Document Type: Final Statistical Analysis Plan

Document Date: 31 January 2024

Study Title: A multicenter, randomized, open-label, 2-arm, Phase II study with a safety lead-in phase evaluating the combination of encorafenib and cetuximab versus irinotecan/cetuximab or infusional 5-fluorouracil (5-FU)/folinic acid (FA)/irinotecan (FOLFIRI)/cetuximab in Chinese patients with BRAF V600E mutant metastatic colorectal cancer.

Protocol Reference Number: W00090GE202

NCT Number: NCT05004350

STATISTICAL ANALYSIS PLAN

Protocol Title:

A multicenter, randomized, open-label, 2-arm, Phase II study with a safety lead-in phase evaluating the combination of encorafenib and cetuximab versus irinotecan/cetuximab or infusional 5-fluorouracil (5-FU)/folinic acid (FA)/irinotecan (FOLFIRI)/cetuximab in Chinese patients with *BRAF* ^{V600E} mutant metastatic colorectal cancer.

Protocol Number: W00090GE202

Compound Number: Encorafenib in combination with Cetuximab

Short Title: The NAUTICAL CRC study.

Sponsor Name: Pierre Fabre Médicament

Legal Registered Address: Les Cauquillous, 81500 Lavaur, France

Regulatory Agency Identifier Number(s):

Registry **ID** 2021LP00253

Signature Page

I have read the statistical analysis plan for the NAUTICAL CRC study dated 31JAN2024 and confirm that to the best of my knowledge it accurately describes the planned analyses for the study.

	Title	Name (First & Last name)	Consistency between SAP and study protocol	Date	Signature
PPD					

Table of Contents

Version History

This Statistical Analysis Plan (SAP) for the NAUTICAL CRC study is based on the protocol Version 5.0 dated 10Jul2023.

SAP Version	Approval Date	Change	Rationale
1	08Jul2021	Not Applicable	Original version
2	21Apr2023	- Cover page (Sponsor's address change, Regulatory Agency Identifier Number(s)) - Updates some sponsor approvers in signature page - Covance changed to Labcorp - Update of Schedule of Activities (Table 3, Table 4, Table 5) - Definition of Screened Participants Set instead of Enrolled - Other minor adjustments, eg to facilitate programming - Additional PFS sensitivity analysis considering interruption due to COVID-19 as intercurrent events - Updates in analysis of vaccine-emergent AEs and deaths - Updates on companion diagnosis analysis	- Reference to protocol version 3.0 dated 10Jan2022 (non substantial PA01) which has following description of changes: "The overall rationale for the changes is to implement an additional exploratory objective to support a companion diagnostic development in China, then the main other changes are the change of End of Study definition for clarification, the addition of Lipase and Amylase as routine tests, the removal of Bicarbonate as routine test, the change of exclusion criteria related to HBV to be in accordance with the Centers for Disease Control and Prevention (cdc.gov), the change of Visual Acuity measurement according to local practice, the revision of the information related to contraception to be in accordance with local labels for Standard of Care Treatment, the addition of instructions for tumor tissue to be sent for central testing, the addition of prescreening for patients with local BRAF positive testing in case the maximum of indeterminate or discordant central results is reached at site or study level". - Reference to protocol version 4.0 dated 22Dec2022 (substantial PA02) which has following description of changes: "The overall rationale of the changes is to add instructions related to COVID19 restriction and the impact on study treatment management when temporary interruption is needed, some clarifications have been added to be consistent between sections or to avoid any misinterpretation, to correct some discrepancies between sections and to correct some typo errors".
3	29Jan2024	1. Labcorp changed to Fortrea.	1. Compagny name changed.

SAP Version	Approval Date	Change	Rationale
		2. Update of definition for cut-off date for main analysis as per protocol V5.0 3. Adjustments following dry-run analysis 4. Handling of participants continuing study treatment beyond progression (as allowed by protocol) 5. Removal of summary tables for ophthalmological examinations 6. PK: replacement of BLQ with BQL to match data & an additional parameter AUC_{tau} is also added. 7. Use of medical review to identify prior and subsequent lines, sub-categories of subsequent lines of therapy. Categories are collected in CRF for the post-study drug. The sub-categories were defined based on medical review. 8. Disease Characteristics, Stage at study entry, items collected in eCRF as Stages IVM1a, IVM1b and IVM1c to be considered as Stage IV and that only categories/items will be: I, II, III and IV. 9. Per Protocol population definition clarification 10. Modification of the definition of AE occurrence	2. Reference to protocol version 5.0 dated 10Jul2023 (substantial PA03) which has following description of changes: "The overall rational of the changes is to adjust the cut-off date definition for primary analysis considering that number of events expected to trigger this analysis (n=74) may not be reached due to a high number of censored patients for primary endpoint. The main other changes are related to the update of the encorafenib Investigator Brochure dated may 2023." 3. Dry-run outputs review and comments implemented in SAP when needed for more clarification. 4. Planned per protocol to continue the treatment after progression analysis have been aligned to take this into consideration. 5. Considering limited number of post-screening assessments for ophthalmological examinations, no summary tables will be provided. 6. Terminology updates. AUC_{tau} added for analysis purposes. 7. Prior and subsequent lines and sub-categories of subsequent lines of therapy are not collected in eCRF. The objective of the medical review was to clarify the number of prior/subsequent lines received before study entry/after study treatment based on the regimens entered in CRF. 8. Medical needs 9. Per Protocol population definition adjusted for more clarity 10. Modification of the definition of AE occurrence for more relevance

1. Introduction

This document describes the statistical analyses and data presentations to be performed for the clinical protocol W00090GE202 entitled "A multicenter, randomized, open-label, 2-arm, Phase II study with a safety lead-in phase evaluating the combination of encorafenib and cetuximab versus irinotecan/cetuximab or infusional 5-fluorouracil (5-FU)/folinic acid (FA)/irinotecan (FOLFIRI)/cetuximab in Chinese patients with *BRAF* ^{V600E} mutant metastatic colorectal cancer".

This Statistical Analysis Plan (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 analyses in population as outlined in the protocol. The summaries/analyses of baseline characteristics and demographics, important protocol deviations, derivations, will be described in appendix. The purpose of this SAP is to ensure the credibility of the study findings by specifying the statistical approaches for all analyses of study data.

This SAP will be approved and signed prior to the clinical database lock for any analyses specified in this SAP.

This 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 1.0 dated 21 April 2021. Any further changes to the protocol or eCRFs may necessitate updates to the SAP (see section 6.2)

The Table Of Contents (TOC) of Tables Figures and Listings (TFLs) will be provided in a separate document, as well as the shell TFLs and specifications.

The study W0090GE202 also called NAUTICAL CRC study is a Phase II, multicenter, randomized, open-label, parallel group, two-arm bridging study with a Safety Lead-In (SLI) phase to evaluate the combination of the BRAF inhibitor encorafenib and Epidermal Growth Factor Receptor (EGFR) inhibitor cetuximab (referred to as the "doublet") versus irinotecan and cetuximab or 5-fluorouracil/folinic acid+irinotecan (FOLFIRI) and cetuximab (referred to as the "control") in Chinese participants with B-RAF proto-oncogene, serine/threonine kinase V600E-mutant *(BRAF* V600E) metastatic Colorectal Cancer (mCRC) whose disease has progressed after one or two prior regimens in the metastatic setting.

1.1. Objectives and Endpoints

Objectives and associated endpoints are provided in Table 1 below for the SLI phase.

Table 1: Objectives and Endpoints in the SLI phase

	Estimands	
Objectives	Endpoints	Other attributes
Primary		
To assess the safety and tolerability of the doublet.	• Incidence of Dose-Limiting Toxicities (DLTs) during the DLT-evaluation period (which is the first 28 days after the first dose of study intervention in the SLI). DLTs are defined in Section 1.2.2.	**Population**: all participants with either a local or a central laboratory confirmed result of BRAF V600E-mutant mCRC, as defined by the screening inclusion/exclusion criteria in Protocol Section 5.1.1 and Section 5.1.2 which correspond to the target population of the SLI part of the study. **Intercurrent events**: hypothetical strategy will be applied for the intercurrent events. The intercurrent event is treatment discontinuation for reasons other than treatment-related toxicity that leads to < 75% of the planned dose of each study intervention during the DLT evaluation period. Participants without DLTs and with the intercurrent event will not be included in the DLT rate calculation. **Population-level summary measure**: DLT rate defined as the number of DLT-evaluable participants with DLTs in the DLT-evaluation period divided by the number of DLT-evaluable participants.

Secondary		
To assess the overall safety and tolerability of the doublet	• Type and severity of Adverse Events (AEs) and Serious Adverse Events (SAEs), changes in physical examinations, vital signs, electrocardiograms (ECGs), clinical safety laboratory assessment values; dermatological examinations and performance status using the Eastern Co-operative Oncology Group (ECOG) performance status scale. • Incidence of dose interruptions, dose modifications and discontinuations due to AEs.	Not applicable
To provide preliminary PK data of the doublet.	• Plasma concentrations of encorafenib and serum concentrations of cetuximab on Day 1 of Cycles 1 and 2 and derived PK parameters (Area Under the Curve (AUC), Minimum Concentration (C_{min}), Maximum Concentration (C_{max})) in Chinese participants.	Not applicable
To provide preliminary antitumor activity data of the doublet as measured by Overall Response Rate (ORR) and Duration Of Response (DOR)	• ORR, defined as the proportion of participants with a best overall best response of either Complete Response (CR) or Partial Response (PR) as determined by Blinded (to treatment received) Independent Central Review (BICR) and investigator assessment per RECIST Version 1.1.	Not applicable

	• DOR, defined as the time from first documented response (CR or PR) to the earliest date of disease progression as determined by BICR and investigator assessment per RECIST Version 1.1, or death due to underlying disease.	
Exploratory		
[redacted]		
• To assess the predictive significance of the Microsatellite Instability (MSI) status.	• MSI status in Formalin-fixed and Paraffin Embedded (FFPE) samples using established Polymerase Chain Reaction (PCR) assays in tumor sample versus germline control at screening.	

Objectives, endpoints and estimands are provided in Table 2 below for the randomized phase.

Table 2: Objectives and Endpoints in the Randomized (Phase II) phase

	Estimands	
Objectives	Endpoints	Other attributes
Primary		
To compare the efficacy of the doublet versus control as measured by Progression Free Survival (PFS) assessed by BICR.	• PFS defined as the time from the date of randomization to the earliest documented date of disease progression as determined by BICRassessment per RECIST Version 1.1, or death due to any cause.	**Population**: all participants with either a local or a central laboratory confirmed *BRAF* V600E-mutant mCRC, as defined by the screening inclusion/ exclusion criteria in Protocol Section 5.1.1 and Section 5.1.2, which corresponds to the target population of the Phase 2 part of the study. The Full Analysis Set will be the population for the primary analysis of the primary endpoint. **Intercurrent events**: hypothetical strategy will be applied for the new anticancer therapy and for more than two missing assessments. Treatment policy will be applied for tolerability and duration of study treatment. Details of the intercurrent events and censoring rules for the primary analysis are summarized in Section 5.3.2.2. **Population-level summary measure**: hazard ratio for PFS and corresponding 2-sided Confidence Interval (CI) based on Cox proportional hazard model stratified by ECOG performance status (0 versus 1) and prior use of irinotecan (yes versus no) at baseline (collected via the Interactive Response Technology (IRT)). PFS will be compared between the two treatment arms using a 1-sided stratified log-rank test.

Secondary		
To compare the efficacy of the doublet versus the control with respect to: o PFS by investigator assessment o Overall Response Rate (ORR). o Duration Of Response (DOR). o Disease Control Rate (DCR). o Time To Response (TTR). o Overall Survival (OS).	• PFS, defined as the interval of time between the date of randomization to the earliest date of disease progression, as determined by investigator assessment per RECIST v1.1, or death due to any cause. • ORR (for confirmed and unconfirmed responses), defined as the proportion of participants with a confirmed (resp. unconfirmed) Best Overall Response (BOR) of either Complete Response (CR) or Partial Response (PR) as determined by BICR and investigator assessment per RECIST Version 1.1. • DOR, defined as the time from the date of the first documented response (CR or PR) to the earliest date of disease progression as determined by BICR and investigator assessment per RECIST Version 1.1, or death due to any cause. • DCR, defined as the proportion of participants with a best overall	Not applicable

	response of either CR, PR or SD, as determined by BICR and investigator assessment per RECIST v1.1 • TTR (for confirmed and unconfirmed responses), defined as the time between the date of randomization until the first documented response of CR or PR per RECIST Version 1.1. • OS, defined as the time from randomization until date of death due to any cause.	
• To characterize the safety and tolerability of the doublet	• Type and severity of adverse events and SAEs according to National Cancer Institute-Common Terminology Criteria for Adverse Event (NCI-CTCAE) Version 4.03, changes in physical examinations, vital signs, ECGs, clinical safety laboratory assessment values, dermatological examinations and performance status using the ECOG performance status scale.	Not applicable

• To assess the effect of the doublet on QoL.	• Change from baseline in the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire for Cancer Patients (EORTC QLQ-C30), EuroQoL Five Dimensions, Five Levels (EQ-5D-5L), Functional Assessment of Cancer Therapy-Colon Cancer (FACT-C) and Patient Global Impression of Change (PGIC) questionnaire scores.	Not applicable
• To characterize the PK of the doublet in this population.	• Plasma concentrations encorafenib and serum concentrations of cetuximab on Day 1 of Cycles 1 and 2 and derived PK parameters (e.g. AUC, C_{min}, C_{max}; 15 participants with serial sampling only) in Chinese participants. • Population PK analysis using the ARRAY-808-302 study PK data.	Details of the population PK analysis will be provided in a specific standalone modelling plan. The modelling results will be reported in a separate report.

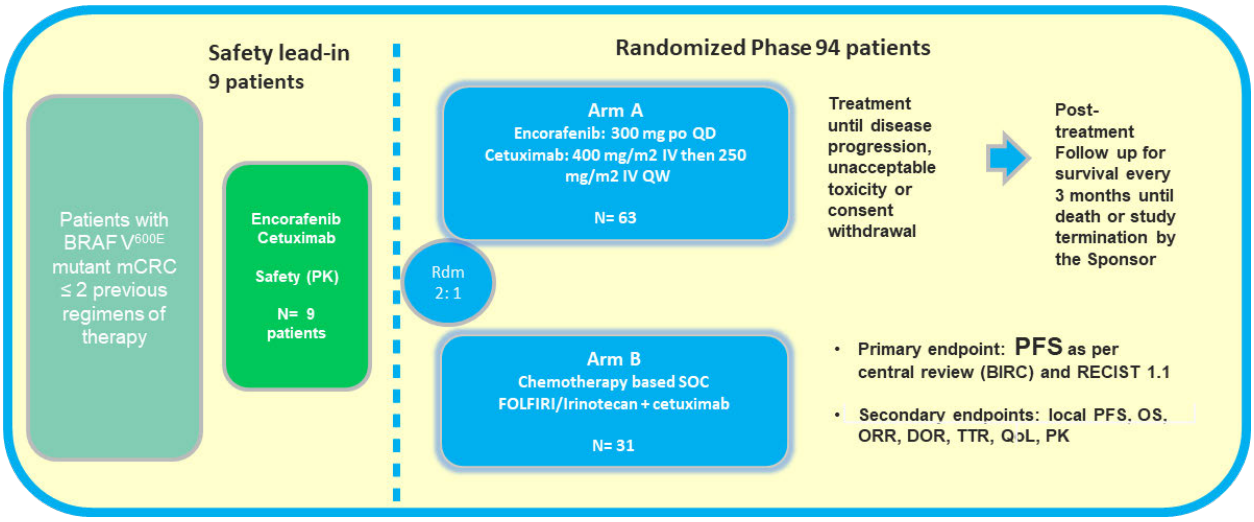
Exploratory		
[redacted]		• Not applicable
• To assess the predictive significance of the MSI status.	• MSI status in FFPE samples using established PCR assays in tumor sample versus germline control at screening.	
• To investigate the concordance of *BRAF* V600E mutation detection in archived or fresh tumor sample used for participant selection with the presence of this mutation as detected by a candidate companion diagnostic.	• Detection of *BRAF* V600E mutation in tumor tissues by a candidate companion diagnostic based on NGS or PCR methods in a central laboratory to explore the concordance of *BRAF* V600E mutation status between the central BRAF testing used for patient selection and the candidate companion diagnostic.	

1.2. Study Design

This is a Phase II, multicenter, randomized, open-label, parallel group, two-arm bridging study with a SLI phase to evaluate the doublet versus irinotecan and cetuximab or FOLFIRI and cetuximab in Chinese participants with *BRAF* V600E mCRC whose disease has progressed after one or two prior regimens in the metastatic setting.

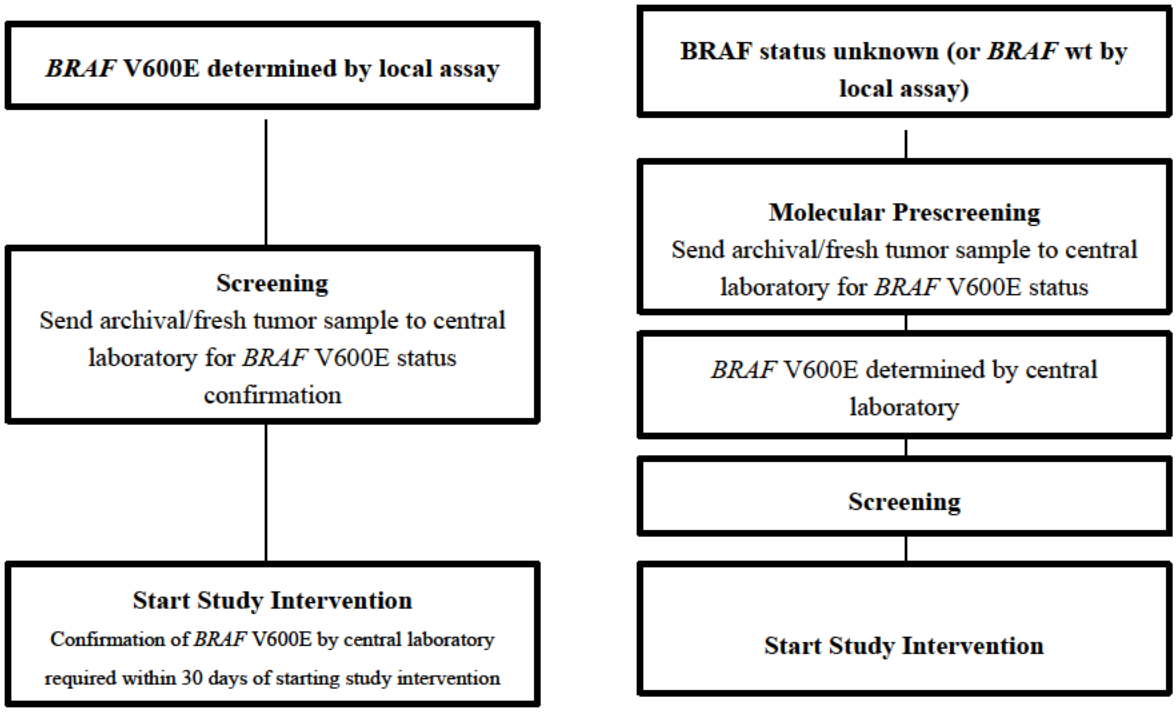
The study schema is shown in the Figure 1 below:

Figure 1: Study Schema for W00090GE202 SLI Phase and Randomized Phase



Participants will be eligible for the study based on identification of a *BRAF* V600E mutation in the tumor as determined by a local assay result obtained any time before screening or by the central laboratory as part of molecular prescreening. If the participant is enrolled based on local assay results, the BRAF V600E status must be confirmed by the central laboratory no later than 30 days from first dose of study intervention (see Figure 2 below). Full details are provided in Protocol Section 8.1.1.

Figure 2: Identification of BRAF V600E Mutation



Abbreviations: *BRAF* V600E = B-RAF proto-oncogene serine/threonine kinase V600E mutant; *BRAF* wt =BRAF wild type.

The study contains a SLI phase to assess the safety and tolerability of the doublet prior enrolling participants in the randomized (Phase II) phase:

- **SLI phase:** The SLI phase will be conducted at a limited number of study sites. A total of nine evaluable participants will be assigned to treatment on a rolling basis in a single cohort. All nine participants will receive 28-day cycles of encorafenib once daily (QD) and cetuximab once weekly, see Protocol Section 6.1.

  DLTs will be evaluated across the first cycle of therapy and the tolerability of the doublet will be assessed by the sponsor or sponsor representative and the investigator by regular communication.

  An independent Data Monitoring Committee (DMC) will review both the DLTs and cumulative toxicity at two timepoint:

  - At the time the first three participants of the SLI has completed the first cycle.
  - at the time the last participant of the SLI has completed the first cycle.

  The randomized phase of the study will start as soon as the last participant recruited has completed at least one full cycle (28 days) of treatment and following a tolerability assessment by an independent DMC.

Due to early recruitment faster than expected, the screening/enrolment have been put on hold while the first 3 participants were under treatment in the SLI phase and an ad-hoc independent DMC was performed to ensure the safety and the tolerability of participants.

Further details are provided in Section 5.8.1.

- **Randomized (Phase II) phase:** Eligible participants will be randomized in a 2:1 ratio to the doublet arm or control arm, respectively:
  - Doublet arm: 63 participants will receive encorafenib QD and cetuximab once weekly in 28-day cycles (see Protocol Section 6.1).
  - Control arm: 31 participants will receive either irinotecan and cetuximab or FOLFIRI and cetuximab (at the investigator's discretion) in 28-day cycles. Irinotecan or FOLFIRI will be given every 2 weeks and cetuximab once weekly (see Protocol Section 6.1). The choice of irinotecan or FOLFIRI must be declared before randomization.

Randomization will be stratified by baseline ECOG performance status (0 versus 1) and prior use of irinotecan (yes versus no). Randomization will be performed using stratified randomization list with blocks.

The independent DMC will review the available safety information after the first 15 participants in the randomized phase of the study treated with the doublet have completed at least one cycle of treatment to confirm tolerability and then will be responsible for reviewing safety data at regular intervals (at a minimum of every 6 months).

Main analysis will be conducted after the occurrence of the 74th event.

Final analysis will be conducted at the end of study which is defined as the timepoint when all participants have been followed for at least 1 year after the start of study intervention of the last participant randomized and when 70% of the expected deaths in the randomized phase are recorded (whichever occurs last). Intervention after the end of the study is discussed in Protocol Section 6.7.

The primary Clinical Study Report (CSR) will be based on the main analysis. An addendum to the CSR will be written once the end of the study has been reached (final analysis). Additional data cutoff dates for other analyses may be established if requested by the competent authorities.

1.2.1. Study Procedures and Assessments

Participants without a locally available *BRAF* V600E result may undergo molecular tumor prescreening with the central laboratory BRAF mutation assay at any time before screening as long as they meet all the molecular prescreening inclusion/exclusion criteria (see Protocol Section 5.1.1 and Section 5.1.2). For participants with a locally available BRAF V600E result, a representative tumor specimen (primary or metastatic, archival or newly obtained) will be provided for central BRAF testing.

Participants will be screened for eligibility in the 28 days before the start of study intervention.

Eligible participants will enter the treatment phase made up of weekly clinic visits. Study intervention will be administered until death, disease progression (in special circumstances, continuation of treatment beyond disease progression may be allowed, see Protocol Section 6.1) or one of the other predefined criteria for study intervention discontinuation is met (see Protocol Section 7.1).

An end of treatment visit will be completed at the time of study intervention discontinuation (as soon as possible and ≤14 days after the last dose of study intervention).

There will be a follow-up period after the end of treatment. A safety follow-up visit will be performed approximately 30 days after the last dose of study intervention or before the initiation of subsequent anticancer therapy, whichever occurs first. Participants will then enter survival follow-up and will be followed for subsequent anticancer therapies, disease progression following the initiation of subsequent therapies and survival status every 3 months (or more frequently as needed) until death, one of the other criteria for participant discontinuation from the study (see Protocol Section 7.2), the participant is lost to follow-up (see Protocol Section 7.3) or the end of study is reached (see Protocol Section 4.4). Any participants still receiving study intervention at the end of the study will be allowed to continue at the discretion of the investigator and as long as none of the criteria for study intervention discontinuation are met.

- For participants who definitely discontinue study intervention due to disease progression, survival follow-up will start after the 30-day safety follow-up.
- For participants who definitely discontinue study intervention for reasons other than disease progression, survival follow-up will start upon disease progression or the start of another anticancer therapy.

Tumor assessments by radiological imaging (e.g. computed tomography [CT], magnetic resonance imaging (MRI), X-ray, different methods of whole-body bone scans) will be performed at baseline (within 28 days of the first dose in the SLI phase or within 28 days of randomization for the randomized phase), every 6 weeks (±7 days) for the first 24 weeks, then every 12 weeks (±7 days) thereafter until disease progression (in special circumstances, continuation of treatment beyond disease progression may be allowed, see Protocol Section 6.1) or death, withdrawal of consent, initiation of subsequent anticancer therapy, participant is lost to follow-up or the end of study is reached (whichever occurs first). Tumor response will be determined locally by the investigator and by blinded (to treatment received) independent central review using RECIST Version 1.1.

Patient reported outcomes will be assessed using the EORTC QLQ-C30, EQ-5D-5L and FACT-C questionnaires and PGIC.

Safety assessments include recording of adverse events and concomitant medications, physical examinations, dermatological assessments, clinical safety laboratory assessments (including pregnancy testing for women of childbearing potential [WOCBP]), vital signs, ECGs, cardiac function tests (where required) and ECOG performance status.

Blood samples for characterization of the PK of encorafenib and cetuximab will be collected on the first day of treatment (Cycle 1 Day 1) and at steady state after 1 month treatment (Cycle 2 Day 1) for participants receiving encorafenib only (SLI phase and doublet arm of the randomized phase). Serial blood samples will be collected from a subset of 24 participants treated at the recommended dose at selected sites (all nine participants in the SLI phase and the first 15 participants at the selected sites in the doublet arm of the randomized phase). Sparse sampling will be performed in all remaining participants in the doublet arm.

Exploratory assessments include CCI [redacted] determination of microsatellite instability (MSI) status. The *BRAF* V600E, *RAS* wt status and deoxyribonucleic acid (DNA) mismatch repair (dMMR) system/MSI-High status of all participant's tumors will be confirmed centrally.

The schedule of activities for molecular prescreening and screening for all participants is shown in Table 3. The schedule of activities for the treatment and follow-up phases is shown in Table 4 for the SLI phase and the doublet arm of the randomized phase, and in Table 5 for the control arm of the randomized phase.

Table 3: Schedule of Events for Molecular Prescreening and Screening

Procedure/Assessment	Molecular Prescreening[a]	Screening	Randomization
	Any time before screening	Day –28 to–1	
Epoch	**PRESCREENING**	**SCREENING**	
Molecular prescreening informed consent	X		
Molecular prescreening inclusion/exclusion criteria	X		
Tumor sample (archival or fresh) for *BRAF* V600E, *RAS* wt status and MSI testing to be submitted for central laboratory testing	X	X	
Register in IRT system	X	X	
Screening informed consent		X	
Screening inclusion/exclusion criteria		X	
Demographics	X	X	
Height		X	
Medical and disease history		X	
Prior medications/therapies/procedures		X	
Complete physical examination [b]		X	
Weight		X	
Vital signs		X	
ECOG performance status		X	
EORTC QLQ-C30, EQ-5D-5L, FACT-C [h]		X	
ECG		X	
FSH (LH and/or estradiol)[c]		X	
HBV deoxyribonucleic acid (DNA), hepatitis B surface antigen (HBsAg), Hepatitis B core antibody, serum HCV ribonucleic acid (RNA), hepatitis C antibody (HCV Ab) and HIV testing		X	
Pregnancy test[c]		X	
Hematology[d]		X	
Clinical chemistry[e]		X	

Procedure/Assessment	Molecular Prescreening[a] Any time before screening	Screening Day –28 to–1	Randomization
Epoch	**PRESCREENING**	**SCREENING**	
Coagulation[f]		X	
Urinalysis[g]		X	
Tumor evaluation (CT scan, MRI)		X	
Blood sample for CRP		X	
CCI			
Concomitant medications/therapies		X	
Adverse events	X	X	X
Randomize participant using IRT			X[h]
Determination of comparator treatment in the control arm			X[i]

Abbreviations: ALT = alanine aminotransferase; ANC = absolute neutrophil count; aPTT = activated partial thromboplastin time; AST = aspartate aminotransferase; *BRAF* V600E = B-raf murine sarcoma viral oncogene homolog B1 V600E mutant; CCI ; CRP = C-reactive protein; CT = computed tomography; ECG = electrocardiogram; ECOG = Eastern Cooperative Oncology Group; EORTC QLQ-C30 = European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire; EQ-5D-5L = EuroQol-5D-5L; FACT C = Functional Assessment of Cancer Therapy-Colon Cancer; FOLFIRI = 5-fluorouracil/folinic acid+irinotecan; FSH = follicle-stimulating hormone; HIV = human immunodeficiency virus; INR = international normalized ratio; IRT = interactive response technology; LDH = lactate dehydrogenase; LH = luteinizing hormone; MRI = magnetic resonance imaging; MSI = microsatellite instability; pH = hydrogen ion concentration; PT = prothrombin time; *RAS* wt = rat sarcoma viral oncogene homolog wild type; RBC = red blood cell; SLI = safety lead-in; WBC = white blood cell.

[a] If local *BRAF* V600E mutation test positive, proceed to screening phase.

[b] Includes cardiovascular, respiratory, gastrointestinal, dermatological, ophthalmological and neurological systems (at a minimum). Ophthalmic examination will include visual acuity (Standard Logarithmic Visual Acuity Chart), tonometry (intraocular pressure), slit lamp examination, and fundoscopy. Physical examinations should be targeted as clinically indicated at subsequent visits.

