STATISTICAL ANALYSIS PLAN

The BEACON CRC Study (Binimetinib, Encorafenib, And Cetuximab COmbiNed to Treat *BRAF*-mutant ColoRectal Cancer):

A Multicenter, Randomized, Open-label, 3-Arm Phase 3 Study of Encorafenib + Cetuximab Plus or Minus Binimetinib vs. Irinotecan/Cetuximab or Infusional 5-Fluorouracil (5-FU)/Folinic Acid (FA) /Irinotecan (FOLFIRI)/Cetuximab with a Safety Lead-in of Encorafenib + Binimetinib + Cetuximab in Patients with *BRAF* V600E-mutant Metastatic Colorectal Cancer

PROTOCOL NUMBER:	ARRAY-818-302
VERSION:	5.0
DATE:	28 January 2019

Prepared by:

PPD PhD

PPD

Array BioPharma Inc.
3200 Walnut Street
Boulder, CO 80301

[redacted]

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SIGNATURE PAGE

I have read the statistical analysis plan for Clinical Study ARRAY-818-302 dated 28 January 2019 and confirm that to the best of my knowledge it accurately describes the planned analyses for the study

PPD

29 Jan 2019

PPD, PhD
Project Statistician PPD
Date

PPD

29 Jan 2019

PPD, MS
Statistical Programmer
Date

PPD

29 Jan 2019

PPD, MD
Clinical Science Lead
Date

29-Jan-2019

PPD, PhD
Clinical Pharmacology Lead
Date

PPD

29-JAN-2019

PPD, MD
Drug Safety Lead
Date

PPD

29 Jan 2019

PPD, PhD
Translational Science Lead
Date

PPD

25 Jan 2019

Medical Writer
Date

LIST OF ABBREVIATIONS

Abbreviation or Special Term	Explanation
5-FU	5-Fluorouracil
AE	adverse event
Array	Array BioPharma Inc.
ATC	Anatomical Therapeutic Chemical
AUC	area under the plasma concentration-time curve
AUC_{0-x}	area under the plasma concentration-time curve from zero to the specified time point
AUC/D	dose-normalized AUC
AUC_{ext}	percentage of AUC_{inf} due to extrapolation from T_{last} to infinity
AUC_{inf}	area under the plasma concentration-time curve extrapolated to infinity
AUC_{last}	area under the plasma concentration-time curve from zero to the last measurable time point
AUC_{tau}	area under the plasma concentration-time curve over the dosing interval
BLQ	below the limit of quantitation
BP	blood pressure
Bpm	beats per minute
C1D1	Day 1 of Cycle 1
C1DX	Day X of Cycle 1
C_{last}	last measured concentration
CL/F	apparent systemic clearance
C_{max}	observed maximum plasma concentration
C_{max}/D	dose-normalized C_{max}
CR	complete response
CRP	C-reactive protein
CTCAE	Common Terminology Criteria for Adverse Events
CV	coefficient of variation
DILI	drug-induced liver injury
DDS	dose-determining set
EDC	electronic data capture
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic case report form
EORTC	European Organization for Research and Treatment of Cancer
EORTC QLQ-C30	European Organization for Research and Treatment of Cancer Quality of Life Questionnaire

Abbreviation or Special Term	Explanation
EQ5D-5L	EuroQol-5D-5L
FA	Folinic Acid
FACT-C	Functional Assessment of Cancer Therapy-Colon cancer
FAS	full analysis set
FDA	United States Food and Drug Administration
f_e	fraction of daily dose excreted
FOLFIRI	5-FU/FA/irinotecan
HR	hazard ratio
IWRS	Interactive Web Response System
k_{el}	terminal elimination rate constant
KM	Kaplan-Meier
LF_{AUC}	linearity factor based on AUC values
LLT	lowest level term
logMAR	logarithm of the minimum angle of resolution
LVEF	left ventricular ejection fraction
MedDRA	Medical Dictionary for Regulatory Activities
MR_{AUC}	metabolite ratio based on AUC values
MR_{Cmax}	metabolite ratio based on C_{max} values
Msec	millisecond(s)
MSI	microsatellite instability
n, N	number
NC	not calculated
NCA	noncompartmental analysis
NCI	National Cancer Institute
OCT	optical coherence tomography
ORR	objective response rate (overall response rate)
OS	overall survival
PAP	pharmacokinetic analysis plan
PD	progressive disease
PFS	progression-free survival
PFS2	progression after next line of therapy
PGIC	Patient Global Impression of Change
PK	pharmacokinetic(s)
PPS	per protocol set
PR	partial response

Abbreviation or Special Term	Explanation
PT	preferred term
QTc	QT interval corrected for heart rate
QTcB	QT interval corrected for heart rate using Bazett's formula
QTcF	QT interval corrected for heart rate using Fredericia's formula
R_{AUC}	accumulation ratio based on AUC values
R_{Cmax}	accumulation ratio based on C_{max} values
RECIST	Response Evaluation Criteria in Solid Tumors
SAE	serious adverse event
SAP	statistical analysis plan
SD	standard deviation
SI units	International System of Units
SOC	system organ class
SS	safety set
$t_{1/2}$	terminal elimination half-life
SAE	treatment-emergent serious adverse event
T_{last}	observed time of C_{last}
TLF	tables, listings and figures
T_{max}	observed time of C_{max}
TTR	time to overall response
V_z/F	apparent volume of distribution
WHO	World Health Organization

1.0 INTRODUCTION

This document describes the statistical analyses and data presentations to be performed for the clinical protocol ARRAY-818-302 entitled "A Multicenter, Randomized, Open-label, 3-Arm Phase 3 Study of Encorafenib + Cetuximab Plus or Minus Binimetinib vs. Irinotecan/Cetuximab or Infusional 5-Fluorouracil (5-FU)/Folinic Acid (FA) /Irinotecan (FOLFIRI)/Cetuximab with a Safety Lead-in of Encorafenib + Binimetinib + Cetuximab in Patients with *BRAF* V600E-mutant Metastatic Colorectal Cancer." This statistical analysis plan (SAP) should be read in conjunction with the study protocol and electronic case report forms (eCRFs). This document has been developed using the protocol Version 7 dated 25 January 2019 and eCRFs dated 10 August 2018. Any further changes to the protocol or eCRFs may necessitate updates to the SAP.

This SAP provides a comprehensive and detailed description of the strategy, rationale and statistical techniques to be used to assess the safety, efficacy, pharmacokinetics (PK) and biomarker analysis in patients with $BRAF^{V600E}$, RASwt mCRC whose disease has progressed after 1 to 2 prior regimens in the metastatic setting as outlined in the protocol. The purpose of this SAP is to ensure the credibility of the study findings by specifying the statistical approaches for the initial and final analysis of study data. This SAP will be finalized and signed prior to the clinical database lock for any analyses specified in this SAP.

Statistical analyses detailed in this SAP will be conducted using SAS®, Version 9.2 or higher (SAS Institute, Inc., Cary, NC USA). Noncompartmental PK analyses will be performed with Phoenix WinNonlin® (Pharsight Corporation, St. Louis, MO USA). Original sample size calculations were conducted using East v6.3.1 (Cytel). For protocol version 7, East v6.4 (Cytel) was used in the power calculations.

1.1 Responsibilities

An Array Biostatistician or Statistical Programmer (or delegate) will perform any statistical analyses required for patient disposition, protocol deviations, patient characteristics, efficacy and safety; an Array Biostatistician (or delegate) is responsible for the production and quality control of all tables, figures and listings associated with these analyses. An Array Clinical Pharmacology representative (or delegate) will perform any PK parameter analyses, and is responsible for the production and quality control of all tables, figures and listings associated with these analyses. An Array Translation Medicine representative (or delegate) will perform any biomarker statistical analyses, and is responsible for the production and quality control of all tables, figures and listings associated with these analyses.

2.0 STUDY OBJECTIVES

2.1 Primary Objectives

- Safety Lead-In: Assess the safety/tolerability of the combination of encorafenib + binimetinib + cetuximab
- Randomized Phase 3:
  - Compare the activity of encorafenib + binimetinib + cetuximab (Triplet Arm) vs. irinotecan/cetuximab or 5-FU/FA/irinotecan (FOLFIRI)/cetuximab (Control Arm) as measured by overall survival (OS)
  - Compare the activity of encorafenib + binimetinib + cetuximab (Triplet Arm) vs. irinotecan/cetuximab or FOLFIRI/cetuximab (Control Arm) as measured by BICR determined objective response rate (ORR)

2.2 Key Secondary Objective

- Randomized Phase 3: Compare the activity of encorafenib + cetuximab (Doublet Arm) vs. irinotecan/cetuximab or FOLFIRI/cetuximab (Control Arm) as measured by OS

2.3 Secondary Objectives

- Safety Lead-In:
  - Assess the activity of encorafenib + binimetinib + cetuximab as measured by Investigator and BICR determined ORR, progression-free survival (PFS), duration of response (DOR) and time to response
  - Characterize the pharmacokinetics (PK) of encorafenib, cetuximab, binimetinib and the active metabolite of binimetinib (AR00426032)
- Randomized Phase 3:
  - Compare the Investigator-determined ORR of encorafenib + binimetinib + cetuximab (Triplet Arm) vs. irinotecan/cetuximab or FOLFIRI/cetuximab (Control Arm)
  - Compare the BICR-determined and Investigator-determined ORR of encorafenib + cetuximab (Doublet Arm) vs. irinotecan/cetuximab or FOLFIRI/cetuximab (Control Arm)
  - Compare the BICR-determined and Investigator-determined PFS of encorafenib + binimetinib + cetuximab (Triplet Arm) vs. irinotecan/cetuximab or FOLFIRI/cetuximab (Control Arm)

- Compare the BICR-determined and Investigator-determined PFS of encorafenib + cetuximab (Doublet Arm) vs. irinotecan/cetuximab or FOLFIRI/cetuximab (Control Arm)
- Compare the activity of Triplet Arm vs. Doublet Arm as measured by OS
- Compare the BICR-determined and Investigator-determined ORR of Triplet Arm vs. Doublet Arm
- Compare the BICR-determined and Investigator-determined PFS of Triplet Arm vs. Doublet Arm
- Compare BICR-determined and Investigator-determined DOR of Triplet Arm vs. Control Arm, of Doublet Arm vs. Control Arm and of Triplet Arm vs. Doublet Arm
- Compare BICR-determined and Investigator-determined time to response of Triplet Arm vs. Control Arm, of Doublet Arm vs. Control Arm and of Triplet Arm vs. Doublet Arm
- Assess the safety/tolerability of Triplet Arm, of Doublet Arm and of Control Arm
- Compare the effect on QoL of Triplet Arm vs. Control Arm, of Doublet Arm vs. Control Arm and of Triplet Arm vs. Doublet Arm
- Characterize the PK of encorafenib, cetuximab, binimetinib and the active metabolite of binimetinib (AR00426032)
- Assess for drug interactions between encorafenib, cetuximab, binimetinib and the active metabolite of binimetinib (AR00426032) based on PK modeling

2.4 Exploratory Objectives

- Safety Lead-In:
  - Assess the activity of encorafenib + binimetinib + cetuximab as measured by OS

[redacted]

3.0 INVESTIGATIONAL PLAN

3.1 Overall Study Design and Plan

This is a multicenter, randomized, open-label, 3-arm Phase 3 study to evaluate encorafenib + cetuximab plus or minus binimetinib versus physician's choice of either irinotecan/cetuximab or FOLFIRI/cetuximab, as controls, in patients with *BRAF*^{V600E} mCRC whose disease has progressed after 1 or 2 prior regimens in the metastatic setting. The study contains a Safety Lead-in Phase in which the safety and tolerability of encorafenib + binimetinib + cetuximab will be assessed prior to the Phase 3 portion of the study. In the Japanese Safety Lead-in portion of the study, a separate cohort of patients will be evaluated at a limited number of sites in Japan. Six patients will be enrolled on a rolling basis in a single cohort to evaluate the triplet dose. The starting dose in Japanese patients will be the dose assessed to be tolerable in the Safety Lead-in for non-Japanese patients. Similar to the Safety Lead-in for non-Japanese patients, the DMC will evaluate the safety data to confirm tolerability.

BRAF Testing

Patients will be eligible for the study based on identification of a *BRAF*^{V600E} mutation in the tumor as determined by the central laboratory as part of the Molecular Prescreening for the trial or by a local assay result obtained any time prior to Screening. Only polymerase chain reaction (PCR) and next generation sequencing (NGS)-based local assays results will be acceptable. If the patient is enrolled based on local assay results, the *BRAF* mutation status must be confirmed by the central laboratory no later than 30 days from first dose of study treatment.

In cases where there is discordance between the local assay and central laboratory results, or if the central laboratory is not able to confirm presence of a *BRAF*^{V600E} mutation due to inadequate or poor sample condition within 30 days of initiating study therapy, patients may only continue treatment if there is no clinical indication of deterioration or disease progression and the investigator determines that the patient is deriving benefit. In such instances, patients must be informed that the *BRAF* mutation status is unconfirmed and must sign a separate informed consent form that includes this information and describes alternative treatment options.

