CL3-95005-007 - Statistical Analysis Plan | VV-TMF-103206 | 1.0

I.R.I.S.

INSTITUT DE RECHERCHES INTERNATIONALES SERVIER

Document title	**STATISTICAL ANALYSIS PLAN (SAP)**
Study official title	**An open-label, randomised, phase III study comparing trifluridine/tipiracil in combination with bevacizumab to trifluridine/tipiracil monotherapy in patients with refractory metastatic colorectal cancer**
Study brief title	**Phase III study of trifluridine/tipiracil in combination with bevacizumab vs trifluridine/tipiracil single agent in patients with refractory metastatic colorectal cancer**
Test drug code	**Trifluridine/tipiracil (also known as S 95005 or TAS-102)**
Indication(s)	**Refractory metastatic colorectal cancer**
Development phase	**Phase III**
Protocol code	**CL3-95005-007**
EudraCT Number	**2020-001976-14**
Universal Trial Number	**Not applicable**
Sponsor	**Institut de Recherches Internationales Servier (I.R.I.S.)**
Date of the document	**02 August 2022**
Version of the document	**Version 2.0**
Author(s) of the document	PPD

CONFIDENTIAL

CL3-95005-007 - Statistical Analysis Plan | VV-TMF-103206 | 1.0

Follow up of versions

Version	Release date (dd/mm/yyyy)	Key modifications(*)	Impact
V1.0	06/11/2020	Not applicable.	
V2.0	02/08/2022	- Section 2.3: update of subgroup categories - Section 3.1: addition of methodology for data Handling according to a survival cut-off date - Section 3.2.1: addition of the description of changes and accommodations du to COVID-19 - Section 3.2.4.1: update of age categories and removal of shift table of geographic region eCRF*IWRS as automatic region completion in both sources - Sections 3.2.4.2 & 4.2.1.3.2: addition of 'Time from first metastasis diagnosis to randomization' definition and addition of precision on staging classifications - Section 3.2.4.5: updates of previous therapies descriptive analyses - Section 3.3.1.1: update of classes for treatment duration and number of cycles - Section 3.3.2: addition of G-CSF treatment intakes analysis - Section 3.4.1.3.2: update of covariate categories - Section 3.4.1.3.1: addition of a sensitivity analysis (supportive 3) for secondary estimand based on PFS (FDA IND CCI [redacted] - Section 3.5.2: addition of TEAE related to disease progression and leading to death analyses - Section 3.6: addition of figures displaying mean changes in QoL from baseline by scheduled assessment time point - Section 3.6.1: update of categories of changes from baseline	

CL3-95005-007 - Statistical Analysis Plan | VV-TMF-103206 | 1.0

		- Section 4.1.5: update of first and last trifluridine/tipiracil intake dates by first and last IMP intake dates - Section 4.2.1.3.4: modification of the definition of a previous metastatic drug treatment - Sections 6 & 6.1: update of the treatment period and after treatment period definitions considering both IMPs and a time window of 30 days - Section 6.3: Total calcium (biochemistry) moved to non-gradable parameter (not defined as per CTCAE) - Sections 7.1 & 7.2: update of scheduled assessment time point definitions to take into account information from ePRO and eCRF; addition of withdrawal questionnaire definition	

(*) Main changes as compared to the statistical analyses planned in the protocol for the first SAP signed version (1.0). Main changes from the previous signed version for the other SAP signed version(s).