[c] Serum test for women of childbearing potential only (FSH [LH and/or estradiol] measurements, if applicable).

[d] Erythrocytes (RBC), hematocrit, hemoglobin, platelets; leukocytes (WBC) count with differential: basophils, eosinophils, lymphocytes, monocytes, neutrophils/ANC.

[e] ALT, albumin, alkaline phosphatase, AST, lipase, amylase, bilirubin (total and indirect), blood urea nitrogen/urea, calcium, chloride, creatine kinase, creatinine, glucose, LDH, magnesium, potassium, sodium, total protein, troponin I or T, uric acid.

[f] aPTT, INR or PT.

[g] Blood, glucose, ketones, leukocytes, pH, protein.

Procedure/Assessment	Molecular Prescreening[a]	Screening	Randomization
	Any time before screening	Day –28 to–1	
Epoch	**PRESCREENING**	**SCREENING**	

[h] Does not apply to participants in the SLI phase. Only for participants in the randomized phase.
[i] For participants in the control arm of the randomized phase, the investigator will determine the best treatment option for the participant (either irinotecan and cetuximab or FOLFIRI and cetuximab). The treatment decision will be recorded in the IRT system.

Table 4: Schedule of Events for SLI Phase and Doublet Arm of the Randomized (Phase II) Phase

Procedure/Assessment (± 3-day window for procedures/assessments)	Treatment Phase									Follow-up Phase	
	Cycle 1				Subsequent Cycles[a]				End of Treatment[b]	30-day Safety Follow-up[c]	Survival Follow-up
	Day 1	Day 8	Day 15	Day 22	Day 1	Day 8	Day 15	Day 22			
Epoch	TREATMENT									FOLLOW-UP	LONGTERM FOLLOW-UP
Baseline/eligibility assessments:											
Verify inclusion/exclusion criteria	X[d]										
Medical history from screening to Cycle 1 Day 1	X										
Dosing:											
BSA	X[e]				X						
Encorafenib	X				X						
Issue new encorafenib dosing diary	X				X						
Review encorafenib dosing diary from previous cycle and assess compliance					X				X		
Cetuximab IV infusion	X	X	X	X	X	X	X	X			
Efficacy/PK/biomarker assessments:											
Tumor evaluation (CT scan, MRI)	Every 6 weeks (±7 days) from first dose for the first 24 weeks, then every 12 weeks[f] (±7 days)										
EORTC QLQ-C30, EQ-5D-5L, FACT-C, PGIC[g]	X				X				X	X	
PK blood samples	X[h]				X[h]						
CCI											

Procedure/Assessment (± 3-day window for procedures/assessments)	Treatment Phase									Follow-up Phase	
	Cycle 1				Subsequent Cycles[a]				End of Treatment[b]	30-day Safety Follow-up[c]	Survival Follow-up
	Day 1	Day 8	Day 15	Day 22	Day 1	Day 8	Day 15	Day 22			
Epoch	TREATMENT									FOLLOW-UP	LONGTERM FOLLOW-UP
CCI											
Blood sample for MSI testing	X										
Safety assessments:											
Targeted physical examination	X[e]				X				X	X	
Weight	X[e]				X				X	X	
Dermatological examination	X				X[i]				X	X	
Vital signs	X	X	X	X	X	X	X	X	X	X	
ECG[j]	X		X		X				X	X	
ECOG performance status	X[e]				X				X	X	
Pregnancy test[k]	X[e]				X				X	X	
Hematology[l]	X[e]		X	X	X				X	X	
Clinical chemistry[m]	X[e]		X		X				X	X	
Coagulation[n]	X[e]				X				X	X	
Urinalysis[o]	X[e]				X				X	X	
Adverse events	Assess continuously[p]										
Concomitant medications/therapies	Assess continuously										
Follow-up assessments:											

Procedure/Assessment (± 3-day window for procedures/assessments)	Treatment Phase									Follow-up Phase	
	Cycle 1				Subsequent Cycles[a]				End of Treatment[b]	30-day Safety Follow-up[c]	Survival Follow-up
	Day 1	Day 8	Day 15	Day 22	Day 1	Day 8	Day 15	Day 22			
Epoch	TREATMENT									FOLLOW-UP	LONGTERM FOLLOW-UP
Survival status										X	Every 3 months[q]
Documentation of subsequent anticancer therapy										X	Every 3 months[q]
Documentation of disease progression after subsequent anticancer therapy											Every 3 months[q]

Abbreviations: ALT = alanine aminotransferase; ANC = absolute neutrophil count; aPTT = activated partial thromboplastin time; AST = aspartate aminotransferase; BSA = body surface area; CCI CT = computed tomography;ECG = electrocardiogram; ECOG = Eastern Cooperative Oncology Group; eCRF = electronic case report form; EORTC QLQ-C30 = European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire for cancer patients; EQ-5D-5L = EuroQol-5D-5L; FACT-C = Functional Assessment of Cancer Therapy-Colon Cancer; INR = international normalized ratio; IV = intravenous(ly); LDH = lactate dehydrogenase; MRI = magnetic resonance imaging; MSI = microsatellite instability; PGIC = Patient Global Impression of Change; pH = hydrogen ion concentration; PK = pharmacokinetic(s); PT = prothrombin time; RBC = red blood cell; SLI = safety lead-in; WBC = white blood cell; WOCBP = women of childbearing potential.

[a] If a cycle initiation (CnD1 visit) is postponed due to toxicity, corresponding safety assessments must be recorded as unscheduled visit(s) in the eCRF. Whatever the reason, the CnD1 will be the day of the treatment initiation of cycle "n". If a visit is missed within the cycle (CnD8, CnD15, CnD22), following visits should be performed according to the original schedule.

[b] To be performed at the time of study intervention discontinuation (as soon as possible and ≤14 days after the last dose of study intervention).

[c] To be performed approximately 30 days after the last dose of study intervention or before the initiation of subsequent anticancer therapy, whichever occurs first. Following the 30-day safety follow-up, when clinically appropriate, it is recommended participants be monitored with dermatological examinations and chest CT scans for cutaneous and non-cutaneous secondary malignancies for up to 6 months after the last encorafenib dose or until initiation of another antineoplastic therapy.

[d] Inclusion/exclusion criteria are to be verified before the first dose of study intervention.

Procedure/Assessment (± 3-day window for procedures/assessments)	Treatment Phase									Follow-up Phase	
	Cycle 1				Subsequent Cycles[a]				End of Treatment[b]	30-day Safety Follow-up[c]	Survival Follow-up
	Day 1	Day 8	Day 15	Day 22	Day 1	Day 8	Day 15	Day 22			
Epoch	TREATMENT									FOLLOW-UP	LONGTERM FOLLOW-UP

[e] Procedures do not have to be repeated if performed within 72 hours before Cycle 1 Day 1 (i.e. first day of dosing).

[f] Tumor evaluations are to be performed every 6 weeks (±7 days) from first dose for the first 24 weeks, then every 12 weeks (±7 days) until disease progression or death, withdrawal of consent, initiation of subsequent anticancer therapy, participant is lost to follow-up or the end of study is reached (whichever occurs first). If a participant discontinues study intervention for reasons other than disease progression, then tumor assessments must continue to be performed according to this schedule.

[g] The questionnaires should be completed only by the participants in randomized phase at the beginning of the study visit before receiving any study intervention, before any other study assessment or consultation with the investigator and before being informed of their current disease status. Note: PGIC is not completed on Cycle 1 Day 1.

[h] Samples for encorafenib and cetuximab PK will be collected during Cycle 1 and Cycle 2 only. Serial PK sampling (predose and at 1, 2, 4 and 6 hours postdose) on Day 1 of both cycles will be performed in 24 participants treated at the recommended dose: all nine participants in the SLI phase and the first 15 participants in the randomised phase at selected sites. Sparse sampling (at 2 and 6 hours postdose on Cycle 1 Day 1 and predose and 2 hours postdose on Cycle 2 Day 1) will be collected in the remaining participants.

[i] Dermatological examinations are to be performed every 8 weeks from Cycle 1 Day 1 (i.e. on Day 1 of Cycles 3, 5, 7…).

[j] Electrocardiograms are to be performed in triplicate predose on Cycle 1 Day 1 (conducted within approximately 5 to 10 minutes total time), followed by a single ECG at remaining timepoints. On Cycle 1 Day 1 and Cycle 2 Day 1, ECGs are to be performed predose and at 2 (+ 0.5) hours after administration of encorafenib and before the start of the cetuximab infusion. Electrocardiograms are performed predose at remaining timepoints. Electrocardiograms should be performed before PK and biomarker blood collection at equivalent nominal timepoints.

[k] Local urine test for WOCBP only.

[l] Erythrocytes (RBC), hematocrit, hemoglobin, platelets; leukocytes (WBC) count with differential: basophils, eosinophils, lymphocytes, monocytes, neutrophils/ANC.

[m] ALT, albumin, alkaline phosphatase, AST, lipase, amylase, bilirubin (total and indirect), blood urea nitrogen/urea, calcium, chloride, creatine kinase, creatinine, glucose, LDH, magnesium, potassium, sodium, total protein, troponin I or T, uric acid.

[n] aPTT, INR or PT.

[o] Blood, glucose, ketones, leukocytes, pH, protein.

[p] All adverse events are collected from when the participant first provides informed consent until the 30-day safety follow-up visit.

Procedure/Assessment (± 3-day window for procedures/assessments)	Treatment Phase									Follow-up Phase	
	Cycle 1				Subsequent Cycles[a]				End of Treatment[b]	30-day Safety Follow-up[c]	Survival Follow-up
	Day 1	Day 8	Day 15	Day 22	Day 1	Day 8	Day 15	Day 22			
Epoch	TREATMENT									FOLLOW-UP	LONGTERM FOLLOW-UP

[q] For participants who discontinue study intervention due to disease progression, the survival follow-up phase will start after the 30-day safety follow-up. For participants who discontinue study intervention for reasons other than disease progression, the survival follow-up phase will start upon disease progression or the start of another anticancer therapy. Participants will be contacted by telephone approximately every 3 months for collection of information during the survival follow-up period. This may be conducted at routine visits as well, and more frequently as needed. Participants will be followed for at least 1 year after the start of study intervention of the last participant randomized. For participants that were lost to follow-up or withdrawal of consent, attempts to determine survival status will be made via access to public records as permitted by local laws.

Table 5: Schedule of Events for Control Arm of the Randomized (Phase II) Phase

Procedure/Assessment (± 3-day window for procedures/assessments)	**Treatment Phase**									**Follow-up Phase**	
	Cycle 1				**Subsequent Cycles[a]**				**End of Treatment[b]**	**30-Day Follow-up[c]**	**Survival Follow-up**
	Day 1	**Day 8**	**Day 15**	**Day 22**	**Day 1**	**Day 8**	**Day 15**	**Day 22**			
Epoch	**TREATMENT**									**FOLLOW-UP**	**LONGTERM FOLLOW-UP**
Baseline/eligibility assessments:											
Verify inclusion/exclusion criteria	X[d]										
Medical history from screening to Cycle 1 Day 1	X										
Dosing:											
BSA	X[e]				X						
Irinotecan IV infusion (all participants)	X		X		X		X				
5-FU[f] and FA IV infusion (FOLFIRI/cetuximab only)	X		X		X		X				
Cetuximab IV infusion	X	X	X	X	X	X	X	X			
Efficacy/ biomarker assessments:											
Tumor evaluation (CT scan, MRI)	Every 6 weeks (±7 days) from first dose for the first 24 weeks, then every 12 weeks[g](±7 days)										
EORTC QLQ-C30, EQ-5D-5L, FACT-C, PGIC[h]	X				X				X	X	
CCI											
Blood sample for MSI testing	X										
Safety assessments:											
Targeted physical examination	X[e]				X				X	X	
Weight	X[e]				X				X	X	
Vital signs	X	X	X	X	X	X	X	X	X	X	
ECG[i]	X		X		X				X	X	
ECOG performance status	X[e]				X				X	X	

Procedure/Assessment (± 3-day window for procedures/assessments)	Treatment Phase									Follow-up Phase	
	Cycle 1				Subsequent Cycles[a]				End of Treatment[b]	30-Day Follow-up[c]	Survival Follow-up
	Day 1	Day 8	Day 15	Day 22	Day 1	Day 8	Day 15	Day 22			
Epoch	TREATMENT									FOLLOW-UP	LONGTERM FOLLOW-UP
Pregnancy test[j]	X[e]				X				X	X	
Hematology[k]	X[e]		X	X	X		X[l]		X	X	
Clinical chemistry[m]	X[e]		X		X				X	X	
Coagulation[n]	X[e]				X				X	X	
Urinalysis[o]	X[e]				X				X	X	
Adverse events	Assess continuously[p]										
Concomitant medications/therapies	Assess continuously										
Follow-up assessments:											
Survival status										X	Every 3 months[q]
Documentation of subsequent anticancer therapy										X	Every 3 months[q]
Documentation of disease progression after subsequent anticancer therapy											Every 3 months[q]

Abbreviations: ALT = alanine aminotransferase; ANC = absolute neutrophil count; aPTT = activated partial thromboplastin time; AST = aspartate aminotransferase; BSA = body surface area; CCI CT = computed tomography; ECG = electrocardiogram; ECOG = Eastern Cooperative Oncology Group; eCRF = electronic case report form; EORTC QLQ-C30 = European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire for cancer patients; EQ-5D-5L = EuroQol-5D-5L; FA = folinic aid; FACT-C = Functional Assessment of Cancer Therapy-Colon Cancer; FOLFIRI = 5-fluorouracil/folinic acid+irinotecan; 5-FU = 5-fluorouracil; INR = international normalized ratio; IV = intravenous(ly); LDH = lactate dehydrogenase; MRI = magnetic resonance imaging; MSI = microsatellite instability; PGIC = Patient Global Impression of Change; pH = hydrogen ion concentration; PK = pharmacokinetic(s); PT = prothrombin time; RBC = red blood cell; WBC = white blood cell; WOCBP = women of childbearing potential.

[a] If a cycle initiation (CnD1 visit) is postponed due to toxicity, corresponding safety assessments must be recorded as unscheduled visit(s) in the eCRF. Whatever the reason, the CnD1 will be the day of the treatment initiation of cycle "n". If a visit is missed within the cycle (CnD8, CnD15, CnD22), following visits should be performed according to the original schedule.

[b] To be performed at the time of study intervention discontinuation (as soon as possible and ⩽14 days after the last dose of study intervention).

Procedure/Assessment (± 3-day window for procedures/assessments)	Treatment Phase									Follow-up Phase	
	Cycle 1				Subsequent Cycles[a]				End of Treatment[b]	30-Day Follow-up[c]	Survival Follow-up
	Day 1	Day 8	Day 15	Day 22	Day 1	Day 8	Day 15	Day 22			
Epoch	TREATMENT									FOLLOW-UP	LONGTERM FOLLOW-UP

[c] To be performed approximately 30 days after the last dose of study intervention or before the initiation of subsequent anticancer therapy, whichever occurs first.
[d] Inclusion/exclusion criteria are to be verified before the first dose of study intervention.
[e] Procedure does not have to be repeated if performed within 72 hours before Cycle 1 Day 1 (i.e. first day of dosing).
[f] The initial 5-FU dose is given as a bolus (not to exceed 15 minutes) on Days 1 and 15, followed by continuous IV infusion over 46 to 48 hours (2 days) or according to study site standards.
[g] Tumor evaluations are to be performed every 6 weeks (±7 days) from first dose for the first 24 weeks, then every 12 weeks (±7 days) until disease progression or death, withdrawal of consent, initiation of subsequent anticancer therapy, participant is lost to follow-up or the end of study is reached (whichever occurs first). If a participant discontinues study intervention for reasons other than disease progression, then tumor assessments must continue to be performed according to this schedule.
[h] The questionnaires should be completed by the participants at the beginning of the study visit before receiving any study intervention, before any other study assessment or consultation with the investigator and before being informed of their current disease status. Note: PGIC is not completed on Cycle 1 Day 1.
[i] Electrocardiograms are to be performed in triplicate predose on Cycle 1 Day 1 (conducted within approximately 5 to 10 minutes total time), followed by a single ECG predose at remaining timepoints. Electrocardiograms should be performed before PK and biomarker blood collection at equivalent nominal timepoints.
[j] Local urine test for WOCBP only.
[k] Erythrocytes (RBC), hematocrit, hemoglobin, platelets; leukocytes (WBC) count with differential: basophils, eosinophils, lymphocytes, monocytes, neutrophils/ANC.
[l] A blood sample for hematology is to be collected on Day 15 during Cycle 2 onwards only while participants continue to receive irinotecan.
[m] ALT, albumin, alkaline phosphatase, AST, lipase, amylase, bilirubin (total and indirect), blood urea nitrogen/urea, calcium, chloride, creatine kinase, creatinine, glucose, LDH, magnesium, potassium, sodium, total protein, troponin I or T, uric acid.
[n] aPTT, INR or PT.
[o] Blood, glucose, ketones, leukocytes, pH, protein.
[p] All adverse events are collected from when the participant first provides informed consent until the 30-day safety follow-up visit.
[q] For participants who discontinue study intervention due to disease progression, the survival follow-up phase will start after the 30-day safety follow-up. For participants who discontinue study intervention for reasons other than disease progression, the survival follow-up phase will start upon disease progression or the start of another anticancer therapy. Participants will be contacted by telephone approximatively every 3 months for collection of information during the survival follow-up period. This may be conducted at routine visits as well and more frequently as needed. Participants will be followed for at least 1 year after the start of study intervention of the last participant randomized. For participants that were lost to follow-up or withdrawal of consent, attempts to determine survival status will be made via access to public records as permitted by local laws.

1.2.2. Assessment of Tolerability in SLI Phase

Participants in the SLI phase will be assessed for DLTs during the first cycle of treatment with the doublet. A DLT is defined as any adverse event or abnormal laboratory value assessed as unrelated to disease, disease progression, intercurrent illness or concomitant medications/therapies resulting in the inability to tolerate at least 75% dose intensity [(administered dose in mg/planned dose in mg) x 100] of encorafenib or cetuximab occurring within the first 28 days of study intervention (Cycle 1) that satisfies at least one of the criteria listed in Table 6.

Table 6: Criteria for Defining DLT

Cardiac disorders • Supraventricular tachycardia - includes, but not limited to, extrasystoles and sinus tachycardia Grade ≥3.
General disorders and administration site conditions • Fatigue Grade 3 for >14 consecutive days.
Respiratory disorders • Interstitial lung disease/pneumonitis Grade ≥2.
Skin and subcutaneous tissue disorders • Rash, HFSR or photosensitivity Grade 3 for >14 consecutive days despite maximal skin toxicity treatment (according to local practice). • Rash, HFSR or photosensitivity Grade 4.
Gastrointestinal disorders • Diarrhea Grade 3 for ≥48 hours despite optimal use of antidiarrheal therapy. • Diarrhea Grade 4. • Nausea/vomiting Grade 3 for ≥48 hours despite optimal use of antiemetic therapy. • Nausea/vomiting Grade 4.
Investigations • Total bilirubin increase Grade ≥3. • AST or ALT increase Grade ≥3 in conjunction with total bilirubin Grade ≥2 of any duration. • AST or ALT increase Grade 3 for >7 consecutive days. • AST or ALT increase Grade 4. • Serum creatinine increase Grade ≥3. • ANC decrease Grade 4 for >7 consecutive days. • Platelet count decrease Grade 3 with signs of clinically significant bleeding. • Platelet count decrease Grade 4. • ECG QTcF prolonged ≥Grade 3[a].

Eye disorders **Uveitis** • Grade 3 for >21 consecutive days confirmed by ophthalmologic examination • Grade 4 confirmed by ophthalmologic examination
Eye disorders –Visual disturbances without ocular (retinal) changes • Blurred vision, flashing lights, floaters: Grade ≥3.
Eye disorders – (other specify) • Grade 3 for >21 consecutive days. • Grade 4 confirmed by ophthalmic examination.
Other hematologic and non-hematologic toxicities[b] Any other Grade ≥3 adverse event except lymphocyte count decreased (lymphopenia) Grade ≥3 unless clinically significant.
Abbreviations: ALT = alanine aminotransferase; ANC = absolute neutrophil count; AST = aspartate aminotransferase; DLT = dose-limiting toxicities; ECG = electrocardiogram; HFSR = hand-foot skin reaction; NCI CTCAE = National Cancer Institute Common Terminology Criteria for Adverse Events; QTcF = QT interval corrected for heart rate using Fridericia's formula. Grades according to NCI CTCAE grading system. [a] QTcF must be prolonged on two separate ECGs. [b] Isolated laboratory changes (e.g. alkaline phosphatase, cholesterol, lipase, serum amylase) or those due to sampling or laboratory errors without associated clinical signs or symptoms may be determined to not be DLTs upon review and agreement by the investigator and sponsor's medical monitor.

A total of nine participants will be assigned to treatment in the SLI phase on a rolling basis in a single cohort. To be evaluable for tolerability assessment by the independent DMC prior to the randomized phase of the study, participants must:

- Experienced an event meeting the DLT criteria or
- Received ≥75% dose intensity [(administered dose in mg/planned dose in mg) x 100] of encorafenib and cetuximab during Cycle 1.

Non-evaluable participants will be replaced in order to achieve nine evaluable participants in the SLI phase.

Enrollment will be put on hold if at least three participants experience DLTs until a discussion with the independent DMC can occur.

If DLTs are observed in less than three of nine evaluable participants, the independent DMC will review data after the ninth participant has been followed for at least one 28-day cycle (Cycle 1). All data available at the time of review will be included (as these participants will be enrolled over time, the independent DMC will have cumulative data available beyond the first 28 days on some participants). If the independent DMC determines the doses to be tolerable in the first nine evaluable participants based on observing DLTs in <33% participants and evaluation of the overall toxicity profile, the study will proceed to the randomized phase. The independent DMC will make a recommendation if the doses are not deemed tolerable.

2. Statistical Hypotheses

Because of the exploratory nature of the SLI phase, all data will be analysed descriptively and not as inferential issues. No formal statistical testing will therefore be performed.

The main aim in the randomized phase is to compare the efficacy of the doublet arm versus the control arm, as measured by the primary endpoint of PFS by BICR where $HR_{PFS\ (doublet\ vs\ control)}$ is the hazard ratio for PFS of the doublet arm versus the control arm. The following statistical hypotheses will be tested to address the primary objectives:

$$H_0: HR_{PFS}\ (doublet\ vs\ control) \geq 1 \text{ vs } H_1: HR_{PFS}\ (doublet\ vs\ control) < 1$$

which corresponds to the rejection of the null hypothesis (the PFS time for the doublet arm is less than or equal to the PFS time of the control arm).

Only the test on the primary endpoint, on which is based the sample size justification, may lead to a causal interpretation. All other statistical results are to be regarded within an exploratory perspective.

2.1. Adjustment for multiplicity

No adjustment for multiplicity is needed in the study.

3. Sample Size Determination

The planned sample size is of 103 participants (nine treated in the SLI phase and approximately 94 randomized in the Phase II part). Additional participants may be included in SLI phase if deemed necessary. Approximately 147 participants will be screened to achieve an estimated total of 103 participants.

3.1.1. SLI Phase

Participants will be assigned to the starting dose of encorafenib 300 mg QD and cetuximab 400 mg/m^2 initial dose then 250 mg/m^2 once weekly thereafter. The dose for encorafenib and cetuximab will be considered acceptable if the observed Cycle 1 DLT rate is <33% (i.e. less than three participants with DLTs) out of nine evaluable participants (see Section 1.2.2 for the definition of a DLT). To be evaluable participants must meet the criteria for the Dose-determining Set (DDS), see Section 4.4.

A comparison of the operating characteristics for this dose-testing rule with nine participants and traditional 3+3 rules is shown in Table 7 below. The results illustrate the benefit of the additional participants as the probability of falsely declaring a dose to be toxic is lowest with a nine-participant cohort when the true DLT rate is ≤20%. Similarly, the probability of correctly declaring a dose to be toxic is higher with the nine-participant cohort when the true DLT rate is ≥40%. In addition, observing no Cycle 1 DLTs in nine participants would be expected to occur with probability 0.040 if the true DLT rate is 30%.

Table 7: Operating Characteristics of SLI Criteria for Nine Participants 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 Nine Participants
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
Abbreviations: DLT = dose limiting toxicity.		

If the independent DMC determines the doses to be tolerable in the nine evaluable participants based on evaluation for potential DLTs and the evaluation of the overall toxicity profile based on accumulated safety data, then the study will proceed to the randomized phase of the study.

3.1.2. Randomized Phase

The objective is to evaluate the consistency of the ARRAY-808-302 study data in the Chinese population. The study design therefore aims to detect a difference between the two arms (doublet arm and control arm) in terms of PFS. In the control arm, the median PFS is estimated to be 2 months. The present study is powered to detect a PFS benefit that would result in a 2-month median difference between arms, i.e. from 2 to 4 months; this corresponds to a hazard ratio of 0.5.

A one-sided log-rank test will be used, with the following parameters: alpha=0.025, beta=0.2.

Applying a 2:1 randomization and considering 10% of lost to follow-up a total number of 94 participants (63 participants in the doublet arm and 31 in the control arm) will be randomized in the study. With a monthly accrual rate of five participants, the duration of recruitment will be approximately 18.8 months. Under these hypotheses, 74 events should occur 21 months after the first randomized participants. The cutoff date for the main analysis will be set at the occurrence of the 74th event or when all participants without PFS event (disease progression per BICR or death) will have had the opportunity to perform their tumor assessment planned at week 36, whichever comes first. PFS events considered here are those as defined in Section 5.3.2.2 (main analytical approach of primary endpoint).

Note that the power will depend on the number of PFS events that will be observed at the time of the main analysis (for example: if only 65 events are observed at the cut-off date due to high number of censored patients, power is estimated at 75%).

The number of participants and timing of the analyses were determined using EAST6.4.

The independent DMC will continue to monitor for DLTs in the first 15 evaluable participants who are randomized to receive encorafenib and cetuximab, as well as continue to evaluate the safety profile of this combination therapy as per the DMC charter.

4. Analysis Sets

The following populations will be analyzed:

4.1. Screened Participants Set

All participants who signed the main (screening) ICF for the study.

4.2. Full Analysis Set (FAS)

- For the SLI phase

All participants who receive at least one dose of study intervention.

- For the Randomized phase

All randomized participants. Participants will be analyzed according to the study intervention assigned at randomization.

4.3. Safety Set (SS)

All participants who receive at least one dose of study intervention. Participants will be analyzed according to treatment received. There will be separate Safety Sets for the SLI phase and the randomized phase.

4.4. Dose-Determining Set (DDS)

All SLI phase participants from the Safety Set who either completed a minimum exposure requirement and have sufficient safety evaluations or experienced a DLT.

A participant is considered to have met the minimum exposure requirement if he received at least 75% of the planned dose of each study intervention during Cycle 1.

Participants 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 investigator to have enough safety data to conclude that a DLT did not occur.

4.5. Efficacy Set (ES)

All participants in the randomized phase and included in the FAS with a centrally confirmed mCRC BRAF V600E mutation. Participants will be analyzed according to the study intervention assigned at randomization.

4.6. Per Protocol Set (PPS)

All participants in the randomized phase and included in the FAS who are considered sufficiently compliant with the protocol requirements. Participants having important protocol deviation with major impact on the primary criterion analyses will be excluded from PPS (See section 6.8).

Additional reasons of exclusion from PPS will be:

- Not treated participants;
- Participants with a negative central BRAF assessment;
- Participants with intercurrent events related to COVID-19 defined as “at least one study drug interruption due to COVID-19 of more than 4 consecutive weeks (28 days) or more than 6 weeks (42 days) overall”.

4.7. PK Set

All participants who receive at least one dose of encorafenib or cetuximab, and who have at least one post-dose PK blood collection with associated bioanalytical results (See Section 5.7.1.1.4).

5. Statistical Analyses

5.1. General Considerations

All analyses will be performed for the SLI and by treatment arm for the randomized phase of the study separately. More precisely, the following columns will be detailed for baseline and characteristics analyses:

- Randomized and SLI (pooled for encorafenib + cetuximab),
- SLI (encorafenib + cetuximab)
- Randomized Phase
    - Encorafenib + Cetuximab
    - Control
    - Total Randomized Phase

For efficacy analyses, the following columns will be displayed for the Randomized phase only:

- Encorafenib + Cetuximab
- Control

For safety analyses, the following columns will be displayed:

- Randomized and SLI (pooled for encorafenib + cetuximab),
- SLI (encorafenib + cetuximab)
- Randomized Phase
    - Encorafenib + Cetuximab
    - Control

Any table displayed on DDS will only show SLI column.

All efficacy analyses will be performed using the FAS, and all safety analyses will be performed using the Safety Set unless otherwise specified. All data will be listed using specific population as specified in the list of TFLs. In listings based on Screened Participants Set or FAS, participants who are excluded from SS (i.e. participants not treated) will be flagged (*).

For summary statistics by visit, only scheduled visits will be summarized. Re-test and unscheduled assessments will not be taken into account for summaries of post-baseline visit. For analyses of worst post-baseline assessment, all on-treatment values are considered regardless scheduled or unscheduled.

For TFLs, visits will be identified as collected in CRF.

In general, following descriptive methods will be used to present all relevant data:

- Continuous data will be presented using number of observations, number of missing (if any), mean, standard deviation, median, Q1, Q3, minimum and maximum. 95% Confidence Intervals (CI) will be presented if relevant.
- Categorical data will be summarized using number of observations, frequencies, percentages and number of missing (if any missing). Unless otherwise specified, the calculation of proportions will be based on the non-missing data. Counts of missing observations will be excluded from the denominator. For time-to-event data, Greenwood formula will be used for CIs of Kaplan-Meier estimates.