If at any time in the Phase 3 portion of the study there is lack of *BRAF*^{V600E} confirmation by the central laboratory (for any reason including discordance and inadequate available tissue) in 37 total patients (6% of the total planned enrollment of the randomized portion of the trial) or discordance (a valid result of "no *BRAF*^{V600E} mutation" as determined by the central laboratory) between the local assay and the central laboratory in 18 patients (3% of the total planned enrollment), all subsequent patients will be required to have *BRAF*^{V600E} determined by the central laboratory prior to enrollment. Central testing cannot be repeated to resolve discordances with a

local result once the central laboratory delivers a definitive result (positive or negative). If the result from the central laboratory is indeterminate or the sample is deemed is inadequate for testing, additional samples may be submitted. Results from local laboratories with more than 1 discordant result leading to patient enrollment will not be accepted for further patient enrollment. Sites with more than 2 randomized patients having indeterminate results will be required to enroll all subsequent patients based only on central laboratory assay results.

Molecular Prescreening

Prior to eligibility assessment for study enrollment/randomization, patients may undergo molecular tumor prescreening with the central laboratory *BRAF* mutation assay at any time prior to Screening as long as they meet all the Molecular Prescreening inclusion/exclusion criteria. Note that tumor samples previously determined to be *BRAF* wild-type by local assessment may be submitted to the central laboratory. In particular, tumors with clinicopathological features of *BRAF* mutations may be considered for testing by central laboratory regardless of the results of prior local *BRAF* mutation testing.

Safety Lead-in

The Safety Lead-in will be conducted at a limited number of sites. Dose-limiting toxicities will be evaluated and the tolerability of the binimetinib, encorafenib, and cetuximab combination will be assessed by the Sponsor and the Investigator in (approximately) weekly communications. The Data Monitoring Committee (DMC) will evaluate the safety data at pre-specified intervals and at additional points during the conduct of the Safety Lead-in, if necessary. The first 9 evaluable patients will be enrolled on a rolling basis in a single cohort to evaluate the combination of encorafenib 300 mg QD + binimetinib 45 mg BID + cetuximab 400 mg/m^2 followed by 250 mg/m^2 IV weekly. Additional patients will be enrolled based on assessments of the safety data by the DMC during the Safety Lead-in. Following study treatment discontinuation, patients who provide informed consent for survival follow-up will continue to be assessed to determine survival status until withdrawal of consent, patient is lost to follow-up, death or defined end of study. If informed consent cannot be obtained due to the patient being lost to follow-up or previous withdrawal of consent, attempts to determine survival status will be made via access to public records where permitted by local laws. The doses for the Triplet Arm in the randomized Phase 3 portion of the study will be determined after a total of 25-30 patients have been treated at the proposed doses and their data evaluated by the DMC.

In order to confirm the triplet dose is tolerable in Japanese patients, a separate Japanese Safety Lead-in cohort will be conducted in 6 patients at a limited number of sites in Japan. The starting dose in Japanese patients will be the dose assessed to be tolerable in the non-Japanese Safety

Lead-in. Similar to the Safety Lead-in for non-Japanese patients, the DMC will evaluate the safety data to confirm tolerability.

Phase 3

Once the tolerability of the proposed doses for the Triplet Arm has been established, the Phase 3 portion of the study will begin and eligible patients will be randomized in a 1:1:1 ratio to the Triplet Arm, Doublet Arm or Control Arm. The number of 3rd line patients (those having received 2 prior regimens in the advanced/metastatic setting) will be limited to 215 (35% of the total enrollment) after which only patients with 1 prior regimen will be randomized. Patients with 2 prior regimens who have entered Screening at the time that the limit has been reached will be permitted to continue into the study if they are otherwise determined to be eligible. Patients randomized to the Control Arm may be treated with either irinotecan + cetuximab or FOLFIRI + cetuximab as per Investigator's choice. The choice of irinotecan or FOLFIRI must be declared prior to randomization. Randomization will be stratified by baseline Eastern Cooperative Oncology Group performance status (ECOG PS; 0 vs. 1), prior use of irinotecan (yes vs. no) and cetuximab source (US-licensed vs. EU-approved). The DMC will review the available safety information after the first 30 patients in the randomized Phase 3 portion of the study (i.e., approximately 10 patients in each arm) have had the opportunity to complete at least 1 cycle of treatment to confirm tolerability. Similar to the Safety Lead-in for non-Japanese patients, the DMC will evaluate the safety data to confirm tolerability for SLI patients enrolled at Japanese sites.

Treatment will be administered in 28-day cycles until disease progression, unacceptable toxicity, withdrawal of consent, initiation of subsequent anticancer therapy or death. In special circumstances (defined in Section 4.5 of the protocol), continuation of treatment beyond disease progression may be allowed. After treatment is discontinued, randomized Phase 3 patients will continue to be followed for OS.

End of Study

End of study will be defined as the point when all patients have the opportunity to be followed for at least 1 year after the randomization date of the last patient enrolled **and** at least 80% of patients have an OS event (or are lost to follow-up). Any patients still receiving study drugs at the end of the study will be allowed to continue at the discretion of the Investigator and as long as none of the treatment discontinuation criteria are met. After the end of the study, access to study drugs will be provided in accordance with local regulations and requirements.

For additional information regarding eligibility criteria, planned dose levels, dose escalation rules, schedules of assessments, and other study design details and procedures, please refer to the protocol.

3.2 Sample Size Considerations

The planned sample size is approximately 646-651 patients (includes 31-36 patients for the non-Japanese and Japanese Safety Lead-in study portions and approximately 615 in the randomized Phase 3 portion of the study). Additional patients may be included in Safety Lead-in if lower dose levels are evaluated.

Safety Lead-In

During the Safety Lead-In, the first 9 evaluable patients will be enrolled to the starting dose of the Triplet Arm. This Triplet dose will be presented to the DMC as tolerable if the observed Cycle 1 DLT rate is < 33% (i.e., < 3 patients with DLTs out of 9 patients). Table 3-1 provides a comparison of the operating characteristics for this dose-escalation rule with 9 patients and traditional 3+3 rules. The results illustrate the benefit of the additional patients as the probability of falsely declaring a dose to be toxic is lowest with a 9-patient cohort when the true DLT rate is ≤20%. Similarly, the probability of correctly declaring a dose to be toxic is higher with the 9-patient cohort when the true DLT rate is ≥40%. In addition, observing no Cycle 1 DLTs in 9 patients would be expected to occur with probability 0.040 if the true DLT rate is 30%.

Table 3-1: Operating Characteristics of Safety Lead-In Criteria for 9 Patients Compared To 3+3 Rules

True Cycle 1 DLT Rate	Probability of dose declared toxic using 3+3 rules	Probability of observed Cycle 1 DLT rate ≥ 33% in 9 patients
10%	0.094	0.053
20%	0.291	0.262
30%	0.506	0.537
40%	0.691	0.768
50%	0.828	0.910

If the DMC determines the doses to be tolerable in the first 9 evaluable patients based on observing DLTs in < 33% of patients and evaluation of the overall toxicity profile, the Safety Lead-in will be expanded by an additional 16-21 patients. A total of 25-30 patients will complete the Safety Lead-in (the initial 9 plus the additional patients in the expansion) at the doses

proposed for the randomized Phase 3 portion of the study. If dose de-escalation is required during the Safety Lead-In, additional patients may be necessary.

As in the non-Japanese Safety Lead-in, a separate safety evaluation of Japanese patients will be conducted once the Japanese Safety Lead-In cohort completes 1 cycle of therapy. For this cohort, 6 patients will be enrolled at a limited number of sites in Japan, resulting in a total of 31-36 patients for the entire Safety Lead-In study portion.

Randomized Phase 3

Overall Survival

Original sample size considerations for the phase 3 portion are based on East® version 6.3.1 (Cytel). For protocol version 7, East® version 6.4 (Cytel) was used for the power calculations.

During the randomized Phase 3 portion of the study, eligible patients will be randomized 1:1:1 such that approximately 205 patients will receive binimetinib + encorafenib + cetuximab (Triplet Arm), approximately 205 patients will receive encorafenib + cetuximab (Doublet Arm) and approximately 205 patients will receive irinotecan/cetuximab or FOLFIRI/cetuximab (Control Arm). The number of 3rd-line patients (those having received 2 prior regimens) randomized to the Phase 3 portion of the study will be limited to 215 (35% of the total randomized) in order to ensure generalizability of the study. Patients with 2 prior regimens who have entered Screening at the time that the limit has been reached will be permitted to continue into the study if they are otherwise determined to be eligible. Permitting up to 35% of the randomized patients to have received 2 prior regimens ensures a sufficient sample size to adequately generalize the results across subgroups when incorporating other relevant stratification factors. Randomization will be stratified by ECOG PS (0 vs. 1), prior use of irinotecan (yes vs. no), and cetuximab source (US-licensed vs. EU-approved).

The primary endpoints are OS and ORR of the Triplet Arm vs. Control Arm. Historical evidence (Peeters et al. 2014) suggests that the median OS of FOLFIRI combined with another EGFR inhibitor (panitumumab) will be approximately 5 months in *BRAF*-mutant mCRC. Existing published data suggest only modest differences between irinotecan single-agent and FOLFIRI efficacy (Clarke et al. 2011). In addition, available data suggests that treatment-eligible patients in the 2nd and 3rd-line settings and with or without prior irinotecan have similar PFS and OS with standard therapy approaches (DeRoock et al. 2010, Morris et al. 2014, Peeters et al. 2014). Based on these findings, it is assumed that both control arm options will have an approximate median OS of 5 months.

The number of patients required for the randomized Phase 3 portion of the study is driven by the key secondary endpoint of the OS of the Doublet Arm vs. Control Arm. For this comparison, the study is powered to detect an improvement of 2.1 months (7.1 months vs. 5 months; HR=0.70). With 338 OS events, the study will have approximately 90% power to detect this improvement using a group-sequential design and 1-sided α=0.025. Assuming accrual to the randomized Phase 3 portion of the study increases over a period of time before reaching a maximum of 25 patients per month (for an accrual duration of approximately 25 months) and 5% loss to follow-up, approximately 615 patients will be randomized to reach 338 events. With 338 events, the null hypothesis that the OS in the Doublet Arm is the same or worse than the Control Arm will be rejected if the hazard ration (HR) is smaller (i.e., better) than 0.808.

The final analysis for OS will occur once at least 268 events are observed in the Triplet Arm + Control Arm and at least 338 events are observed in the Doublet Arm + Control Arm. This is expected to occur approximately 33 months after the first patient is randomized (i.e., approximately 8 months after randomization is complete). At the time of the final analysis, it is anticipated that approximately 333 OS events will be observed in the Triplet Arm + Control Arm, but only the first 268 OS events will be included in the primary analysis. With 268 events, there will be approximately 90% power to detect a HR=0.67 (median OS of 7.5 months vs. 5 months). The null hypothesis that the OS in the Triplet Arm is the same or worse than the Control Arm will be rejected if the HR is smaller (i.e., better) than 0.79.

Overall Response

During the initial analysis of the study, the primary analysis of Triplet vs. Control ORR by BICR will be conducted based on the first 330 randomized patients. Based on historical data, the ORR in the control arm is expected to be approximately 10% (Kopetz et al. 2017; Seymour et al. 2013). A sample size of 110 patients per arm provides 88% power to detect a 20% absolute difference in ORR, assuming an ORR of 10% in the Control Arm and an ORR of 30% in the Triplet Arm at a 1-sided alpha of 0.005.

