECIST v1.1

Progression-free survival (PFS) is defined as the time from the start of treatment to the date of the first documented disease progression or death due to any cause. Patients who have not progressed or died at the time of the data cut-off will be censored at the date of last adequate tumor assessment. If progression or death due to any cause has occurred after two or more missing assessments, or after new anticancer therapy is given then the patient is censored at the date of last adequate tumor assessment. Patients who do not have baseline or post-baseline tumor assessments will be censored at the first dose date.

Last adequate tumor assessment is defined as a tumor assessment with a response assessment that is not missing or unknown.

In the case where a patient does not have documented progression but has disease progression as the reason for end of treatment/study evaluation completion, it will not be considered as a PFS event.

A Kaplan-Meier summary and plot for PFS will be presented for each phase, as described for DOR in Section 3.3.2. In addition, time to progression will be listed by patient and treatment group.

Frequency counts and percentages of patients with each event type (PD or death) and censoring reasons will be summarized. Censoring reasons are as follows:

- Ongoing in the study without an event;
- No baseline assessment;
- No post-baseline assessment;
- New anti-neoplastic therapy was given;
- Progression after 2 or more missed assessments;
- Death after 2 or more missed assessments;
- Adequate assessment no longer available;
- Withdrawal of consent;
- Lost to follow-up.

### 3.4.2. Overall Survival

Overall Survival (OS) is defined as the duration from the start of treatment to the time of death due to any cause. If a death has not been observed by the date of analysis cut-off, OS will be censored at the date of last contact.

A Kaplan-Meier summary and plot for OS will be presented for each phase, as described for DOR in Section 3.3.2.

Frequency counts and percentages of patients with an event (death) and censoring reasons will be summarized. Censoring reasons are as follows:

- Ongoing and no death;
- Withdrawal of consent;
- Lost to follow-up.

CCI

CCI

[redacted]

### 3.5. Safety Analysis

All safety data will be listed and summarized as detailed in the sections below. The assessment of safety will be based mainly on the type and frequency of adverse events, and on the number of laboratory values that fall outside pre-determined ranges (CTCAE version 4.03 grading limits or normal ranges, as appropriate). Other safety data (eg, ECGs, vital signs, and special tests) will be presented, as detailed below. All safety data will be listed. The Safety Set will be used for summaries of safety data, with the exception of DLTs for which the DDS will be used.

The safety summary tables will include assessments from baseline through those collected no later than 30 days after study treatment discontinuation.

#### 3.5.1. Adverse Events

An adverse event is considered treatment-emergent adverse event (TEAE) if the event starts on or after the first dosing day and time/start time, if collected, but before the last dose +30 days, or start of subsequent anticancer therapy minus 1 day, whichever occurs first. Unless otherwise stated, only TEAEs will be presented in summary tables.

All AEs (treatment emergent or not) and their attributes will be presented in data listings sorted by treatment group, patient identifier, AE and date of onset of the AE. Separate listings for AEs, serious AEs (SAEs), AEs requiring dose interruption, AEs leading to any study drug discontinuation, immune-mediated AEs, and death recorded during the study will be provided.

AEs will be coded using the latest version of MedDRA available prior to clinical database lock and will be graded using CTCAE version 4.03. Patients with a Grade 5 AE will be included in the “All grades” category, but will not be included in the “Grade 3/4" category.

All TEAE summaries will be summarized (frequency counts and percentages) by system organ class and preferred term, if not otherwise noted. The following TEAE summaries will be produced:

- Overview of TEAEs (overall and grade 3 and 4), including the number of patients with at least one adverse event, serious adverse event, adverse event leading to any study drug discontinuation, adverse event leading to binimetinib discontinuation, adverse event leading to nivolumab discontinuation, adverse event leading to ipilimumab discontinuation, adverse event leading to study treatment discontinuation,

adverse event leading to binimetinib dose reduction, adverse event leading to binimetinib dose interruption, adverse event leading to nivolumab dose interruption, adverse event leading to ipilimumab dose interruption, adverse event requiring additional therapy, dose limiting toxicities, and dose limiting toxicities leading to discontinuation. Summaries will include both all adverse events and those that are treatment related. In addition, serious adverse events, adverse events leading to discontinuation, dose limiting toxicities, and dose limiting toxicities leading to discontinuation will be also be summarized by those that are related to binimetinib, those that are related to nivolumab, and those that are related to ipilimumab. The number of deaths and on-treatment deaths will also be summarized.