CL3-95005-007 - Statistical Analysis Plan | VV-TMF-103206 | 1.0

Table of contents

CL3-95005-007 - Statistical Analysis Plan | VV-TMF-103206 | 1.0

List of tables

List of figures

List of abbreviations

%	:	Percentage
AE	:	Adverse Event
ALT	:	ALanine (Amino)Transferase
ALP	:	ALkaline Phosphatase
AST	:	ASpartate (Amino)Transferase
BID	:	bis in die (twice a day)
BOR	:	Best Overall Response
bpm	:	beats per minute (heart rate unit)
BSA	:	Body Surface Area
CHMP	:	Committee for Medicinal Products for Human Use
CI	:	Confidence Interval
CL	:	Total CLearance
cm	:	Centimetre
CMH	:	Cochran-Mantel-Haenszel
CR	:	Complete Response
e-CRF	:	electronic-Case Report Form
CTCAE	:	Common Terminology Criteria for Adverse Events
DBP	:	Diastolic Blood Pressure
DCR	:	Disease Control Rate
DI	:	Dose Intensity
EAE	:	Emergent Adverse Event
ECG	:	ElectroCardioGram
ECOG PS	:	Eastern Cooperative Oncology Group Performance Status
EMA	:	European Medicines Agency
EORTC	:	European Organisation for Research and Treatment of Cancer
FAS	:	Full Analysis Set
g	:	Gram
G/L	:	Giga (10^9) per litre
GGT	:	Gamma-Glutamyl Transferase (Gamma-Glutamyl Transpeptidase)
h	:	Hour
HR	:	Heart Rate
HR	:	Hazard Ratio
ICH	:	International Conference on Harmonization

i.e.	:	id est
IME	:	Important Medical Event
IMP	:	Investigational Medicinal Product
INR	:	International Normalized Ratio
I.R.I.S.	:	Institut de Recherches Internationales Servier
IS	:	Included Set
IU	:	International Unit
IV	:	IntraVenous (route)
IWRS	:	Interactive Web Response System
kg	:	Kilogram
KRAS	:	V-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog
L	:	Litre
LDH	:	Lactate DeHydrogenase
LLN	:	Lower Limit of Normal laboratory reference range
m	:	Metre
mCRD	:	Metastatic colorectal cancer
MedDRA	:	Medical Dictionary for Regulatory Activities
MSI	:	Microsatellite Instability
mg	:	Milligram
min	:	Minute
Min	:	Minimum
mL	:	Millilitre
mm	:	Millimetre
mmol	:	Millimole
MMRM	:	Mixed-effects Model for Repeated Measures
msec	:	millisecond
NA	:	Not Applicable
NCI	:	National Cancer Institute
ng	:	Nanogram
NLR	:	Neutrophils lymphocytes ratio
ORR	:	Overall Response Rate
OS	:	Overall Survival
PD	:	Progressive Disease
PFS	:	Progression Free Survival

PR	:	Partial Response
PT	:	Preferred Term
PV	:	PharmacoVigilance
QLQ	:	Quality of Life Questionary
QoL	:	Quality of Life
QTc	:	QT interval corrected for heart rate
QTcB	:	Bazett's corrected QT interval
QTcF	:	Fridericia's corrected QT interval
RBC	:	Red Blood Cells
RDI	:	Relative Dose Intensity
s	:	Second
SAE	:	Serious Adverse Event
SAP	:	Statistical Analysis Plan
SBP	:	Systolic Blood Pressure
SD	:	Standard Deviation
SDTM	:	Study Data Tabulation Model
SD	:	Stable Disease
SEAE	:	Serious Emergent Adverse Event
SOC	:	System Organ Class
SS	:	Safety Set
TLG	:	Tables, Listings and Graphs
TR	:	Tumour Response
TUDD	:	Time until definitive deterioration
ULN	:	Upper Limit of Normal laboratory reference range
WBC	:	White Blood Cells
WHO	:	World Health Organization

CL3-95005-007 - Statistical Analysis Plan | VV-TMF-103206 | 1.0

1. INTRODUCTION

This Statistical Analysis Plan (SAP) details the planned analyses to be performed, in accordance with the main characteristics of the study protocol.
The templates for Tables, Listings and Graphs (TLG) are described in a separate document.

1.1. Study objectives

The primary objective is to demonstrate the superiority of trifluridine/tipiracil in combination with bevacizumab over trifluridine/tipiracil monotherapy in terms of Overall Survival (OS) in patients with refractory metastatic colorectal cancer (mCRC).