4.0 STUDY ENDPOINTS

4.1 Primary Endpoints

- Safety Lead-In:
  - Incidence of dose-limiting toxicities (DLTs)
  - Incidence and severity of adverse events (AEs), graded according to the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE), version 4.03 (v.4.03), and changes in clinical laboratory parameters, vital signs, electrocardiograms (ECGs), echocardiogram (ECHO)/multi-gated acquisition (MUGA) scans and ophthalmic examinations
  - Incidence of dose interruptions, dose modifications and discontinuations due to AEs
- Randomized Phase 3:
  - Confirmed ORR (by BICR)) per RECIST, v1.1 of Triplet Arm vs. Control Arm
  - OS, defined as the time from randomization to death due to any cause, of Triplet Arm vs. Control Arm

4.2 Key Secondary Endpoints

- Randomized Phase 3
  - OS of Doublet Arm vs. Control Arm

4.3 Secondary Endpoints

- Safety Lead-In:
  - ORR (by BICR and Investigator) per the Response Evaluation Criteria in Solid Tumors (RECIST), version 1.1 (v1.1), 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
  - PFS (by BICR and Investigator), defined as the time from first dose to the earliest documented disease progression or death due to any cause
  - DOR (by BICR and Investigator), defined as the time from first radiographic evidence of response to the earliest documented disease progression or death due to underlying disease
  - Time to response (by BICR and Investigator), defined as the time from first dose to first radiographic evidence of response

- PK parameters of encorafenib, cetuximab, binimetinib and the active metabolite of binimetinib (AR00426032)

o Randomized Phase 3:

- Confirmed ORR (by Investigator) per RECIST, v1.1 of Triplet Arm vs. Control Arm
- Confirmed ORR (by BICR and Investigator) per RECIST, v1.1 of Doublet Arm vs. Control Arm
- PFS (by BICR and Investigator), defined as the time from randomization to the earliest documented disease progression or death due to any cause, of Triplet Arm vs. Control Arm
- PFS (by BICR and Investigator) of Doublet Arm vs. Control Arm
- OS of Triplet Arm vs. Doublet Arm
- Confirmed ORR (by BICR and Investigator) per RECIST, v1.1 of Triplet Arm vs. Doublet Arm
- PFS (by BICR and Investigator) of Triplet Arm vs. Doublet Arm
- DOR (by BICR and Investigator) of Triplet Arm vs. Control Arm, of Doublet Arm vs. Control Arm and of Triplet Arm vs. Doublet Arm
- Time to response (by BICR and Investigator), defined as the time from randomization to first radiographic evidence of response, of Triplet Arm vs. Control Arm, of Doublet Arm vs. Control Arm and of Triplet Arm vs. Doublet Arm
- Incidence and severity of AEs, graded according to NCI CTCAE, v 4.03, and changes in clinical laboratory parameters, vital signs, ECGs, ECHO/MUGA scans and ophthalmic examinations
- Change from baseline in the European Organization for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire for Cancer Patients (QLQ-C30), Functional Assessment of Cancer Therapy-Colon Cancer (FACT-C), EuroQol-5D-5L (EQ-5D-5L), and Patient Global Impression of Change (PGIC) of Triplet Arm vs. Control Arm, of Doublet Arm vs. Control Arm and of Triplet Arm vs. Doublet Arm
- Model-based PK parameters of encorafenib, cetuximab, binimetinib and the active metabolite of binimetinib (AR00426032)
- Model-based PK assessment of drug-drug interactions between encorafenib, cetuximab, binimetinib and the active metabolite of binimetinib (AR00426032)

4.4 Exploratory Endpoints

o Safety Lead-In:

- OS, defined as the time from first dose to death due to any cause

[redacted]

5.0 ANALYSIS SETS

5.1 Full Analysis Set

For patients in the Safety Lead-in, the Full Analysis Set (FAS) will consist of all patients who receive at least 1 dose of study drug and have at least 1 post-treatment assessment, which may include death.

For patients in the Phase 3 portion, the FAS will consist of all randomized Phase 3 patients. Patients will be analyzed according to the treatment arm and stratum they were assigned to at randomization.

5.2 Safety Set

The Safety Set (SS) will consist of all patients who receive at least 1 dose of study drug and have at least 1 post-treatment assessment, which may include death. Patients will be analyzed according to treatment received.

Patients who received the wrong study treatment for only a part of the treatment period will be analyzed according to the randomized treatment. Patients who received a wrong study treatment (i.e. a treatment different from the one assigned by randomization) during the whole treatment period will be analyzed according to the actual treatment received.

5.3 Dose-Determining Set

The dose-determining set (DDS) includes all Safety Lead-In patients from the safety set who either completed a minimum exposure requirement and have sufficient safety evaluations or experienced a dose limiting toxicity (DLT).

A patient is considered to have met the minimum exposure requirement if having received at least 75% dose intensity ([Cumulative administered dose in mg in Cycle 1/Cumulative planned dose in mg in Cycle 1] x 100) of the planned dose for each of the three agents; binimetinib (i.e., 1890 of the planned 2520 mg dose), encorafenib (i.e., 6300 of the planned 8400 mg dose) and cetuximab doses (i.e., 750 of the planned 1000 mg/m2 dose) within the first Cycle of dosing. The length of a cycle is 28 days.

Patients who do not experience a 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.

5.4 Safety Lead-in Efficacy Set

The Safety Lead-in Efficacy Set will consist of all Safety Lead-in patients in the FAS who were identified at screening as having a *BRAF*^{V600E} mutation (per local or central testing).

5.5 Phase 3 Response Efficacy Set

The Phase 3 Response Efficacy Set will consist of the first 330 patients randomized into the Phase 3 portion of the study.

5.6 Per Protocol Set

The Per-Protocol Set (PPS) consists of all Phase 3 patients from the FAS without any major protocol deviations (or other criteria that could largely impact efficacy results) and who received at least one dose of study medication.

The reasons that will lead to exclusion of patients from the PPS are listed below:

- No histologically or cytologically confirmed CRC that is metastatic;
- Not positive for *BRAF* V600 mutation per central assessment;
- Prior treatment with any RAF inhibitor, MEK inhibitor, cetuximab, panitumumab or other epidermal growth factor receptor (EGFR) inhibitor;
- Baseline ECOG PS greater or equal to 3 (i.e. at least 2 categories worse than the defined inclusion criteria);
- Study treatment received different from treatment assigned by randomization [patients in randomized Phase 3 only].

5.7 Pharmacokinetic Set

The PK set will include all patients in the SS who had at least 1 post-dose blood collection for PK with associated bioanalytical results. Patients will be analyzed according to the actual treatment and dose received.

6.0 CHANGES FROM THE STUDY PROTOCOL

This SAP incorporates the following changes from the statistical analyses described in the study protocol:

- N/A

7.0 STATISTICAL METHODS

7.1 Reporting Conventions and Definitions

7.1.1 Reporting Conventions

Durations of events (e.g., duration of treatment) will be calculated in days as (stop date – start date + 1). Conversions to weeks, months, years will be days/7, days/30.4375, days/365.25, respectively. Handling of missing or partial dates will depend on the specific event and will be described in subsequent sections.

In general, missing values will be handled as follows unless otherwise specified. For continuous variables at Baseline, missing values will be excluded from calculation of summary statistics, and the number and percent of patients with missing values will be displayed. For categorical values at Baseline, the number and percent of patients with a missing value will be displayed. For missing post-Baseline values, the method for reporting missing values will depend on the summary table.

All pre-treatment and safety tables will be presented by the following 5 treatment groups:

ARRAY-818-302 Randomized and Safety Lead-in (Pooled)	**ARRAY-818-302 Safety Lead-in**	**ARRAY-818-302 Randomized Portion**		
Enco+Cetux+Bini	**Enco+Cetux+Bini**	**Enco+Cetux+Bini**	**Enco+Cetux**	**Control**

All SLI efficacy tables, will be presented by the following 3 groups:

ARRAY-818-302 Safety Lead-In (SLI)-		
Non-Japan SLI	**Japan SLI**	**SLI Total**

All Phase 3 efficacy tables, unless otherwise specified will be presented by the following 3 treatment groups:

ARRAY-818-302 Randomized Portion		
Enco+Cetux+Bini	**Enco+Cetux**	**Control**

Data listings will be sorted by treatment group, patient identifier, parameter and the corresponding date of assessment.

7.1.2 Definitions

Study drug will refer to encorafenib, binimetinib, cetuximab, 5-FU, FA, or irinotecan.

Study treatment will refer to the Triplet Combination (binimetinib + encorafenib + cetuximab), Doublet Combination (encorafenib + cetuximab) or Control (irinotecan + cetuximab or FOLFIRI + cetuximab).

Treatment start date is defined as the first date a non-zero dose of study drug was received.

For safety assessments, *study day* is defined in the following manner:

- On or after the start date of study treatment: (date of assessment) – (treatment start date) + 1. Study day 1 will therefore be the first day of study treatment.
- Before the start date of study treatment: (date of assessment) – (treatment start date).

For non-safety assessments, *study day* is defined in the following manner:

- On or after the start date of study treatment: (date of assessment) – (randomization rate) + 1. Study day 1 will therefore be the day of randomization.
- Before the start date of study treatment: (date of assessment) – (randomization rate).

For safety assessments, *Baseline* is defined as the last assessment prior to the *treatment start date/time*. If an assessment is planned to be performed prior to the first dose of study drug in the protocol and the assessment is performed on the same day as the first administration of study drug and the time is unknown, it will be assumed that it was performed prior to study drug administration. Unscheduled assessments will be used in the determination of baseline. Data reported at the End of Treatment (EOT) visit are not eligible for baseline selection.

For non-safety assessments, *Baseline* is defined as the last assessment prior to (or on the day of, for ambiguous cases) randomization.

For analyses of assessments at a specified study visits, each study visit is defined using the protocol-specified window for the assessment.

7.2 Patient Disposition

Enrollment and the extent of participation in the study will be summarized by treatment group (5 groups). The number and percent of patients randomized will be presented, the number and percent of patients in each analysis set, the number and percent of patients who discontinued study treatment and the reason for treatment discontinuation will be presented.

Number of patients screened and randomized will be summarized respectively by region, country, center, and treatment arm. Number of patients randomized will also be summarized by stratification factor [ECOG PS (0 vs. 1), prior use of irinotecan (yes vs. no), cetuximab source (US-licensed vs EU-approved)] and treatment arm, based on the IVRS data used for randomization. In addition, separate by-treatment summaries will be generated for the subgroups of patients in Japan and in Korea.

Patient disposition data will be provided in a data listing. A separate listing will describe each patient's inclusion or exclusion status for each of the analysis sets defined in Section 5.0.

7.3 Protocol Deviations

Protocol deviations will be determined and documented prior to database lock. Frequency counts and percentages of patients in the FAS with any protocol deviations will be tabulated by the deviation type, category and treatment arm. Protocol deviations leading to exclusion of patients from the PPS will be tabulated separately by treatment arm.

7.4 Patient Characteristics

7.4.1 Demographics and Pretreatment Characteristics

All demographic and pretreatment characteristics will be presented in by-patient listings by treatment group using the FAS. All summary tables will be presented by treatment group.

Demographic and Baseline patient characteristics that are only measured or assessed prior to the first dose of study treatment will be summarized using descriptive statistics (n, mean, standard deviation, median, minimum, maximum for continuous variables; n, percent for categorical variables); however, no formal comparisons between treatment groups will be performed.

MSI status at baseline will also be summarized by treatment arm. MSI status will be determined by Immunohistochemistry (IHC) for the SLI patients and by PCR for the phase 3 randomized patients.

Patients with CEA levels above the upper limit of normal (ULN) at baseline, as well as patients with CA19.9 levels above the ULN at baseline, will be summarized by treatment arm.

In addition, separate by-treatment summaries will be generated for the subgroups of patients in Japan and in Korea.

7.4.2 Medical History, Baseline Signs and Symptoms

Medical history and Baseline signs and symptoms will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) terminology and presented in data listings and summarized by treatment group using the FAS.

7.4.3 Prior Anticancer Therapy

Prior anticancer treatments will be coded using World Health Organization (WHO) Drug Dictionary (Uppsala Monitoring Centre, Uppsala, Sweden) and presented in a data listing [by Anatomical Therapeutic Chemical (ATC) code (levels 2 and 4)]. The number of prior regimens (if prior treatments are grouped by regimen) and the types of prior treatments (e.g. medications, radiotherapy, surgery, etc.) will be summarized by treatment group using the FAS.

In addition, separate by-treatment summaries will be generated for the subgroups of patients in Japan and in Korea.

7.4.4 Other Prior Medications

Other prior medications will also be coded using WHO Drug Dictionary and presented in a data listing and summarized by Anatomical Therapeutic Chemical (ATC) code (levels 2 and 4) and treatment group using the FAS.

7.5 Efficacy Analysis

Efficacy analyses for the SLI patients will be performed using the SLI Efficacy Set and results will be presented by group (non-Japan SLI, Japan SLI, Total SLI). The primary analysis of the Phase 3 primary endpoint of Triplet vs Control ORR will be analyzed using the Phase 3 Response Efficacy Set. Unless otherwise stated, other efficacy analyses for Phase 3 patients will be performed using the FAS.

The primary efficacy endpoints are ORR (by BICR; Triplet Arm, vs. Control Arm) and OS (Triplet Arm vs. Control Arm).

The key secondary efficacy endpoint is OS (Doublet Arm vs. Control Arm) and other secondary endpoints include ORR (by Investigator; Triplet Arm vs Control Arm); ORR (by BICR and Investigator; Doublet Arm vs. Control Arm), and PFS (by BICR and Investigator; Triplet Arm vs. Control Arm and Doublet Arm vs. Control Arm), OS (Triplet Arm vs. Doublet Arm), ORR (by BICR and by Investigator; Triplet Arm vs. Doublet Arm), DOR (by BICR and by Investigator; Triplet Arm vs. Control Arm, Doublet Arm vs. Control Arm, Triplet Arm vs. Doublet Arm), and time to response (by BICR and by Investigator; Triplet Arm vs. Control Arm, Doublet Arm vs. Control Arm, Triplet Arm vs. Doublet Arm).