- TEAEs from all causalities (overall and maximum grade 3 and 4), also summarized by preferred term only.
- TEAEs that are treatment related (overall and maximum grade 3 and 4), also summarized by preferred term only.
- TEAEs that are binimetinib related (overall and maximum grade 3 and 4).
- Treatment-emergent serious adverse events from all causalities (overall and maximum grade 3 and 4), also summarized by preferred term only.
- Treatment-emergent serious adverse events that are treatment related (overall and maximum grade 3 and 4).
- Treatment-emergent serious adverse events that are binimetinib related (overall and maximum grade 3 and 4).
- Treatment-emergent non-serious adverse events from all causalities (overall and maximum grade 3 and 4).
- TEAEs leading to binimetinib discontinuation from all causalities (overall and maximum grade 3 and 4).
- TEAEs leading to nivolumab discontinuation from all causalities (overall and maximum grade 3 and 4).
- TEAEs leading to ipilimumab discontinuation from all causalities (overall and maximum grade 3 and 4).
- TEAEs requiring binimetinib dose reduction from all causalities (overall and maximum grade 3 and 4).
- TEAEs requiring binimetinib dose interruption from all causalities (overall and maximum grade 3 and 4).

- TEAEs requiring nivolumab dose interruption from all causalities (overall and maximum grade 3 and 4).
- TEAEs requiring ipilimumab dose interruption from all causalities (overall and maximum grade 3 and 4).
- TEAEs experienced by ≥10% of patients in at least one treatment arm from all causalities, by preferred term.
- TEAEs experienced by ≥10% of patients in at least one treatment arm, that are treatment related, by preferred term.
- Immune-mediated adverse events from all causalities (overall and maximum grade 3 and 4).
- Immune-mediated adverse events that are treatment related (overall and maximum grade 3 and 4).
- Deaths (on-treatment), by primary reason.

Incomplete dates will be handled as described for AE in Section 2.7.

### 3.5.2. Laboratory Data

The severity grade will be calculated for laboratory values as applicable using CTCAE, version 4.03. For laboratory tests covered by CTCAE, version 4.03, a grade 0 will be assigned for all non-missing values not graded as 1 or higher. If the value is graded ≥1 but falls within the normal range (NR), the grade is reset to 0. Grade 5 will not be used. For laboratory tests where grades are not defined by CTCAE, results will be graded by the low/normal/high classifications based on laboratory normal ranges.

The following rules will be used for grading elevated glucose:

- If a glucose value falls within normal ranges it should be graded as zero, irrespective of fasting status.
- If fasting status is unknown and the result falls within the ranges for hyperglycemia grades 1 or 2, grade should be set to missing.
- If fasting status is known and fasting did not occur, and the result falls within the ranges for hyperglycemia grades 1 or 2, grade should be set to 0.

Corrected calcium will be derived from calcium and albumin results as per the following formula:

$$\textit{Corrected calcium (mmol/L)} = [4*\textit{calcium(mmol/L)} - 0.8*(0.1*\textit{albumin(g/L)} - 4)]/4$$

Values entered as "<x" or ">x", with x a numerical value, will be considered as x for the analysis.

The following by-treatment summaries will be generated separately for hematology and biochemistry laboratory tests:

- Shift tables using CTCAE grades to compare baseline to the worst on-treatment value
- For laboratory tests where CTCAE grades are not defined, shift tables using the low/normal/high/(low and high) classification to compare baseline to the worst on-treatment value.

The following listings will be produced:

- Listing of all laboratory data, presented separately by chemistry, hematology, urinalysis, and other laboratory test with values flagged to show the corresponding CTCAE grades and the classifications relative to the laboratory normal ranges.
- Listing of patients with laboratory abnormalities of CTCAE grade 3 and 4.

Table 4 and Table 5 list all laboratory parameters that will be summarized.