Secondary objectives are to estimate the effect of trifluridine/tipiracil in combination with bevacizumab versus trifluridine/tipiracil monotherapy in terms of Progression Free Survival (PFS), Overall Response Rate (ORR), and Disease Control Rate (DCR) in patients with refractory mCRC. Other secondary objectives are to compare the safety and tolerance, and the impact on Quality of Life (QoL) of trifluridine/tipiracil in combination with bevacizumab to trifluridine/tipiracil monotherapy in patients with refractory mCRC.

1.2. Study design

This present study is designed as an international, open-label, controlled two-arm, randomised phase III comparison study evaluating the efficacy and safety of trifluridine/tipiracil in combination with bevacizumab versus trifluridine/tipiracil monotherapy in patients with refractory mCRC.

1.2.1. Study plan

The analysis will be done after 331 events are reported. In order to observe this number of events, 490 patients will be randomised (1:1) to receive trifluridine/tipiracil in combination with bevacizumab (experimental arm) or trifluridine/tipiracil monotherapy (control arm).

Considering the anticipated rapid enrolment and event accumulation in this population, there is no planned interim analysis for efficacy or futility.

Randomisation will take place once the consented patient has completed all the necessary baseline procedures and is deemed eligible for study entry. The stratification factors will be:
- Geographic region (North America, European Union, Rest of the World).
- Time since diagnosis of first metastasis (< 18 months, ≥ 18 months).
- RAS status (wild-type, mutant).

The study will be divided into the following periods for each patient:
- **Screening visit** (up to 28 days prior to randomisation): to obtain informed consent.
- **Screening period/Inclusion**: to check the eligibility of the patient to be included and randomised in the study.
- **Randomisation**: patients will be randomly assigned to one of the two treatment arms:
  - Trifluridine/tipiracil in combination with bevacizumab: trifluridine/tipiracil (starting dose at 35 mg/m2/dose) will be administered orally BID for 5 days on/2 days off, for 2 weeks, followed by a 14-day rest, with bevacizumab (5 mg/kg) administered IV every 2 weeks (day 1 and day 15). This treatment cycle will be repeated every 4 weeks.

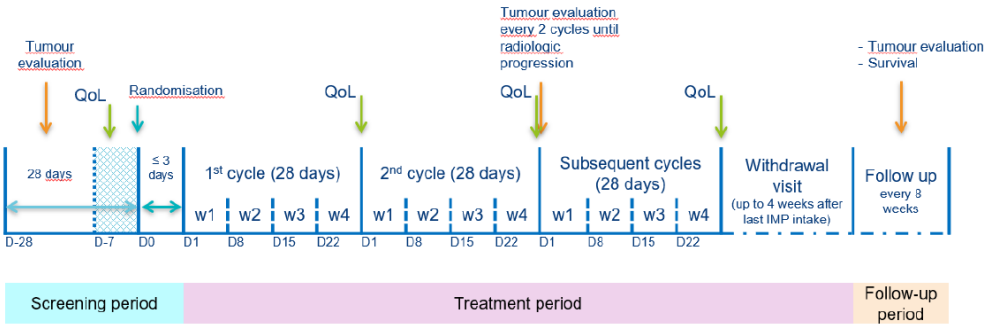
- Trifluridine/tipiracil: trifluridine/tipiracil (starting dose at 35 mg/m2/dose) will be administered orally BID for 5 days on/2 days off, for 2 weeks, followed by a 14-day rest. This treatment cycle will be repeated every 4 weeks.

- **Treatment period:** randomised patients should receive the first dose of study treatments (day 1 of cycle 1) within 3 days after randomisation.
  Patients will be treated until they meet a discontinuation criterion. Patients will be considered on treatment as long as the patients continue with trifluridine/tipiracil. Bevacizumab monotherapy is not allowed.
- **Withdrawal visit:** within 4 weeks following the date of IMPs withdrawal and prior to the start of a new anticancer therapy.
- **Follow-up period:** after the withdrawal visit, a follow-up will be done every 8 weeks until death or end of study.