For primary PFS analysis (described in Section 5.3.2.2) and OS analysis (described in Section 5.4.1.2.2), hypothesis of proportional hazards will be checked through the following graphical methods:

- Visually by means of plots of log(-log(survival)) versus log of survival time, and look for parallelism;
- And/or by Schoenfeld residuals, including a LOESS curve. This will be plotted to investigate graphically violations of the proportional hazards assumption. Schoenfeld residuals will be computed in SAS using the PHREG procedure and using the OUTPUT statement and the keyword RESSCH. With proportional hazards the LOESS curve should be parallel to the X-axis.

If the performed checks do not suggest the validity of the proportional hazards assumption, the Wilcoxon test will be used instead of the Log-rank test. In this case Cox models will not be performed and hazard ratio will not be provided.

Formal statistical testing will be performed on the primary endpoint in the randomized phase. The overall significance level will be 0.025 (1-sided) and all statistical testing will be performed as 1-sided. CIs will be 2-sided with a confidence level of 95%, if not otherwise specified. The SLI phase results, the secondary and exploratory endpoints, sensitivity analyses and subgroup analyses in the randomized phase will be reported with descriptive statistics only.

The statistical analysis will be performed after the database lock or snapshot using Statistical Analysis Software (SAS®) version 9.4.

5.1.1. Pooling of Centers

In order to provide overall estimates of treatment effects, data will be pooled across study centers. The “center” factor will not be considered in statistical models or for subgroup analyses due to the high number of participating centers in contrast to the anticipated small number of participants randomized at each center.

5.1.2. Stratification Factors

The analysis of the randomized phase will be stratified using stratification factors collected via the IRT: ECOG performance status (0 versus 1) and prior use of irinotecan (yes versus no).

For the randomized phase, both stratification factors will be summarized on the FAS population based on values collected via the IRT and also based on values collected in CRF, along with investigator's choice of control arm collected in IRT.

A shift table will be provided for stratification factors in randomized phase (IRT versus CRF values).

CRF value of prior use of irinotecan (yes versus no) will be derived from the CRF form "Prior Anti-neoplastic Therapy – Medication" using WHO Drug coding.

CRF value of ECOG performance status at baseline will be derived as specified in Section 5.1.3 (safety assessment).

5.1.3. Baseline Definition

For all endpoints assessments in the SLI phase and all safety assessments in the randomized phase of the study, baseline is defined as the last completed and available assessment prior to date of first dose in the SLI phase and randomized phase, respectively. If an assessment that is planned to be performed prior to the first dose of study intervention in the protocol is performed on the same day as the first dose of study intervention and the time is unknown, it will be assumed that it was performed prior to study intervention administration and will be considered as the baseline assessment. For a participant who was not dosed, baseline is defined as the last completed and available assessment.

For efficacy assessments in the randomized phase of the study, baseline is defined as the last assessment prior to randomization. If the assessment that is planned to be performed prior to randomization in the protocol is performed on the same day as the date of randomization and assessment time point is missing, it will be assumed that it was performed prior to randomization and will be considered as the baseline assessment. If such a value is missing, the last measurement prior to the first dose of study intervention will be used as the baseline measurement for the efficacy analysis, except for analyses of tumor assessments data where the baseline assessment would be considered missing.

Unscheduled assessments will be used in the determination of baseline. Data reported at the end of treatment visit are not eligible for baseline selection.

ECG Baseline will be the mean of the last triplicate measurements performed before the first study intervention administration. If no triplicate available, then baseline will be the last single ECG measurement performed before the first study intervention administration. For the interpretation parameter (normal/abnormal and its related clinical significance), the baseline will be the last

completed and available assessment performed before the first study intervention administration, regardless triplicate or single ECG.

5.1.4. Definition of Treatment Period

The on treatment period is defined as the time from the first dose of study intervention to last dose of study intervention + 30 days, using below definitions:

First dose of study intervention is defined, according to the treatment arm, as

- Doublet arm: the earliest date between first administration of encorafenib and first administration of cetuximab
- Control arm, depending on which control regimen is administrated:
    - the earliest date between first administration of irinotecan and first administration of cetuximab
    - or, the earliest date between first administration of FOLFIRI (i.e. earlier date between Folinic acid, 5-FU and irinotecan) and first administration of cetuximab

Last dose of study intervention is defined according to the treatment arm, as:

- Doublet arm: the latest date between last administration of encorafenib and last administration of cetuximab.
- Control arm, depending on which control regimen is administrated:
    - the latest date between last administration of irinotecan and last administration of cetuximab
    - the latest date between last administration of FOLFIRI (i.e. latest date between Folinic acid, 5-FU and irinotecan) and last administration of cetuximab

Of note, for any participant continuing the study treatment beyond progression (as allowed by protocol), date of last dose will be the date of treatment discontinuation reported in the CRF End of Treatment visit corresponding to the treatment discontinuation due to progressive disease observed under study treatment and leading to EoT visit.

Data collected after progressive disease will be listed in the appropriate listings with a flag to identify the post progressive disease data.

Treatment period of each drug separately is described in Section 5.6.1.

5.2. Participant Dispositions

Participant dispositions will be summarized by treatment group and overall.

A first disposition table will include the number and percentage of participants among the Prescreened Participants for:

- Prescreened failure, with tabulation of the reasons.

A second disposition table will include the number and percentage of participants among the Screened analysis set for Screened failure, with tabulation of the reasons.

A third disposition table will display the number and percentage of participants included in each study population (analysis set) among the FAS.

Another disposition table will display the number and percentage of participants by reason of discontinuation from the treatment phase among the FAS. Same table will include disposition of 30 Days Safety Follow-up and Survival Follow-up.

Corresponding data will also be listed on FAS.

Among Screened Participants Set, disposition overall and by center will also be provided, analyzing number of screened participants and FAS participants.

Study Dates and follow-up times will be summarized:

- FPFV - LPFV dates: FPFV is the treatment start date for the first treated participant in SLI and the randomization date for the first randomized participant in Randomized phase. LPFV is the treatment start date for the last treated participant in SLI and the randomization date for the last randomized participant in Randomized phase.
- Randomization dates: As SLI are not randomized, treatment assignment period is provided.
- Randomization period (months): Randomization period = (treatment start date for last participant treated - treatment start date for first participant treated + 1)/30.4375 for SLI, and (randomization date for last participant randomized - randomization date for first participant randomized + 1)/30.4375 for Randomized Phase.
- Analysis Cutoff date
- Duration between randomization (or first dose of study drug for SLI participants) and cutoff date (months)
- For randomized phase only, Follow-up time for PFS per BICR (months): Follow-up time = (Date of event or censoring - Date of randomization + 1)/30.4375. Date of censoring is the same as defined for the main analyses.
- For randomized phase only, Follow-up time for OS (months): Follow-up time = (Date of event or censoring - Date of randomization + 1)/30.4375. Date of censoring is the same as defined for the main analyses.

Missed visits will also be presented (i.e. visits reported as not performed in CRF), overall and by visit label.

Data collected in the CRF form “Covid-19 Impact” will be listed. Summary of Covid-19 (SARS-CoV-2) impact will be provided with number and percentages of FAS participants impacted by: Visit, Case Report Form, Category.

A summary table will be provided for:

- Treatment discontinuations related to Covid-19 (SARS-CoV-2): This will be defined as an event reported in “Covid-19 Impact form” having CRF form entered as “Disposition - End of Treatment” and visit name entered as “End of Treatment”. It will be presented overall and by reason for treatment discontinuation.
- Missed visits due to Covid-19 (SARS-CoV-2), overall and by visit label: This will be defined as a visit not performed with reason verbatim containing “COVID-19”.

Number of participants by Inclusion/ Exclusion criteria not met will be provided for:

- Prescreen failure participants,
- Screen failure participants.

The following listings will also be provided:

- Inclusion/ Exclusion criteria not met in prescreen failure participants and in screen failure participants.
- Informed Consent in prescreen failure participants and screened participants.
- Vendor (Clario) comments in prescreen failure participants and screened participants.

5.3. Primary Endpoint(s) Analysis

For the SLI Phase, the primary endpoint is the incidence of DLTs.

For the randomized Phase, the primary endpoint is the PFS assessed by BICR.

5.3.1. Definition of Endpoint(s)

5.3.1.1. SLI Phase

DLT will be defined as per Table 6 in section 1.2.2. DLTs will be identified using the DLT flag reported in the Adverse Event CRF form.

5.3.1.2. Randomized Phase

PFS will be defined as the time from the date of randomization to the earliest documented date of disease progression as determined by BICR assessment per RECIST Version 1.1, or death due to any cause, whichever occurs first. PFS will be calculated in months:

PFS_{BICR} (months) = (event date or censor date – randomization date + 1) / 30.4375

Derivation of event and censor dates is presented in section 5.3.2.2.

5.3.2. Main Analytical Approach

5.3.2.1. SLI Phase

The occurrence of DLTs will be summarized descriptively for the DDS:

- The number and proportion of participants experiencing DLTs during the DLT evaluation period (defined as the first 28 days after the first dose of study intervention in the SLI, i.e. from Day 1 to Day 28 inclusive) will be summarized and listed for the DDS.
- The number and proportion of participants by Type of DLT, by Preferred Term, by Grade, by Action Taken.

All reported DLT will be listed for the DDS, including also any DLT occurring outside of DLT evaluation period (i.e. prior to first administration of study intervention or posterior to Day 28).

Analyses of other safety parameters will be performed as described in Section 0.

5.3.2.2. Randomized Phase

As preliminary, dates of all tumor assessments collected in CRF will be listed on FAS.

The primary estimand is built to answer to the primary objective and evaluate PFS by BICR. The intercurrent events will be taken into account for the primary analysis as defined in Table 2.

If no event (disease progression or death) is observed or if an event is observed two or more consecutive missing assessments after the last adequate tumor assessment, the participant will be censored on the date of the last adequate tumor assessment, prior to cutoff date, that documented no progression.

Given the scheduled visit assessment scheme (every 6 weeks (±7 days) for the first 24 weeks and every 12 weeks (±7 days) thereafter) the definition of "two or more consecutive missing assessments" will change, based on the number of days since the last adequate tumor assessment before the event (referred as "previous adequate tumor assessment"). Following rules will apply, based on the largest distance (relative to the last adequate tumor assessment date) that a participant can have a tumor assessment and not be classified as having missed 2 assessments:

- If the previous adequate tumor assessment is <= study day 21 then the largest distance defining accepted missing assessments will equate to 91 days since the previous adequate tumor assessment, allowing for a late assessment (i.e. 2 x 6 weeks + 1 week for a late assessment = 13 weeks = 91 days).
- If the previous adequate tumor assessment is > study day 21 and < = study day 118 then the largest distance defining accepted missing assessments will equate to 98 days since the previous adequate tumor assessment, allowing for early and late visits (i.e. 2 x 6 weeks + 1 week for an early assessment + 1 week for a late assessment = 14 weeks = 98 days.
- If the two missing assessments occur over the period when the scheduled frequency of tumor assessments changes from every 6 weeks to every 12 weeks, i.e. if the previous adequate tumor assessment is > study day 118 and < = study day 160, the largest distance defining accepted missing assessments will equate to 140 days (i.e. take the average of 6 and 12 weeks which gives 9 weeks and then apply same rationale, hence 2 x 9 weeks + 1 week for an early assessment + 1 week for a late assessment = 20 weeks = 140 days).
- If the previous adequate tumor assessment is day 161 onwards (when the scheduling changes to twelve-weekly assessments), the largest distance defining accepted missing assessments will equate to 182 days (i.e. 2 x 12 weeks + 1 week for an early assessment + 1 week for a late assessment = 26 weeks = 182 days).

Similarly, the definition of "one or more consecutive missing assessment" (used for PFS sensitivity analysis defined in Section 5.3.3) will be based on following rules:

- If the previous adequate tumor assessment is <= study day 21 then the largest distance defining accepted missing assessments will equate to 49 days since the previous adequate tumor assessment, allowing for a late assessment (i.e. 1 x 6 weeks + 1 week for a late assessment = 7 weeks = 49 days).
- If the previous adequate tumor assessment is > study day 21 and < = study day 160 then the largest distance defining accepted missing assessments will equate to 56 days since the previous adequate tumor assessment, allowing for early and late visits (i.e. 1 x 6 weeks + 1 week for an early assessment + 1 week for a late assessment = 8 weeks = 56 days.

- If the previous adequate tumor assessment is day 161 onwards (when the scheduling changes to twelve-weekly assessments), the largest distance defining accepted missing assessments will equate to 98 days (i.e. 1 x 12 weeks + 1 week for an early assessment + 1 week for a late assessment = 14 weeks = 98 days).

The below table illustrates how to define if there were one or more (resp. two or more) missing assessments since the last adequate assessment. Some examples are also provided in below table showing the decision rules to classify a participant as having missed one or two assessments prior to event:

If the last adequate tumor assessment before the event date occurs between Study Days	**D1***	**D2***
1 – 21	49	91 See example 2A
22 – 118	56 See example 1A	98
119 – 160	56	140 See example 2B
>=161	98 See example 1B	182
Examples		
	Example 1A: ATA performed at Day 60 shows SD, at Day 110 shows disease progression. This PFS event will not be considered occurring after one or more missing tumor assessments because it is not far enough from previous ATA. Indeed Day 110 ≤ Day 60+56.	Example 2A: First post-baseline ATA at Day 10 shows SD. Subsequent ATA performed at Day 100 shows disease progression. This PFS event will not be considered occurring after two or more missing tumor assessments because it is not far enough from previous ATA. Indeed Day 100 ≤ Day 10+91.
	Example 1B: ATA performed at Day 56, Day 120 and Day 166 shows SD, at Day 270 shows disease progression. This PFS event will be considered occurring after one or more missing tumor assessments because it is too far from previous ATA. Indeed Day 270 > Day 166+98	Example 2B: ATA performed at Day 56 shows SD, at Day 120 shows SD, at Day 270 shows disease progression. This PFS event will be considered occurring after two or more missing tumor assessments because it is too far from previous ATA. Indeed Day 270 > Day 120+140.

ATA = Adequate Tumor Assessment;
*Di: largest distance (relative to the last adequate tumor assessment date) that a participant can have a tumor assessment and not be classified as having missed i assessment(s)

In addition, if a new anticancer therapy (medications or procedures) is started prior to an event, the participant will be censored on the date of the last adequate tumor assessment that documented no progression prior to the start of the new anticancer therapy. If a new anticancer therapy (medications or procedures) is started and no progression or death is observed, the participant will

be censored on the date of the last adequate tumor assessment prior to the start of the new anticancer therapy.

An adequate post-baseline assessment is defined as an assessment where an overall response of CR, PR, SD, non-CR/non-Progressive Disease (non CR/non-PD), or Progressive Disease (PD) is reported. Timepoints where the response is not evaluable or no assessment was performed are considered as missed/inadequate tumor assessments and will not be used for determining the censoring date.

Participants with no baseline tumor assessment (including participants with an inadequate baseline assessment, missing baseline assessment or with a first tumor assessment post-randomization but prior to treatment start) or with no adequate post-baseline tumor assessments within 12 weeks after the date of randomization will be censored on the day of randomization, unless the participant dies within 12 weeks of randomization, in which case, death will be an event on date of death.

Participants with several censoring reasons will be censored whichever occurs first.

The censoring and event date options to be considered for the PFS primary analysis are presented in **Table 8**.

If a PFS event is observed after a single missing or non-adequate tumor assessment, the actual date of event will be used. Participants without post-baseline tumor assessments but known to be alive will be censored at the time of the first administration of the study intervention.

PFS will be described in tabular with number of events and graphical format using Kaplan-Meier (KM) methods, reporting estimated median (in months) with 95% CI, 25^{th} and 75^{th} percentiles [Brookmeyer and Crowley 1982, Klein and Moeschberger 1997] and KM estimated probabilities (in percentages) with corresponding 95% CIs [Kalbfleisch and Prentice 2002] at selected timepoints (every 2 months). A Cox regression model stratified by randomization strata (collected via the IRT) will be used to estimate the hazard ratio and the corresponding 95% CI based on the Wald test. The distribution of PFS will be compared using a stratified log-rank test.

Furthermore, the pattern of censored data will be examined between the treatment arms by summarizing the number of PFS observations that are censored and by tabulating the reasons for censored observations:

- No Baseline Assessment
- No Adequate Post-baseline Assessment
- Subsequent Therapy Given (therapy ie medication or surgery or radiotherapy)
- Progression After 2 or more Missed Assessments
- Death After 2 or more Missed Assessments
- Last Adequate Assessment
- Withdrawal of Consent
- Ongoing Tumor Assessments

PFS and censoring reasons will also be listed on FAS.

Formal statistical testing will be performed as described in Section 2.

In the below table and for efficacy analyses, "new" or "subsequent" "antineoplastic therapy" or "anticancer therapy" both refers to any additional secondary antineoplastic medication or anti-cancer procedure (surgery or radiotherapy) prescribed for cancer treatment (palliative treatment conducted for AE will not be considered), as collected in following CRF forms: "Anti-Neoplastic Therapy Since Study Treatment Discontinuation – Medication", "Anti-Neoplastic Therapy Since Study Treatment Discontinuation – Surgery", "Anti-Neoplastic Therapy Since Study Treatment Discontinuation - Radiotherapy". Concomitant medications and concomitant procedures will however be reviewed by medical monitor, similarly to the subsequent therapies, in order to detect any curative anticancer therapy reported among the general concomitant therapies that would bias PFS analysis. If any is found, it will be identified and will impact efficacy analyses similarly to the therapies reported in the forms "Anti-Neoplastic Therapy Since Study Treatment Discontinuation".

Table 8: PFS Outcome and Event Dates – Primary Analysis

	Situation	**Event/Censoring Date** Participants with several censoring reasons will be censored whichever occurs first.	**Outcome**
No baseline tumor assessments (e.g. missing or inadequate baseline assessment, or when first tumor assessment is post-randomization but prior to treatment start)			
A1	No baseline assessment and death after the second scheduled post-baseline tumor assessment, or no death	Date of randomization	Censored
A2	No baseline assessment and death on or before the second scheduled post-baseline tumor assessment	Date of death	Event
No adequate post-baseline tumor assessments (without new antineoplastic therapy given)			
B1	Death on or before the second scheduled post-baseline tumor assessment	Date of death	Event
B2	Death after the second scheduled post-baseline tumor assessment or no death	Date of randomization	Censored
With baseline and adequate post-baseline tumor assessments (without new antineoplastic therapy given prior [or without] progression or death without progression observed)			
C1	Progression with zero or one missed/inadequate tumor assessment prior to progression	Date of progression	Event

C2	Death but no progression, with zero or one missed/inadequate tumor assessment prior to death	Date of death	Event
D1	Progression but two or more consecutively missed/inadequate tumor assessments prior to progression	Date of last adequate tumor assessment*	Censored
D2	Death without progression, but two or more consecutively missed/inadequate tumor assessments prior to death	Date of last adequate tumor assessment*	Censored
E	No progression, no death	Date of last adequate tumor assessment*	Censored
F	Treatment discontinuation due to 'Disease progression' without documented progression, i.e. clinical progression based on investigator claim	N/A (not considered as an event, participant without documented PD should be followed for progression after discontinuation of treatment)	Information ignored
Initiation of new antineoplastic therapy			
G	New antineoplastic therapy started before progression or death	Date of last adequate tumor assessment* prior to initiation of new antineoplastic therapy**	Censored
H	No adequate post-baseline tumor assessment and new antineoplastic therapy started prior to a death	Date of randomization	Censored
I	New antineoplastic therapy started without progression observed or death	Date of last adequate tumor assessment* prior to initiation of new antineoplastic therapy**	Censored

* Adequate tumor assessment has non-missing and non-unknown overall lesion response, as defined earlier in this section.

** In absence of such last adequate tumor assessment, date of randomization is used.

The terms 'adequate Tumor Assessments (TA)' and 'missing adequate TA' are defined earlier in this section, together with rules to determine whether there is one or two missing/inadequate TAs.

A bar chart of the censoring distributions (events and censor types) over time (]0-1],]1-2],]2-3] months, …) will also be provided.

In addition, for the FAS (by treatment group), a reverse KM analysis will be performed for PFS to estimate the median duration of potential follow-up (with 25th and 75th percentiles) as described

by Schemper and Smith (Schemper and Smith 1996). Number of events will be provided. Participants who had a PFS event for the purposes of the PFS analysis will be censored at the date of the event. Participants who were censored (e.g., lost to follow-up, withdrew consent, ongoing, etc.) in the PFS 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 PFS analysis will be used in this analysis. KM estimated probabilities (in percentages) with corresponding 95% CIs (Kalbfleisch and Prentice 2002) will be presented at the same time points as in the PFS analysis.

Listing of individual PFS by BICR will be provided.

Assessments by BICR will also be listed on FAS: RECIST assessment, assessment of target lesions, assessment of non-target lesions, assessment of new lesions.
Listing of RECIST assessments, by participant and by visit, will present the BOR only at the time it was first assessed for each participant.

5.3.3. Sensitivity Analysis

The following sensitivity analyses will be performed to further assess the comparisons described by the primary objective of the randomized phase.

- The PFS analysis will be repeated in the ES.
- The PFS analysis will be repeated in the PPS.
- The distribution of PFS in FAS will be compared between the treatment arms using an unstratified log-rank test and the hazard ratio (with associated 95% CI) resulting from an unstratified Cox model will be presented.
- Treatment policy estimand strategy will be used for intercurrent event of starting new anticancer therapy. PDs or deaths that occur after the start of new anticancer therapy will be considered as events.
- Treatment policy estimand strategy will be used for intercurrent event of missing two or more tumor assessments (see **Table 8**). PDs or deaths that occur after two or more missing assessments will be considered as events.
- The analyses for PFS will be repeated with a censoring rule that backdates events occurring after one or more missing tumor assessments. This backdating analysis assumes PD occurs at the next assessment, so the PFS event date will be set to be at 6 weeks (or 12 weeks) after the last adequate tumor assessment for events occurring after one or more missing assessments. Backdating process applies to PFS events being progressions, deaths are not impacted.
- If a PFS event does not occur exactly or on within the 7-day window of the scheduled day of assessment, the event date will be moved to the next scheduled assessment. If a participant is censored on a day that does not coincide with the scheduled assessment day (exactly on or within the ±7-day window), the censoring date will be moved back to the

previous scheduled assessment. This process applies to PFS events being progressions, deaths are not impacted.

- Hypothetical strategy will be applied for the intercurrent events related to COVID-19 defined as "at least one study drug interruption due to COVID-19 of more than 4 consecutive weeks (28 days) or more than 6 weeks (42 days) overall". Participants will be censored at the date of last adequate tumor assessment prior the end of 4th or 6th week of study drug interruption (i.e prior or equal to the 28th or the 42nd day of interruption). Censoring reasons will be summarized. If both an interruption of >28 consecutive days and a cumulative interruption of >42 days are reported, the first one which occurs will be censored.

Besides, the distribution of PFS in the FAS will be compared by adjusting with potential prognostic factors. A multivariate Cox proportional hazards models will be performed to evaluate the effect of confounding variables on PFS, adding the two randomization strata as stratification variables in the model. Following factors will be tested:

- Treatment group (Doublet versus Control)
- Gender (male versus female);
- Age (< 65yrs versus ≥ 65years);
- MSI status (high versus normal);
- Removal status of primary tumor (Complete Resection vs. Partial Resection/Unresected);
- C-reactive protein (CRP) baseline level (≤ Upper Limit of Normal (ULN) versus > ULN). ULN for CRP will be set to 0.01 g/L.
- [redacted]
- Side of tumor (left versus right, as per classification rules reported in Section 5.7.2.1)
- Number of organs involved based on Target and Non-target lesion assessment (<=2 versus >=3)
- Presence of liver metastases at baseline, based on Target and Non-target lesion assessment (yes versus no)
- Number of prior regimens in the metastatic setting (1 versus 2)
- Prior oxaliplatin (yes versus no) derived using WHO Drug coding of medications reported in CRF form "Prior Anti-Neoplastic Therapy – Medication"
- Any other covariate as discussed and approved by the Sponsor.

These factors will be first analyzed in separate univariate analyses. More specifically, each potential prognostic factor will be tested in an unstratified Cox model where the considered predictor will be the only covariate included (treatment group will not be included in the model).

Unstratified univariate analyses will be presented in a single table. Randomization strata (ECOG performance status (1 versus 0) / prior use of irinotecan (Yes versus No)) will be added in the list of factors for univariate analyses.

As this comparison is not part of the testing strategy, no formal testing will be performed.

5.3.4. Supplementary Analyses

Not Applicable

5.4. Secondary Endpoint(s) Analysis

For the SLI Phase, the secondary endpoints are:

- Type and severity of adverse events and SAEs,
- Changes in physical examinations, vital signs, ECGs, clinical safety laboratory assessment values, graded according to NCI-CTCAE Version 4.03;
- Dermatological examinations
- Performance status using the ECOG performance status scale.
- Incidence of dose interruptions, dose modifications and discontinuations due to adverse events.
- PK endpoints: Plasma concentrations of encorafenib and serum concentrations of cetuximab on Day 1 of Cycles 1 and 2 and derived PK parameters (AUC, C_{min}, C_{max}).
- ORR
- DOR

For the randomized Phase, the secondary endpoints are:

- PFS assessed by investigator assessment
- OS
- ORR
- DOR
- DCR
- TTR
- AE and SAE
- Physical Examinations
- Vital Signs
- ECGs
- Clinical safety laboratory
- Dermatological examinations
- ECOG performance status
- Quality of life with EORTC QLQ-C30, EQ-D-5L, FACT-C and PGIC questionnaire scores

5.4.1. Key Secondary Endpoint(s)

5.4.1.1. Definition of Endpoint(s)

5.4.1.1.1. SLI Phase

- ORR (including both confirmed and unconfirmed responses) will be defined as the proportion of participants with a BOR of either CR or PR as determined by BICR and investigator assessment per RECIST Version 1.1. Two sets of ORR will be considered, one for confirmed and one considering confirmed and unconfirmed responses, as defined in section 5.4.1.2.3.
- DOR will be defined for confirmed responders (CR or PR) only, as the time from the date of the first documented response (CR or PR) to the earliest date of disease progression as determined by BICR and investigator assessment per RECIST Version 1.1, or death due to any cause (and occurring before two or more consecutive missing assessments).

  Given the scheduled visit assessment scheme the definition of "two or more consecutive missing assessments" will change, based on the number of days since the last adequate tumor assessment before the event (referred as "previous adequate tumor assessment"). Rules specified in section 5.3.2.2 will apply, based on the largest distance (relative to the last adequate tumor assessment date) that a participant can have a tumor assessment and not be classified as having missed 2 assessments prior to event.

  Date of first documented response of CR or PR will be derived using assessment per RECIST Version 1.1 collected at each time point, either by BICR data, or by investigator assessment.

  Confirmed responders who do not have a PD or death date by the data cutoff date will be censored for DOR at their last adequate tumor assessment of CR, PR or SD before the cutoff date or last adequate tumor assessment of CR, PR or SD before date a subsequent anticancer therapy for mCRC is started.

  DOR (months) = (Progression date or death date / censor date – First documented response date + 1) / 30.4375
- Grades of clinical safety laboratory assessments and abnormal values for blood pressure will be derived using NCI-CTCAE Version 4.03.
- For vital signs, changes from baseline of continuous endpoints (SBP/ DBP/ temperature/ Heart rate/ Respiratory rate) will be derived: Value at the visit – Value at baseline.
- For all other endpoints, raw values will be used.

5.4.1.1.2. Randomized Phase

- PFS will be defined as the time from the date of randomization to the earliest documented date of disease progression, as determined by investigator assessment per RECIST Version 1.1, or death due to any cause, whichever occurs first.

  PFS_{inv} (months) = (event date or censor date – randomization date + 1) / 30.4375
- OS, defined as the time from randomization until date of death due to any cause, in months.

If a participant is not known to have died, survival will be censored at the date of last known date the participant was alive or at their last contact date whatever is earlier.

If a participant is known to have died, survival will be analyzed as event. Incomplete or missing death dates will be imputed as per section 6.9.5.

- ORR (including both confirmed and unconfirmed responses) will be defined as the proportion of participants with a BOR of either CR or PR as determined by BICR and investigator assessment per RECIST Version 1.1. Two sets of ORR will be considered, one for confirmed and one considering confirmed and unconfirmed responses, as defined in section 5.4.1.2.3.
- DOR will be defined for confirmed responders (CR or PR) only as the time from the date of the first documented response (CR or PR) to the earliest date of disease progression as determined by BICR and investigator assessment per RECIST Version 1.1, or death due to any cause (and occurring before two or more consecutive missing assessments).

  Given the scheduled visit assessment scheme the definition of "two or more consecutive missing assessments" will change, based on the number of days since the last adequate tumor assessment before the event (referred as "previous adequate tumor assessment"). Rules specified in section 5.3.2.2 will apply, based on the largest distance (relative to the last adequate tumor assessment date) that a participant can have a tumor assessment and not be classified as having missed 2 assessments prior to event.

  Date of first documented response of CR or PR will be derived using assessment per RECIST Version 1.1 collected at each time point, either by BICR data, or by investigator assessment.

  Confirmed responders who do not have a PD or death date by the data cutoff date will be censored for DOR at their last adequate tumor assessment of CR, PR or SD before the cutoff date or last adequate tumor assessment of CR, PR or SD before date a subsequent anticancer therapy for mCRC is started.

  DOR (months) = (Progression date or death date / censor date – First documented response date + 1) / 30.4375
- DCR will be defined as the proportion of participants with a BOR of either CR, PR or SD, as determined by BICR and investigator assessment per RECIST Version 1.1. Two sets of DCR will be considered, one for confirmed and one considering confirmed and unconfirmed responses.
- TTR (including both confirmed and unconfirmed responses), will be defined as the time in months between the date of randomization until the first documented response of CR or PR as determined by BICR and investigator assessment per RECIST Version 1.1. Two sets of TTR will be considered, one for confirmed and one considering confirmed and unconfirmed responses. Date of first documented response of CR or PR will be derived using assessment per RECIST Version 1.1 collected at each time point, either by BICR data, or by investigator assessment.