The Type I error rate for the primary endpoints will be controlled using a fallback procedure described by Wiens and Dmitrienko (2005). One-sided alpha of 0.005 will be assigned to the Triplet vs Control ORR endpoint. The remaining 0.020 will be assigned to the Triplet vs Control OS endpoint. If the p-value of the Triplet vs Control comparison of ORR at the primary analysis is <0.005, then the Triplet vs Control OS comparison will be assigned a total 1-sided alpha of 0.025. Otherwise, it will remain at 1-sided 0.020.

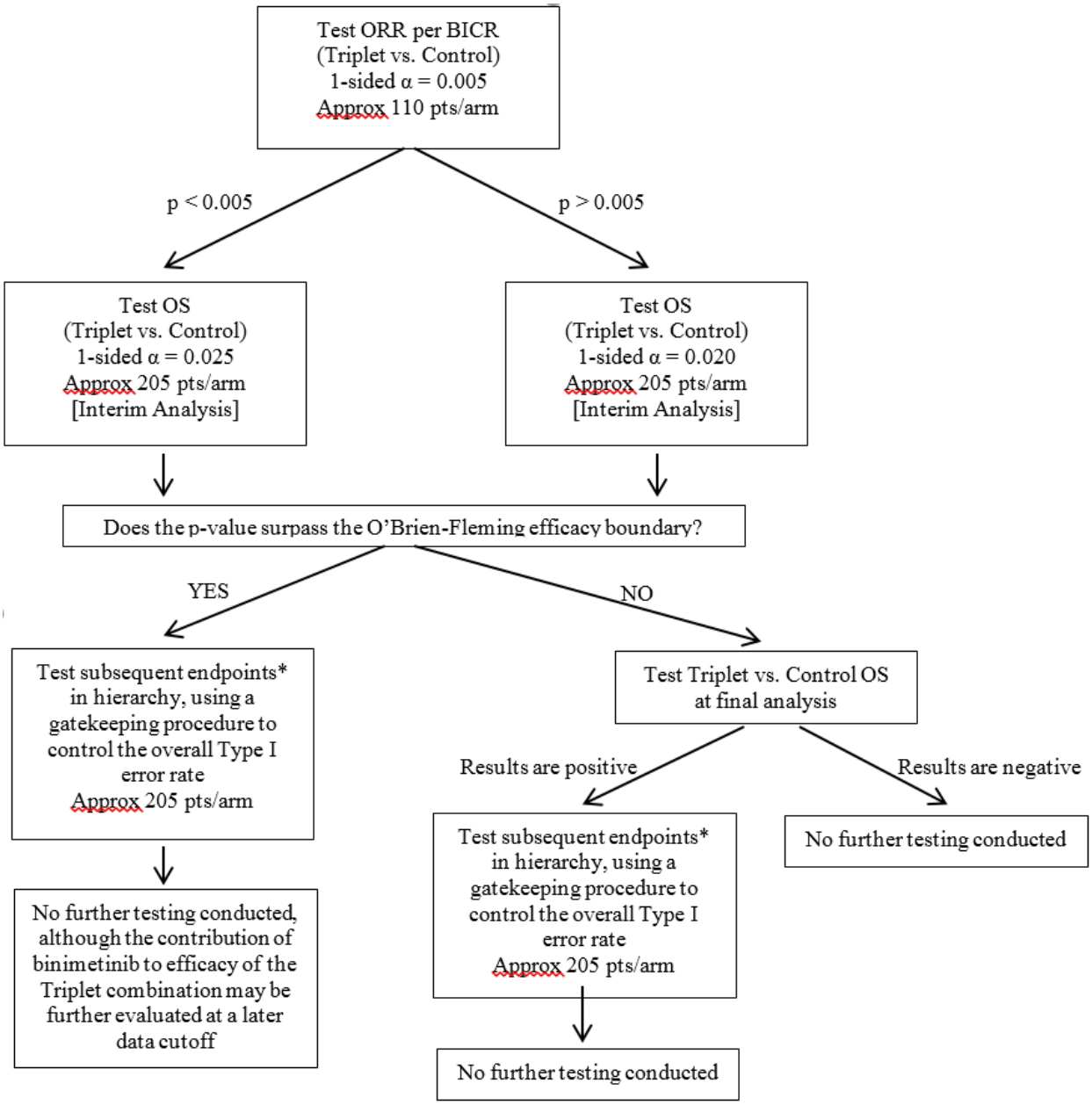
The key secondary endpoint and several secondary endpoints may also be formally tested. To control the overall Type I error rate, a gatekeeping procedure using hierarchical testing will be performed. If the OS of the Triplet Arm vs. Control Arm is found to be significant at either the interim analysis or the final analysis, the following tests will then be conducted sequentially, each at the same total alpha assigned to the Triplet vs Control OS endpoint:

1. OS of Doublet Arm vs. Control Arm
2. ORR (per BICR) of Doublet Arm vs. Control Arm
3. PFS (per BICR) of Triplet Arm vs. Control Arm
4. PFS (per BICR) of Doublet Arm vs. Control Arm

If any of the above tests is found to not be statistically significant, all subsequent comparisons will only be summarized using descriptive statistics, including nominal p-values.

The overall testing strategy of the study is summarized in Figure 7-1.

Figure 7-1: Testing Strategy for ARRAY-818-302



* Subsequent endpoints would be tested in the following order: Doublet vs. Control OS, Doublet vs. Control ORR per BICR, Triplet vs. Control PFS per BICR, and then Doublet vs. Control PFS per BICR.

7.5.1 Objective Response Rate

The overall best response (i.e., CR or PR) as assessed by BICR per RECIST v1.1 will be determined for each patient. The ORR will be calculated within each treatment arm, where ORR

is defined as the number of patients achieving an overall best response of CR or PR divided by the total number of patients in that treatment arm.

The ORR will be tested for the primary endpoint of Triplet Arm vs. Control Arm based on the Phase 3 Response Efficacy Set and using the Cochran-Maentel-Haenzel test at a one-sided alpha of 0.005. Both confirmed and unconfirmed ORR will be summarized but, for purposes of formal testing, the analysis of the confirmed responses will be used. The stratification factors used in the test will be those used for randomization and will be based on the actual randomization (IWRS) information. For the primary analysis, ORR will be presented by arm and stratum, along with a 95% and 99% CI.

The secondary ORR endpoints will be analyzed in a similar manner based on the Phase 3 Response Efficacy Set and the FAS.

7.5.2 Overall Survival

For the Phase 3 portion of the study, the primary and key secondary efficacy endpoints are OS, defined as the time from randomization to death due to any cause. Patients who do not have a death date by the data cutoff date will be censored for OS at their last contact date. Overall survival will be calculated for all patients in the FAS and summarized by treatment arm using the Kaplan-Meier (KM) method. In the SLI portion of the study, OS will be analyzed as an exploratory endpoint using the Safety Lead-in Efficacy Set and is defined as the time from first dose of study drug to death due to any cause. Patients who are alive as of the data cutoff date will be censored for OS at their last contact date.

Overall treatment arm estimates as well as treatment arm estimates by stratum will be provided for the Phase 3 randomized patients.

For the OS primary endpoint, the null hypothesis of the primary objective is that the OS for the triplet combination is less than or equal to the OS of the control, i.e.,

$$H_0: S_{OS,A}(t) \leq S_{OS,C}(t)$$

where $S_{OS,A}(t)$ is the OS survival distribution function for the triplet combination and $S_{OS,C}(t)$ is the OS survival distribution function for the control arm. The null hypothesis will be tested using a stratified log-rank test against the alpha assigned to the endpoint based on the Fallback approach. The stratification factors used in the test will be precisely those used for randomization, and will be based on the actual randomization (interactive web response system [IWRS]) information.

The distribution of OS will be described in tabular and graphical format by treatment group (4 groups) using KM methods, reporting estimated median (in months) with 95% CI, 25th and 75th percentiles (Brookmeyer and Crowley 1982, Klein and Moeschberger 1997) and KM estimated probabilities with corresponding 95% CIs (Kalbfleisch and Prentice 2002) at several time points (including at least 2, 4, 6, 8, 10, 12, and 14 months).

For the Phase 3 randomized part, a Cox regression model stratified by randomization stratification factors will be used to estimate the hazard ratio (HR) of OS, along with 95% CI based on the Wald test. The PHREG procedure in SAS with option TIES=EXACT will be used to fit the Cox proportional hazards model.

At the initial analysis of the study, an interim analysis for superiority or (non-binding) futility of the Triplet vs Control OS endpoint will be performed based on all available data (i.e., using the FAS). Futility and superiority boundaries for both the OS interim and final analyses will be determined using a Lan-DeMets spending function (Lan and DeMets, 1983) that approximates O'Brien-Fleming stopping boundaries (see Section 7.11).

For the purposes of the final analysis of the primary endpoint (i.e. Triplet Arm vs Control Arm OS), only the first 268 OS events will be included in the analysis.

The same method of analysis and supportive analyses as for the primary endpoint will be used for the key secondary OS endpoint.

In addition to the primary analysis, the analyses for OS will also be conducted at the time of the end of study.

7.5.2.1 OS Sensitivity Analyses

The following sensitivity analyses will be conducted to support the analyses of OS. Nominal p-values will be displayed for the supportive and sensitivity analyses, so these should be considered as descriptive only.

The analyses of OS will be repeated using the PPS.

As the Triplet Arm and Control Arm will likely have more than 268 combined OS events by the time the required number of OS events are observed in the Double Arm and Control Arm, the OS analysis of the triplet versus control will be repeated using all available OS events.

The distribution of OS in the FAS will be compared between treatment arms using an unstratified log-rank test and the HR (with associated 95% confidence interval) resulting from an unstratified Cox model will be presented.

7.5.2.2 Other Supportive Analyses for OS

For the randomized Phase 3 part, the effect of potential prognostic factors will be investigated by using multivariate Cox regression. In addition to covariates for the stratification factors considered for randomization (ECOG status, prior irinotecan, cetuximab source), the following factors measured at baseline might be included as covariates:

1. Gender (male vs. female)
2. Age (continuous)
3. Region (North America, Europe, Rest of the World)
4. MSI status (high vs. stable)
5. Removal status of primary tumor (no resection, partial resection, complete resection)
6. C-reactive protein (CRP) baseline level (continuous)
7. Side or tumor (left vs right)
8. Number of metastatic sites based on Target and Non-target lesion assessment (1 vs 2+)
9. Presence of liver metastases at baseline, based on Target and Non-target lesion assessment (yes vs. no)

To avoid model instabilities, the above categorized covariates will only be included in the model if the number of observations allows i.e. if there are at least 10 events in each category for binary covariates.

Subgroup analyses will be performed for each of the three baseline stratification factors and other relevant baseline variables provided the number of patients randomized with these particular covariates allows (i.e. at least 10 events are available in the considered sub-group). Section 7.13 describes the subgroups that will be considered. The analyses will include KM summaries, and HRs (together with associated 95% confidence interval) from unstratified Cox models. A forest plot representation will be provided.

In addition, for the entire FAS population, a reverse KM analysis will be performed for both OS to estimate the median duration of potential follow-up as described by Schemper and Smith (1996). A plot of the censoring distributions will also be provided. Patients who had an OS event for the purposes of the OS analysis will be censored at the date of the event. Patients who were

censored (e.g. lost to follow-up, withdrew consent, ongoing, etc.) in the OS analysis will be considered as events for the purposes of estimating duration of potential follow-up. The same duration of time values used in the OS analysis will be used in this analysis. KM estimated probabilities with corresponding 95% CIs (Kalbfleisch and Prentice 2002) will be presented at the same time points as in the OS analysis.

Furthermore, the pattern of censored data will be examined between the treatment arms by summarizing the number of OS observations that are censored and by tabulating the reasons for censored observations. Censoring reasons will also be listed.

7.5.3 Efficacy Analyses of Secondary Endpoints

7.5.3.1 Progression-free Survival

Progression-free survival (PFS) is defined as the time from randomization to the earliest documented progression date or death due to any cause. Progression-free survival will be calculated for all patients in the FAS and summarized by treatment arm using the KM method. Both, PFS data determined by BICR and by Investigator will be analyzed. Overall treatment arm estimates as well as treatment arm estimates by stratum will be provided. PFS in the subgroups of patients in Japan and in Korea will be examined separately using KM plots.

For the primary PFS analysis, disease progression and death (from any cause) will be considered as events. If death or disease progression is not observed, the PFS will be censored at the date of last adequate tumor assessment (i.e., at the date of last tumor assessment of complete response [CR], partial response [PR] or stable disease) prior to cutoff date or date a subsequent therapy is started (e.g., systemic therapy, surgery, radiotherapy). However, if a PFS event is observed after more than 1 missing or inadequate tumor assessment, PFS will be censored at the last adequate tumor assessment. If a PFS event is observed after a single missing or non-adequate tumor assessment, the actual date of event will be used.

Censoring rules to be applied to the PFS endpoint are described in Table 7-1.

In addition to the primary analysis, the analyses for PFS will also be conducted at the time of the end of study.