The end of study will occur 19 months after the first IMP intake of the last patient randomised and is defined as the date of the last follow-up of the last patient (including a contact phone) or the date of the last contact attempt if the last patient is declared lost to follow-up.

The study plan is shown in Figure (1.2.1) 1.

Figure (1.2.1) 1 - Study plan



1.2.2. Randomisation

Patients will be randomly assigned in a (1:1) ratio to receive trifluridine/tipiracil plus bevacizumab (experimental arm) or trifluridine/tipiracil as monotherapy (control arm). Randomisation will take place once the consented patient has completed all the necessary baseline procedures and deemed eligible for study entry. Treatment assignment will be done centrally via an Interactive Web Response System (IWRS) stratified by:

- Geographic region (North America, European Union, Rest of the World).
- Time since diagnosis of first metastasis (< 18 months, ≥ 18 months).
- RAS status (wild type, mutant).

Due to the difference between the two arms in term of bevacizumab injection (1 vs 0), this study must be performed in open-label.

1.3. Determination of sample size

Calculations are done with EAST 6.4 software.

A maximum of 331 events (deaths for any cause) will be required for the primary analysis to detect a hazard ratio of 0.70 with 90% power using a log-rank test at one-sided cumulative 2.5% level of significance. Based on the data from the RECOURSE study (Mayer, 2015) the median duration of OS in the control group is expected to be around 7.1 months. A hazard ratio of 0.70 translates into 3 months increase of the median OS in the experimental arm (10.1 months) compared to the control arm. Based on the assumption that enrolment will continue for approximately 12 months, and that about 5%/year of the subjects will drop out, a total of 490 patients randomised in a 1:1 ratio will be needed to observe the 331th OS events approximately 9 months after the last patient randomisation.

2. ANALYSIS SETS AND SUBGROUPS / TREATMENT GROUPS

2.1. Analysis sets

Screened set: All patients screened.

Included set (IS): All included patients.

Full analysis set (FAS): In accordance with the intention-to-treat principle and the Section 5.2.1 of ICH E9 guideline, all patients to whom a therapeutic unit was randomly assigned using IWRS. Patients in the FAS will be analysed in the arm they were assigned by randomisation.

Safety set (SS): All patients having taken at least one dose of trifluridine/tipiracil. Patients will be analyzed according to the treatment actually received.

2.2. Treatment groups

- Trifluridine / tipiracil (starting dose at 35 mg/m²/dose) BID + bevacizumab (5 mg/kg IV).
- Trifluridine / tipiracil (starting dose at 35 mg/m²/dose) BID.

2.3. Subgroups

Depending on the sample size, the following subgroups will be examined: region IWRS (North America, European Union, Rest of the World), time since diagnosis of 1st metastasis IWRS (< 18, ≥ 18 months), RAS IWRS (wild, mutant), location of primary disease (right, left), ECOG performance status (0, ≥1), gender (female, male), age (< 65, ≥ 65 years), prior surgical resection (yes, no), number of metastatic sites (1-2, ≥ 3), neutrophils to lymphocytes ratio (NLR < 3, NLR ≥ 3), number of prior metastatic drug regimens (1, ≥2) , BRAF (wild, mutant) and MSI status (MSI-H, MSS/MSI-L), prior bevacizumab (yes, no), subsequent regorafenib (yes, no).

Subgroup analyses are detailed in Section 3.4.1.3.1.

3. STATISTICAL METHODS

3.1. General considerations

The following descriptive statistics will be provided depending on the nature of considered data:

Qualitative data: number of observed values, number and percentage of patients per class.

Quantitative data: number of observed values, mean and standard deviation, median, first and third quartiles, minimum and maximum.

Descriptive summary statistics (n, mean (SD), median, min and max) will be provided for variables measured on a continuous scale.

The frequency distribution (n, %) will be provided for variables measured on a nominal scale.

Handling of data according to a survival cut-off date

A survival cut-off date will be established for the analyses of the primary estimand based on the OS, the additional estimand based on the OS, the secondary estimands based on the PFS, the time from randomisation to worsening ECOG performance status of ≥ 2 and the time until definitive deterioration of 10 points in EORTC QLQ-C30 questionnaire subscales.