Participants who do not achieve a CR or PR by the data cutoff date will be censored as follows:

- Censored at the last adequate tumor assessment date when they do not have a PFS event. In this case, participants have not yet progressed so they theoretically still have a chance of responding;
- Censored at maximum follow-up (i.e., a unique date common to all participants, equal to time from 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 participant 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). LPLV is the latest date of last contact. FPFV and LPLV refers to randomized part only.

Participants who receive subsequent anticancer therapy prior to response will be censored at their last radiological assessment prior to initiation of subsequent anticancer therapy (medication or procedure).

- Grades of clinical safety laboratory assessments and abnormal values for blood pressure will be derived using NCI-CTCAE Version 4.03.
- For vital signs, changes from baseline of continuous endpoints (SBP / DBP / Temperature / Heart rate / respiratory rate) will be derived: Value at the visit – Value at baseline.
- Patient Reported Outcome (PRO)

  The QoL questionnaires (EORTC QLQ-C30, FACT-C, EQ-5D-5L, PGIC) will be scored according to their respective user guides/scoring manuals.

  - The EORTC QLQ-C30 incorporates nine multi-item scales: five functional scales (physical, role, cognitive, emotional and social); three symptom scales (fatigue, pain, nausea and vomiting); and a global health and QoL scale. Several single-item symptom measures are also included [Aaronson 1993]. English sample of Version 3.0 is shown in Section 6.12, along with scoring instructions. Of note, for functional scales, a higher value reflects a better level of function; but for symptoms scales, items with a higher value reflects worse symptoms. Moreover, high score for the global health/QoL status represents a high health/QoL.
  - The EQ-5D-5L essentially consists of the EQ-5D descriptive system and the EQ visual analogue scale (VAS). An English sample is provided in Section 6.13. The descriptive system has five dimension (mobility, self-care, usual activities, pain/discomfort and anxiety/depression), each is rated according to a five-point verbal rating scale (VRS) (1. no problems, 2. slight problems, 3. moderate problems, 4. severe problems and 5. extreme problems) and translated into a five-digit number that describes the participant's health state. For example, state 11111 indicates no problems on any of the 5 dimensions, while state 12345 indicates no problems with mobility, slight problems with washing or dressing, moderate problems with doing usual activities, severe pain

or discomfort and extreme anxiety or depression. There should be only one response for each dimension. For dimensions, ambiguous values (e.g, several boxes are ticked for a single dimension) should be treated as missing values.

An EQ-5D-5L index value is obtained based on participant responses to the 5 dimensions and applying population-assessed weights (coefficients) to each set of responses that define a unique health state. Coefficients for China will be used [Luo 2017, Table 4, recommended MULT8r: 8-parameter multiplicative model with a random intercept] to compute the health state index as follows: index = 1 – [coefficient of mobility + coefficient of self-care + coefficient of usual activities + coefficient of pain/discomfort + coefficient of anxiety/depression].

The EQ-5D-5L index score will not be calculated when responses are missing for one or more of the 5 dimensions.

The EQ VAS records the participant's self-rated health on a vertical VAS and used as a quantitative measure of health outcome, ranging from 0 (worst imaginable health state) to 100 (best imaginable state). Missing EQ VAS values will be set to missing. If there is a discrepancy between where the participant has placed the X and the number he/she has written in the box, the number in the box should be used.

- The FACT-C consists of 36 items, presented on a five-point Likert scale, in four domains of well-being (27 items from the FACT-G: physical, emotional, social and functional) and the Colorectal Cancer Subscale (CCS). From the FACT-C questionnaires, following scores will be defined, where higher score reflects a better quality of life:
  - Physical well-being (PWB, range 0-28)
  - Social/Family well-being (SWB, range 0-28)
  - Emotional well-being (EWB, range 0-24)
  - Functional well-being (FWB, range 0-28)
  - FACT-G (range 0-108)
  - CCS, where negatively stated items are reversed by subtracting the response from "4". After reversing proper items, all subscale items are combined to produce CCS score (range 0-28).
  - The FACT–C Total score, computed by summing the FACT-G physical and functional domains and the CCS (range 0-136).

  An English sample of FACT-C Version 4 is provided in Section 6.14, as well as scoring instructions.
- The PGIC will ask participants to evaluate their CRC symptoms since starting study intervention according to a seven-point VRS (1. very much improved, 2. much improved, 3. minimally improved, 4. no change, 5. minimally worse, 6. much worse, 7. very much worse).

- The time to definitive deterioration will be derived for the following QoL scales: EORTC QLQ-C30 Global health status / QoL score, FACT-C Total Score, EQ VAS score.

  The time to definitive deterioration is defined as the time from the date of randomization to the date of event (in months), 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 participant 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. No cutoff other than 10% will be tested.

- For all other endpoints (AE, SAE, Physical examination), raw values will be used.

Method for analysis of efficacy secondary endpoints as well as PRO is detailed in Section 5.4.1.2.

Method for analysis of safety endpoints is detailed in Section 5.6.

5.4.1.2. Main Analytical Approach

5.4.1.2.1. PFS based on Investigator Assessments

The primary analysis will be repeated based on investigator assessments using the same analysis method (Kaplan-Meier estimates, stratified Cox model, stratified log-rank test) and the same censoring rules described in Section 5.3.2.2**.** However, as these comparisons are not part of the testing strategy, no formal testing will be performed to compare these two treatment arms. The PFS will be described for the randomized phase only in participants from the FAS.

A table will be provided for comparison of PFS and censoring between investigator assessment and central review (BICR).

In addition and similarly to primary endpoint, for the FAS (by treatment group), a reverse KM analysis will be performed for PFS to estimate the median duration of potential follow-up (with 25th and 75th percentiles) as described by Schemper and Smith (Schemper and Smith 1996). Number of events will be provided. Participants who had a PFS event for the purposes of the PFS analysis will be censored at the date of the event. Participants who were censored (e.g., lost to follow-up, withdrew consent, ongoing, etc.) in the PFS 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 PFS analysis will be used in this analysis. KM estimated probabilities (in percentages) with corresponding 95% CIs (Kalbfleisch and Prentice 2002) will be presented at the same time points as in the PFS analysis. Kaplan-Meier plot will be provided.

Listing of individual PFS based on investigator assessments will be provided with reasons for censored observations.

Investigator assessments will also be listed on FAS for RECIST evaluation, assessment of target lesions, assessment of non-target lesions, assessment of new lesions.

5.4.1.2.2. OS

The KM method will be used to estimate OS time. The median OS with 95% CI and 25th and 75th percentiles will be presented by treatment arm. OS rates at different time points (every 2 months) will be estimated with corresponding 2-sided 95% CIs. The OS of the doublet arm versus the control arm will also be analyzed using the same analysis method described for PFS in Section 5.3.2.2 (Kaplan-Meier estimates, stratified Cox model, stratified log-rank test). However, as this comparison is not part of the testing strategy, no formal testing will be performed to compare OS of these two treatment arms. The OS will be analyzed in a descriptive manner and described in tabular with number of events and graphical format for the randomized phase only in participants from the FAS.

According to the Pharmaceuticals and Medical Devices Agency (PMDA) guidelines Method 1 (PMDA 2007), a positive OS trend could be defined as observing at least half of the treatment effect obtained in BEACON study. Based on De Roock 2010, Ulivi 2012 and Saridaki 2013, a 6 months median OS is expected to be observed in the control arm (value also supported by BEACON study results), and a 9.3 months median OS was observed for Doublet in Beacon study. With limited survival data, OS analysis will be performed 21 months after the first randomized participant. With 94 randomized participants, the probability to observe a positive OS trend, i.e. an HR less than 0.8225, knowing that the true HR is 0.645 (6 months vs 9.3 months) and based on 56 deaths observed, is 80.3%.

Reasons for censored observations (Ongoing without event, Lost to follow-up, Withdrawal of consent) will also be described. Moreover, the OS will be listed with reasons for censored observations, on FAS.

A bar chart of the censoring distributions (events and censor types) over time (]0-1],]1-2],]2-3] months, …) will also be provided.

In addition, for the FAS (by treatment group), a reverse KM analysis will be performed for OS to estimate the median duration of potential follow-up (with 25th and 75th percentiles) as described by Schemper and Smith (Schemper and Smith 1996). Number of events will be provided. Participants who had an OS event for the purposes of the OS analysis will be censored at the date of the event. Participants 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 (in percentages) with corresponding 95% CIs (Kalbfleisch and Prentice 2002) will be presented at the same time points as in the OS analysis. Kaplan-Meier plot will be provided.

A table for Overall Survival Gap Analysis (FAS) will also be provided, showing time between last contact date and the cutoff date, in category (months), for censored participants: first for censored participants who are alive and ongoing OS follow-up, and then for censored participants who are lost to follow-up.

5.4.1.2.3. ORR

The **ORR** will be described for the SLI phase and randomized phase in participants from the FAS, as determined by BICR and investigator assessment per RECIST Version 1.1.
ORR analysis using "confirmed responses" and "confirmed + unconfirmed responses" will be sequentially displayed in the same TFL.

The **BOR** is the best response obtained among all tumor assessment visits after the date of first dose (SLI) or the date of randomization (randomized phase) until documented disease progression, death, start of subsequent antineoplastic therapy, and performed not later than 30 days after last dose intake.
BOR analysis using "confirmed responses" and "confirmed + unconfirmed responses" will be sequentially displayed in the same TFL.

Clinical deterioration or clinical progression noted on the end of treatment visit eCRF will not be considered as documented disease progression.

BOR (confirmed best response and confirmed + unconfirmed best response) will be derived as per computation method specified in this section, using assessment per RECIST Version 1.1 collected at each time point, either by BICR data either by investigator assessment.

For confirmed BOR, confirmation of the response will be performed per RECIST Version 1.1, preferably at the regularly scheduled assessment interval, but no sooner than 4 weeks after the initial documentation of CR or PR. Confirmation of PR or CR can be confirmed at an assessment later than the next assessment after the initial documentation of PR or CR, respectively (see Appendix 11).
Two sets of BOR will be considered, one for confirmed and one for confirmed + unconfirmed responses. BOR will be assessed based on BICR and investigator assessment. BOR is determined from the sequence of overall (lesion) responses according to the following rules:

Complete Response (CR) = at least 2 determinations of CR (consecutive or separated by one single NE or by one single not done assessment) at least 4 weeks apart before disease progression, where confirmation required, or one determination of CR prior to disease progression, where confirmation not required.

Partial Response (PR) = at least 2 determinations of PR or better (consecutive or separated by one single NE or by one single not done assessment) at least 4 weeks apart before disease progression (and not qualifying for a CR), where confirmation required, or one determination of PR prior to disease progression, where confirmation not required.

Stable Disease (SD) (applicable only to participant with measurable disease at baseline) = at least one SD assessment (or better) ≥ 6 weeks after randomization/start of treatment (and not qualifying for CR or PR).

Non-CR/non-PD (applicable only to participants with non-measurable disease at baseline) = at least one non-CR/non-PD assessment (or better) after date of randomization/start of treatment and before progression (and not qualifying for CR or PR).

Progressive Disease (PD) = disease progression ≤ 9 weeks after date of randomization/ first dose of study drug (and not qualifying for CR, PR or SD).

Not Evaluable (NE) = all other cases (i.e., not qualifying for CR or PR and without SD after at least 6 weeks nor early progression within the first 9 weeks). Participants with a first tumor assessment post randomization but prior to treatment start will be NE.

Note: This sequential assessment of responses is not exhaustive. RECIST 1.1 should be followed for some special situations as described in Table 9 thereafter, where "minimum criteria for SD duration met" means at least 6 weeks (ie 42 days) after date of randomization/start of treatment. Besides, other situations as described in Table should be followed.

Table 9: Best overall response when confirmation of CR and PR required

Overall Response First Time point	Overall Response Subsequent Time point	BEST confirmed overall response
CR	CR	CR
CR	PR	SD, PD, or PR[a]
CR	SD	SD provided minimum criteria for SD duration met, otherwise, PD
CR	PD	SD provided minimum criteria for SD duration met, otherwise, PD
CR	NE*	SD provided minimum criteria for SD duration met, otherwise, NE
PR	CR	PR
PR	PR	PR
PR	SD	SD
PR	PD	SD provided minimum criteria for SD duration met, otherwise, PD
PR	NE*	SD provided minimum criteria for SD duration met, otherwise, NE
NE	NE	NE

Abbreviations: CR = complete response; NE=not evaluable; PD=progressive disease; PR=partial response; SD=stable disease.

[a] If a CR is truly met at first time point, then any disease seen at a subsequent time point, even disease meeting PR criteria relative to baseline, makes the disease PD at that point (since disease must have reappeared after CR). Best response would depend on whether minimum duration for SD was met. However, sometimes 'CR' may be claimed when subsequent scans suggest small lesions were likely still present and in fact the patient had PR, not CR at the first time point. Under these circumstances, the original CR should be changed to PR and the best response is PR.
Source: Eisenhauer 2009

*Subsequent documentation of CR may provide confirmation of a previously identified CR for subjects where the second integrated response was NE (ie sequence of CR-NE-CR). Subsequent documentation of PR may provide confirmation of a previously identified PR for subjects where the second integrated response was NE (ie sequence of PR-NE-PR) If the third time point response confirms the CR (resp. PR) then the Confirmed Best Response will be CR (resp. PR). For this study, only one (1) intervening NE is allowed between responses: CR NE CR leads to best confirmed response being CR; PR NE PR leads to best confirmed response being PR.

Table 10: Other situations for confirmed best overall response

Overall Response First Time point	Overall Response Subsequent Time point	BEST confirmed overall response
SD	SD	SD
SD	PD	SD or PD or NE • SD, if SD assessment ≥ 6 weeks, i.e. 42 days after date of randomization/start of treatment, • otherwise PD if PD assessment ≤ 9 weeks, i.e. 63 days after date of randomization/start of treatment, • otherwise NE
SD	NE	SD or NE • SD, if SD assessment ≥ 6 weeks, i.e. 42 days after date of randomization/start of treatment, • otherwise NE
NE, -	SD	SD or NE • SD, if SD assessment ≥ 6 weeks, i.e. 42 days after date of randomization/start of treatment, • otherwise NE
CR, PR, SD	-	SD or NE • SD, if assessment ≥ 6 weeks, i.e. 42 days after date of randomization/start of treatment • otherwise NE.
PD		PD or NE • PD if PD assessment ≤ 9 weeks, i.e. 63 days after date of randomization/start of treatment, • otherwise NE Ignore all assessments after initial overall response of PD.

Participants with BOR 'Not Evaluable' will be summarized by their reason for having not evaluable status. The following reasons will be used (in the following hierarchy):

- No baseline assessment.
- SD occurred < 6 weeks after date of randomization (OR first dose of study drug for SLI participants).
- Progression > 9 weeks after date of randomization (OR first dose of study drug for SLI participants).
- New antineoplastic therapy started before first post-baseline assessment.

- All post-baseline assessments have an overall response of "NE".
- No adequate post-baseline assessment.

The ORR will be provided with a corresponding Clopper-Pearson (exact) binomial 95% CI [Clopper and Pearson 1934].

5.4.1.2.4. DCR

The DCR will be described for the randomized phase only, as determined by BICR and investigator assessment per RECIST Version 1.1, in participants from the FAS.

Analyses will be duplicated for confirmed BOR and for confirmed + unconfirmed BOR.
DCR analysis using "confirmed responses" and "confirmed + unconfirmed responses" will be sequentially displayed in the same TFL.

5.4.1.2.5. DOR

The KM method will be used to estimate DOR time. The median DOR and its 95% CI with 25th and 75th percentiles will be presented by treatment arm. DOR rates at different timepoints (every 2 months) will be estimated with corresponding 2-sided 95% CIs.

The DOR will be described in tabular with number of events for the SLI phase and randomized phase in participants from the FAS, according to BICR and investigator assessment. Graphical representation of DOR (Kaplan-Meier plot) will be provided only for randomized phase.
The proportion of participants with a DOR of 6 months or more, and participants with DOR < 6 months and response ongoing will be also summarized by treatment arm.

5.4.1.2.6. TTR

The KM method will be used to estimate TTR. The median TTR and its 95% CI with 25th and 75th percentiles will be presented by treatment arm. TTR rates at different timepoints (every 2 months) will be estimated with corresponding 2-sided 95% CIs.

The TTR will be described in tabular with number of events and graphical format for the randomized phase only in participants from the FAS, according to BICR and investigator assessment. Plots will be provided for time to confirmed responses only.

Two sets of TTR will be considered, one for confirmed and one considering confirmed and unconfirmed responses.
TTR analysis using "confirmed responses" and "confirmed + unconfirmed responses" will be sequentially displayed in the same TFL.

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

5.4.1.2.7. Patient Reported Outcomes (PRO)

PRO assessments will be collected only for randomized Phase II participants, using the QoL questionnaires EORTC QLQ-C30, EQ-5D-5L, FACT-C at screening and on Day 1 of each cycle, the end of treatment visit and the 30-day safety follow-up visit. The PGIC will also be assessed on Day 1 of each cycle from Cycle 2 onwards, the end of treatment visit and the 30-day safety follow-up visit.

Of note, some QoL assessments were performed using paper questionnaires, which were then retrospectively entered in electronic PRO device. Thus, 'data and time of completion' timestamp was corresponding to the date of retrospective entry rather than actual date of completion. Therefore, an additional 'date of completion' was added in CRF for site entry. It is to be used as assessment date in the statistical analysis instead of the original one, when available.

All PRO analyses will be performed on the FAS and provided for the randomized phase only.

Individual responses to each question will be listed on FAS.

The QoL questionnaires (EORTC QLQ-C30, FACT-C, EQ-5D-5L, PGIC) will be scored according to their respective user guides/scoring manuals. The completion of each questionnaire will be summarized at each timepoint by treatment arm, with:

- One table showing number of participants still on treatment (noted m), along with number and percentage of participants who filled out instrument (n(%)). Participants still on treatment at visit ViDi are patients with at least one study drug not definitely discontinued at the visit CiDi.

  On-treatment means between randomisation date and last treatment date + 30.

  - For EOT, m is the number of patients still on treatment at EOT visit.
  - For 30-day Safety follow-up, m will be empty and only the "Number of patients who filled out instrument" will be displayed without %.

- One table showing number and percentage of participants by type of completion (complete or partially complete) computed as n/m(%):

  - Partial completion will be defined when at least one item of the given questionnaire is {"Not answered", "Missing" or empty} while this item is expected to be completed, when participant filled out instrument.
  - Else, if participant filled out instrument, then the participant will be considered as having fully completed the questionnaire.

Of note, in case two answers have been entered for a same item, conservative approach will be applied:

- If participant was randomized in the doublet arm, the less favourable answer will be retained for analysis.

- If participant was randomized in the control arm, the more favourable answer will be retained for analysis.

Descriptive statistics will be used to summarize the scored scales by treatment arm at each scheduled assessment:

- EORTC QLQ-C30:
    - Each of the five functional scales (physical, role, cognitive, emotional and social)
    - Each of the nine symptoms scale/items (nausea and vomiting, pain, fatigue, dyspnoea, insomnia, appetite loss, constipation, diarrhoea and financial difficulties)
    - The global health status/QoL scale
- EQ-5D-5L as continuous variable: EQ-5D-5L Index Value, EQ VAS.
- FACT-C as continuous scores (raw scores): PWB, FWB, EWB, SWB, FACT-G, CCS, FACT-C Total score.
- PGIC as categorical variable: Number and percentages of participants in each category will be provided.

For EORTC QLQ-C30 and FACT-C, derived scores will be listed on FAS.

Additionally, change from baseline in the domain scores (except the PGIC) at the time of each assessment will be summarized. Participants with an evaluable baseline score and at least one evaluable post-baseline score during the treatment period will be included in the change from baseline analyses. For those same assessments, change from baseline will also be plotted over time, displaying mean and standard deviations.

No analysis will be performed on the scores of the five dimensions of the EQ-5D-5L.

In addition, a repeated measurement analysis model will be used to compare the two treatment arms with respect to changes in the domain scores longitudinally over time, for the three following scores: EORTC QLQ-C30 Global health status / QoL score, FACT-C Total score, EQ VAS score. Model will be adjusted by using baseline score and the two stratification factors (IRT values) as covariates. Each model will have an intercept term, a linear time trend term (in weeks), a term for treatment group, and a term for treatment-by-time interaction. The intercept and slope terms for time will be random effects with an unstructured variance/covariance matrix. In addition, each observation is assumed to be measured with error and the error terms are independent of each other. All parameter estimates will be obtained using restricted maximum likelihood estimation.

The significance of the treatment-by-time interaction will be evaluated first. The overall significance of the difference between the trajectories for the treatment arms will be tested at a significance level of 0.05 (two-sided) within the model without the interaction. Adjusted means will be provided for each treatment group as well as their 95% confidence interval.

For the above analyses of QOL questionnaires (summary by visit, change from baseline by visit, repeated measures), post-baseline values will be analysed only if participant is still on treatment at the given visit.

Time to definitive deterioration.

Time to definitive deterioration of EORTC QLQ-C30 Global health status and Time to definitive deterioration of FACT-C Total Score will be assessed in the treatment arms. Number of events will be provided. The distribution will be presented descriptively using KM curves. Median time to definitive deterioration along with 2-sided 95% CI will be provided as well as 25^{th} and 75^{th} percentiles.

A stratified Cox model will be fit with treatment arm as covariate to obtain a hazard ratio 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 (IRT) information. No p-value will be provided.

5.4.1.3. Sensitivity Analysis

For the randomized phase, the following sensitivity analyses will be performed to further assess the comparisons described by the secondary objective on the PFS based on investigator assessment:

- The PFS analysis will be repeated in the ES.
- The PFS analysis will be repeated in the PPS.

In addition, all secondary endpoints of the randomized phase will also be analyzed in other populations:

- OS: ES and PPS
- ORR: ES
- DCR: ES
- DOR: ES
- TTR: ES.

5.4.1.4. Supplementary Analyses

Not Applicable.

5.4.2. Supportive Secondary Endpoint(s)

Not Applicable since all secondary endpoints are considered in section 5.4.1.

5.5. Tertiary/Exploratory Endpoint(s) Analysis

Exploratory endpoints are only defined for the Randomized phase:

CCI

- Markers from central laboratory
    - MSI Status in FFPE samples using established PCR assays
    - Detection of *BRAF* V600E mutation in tumor tissues by a candidate companion diagnostic based on NGS or PCR methods

CCI

5.5.2. Central Laboratory: MSI Status in FFPE samples using established PCR assays

The MSI is a common test done for colorectal cancer participants in order to assess the disease subtype (MSI-High (MSI-H) vs Microsatellite stable (MSS)). Descriptive statistics will be provided in participants from the FAS.

A data listing will be also presented.

5.5.3. Central Laboratory: Detection of *BRAF* V600E mutation by a candidate companion diagnostic

The FFPE tumor tissue source collected for central determination of BRAF V600E status at molecular prescreening or screening will also be used to retrospectively perform a comparison analysis between the BRAF V600E mutation as determined centrally in this clinical study with a companion diagnostic candidate.

The analyses listed below will be provided upon data availability from vendor, and at the latest at time of final analysis:

- Presence of *BRAF* V600E mutation as detected by a candidate companion diagnostic will be summarized for both treatment groups combined, in participants from the FAS. Related listing will be provided.

- To investigate the concordance with the BRAF V600E mutation result used for participant selection, a shift table will be provided for both treatment groups combined in participants from the FAS.

- Demographics characteristics (Age, Sex and Race) will be summarized on the following participants sets based on FAS, for both treatment groups combined:

    - Participants with positive result on central test [CTA+]

    - Participants with positive result on central test and an evaluable result on CDx [CTA+/CDx evaluable]

    - Participants with positive result on central test result and CDx [CTA+/CDx+]

    - Participants with positive result on central test result but tested negative with CDx [CTA+/CDx-]

    - Participants with positive result on central test but CDx testing unevaluable (including CDx testing failure or not available for CDx testing) [CTA+/CDx unevaluable]

    - Participants with negative result on central test [CTA-]

  - Participants with negative result on central test result and tested positive with CDx [CTA-/CDx+]

- Another table will be provided showing demographics in both SLI and randomized phase for participants according to their CTA/CDx status

- Demographics characteristics (Age, Sex and Race) will also be summarized by treatment group for FAS participants of the randomized phase with positive result on central test result and CDx [CTA+/CDx+]

- Primary and secondary criteria (PFS, cORR, DCR, DOR and OS) will be summarized by treatment group for CTA+/CDx+ participants, and CTA+ participants from FAS.

5.6. (Other) Safety Analyses

5.6.1. Study Intervention Exposure

Exposure to the study intervention is defined as the time interval in days between the actual date of first study intervention administration (included) and the actual date of last study intervention administration (included), derived as:

- For encorafenib (daily administration):

[Date of last dose of encorafenib or date of cutoff*] – [date of first (non-zero) dose of encorafenib] + 1.

* If participant is off treatment, this is the last non-zero dose of the study drug. If the participant is on treatment at the cutoff date, this is the cutoff date.

- For cetuximab (weekly injection):

If participant is off treatment: [Date of last dose of cetuximab] – [date of first (non-zero) dose of cetuximab] + 7**.

If participant is on treatment: [Date of cutoff] – [date of first (non-zero) dose of cetuximab] + 1.

- For IV drugs of the control arm (irinotecan, 5-FU and FA) (injections every 2 weeks):

If participant is off treatment: [Date of last dose of IV drug] – [date of first (non-zero) dose of IV drug] + 14**.

If participant is on treatment: [Date of cutoff] – [date of first (non-zero) dose of IV drug] + 1.

**7 and 14 days correspond to the number of days between two doses for intermittent dosing of IV drugs.

Of note, for any participant continuing the study treatment beyond progression (as allowed by protocol), date of last dose will be the date of treatment discontinuation reported in the CRF End of Treatment visit corresponding to the treatment discontinuation due to progressive disease observed under study treatment and leading to EoT visit.

The duration will be also computed in weeks (dividing by 7 the above computations), and summarized by study drug within each treatment arm, and by treatment arm*, for both phases (SLI and randomized) in participants from the SS for following variables:

- Duration as continuous (in weeks) using descriptive statistics,
- Categorized by time intervals (< 4 weeks, 4-<8 weeks, …) for which frequency counts and percentages of participants will be provided.

*Duration of exposure by treatment arm will be computed as the maximum of all durations of exposure for each drug (as defined above) of the assigned regimen.

Kaplan-Meier plot of duration of exposure (in weeks) will be also presented by treatment arm (as defined there above) for the randomized phase (doublet versus control on the same plot) for the Safety Set. Participants who have not discontinued study drugs prior to the data cutoff date [ie who did not complete End of Treatment prior to cut-off date] will be censored at the cut-off date.

Corresponding table of Kaplan-Meier estimates will be provided.

Kaplan-Meier analysis of duration of exposure described above (figure and table) will be repeated by number of prior regimens for metastatic disease.

Total exposure time, expressed in patient-months, will be provided for encorafenib in participants from the SS (SLI and randomized phases pooled). Total exposure time will be a unique number, calculated as the sum of the duration of encorafenib exposure (in days) for each participant divided by 30.4375. The total exposure time (aside the number of participants) will be provided according to:

- The duration of encorafenib exposure (i.e. for participants with at least 1, 3, 6, 12, 18, 24 or 30 months of encorafenib exposure)
- The age group as per Risk Management Plan (18 - 64 years, 65 - 74 years, 75 - 84 years, 85+ years and, Total) cross tabulated with gender (Female vs Male)

The actual Dose Intensity (DI) and Relative Dose Intensity (RDI) will be defined for each study intervention separately (encorafenib, cetuximab, irinotecan, 5-FU, FA), as below.

The actual Dose intensity (DI) will be defined as follows.

- For encorafenib (daily administration):

DI (mg/day) = Cumulative actual dose*/duration of exposure in days as computed above

* If the participant is on treatment at the cutoff date, last dose is extrapolated until the cutoff date.

- For IV drugs (weekly injection for cetuximab, biweekly injections for irinotecan, 5-FU and FA):

DI (mg/m²/dose) = Cumulative actual dose (in mg/m²)/number of planned doses

with Cumulative actual dose being the sum of the { actual dose (mg) / BSA at the given visit (or, if missing, the last non-missing BSA prior to the visit) }

and Number of planned doses = number of non-zero doses (regardless if interrupted during infusion or not) plus the number of missed doses. Dose reported as 0 mg will be considered as a missed dose.

For each drug, the RDI is defined as:

$$RDI\ (\%) = (DI/PDI)*100$$

where:

- DI is the dose intensity as defined above,
- PDI is the Planned Dose Intensity during the study for the given study drug.
    - For encorafenib, PDI = 300 mg/day.
    - For IV drugs, PDI (in mg/m²/dose) = (Cumulative planned dose)/(number of planned doses, as defined above for DI), where cumulative planned dose is the sum of all protocol-specified doses* (in mg/m²) across each planned day of dosing.

*Protocol-specified doses for IV drugs are: For cetuximab: 400 mg/m² once then 250 mg/m²; for irinotecan: 180 mg/m²; for 5-FU: 400 mg/m2 of Bolus IV and 2400 mg/m2 of Infusion IV; for Folinic acid: 400 mg/m2. For 5-FU (bolus+infusion IV) and FA, it is superseded by the actual dose level taken at C1D1 (in mg/m²), provided this *actual dose level taken at C1D1* differs from the protocol dose level (ie 400 for FA and 2800 for FU (400 bolus + 2400 infusion IV) and only if this *actual dose level taken at C1D1* is equal to the "maximal dose tolerated in a prior regimen (mg/m²)" as recorded in the CRF.

For 5-FU, bolus and injections are to be analysed on a combined approach (ie summing up the doses).

Actual DI, PDI and RDI will be calculated with no more than one decimal.

Actual DI and RDI (%) will be summarized by study drug within each treatment arm for both phases (SLI and randomized) in participants from the SS. Descriptive statistics will be provided.

For RDI, number and percentage of participants will be provided for following categories: <50%, 50 - <80%, 80 - <100%, 100%, >100%.