Progression after next line of therapy (PFS2) will be calculated for patients who receive subsequent anticancer therapy. PFS2 is defined as time from randomization (on the ARRAY-818-302 study) to PD or death after the start of subsequent anticancer therapy; patients alive and without progression on their subsequent anticancer therapy will be censored at last known

survival. PFS2 will be summarized by treatment arm (as randomized on the ARRAY-818-302 study) using the KM method.

Table 7-1: Options for PFS Analysis

	Situation	Event Date	Outcome
A[a]	No baseline assessment	Date of randomization	Censored
B	Progression or death at or before next scheduled assessment	Date of progression (or death)	Progressed
C1	Progression or death after exactly one missing assessment	Date of progression (or death)	Progressed
C2	Progression or death after two or more missing assessments	Date of last adequate tumor assessment*	Censored
D	No progression	Date of last adequate tumor assessment*	Censored
E	Treatment discontinuation due to 'Disease progression' without documented progression, i.e. clinical progression based on investigator claim	N/A (not considered as an event, patient without documented PD should be followed for progression after discontinuation of treatment)	Information ignored
F	New antineoplastic therapy given	Date of last adequate tumor assessment*	Censored

[a] The rare exception to this is if the patient dies no later than the time of the second scheduled assessment as defined in the protocol in which case a PFS event at the date of death is counted

* tumor assessment with non-missing and non-unknown overall lesion response

7.5.3.2 Determination of Missing Adequate Tumor Assessments

The term 'adequate Tumor Assessments (TA)' is defined as TA with overall lesion response different from missing and 'Unknown'. The term 'missing adequate TA' is defined as TA not done or TA with overall lesion response equal to 'Unknown'. For the sake of simplicity, the 'missing adequate TA' will also be referred as 'missing TA'. An exact rule to determine whether there is one or two missing TAs is therefore needed.

As per protocol, tumor assessments are expected to be performed every 6 weeks (±7 days) for the first 24 weeks and every 12 weeks (±7 days) thereafter.

The threshold D_1 will be defined as the protocol specified interval between the tumor assessments plus the protocol allowed window around the assessments. Similarly, the threshold D_2 is defined as two times the protocol specified interval between the tumor assessments plus the protocol allowed window around the assessments. If the distance is larger than threshold D_1 or D_2 then the analysis will assume one or two missing tumor assessments, respectively.

Therefore:

- if the last adequate tumor assessment is performed on or before study day 21 then D_1=49 days and D_2=91 days;
    - This accounts for the scenario where the last adequate tumor assessment is the baseline assessment. This also allows for the rare instance where an additional assessment occurred early in the study.
- if the last adequate tumor assessment is performed between study day 22 and study day 105 then D_1= 56 days and D_2=98 days;
    - This accounts for the scenario where both of the next two scheduled assessments occur prior to week 24.
- if the last adequate tumor assessment is performed between study day 106 to study day 147 then D1=56 days and D2=140 days;
    - This accounts for the scenario where next scheduled assessment occurs prior to week 24, but the second subsequent assessment occurs after week 24.
- otherwise, when the last adequate tumor assessment is performed after study day 148, then D_1=98 days and D_2=182 days.
    - This accounts for the scenario where both scheduled assessments occur after week 24.

Using the D_2 definition above, an event is censored as occurring after ≥2 missing tumor assessments if the distance between the last adequate tumor assessment date and the PFS event date is larger than D_2.

7.5.3.3 **PFS Sensitivity Analyses**

The following sensitivity analyses will be conducted to support the analyses of PFS. Nominal p-values will be displayed for the supportive and sensitivity analyses, these should be considered as descriptive only.

- The analyses for PFS will also be repeated for the PPS.

- The distribution of PFS in FAS will be compared between the treatment arms using an unstratified log-rank test and the HR (together with associated 95% confidence interval) resulting from an unstratified Cox model will be presented.

- The analyses for PFS will be repeated with a censoring rule that includes a PFS event even if the event is recorded after 2 or more missing tumor assessments (i.e. counting events for line C2 of Table 7-1)

- The analyses for PFS will be repeated with a censoring rule that backdates events occurring after one or more missing tumor assessments (i.e. backdating events for line C1 and C2 of Table 7-1).

- If at least 15 patients have tumor assessments which occurred after the start of another anticancer therapy, an additional sensitivity analysis for PFS will be performed to include these tumor assessments (Table 7-1 situation E).

7.5.3.4 Other Supportive Analyses for PFS

Similar to OS, the effect of potential prognostic factors will be investigated by using multivariate Cox regression. The same covariates will be used as the OS analysis. The same subgroup analyses will be performed as the OS analysis.

The censoring distribution for PFS will be summarized using the same methods as OS.

To assess the balance of the timing of the tumor assessment between the treatment arms, the timing of all tumor assessments will be depicted graphically by treatment arm.

7.5.3.5 Response Rate Sensitivity Analyses

The following sensitivity analyses will be conducted to support the analyses of ORR. Nominal p-values will be displayed for the supportive and sensitivity analyses, so these should be considered as descriptive only.

For the Phase 3 patients, analyses of ORR will be repeated for patients in the Phase 3 Response Efficacy Set who had measurable disease at baseline.

The ORR will also be compared between treatment arms using an unstratified Chi-squared test.

7.5.3.6 Other Supportive Analyses for Response Rate

For Phase 3 patients, using the Phase 3 Response Efficacy Set, a multivariate logistic regression model that includes the same covariates as the OS supportive analysis will be used to further assess the response rate. The same subgroup analyses will also be performed.

7.5.4 Duration of Response

Duration of response is defined as the time from first radiographic evidence of response to the earliest documented progressive disease (PD) or death, and is calculated for responders only. Responders who do not have a PD or death date by the data cutoff date will be censored for DOR at their last adequate radiological assessment (i.e., at the date of last tumor assessment of CR, PR or SD) prior to cutoff date or date a subsequent anticancer therapy for mCRC is started. Duration of response will be summarized by arm based on both the Phase 3 Response Efficacy Set and the FAS using the KM method in a tabular format.

The proportion of patients with a DOR of 6 months or more will be summarized by treatment arm.

Both DOR based on BICR and on Investigator's assessment will be summarized secondary endpoints.

7.5.5 Time to Response

Time to response (TTR) is the time between date of randomization until first documented response of CR or PR. Confirmed responses and unconfirmed responses will be summarized separately. Patients who do not achieve a PR or CR will be censored as follows:

- Censored at the last adequate tumor assessment date when they do not have a PFS event. In this case patients have not yet progressed so they theoretically still have a chance of responding;
- Censored at maximum follow-up (i.e. FPFV to LPLV used for the analysis) when they have a PFS event (i.e. progressed or died due to any cause). In this case, the PFS event is the worst possible outcome as it means the patient cannot subsequently respond. Since the statistical analysis makes use of the ranking of times to response it is sufficient to assign the worst possible censoring time which could be observed in the study which is equal to the maximum follow-up time (i.e. time from FPFV to LPLV).

TTR will be described in tabular by treatment arm based on both the Phase 3 Response Efficacy Set and the FAS using KM methods including estimated median (in months) with 95% CI, 25th and 75th percentiles (Brookmeyer and Crowley 1982, Klein and Moeschberger 1997) and KM estimated probabilities with corresponding 95% CIs (Kalbfleisch and Prentice 2002) at several timepoints (including at least 2, 4, 6, 8, 10 and 12 months).

Using the same summaries described above, TTR for responders only (i.e. patients achieving at least one CR or PR) will also be presented.

The primary analysis of time to response will be based on the assessments per BICR. Time to response based on Investigator's assessment will also be summarized as a supportive analysis.

No formal statistical test will be performed.

7.6 Assessment of the Binimetinib Contribution to Efficacy

In order to adequately assess the efficacy benefit that binimetinib contributes to the triplet combination, comparisons of the Triplet Arm to the Doublet Arm will be conducted for each efficacy endpoint using the FAS.

A comparison of the OS distributions for the Triplet Arm and the Doublet Arm will be described in tabular and graphical format by treatment arm using KM methods. A Cox regression model stratified by randomization stratification factors will be also used to estimate the HR of OS, along with 95% CI based on the Wald test. Nominal p-values will be provided.

In addition to the response rates reported by treatment arms with exact binomial confidence intervals, the Cochran-Mantel-Haenszel test will be used to compare ORR between the Triplet Arm and the Doublet Arm. The stratification factors used in the test will be those used for randomization, and will be based on the actual randomization (IWRS) information. Nominal p-values will be provided. Separate summaries for confirmed responses and all responses (confirmed and unconfirmed) will be generated.

The PFS of the Triplet Arm and the Doublet Arm will be summarized using the KM method. Nominal p-values will be provided. TTR and DOR of the two arms will also be summarized using the KM method.

The ORR and PFS analyses will be conducted on both BICR and Investigator tumor assessment data. In addition to the primary analysis, the analyses described here will also be conducted at the time of the end of study.

7.7 Safety Analysis

All safety analyses will be performed using the appropriate data for all patients in the SS.

Key safety analyses will also be conducted at the time of the end of study.

7.7.1 Extent of Study Drug Exposure

Definitions:

Definitions as well as intermediate calculations of duration of exposure, cumulative dose, average daily dose, actual dose intensity (DI), relative dose intensity (RDI) are provided below.

These definitions will be applied for each study drug individually. For drugs administered by IV, the dose in mg will be taken from the CRF.

Duration of exposure

This is the duration that the patient was exposed to the drug and is summarized in the exposure tables. The duration will be reported in weeks.

For daily dosing of binimetinib and encorafenib:

- duration of exposure = [[Date of last (non-zero) dose of study drug] – [date of first dose of study drug] + 1]/7

For intermittent dosing of IV drugs, such as cetuximab:

- duration of exposure = [[Date of last (non-zero) dose of study drug] – [date of first dose of study drug] + [days until next dose]]/7

  As an example, for drugs that are dosed weekly, duration of exposure = [[Date of last (non-zero) dose of study drug] – [date of first dose of study drug] + 7]/7

Relative Dose Intensity

- Duration of planned exposure (this is not summarized in any exposure table) = [[Date of EOT or data cutoff (the earlier one)] – [date of first dose of study drug] + 1]/7

For daily dosing of binimetinib and encorafenib:

- Dose intensity (DI)= [Cumulative actual dose]/[duration of planned exposure]
- Cumulative planned dose = sum (protocol specified dose across each day of planned exposure)
- Planned dose intensity (PDI) = [Cumulative planned dose]/[duration of planned exposure]
- Relative dose intensity (RDI) = 100 x [Dose intensity]/[Planned dose intensity]

For intermittent dosing of IV drugs, such as cetuximab:

- Number of planned doses = number of non-zero doses plus the number of interrupted doses.
  - May be calculated from the duration of planned exposure as floor([duration of planned exposure – 1]/[days between doses]) + 1
- Dose intensity (DI) = [Cumulative actual dose]/[number of planned doses]
- Cumulative planned dose = sum (protocol specified dose across each planned day of dosing)
- Planned dose intensity (PDI) = [Cumulative planned dose]/[number of planned doses]
- Relative dose intensity (RDI) = 100*[Dose intensity]/ [number of planned doses]

A dose interruption will be indicated in the eCRF by a dosing record with a total daily dose of 0 mg for one or more days.

In order not to over count interruptions, dosing records with 0 mg entered as last dosing record will not be counted as interruptions. Those represent the reason for permanent discontinuation and will therefore be presented in the reason for treatment discontinuation analysis.

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 encorafenib, in the sequence of total daily dose 300 mg – 0 mg – 200 mg, the 200 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 to 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.

Reporting

Information about administration of each study drug (encorafenib, binimetinib, cetuximab, 5-FU, FA, irinotecan) will be presented in a data listing. For patients who have not discontinued study drug prior to the data cutoff date, the patient's last dose date will be the date of the last administration or infusion.

Duration of exposure to study drug, average daily dose, DI and RDI will be summarized by study drug and treatment arm. Duration of exposure will also be categorized by time intervals (e.g. < 1 months, 1-3 months, etc. as appropriate for the protocol) for which frequency counts and percentages of patients will be provided. Duration of exposure to study treatment (i.e. Triplet Arm, Doublet Arm, or Control Arm) will also be presented.

Frequency counts and percentages of patients who have dose reductions or interruptions, and the corresponding reasons, will be summarized by study drug and treatment arm. Dose reductions and interruptions will be tabulated both separately and in a combined fashion.

The percentages of days between the first and last non-zero dose with reduced dose i.e. below the protocol planned dose will be summarized. Note, in this calculation a dose of 0 mg will also be considered as a reduced dose. The percentages of days between the first and last non-zero dose with dose interruption will also be summarized.

Separate summaries will be generated for the subgroups of patients in Japan and in Korea.

For each patient, listings of each dose of the study drug administered along with dose change/interruption reasons will be produced.

7.7.2 Concomitant Medications

Concomitant therapies are defined as any medications (excluding study drug and prior antineoplastic treatments) administered in the study.

Concomitant medications will be coded using the WHO Drug Dictionary and presented in a data listing. Concomitant medications will be summarized by Anatomical Therapeutic Chemical (ATC) code (levels 2 and 4) and treatment group.