The survival cut-off date will be defined based on the actual date of the occurrence of the 331st OS event.

Patients having an OS event after the survival cut-off date will be censored at the survival cut-off date.

Patients having a PFS event after the survival cut-off date will be censored at the last evaluable tumour assessment date prior to the survival cut-off date.

Patients having a worsening ECOG performance status of ≥ 2 event after the survival cut-off date will be censored at the last ECOG PS assessment date prior to the survival cut-off date.

Patients having a definitive deterioration of 10 points in EORTC QLQ-C30 questionnaire subscales event after the survival cut-off date will be censored at the last QoL assessment date prior to the survival cut-off date.

Clopper-Pearson Confidence interval (CI)

To provide CI for estimates of ORR and DCR, Clopper-Pearson CI will be used. The exact or Clopper-Pearson confidence limits for the binomial proportion are constructed by inverting the equal-tailed test based on the binomial distribution. This method is attributed to Clopper and Pearson (1934).

The Clopper-Pearson interval can be written as follows for n1 = 1, 2… n-1:

$$\left(\left(1+\frac{n-n1+1}{n1\,F\left(1-\alpha/2,2n1,2(n-n1+1)\right)}\right)^{-1},\left(1+\frac{n-n1}{(n1+1)F\left(\alpha/2,2(n1+1),2(n-n1)\right)}\right)^{-1}\right)$$

Where F(α,b,c) is the αth percentile of the F distribution with b and c degrees of freedom.

Where $z^2_{\alpha/2}$ is the 100(1-α/2) th percentile of the standard normal distribution, α is the confidence level, $\hat{p} = \frac{n1}{n}$ is the proportion of observations and n is the size of sample.

Kaplan Meier survival analysis

Kaplan Meier survival analysis (Kaplan and Meier, 1958) will be used for time dependent variables. The objective is to describe time-to-event variable, without covariates, taking into account censored observations (for which the event was not experienced during the observation period). Event probabilities depend only on time. All subjects are assumed to behave similarly and computed survival functions are assumed to describe all subjects. Censored and uncensored observations behave the same.

The survival function *S* (t) is the probability that a patient survives (has no event) at least up to and including time t.

S (t) = Pr (T > t) where *S*(t) denotes the survival distribution function and T is the time to the event of a selected patient.

The median time to event is the time t so that P (T > t) = 0.5

The Kaplan-Meier method is based on conditional probabilities calculations and enables to compute non-parametric estimates of the survival function.

The survival function is estimated at each time-point as:

$$\frac{\text{The number of patients with no Event on that time}}{\text{The total amount of patients who can have the event}}$$

Taking into account censored observations.

The Kaplan-Meier estimate of *S* (t) is

$$\hat{S}(t) = \prod_{t_{(i)} \le t} \frac{n_{(i-)} - d_{(i)}}{n_{(i-)}}$$

With:

- $t_{(i)}$ is the i^{th} ordered event time.
- $d_{(i)}$ is the number of events at $t_{(i)}$.
- $n_{(i-)}$ is the number of observations without event an instant before $t_{(i)}$.

It is a non-parametric analysis; there is no hypothesis to check on distribution.

In case of the censored and uncensored observations are different, results may be biased. In the same way, if covariates other than time are thought to be important in determining duration to outcome, results reported may not show important differences in groups formed by the covariates.

Cox regression models can then be used to study the impact of covariates.

Cox proportional hazards model

The cox model (Cox, 1972) explores the relationship and the effect of several risk factors on survival.

It estimates and tests the hazard ratio between different groups. In the cox regression model, explanatory variables could be continuous or categorical.

The cox regression model does not make any parametric assumption about the survival probability distribution of each group, however it does assume a proportional hazard between two groups.