For SLI phase only, the relative dose intensity in Cycle 1 will be defined for encorafenib and cetuximab, as 100*administered dose (mg) through Cycle 1 divided by planned dose (mg), with planned dose being:

- 8400 mg for encorafenib (28 daily doses of 300 mg)
- the sum of BSA (m²) * planned dose (in mg/m²), across all injections of cetuximab. BSA is taken at the given injection visit (or, if missing, the last non-missing BSA prior to the visit). Planned dose is the protocol-specified dose at the given visit, as defined above (ie 400 then 250 mg/m²).

Relative dose intensity in Cycle 1 will be listed in the listing of dose intensities, for SLI participants only.

Dose modifications will be defined as below:

- 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, for encorafenib. For IV drugs, it is defined as an injection reported as not performed.

To avoid over-counting interruptions, dose interruptions (as defined above) 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 permanent discontinuation analysis.

- A dose reduction is defined as a decrease in dose level from the protocol-planned dose (protocol-planned doses for IV drugs are defined above in the PDI definition) 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 participant 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.

For 5-FU, one dose reduction is when either bolus or infusion is reduced at one given visit.

Of note, for any participant continuing the study treatment beyond progression (as allowed by protocol), administrations periods (including interruptions and reductions) reported after the date of 'last dose' as defined on top of this section, will be considered as subsequent anticancer therapy and therefore will not be used for computation of exposure parameters and consequently not included in Exposure summary data.

Frequency and percentages of participants who have dose reductions or interruptions, and corresponding reasons, will be summarized by study drug within each treatment arm for participants from the SS. Dose reductions and interruptions will be tabulated both separately and in a combined fashion. For IV drugs (Cetuximab and control drugs), number of infusion interruptions will be presented.

For participants with at least one dose reduction as defined above, the percentage of days exposed to any 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 a reduced dose. Denominator will be the exposure duration as defined in top of the present section. Numerator will sum the durations of intervals (from SDTM.EC) corresponding to reduced dose(s). If last administration is with a reduced dosage, interval duration will be extrapolated until last day of exposure duration (exposure duration as defined in top of the present section).
For participants with dose interruption as defined above, the percentages of days between the first and last null dose of dose interruption will also be summarized. For IV drugs, one missed injection of cetuximab is assumed to create an interval of 7 days, and one missed injection of the other drugs (irinotecan, 5-FU or folinic acid) is assumed to create an interval of 14 days.

Treatment interruptions due to Covid-19 (SARS-CoV-2) will be described as part of the data collected by the "Covid-19 impact" CRF form.

Information about administration and exposure of each study drug (encorafenib, cetuximab, 5-FU, FA, irinotecan) will be presented in data listings, with duration of exposure presented in days.

A listing of drug accountability will also be provided.

5.6.2. Adverse Events

The occurrence of an AE is defined by the appearance of a new single event, or the reappearance of a previously recovered event. A continuous event with change in severity or seriousness (relative to its previous status) will be considered as one occurrence.

A continuous AE started prior first dose and reported at a higher grade than Baseline should be counted as occurrence. Same logic should be followed for lower (resp. higher) seriousness.
Note:

- ✓ A continuous AE started prior first dose and reported at a lower grade than Baseline should be not counted as occurrence of TEAE it will be an AE occurrence.
- ✓ If a subject experienced the same adverse event at more than 1 severity, or seriousness, the most severe rating or the stronger seriousness was given precedence.
- ✓ The definition of occurrence should be considered independently of timing of start date of prior first administration, the maximum grade should be recorded irrespectively of the administration. Example, event was grade 2 before first administration and dropped to grade 1 post first administration and resolved, we would flag the grade 2. And it will be counted as AE occurrence but not a TEAE occurrence.

AEs will be coded using MedDRA (Version 26.0 at set-up, then latest version in use).

A treatment emergent adverse event (TEAE) is 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

*Same rule applies for any participant continuing the study treatment beyond progression (as allowed by protocol), with ‘treatment discontinuation’ being the last dose as defined in Section 5.1.4.

As conservative approach, if year of AE start is missing, and when end date is missing or >= first administration date, AE will be considered as a TEAE.

A participant with more than one occurrence of the same adverse event in a particular System Organ Class (SOC) will be counted only once in the total of those experiencing adverse events in that particular SOC.

If a participant experiences the same adverse event at more than one severity, or with more than one relationship to study intervention, the most severe rating or the stronger causal relationship to study intervention will be given precedence.

The denominator used to calculate incidence percentages consists of participants in the SS.

Exposure adjusted incidence rate (EAIR) analyses will also be presented for some analyses:

EAIR per 100 patient-months = (n*100)/(total exposure time).

- n is the number of participants reporting the event
- total exposure time is the sum of exposure time in months (of the given treatment arm, i.e. doublet or control) for all participants: Exposure time for a participant without the specific event is the duration of exposure by treatment arm (as defined in Section 5.6.1), whereas the exposure time for a participant with the specific event is the treatment duration up to the start date (inclusive) of the first occurrence of the specific event, computed as: date of AE onset – first dose of study intervention as defined in Section 5.1.4 + 1.

AEs will be grouped by SOC and Preferred Term (PT) and sorted in descending frequency (or descending EAIR when appropriate) in the Doublet group of the Randomized phase.

For tables by NCI-CTCAE grade, if a participant has more than one AE the worst grade will be summarized.

Following summaries with number and percentage of participants and number of events will be presented:

- Summary of AE (Overall, Related, Grade 3+, Related Grade 3+):
    - AEs
    - TEAEs,
    - Serious TEAE,
    - TEAE by severity grade,
    - TEAE leading to discontinuation of all study drug regardless of causality
    - TEAE leading to discontinuation of any study drug regardless of causality
    - TEAEs leading to study intervention dose modification (interruption or reduction) (any drug),
    - TEAEs leading to Study Intervention interruption (any drug),
    - TEAEs leading to Study Intervention dose reduction (any drug),
    - TEAEs requiring additional therapy,
    - TEAE worst outcome,
        - For each participant, the worst outcome will be considered over all TEAEs, ranging: 'Fatal', 'Not recovered/Not resolved', missing/'Unknown (Lost to Follow-up)', 'Recovering/Resolving', 'Recovered/Resolved with sequelae', 'Recovered/Resolved'.
    - TEAE worst action taken with study intervention (any drug).
        - For each participant, the worst action taken with study intervention (any drug) will be considered over all TEAEs, ranging: ‘DRUG WITHDRAWN (Permanently discontinued)’, ‘DOSE REDUCED’, ‘DRUG INTERRUPTED (Temporarily interrupted)’, ‘UNKNOWN’, ‘DOSE NOT CHANGED (No action taken - study drug ongoing)’, ‘DOSE INCREASED’, ‘NOT APPLICABLE’
- Summary of SAEs (Overall, Related, Grade 3+, Related Grade 3+):
    - SAEs,
    - Serious TEAEs,
    - Serious TEAE by severity grade,
    - Serious TEAE leading to discontinuation of all study drug regardless of causality
    - Serious TEAE leading to discontinuation of any study drug regardless of causality
    - Serious TEAEs leading to study intervention dose modification (interruption or reduction) (any drug),
    - Serious TEAEs leading to Study Intervention interruption (any drug),

- Serious TEAEs leading to Study Intervention dose reduction (any drug),
- Serious TEAEs requiring additional therapy,
- Serious TEAE worst outcome,
  - For each participant, the worst outcome will be considered over all serious TEAEs, ranging: 'Fatal', 'Not recovered/Not resolved', missing/'Unknown (Lost to Follow-up)', 'Recovering/Resolving', 'Recovered/Resolved with sequelae', 'Recovered/Resolved'.
- Serious TEAE worst action taken with study intervention (any drug).
  - For each participant, the worst action taken with study intervention (any drug) will be considered over all serious TEAEs, ranging: 'DRUG WITHDRAWN (Permanently discontinued)', 'DOSE REDUCED', 'DRUG INTERRUPTED (Temporarily interrupted)', 'UNKNOWN', 'DOSE NOT CHANGED (No action taken - study drug ongoing)', 'DOSE INCREASED', 'NOT APPLICABLE'
- Seriousness criteria for Serious TEAE

In addition, AE will be described on the Safety Set for all grades and for grade 3+ by age subgroup as per Risk Management Plan (18 - 64 years, 65 - 74 years, 75 - 84 years, 85+ years) for Doublet group in the following terms:

- At least one TEAEs (overall, related)
- At least one Serious TEAEs (overall, related),
- At least one Fatal TEAE (i.e. Grade 5) (overall, related),
- At least one Serious TEAE requiring Hospitalization (overall, related),
- At least one TEAE leading to discontinuation (any drug withdrawn) (overall, related)
- At least one TEAE leading to study intervention dose modification (interruption or reduction) (any drug) (overall, related)
- At least one TEAE leading to Study Intervention dose reduction (any drug) (overall, related)
- SOC and Standardised MedDRA Queries (SMQ) of interest:
  - SOC Psychiatric disorders,
  - SOC Nervous system disorders,
  - Accidents and injuries SMQ,
  - SOC Cardiac disorders,
  - SOC Vascular disorders,
  - Central nervous system vascular disorders (SMQ),
  - SOC Infections and infestations

- PTs of interest (all grades), overall and for:
  - Dizziness (includes following PTs: Dizziness, Dizziness postural),
  - Fracture (includes following PTs: Femur fracture, Hip fracture, Humerus fracture, Lumbar vertebral fracture, Pathological fracture, Spinal fracture),
  - Loss of consciousness,
  - Orthostatic Hypotension,
  - Syncope,
  - Fall
  - Ataxia.
- PT Anticholinergic syndrome,
- PT Quality of life decreased,
- Other AEs appearing more frequently in older participants, i.e. at least 3 participants with this specific PT in the elderly population (>=65 years) and a difference >=10 point on percent between the elderly and the other population (<65 years).

The following description of TEAE will also be provided (incidence of) with number and percentage of participants, the number and percentage of participants with Grade >=3 (i.e. Grade 3+):

- TEAE by SOC and PT, also providing number of events
- Serious TEAEs by SOC and PT, also providing number of events
- Related Treatment Emergent Adverse Events by SOC and PT, also providing number of events
- Related Serious Treatment Emergent Adverse Events by SOC and PT, also providing number of events
- TEAE Leading to Study Intervention Discontinuation of all study drugs by SOC and PT
- TEAE Related to Covid-19 (SARS-CoV-2) by SOC and PT
- Most frequent (≥10% of participants in any treatment group*) TEAEs by PT in Order of Frequency

  *>=10% of participants in any of the treatment groups shown in the table (for the given category of AEs, considering “all grades”) – within Safety population.
- Most frequent (≥10% of participants in any treatment group*) Serious TEAEs by PT in Order of Frequency

  *>=10% of participants in any of the treatment groups shown in the table (for the given category of AEs, considering “all grades”) – within Safety population.
- TEAE Leading to Study Intervention Discontinuation of any study drug by PT in Order of Frequency
- Grade 5 TEAEs by PT in Order of Frequency

- TEAE Leading to Dose Modification (Action Taken = "Dose reduced" or "Drug interrupted") by PT in Order of Frequency (to be presented for each molecule)
- TEAE Leading to Dose Reduction of any study drug by SOC and PT
- TEAE Leading to Dose Reduction by PT in Order of Frequency (to be presented for each molecule)
- TEAE Leading to Drug Interruption of any study drug by SOC and PT
- TEAE Leading to Drug Interruption by PT in Order of Frequency (to be presented for each molecule)
- TEAE by SOC and PT and Worst NCI-CTCAE Grade
- Serious TEAE by SOC and PT and Worst NCI-CTCAE Grade
- Related Treatment Emergent Adverse Events by SOC and PT and Worst NCI-CTCAE Grade
- Related Serious Treatment Emergent Adverse Events by SOC and PT and Worst NCI-CTCAE Grade
- TEAE Requiring Additional Therapy by SOC and PT and Worst NCI-CTCAE Grade
- EAIR analysis: AEs, regardless of study drug relationship, adjusted for patient-month of exposure, by PT and treatment arm (i.e., number of participants with at least one occurrence of an AE per 100 patient-months of exposure). Only Treatment Emergent Adverse Events will be analysed.
- Adverse events, regardless of study drug relationship, occurring within the first 3 months of study treatment by PT and treatment (overall and maximum Grade 3 or higher). This will be repeated for 6 and 12 months.

In case of a participant experiencing several episodes of a same continuous AE (ie, a first episode not recovered followed by another episode with a higher severity grade, or with a recovery date), in the table by SOC and PT, the event will be considered only once.
For the table by SOC, PT and Worst NCI-CTCAE Grade, the worst episode (the one with the highest severity grade) will be considered.
See Example below for the same participant and the same PT:

PT	AE Number	Episode Number	Start Date	Stop Date	Description
Headache	1	1.01	20Jan2022		Grade 1
	1	1.02	01Feb2022	10Feb2022	Grade 3
Headache	2	2.01	15Mar2022	17Mar2022	Grade 2
Headache	3	3.01	11Apr2022		Grade 2
	3	3.02	16Apr2022	18Apr2022	Grade 3

With the example above, for the tables by SOC and PT, for this participant, we will consider that he experiences 3 Events of Headache (due to AE number), without considering the details of episodes.

For the table by SOC, PT and Worst NCI-CTCAE grade, the participant will be counted only for worst grade equals to 3 with 2 Events due to episodes 1.02 and 3.02. Headache whose AE number equals to 2 will not be considered to this analysis by Worst NCI-CTCAE grade as this is a grade 2 event.

Number and percentage of participants with at least one event of any AE of special interest (AESI) of encorafenib, and of each category of AESI of encorafenib will be reported in participants from the SS. Corresponding summary table will be provided, for all grades as well as for Grade 3+. Such categories consist of grouping one or more well-defined safety events which are similar in nature and for which there is a specific clinical interest in connection with the study intervention. The following AE of special interest of encorafenib will be analysed:

- Acute renal failure, cutaneous non-squamous cell carcinoma, cutaneous squamous cell carcinoma, melanomas, palmar-plantar erythrodysaesthesia syndrome, photosensitivity, rash, severe cutaneous adverse reactions, facial paresis, tachycardia, haemorrhage, hepatic failure, liver function test abnormalities, myopathy and uveitis-type events. Definition of those AESIs will be based on encorafenib case retrieval strategy as provided in Section 6.15.
- Additional AESI or updated case retrieval strategy could be used as deemed relevant during the course of the study.

In addition, EAIR analysis of AESI will be provided: AESIs, regardless of causality, adjusted for patient-month of exposure (i.e. number of participants with at least one occurrence of an AESI per 100 patient-months exposure) will be analysed, by AESI grouping and treatment arm.

For all analyses of AESIs, only Treatment Emergent Adverse Events will be analysed.

Vaccine-emergent AEs will also be analysed, distinguishing which COVID-19 (SARS-CoV-2) vaccine dose it corresponds. Vaccine-emergent AE for vaccine dose i will be defined as any event that first occurred or worsens from the COVID-19 (SARS-CoV-2) vaccine dose i date up to 30 days (inclusive) thereafter, and regardless if another vaccine dose was administrated during those 30 days, for participants having received the i^{th} dose of a COVID-19 (SARS-CoV-2) vaccine.

Analysis will be a summary table showing Vaccine-emergent AEs by SOC and PT (using number of participants), overall then repeated for each COVID-19 (SARS-CoV-2) vaccine type. Those tables will be provided for Vaccinedose1-emergent AEs (resp. Vaccinedose2-emergent AEs and Vaccinedose3-emergent AEs), for participants having received the 1^{st} (resp. 2^{nd} and 3^{rd}) dose of a COVID-19 (SARS-CoV-2) vaccine. Vaccines and vaccine types will be identified as stated in Section 6.7.2.2.

Vaccine-emergent AEs will also be listed in a separate listing, identifying if Vaccinedose1-emergent and/or Vaccinedose2-emergent and/or Vaccinedose3-emergent.

In addition, vaccine-emergent AE resulting in death (ie grade 5) will be listed as part of the vaccine-emergent AEs.

Analyses of time to onset of first TEAE will be provided for all participants (only for randomized phase) using KM methods, reporting estimated median (in months) with 95% CI, 25th and 75th percentiles. Number of participants with events and number of censored participants will also be described. If a participant has no event, he will be censored using date of last administration + 30 days (if discontinued) or cut-off date (if ongoing). A graph of KM curves will also be displayed. Same analysis will be repeated in participants with the following events (so no censored participants):

- First Serious TEAE
- First TEAE with a Grade Greater Than or Equal to 3
- First TEAE Resulting in Discontinuation of Any Study Drug
- First TEAE Resulting in Discontinuation of All Study Drug
- First AESI

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.

Summaries for deaths on-study treatment and off-treatment will be provided showing number and percentage of participants for each, regardless cause of death (AE, progressive disease…). On-treatment deaths are defined when date of death is occurring within 30 days after last dose of study treatment, off-treatment when date of death is occurring >30 days after last dose of study treatment.

In addition, AEs resulting in on-treatment death (i.e. grade 5) and occurring within 30 days of first dose of study intervention will be summarized for each arm by SOC and PT.

An EAIR summary of AEs resulting in on-treatment death (ie grade 5) by PT and treatment arm adjusted for patient-month exposure (i.e., number of participants with at least one occurrence of a PT resulting in death per 100 patient-months exposure) will also be presented, regardless treatment emergent or not.

The following listings will be provided with participant's code, sex and age, verbatim term, PT, action taken with study intervention, use of a corrective treatment or procedure, outcome and relationship to the study intervention in the investigator's opinion:

- AE
- Grade 5 Adverse Events
- AE related to Covid-19 (SARS-CoV-2)
- AEs defined as vaccine-emergent (identifying in the listing if AE is Vaccinedose1-emergent and/or Vaccinedose2-emergent and/or Vaccinedose3-emergent)

A listing will present all reported SAEs.

AE listing will be repeated in prescreen failure participants, screen failure participants and other participants not in the SS.

Notes:

- Missing severity, relationship or outcome will be classed as unknown.
- Missing seriousness will be imputed as serious=Yes, per conservative approach
- "Related" AE (unless otherwise specified) means related to any drug.
- "Leading to discontinuation/ interruption/ dose reduced" refers to any drug, unless otherwise specified
- "AE requiring additional therapy" corresponds to the CRF item "Corrective treatment and/ or corrective procedure".
- Grade 3+ is defined as maximum grade 3 or higher and is marked as "G3+" in the TFL shells
- AE related to Covid-19 (SARS-CoV-2) will be identified when AE verbatim includes "COVID-19" as prefix or when AE verbatim includes "CORONAVIRUS".

5.6.3. Clinical Safety Laboratory Evaluation

Viral serology and tests to confirm postmenopausal status for females will be performed before the start of study treatment to determine eligibility or baseline status only.
Blood samples for hematology, clinical chemistry, coagulation and urine samples for dipstick urinalysis will be taken at screening and repeated on Cycle 1 Day 1 (if not performed within 72 hours before Cycle 1 Day 1 (i.e. first day of study intervention), on Day 1 of each subsequent cycle, the end of treatment visit and the 30-day safety follow-up visit. Additional blood samples will be taken as follows:

- For hematology:
  - Cycle 1 Days 15 and 22 for all participants.
  - Day 15 of each cycle from Cycle 2 onwards for participants in the control arm of the randomized phase continuing to receive irinotecan only.
- For clinical chemistry: on Cycle 1 Day 15 for all participants.

All assessments during the treatment phase should be made predose. All clinical chemistry assessments should be made after the participant has fasted for at least 8 hours.

The parameters to be assessed at each timepoint are listed in Table 11.

Unscheduled clinical laboratory tests may be obtained at any time during the study at the investigator's discretion. Laboratory test results required to make decisions regarding potential dose modifications (as specified in Protocol Section 6.6) should be reviewed before study intervention administration.

Table 11: Protocol-required Clinical Laboratory Safety Assessments

Hematology	Erythrocytes (RBC), hematocrit, hemoglobin, platelets Leukocytes (WBC) count with differential: basophils, eosinophils, lymphocytes, monocytes, neutrophils/ANC
Chemistry[a]	Albumin, Alkaline phosphatase (ALP), ALT, AST, lipase, amylase, bilirubin (total and indirect), blood urea nitrogen/urea, calcium, chloride, creatine kinase, creatinine glucose, Lactate Dehydrogenase (LDH), magnesium, potassium, sodium, total protein, troponin I or T, uric acid (uric acid will be referred as urate in statistical analysis)
Coagulation	activated Partial Thromboplastin Time (aPTT), International Normalized Ratio (INR) or Prothrombin Time
Urinalysis	Specific gravity Blood, glucose, leukocytes, ketones, pH, protein by dipstick Microscopic examination (if blood or protein abnormal)
Viral Serology (screening only)	HBsAg (hepatitis B surface antigen), HBcAb, HBV DNA when HBsAg and/or HBcAb positive and hepatitis C antibody, HCV RNA when HCV Ab positive, and HIV. If required, analysis of follicle stimulating hormone (FSH), luteinizing hormone and/or estradiol to confirm postmenopausal status in females

Abbreviations: aPPT = activated partial thromboplastin time; ALT = alanine aminotransferase; ANC = absolute neutrophil count; AST = aspartate aminotransferase; INR = international normalized ratio; LDH = lactate dehydrogenase; pH = hydrogen ion concentration; RBC = red blood cell(s); SAE = serious adverse event; ULN = upper limit of normal; WBC = white blood cell.

[a] Details of liver clinical chemistry stopping criteria and required actions and follow-up assessments after liver stopping or monitoring event are given in Table 28 from protocol section 10.6.1. All events of ALT >3 × Upper Limit Of Normal (ULN) and bilirubin >2 × ULN (>35% direct bilirubin) or ALT >3 × ULN and INR >1.5, if INR measured which may indicate severe liver injury (possible Hy's Law), must be reported as an SAE.

Blood samples will be collected, handled and analyzed at the local laboratory according to the study site's standard procedures and the latest updated references. If needed, lipase and amylase will be collected locally and sent to central laboratory for analysis. Ranges from the laboratory will be used to identify abnormal values. Urinalysis will be performed in the clinic or at the local laboratory, according to study site standards.

Clinically significant abnormal laboratory findings are those which are not associated with the underlying disease, unless judged by the investigator to be more severe than expected for the participant's condition.

Listings of all laboratory evaluations (hematology, chemistry, urinalysis including microscopic evaluation, coagulation, serology) will be provided.

A listing of pregnancy test results will also be provided. Reproductive status will also be listed.

5.6.3.1. Hematology and Clinical Chemistry

Descriptive statistics of all parameters over time will be performed. Values at each visit and changes from baseline will be calculated and tabulated. Unscheduled visits or retest will not be analyzed for the summary by visit.

Shift tables of maximum (worst on-treatment as per section 5.1.4) post baseline NCI-CTCAE grades according to baseline NCI-CTCAE grade (version 4.03) will be provided for the following parameters:

- Hematology
    - Hemoglobin (Low)
    - Hemoglobin (High)
    - Leukocytes (Low)
    - Leukocytes (High)
    - Neutrophils (Low)
    - Platelets (Low)
    - Lymphocytes (Low)
    - Lymphocyte (High)
- Biochemistry
    - Alanine Aminotransferase (High)
    - Albumin (Low)
    - Alkaline Phosphatase (High)
    - Amylase (High)
    - Aspartate Aminotransferase (High)
    - Bilirubin (High)
    - Calcium Corrected (Low)
    - Calcium Corrected (High)
    - Creatinine (High)
    - Creatine Kinase (High)
    - Glucose (Low)
    - Glucose (High)
    - Lipase (High)
    - Magnesium (Low)
    - Magnesium (High)
    - Potassium (Low)

- Potassium (High)
- Sodium (Low)
- Sodium (High)
- Troponin I (high)
- Troponin T (high)
- Urate (High)

Laboratory data will be graded using NCI-CTCAE Version 4.03. Grade 0 will be assigned for all non-missing values not graded as ≥1. Grade 5 will not be used.

For some parameters, grades are defined for low and high values.

In addition, incidence of clinically notable shifts in laboratory parameters based on NCI-CTCAE grade for the same list of parameters (as above) will be summarized, i.e. number and percentage of participants with at least one clinically notable shift during study. Clinically notable shift is defined as a worsening from baseline by at least 2 grades, or to grade 3 or above.

Moreover, for parameters that cannot be graded, shift tables of worst post baseline normal/abnormal values according to baseline values will also be displayed (worst on-treatment as per section 5.1.4).

For chemistry parameters, all parameters will be considered for analysis without differentiation on fasting (baseline flag, summary by visit, worst grade by participant, shifts, Hy's law). Fasting glucose parameter will be created, this parameter will include records with fasting sampling and will be used for grading and also for values overtime. Single glucose parameter will be analyzed over time but it will not be graded.

Hepathic toxic

Possible Hy's Law cases will be summarized by treatment group as part of the analysis below (last category of Table 12).

Hepatic toxicity will be assessed based on the following Liver Function Tests (LFTs): Albumin, ALT, AST, ALP and Total Bilirubin (TBL). For these parameters, NCI-CTCAE grades are defined.

LFTs will be summarized by treatment arm and liver metastasis at baseline (overall and yes vs. no). Presence of liver metastases at baseline (yes versus no) will be defined based on Target and Non-target lesion locations according to BICR. Frequency counts and percentages of participants having a newly occurring hepatic value in the categories presented in Table 12 will be provided, for each category:

- Numerator will be the number of participants not meeting the criterion at baseline and meeting criterion post-baseline ("newly occurring"). Of note, If a participant has a post-baseline ALT value ">10xULN", he is also considered as meeting the ">3xULN"

">5xULN" and ">8xULN" criteria. All post-baseline values are considered regardless scheduled or unscheduled.

- Denominator used to calculate the percentage will be Number of participants at risk, i.e. with non-missing values at baseline and with a baseline value that does not already meet the abnormality.

Table 12: Hepatic Toxicity Criteria

Parameter	Criterion
Albumin	< LLN
ALT	>3×ULN; >5×ULN; >8×ULN;>10×ULN;>20×ULN
AST	>3×ULN; >5×ULN; >8×ULN >10×ULN;>20×ULN
AT (ALT or AST)	>3×ULN; >5×ULN; >8×ULN >10×ULN;>20×ULN
TBL	>1.5×ULN, >2×ULN
ALP	>2×ULN, >3×ULN
AT & TBL	AT >3×ULN & TBL >2×ULN; AT >5×ULN & TBL >2×ULN; AT >10×ULN & TBL >2×ULN
ALP & TBL	ALP >3×ULN & TBL >2×ULN
AT & TBL & ALP (i.e., Hy's Law)	AT >3×ULN & TBL >2×ULN & ALP <2×ULN

LFTs will also be listed by visit.

5.6.3.2. Coagulation

Descriptive statistics of all parameters over time will be performed. Values at each visit and changes from baseline (for quantitative parameters) will be calculated and tabulated. Unscheduled visits or retest will not be analyzed for the summary by visit.

Shift tables of maximum post baseline NCI-CTCAE grades according to baseline NCI_CTCAE grade will be provided for INR (high values): It will be graded using NCI-CTCAE Version 4.03. Grade 0 will be assigned for all non-missing values not graded as ≥1. Grade 5 will not be used.

In addition, for parameters that cannot be graded, shift tables of post baseline normal/abnormal values according to baseline values will also be displayed.

5.6.3.3. Urinalysis

Descriptive statistics of all parameters over time will be performed. Values at each visit and changes from baseline (for quantitative parameters) will be calculated and tabulated. Unscheduled visits or retest will not be analyzed for the summary by visit.

Microscopy findings (normal/abnormal and clinical significance) will be listed as part of the urinalysis listing.

5.6.3.4. Viral Serology and tests to confirm postmenopausal status for females

All parameters will be summarized at screening and presented in a data listing. Unscheduled visits or retests will not be analyzed but listed only.

5.6.4. Physical Examinations

A physical examination will be carried out on each body system at screening and repeated on Cycle 1 Day 1 (if not performed within 72 hours before Cycle 1 Day 1 (i.e. first day of study intervention), on Day 1 of each subsequent cycle, the end of treatment visit and the 30-day safety follow-up visit. Body weight will also be measured as part of the physical examination.

A complete physical examination including assessments of the cardiovascular, respiratory, gastrointestinal, dermatological, ophthalmological and neurological systems (at a minimum) will be performed at screening. Physical examinations should be targeted as clinically indicated at subsequent visits.

Measurements of BSA will be made on Cycle 1 Day 1 and Day 1 of each subsequent cycle to calculate the dose for study intervention infusions. All assessments during the treatment phase should be made predose.

Descriptive statistics of all parameters at each scheduled visit will be performed in participants from the SS:

- BSA (m^2)
- Weight (kg)
- Frequency of abnormal results of physical examination (normal/abnormal for the global or targeted examination).

Physical examination results (normal/abnormal and clinical significance) and details of abnormalities will also be listed for each participant and visit (scheduled or not).

BSA and Weight will be listed as part of the Vital Signs (see Section 5.6.5).

Ophthalmic examination

All ophthalmic examination results (i.e., tonometry, visual acuity, fundoscopy, slit lamp) and changes from baseline will be listed.

Considering limited number of post-screening assessments, no summary tables will be provided.

Visual Acuity (Standard Logarithmic Visual Acuity Chart: “L Value” expected to be between 4.0 and 5.3):

L values ≥5 correspond to good vision, L values <5 correspond to poor vision.

Visual acuity scores will be listed by eye and treatment arm.

Tonometry (Intraocular Pressure):

Intraocular pressure above 30 mmHg (which is considered as clinically significant) will be flagged into the individual data listing.

Fundoscopy and slit lamp

Results from fundoscopy and slit lamp examinations will be listed.

5.6.5. Vital Signs

Tympanic temperature (°C), heart rate (beats/min), respiratory rate (breaths/min), Systolic Blood Pressure (SBP, in mmHg) and Diastolic Blood Pressure (DBP, in mmHg) will be measured using study site standard techniques at screening and repeated on each visit during the treatment phase, the end of treatment visit and the 30-day safety follow-up visit.