For partial concomitant medication start date, the date imputation will be based on the temporal relation between the partial date and start of treatment date. No imputation will be performed when the year is missing for the concomitant start date.
For partial concomitant medication end date or completely missing end date (concomitant medication is ongoing), the date imputation will be based on the temporal relation between the partial date, the last contact date and the 30-day follow-up date.

Table 7-2 and Table 7-3 provide examples of the different considered imputations for concomitant medication start and end dates, respectively.

Table 7-2: Concomitant Medication Start Date Imputation Example Scenarios

Partial conmed 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
ddSEP2001	20OCT2001	Before	15SEP2001
ddNOV2001	20OCT2001	After	01NOV2001

Table 7-3: Concomitant Medication End Date Imputation Example Scenarios

Partial conmed end date	Minimum (Last contact date, 30-day FU date)	Ongoing	Imputed Date
Missing	20OCT2001	Yes	20OCT2001

Partial conmed end date	Minimum (Last contact date, 30-day FU date)	Ongoing	Imputed Date
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

7.7.3 Adverse Events

Adverse events (AEs) and serious adverse events (SAEs) will be coded by Lowest Level Term (LLT) using the most recent version of the Medical Dictionary for Regulatory Activities (MedDRA) at the time of the database lock (Northrop Grumman Corporation, Chantilly, VA USA).

The severity of an AE will be assessed by the Investigator using the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE), Version 4.03. For patients with more than 1 AE within a SOC or PT, only the highest grade will be included in by-severity summaries.

The imputation of the start date and end date of AEs will follow the same conventions as for concomitant medication start date and end date.

All dates of death must be completed with day, month and year. If the day or month is missing, death will be imputed to the maximum of the full (non-imputed) last contact date (excluding the date of death) and the following:

- Missing day: 15th day of the month and year of death
- Missing day and month: July 1st of the year of death

For summaries by SOC and PT, each patient will be counted at most once per SOC and at most once per PT. For summaries by PT, each patient will be counted at most once per PT.

Treatment-emergent AEs (AEs) will be defined as:

- Any new event that starts after administration of study drug and ≤ 30 days after treatment discontinuation.

- Any event that was ongoing when treatment with study drug started and the severity/grade after treatment was higher than the Baseline value (fluctuations below the Baseline severity/grade are not considered as treatment emergent).
- Any new event that starts > 30 days after treatment discontinuation and is assessed by the Investigator as related to study treatment.

Unless otherwise stated, only AEs 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. All deaths will be listed and deaths within 30 days of last dose of study drug will be flagged.

An overall summary of safety table will include numbers and percentages of the following:

- Patients who died on study or within 30 days of last treatment.
- Patients with at least 1 AE, regardless of causality
- Patients with at least 1 AE with suspected study drug relationship
- Patients with at least 1SAE, regardless of causality
- Patients with at least 1 SAE with suspected study drug relationship
- Patients with at least 1 AE leading to discontinuation of study drug. regardless of causality
- Patients with at least 1 AE leading to discontinuation of study drug with suspected study drug relationship
- Patients with at least 1 AE requiring dose reduction regardless of causality
- Patients with at least 1 AE requiring dose reduction with suspected study drug relationship
- Patients with at least 1 AE requiring dose interruption regardless of causality
- Patients with at least 1 AE requiring dose interruption with suspected study drug relationship
- Patients with at least 1 AE requiring additional therapy regardless of causality
- Patients with at least 1 AE requiring additional therapy with suspected study drug relationship

Individual summary tables showing the incidence of patients with the following subsets of AEs will be generated:

- AEs, regardless of causality, by SOC and PT (all grades and grade 3+)
- AEs, regardless of causality, by PT (all grades and grade 3+)
- AEs, regardless of causality, by PT (all grades and grade 1, 2, 3, 4, 5)

- AEs, with suspected study drug relationship, by SOC and PT (all grades and grade 3+)
- AEs, with suspected study drug relationship, by PT (all grades and grade 3+)
- SAEs, regardless of causality, by SOC and PT (all grades and grade 3+)
- SAEs, regardless of causality, by PT (all grades and grade 3+)
- SAEs, with suspected study drug relationship, by SOC and PT (all grades and grade 3+)
- SAEs, with suspected study drug relationship, by PT (all grades and grade 3+)
- Non-serious AEs, regardless of causality, by SOC and PT (all grades and grade 3+)
- AEs that led to discontinuation of study drug, regardless of causality, by SOC and PT (all grades and grade 3+)
- AEs that led to discontinuation of study drug, regardless of causality, by PT (all grades and grade 3+)
- AEs that led to dose reduction of study drug, regardless of causality, by SOC and PT (all grades and grade 3+)
- AEs that led to dose reduction of study drug, regardless of causality, by PT (all grades and grade 3+)
- AEs that led to dose interruption of study drug, regardless of causality, by SOC and PT (all grades and grade 3+)
- AEs that led to dose interruption of study drug, regardless of causality, by PT (all grades and grade 3+)
- AEs that required additional therapy, regardless of causality, by SOC and PT (all grades and grade 3+)
- AEs that required additional therapy, regardless of causality, by PT (all grades and grade 3+)
- AEs that resulted in death on study or within 30 days of last study treatment, regardless of causality, by SOC and PT

Selected tables may be presented for specific subsets of AEs.

Time to onset (i.e. first occurrence) of an AE is defined as time from start of study treatment to the date of first occurrence of an AE within the grouping, i.e. time in days is calculated as (start date of first occurrence of AE) – (date of first dose of study treatment) +1. In the absence of an event, the censoring date applied will be the earliest from the following dates: end of treatment + 30 days, analysis cut-off, new anti-cancer therapy start date, death and last contact date. Ascending KM curves will be constructed by treatment arm. Median together with 95%

confidence interval will be presented for each treatment arm. Hazard ratio will be obtained from an unstratified Cox model.

The time to onset and the duration of AEs will be performed for

- first SAE
- first Grade 3 or above AE
- AEs resulting in discontinuation of study drug

7.7.4 Adverse Events of Special Interest

Specific groupings of AEs of special interest will be considered and the number of patients with at least one event in each grouping will be reported. Such groups consist of AEs for which there is a specific clinical interest in connection with encorafenib and/or binimetinib treatment (i.e. where encorafenib and/or binimetinib may influence a common mechanism of action responsible for triggering them) or adverse events which are similar in nature (although not identical). Separate summaries will be generated for AEs of special interest attributed to encorafenib and those attributed to binimetinib.

All AEs of special interest (AESI) groupings are defined through the use of Preferred Terms (PT), High Level Terms (HLT), System Organ Classes (SOC), Standardized MedDRA Queries (SMQ) or through a combination of these components. The MedDRA terms that define each AESI grouping will be outlined in the Case Retrieval Strategy (CRS). It will be used to map reported AEs to the notable AEs groupings. The list of AESIs may be updated during the course of the trial based on accumulating safety data. Therefore, each clinical study report will list the AE groupings used and provide a listing of the corresponding case retrieval terms.

All events occurring between the start of study treatment and last treatment +30 days will be taken into account. Grouping of AEs will be summarized, as follows:

- Adverse events of special interest, regardless of causality, by grouping and PT (overall and grade 3+)
- Adverse events of special interest, regardless of causality, by grouping and PT (overall and grade 1, 2, 3, 4, 5)
- Adverse events of special interest, with suspected study drug relationship, by grouping and PT (overall and grade 3+)
- Serious adverse events of special interest, regardless of study drug relationship, by grouping and PT (overall and grade 3+)

- Adverse events of special interest that led to discontinuation of study drug, regardless of causality, by grouping and PT (overall and grade 3+)
- Adverse events of special interest that led to dose reduction of study drug, regardless of causality, by grouping and PT (overall and grade 3+)
- Adverse events of special interest that led to dose interruption of study drug, regardless of causality, by grouping and PT (overall and grade 3+)
- Adverse events of special interest that required additional therapy, regardless of causality, by grouping and PT (overall and grade 3+)

Note that depending on the project level information available at the time of the analysis the grouping might be updated to reflect the new AEs of interest.

The CRS grouping terms and corresponding MedDRA elements will be listed.

7.7.5 Clinical Laboratory Evaluations

Required hematology, coagulation, urinalysis and clinical chemistry tests are described in the protocol. Hematology, coagulation and clinical chemistry test results will be presented using the International System of Units (SI units) and, where appropriate, will be graded using NCI CTCAE, Version 4.03.[18]

Incidence of clinically notable shifts in laboratory parameters based on CTCAE grade for specific laboratory evaluations will be summarized, where clinically notable is defined as worsening from baseline by at least 2 grades or to ≥ Grade 3.

Laboratory parameters not graded by NCI CTCAE (e.g. basophils, hematocrit, etc.) , will be summarized by normal range.

All lab test results will be presented in listings sorted by treatment group, patient identifier, lab test and date/time of collection. Values outside laboratory normal ranges will be flagged where appropriate. Separate listings will be provided for patients with any clinically notable shifts in laboratory parameters as defined above.

Graded hematology, coagulation, urinalysis, and clinical chemistry test results will be summarized in shift tables from Baseline to worst grade on study. For shift tables, the number of patients with missing data at Baseline or on study will be displayed. Separate shift tables will be generated for the subgroups of patients in Japan and in Korea.

Possible Hy's Law and drug-induced liver injury (DILI) cases will be summarized by treatment group.

A shift table summarizing the worst post-baseline CK CTC grade vs baseline CTC grade will be generated.

7.7.6 Pregnancy Tests

Results of all pregnancy tests will be presented in a data listing.

7.7.7 Physical Examinations

Physical examination dates will be presented in a data listing. Any abnormal findings were to have been reported as AEs and will be reported in the appropriate AE listings and summary tables.

7.7.8 ECOG Performance Status

ECOG performance status is used to assess the physical health of patients, and ranges from 0 (most active) to 5 (dead).

Frequency counts and percentages of patients in each score category will be provided by treatment arm and assessment visit.

ECOG PS at each time point will be listed.

7.7.9 Vital Signs

The following criteria define clinically notable abnormalities:

- Clinically notable elevated values
  - Systolic BP: ≥ 160 mmHg and an increase ≥ 20 mmHg from baseline;
  - Diastolic BP: ≥ 100 mmHg and an increase ≥ 15 mmHg from baseline;
  - Heart rate (collected as pulse rate in the vital signs eCRF): ≥ 120 bpm with increase from baseline of ≥15 bpm;
  - Weight: increase from baseline of ≥ 10%;
  - Body temperature [C]: >= 37.5 C.
- Clinically notable low values
  - Systolic BP: <=90 mmHg with decrease from baseline of >=20 mmHg;
  - Diastolic BP: <=50 mmHg with decrease from baseline of >=15 mmHg;

- Heart rate (collected as pulse rate in the vital signs eCRF): <=50 bpm with decrease from baseline of >=15 bpm;
- Weight: >=20% decrease from baseline;
- Body temperature [C]: <=36 C.

Number and percentage of patients with at least one post-baseline vital sign abnormality will be summarized by treatment arm.

In addition, blood pressure (systolic, diastolic and both combined) shift table based on CTCAE grade for hypertension at baseline and worst post-baseline will be produced.

Patients with clinically notable vital sign abnormalities will be listed by treatment arm. All vital sign assessments will be listed by treatment arm, patient and vital sign parameter.

In the listings, clinically notable values will also be flagged.

7.7.10 Electrocardiograms

As per protocol, baseline ECG will be obtained as the triplicate measurements at Cycle 1 Day predose.

The following summaries will be produced for QT, QTcF, and heart rate by treatment arm:

- Frequency counts and percentages of patients having clinically notable ECG values (Table 7-4).

Patients with clinically notable ECG values will be listed by treatment arm including the corresponding notable values and abnormality findings. All ECG values will be listed by treatment arm, patient and ECG parameter.

Table 7-4: Clinical Notable ECG Criteria

Parameter	Criterion
QT, QTcF, QTcB	increase from baseline > 30 ms increase from baseline > 60 ms new > 450 ms new > 480 ms new > 500 ms
Heart rate	Increase from baseline >25% to a value >100 Decrease from baseline >25% and to a value <50

7.7.11 MUGA/Echocardiogram

LVEF abnormalities are defined according to CTCAE version 4.03. Patient will be considered as having a LVEF abnormality if the worst post baseline value is grade 2, 3 or 4 according to the following classification:

- Grade 0: Non-missing value below Grade 2
- Grade 2: LVEF between 40% and 50% or absolute change from baseline between -10% and < -20%
- Grade 3: LVEF between 20% and 39% or absolute change from baseline lower than or equal to -20%
- Grade 4: LVEF lower than 20%.

The following summaries will be produced for LVEF by treatment arm:

- Incidence of worst post-baseline LVEF value with CTC grades
- Shift tables using CTC grades to compare baseline to the worst post-baseline LVEF value with CTC grades.

All LVEF assessments will be listed.

Patients with grade 2 or above CTCAE abnormalities will be listed by treatment arm.