The proportional hazard model can be expressed as:

$$h(t) = h_0(t) \exp(\beta X)$$

Where X is the explanatory variables matrix, β is the regression coefficients matrix and $h_0(t)$is the baseline hazard function.
By dividing both sides of the above equation by $h_0(t)$ and taking logarithms, we obtain:

$$\ln\left(\frac{h(t)}{h0(t)}\right) = \beta X$$

We call h(t) / $h_0(t)$ the hazard ratio.
Because cox regression model assumes the proportional hazard between different groups, it's necessary to check this assumption before concluding the estimations. If the proportional hazard assumption does not fit the data, other models should be considered.

Kalbfleisch and Prentice confidence interval for S (t)
Kalbfleisch and Prentice (2002) got around the problem of having confidence intervals of > 1 or < 0. The transformation is done by adopting:

The estimated variance of log(-log($\hat{S}(t)$)) is:
$\tau 2(t) = \sigma 2[\hat{S}(t)] / [\hat{S}(t)\log(\hat{S}(t))]^2$
The 100(1-α)% confidence interval for S(t) is given by :

$$[\hat{S}(t)]^{\exp(z\propto/2\sigma(t))} \leq S(t) \leq [\hat{S}(t)]^{\exp(-z\propto/2\sigma(t))}$$

Restricted mean survival time (RMST)
The RMST (Irwin, 1949) is defined as the area under the curve of the survival function S(t) up to a time τ $(<\infty)$:

$$\mu_\tau = \int_0^\tau S(t)dt\,,$$

where S(t) is the survival function of a time-to-event variable of interest.

A natural estimator for μ_τ is:

$$\hat{\mu}_\tau = \int_0^\tau \hat{S}(t)dt\,,$$

where $\hat{S}(t)$ is the KM estimator for S(t).

τ will be the minimum of the largest observed event time in each of the two groups.

In addition, to take into account the stratification factor used at randomisation (region (North America, European Union, Rest of the World), time since diagnosis of 1st metastasis (< 18, ≥ 18 months), RAS status (wild, mutant)), an ANCOVA type adjusted analysis proposed by Tian et al. will be performed:
Let Y be the restricted mean survival time, and let Z be the treatment indicator. Also, let X denote a q-dimensional baseline covariate vector. Tian's method consider the following regression model:

$$E(Y|Z,X) = \alpha + \beta Z + \gamma X$$

where (α,β,γ) is a (q+2)-dimension unknown parameter vector. Tian's method (Tian, 2014) utilizes an inverse probability censoring weighting technique to handle censored observations.

3.2. Disposition and baseline characteristics

Description of disposition of patients and baseline characteristics will be performed by treatment arm and overall.
Specific definitions on disposition of patients and baseline characteristics are provided in 4 Appendix A.

3.2.1. Disposition of patients

Disposition of patients will be described in the Screened Set before and at inclusion and in the FAS after the inclusion.

Changes and accommodations due to COVID-19 will also be described.

3.2.2. Protocol deviations

All important protocol deviations before or at inclusion, as well as after inclusion, will be described in the FAS, by category of deviations based on ICH E3 guideline and ICH E3 Q&A.
The following subset of important protocol deviations will also be described:

- Has histologically confirmed unresectable adenocarcinoma of the colon or rectum.
- Has received a maximum of 2 prior chemotherapy regimens for the treatment of advanced colorectal cancer and had demonstrated progressive disease or intolerance to their last regimen.
- Has measurable or non-measurable metastatic lesion(s) as defined by Response Evaluation Criteria in Solid Tumours (RECIST) version 1.1.
- Has an Eastern Cooperative Oncology Group (ECOG) performance status ≤ 1.
- More than 2 prior chemotherapy regimens for the treatment of advanced colorectal cancer.
- Has previously received trifluridine/tipiracil.
- Does not currently received or has not received anticancer therapies with 4 weeks prior to randomisation.

A first analysis will be made considering all important protocol deviations and a second analysis considering only the subset of important protocol deviations.

3.2.3. Study population

The number of patients in each study population and the reasons for exclusion will be summarized.

Moreover, a listing of patients' stratification factors and a listing of patients with discrepancies between IWRS and eCRF stratification factors will be provided in the IS.