All parameters will be described at each scheduled visit with changes from baseline in participants from the SS.

In addition, shift tables of worst post baseline grade (worst on-treatment as per section 5.1.4) according to the baseline grade will be provided for blood pressure hypertension criteria (systolic, diastolic, and both combined). It will be derived as follows:

- Systolic BP (SBP) only (mmHg):
  - Grade 0: < 120
  - Grade 1: 120-139
  - Grade 2: 140-159
  - Grade 3: ≥ 160
  - No Grade 4 or 5

- Diastolic BP (DBP) only (mmHg):
    - Grade 0: < 80
    - Grade 1: 80-89
    - Grade 2: 90-99
    - Grade 3: ≥ 100
- Systolic and diastolic BP combined (mmHg). Maximum grade based on SBP and DBP values at the same assessment will be considered:
    - Grade 0: SBP < 120 and DBP < 80
    - Grade 1: SBP in 120-139 or DBP in 80-89
    - Grade 2: SBP in 140-159 or DBP in 90-99
    - Grade 3: SBP ≥ 160 or DBP ≥ 100
    - No Grade 4 or 5

The following criteria define clinically notable abnormalities:

- Clinically notable elevated values
    - Systolic blood pressure (BP): ≥ 160 mmHg and an increase ≥ 20 mmHg from baseline;
    - Diastolic BP: ≥ 100 mmHg and an increase ≥ 15 mmHg from baseline;
    - Heart rate): ≥ 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: ≤ 50 bpm with decrease from baseline of ≥ 15 bpm;
    - Weight: ≥ 20% decrease from baseline;
    - Body temperature [C]: ≤ 36 C.

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

Participants with clinically notable vital sign abnormalities will be listed by treatment arm.

All vital signs parameters and clinically notable values flag will be also presented in a data listing.

5.6.6. ECG

12-lead ECGs will be collected as follows:

- A single ECG measurement will be performed at screening to determine eligibility.
- During Cycle 1, ECGs will be performed as follows:
  - A triplicate ECG predose on Day 1 (conducted within approximately 5 to 10 minutes total time). The mean of the triplicate ECG measurements will serve as the participant's baseline value for all postdose comparisons.
  - For participants receiving the doublet only (SLI phase and doublet arm of the randomized phase): a single ECG at 2 (±0.5) hours after administration of encorafenib and before the start of the cetuximab infusion on Day 1.
  - A single ECG predose on Day 15.
- During Cycle 2, ECGs will be performed as follows:
  - A single ECG predose on Day 1.
  - For participants receiving the doublet only (SLI phase and doublet arm of the randomized phase): a single ECG at 2 (±0.5) hours after administration of encorafenib and before the start of the cetuximab infusion on Day 1.
- Further ECGs will be performed on Day 1 predose in all subsequent cycles, the end of treatment visit and the 30-day safety follow-up visit.

12-lead ECGs will be performed before blood collection, where applicable.

ECG parameters (including interpretation) will be described at each scheduled visit by applicable treatment group and listed in participants from the SS, with a flag for clinically notable values.

Frequency counts and percentages of participants having at least one post-baseline clinically notable ECG values during study (Table 13) will also be described and listed.

Table 13: Clinical Notable ECG Criteria

Parameter	Criterion
QT, QTcF	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 bpm
	Decrease from baseline > 25% and to a value < 50 bpm

5.6.7. Dermatological Examinations

Dermatological examinations will be performed for all participants at screening then for participants receiving encorafenib (SLI phase and doublet arm of the randomized phase) only. This is to monitor for the possible development of keratoacanthoma and/or squamous cell carcinoma, as these have been reported to occur with selective BRAF inhibitor treatment. This assessment will be performed at screening and every 8 weeks from Cycle 1 Day 1 (i.e. on Day 1 of Cycles 1, 3, 5, 7…), the end of treatment visit and the 30-day safety follow-up visit. All assessments during the treatment phase should be made predose.

Following the 30-day safety follow-up (when clinically appropriate) it is recommended further dermatological examinations be performed for up to 6 months after the last encorafenib dose or until initiation of another antineoplastic therapy.

All dermatological parameters will be described at each scheduled visit in participants by applicable treatment group from the SS and will be presented in a data listing.

5.6.8. ECOG Performance Status

An assessment of ECOG performance status will be made at screening. All participants must have a score of 0 or 1 for inclusion. Assessments will be repeated on Cycle 1 Day 1 (if not performed within 72 hours before Cycle 1 Day 1 (i.e. first day of study intervention), on Day 1 of each subsequent cycle, the end of treatment visit and the 30-day safety follow-up visit.

ECOG performance status will be described at each visit in participants from the SS and presented in a data listing.

In addition, shift tables in ECOG performance status will be also provided for worst on-treatment value by treatment group on the SS (worst on-treatment as per section 5.1.4).

5.7. Other Analyses

5.7.1. Other Variables and/or Parameters

5.7.1.1. PK

5.7.1.1.1. PK samples

For participants receiving the doublet only (SLI phase and doublet arm of the randomized phase), blood samples for characterization of the PK of encorafenib and cetuximab will be collected on the first day of treatment (Cycle 1 Day 1) and at steady state after 1 month treatment (Cycle 2 Day 1).

Serial blood samples will be collected from a subset of 24 participants treated at the recommended dose at selected sites (all nine participants in the SLI phase and the first 15 participants at the selected sites in the doublet arm of the randomized phase). Sparse sampling will be performed in all remaining participants in the doublet arm.

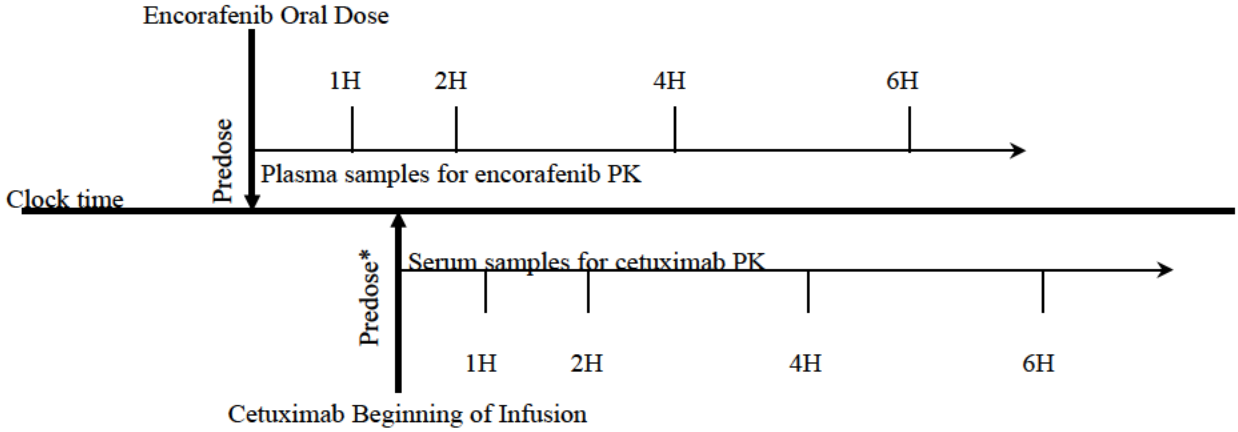
The timepoints for the serial and sparse sampling schedules are shown in Table 14.

Table 14: Sampling for Serial and Sparse Pharmacokinetic Schedules

	N	Cycle 1 Day 1					Cycle 2 Day 1				
		Pre	1 H	2 H	4 H	6 H	Pre	1 H	2 H	4 H	6 H
SLI Phase – Serial Sampling											
Encorafenib sampling[a]	9	X	X	X	X	X	X	X	X	X	X
Cetuximab sampling[b]		X	X	X	X	X	X	X	X	X	X
Randomized Phase – Doublet Arm Serial Sampling											
Encorafenib serial sampling[a]	15	X	X	X	X	X	X	X	X	X	X
Cetuximab sampling[b]		X	X	X	X	X	X	X	X	X	X
Randomized Phase – Doublet Arm Sparse Sampling											
Encorafenib spare sampling[a]	48			X		X	X		X		
Cetuximab sparse sampling[b]				X		X	X		X		

Abbreviations: H = hours postdose; pre = predose; SLI = safety lead-in, N = number of participants.
[a] Timepoints are in relation to the encorafenib dose.
[b] The timepoints are in relation to the start of the cetuximab infusion.

Figure 3: Graphic Representation of Pharmacokinetic Sampling Plan



**Predose PK samples for cetuximab analysis should be collected just prior the beginning of the cetuximab infusion*

Abbreviations: H = hours postdose; PK = pharmacokinetics

If a participant experiences an adverse event that results in an unscheduled visit or meets the criteria for an SAE, a further blood sample should be collected (if feasible and if less than 24 hours have elapsed since the last dose).

Study visits for PK sampling should be scheduled in the morning so that correct predose and postdose PK blood samples can be collected. The encorafenib dose on PK visit days should be taken at the study site, only after collecting the predose PK sample.

5.7.1.1.2. Non-compartmental PK Analysis

Encorafenib and cetuximab concentrations will be transmitted by the bioanalytical laboratory to Fortrea.

PK analyses will be performed on the PK Set.

The following PK parameters for encorafenib and cetuximab will be determined where possible for the individuals with serial blood samples using non-compartmental methods in the validated software program Phoenix WinNonlin (Certara, Version 8.1 or higher):

Parameter	Definition
AUC_{0-6}	area under the concentration-time curve from time 0 to 6 hours post dose
AUC_{tau}	area under the plasma concentration-time curve over a dosing interval at steady-state (encorafenib, Cycle 2 Day 1 only) *
$AUC_{0-tlast}$	area under the concentration-time curve from time 0 to the time of the last quantifiable concentration (t_{last})
C_{max}	maximum observed concentration
t_{max}	time of the maximum observed concentration
C_{min}	minimum observed concentration following multiple dosing (taken as the predose concentration on Cycle 2 Day 1) – determine for Cycle 2 Day 1 only
C_{last}	last quantifiable concentration
t_{last}	time of last quantifiable concentration
AR_{AUC}	observed accumulation ratio based on AUC_{0-6}

* Derived by imputing the Cycle 2 Day 1 predose concentration as the concentration at the end of the dosing interval (i.e 24 hours)

Additional PK parameters may be determined where appropriate.

During the study, a preliminary PK analysis will be carried out after the last SLI participant C2D1 PK sampling, using the theoretical blood sampling times (planned time or scheduled time) and the validated concentration data.

After the data base lock, the final PK analysis will be carried out using the validated concentrations data and the actual blood sampling times postdose where possible. If an actual time is missing, nominal time (i.e. theoretical time) will be assigned to the sample concentration.

The parameters C_{max}, C_{min}, C_{last}, t_{last} and t_{max} will be obtained directly from the concentration-time profiles. If C_{max} occurs at more than 1 timepoint, t_{max} will be assigned to the first occurrence of C_{max}.

AUC_{tau} will be derived by setting the predose concentration as equivalent to concentration at the end of the dosing interval (24 hours postdose) based on the assumption that steady-state has been attained by Cycle 2 Day 1 and that the concentrations decline in a monophasic manner from 6 to 24 hours postdose.

All AUC values will be calculated using the linear trapezoidal rule for increasing concentrations and the logarithmic rule for decreasing concentrations (linear up/log down rule).

AR_{AUC} will be derived by dividing AUC_{0-6} from Cycle 2 Day 1 by AUC_{0-6} from Cycle 1 Day 1.

5.7.1.1.2.1. Criteria for Handling Concentrations Below the Limit of Quantification

Handling of PK concentrations below quantification limit (BQL) will be applied as follows:

	Substitution value if BQL occurs:		
Type of analysis	Before t_{max}	After t_{max} & between 2 quantifiable concentrations	After t_{max} & NOT between 2 quantifiable concentrations
PK Non-compartmental analysis	0	Missing	Missing
Summary statistics	0	0	0
Plotting of individual data	0	Missing	Missing
Listing of individual data	BQL	BQL	BQL

If an entire concentration-time profile is BQL, it will be excluded from PK analysis.

If a predose concentration on Cycle 1 Day 1 is missing, it will be set to zero by default for PK analysis.

5.7.1.1.2.2. Treatment of Outliers in PK Analysis

If a value is considered to be anomalous due to being inconsistent with the expected PK profile, it may be appropriate to exclude the value from the PK analysis. However, the exclusion of any data must have strong justification and will be documented in the CSR.

Any quantifiable predose concentration value on Cycle 1 Day 1 will be considered anomalous and set to missing for the PK analysis.

5.7.1.1.2.3. Criteria for Calculation of AUC

The minimum requirement for the calculation of AUC values will be the inclusion of at least three consecutive plasma concentrations above Lower Limit Of Quantification (LLOQ). If there are only 3 consecutive concentrations, at least one should follow C_{max}.

5.7.1.1.3. Population PK Analysis

A population PK analysis will be performed, pooling data from at least the ARRAY-808-302 study (see Protocol Section 2.1.4). Race as a covariate for encorafenib and the potential interaction between encorafenib and cetuximab in the Chinese population will be explored. Individual PK parameters (AUC, C_{min} and C_{max}) of Chinese participants enrolled in the current study will be calculated using the population PK model mentioned above. If appropriate, exploratory and descriptive exposure-response relationships will be performed. Details of these analyses will be provided in a specific standalone modelling plan.

The modelling results will be reported in a separate report.

5.7.1.1.4. Presentation of PK Data

- Preliminary analysis

Preliminary individual PK encorafenib and cetuximab parameters will be listed and summarized.

Individual and mean concentration-time profiles will be presented graphically for encorafenib and cetuximab on both linear and semi-logarithmic scales.

Only Phoenix WinNonlin output will be used at this stage.

- Final analysis

All serial and sparse encorafenib plasma concentrations and cetuximab serum concentrations will be listed by cycle and summarised by cycle. Concentrations from participants with serial PK sampling will be presented separately from participants with sparse PK sampling in the tables, figures and listings.

Listing of samples collection (both serial and sparse PK sampling) and concentrations will be provided on Safety set (SS), where participants excluded from PK Set will be flagged (**).

Listing of PK parameters will be provided on PK Set.

Figures and tables will be provided on PK Set.

If the actual time of sample collection deviates from the nominal time by more than ±10%, the concentration will be flagged and excluded from the summary statistics. Individual concentrations deemed to be anomalous will be flagged in the listings, with the reason for exclusion footnoted, and excluded from the summary statistics.

PK concentrations will be summarised by cycle using the description statistics: geometric mean, geometric coefficient of variation (CV%), arithmetic mean, arithmetic standard deviation (SD), median, minimum, maximum, number of observations (n), and number of participants (N). The following rules will apply for summarising of PK concentrations:

- At least 3 evaluable data (N≥3), otherwise only the minimum and the maximum are displayed and other statistics are indicated as "Not Applicable" (NA)
- Values that are BQL will be set to 0 for the calculation of summary statistics.
- Arithmetic mean or median values that are BQL will be presented as 0.
- If any BQL results (treated as 0) are in a series of summarized data, geometric mean and CV% of geometric mean will be reported as not calculated (NC).

PK parameters determined from serial PK sampling will be listed by cycle and summarised by cycle using the description statistics: geometric mean, geometric coefficient of variation (CV%), arithmetic mean, arithmetic standard deviation (SD), median, minimum, maximum, number of observations (n), and number of participants (N), with the exception of t_{max}, t_{last} where only N, n, median, minimum and maximum will be presented.

For derived PK parameters, descriptive statistics are calculated according to the following rules:

- at least 3 evaluable data (N≥3) are available, otherwise only the N, n, minimum and the maximum are presented and other statistics are shown as NA
- at least 50% of the participants have evaluable data, otherwise only N, n, minimum and maximum are presented, and other statistics are shown as NC

For encorafenib only, participant's PK concentrations and parameters will be flagged and excluded from descriptive statistics if:

- vomiting occurs within the first 4 hours postdose on Cycle 1 Day 1 and Cycle 2 Day 1
- vomiting episode occurs within the first 4 hours postdose on the day of the last encorafenib dose prior to collection of PK samples (i.e relating to the day before Cycle 2 Day 1)
- a participant missed any encorafenib dose or dose reduced within the 3 days immediately prior or the day of Cycle 2 Day 1 PK sample collection (i.e for PK parameters and concentrations)

Individual and summary serial concentration-time profiles will be presented graphically for encorafenib and cetuximab on both linear and semi-logarithmic scales, using the following 3 types of plots:

- Graphical display of the concentration-time profiles (spaghetti plot), overlaying individual profiles (using actual sampling times) and the median concentration-time profile (using nominal sampling times). Figure will be repeated by analyte (encorafenib and cetuximab)

and by cycle (Cycle 1 Day 1 and Cycle 2 Day 1). Participants excluded from statistical analysis will be flagged.

- Individual profiles are by-participants plots and will use the actual sampling times. Cycle 1 Day 1 and Cycle 2 Day 1 concentration-time profiles will be overlaid on the same plot. Figure will be repeated by analyte (encorafenib and cetuximab). Participants excluded from statistical analysis will be flagged.
- Arithmetic mean concentration-time profiles with error (+/-SD) bars using the nominal sampling times. Cycle 1 Day 1 and Cycle 2 Day 1 concentration-time profiles will be overlaid on the same plot. Figure will be repeated by analyte (encorafenib and cetuximab).

Linear and semi logarithmic plots will be displayed on the same page side by side as appropriate.

For sparse sampling: concentrations of all the participants will be plotted in a single graph by cycle and repeated by analyte and using linear and semi logarithmic plots. Participants excluded from statistical analysis will be flagged.

5.7.2. Subgroup Analyses

5.7.2.1. Efficacy Analyses

The following subgroups will be studied for primary endpoint of the randomized phase, provided that the number of participants randomized with these particular covariates allowed (i.e., at least 10 events available in the considered subgroup, for both treatment groups combined):

1. ECOG performance status (0 versus 1) (randomization strata)
2. Prior use of irinotecan (Yes versus No) (randomization strata)
3. Number of prior regimens in the metastatic setting (1 vs. 2)
4. Age (< 65 vs. ≥ 65 years)
5. Gender (Male versus Female)
6. Number of organs involved at baseline based on Target and Non-target lesion assessment (≤ 2 versus ≥ 3)
7. MSI (high versus normal)
8. BRAF V600E mutation per central assessment (positive versus negative/indeterminate)
9. Baseline CEA (≤ 5 versus > 5 µg/L)
10. Baseline CA 19-9 (≤ 35 versus > 35 U/mL)
11. Baseline CRP (≤ 0.01 versus > 0.01 g/L)
12. Removal status of primary tumor (no resection/partial resection, complete resection)
13. Side of tumor ("Colon, left" versus "Colon, right") as per following classification rules:
    - Primary tumor location reported as "COLON, TRANSVERSE" will be considered as right-sided.
    - Primary tumor location reported as "Other" will be considered into existing category when appropriate. In particular:
        - "CAECUM"," CECUM", "ASCENDING COLON", "HEPATIC FLEXURE", "COLON, RIGHT COLON TRANSVERSUM", "ILEOCECAL BOWEL" or "ADENOCARCINOMA OF COLON AND LIVER" will be considered as right-sided
        - "SIGMOID", "RECTUM", "RECTAL", "DESCENDING COLON", "ANUS" or "SPLENIC FLEXURE" will be considered as left-sided.
    - For subgroup analysis purpose only, participants with primary tumor location reported as "Other" (except modalities considered as right-sided or left-sided, as per above rules) or "Unknown" will not be included.
14. Presence of liver metastases at baseline, based on Target and Non-target lesion assessment according to BICR (yes versus no)

Subgroups selected for analysis will be described in tabular with number of PFS events and using KM methods, reporting estimated median PFS (in months) with 95% CI, 25th and 75th percentiles. Within each subgroup, an unstratified Cox regression model will be used to estimate the hazard ratio of the treatment effect and the corresponding 95% CI based on Wald test. No p-value will be provided.

The effects of the primary endpoint (PFS) will be displayed using a forest plot of the treatment effect hazard ratios by subgroups (unstratified Cox model).

5.7.2.2. Safety Analyses

Selected safety analyses will be performed on various subgroups:

- For TEAEs: Following selected TEAE tables will be repeated in the subgroups:
    - TEAE by SOC and PT
    - Related TEAE by SOC and PT
    - Serious TEAEs by SOC and PT
    - Related Serious TEAEs by SOC and PT
    - TEAE Leading to Dose Reduction of any study drug by SOC and PT
    - TEAE Leading to Drug Interruption of any study drug by SOC and PT
    - TEAE Leading to Study Intervention Discontinuation of all study drug by SOC and PT
    - TEAE by descending PT
    - Related TEAE by descending PT
    - Serious TEAEs by descending PT
    - Related Serious TEAEs by descending PT
    - TEAE resulting in on-treatment death by descending PT
    - TEAE Leading to Dose Reduction of any study drug by descending PT
    - TEAE Leading to Drug Interruption of any study drug by descending PT
    - TEAE Leading to Study Intervention Discontinuation of all study drug by descending PT

  The following subgroups will be analyzed:

    - Age category (< 65, ≥ 65, <75, ≥ 75 years) at screening visit
    - Gender (Male versus Female)
    - Number of prior regimens in the metastatic setting (1 versus 2)
    - Removal status of primary tumor (no resection/partial resection, complete resection)
    - Liver metastases at baseline based on either disease characteristics or tumor lesion eCRFs:
        - No liver metastases
        - Liver metastases and both AST and ALT < 3 x ULN
        - Liver metastases and AST and/or ALT ≥ 3 x ULN
- For duration of exposure (in weeks): Analyses to be repeated in all following subgroups in each drug of each arm:
    - Age category (< 65, 65-74, ≥ 75 years) at screening visit;
    - Gender (Male versus Female);
    - Number of prior regimens in the metastatic setting (1 versus 2);
    - Removal status of primary tumor (no resection/partial resection, complete resection);
    - Liver metastases at baseline based on either disease characteristics or tumor lesion eCRFs:

- No liver metastases
- Liver metastases and both AST and ALT < 3 × ULN
- Liver metastases and AST and/or ALT ≥ 3 × ULN

- Baseline cardiac risk (Yes, No), based on Medical History as defined in Section 6.6.
- Baseline renal impairment based on estimated Creatinine Clearance (CrCl):
  - Normal (CrCl > 90 mL/min)
  - Mild impairment (CrCl between 60 and 90 mL/min)
  - Moderate impairment (CrCl between 30 and < 60 mL/min)
  - Severe impairment (CrCl < 30 mL/min)

5.8. Interim Analyses

There is no interim analysis planned.

5.8.1. Data Monitoring Committee

An independent DMC will review the available safety information in the SLI phase in order to evaluate the tolerability of the doublet prior to the start of the randomized phase. If the independent DMC determines the doses to be tolerable in the first nine evaluable participants the study will proceed to the randomized phase. The independent DMC will make a recommendation if the doses are not deemed tolerable (see Protocol Section 4.1.3).

To be evaluable for tolerability assessment by the independent DMC prior to the randomized phase of the study, participants must:

- Experienced an event meeting the DLT criteria or
- Received ≥75% dose intensity [(administered dose in mg/planned dose in mg) x 100] of each of the two drugs (encorafenib and cetuximab) during Cycle 1.

Non-evaluable participants will be replaced in order to achieve nine evaluable participants in the SLI phase.

Enrollment will be put on hold if at least three participants experience DLTs until a discussion with the independent DMC can occur.

If DLTs are not observed in at least three of nine evaluable participants, the independent DMC will review data after the ninth participant has been followed for at least one 28-day cycle (Cycle 1). All data available at the time of review will be included (as these participants will be enrolled over time, the independent DMC will have cumulative data available beyond the first 28 days on some participants). If the independent DMC determines the doses to be tolerable in the first nine evaluable participants based on observing DLTs in <33% participants and evaluation of the overall toxicity profile, the study will proceed to the randomized phase. The independent DMC will make a recommendation if the doses are not deemed tolerable.

The independent DMC will further review the available safety information after the first 15 participants of the randomized phase treated with encorafenib and cetuximab have completed at least one cycle of treatment to confirm tolerability and will then be responsible for reviewing all safety data at regular intervals (every 6 months at a minimum).

The independent DMC membership, data to be reviewed, timing of the planned reviews as well as the operating procedures will be described in the DMC Charter.

6. Supporting Documentation

6.1. Appendix 1: List of Abbreviations

Abbreviation	*Definition*
5-FU	*5-fluorouracil*
AE	*Adverse Event*
AESI	*Adverse Event of Special Interest*
ALP	*Alkaline Phosphatase*
ALT	*Alanine Aminotransferase*
ANC	*Absolute Neutrophil Count*
aPTT	*activated Partial Thromboplastin Time*
AST	*Aspartate Aminotransferase*
ATA	*Adequate Tumor Assessment*
ATC	*Anatomical Therapeutic Classification*
AUC	*Area Under the Curve*
BICR	*Blinded (to treatment received) Independent Central Review*
BQL	*Below Quantification Limit*
BRAF	*B-RAF Proto-oncogene, Serine/threonine Kinase*
BRAF V600E	*B-RAF Proto-oncogene, Serine/threonine Kinase V600E Mutant*
BRAF wt	*B-RAF Proto-oncogene, Serine/threonine Kinase Wild Type*
BMI	*Body Mass Index*
BOR	*Best Overall Response*
BSA	*Body Surface Area*
CCI	

CCI

CI	*Confidence Interval*
C_{last}	*last quantifiable concentration*
C_{max}	*Maximum Concentration*
C_{min}	*Minimum Concentration*
CR	*Complete Response*
CRC	*Colorectal Cancer*
CrCl	*Creatinine Clearance*
CSR	*Clinical Study Report*
CTCAE	*Common Terminology Criteria for Adverse Event*
DBP	*Diastolic Blood Pressure*
DCR	*Disease Control Rate*
DDS	*Dose-Determining Set*
DI	*Dose Intensity*
DILI	*Drug-Induced Liver Injury*
DLT	*Dose-Limiting Toxicity*
DMC	*Data Monitoring Committee*
DOR	*Duration Of Response*
ECG	*Electrocardiogram*
ECOG	*Eastern Co-operative Oncology Group*
eCRF	*Electronic Case Report Form*
EAIR	*Exposure-Adjusted Incidence Rate*
EGFR	*Epidermal Growth Factor Receptor*
EMA	*European Medicines Agency*

EORTC QLQ-C30	*European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire for Cancer Patients*
EQ-5D-5L	*EuroQoL Five Dimensions, Five Levels*
ES	*Efficacy Set*
EuroQoL	*Euro Quality of Life*
FFPE	*Formalin-fixed and Paraffin Embedded*
FACT-C	*Functional Assessment of Cancer Therapy-Colon Cancer*
FA	*Folinic Acid*
FAS	*Full Analysis Set*
FDA	*US Food and Drug Administration*
FOLFIRI	*5-fluorouracil/Folinic Acid+Irinotecan*
INR	*International Normalized Ratio*
IRT	*Interactive Response Technology*
KM	*Kaplan-Meier*
LDH	*Lactate Dehydrogenase*
LFT	*Liver Function Test*
LLN	*Lower Limit of Normal*
LLoQ	*Lower Limit of Quantification*
mCRC	*Metastatic Colorectal Cancer*
MedDRA	*Medical Dictionary for Regulatory Activities*
MRI	*Magnetic Resonance Imaging*
MSI	*Microsatellite Instability*
MSI-H	*MSI-High*
MSS	*Microsatellite stable*

NCI	*National Cancer Institute*
NCI-CTCAE	*National Cancer Institute-Common Terminology Criteria for Adverse Event*
NE	*Not Evaluable*
ORR	*Objective Response Rate*
OS	*Overall Survival*
PCR	*Polymerase Chain Reaction*
PD	*Progressive Disease*
PDI	*Planned Dose Intensity*
PFS	*Progression Free Survival*
PGIC	*Patient Global Impression of Change*
PK	*Pharmacokinetic*
PPS	*Per Protocol Set*
PR	*Partial Response*
PRO	*Patient Reported Outcomes*
PT	*Preferred Term*
RBC	*Red Blood Cells*
RDI	*Relative Dose Intensity*
RECIST v1.1	*Response Evaluation Criteria In Solid Tumors version 1.1*
SAE	*Serious Adverse Event*
SAP	*Statistical Analysis Plan*
SAS®	*Statistical Analysis Software*
SBP	*Systolic Blood Pressure*
SD	*Stable Disease*
SI	*System International*

SLD	*Sum of Lesion Diameters*
SLI	*Safety Lead-In*
SOC	*System Organ Class*
SMQs	*Standardized MedDRA Queries*
SS	*Safety Set*
TBL	*Total Bilirubin*
t_{last}	*time of last quantifiable concentration*
t_{max}	*time of the maximum observed concentration*
TEAE	*Treatment Emergent Adverse Event*
TFL	*Tables Figures and Listings*
TOC	*Table Of Contents*
TTR	*Time To Response*
ULN	*Upper Limit of Normal*
VAS	*Visual Analogue Scale*
WBC	*White Blood Cells*
WHO	*World Health Organization*
WOCBP	*Women of childbearing potential*

6.2. Appendix 2: Changes to Protocol-planned Analyses

Patient Reported Outcome: The reasons for non-completion of the questionnaires will not be summarized.

6.3. Appendix 3: List of TFLs

The Table Of Contents (TOC) of Tables Figures and Listings (TFLs) will be provided in a separate document, as well as the shell TFLs and specifications.

In the TOC of TFLs, the outputs selected for independent DMC will be flagged.

6.4. Appendix 4: Disease Characteristics and demographics

All baseline and demographic characteristics will be summarized by treatment group and overall. This will include gender, race, age groups, weight, height, Body Mass Index (BMI), BSA and ECOG performance status. In addition, number of participants in each studied subgroup will be presented (see section 5.7.2).

All demographic data and baseline characteristics will be summarized in output tables by treatment groups in all participants from the FAS, the SS and the DDS and in all other populations with a difference of more than 15% with FAS (the ES, the PPS), and presented in data listings on FAS.