**Table 4. Laboratory Parameters (and directions) for which CTCAE Grades are Defined**

| Hematology and Coagulation | | Chemistry | |
|---|---|---|---|
| White Blood Cells (WBC) | ↑↓ | Creatinine | ↑ |
| Hemoglobin | ↑↓ | Sodium | ↑↓ |
| Platelets counts | ↓ | Potassium | ↑↓ |
| Absolute neutrophils | ↓ | Glucose | ↑↓ |
| Absolute lymphocytes | ↑↓ | Corrected calcium | ↑↓ |
| INR | ↑ | Magnesium | ↑↓ |
| Activated partial thromboplastin time (aPTT) | ↑ | Albumin | ↓ |
| | | AST | ↑ |
| | | ALT | ↑ |
| | | Total bilirubin | ↑ |
| | | Creatine kinase | ↑ |
| | | Alkaline phosphatase | ↑ |
| | | Lipase | ↑ |
| | | Amylase | ↑ |

↑ Indicates that CTC grade increases as the parameter increases.
↓ Indicates that CTC grade increases as the parameter decreases.

**Table 5. Laboratory Parameters (without CTCAE grades) for which Lab Reference Ranges are Defined**

| Hematology and Coagulation | Chemistry | Thyroid Function Tests |
|---|---|---|
| Prothrombin time (PT) | Urea or BUN | Free T3 |
| Hematocrit | Total protein | Free T4 |
| Basophils | Chloride | TSH |
| Eosinophils | CA 19-9 | |
| Monocytes | BNP | |
| RBC | Bicarbonate | |
| | Direct bilirubin | |
| | CEA | |
| | LDH | |
| | Uric acid | |
| | Troponin I | |

Other clinical laboratory tests that will be performed include:

- At Screening only: Hepatitis B surface antigen, Hepatitis C antibody, human immunodeficiency virus (HIV) (as applicable), and C-reactive protein (CRP);
- If applicable: serum pregnancy test; luteinizing hormone (LH), follicle-stimulating hormone (FSH), and/or estradiol; serum creatine kinase (CK) isoenzymes, myoglobin.

Hepatic toxicity will be assessed based on the following Liver Function Tests: Alanine aminotransferase (ALT), Aspartate aminotransferase (AST), Alkaline phosphatase (ALP), and total bilirubin (TBL). Number and percentage of patients will be provided for predefined categories of notable hepatic lab values (Table 6).

**Table 6. Hepatic Toxicity Criteria**

| Parameter | Criterion |
|---|---|
| ALT | >3xULN; >5xULN; >8xULN;>10xULN;>20 ULN |
| AST | >3xULN; >5xULN; >8xULN >10xULN;>20 ULN |
| AT (ALT or AST) | >3xULN; >5xULN; >8xULN >10xULN;>20 ULN |
| TBL | >1.5xULN, >2xULN |
| ALP | >2xULN, >3xULN |
| AT & TBL | AT >3xULN & TBL >2xULN;<br>AT >5xULN & TBL >2xULN;<br>AT >10xULN & TBL >2xULN |
| ALP & TBL | ALP >3xULN & TBL >2xULN |
| AT & TBL & ALP | AT >3xULN & TBL >2xULN & ALP <2xULN |

A by-patient listing will also be presented of patients who met the following criteria: AT >3xULN & TBL >2xULN & ALP <2xULN.

### 3.5.3. Vital Signs

Vital sign parameters collected are systolic and diastolic blood pressure, pulse rate, temperature, height, and weight. Vital sign values considered notably abnormal are defined in Table 7.

**Table 7. Criteria for Notable Vital Sign Values**

| Vital Sign | Criteria for Clinically Notable Vital Sign Values |
|---|---|
| Systolic blood pressure [mmHg] | ≥160 mmHg/≤90 mmHg with increase/decrease from baseline of ≥20 mmHg |
| Diastolic blood pressure [mmHg] | ≥100 mmHg/≤50 mmHg with increase/decrease from baseline of ≥15 mmHg |
| Pulse rate [bpm] | ≥120 bpm/≤50 bpm with increase/decrease from baseline of ≥15 bpm |
| Weight [lb] | ≥20% decrease from baseline or >=10% increase from baseline |
| Temperature [°F] | ≥99.5°F/≤96.8°F |

A summary of notably abnormal vital signs will be produced.