3.2.4. Demographic data and other baseline characteristics

3.2.4.1. Demographic characteristics

The following demographic and baseline variables will be summarized for patients in the FAS:

- Age (years).
- Age (< 65 years, ≥ 65 years).
- Age (< 70 years, ≥ 70 years).
- Age (< 75 years, ≥ 75 years).
- Age (Adults (18-64 years, from 65 to 74 years, from 75 to 84 years, 85 years and over)).

- Gender (Female, Male).
- Ethnic origin (White, American Indian or Alaska native, Asian, Black or African American, Native Hawaiian/ Other Pacific Islander, Other).
- Geographic region (North America, European Union, Rest of the World).
- Country.
- Pregnancy test.

3.2.4.2. History of cancer

Nature and duration of cancer will be summarized in the FAS according to the following variables:

- Primary diagnosis in classes (Colon, Rectal, Other) (4 Appendix A).
- Primary tumour site (Right colon, Left colon, Rectum, Transverse colon, Other) (4 Appendix A).
- Primary tumour localisation (Right, Left) (4 Appendix A).
- Stage at diagnosis according to AJCC, Dukes or Astler-Coller modified Dukes' classification.
- Stage at diagnosis in classes (I-II, III, IV).
- Current tumour status: in relapse (Yes, No).
- Current tumour status: metastasis (Yes, No).
- Site of metastasis (lymph node, lung, liver, bone, brain, skin, peritoneal, soft tissue, other).
- Number of metastatic sites.
- Number of metastatic sites in classes (1-2, ≥ 3).
- Disease duration (years).
- Disease duration in classes (≤ 1,]1, 2],]2, 4], > 4 years).
- Time from first metastasis diagnosis to randomization (months) documented in the e-CRF.
- Time from first metastasis diagnosis to randomization in classes (< 18, ≥ 18 months) documented in the e-CRF.
- Time from first metastasis diagnosis to randomization in classes (< 18, ≥ 18 months) documented in the IWRS:
  - Progression free interval (months).
  - Progression free interval in classes (≤ 1 (immediate),]1;6] (early),]6 to 12] (intermediate), > 12 months (late)).
  - Treatment free interval (months).
  - Treatment free interval in classes (≤1 (immediate),]1;6] (early),]6 to 12] (intermediate), > 12 months (late)).
  - RAS mutational status performed (Yes, No).
  - Site of the tumour biopsy (primary tissue, metastatic (lymph node, lung metastasis, liver metastasis, other metastasis)).
  - RAS status documented in the e-CRF (wild type, mutant type, NE).
  - RAS status documented in the IWRS (wild type, mutant type).
  - KRAS status (wild type, mutant type).
  - KRAS mutant status:
    - Mutation type Codon 12 (Yes, No, Not done, Unknown):
      - If Mutation type Codon 12: Gly12Asp (G12D), Gly12Val (G12V), Gly12Cys (G12C), Gly12Ser (G12S), Gly12Ala (G12A), Gly12Arg (G12R), Other.

- Mutation type Codon 13 (Yes, No, Not done, Unknown):
  - If Mutation type Codon 13: Gly13Asp (G13D), Other.
- Mutation type Codon 59 (Yes, No, Not done, Unknown).
- Mutation type Codon 61 (Yes, No, Not done, Unknown).
- Mutation type Codon 117 (Yes, No, Not done, Unknown).
- Mutation type Codon 146 (Yes, No, Not done, Unknown).
- Mutation other type (Yes, No).