Different modalities to assess LVEF might be used for the same patient, change from baseline and shift table will be provided regardless of the modality.

7.7.12 Ophthalmic Examination

All ophthalmic examinations (i.e. Tonometry, Visual acuity, Visual field, Fundoscopy, Slit lamp, Optical Coherence Tomography (OCT), Fluorescein Angiography) will be listed.

Visual acuity

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).

Total visual acuity score will also be assessed identifying clinically meaningful deterioration in LogMAR. The number of patients reporting at least once a decrease in score of <=0, 0 to <0.1, 0.1 to <0.2, 0.2 to <0.3 and ≥0.3 LogMAR will be summarized in shift table.

Ophthalmic events occurring within -15 days to 30 days of the worst logMAR score (<=0, 0 to <0.1, 0.1 to <0.2, 0.2 to <0.3 or ≥0.3) will be presented by preferred term and worst logMAR score.

Patients with ≥0.2 LogMAR loss will be listed by treatment arm, patient and timepoint including the Snellen fraction and logMAR score. In addition, all Snellen fraction and logMAR scores will be listed by eye and treatment arm.

Intraocular pressure

Number and percentage of patients with clinically intraocular pressure above 30 mmHg will be summarized by treatment arm. For patients with such values, a listing will also be provided.

Optical Coherence Tomography (OCT) and Fluorescein Angiography (FA)

The clinical outcome presented in the listing combine the clinical observation (not present, clinically insignificant and clinically significant) and the eye involvement (unilateral or bilateral). Since these are optional assessments, all the patients are considered at risk at baseline of developing a new abnormality.

Visual field, fundoscopy and slit lamp

Visual field, fundoscopy and slit lamp will be summarized by treatment arm. For each considered category in each assessment, newly occurring abnormalities will be provided by treatment arm.

7.7.13 Dermatologic Examination

All dermatologic examination data will be listed.

7.7.14 Follow-up

Information obtained during a Follow-up Visit that is not otherwise captured in data listings (e.g., date of Follow-up Visit, start date of subsequent therapy) will be presented in a data listing. A summary of subsequent anticancer therapies will be provided.

7.7.15 Comments

All comments entered into the electronic data capture (EDC) system by personnel at the study sites will be presented in a data listing.

7.8 Pharmacokinetic Analysis

Systemic concentrations and PK parameters will be determined for cetuximab, encorafenib, binimetinib and the primary metabolite for binimetinib, AR00426032. PK summaries will be presented by treatment group for patients in the PK set. In the following, Day y of Cycle x will be abbreviated as CxDy. C1D1 is defined as single dose (i.e., first dose) and C2D1 is defined as steady-state.

7.8.1 Concentrations of Cetuximab, Encorafenib, Binimetinib, and AR00426032

Plasma concentrations of encorafenib, binimetinib, and AR00426032 will be quantitated at the time points indicated in the schedules of assessments in the protocol. Serum concentrations of cetuximab will be quantitated at equivalent time points with respect to the start of the cetuximab infusion. All concentration values for each patient in the PK set will be included in the bioanalytical concentration listings.

The concentrations of cetuximab, encorafenib, binimetinib, and AR00426032 will be summarized for all nominal time points, including predose (trough) concentrations, using the following descriptive statistics: n, arithmetic mean, SD, coefficient of variation (CV), geometric mean, geometric SD, geometric CV, minimum, median and maximum. Detailed information regarding handling of data with respect to collection time window specified in the protocol, dose reductions, presentation and handling of data below the limit of quantitation (BLQ) and additional issues related to PK collection and analysis will be discussed in detail in a standalone Pharmacokinetic Analysis Plan (PAP). Separate outputs of concentration summaries will be generated for the Safety Lead-In and Phase 3 portions of the study.

Due to the sparse PK sample collection of the Phase 3 portion of the study, the graphical evaluation of PK data described below will only be conducted for the Safety Lead-In portion of the study. The geometric mean with geometric SD concentration-versus-time profiles will be presented graphically for each analyte using both linear and semi-logarithmic scales by treatment group on C1D1 and C2D1. Individual plasma concentration-time profiles will also be presented graphically using linear and semi-logarithmic scales by treatment group and study day. For ease of presentation, nominal times will be used to present results in summary and individual figures. Handling of plotted data where BLQ values are present will be specified in a standalone PAP.

7.8.2 Pharmacokinetic Parameters for Cetuximab, Encorafenib, Binimetinib, and AR00426032

Pharmacokinetic parameters for patients in the PK set will be determined by treatment group for cetuximab, encorafenib, binimetinib, and AR00426032 when possible and appropriate.

7.8.2.1 PK Analysis for Safety Lead-In

The individual concentration-time data for each analyte will be evaluated with noncompartmental analysis (NCA) using Phoenix WinNonlin®, Version 6.4 or higher. Actual blood collection times and doses will be used for PK calculations. Additional details regarding the handling of data and presentation of parameters for the NCA analysis will be provided in a standalone PAP.

The following parameters will be calculated for cetuximab, encorafenib, binimetinib, and AR00426032 on C1D1 and C2D1:

PK Parameter	Definition
AUC_{ext}	Percentage of AUC_{inf} due to extrapolation from T_{last} to infinity, calculated as: $\frac{AUC_{inf} - AUC_{last}}{AUC_{inf}} \times 100$
AUC_{inf}	Area under the plasma concentration-time curve extrapolated to infinity
AUC_{last}	Area under the plasma concentration-time curve from zero to the last measurable time point
AUC_{tau}	Area under the plasma concentration-time curve over the dosing interval
AUC/D	Dose-normalized AUC, calculated as: $\frac{AUC}{Dose}$ Note: an appropriate AUC parameter for use in this will be specified depending on data quality and will be discussed in the final report.
C_{max}	Observed maximum plasma concentration
C_{max}/D	Dose-normalized C_{max}, calculated as: $\frac{C_{max}}{Dose}$
C_{last}	Last measured concentration
T_{max}	Observed time of C_{max}
T_{last}	Observed time of C_{last}
k_{el}	Terminal elimination rate constant

PK Parameter	Definition
$t_{1/2}$	Terminal elimination half-life, calculated as: $\frac{ln(2)}{k_{el}}$
CL/F (oral) or CL (intravenous)	Apparent systemic clearance, calculated as: $\frac{Dose}{AUC_{inf}}$ for single dose and $\frac{Dose}{AUC_{tau}}$ at steady state
V_{ss} (intravenous)	Apparent steady-state volume of distribution, calculated as: $\frac{Dose\ x\ AUMC_{inf}}{(AUC_{inf})^2}$
V_z/F (oral) or V_z (intravenous)	Apparent volume of distribution, calculated as: $\frac{Dose}{AUC_{inf} \times k_{el}}$ for single dose and $\frac{Dose}{AUC_{tau} \times k_{el}}$ at steady state
MR_{AUC}, MR_{Cmax} (binimetinib only)	Metabolite ratios, calculated as: $\frac{AUC_{metabolite}}{AUC_{parent}}$, $\frac{C_{max,\ metabolite}}{C_{max,\ parent}}$ Note: an appropriate AUC parameter for use in this calculation will be specified depending on data quality and will be discussed in the final report.
R_{AUC}, R_{Cmax}	Accumulation ratios, calculated as: $\frac{AUC_{tau,\ steady\ state}}{AUC_{tau,\ single\ dose}}$, $\frac{C_{max,\ steady\ state}}{C_{max,\ single\ dose}}$ Note: an appropriate AUC parameter for use in this calculation will be specified depending on data quality and will be discussed in the final report.

The AUC parameters will be calculated according to the linear-up log-down trapezoidal rule. All AUC and C_{max} values will also be provided as dose-normalized values based on the actual dose administered prior to PK sampling. Additional PK parameters may be calculated at the discretion of the pharmacokineticist, and reporting of PK parameters listed above is subject to data quality.

All PK parameter values will be presented in data listings by analyte, treatment group, cycle and study day. Each parameter will be summarized in tables by analyte, treatment group, cycle and study day using the following descriptive statistics: n, arithmetic mean, standard deviation, CV, geometric mean, geometric standard deviation, geometric CV, minimum, median and maximum. Summary descriptive statistics for in-text summary tables will include geometric mean with geometric CV for AUC, AUC/D, C_{max}, C_{max}/D, R_{AUC}, MR_{AUC}, CL, CL/F, V_{ss} and V_z/F values as appropriate. For T_{max} and $t_{1/2}$ values, median, minimum and maximum will be presented. If either the extent of extrapolation for an AUC_{inf} is not acceptable and/or the goodness of fit value for

terminal phase regression is not acceptable, then only AUC_{tau} or AUC_{last} may be presented in the summary statistics rather than any terminal phase-derived parameters.

Differences in PK parameter estimates for non-Japanese and Japanese Safety Lead-in cohorts will be assessed by generating separate summary tables of PK parameters for non-Japanese and Japanese Safety Lead-in cohorts. Additional comparison of non-Japanese and Japanese Safety Lead-in cohorts using dose-normalized AUC and C_{max} values pooled across dose levels. Assessments of dose-normalized AUC and C_{max} between non-Japanese and Japanese Safety Lead-in cohorts will be conducted for cetuximab, encorafenib, binimetinib, and AR00426032 based on geometric mean ratios and 90% confidence intervals for Cycles 1 and 2 (Day 1).

7.8.2.2 PK Analysis for Randomized Phase 3 Portion

PK analysis based on population PK modeling will be performed and described in a separate stand-alone modeling plan and a specific report will be produced.

7.8.2.3 Drug-drug Interactions

The potential effect of co-administration of one study drug on the PK of another study drug will be assessed (1) by comparison of PK results to historical PK results, and (2) by covariate analysis using nonlinear mixed effects analysis (i.e., population PK analysis). Comparisons to historical PK results will be qualitative and only done as data permit. Covariate analysis will be conducted as appropriate for the following potential interactions:(1) the effect of coadministration of encorafenib on the PK of cetuximab, (2) the effect of coadministration of encorafenib+binimetinib on the PK of cetuximab, (3) the effect of coadministration of binimetinib on the PK of encorafenib. The source of cetuximab drug product may also be included as a covariate. Analyses will be described in a separate stand-alone modeling plan and a specific report will be produced.

7.8.2.4 Exposure-response Analysis

Relationships between exposure of study drugs and response or safety will be conducted using post hoc exposure estimates (e.g., C_{max}, AUC or C_{min}) from population PK modeling, if possible and appropriate. Measures of efficacy may include but are not limited to OS, PFS and ORR. Selection of safety measures will be based on frequency of observations or by selection of safety measure of interest. Analyses will be described in a separate stand-alone modeling plan and a specific report will be produced.

CC

[redacted]

7.10 Patient Reported Outcomes

Quality-of-life data will be analyzed on the FAS.

The EORTC QLQ-C30, FACT-C, EQ-5D-5L, and PGIC will be used to assess QoL. For EORTC-QLQ-C30, the global health status/QoL scale score is identified as the primary patient-reported outcome variable of interest; physical functioning, emotional functioning and social functioning scale scores of the QLQ-C30 are considered as secondary. For FACT-C, the functional well-being score is the primary patient-reported outcome variable of interest; the physical well-being, social/family well-being, emotional well-being, and additional concern scores are considered as secondary. The EQ-5D-5L contains 1 item for each of 5 dimensions health-related quality of life (i.e., mobility, self-care, usual activities, pain or discomfort, and anxiety or depression). Response options for each item vary from having no problems, moderate problems, or extreme problems.

The EORTC QLQ-C30, FACT-C, EQ-5D-5L, and PGIC will be scored according to respective scoring manuals, respectively. The number of patients completing the questionnaires and the number of missing or incomplete assessments will be summarized by each treatment arm for each scheduled assessment time point.

Descriptive statistics will be used to summarize the scored scales at each scheduled assessment. Additionally, change from baseline in the domain scores at the time of each assessment will be summarized. Patients with an evaluable baseline score and at least 1 evaluable post-baseline score during the treatment period will be included in the change from baseline analyses. In addition, a repeated measurement analysis model may be used to compare the treatment arms with respect to changes in the domain scores longitudinally over time.

Time to definitive deterioration in the QoL domains will be assessed in the 3 treatment arms in the FAS. The time to definitive deterioration is defined as the time from the date of randomization to the date of event, which is defined as at least 10% worsening relative to Baseline of the corresponding scale score with no later improvement above this threshold

observed during the course of the study or death due to any cause. If a patient has not had an event prior to analysis cutoff or start of another anticancer therapy, time to deterioration will be censored at the date of the last adequate QoL evaluation. The distribution will be presented descriptively using KM curves. Median time to definitive deterioration along with 2-sided 95% CI will be provided. A Cox model will be fit with treatment arm and stratification factors as the covariates to obtain a HR estimate of the treatment effect along with 95% CI. The stratification factors used in the test will be precisely those used for randomization, and will be based on the actual randomization (IWRS) information.