All vital signs measures will by listed. Clinically notable vital sign values will be listed separately.

### 3.5.4. Electrocardiograms (ECGs)

Baseline ECG will be obtained as the mean of triplicate measurements taken pre-dose on C1D1. Scheduled C1D1 pre-dose ECGs will be considered to have been obtained prior to study drug administration if dosing time is missing.

Data from ECGs will be listed, notable values will be flagged, and any other information collected will be listed as appropriate. Furthermore, combined QTc values from the triplicate ECGs will be averaged to provide a single value for each patient to be used in the summary tables.

The following summaries will be provided for each applicable ECG parameter:

- Frequency counts and percentages of patients having notable ECG values according to Table 8.
- For each QT interval (QT, QTcF, QTcB), shift tables based on notable QT interval categories (≤450, >450 - ≤480, >480 - ≤500, >500 msec) at baseline to the worst post-baseline value observed.

**Table 8. Criteria for Notable ECG Values**

| ECG Parameter | Criteria for ECG Notable Values |
|---|---|
| QT, QTcF, QTcB (msec) | Increase from baseline >30 msec, >60 msec<br>New QT interval >450, >480, >500 msec |
| HR (bpm) | Increase from baseline >25% and value >100 bpm<br>Decrease from baseline >25% and value <50 bpm |
| PR (msec) | Increase from baseline >25% and value >200 msec |
| QRS (msec) | Increase from baseline >25% and value >110 msec |

### 3.5.5. Left Ventricular Ejection Fraction

The analysis of cardiac assessments will include left ventricular ejection fraction (LVEF) (multi-gated acquisition [MUGA] scan or echocardiogram [ECHO]). Cardiac imaging results, including LVEF (%) and overall interpretation will be listed by patient. In addition, a shift table using CTCAE grades will be presented to compare baseline to the worst post-baseline LVEF value.

### 3.5.6. ECOG Performance Status

Performance status will be scored using the Eastern Cooperative Oncology Group (ECOG) performance scale (see Table 9) (Oken et al, 1982).[8] ECOG values collected over the study period will be listed by patient.

**Table 9. ECOG Performance Status Scale**

| Grade | ECOG |
|---|---|
| 0 | Fully active, able to carry on all pre-disease performance without restriction |
| 1 | Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, eg, light house work, office work |
| 2 | Ambulatory and capable of all self-care but unable to carry out any work activities. Up and about more than 50% of waking hours |
| 3 | Capable of only limited self-care, confined to bed or chair more than 50% of waking hours |
| 4 | Completely disabled. Cannot carry on any self-care. Totally confined to bed or chair |
| 5 | Dead |

### 3.5.7. Pregnancy Test

A listing will be produced for the serum pregnancy test to include all female patients regardless of test result.

### 3.5.8. Other Safety Data

**Ophthalmic examinations**

Visual acuity will be measured using the Snellen visual acuity. This is determined by establishing the smallest optotypes that can be identified correctly by the patient at a given observation distance. Snellen visual acuity can be reported as a Snellen fraction (m/M) in which the numerator (m) indicates the test distance and the denominator (M) indicates the distance at which the gap of the equivalent Landolt ring subtends 1 minute of arc.

For each timepoint, the LogMAR score will be calculated as –log(m/M).

The following summaries of ophthalmic examination data will be presented:

- Snellen visual acuity LogMAR shift table;
- Newly occurring intraocular pressure above 30 mmHg;
- Newly occurring abnormality in optical coherence tomography;
- Newly occurring abnormality in fluorescein angiography;
- Newly occurring abnormality in fundoscopy;
- Newly occurring abnormality in slit lamp.

Listings of all ophthalmic examination data, as well as notable LogMAR decreases, and newly occurring abnormalities in each ophthalmic assessment, will be presented by patient.

**Dermatologic evaluations**

A listing of all dermatologic examination data will be presented by patient.

## 3.6. Pharmacokinetics and Biomarkers

### 3.6.1. Pharmacokinetics

All PK analyses will be performed based on the PAS unless otherwise specified. Only valid PK concentrations will be used in the analyses.

Valid pre-dose samples will be:

- Taken at steady state (except for Day 1);
- Collected 10-14 hours after the last dose (except for Day 1);
- With no vomiting within the first 4 hours following the last dose (considering information collected in the eCRF).