The Baseline Characteristics and demographics data are described in the Table 15 below:

Table 15: Baseline Characteristics and demographic data

Variable	Quantitative parameters	Qualitative parameters	Derived variable	FAS	SS	ES **	PPS **	DDS
Country		X		X	X	X	X	X
Sex (M/F)		X		X	X	X	X	X
Race		X		X	X	X	X	X
Age (years)	X			X	X	X	X	X
Age group		X	X	X	X	X	X	X
Weight (kg)	X			X	X	X	X	X
Height (cm)	X			X	X	X	X	X
BMI (kg/m²)	X		X	X	X	X	X	X
BSA (m²)				X	X	X	X	X
ECOG performance status at Baseline (eCRF value) – displayed in the stratification factors table		X		X	X	X	X	X
Prior use of irinotecan (eCRF value) – displayed in the stratification factors table				X	X	X	X	X
RASwt (Local)		X		X	X	X	X	X
BRAF V600E mutation per local assessment		X		X	X	X	X	X
BRAF V600E mutation per central assessment		X		X	X	X	X	X

*: Only for randomized phase columns
**: If a difference more than 15% with FAS

Only for FAS, following tables will also be provided:

- Local versus Central BRAF Status
- Stratification factors (eCRF vs IRT)
- Cross-classification for Randomization by stratification factors (using IRT values)

Notes:

- The Age described in the table is the age reported in the eCRF.
- Unless otherwise specified, Age will also be described by classes, using following classes: < 65, 65-74, ≥75 years.
- Listing will also present age, weight (kilograms (kg)), BSA (m^2), BMI (kg/m^2).
- BMI will be calculated as weight(kg)/height^2(m)

6.5. Appendix 5: CRC Cancer History

CRC Cancer History (Disease Characteristics) will be summarized by treatment groups in participants from the FAS, the SS, the DDS and in all other populations with a difference of more than 15% with FAS (the ES and the PPS).
CRC Cancer History will also be listed on FAS.

Note:

- In case of missing initial diagnosis date, substituting rules defined in appendix 6.9 (section 6.9.3) will be applied.

The CRC Cancer History data are described in the Table 16 below:

Table 16: CRC Cancer History data

Variable	Quantitative parameters	Qualitative parameters	Derived variable	FAS	SS	ES*	PPS*	DDS
Time since initial diagnosis of primary site (months) computed as [reference date minus the date of initial diagnosis + 1]/30.4375 *With Reference date= Randomization date for randomized participants, Treatment assignment date for SLI patients, Date of main (screening) ICF for screened participants not assigned to treatment.*	X		X	X	X	X	X	X
Primary tumor location as per classification rules reported in Section 5.7.2.1 (Left, Right, Other, Unknown)		X		X	X	X	X	X
Removal status (no resection, partial resection, complete resection)		X		X	X	X	X	X
Delay between diagnosis of the primary tumor and metastasis (months)	X		X	X	X	X	X	X
Delay between diagnosis of the primary tumor and metastasis (Simultaneous, 0.1–12 months, > 12 months)		X	X	X	X	X	X	X

T, N and M at study entry		X	X (Grouped TNM e.g. T2N0M1)	X	X	X	X	X
Stage at study entry			X (Group IV includes all items such as IV, IVM1a, IVM1b and IVM1c)	X	X	X	X	X
Number of organs affected by metastases ("Number of metastatic organs at study entry") (1, 2, >2)		X		X	X	X	X	X
Number of Prior Regimens in the metastatic setting		X		X	X	X	X	X
Baseline CEA (≤ 5 versus > 5 µg/L)		X	X	X	X	X	X	X
Number of Organs involved at baseline based on Target and Non-target lesion assessment per BICR (≤ 2 versus ≥ 3)		X	X	X	X	X	X	X
Baseline CA 19-9 (≤ 35 versus > 35 U/mL)		X	X	X	X	X	X	X
CRP (mg/L) at Baseline	X			X	X	X	X	X
Baseline CRP (≤ 0.01 versus > 0.01 g/L)		X	X	X	X	X	X	X
Presence of liver metastases at baseline, based on Target and Non-target lesion assessment per BIRC (yes versus no)		X	X	X	X	X	X	X
Liver metastases at baseline based based on Target and Non-target lesion assessment per BIRC (No liver metastases/ Liver metastases and both AST and ALT < 3 x ULN/ Liver metastases and AST and/or ALT ≥ 3 x ULN)		X	X	X	X	X	X	X
Baseline cardiac risk (Yes / No)		X	X	X	X	X	X	X
Baseline Renal Impairment based on estimated CrCl		X	X	X	X	X	X	X

(Normal/ Mild/ Moderate/ Severe)								

*: If a difference more than 15% with FAS

6.6. Appendix 6: Medical History, Baseline Signs and Symptoms

Medical history data will be coded using MedDRA dictionary (Version 26.0 at set-up, then latest version in use) and summarized by SOC and PT with the number and percentage presented by treatment arm and overall for the FAS, SS and DDS. A by-participant listing of medical history information will be provided on FAS.

For defining the "Baseline cardiac risk" subgroup, medical histories will be selected as per Table 17 below.

Table 17: Medical history terms for cardiac risk

Disease	Terms
Hypertension	SMQ: Hypertension [broad]
Diabetes	PT: Diabetic vascular disorder
	HLT: Diabetes mellitus (incl. subtypes)
	HLT: Diabetic complications cardiovascular
Hyperlipidemia	SMQ: Dyslipidaemia [broad]
Cardiac Disorders	SOC: Cardiac disorders
Arteriosclerosis	PT: Aortic arteriosclerosis
	PT: Arteriosclerosis
	PT: Peripheral arterial occlusive disease
Ischaemic heart disease	SMQ: Ischaemic heart disease [narrow]

Incomplete dates will be handled as described in Appendix 6.9.

6.7. Appendix 7: Prior/concomitant/follow-up therapies

Medications will be coded using the latest version of the World Health Organization (WHO) Drug Dictionary (WHO Drug Version March 2021 Enhanced Dictionary Version Global B3 at set-up, then latest version in use) and recorded on the "Prior/Concomitant Medications" and "Prior Anti-Neoplastic Therapy – Medication" and "Anti-neoplastic therapy since study treatment discontinuation – Medication" eCRF pages.

Medications used to treat Covid-19 (SARS-CoV-2) will be identified using WHO Drug Standardised Drug Groupings (SDG) "Drugs and vaccines for COVID-19", downloadable on the WHODrug Insight tools.

Therapeutic / Diagnostic Procedures will be coded using the MedDRA dictionary (Version 24.0 at set-up, then latest version in use) and recorded on the "Medical and Surgical Procedures" eCRF pages.

"Prior Anti-Neoplastic Therapy – Surgery" and "Anti-neoplastic therapy since study treatment discontinuation – Surgery" will be coded similarly.

Procedures used to treat Covid-19 (SARS-CoV-2) will be flagged when the "Adverse Event ID linked to the procedure" or "Medical History ID linked to the procedure" is coded as related to Covid-19 (SARS-CoV-2), meaning when AE/MH verbatim includes "COVID-19" as prefix, or when the "Reason (Indication)" provided in CRF is related to Covid-19 (SARS-CoV-2).

The FAS, the SS and the DDS will be used to describe all Prior/concomitant/follow-up therapies summaries, unless otherwise specified. Listings will be provided on the FAS.

Where a medication start date is partially or fully missing, and it is unclear as to whether the medication is prior or concomitant, it will be assumed that it is concomitant.

6.7.1. Prior Therapies

Prior therapies are defined as any therapies starting and ending before the start of study treatment.

Prior therapies include:

- prior antineoplastic treatments
- prior medication (excluding prior anticancer treatments).
- diagnostic/therapeutic procedures (excluding prior anticancer treatments).

6.7.1.1. Prior Anticancer Therapy

Prior antineoplastic therapy will be classified in three separate categories: medications, radiotherapy, and surgery.

Prior antineoplastic medications will be summarized by anatomical therapeutic classification (ATC) code level 4 and Preferred Term and presented by treatment arm using the using FAS, SS and DDS, for each of the following groups of settings: Adjuvant/Neo-adjuvant; Metastatic/Locally Advanced; Unknown.

In addition, summaries of characteristics will be provided by type of Prior Antineoplastic Therapies and presented by treatment arm using FAS, SS and DDS, including:

- Number and percentage of participants with prior antineoplastic medication: overall, by setting, by best response in metastatic setting, by type of treatment in metastatic setting, by type of treatment in locally advanced setting

- Number and percentage of participants with prior antineoplastic radiotherapy: overall, by intent, by setting, by best response

- Number and percentage of participants with prior antineoplastic surgery: overall.

Listings will be provided for each category separately (medication, radiotherapy, surgery) and presented in FAS showing at least data detailed below

- Prior antineoplastic therapy – medications:
    - Participant reported with medications (Yes/No)
    - Drug Name
    - ATC code level 4 and Preferred Drug Name
    - Setting
    - Best overall response
    - Type of Treatment
    - Time from medication start to progression (months), using the “Date of Relapse/Progression” as collected in the medication form
    - Time from medication end to randomization (OR first dose of study drug for SLI participants) (<1 month, 1-<6 months, 6-<12 months, >=12 months)
- Prior antineoplastic therapy - radiotherapy:
    - Participant reported with radiotherapy (Yes/No)
    - Location/Site
    - Treatment Intent (Palliative, Curative, Preventative)
    - Best Overall Response
    - Treatment Setting
    - Time from radiotherapy end to randomization (OR first dose of study drug for SLI participants) (<1 month, 1-<6 months, 6-<12 months, >=12 months)

- Prior antineoplastic therapy – surgery (excluding biopsies):
    - Participant reported with prior surgeries (Yes/No)
    - Site
    - Location
    - SOC and PT
    - Result of Surgery
    - Time from surgery to randomization (OR first dose of study drug for SLI participants) (<1 month, 1-<6 months, 6-<12 months, >=12 months)

Incomplete dates will be handled as described in Appendix 6.9 (section 6.9.4).

Note:

- Time from medication start to progression will be derived if start of medication and PD dates ("Date of Relapse/Progression" for the given prior antineoplastic regimen) are not completely missing. In that case, participant will be counted in the missing category. It will be computed in months as [date of PD - start date of prior medication + 1] / 30.4375.

- Time from medication end to randomization (OR first dose of study drug for SLI participants) (respectively time from radiotherapy end, time from surgery) will be derived if last intake date of the medication (respectively radiotherapy end date, surgery date) is not completely missing. In that case, participant will be counted in the missing category. It will be computed in months as [date of randomization (OR first dose of study drug for SLI participants) - end date of medication (resp. end date of radiotherapy, date of surgery) + 1] / 30.4375.

6.7.1.2. Prior medications (excluding antineoplastic therapy)

Prior medications (excluding antineoplastic therapy) will be summarized by ATC code level 4, Preferred Drug Name and presented by treatment arm using the FAS, the SS and the DDS. Listing will be provided on FAS (same listing as the concomitant medications, see Section 6.7.2.2).

Incomplete dates will be handled as described in Appendix 6.9 (section 6.9.2).

6.7.1.3. Prior Therapeutic / Diagnostic Procedures

Prior Therapeutic / Diagnostic Procedures (excluding prior anticancer radiotherapies and surgeries) will be summarized by SOC and PT with the number and percentage presented by treatment arm for the FAS, the SS and the DDS. A by-participant listing will be provided on FAS (same listing as the concomitant procedures, see Section 6.7.2.3).

6.7.2. Concomitant therapies

Concomitant therapies are defined as any therapies starting on or after the start of study treatment or medications starting prior to the start of study treatment and continuing after the start of study treatment.

6.7.2.1. Concomitant Anticancer Therapies

Concomitant antineoplastic radiotherapies as collected in the dedicated CRF form will be listed separately.

Other concomitant antineoplastic therapies will be analysed as part of the other general therapies (medications see Section 6.7.2.2, or procedures see Section 6.7.2.3).

6.7.2.2. Concomitant medications

Concomitant Medications will be summarized by ATC code level 4, Preferred Drug Name and presented by treatment arm using the FAS, the SS and the DDS.

Listing will be provided on FAS, gathering both prior and concomitant medications. Indeed, the listing will include all medications reported in the CRF form “Prior and Concomitant Medication”, regardless start and end date. Medication status will be defined as either prior or concomitant.

Concomitant Medications to treat Covid-19 (SARS-CoV-2) will also be summarized by ATC code level 4, Preferred Drug Name and presented by treatment arm using the FAS.

A summary table of Covid-19 (SARS-CoV-2) vaccines will be provided on SS, with:

- Number and proportion of participants vaccinated. Participant analyzed as Vaccinated are participants having received at least one dose of a COVID-19 (SARS-CoV-2) vaccine within 28 days before the first study drug administration, up to study discontinuation.
- Number and proportion of participants vaccinated with 1, 2, 3 doses. Participant analyzed as Vaccinated with X dose(s) means that he received X dose(s) within 28 days before the 1st study drug administration, up to study discontinuation.
- Summary of Covid-19 (SARS-CoV-2) vaccines by ATC code level 4 and Preferred Drug Name. The selection of Covid-19 (SARS-CoV-2) vaccines will be identified using WHODrug SDG “Vaccines for COVID-19” (narrow scope, ATC J07BX).

Types of Covid-19 (SARS-CoV-2) vaccine are based on available vaccines in China:

- 3 of Inactivated vaccines (Vero cell) (2 vaccines from Sinopharm, 1 vaccine from Sinovac) - 2 doses required: Will be identified based on SDG subset “Inactivated vaccines for COVID-19”
- 1 of Recombinant vaccines (adenovirus type 5 vector) (CanSinoBio) - 1 dose required: Will be identified based on drug name confirmed by coder

- 1 of Recombinant vaccines (CHO cell) (Institute of Microbiology, Chinese Academy of Sciences) - 3 doses required: Will be identified based on drug name confirmed by coder

All prior and concomitant medications to treat Covid-19 (SARS-CoV-2) will be listed, including the Covid-19 (SARS-CoV-2) vaccines with their type of vaccine specified, when appropriate.

Incomplete dates will be handled as described in Appendix 6.9 (section 6.9.2).

6.7.2.3. Concomitant Therapeutic / Diagnostic Procedures

Concomitant Therapeutic and Diagnostic Procedures will be summarized by SOC and PT with the number and percentage presented by treatment arm for the FAS, the SS and the DDS.

A listing will be provided on the FAS.

Diagnostic Procedures to treat Covid-19 (SARS-CoV-2) will also be summarized by SOC and PT with the number and percentage presented by treatment arm for the FAS.

A by-participant listing will be provided on FAS, gathering both prior and concomitant procedures. Indeed, the listing will include all procedures reported in the CRF form “Medical and Surgical Procedures”, regardless start and end date. Procedure status will be defined as either prior or concomitant.

Same listing will also be provided for concomitant diagnostic procedures to treat Covid-19 (SARS-CoV-2).

6.7.3. Follow-up Medications

6.7.3.1. Subsequent Anticancer Therapy

The FAS will be used for listing of anticancer therapies initiated after discontinuation of study treatment. Subsequent anticancer therapies include anticancer systemic therapies (medications), anticancer radiation or anticancer surgery.

All summaries will be presented by treatment arm for the FAS.

Subsequent anticancer systemic therapies will be coded using the WHO Drug Dictionary and presented in a data listing based on medications reported in CRF form “Anti-Neoplastic Therapy Since Study Treatment Discontinuation – Medication”. This can include treatment with the study intervention, when reported for participants continuing the study treatment beyond progression (as allowed by protocol).

Subsequent anticancer systemic therapies will be summarized by ATC code level 4 and Preferred Drug Name, by treatment group.

Subsequent anticancer medications will also be summarized by Therapy category as collected (MONOTHERAPY CYTOTOXIC CHEMOTHERAPY; COMBINATION CYTOTOXIC CHEMOTHERAPY; IMMUNOTHERAPY; TARGETED THERAPY; COMBINATION CYTOTOXIC CHEMOTHERAPY + IMMUNOTHERAPY; COMBINATION CYTOTOXIC CHEMOTHERAPY + TARGETED THERAPY; OTHER, SPECIFY), and therapy sub-categories according to Fortrea and Pierre Fabre clinician assessment after the review of data collected in the CRF. This summary will be provided separately for the first treatment line and the second treatment line after Study Treatment Discontinuation.

The number of subsequent treatment lines will be described according to Fortrea and Pierre Fabre clinician assessment after the review of data collected in the CRF, the subsequent treatment line is derived and included in the analysis.

Subsequent radiation therapy will be summarized according to treatment intent and best overall response, by treatment group.

Subsequent surgery will be summarized according to location/site, result of surgery, treatment intent, by treatment group.

There will be no imputation for missing end dates.

6.8. Appendix 8: Protocol Deviations

All protocol deviations will be evaluated at least prior to database lock or any snapshot.

Protocol deviations and additional reasons of exclusion from analysis sets (such as missing primary criterion or insufficient exposure to the treatment due to premature withdrawal), will be described in the dedicated protocol deviations list based on TransCelerate PDAP template, separately from the statistical analysis plan and classified as important/non-important deviation along with its impact on the analysis sets, before the first participant inclusion or prior to any snapshot.

An important protocol deviation with a major impact on the primary criterion analyses for statistical analysis is any event likely to bias significantly the interpretation of the primary efficacy criterion results, especially the primary analysis and therefore, leading to the exclusion of the corresponding participants from the PPS.

A listing of all deviations and a listing of additional reasons of exclusion for statistical analysis will be provided for all randomized/assigned to treatment participants, including the type of impact (Important/Non important) as validated during the data review/ Validation Committee meeting.

Number and percentage of participants with at least one deviation related to Covid-19 (SARS-CoV-2) will also be described by treatment group and category of deviations and detailed in a participant data listing. Otherwise specified, protocol deviations related to Covid-19 (SARS-CoV-2) will be identified when prefix “COVID-19” is used in the deviation description.

Among the FAS population, the number and percentage of participants excluded from analysis sets will be tabulated by treatment group and type of reason of exclusion from each population (SS, DDS, ES, PPS, PK Set). The number and percentage of participants with at least one important deviation with a major impact on the primary criterion analyses will be tabulated by treatment group and type of deviations on the FAS.

6.9. Appendix 9: Missing Values – Missing Visits

For partially missing dates for efficacy endpoints, the following imputation rules will be used: if the day of the month is missing, but month and year are known (UN-MMM-YYYY), it will be imputed by the 1st of the month (01-MMM-YYYY). If this implementation rule produces a date before start of treatment, then the date of start of treatment is used.

6.9.1. Adverse Events

In case of missing information for AEs, this will be treated as described in section 5.6.2.

For partial AE start dates or completely missing AE start dates, the date imputation will be based on the temporal relation between the partial date and the treatment start date as detailed in the Table below. No imputation will be performed when the year is missing for the start date.

Table 18: AE Start Date Imputation Example Scenarios

Partial 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
ddOCT2001	20OCT2001	Uncertain	20OCT2001
ddNOV2001	20OCT2001	After	01NOV2001

*If participant is randomized but not treated, randomization date is used instead of treatment start date.

For missing or partially missing stop dates of AE, 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 as detailed in the table below.

Table 19: End Date Imputation Example Scenarios

Partial end date	Minimum (Last contact date, 30-day follow-up date)	Ongoing	Imputed Date
Missing	20OCT2001	Yes	20OCT2001
ddmmm2000	20OCT2001	No	31DEC2000
ddmmm2002	20OCT2001	No	31DEC2002
ddmmm2001	20OCT2001	No	20OCT2001
ddmmm2001	20OCT2001	Yes	31DEC2001

Partial end date	Minimum (Last contact date, 30-day follow-up date)	Ongoing	Imputed Date
ddSEP2001	20OCT2001	No	30SEP2001
ddOCT2001	20OCT2001	No	20OCT2001
ddOCT2001	20OCT2001	Yes	31OCT2001

6.9.2. Prior and Concomitant Medication

In case of missing or partially missing start dates (resp. stop date) for prior/concomitant medication, same rules as those described in Table 18 (resp. Table 19) will be applied.

6.9.3. Initial Diagnosis Date

In case of missing or partially missing dates for initial diagnosis, the following rules will be applied:

- If initial diagnosis date=../mmm/yyyy (= missing day), then it will be substituted by **01/mmm/yyyy**
- If initial diagnosis date=.../.../yyyy (= missing day and month), then it will be substituted by **01/ /yyyy**
- If initial diagnosis date=.../.../…. (= completely missing), then it won't be substituted.

Same applies for missing or partially missing dates for first metastasis.

6.9.4. Antineoplastic therapies

Prior Antineoplastic therapies

Incomplete start dates will be handled by applying the same imputation rules as described for AE/concomitant medications start date (Sections 6.9.1 and 6.9.2) except:

- If only day is missing, and month and year match that of the treatment start date, impute as date of randomization (OR first dose of study drug for SLI participants) -1;

- If both day and month are missing, and the year matches that of the treatment start date, then impute as date of randomization (OR first dose of study drug for SLI participants) -1.

Completely missing start date is imputed as date of randomization (OR first dose of study drug for SLI participants) - 1;

Incomplete end dates and dates of progression on the prior antineoplastic therapies will be handled using the following rule:

- If only day is missing, imputed date = min (date of randomization (OR first dose of study drug for SLI participants) -1, last day of the month);

- If both month and day are missing, imputed date = min (randomization date (OR first dose of study drug for SLI participants) -1, 31DEC).

- If the end date or the date of progression is not missing and the imputed start date is after the end date or after the date of progression, use the min (end date, date of progression) as the imputed start date.

Post-treatment therapies:

Incomplete subsequent antineoplastic therapy start dates will be imputed using the following rules:

- Imputed date = max (End of Study Intervention Date + 1, first day of the month), if day is missing
- Imputed date = max (End of Study Intervention Date + 1, 01JAN), if day and month are missing
- Imputed date = End of Study Intervention Date + 1, if the date is completely missing

End of Study Intervention Date is defined according to section 5.1.4.

There will be no imputation for missing end dates.

6.9.5. Death 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

If the date of death is fully missing, death will be imputed to the last date the participant was known to be alive or at their last contact date, whatever is earlier.

6.10. Appendix 10: Reporting conventions

6.10.1. P-values presentation

P-values will be rounded to four decimal places. P-values will be reported in tables as:

if p-value < 0.0001: reported as <0.0001
if 0.0001 ≤ p-value < 0.001: reported as <0.001
if p-value > 0.999: reported as >0.999.

6.11. Appendix 11: Response Evaluation Criteria in Solid Tumors Version 1.1

Taken from Eisenhauer 2009

6.11.1. Methods of Measurement

All measurements should be recorded in metric notation, using calipers if clinically assessed. All baseline evaluations should be performed as close as possible to the start of study intervention and never more than 28 days before the first dose.

The same method of assessment and the same technique should be used to characterize each identified and reported lesion at baseline and during the treatment phase and follow-up phase (if applicable). Imaging based evaluation should always be done rather than clinical examination unless the lesion(s) being followed cannot be imaged but are assessable by clinical examination.

Clinical lesions: Clinical lesions will only be considered measurable when they are superficial and ≥10 mm diameter as assessed using calipers (e.g. skin nodules). For the case of skin lesions, documentation by color photography including a ruler to estimate the size of the lesion is suggested. As noted above, when lesions can be evaluated by both clinical examination and imaging, imaging evaluation should be undertaken since it is more objective and may also be reviewed at the end of the study.

Chest X-ray: Chest CT is preferred over chest X-ray, particularly when progression is an important endpoint, since CT is more sensitive than X-ray, particularly in identifying new lesions. However, lesions on chest X-ray may be considered measurable if they are clearly defined and surrounded by aerated lung.

CT, MRI: CT is the best currently available and reproducible method to measure lesions selected for response assessment. This guideline has defined measurability of lesions on CT scan based on the assumption that CT slice thickness is 5 mm or less. When CT scans have slice thickness greater than 5 mm, the minimum size for a measurable lesion should be twice the slice thickness. MRI is also acceptable in certain situations (e.g. for body scans).

Ultrasound: Ultrasound is not useful in assessment of lesion size and should not be used as a method of measurement. Ultrasound examinations cannot be reproduced in their entirety for independent review at a later date and because they are operator dependent, it cannot be guaranteed that the same technique and measurements will be taken from one assessment to the next. If new lesions are identified by ultrasound in the course of the study, confirmation by CT or MRI is advised. If there is concern about radiation exposure at CT, MRI may be used instead of CT in selected instances.

Endoscopy, laparoscopy: The utilization of these techniques for objective tumor evaluation is not advised. However, they can be useful to confirm complete pathological response when biopsies

are obtained or to determine relapse in studies where recurrence following complete response or surgical resection is an endpoint.

Tumor markers: Tumor markers *alone* cannot be used to assess *objective* tumor response. If markers are initially above the ULN, however, they must normalize for a participant to be considered in CR. Because tumor markers are disease specific, instructions for their measurement should be incorporated into protocols on a disease specific basis.

Cytology, histology: These techniques can be used to differentiate between PR and CR in rare cases if required by protocol (for example, residual lesions in tumor types such as germ cell tumors, where known residual benign tumors can remain). When effusions are known to be a potential adverse effect of treatment (e.g. with certain taxane compounds of angiogenesis inhibitors), the cytological confirmation of the neoplastic origin of any effusion that appears or worsens during treatment can be considered if the measurable tumor has met criteria for response or stable disease in order to differentiate between response (or stable disease) and progressive disease.

6.11.2. Measurability of Tumour at Baseline

6.11.2.1. Definitions

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

- Measurable

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

  – 10 mm by CT scan (CT scan slice thickness no greater than 5 mm).

  – 10 mm caliper measurement by clinical examination (lesions which cannot be accurately measured with calipers should be recorded as non-measurable).

  – 20 mm by chest X-ray.

  Malignant lymph nodes: To be considered pathologically enlarged *and* measurable, a lymph node must be ≥15 mm in *short* axis when assessed by CT scan (CT scan slice thickness recommended to be no greater than 5 mm). At baseline and in follow-up, only the *short* axis will be measured and followed.

- Non-measurable

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

6.11.2.2. Special Considerations Regarding Lesion Measurability

Bone lesions, cystic lesions and lesions previously treated with local therapy require particular comment:

Bone lesions:

- Bone scan, PET scan or plain films are not considered adequate imaging techniques to measure bone lesions. However, these techniques can be used to confirm the presence or disappearance of bone lesions.
- Lytic bone lesions or mixed lytic-blastic lesions, with *identifiable soft tissue components*, that can be evaluated by cross sectional imaging techniques such as CT or MRI can be considered measurable lesions if the *soft tissue component* meets the definition of measurability described above.
- Blastic bone lesions are non-measurable.

Cystic lesions:

- Lesions that meet the criteria for radiographically defined simple cysts should not be considered as malignant lesions (neither measurable nor non-measurable) since they are, by definition, simple cysts.
- 'Cystic lesions' thought to represent cystic metastases can be considered as measurable lesions, if they meet the definition of measurability described above. However, if non-cystic lesions are present in the same participant, these are preferred for selection as target lesions.

Lesions with prior local treatment:

- Tumor lesions situated in a previously irradiated area, or in an area subjected to other loco-regional therapy, are usually not considered measurable unless there has been demonstrated progression in the lesion. Study protocols should detail the conditions under which such lesions would be considered measurable.

6.11.3. Tumor Response Evaluation

To assess objective response or future progression, it is necessary to estimate the *overall tumor burden at baseline* and use this as a comparator for subsequent measurements. Measurable disease is defined by the presence of at least one measurable lesion (see Appendix 6.11.2.1).

6.11.3.1. Baseline Documentation of 'Target' and 'Non-target' Lesions

When more than one measurable lesion is present at baseline all lesions up to a maximum of five lesions total (and a maximum of two lesions per organ) representative of all involved organs should be identified as *target lesions* and will be recorded and measured at baseline (this means in

instances where participants have only one or two organ sites involved, a *maximum* of two and four lesions respectively will be recorded).

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

Lymph nodes merit special mention since they are normal anatomical structures which may be visible by imaging even if not involved by tumor. As previously noted, pathological nodes which are defined as measurable and may be identified as target lesions must meet the criterion of a short axis of ≥15 mm by CT scan. Only the *short* axis of these nodes will contribute to the baseline sum. The short axis of the node is the diameter normally used by radiologists to judge if a node is involved by solid tumor. Nodal size is normally reported as two dimensions in the plane in which the image is obtained (for CT scan this is almost always the axial plane; for MRI the plane of acquisition may be axial, sagittal or coronal). The smaller of these measures is the short axis.

For example, an abdominal node which is reported as being 20 mm × 30 mm has a short axis of 20 mm and qualifies as a malignant, measurable node. In this example, 20 mm should be recorded as the node measurement. All other pathological nodes (those with short axis ≥10 mm but <15 mm) should be considered non-target lesions. Nodes that have a short axis <10 mm are considered non-pathological and should not be recorded or followed.

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

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

6.11.3.2. Response Criteria

Tumour Response for Target and Non-target Lesions

- Evaluation of Target Lesions

This section provides the definitions of the criteria used to determine objective tumor response for target lesions.

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

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

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

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

Special notes on the assessment of target lesions:

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

Target lesions that become 'too small to measure'. While on study, all lesions (nodal and non-nodal) recorded at baseline should have their actual measurements recorded at each subsequent evaluation, even when very small (e.g. 2 mm). However, sometimes lesions or lymph nodes which are recorded as target lesions at baseline become so faint on CT scan that the radiologist may not feel comfortable assigning an exact measure and may report them as being 'too small to measure'. When this occurs, it is important that a value be recorded on the eCRF. If it is the opinion of the radiologist that the lesion has likely disappeared, the measurement should be recorded as 0 mm. If the lesion is believed to be present and is faintly seen but too small to measure, a default value of 5 mm should be assigned (Note: It is less likely that this rule will be used for lymph nodes since they usually have a definable size when normal and are frequently surrounded by fat such as in the retroperitoneum; however, if a lymph node is believed to be present and is faintly seen but

too small to measure, a default value of 5 mm should be assigned in this circumstance as well). This default value is derived from the 5 mm CT slice thickness (but should not be changed with varying CT slice thickness). The measurement of these lesions is potentially non-reproducible, therefore providing this default value will prevent false responses or progressions based upon measurement error. If the radiologist is able to provide an actual measurement, that should be recorded, even if below 5 mm.