7.11 Interim Safety Reviews

The independent DMC will review the available safety information after the first 30 patients in the randomized Phase 3 portion of the study (i.e., approximately 10 patients in each arm) have had the opportunity to complete at least 1 cycle of treatment to confirm tolerability. Subsequent DMC safety data reviews will occur at regular intervals. Additional details will be outlined in the DMC Charter.

7.12 Initial Efficacy Analyses

An initial analysis of the study will be performed when all three of the following criteria have been met:

- approximately 9 months after randomization of the 330th patient (i.e., approximately 110 patients per arm), to allow a majority of responders among the 330 Phase 3 patients to have the opportunity to be followed for approximately 6 months or longer after their first response
- at least 188 OS events have occurred in the Triplet and Control arms combined (i.e., approximately 70% information)
- at least 169 OS events have occurred in the Doublet and Control arms combined (i.e., approximately 50% information)

The primary endpoint of Triplet vs. Control ORR (per BICR) will be formally tested first at this analysis and will be based on the first approximately 330 randomized patients (any additional patient[s] randomized on the same day as the 330th randomized patient will be included in the analysis). If the Triplet vs. Control ORR comparison is positive (i.e., $p < 0.005$), then based on the fallback procedure, the OS endpoint of Triplet vs. Control will be assigned a total 1-sided alpha = 0.025. Otherwise, the Triplet vs. Control OS will be assigned a total 1-sided alpha = 0.020.

An interim analysis for superiority or (non-binding) futility of the Triplet vs Control OS endpoint will be also performed at the time of the primary ORR analysis based on all available data. Futility and superiority boundaries for both the OS interim and final analyses will be determined using a Lan-DeMets spending function that approximates O'Brien-Fleming stopping boundaries.

For the Triplet vs. Control comparison, if the entire alpha=0.025 can be applied to the OS endpoint, and assuming the interim analysis will occur at 70% information, the overall power will be 88.7%. The critical HR for efficacy at the interim analysis will be 0.701, the critical HR for futility at the interim analysis will be 0.842, and the critical HR at the final analysis will be 0.783. If alpha=0.020 is applied to the OS endpoint with the interim analysis occurring at 70% information, the overall power will be 86.6%. The critical HR for efficacy at the interim analysis will be 0.690, the critical HR for futility at the interim analysis will be 0.828, and the critical HR at the final analysis will be 0.775.

For the Doublet vs. Control comparison, if the entire alpha=0.025 can be applied to the OS endpoint, and assuming the interim analysis occurs at 50% of information, the overall power will be 90.5%. The critical HR for efficacy at the interim analysis will be 0.634, and at the final analysis, the critical HR will be 0.807. If alpha=0.020 is applied to the OS endpoint with the interim analysis occurring at 50% information, the overall power will be 88.9%. The critical HR for efficacy at the interim analysis will be 0.622, and at the final analysis, the critical HR will be 0.799.

The estimated power and critical values described above assume an exponential survival distribution in each arm. In addition, the actual alpha spent will be determined by the observed information fraction at the interim OS analysis, and the alpha remaining will be adjusted accordingly for final analyses.

If the p-value for the Triplet Arm vs Control Arm OS comparison exceeds the superiority boundary at the interim analysis, testing of the endpoints included in the hierarchical approach described in Section 11.3.3 will be conducted. For each endpoint in the hierarchy, a Lan-DeMets spending function that approximates the O'Brien-Fleming stopping boundaries will be used with data from all available patients. The total alpha assigned to each endpoint will be either 0.025 (if the Triplet Arm vs. Control Arm ORR comparison was $p < 0.005$) or 0.020 (if the Triplet Arm vs. Control Arm ORR comparison was $p \geq 0.005$).

If the p-value for the Triplet Arm vs Control Arm OS comparison does not exceed the superiority boundary at the interim analysis, OS of Triplet Arm vs. Control Arm will be tested again at the final analysis (i.e., when at least 268 OS events in Triplet and Control and at least 338 OS events in Doublet and Control have occurred).

Table 7-5 provides the number of expected events, the cumulative alpha spent, and the cumulative power to reject the null hypothesis at the expected analysis timepoints for the Triplet vs Control OS analysis.

Table 7-5: Expected Number of OS Events and Cumulative Power at Expected OS Analysis Timepoints for the Triplet vs Control Comparison

		OS Analysis		
	Triplet vs Control OS Analysis	**Cumulative Number of OS Events**	**Cumulative Alpha Spent on OS**	**Cumulative Power to Reject H0 for OS (%)**
Triplet vs Control ORR is statistically significant (i.e., one-sided p<0.005)	Interim	188	0.0074	61.5
	Final	268	0.0250	88.3
Triplet vs Control ORR is not statistically significant (i.e., one-sided p>0.005)	Interim	188	0.0055	57.4
	Final	268	0.0200	86.3

Note: All values were calculated using East® v6.4, assuming 70% information at the OS interim analysis. Boundaries will be adjusted according to the actual information fraction observed at the interim analysis. Cumulative power values were estimated using simulations within East® under the alternative hypothesis.

Efficacy and futility are determined using Lan-DeMets approximation of O'Brien-Fleming alpha- and (non-binding) beta-spending boundaries, respectively.

7.13 Subgroup Analysis

The following subgroup analyses will be performed for efficacy endpoints for Phase 3 patients:

For OS on the FAS, and for ORR on the Phase 3 Response Efficacy Set:

- ECOG PS (0 vs. 1)
- Prior use of irinotecan (yes vs. no)
- Cetuximab source (US-licensed vs EU-approved)
- Region (North America vs. Europe vs Rest of World)
- Number of prior regimens (1 vs 2)
- Race (Caucasian vs non-Caucasian)
- Race (Asian vs non-Asian)
- Age (<65 vs ≥65 years)
- Gender (male vs female)
- Number of organs involved at baseline (≤2 vs ≥3)

- MSI (high vs. normal)
- Baseline CRP (≤ULN vs >ULN)
- Removal status of primary tumor (no resection, partial resection, complete resection)
- Side or tumor (left vs right)
- Number of metastatic sites based on Target and Non-target lesion assessment (1 vs 2+)
- Presence of liver metastases at baseline, based on Target and Non-target lesion assessment (yes vs. no)
- Japanese subgroup (patients randomized in Japan)
- Korean subgroup (patients randomized in Korea)

For PFS, ORR, DOR, and TTR on the FAS:

- Japanese subgroup (patients randomized in Japan)
- Korean subgroup (patients randomized in Korea)

The following subgroup analyses will be performed for the AE analyses:

- Race (Asian vs non-Asian)
- Race (Caucasian vs non-Caucasian)
- Age (<65 vs ≥65 years)
- Gender (male vs female)
- Japanese subgroup (patients randomized in Japan)
- Korean subgroup (patients randomized in Korea)

8.0 DATA AND ANALYSIS QUALITY ASSURANCE

This protocol was conducted under the sponsorship of Array BioPharma. Personnel within Array BioPharma provided statistical and data management input for the design of the clinical trial protocol; data management, statistical analysis and generation of tables, listings and figures; and medical writing support for the clinical study report.

All parties mentioned above will work diligently and collaboratively to ensure that data collection and analysis for this study are of the highest quality. This will be accomplished through programmed edit checks, quality control processes, and clinical and statistical review of data displays. Quality and accuracy of statistical analyses will be verified though established statistical programming validation processes.

9.0 REFERENCES

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

2. Cella D. Manual for the Functional Assessment of Chronic Illness Therapy (FACIT) Measurement System. FACIT.org; Elmhurst, IL, 2012.

3. Clarke S, Yip S, Brown C, et al. Single-agent irinotecan or 5-fluorouracil and leucovorin (FOLFIRI) as second-line chemotherapy for advanced colorectal cancer; results of a randomised phase II study (DaVINCI) and meta-analysis. Eur J Cancer 2011;47:1826-36.

4. Clopper CJ, Pearson ES. (1934) The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial. Biometrika; 26(4): 404-413.

5. De Roock W, Claes B, Bernasconi D, et al. Effects of KRAS, BRAF, NRAS, and PIK3CA mutations on the efficacy of cetuximab plus chemotherapy in chemotherapy-refractory metastatic colorectal cancer: a retrospective consortium analysis. Lancet Oncol 2010;11:753-762. 20.

6. East® version 6.3 manual (2014). Cytel Inc. Cambridge, MA USA.

7. East® version 6.4 manual (2016). Cytel Inc. Cambridge, MA USA.

8. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer 2009;45(2):228-47

9. Fayers PM. (2001) The EORTC QLQ-C30 Scoring Manual (3rd Edition). Brussels: European Organisation for Research and Treatment of Cancer.

10. Guidance for Industry: E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarryhthmic Drugs, Center for Drug Evaluation and Research, Food and Drug Administration, Rockville, MD. http://www.fda.gov/RegulatoryInformation/Guidances/ucm129335.htm

11. Kalbfleisch JD, Prentice RL. (2002) The Statistical Analysis of Failure Time Data. Wiley Series in Probability and Statistics.

12. Klein, J. P. and Moeschberger, M. L. (1997), Survival Analysis: Techniques for Censored and Truncated Data, New York: Springer-Verlag.

13. Kopetz S, McDonough SL, Lenz H-J et al. Randomized trial of irinotecan and cetuximab with or without vemurafenib in BRAF-mutant metastatic colorectal cancer (SWOG S1406). J Clin Oncol. 2017;35:4s (suppl abstr 3505).

14. Lan KKG, DeMets DL (1983). Discrete sequential boundaries for clinical trials. Biometrika 70: 649-53.

15. Morris V, Overman MJ, Jiang Z-Q, et. al. Progression-free survival remains poor over sequential lines of systemic therapy in patients with BRAF-mutated colorectal cancer. Clin Col Cancer Res 2014;13:164-71.

16. Peeters M, Oliner KS, Price TJ, et al. Updated analysis of KRAS/NRAS and BRAF mutations in study 20050181 of panitumumab (pmab) plus FOLFIRI for second-line treatment (tx) of metastatic colorectal cancer (mCRC). J Clin Oncol 32:15s, 2014 (suppl; abstr 3568).

17. Schemper M, Smith TL (1996). A note on quantifying follow-up in studies of failure time. Controlled Clinical Trials; 17: 343-346.

18. U.S. Department of Health and Human Services. Common Terminology Criteria for Adverse Events (CTCAE). Revised May 29, 2009. Accessed July 14, 2009. http://ctep.cancer.gov/protocolDevelopment/electronic_applications/docs/ctcaev4.pdf

19. Wiens BL, Dmitrienko A. The fallback procedure for evaluating a single family of hypotheses. J Biopharm Stat. 2005; 15:929-42.

20. Seymour MT, Brown SR, Middleton G et al. Panitumumab and irinotecan versus irinotecan alone for patients with KRAS wild-type, fluorouracil-resistant advanced colorectal cancer (PICCOLO): a prospectively stratified randomised trial. Lancet Oncol. 2013;14:749-59

10.0 CHANGES TO THE STATISTICAL ANALYSIS PLAN

Major changes in Version 2 include:

- Aligned the SAP with protocol version 4
- Addition of the secondary endpoint of PK modeling for drug interactions between encorafenib, cetuximab, and binimetinib to the Phase 3.
- CCI [redacted]
- Addition of a Japanese SLI cohort of 6 patients.

Major changes in Version 3 include:

- Aligned the SAP with protocol version 5
- Addition of OS to the SLI portion as an exploratory endpoint.
- Subgroup analyses for Japan added.

Major changes in Version 4 include:

- Aligned the SAP with protocol version 6
- Addition of a primary objective and endpoint of confirmed ORR by BICR of Triplet Arm vs. Control Arm to the randomized portion of the study.
- Modification of the planned interim analysis of OS (Triplet vs Control) to include boundaries for both superiority and (non-binding) futility; the timing of this analysis was modified to occur at the same time as the primary analysis of the ORR endpoint specified above.
- Addition of BICR of patients' tumor imaging data to support the Phase 3 primary endpoint of confirmed ORR per BICR (Triplet Arm vs. Control Arm) as well as secondary efficacy analyses of ORR, progression-free survival (PFS), duration of response (DOR) and time to response.
- Addition of PFS (per BICR and Investigator assessment) as a secondary objective and endpoint in the SLI portion of the study.
- Modification of the threshold for requiring central laboratory confirmation of *BRAF*^{V600E} mutation status prior to randomization.

- Modification of algorithm to determine number of missing tumor assessments required for PFS analyses
- Updated format of summary tables
- Clarified analysis sets to be used for efficacy analyses
- Updated variables to be included in multivariate regression models of efficacy
- Updated AESI analyses
- Subgroup analyses for Korea added.

Major changes in Version 5 include:

- Aligned the SAP with protocol version 7
- All references to "co-primary" objectives and endpoints have been changed to "primary" objectives and endpoints.
- Table 7-5 was added to detail the number of expected Triplet vs. Control OS events, the cumulative alpha spent, and the cumulative power to reject the null hypothesis at the expected analysis timepoints for OS.
- The qualifier "approximately" was added to the citation of 90% power regarding the final Triplet vs. Control analysis of OS; this qualifier had been inadvertently omitted from this section.