Valid post-dose samples will be:

- Taken at steady state (except for Day 1);
- Collected within the 1.5 hours ± 15 minutes window;
- With no vomiting before the PK sample is taken (considering information collected in the eCRF).

Patients are considered to be at steady state if binimetinib is administered for at least 4 consecutive days at approximately the same total daily dose prior to the PK sampling.

Unscheduled samples will not be included in the descriptive analysis by timepoint but these samples are flagged in the corresponding concentration listing. Missing concentrations will not be imputed. Concentrations below the Lower Limit of Quantification (LLOQ) will be labeled as such in the concentration data listings. Concentrations below the limit of quantitation will treated as zero in summary statistics and excluded from geometric mean and geometric coefficient of variation (CV%) computation. All individual data will be listed and values excluded from the analysis will be flagged.

**Drug concentrations**

For each phase of the study, plasma concentrations of binimetinib and its active metabolite AR00426032 will be summarized by arm, timepoint, and actual leading dose, ie, the dose taken prior to the PK sampling. Descriptive statistics will include arithmetic and geometric mean, median, standard deviation, CV%, geometric CV%, minimum and maximum. If the number of samples is below 3, only *n*, minimum, and maximum is presented.

**Other pharmacokinetic analysis**

For both binimetinib and its active metabolite AR00426032, the accumulation ratio for $C_{min}$ (Day 15 pre-dose concentration for later cycles, 2 through 5, divided by the Cycle 1 Day 15 pre-dose concentration), $R_{acc\_cmin}$, dose-normalized when necessary to account for a different dose in later cycles, will be calculated for each patient. These $R_{acc\_cmin}$ values will also be summarized by cycle and leading dose, as well as over all patients regardless of dose. Graphical displays may also be generated.

If appropriate, others pharmacokinetic analysis based on population PK modeling will be performed and described in a separate standalone modeling plan and a specific report will be produced. CCI [redacted]

### 3.6.2. Biomarkers

All biomarker analyses will be performed based on the Safety Set.

The protocol-specified biomarkers CRP, CEA, and CA19-9 will be listed by patient and visit. Values will be flagged to indicate low or high values and the corresponding normal range given. These biomarkers will also be summarized using descriptive statistics by visit.

In addition, the protocol specifies that MSS disease assessment and RAS mutation will be assessed at Prescreening/Screening. The MSS disease assessment, RAS mutation, and RAS alteration will be listed by patient and summarized by treatment group.

## 4. REFERENCES

1. Brookmeyer R and Crowley J (1982). A Confidence Interval for the Median Survival Time, *Biometrics* 38, 29 - 41.

2. Clopper J and Pearson S (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* 26.4: 404-413.Gough et al. (1995). Assessment of dose proportionality: report from the statisticians in the pharmaceutical industry. Pharmacokinetic UK Joint Working Party. *Drug information Journal* 29: 1039-1040.

3. Collett D. (1944). *Modeling Survival Data in Medical Research*. London: Chapman & Hall.

4. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, Dancey J, Arbuck S, Gwyther S, Mooney M, Rubinstein L, Shankar L, Dodd L, Kaplan R, Lacombe D, Verweij J. (2009). New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *European Journal of Cancer* 45:228-247.

5. Guo W, Wang SJ, Yang S, et al. (2017) A Bayesian interval dose-finding design addressing Ockham's razor: mTPI-2. *Contemp Clin Trials* 58:23-33.

6. Ji Y, Wang SJ (2013). Modified toxicity probability interval design: a safer and more reliable method than the 3 + 3 design for practical phase I trials. *Journal of Clinical Oncology* 31: 1785-1791.

7. Kaplan EL, Meier P. (1958). Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association* 53:457-481.

8. Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, Carbone PP. (1982). Toxicity And Response Criteria Of The Eastern Cooperative Oncology Group. *Am J Clin Oncol* 5:649-655.

C
C
I
[redacted]

10. Smith BP, Vandenhende FR, DeSante KA, Farid NA, Welch PA, Callaghan JT, Forgue ST. (2000). Confidence interval criteria for assessment of dose proportionality. *Pharm Res* 17:1278–1283.