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

- Evaluation of Non-target Lesions

This section provides the definitions of the criteria used to determine the tumor response for the group of non-target lesions. While some non-target lesions may actually be measurable, they need not be measured and instead should be assessed only *qualitatively* at the time points specified in the protocol.

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

Progressive Disease (PD): *Unequivocal progression* of existing non-target lesions. The appearance of one or more new lesions is also considered progression.

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

Special Notes on Assessment of Progression of Non-target Disease

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

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

When the participant has only non-measurable disease. This circumstance arises in some Phase III studies when it is not a criterion of study entry to have measurable disease. The same general concepts apply here as previously noted, however, in this instance there is no measurable disease assessment to factor into the interpretation of an increase in non-measurable disease burden. Because worsening in non-target disease cannot be easily

quantified, a useful test that can be applied when assessing participants for unequivocal progression is to consider if the increase in overall disease burden based on the change in non-measurable disease is comparable in magnitude to the increase that would be required to declare PD for measurable disease i.e. an increase in tumor burden representing an additional 73% increase in 'volume' (which is equivalent to a 20% increase in diameter in a measurable lesion). Examples include an increase in a pleural effusion from 'trace' to 'large', an increase in lymphangitic disease from localized to widespread or may be described in protocols as 'sufficient to require a change in therapy'. If 'unequivocal progression' is seen, the participant should be considered to have had overall PD at that point. While it would be ideal to have objective criteria to apply to non-measurable disease, the very nature of that disease makes it impossible to do so, therefore the increase must be substantial.

- New Lesions

The appearance of new malignant lesions denotes disease progression. There are no specific criteria for the identification of new radiographic lesions; however, the finding of a new lesion should be unequivocal i.e. not attributable to differences in scanning technique, change in imaging modality or findings thought to represent something other than tumor. This is particularly important when the participant's baseline lesions show PR or CR.

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

If a new lesion is equivocal, for example because of its small size, continued therapy and follow-up evaluation will clarify if it represents truly new disease. If repeat scans confirm there is definitely a new lesion, then progression should be declared using the date of the initial scan.

While FDG-PET response assessments need additional study, it is sometimes reasonable to incorporate the use of FDG-PET scanning to complement CT scanning in assessment of progression (particularly possible 'new' disease). New lesions on the basis of FDG-PET imagine can be identified according to the following algorithm:

- Negative FDG-PET at baseline, with a positive FDG-PET at follow-up is a sign of PD based on a new lesion. A 'positive' FDG-PET scan lesion is one which is FDG avid with an uptake greater than twice that of the surrounding tissue on the attenuation corrected image.
- No FDG-PET at baseline and a positive FDG-PET at follow-up:
  - If the positive FDG-PET at follow-up corresponds to a new site of disease confirmed by CT, this is PD.
  - If the positive FDG-PET at follow-up is not confirmed as a new site of disease on CT, additional follow-up CT scans are needed to determine if there is truly

progression occurring at that time (if so, the date of PD will be the date of the initial abnormal FDG-PET scan).

– If the positive FDG-PET at follow-up corresponds to a pre-existing site of disease on CT that is not progressing on the basis of the anatomic images, this is not PD.

Timepoint Response

It is assumed that at each protocol specified timepoint, a response assessment occurs. A summary of the overall response status calculation at each timepoint for participant who have measurable disease at baseline is shown in Table 20.

Table 20: Timepoint Response: Participants with Target (± Non-target) Disease

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

Abbreviations: CR = complete response; NE=not evaluable; PD=progressive disease; PR=partial response; SD=stable disease.

When no imaging/measurement is done at all at a particular timepoint, the participant is not evaluable at that timepoint. If only a subset of lesion measurements are made at an assessment, usually the case is also considered not evaluable at that timepoint, unless convincing argument can be made that the contribution of the individual missing lesion(s) would not change the assigned timepoint response. This would be most likely to happen in the case of PD.

Evaluation of Best Overall Response

The BOR is the best response across all timepoints recorded from the start of the study intervention until the end of intervention (taking into account any requirement for confirmation). On occasion a response may not be documented until after the end of therapy so protocols should be clear if postintervention assessments are to be considered in determination of BOR. Protocols must specify how any new therapy introduced before progression will affect best response designation. The participant's BOR assignment will depend on the findings of both target and non-target disease and will also take into consideration the appearance of new lesions. Furthermore, depending on the nature of the study and the protocol requirements, it may also require confirmatory

measurement. Specifically, in non-randomized trials where response is the primary endpoint, confirmation of PR or CR is needed to deem either one the 'best overall response'.

The BOR is determined once all the data for the participant are known.

Special Notes on Response Assessment

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

Participants with a global deterioration of health status requiring discontinuation of study intervention without objective evidence of disease progression at that time should be reported as 'symptomatic deterioration'. Every effort should be made to document objective progression even after discontinuation of study intervention. Symptomatic deterioration is *not* a descriptor of an objective response: it is a reason for stopping study intervention. The objective response status of such participants is to be determined by evaluation of target and non-target disease.

Conditions that define 'early progression, early death and non-evaluability' are study specific and should be clearly described in each protocol (depending on treatment duration, treatment periodicity).

In some circumstances it may be difficult to distinguish residual disease from normal tissue. When the evaluation of CR depends upon this determination, it is recommended that the residual lesion be investigated (fine needle aspirate/biopsy) before assigning a status of CR. FDG-PET may be used to upgrade a response to a CR in a manner similar to a biopsy in cases where a residual radiographic abnormality is thought to represent fibrosis or scarring. The use of FDG-PET in this circumstance should be prospectively described in the protocol and supported by disease specific medical literature for the indication. However, it must be acknowledged that both approaches may lead to false positive CR due to limitations of FDG-PET and biopsy resolution/sensitivity.

For equivocal findings of progression (e.g. very small and uncertain new lesions; cystic changes or necrosis in existing lesion), treatment may continue until the next scheduled assessment. If at the next scheduled assessment, progression is confirmed, the date of progression should be the earlier date when progression was suspected.

6.12. Appendix 12: Questionnaire and Scoring instructions for EORTC QLQ-C30 v3.0

EORTC QLQ-C30 (version 3)

We are interested in some things about you and your health. Please answer all of the questions yourself by circling the number that best applies to you. There are no "right" or "wrong" answers. The information that you provide will remain strictly confidential.

Please fill in your initials: ⌊ | | | ⌋
Your birthdate (Day, Month, Year): ⌊ | | | | | | | | ⌋
Today's date (Day, Month, Year): 31 ⌊ | | | | | | | | ⌋

	Not at All	A Little	Quite a Bit	Very Much
1. Do you have any trouble doing strenuous activities, like carrying a heavy shopping bag or a suitcase?	1	2	3	4
2. Do you have any trouble taking a long walk?	1	2	3	4
3. Do you have any trouble taking a short walk outside of the house?	1	2	3	4
4. Do you need to stay in bed or a chair during the day?	1	2	3	4
5. Do you need help with eating, dressing, washing yourself or using the toilet?	1	2	3	4

During the past week:	Not at All	A Little	Quite a Bit	Very Much
6. Were you limited in doing either your work or other daily activities?	1	2	3	4
7. Were you limited in pursuing your hobbies or other leisure time activities?	1	2	3	4
8. Were you short of breath?	1	2	3	4
9. Have you had pain?	1	2	3	4
10. Did you need to rest?	1	2	3	4
11. Have you had trouble sleeping?	1	2	3	4
12. Have you felt weak?	1	2	3	4
13. Have you lacked appetite?	1	2	3	4
14. Have you felt nauseated?	1	2	3	4
15. Have you vomited?	1	2	3	4
16. Have you been constipated?	1	2	3	4

During the past week:	**Not at All**	**A Little**	**Quite a Bit**	**Very Much**
17. Have you had diarrhea?	1	2	3	4
18. Were you tired?	1	2	3	4
19. Did pain interfere with your daily activities?	1	2	3	4
20. Have you had difficulty in concentrating on things, like reading a newspaper or watching television?	1	2	3	4
21. Did you feel tense?	1	2	3	4
22. Did you worry?	1	2	3	4
23. Did you feel irritable?	1	2	3	4
24. Did you feel depressed?	1	2	3	4
25. Have you had difficulty remembering things?	1	2	3	4
26. Has your physical condition or medical treatment interfered with your family life?	1	2	3	4
27. Has your physical condition or medical treatment interfered with your social activities?	1	2	3	4
28. Has your physical condition or medical treatment caused you financial difficulties?	1	2	3	4

For the following questions please circle the number between 1 and 7 that best applies to you

29. How would you rate your overall health during the past week?

1	2	3	4	5	6	7
Very poor						Excellent

30. How would you rate your overall quality of life during the past week?

1	2	3	4	5	6	7
Very poor						Excellent

The European Organisation for Research and Treatment of Cancer (EORTC) score questionnaire (QLQ-C30) will be used for quality of life (QoL) evaluation.

Following the EORTC recommendations, fifteen scales can be derived from the initial 30 questions:

- A global health status/QoL scale,
- Five functional, scales (physical, role, cognitive, emotional and social),
- Nine "symptoms" scales /items (nausea and vomiting, pain, fatigue, dyspnoea, insomnia, appetite loss, constipation, diarrhoea and financial difficulties).

Of note, for functional scales, a higher value reflects a better level of function, but for symptoms scales /items a higher value reflects worse symptoms; moreover high score for the global health status represents a high QoL.
Each scale in the questionnaire will be scored (0 to 100) according to the EORTC recommendations in the EORTC QLQ-C30 Scoring Manual [Fayers 2001]. The scoring method is summarized below. In this summary, Qi refers to the i^{th} question on the EORTC QLQ-C30.

Scoring and scale dimension:

	Number of items (range*)	**Item number**
Global health status / QoL		
Global health status/QoL	2 (6)	29,30
Functional scales		
Physical functioning	5 (3)	1 to 5
Role functioning	2 (3)	6, 7
Emotional functioning	4 (3)	21 to 24
Cognitive functioning	2 (3)	20, 25
Social functioning	2 (3)	26, 27
Symptom scales / items		
Fatigue	3 (3)	10, 12, 18
Nausea and vomiting	2 (3)	14, 15
Pain	2 (3)	9, 19
Dyspnoea	1 (3)	8
Insomnia	1 (3)	11
Appetite loss	1 (3)	13
Constipation	1 (3)	16
Diarrhoea	1 (3)	17
Financial difficulties	1 (3)	28

* Item range is the difference between the possible maximum and the minimum response to individual items; most items take values from 1 to 4, giving range = 3.

For all scales, the RawScore, RS, is the mean of the component items:

$$RawScore\ RS = (I_1 + I_2 + ... + I_n)/n$$

Then for **Functional scales**:

$$\text{Score} = \left\{1 - \frac{(RS - 1)}{range}\right\} \times 100$$

and for **Symptom scales / items** and **Global health status / QoL**:

$$Score = \left\{\frac{(RS - 1)}{range}\right\} \times 100$$

Missing value (item) consideration for scoring:

- The scale scores will only be calculated if at least half of the items from the scale have been answered. Otherwise, no score will be calculated and the scale score will be set to missing.
- For single-item measures, the score will be missing if the question is not answered.

6.13. Appendix 13: EQ-5D-5L

Source: EuroQol Research Foundation. *EQ-5D-5L User Guide, 2019.*

Latest version available from: https://euroqol.org/publications/user-guides

Figure 1: EQ-5D-5L (UK English sample version)

Under each heading, please tick the **ONE** box that best describes your health **TODAY**

MOBILITY

I have no problems in walking about	❑
I have slight problems in walking about	❑
I have moderate problems in walking about	❑
I have severe problems in walking about	❑
I am unable to walk about	❑

SELF-CARE

I have no problems washing or dressing myself	❑
I have slight problems washing or dressing myself	❑
I have moderate problems washing or dressing myself	❑
I have severe problems washing or dressing myself	❑
I am unable to wash or dress myself	❑

USUAL ACTIVITIES *(e.g. work, study, housework, family or leisure activities)*

I have no problems doing my usual activities	❑
I have slight problems doing my usual activities	❑
I have moderate problems doing my usual activities	❑
I have severe problems doing my usual activities	❑
I am unable to do my usual activities	❑

PAIN / DISCOMFORT

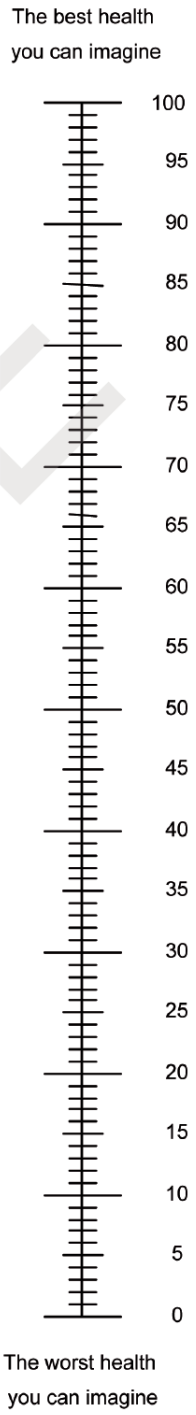
I have no pain or discomfort	❑
I have slight pain or discomfort	❑
I have moderate pain or discomfort	❑
I have severe pain or discomfort	❑
I have extreme pain or discomfort	❑

ANXIETY / DEPRESSION

I am not anxious or depressed	❑
I am slightly anxious or depressed	❑
I am moderately anxious or depressed	❑
I am severely anxious or depressed	❑
I am extremely anxious or depressed	❑

- We would like to know how good or bad your health is **TODAY**.
- This scale is numbered from **0** to **100**.
- **100** means the best health you can imagine.
  0 means the worst health you can imagine.
- Mark an **X** on the scale to indicate how your health is **TODAY**.
- Now, please write the number you marked on the scale in the box below.

YOUR HEALTH TODAY = ☐



6.14. Appendix 14: FACT-C

FACT-C (Version 4)

Below is a list of statements that other people with your illness have said are important. **Please circle or mark one number per line to indicate your response as it applies to the past 7 days.**

	PHYSICAL WELL-BEING	**Not at all**	**A little bit**	**Some-what**	**Quite a bit**	**Very much**
GP1	I have a lack of energy	0	1	2	3	4
GP2	I have nausea	0	1	2	3	4
GP3	Because of my physical condition, I have trouble meeting the needs of my family	0	1	2	3	4
GP4	I have pain	0	1	2	3	4
GP5	I am bothered by side effects of treatment	0	1	2	3	4
GP6	I feel ill	0	1	2	3	4
GP7	I am forced to spend time in bed	0	1	2	3	4

	SOCIAL/FAMILY WELL-BEING	**Not at all**	**A little bit**	**Some-what**	**Quite a bit**	**Very much**
GS1	I feel close to my friends	0	1	2	3	4
GS2	I get emotional support from my family	0	1	2	3	4
GS3	I get support from my friends	0	1	2	3	4
GS4	My family has accepted my illness	0	1	2	3	4
GS5	I am satisfied with family communication about my illness	0	1	2	3	4
GS6	I feel close to my partner (or the person who is my main support)	0	1	2	3	4
Q1	*Regardless of your current level of sexual activity, please answer the following question. If you prefer not to answer it, please mark this box* ☐ *and go to the next section.*					
GS7	I am satisfied with my sex life	0	1	2	3	4

English (Universal)

16 November 2007
Page 1 of 3

FACT-C (Version 4)

Please circle or mark one number per line to indicate your response as it applies to the past 7 days.

	EMOTIONAL WELL-BEING	Not at all	A little bit	Some-what	Quite a bit	Very much
GE1	I feel sad	0	1	2	3	4
GE2	I am satisfied with how I am coping with my illness	0	1	2	3	4
GE3	I am losing hope in the fight against my illness	0	1	2	3	4
GE4	I feel nervous	0	1	2	3	4
GE5	I worry about dying	0	1	2	3	4
GE6	I worry that my condition will get worse	0	1	2	3	4

	FUNCTIONAL WELL-BEING	Not at all	A little bit	Some-what	Quite a bit	Very much
GF1	I am able to work (include work at home)	0	1	2	3	4
GF2	My work (include work at home) is fulfilling	0	1	2	3	4
GF3	I am able to enjoy life	0	1	2	3	4
GF4	I have accepted my illness	0	1	2	3	4
GF5	I am sleeping well	0	1	2	3	4
GF6	I am enjoying the things I usually do for fun	0	1	2	3	4
GF7	I am content with the quality of my life right now	0	1	2	3	4

FACT-C (Version 4)

Please circle or mark one number per line to indicate your response as it applies to the past 7 days.

	ADDITIONAL CONCERNS	Not at all	A little bit	Some-what	Quite a bit	Very much
C1	I have swelling or cramps in my stomach area	0	1	2	3	4
C2	I am losing weight	0	1	2	3	4
C3	I have control of my bowels	0	1	2	3	4
C4	I can digest my food well	0	1	2	3	4
C5	I have diarrhea (diarrhoea)	0	1	2	3	4
C6	I have a good appetite	0	1	2	3	4
C7	I like the appearance of my body	0	1	2	3	4
Q2	Do you have an ostomy appliance? (Mark one box)	☐ No		or	☐ Yes	
	If yes, please answer the next two items:					
C8	I am embarrassed by my ostomy appliance	0	1	2	3	4
C9	Caring for my ostomy appliance is difficult	0	1	2	3	4

FACT-C Scoring Guidelines (Version 4) – Page 1

Source: FACT-C scoring template 05.21.03

FACT-C Scoring Guidelines (Version 4) – Page 1

Instructions:* 1. Record answers in "item response" column. If missing, mark with an X
2. Perform reversals as indicated, and sum individual items to obtain a score.
3. Multiply the sum of the item scores by the number of items in the subscale, then divide by the number of items answered. This produces the subscale score.
4. Add subscale scores to derive total scores (TOI, FACT-G & FACT-C).
5. **The higher the score, the better the QOL.**

Subscale	Item Code	Reverse item?		Item response	Item Score
PHYSICAL WELL-BEING (PWB)	GP1	4	-	_____	=_____
	GP2	4	-	_____	=_____
	GP3	4	-	_____	=_____
	GP4	4	-	_____	=_____
Score range: 0-28	GP5	4	-	_____	=_____
	GP6	4	-	_____	=_____
	GP7	4	-	_____	=_____

Sum individual item scores: _____
Multiply by 7: _____
Divide by number of items answered: _____=**PWB subscale score**

SOCIAL/FAMILY WELL-BEING (SWB)	GS1	0	+	_____	=_____
	GS2	0	+	_____	=_____
	GS3	0	+	_____	=_____
	GS4	0	+	_____	=_____
Score range: 0-28	GS5	0	+	_____	=_____
	GS6	0	+	_____	=_____
	GS7	0	+	_____	=_____

Sum individual item scores: _____
Multiply by 7: _____
Divide by number of items answered: _____=**SWB subscale score**

EMOTIONAL WELL-BEING (EWB)	GE1	4	-	_____	=_____
	GE2	0	+	_____	=_____
	GE3	4	-	_____	=_____
	GE4	4	-	_____	=_____
Score range: 0-24	GE5	4	-	_____	=_____
	GE6	4	-	_____	=_____

Sum individual item scores: _____
Multiply by 6: _____
Divide by number of items answered: _____=**EWB subscale score**

FUNCTIONAL WELL-BEING (FWB)	GF1	0	+	_____	=_____
	GF2	0	+	_____	=_____
	GF3	0	+	_____	=_____
	GF4	0	+	_____	=_____
Score range: 0-28	GF5	0	+	_____	=_____
	GF6	0	+	_____	=_____
	GF7	0	+	_____	=_____

Sum individual item scores: _____
Multiply by 7: _____
Divide by number of items answered: _____=**FWB subscale score**

FACT-C Scoring Guidelines (Version 4) – Page 2

Subscale	Item Code	Reverse item?		Item response	Item Score
COLORECTAL	C1	4	-	_____	=_____
CANCER	C2	4	-	_____	=_____
SUBSCALE	C3	0	+	_____	=_____
(CCS)	C4	0	+	_____	=_____
Score range: 0-28	C5	4	-	_____	=_____
	C6	0	+	_____	=_____
	C7	0	+	_____	=_____
	C8	NOT CURRENTLY SCORED			
	C9	NOT CURRENTLY SCORED			

Sum individual item scores: _____
Multiply by 7: _____
Divide by number of items answered: _____**=CC Subscale score**

To derive a FACT-C Trial Outcome Index (TOI):
Score range: 0-84

_____ + _____ + _____ =_____**=FACT-C TOI**
(**PWB** score) (**FWB** score) (**CCS** score)

To Derive a FACT-G total score:
Score range: 0-108

_____ + _____ + _____ + _____=_____**=FACT-G Total score**
(**PWB** score) (**SWB** score) (**EWB** score) (**FWB** score)

To Derive a FACT-C total score:
Score range: 0-136

_____ + _____ + _____ + _____ + _____ =_____**=FACT-C Total score**
(**PWB** score) (**SWB** score) (**EWB** score) (**FWB** score) (**CCS** score)

*For guidelines on handling missing data and scoring options, please refer to the Administration and Scoring Guidelines in the manual or on-line at www.facit.org.

6.15. Appendix 15: Encorafenib Case Retrieval Strategy

Adverse Event of Special Interest	MedDRA Preferred Term (From MedDRA 26.0, this list could be updated during the study and before database lock, based on a later version of MedDRA)
Cutaneous non-squamous cell carcinoma	Basal cell carcinoma
Cutaneous non-squamous cell carcinoma	Carcinoma in situ of skin
Cutaneous non-squamous cell carcinoma	Dysplastic naevus syndrome
Cutaneous non-squamous cell carcinoma	Eccrine carcinoma
Cutaneous non-squamous cell carcinoma	Epidermal naevus syndrome
Cutaneous non-squamous cell carcinoma	Extramammary Paget's disease
Cutaneous non-squamous cell carcinoma	Hidradenocarcinoma
Cutaneous non-squamous cell carcinoma	Malignant sweat gland neoplasm
Cutaneous non-squamous cell carcinoma	Mastocytoma
Cutaneous non-squamous cell carcinoma	Neoplasm skin
Cutaneous non-squamous cell carcinoma	Neuroendocrine carcinoma of the skin
Cutaneous non-squamous cell carcinoma	Pilomatrix carcinoma
Cutaneous non-squamous cell carcinoma	Porocarcinoma
Cutaneous non-squamous cell carcinoma	Sebaceous carcinoma
Cutaneous non-squamous cell carcinoma	Skin angiosarcoma
Cutaneous non-squamous cell carcinoma	Skin cancer
Cutaneous non-squamous cell carcinoma	Skin cancer metastatic
Cutaneous non-squamous cell carcinoma	Skin neoplasm bleeding
Cutaneous non-squamous cell carcinoma	Trichoblastic carcinoma
Cutaneous squamous cell carcinoma	Atypical fibroxanthoma
Cutaneous squamous cell carcinoma	Basosquamous carcinoma of skin
Cutaneous squamous cell carcinoma	Bowen's disease
Cutaneous squamous cell carcinoma	Keratoacanthoma
Cutaneous squamous cell carcinoma	Marjolin's ulcer
Cutaneous squamous cell carcinoma	Skin squamous cell carcinoma metastatic
Cutaneous squamous cell carcinoma	Squamous cell carcinoma of skin
Cutaneous squamous cell carcinoma	Trichoblastic carcinoma
Melanomas	Acral lentiginous melanoma
Melanomas	Acral lentiginous melanoma stage I
Melanomas	Acral lentiginous melanoma stage II
Melanomas	Acral lentiginous melanoma stage III
Melanomas	Acral lentiginous melanoma stage IV
Melanomas	Desmoplastic melanoma
Melanomas	Lentigo maligna
Melanomas	Lentigo maligna recurrent
Melanomas	Lentigo maligna stage I
Melanomas	Lentigo maligna stage II
Melanomas	Lentigo maligna stage III
Melanomas	Lentigo maligna stage IV
Melanomas	Malignant blue naevus

Adverse Event of Special Interest	MedDRA Preferred Term (From MedDRA 26.0, this list could be updated during the study and before database lock, based on a later version of MedDRA)
Melanomas	Malignant melanoma
Melanomas	Malignant melanoma in situ
Melanomas	Malignant melanoma stage I
Melanomas	Malignant melanoma stage II
Melanomas	Malignant melanoma stage III
Melanomas	Malignant melanoma stage IV
Melanomas	Melanoma recurrent
Melanomas	Naevoid melanoma
Melanomas	Nodular melanoma
Melanomas	Superficial spreading melanoma stage I
Melanomas	Superficial spreading melanoma stage II
Melanomas	Superficial spreading melanoma stage III
Melanomas	Superficial spreading melanoma stage IV
Melanomas	Superficial spreading melanoma stage unspecified
Facial paresis	Brow ptosis
Facial paresis	Crocodile tears syndrome
Facial paresis	Facial nerve disorder
Facial paresis	Facial paralysis
Facial paresis	Facial paresis
Facial paresis	Facial spasm
Facial paresis	Hyperacusis
Facial paresis	Melkersson-Rosenthal syndrome
Facial paresis	Oculofacial paralysis
Facial paresis	VIIth nerve injury
Uveitis type events	Autoimmune uveitis
Uveitis type events	Bacterial iritis
Uveitis type events	Blau syndrome
Uveitis type events	Chemical burns of eye
Uveitis type events	Ciliary hyperaemia
Uveitis type events	Cogan's syndrome
Uveitis type events	Cyclitic membrane
Uveitis type events	Cyclitis
Uveitis type events	Diabetic uveitis
Uveitis type events	Heerfordt's syndrome
Uveitis type events	Immune recovery uveitis
Uveitis type events	Infectious iridocyclitis
Uveitis type events	Infective iritis
Uveitis type events	Iridocyclitis
Uveitis type events	Iritis
Uveitis type events	Sympathetic ophthalmia
Uveitis type events	Traumatic iritis
Uveitis type events	Tubulointerstitial nephritis and uveitis syndrome

Adverse Event of Special Interest	MedDRA Preferred Term (From MedDRA 26.0, this list could be updated during the study and before database lock, based on a later version of MedDRA)
Uveitis type events	Uveitis
Uveitis type events	Viral uveitis
Uveitis type events	keratouveitis
QT prolongation	Electrocardiogram QT interval abnormal
QT prolongation	Electrocardiogram QT prolonged
QT prolongation	Long QT syndrome
QT prolongation	Long QT syndrome congenital
QT prolongation	Torsade de pointes
QT prolongation	Ventricular tachycardia
QT prolongation	Cardiac arrest
QT prolongation	Cardiac death
QT prolongation	Cardiac fibrillation
QT prolongation	Cardio-respiratory arrest
QT prolongation	Electrocardiogram repolarisation abnormality
QT prolongation	Electrocardiogram U wave inversion
QT prolongation	Electrocardiogram U wave present
QT prolongation	Electrocardiogram U-wave abnormality
QT prolongation	Loss of consciousness
QT prolongation	Sudden cardiac death
QT prolongation	Sudden death
QT prolongation	Syncope
QT prolongation	Ventricular arrhythmia
QT prolongation	Ventricular fibrillation
QT prolongation	Ventricular flutter
QT prolongation	Ventricular tachyarrhythmia
Non-cutaneous malignancies with RAS mutation	H-RAS gene mutation
Non-cutaneous malignancies with RAS mutation	K-RAS gene mutation
Non-cutaneous malignancies with RAS mutation	N-RAS gene mutation

7. References

Aaronson NK, Ahmedzai S, Bergman B et al.

The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology.

J Natl Cancer Inst 1993;85(5):365-376

Brookmeyer R and Crowley J.

A confidence interval for the median survival time.

Biometrics 1982;38:29-41.

Clopper CJ, Pearson ES.

The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial.

Biometrika 1934;26(4):404-413.

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(8):753–762.

Eisenhauer EA, Therasse P, Bogaerts J.

New response evaluation criteria in solid tumours: revised RECIST guideline (Version 1.1).

Eur J Cancer 2009; 45(2):228-247.

Kalbfleisch JD and Prentice RL.

The statistical analysis of failure time data.

Hoboken: John Wiley& Sons, 2002.

Klein JP and Moeschberger ML.

Survival analysis: Techniques for censored and truncated data.

New York: Springer-Verlag, 1997.

Luo N, Liu G, Li M, Guan H, Jin X, Rand-Hendriksen K.

Estimating an EQ-5D-5L Value Set for China.

Value in Health. 2017 Apr;20(4):662-669. doi: 10.1016/j.jval.2016.11.016. Epub 2017 Feb 9

P.M. Fayers, N.K. Aaronson, K. Bjordal, M. Groenvold, D. Curran, A. Bottomley, et al.

EORTC QLQ-C30 Scoring Manual (3rd ed.),

EORTC, Brussels (2001)

Pharmaceuticals and Medical Devices Agency (PMDA)

Guidelines for basic principles on global clinical trials

PMDA Notification No.0928010, 28 September 2007)

Saridaki Z, Tzardi M, Sfakianaki M, Papadaki C, Voutsina A, et al.

BRAFV600E Mutation Analysis in Patients with Metastatic Colorectal Cancer (mCRC) in Daily Clinical Practice: Correlations with Clinical Characteristics, and Its Impact on Patients' Outcome.

PLoS ONE 8(12): e84604. doi:10.1371/journal.pone.0084604

Schemper M, Smith TL.

A note on quantifying follow-up in studies of failure time.

Control Clin Trials, 1996 Aug;17(4):343-6. doi: 10.1016/0197-2456(96)00075-x.

Ulivi et al.

Predictive role of multiple gene alterations in response to cetuximab in metastatic colorectal cancer: A single center study.

Journal of Translational Medicine 2012 10:87.