- NRAS Status (wild type, mutant type).
- NRAS mutant status:
  - Mutation type Codon 12 (Yes, No, Not done, Unknown):
    - If Mutation type Codon 12: Gly12Asp (G12D), Gly12Cys (G12C), Gly12Ser (G12S), Other.
  - Mutation type Codon 13 (Yes, No, Not done, Unknown):
    - If Mutation type Codon 13 : Gly13Val (G13V), Gly13Arg (G13R), Other.
  - Mutation type Codon 59 (Yes, No, Not done, Unknown).
  - Mutation type Codon 61 (Yes, No, Not done, Unknown).
  - Mutation type Codon 117 (Yes, No, Not done, Unknown).
  - Mutation other type (Yes, No).
- BRAF mutational status determined (Yes, No).
- BRAF status (wild type, mutant type, unknown/not interpretable):
  - If BRAF mutant type (Val600Glu (V600E), Unknown, Other)).
- MMR or MSI performed (Yes, No).
- MMR/MSI status (MSI-H/Mismatch repair deficient, MSS/MSI-L/Mismatch repair proficient, Unknown/Not interpretable).

A shift table crossing RAS status as documented in the e-CRF versus BRAF status will be provided.

A shift table crossing RAS status as documented in the IWRS versus RAS status in the e-CRF will be provided. The same analysis will be provided for the time from diagnosis to first metastasis (< 18, ≥ 18 months) as documented in the IWRS versus in the e-CRF.

A listing of the history of colorectal cancer will also be provided. It will include the following information: Disease duration, time from first metastasis diagnosis to randomization, primary diagnosis, primary tumour site and metastatic sites.

3.2.4.3. Signs and symptoms related to colorectal cancer

Number and percent of patients with existing signs and symptoms related to mCRC at baseline will be summarized by SOC and PT according to the MedDRA dictionary (latest available version at the analysis date) for patients in the FAS. The same analysis will be duplicated according the NCI-CTCAE grade.

3.2.4.4. Medical and surgical history not related to colorectal cancer

Medical and surgical history not related to colorectal cancer will be summarized by SOC and PT according to the MedDRA dictionary (latest available version at the analysis date) for patients in the FAS.

3.2.4.5. Previous therapies related to colorectal cancer

Previous drug treatments will be described by pharmacological class, pharmacological sub-class, therapeutic class and preferred name (WHO-DD classification, latest available version at the analysis date). Previous surgery procedures related to colorectal cancer will be summarized by SOC and PT according to the MedDRA dictionary (latest available version at the analysis date). All previous therapies for the colorectal cancer will be summarized according to the following variables for patients in the FAS:

- Previous surgery (Yes, No).
- Previous palliative surgery (Yes, No).
- Previous curative surgery (Yes, No).
- Previous radiotherapy (Yes, No).
- Previous curative radiotherapy (Yes, No).
- Previous palliative radiotherapy (Yes, No).
- Previous drug treatments (Yes, No).
- Combination of previous therapies.
- Previous neo-adjuvant drug treatment (Yes, No).
- Previous adjuvant drug treatment (Yes, No).
- Previous adjuvant and neo-adjuvant drug treatment (Yes, No).
- Previous adjuvant or neo-adjuvant drug treatment (Yes, No).
- Previous metastatic drug treatment (Yes, No).
- Previous metastatic fluoropyrimidine (Yes, No).
- Previous metastatic irinotecan (Yes, No).
- Previous metastatic oxaliplatin (Yes, No).
- Previous metastatic anti-VEGF monoclonal Anti Body (Yes, No).
- Previous metastatic anti-EGFR monoclonal Anti Body (Yes, No).
- Previous metastatic anti-EGFR monoclonal Anti Body (Yes, No) in RAS wild type patients only as documented in the e-CRF.
- Number of prior drug regimens.
- Number of prior drug regimens in classes (1, 2, 3, ≥4).
- Number of prior adjuvant or neoadjuvant drug regimens.
- Number of prior metastatic drug regimens.
- Number of prior metastatic drug regimens in classes (1, 2, ≥3).
- Reason for treatment discontinuation of the last drug regimen prior to randomisation (Toxicity, Progressive disease, Other).

3.2.4.6. Vital signs at baseline

Vital signs and clinical examination will be described in the FAS, in terms of quantitative values for Supine Systolic Blood Pressure - SBP (mmHg), Supine Diastolic Blood Pressure - DBP (mmHg), Heart Rate - HR (bpm), Weight (Kg), Temperature (°C), BSA (m^2) and by classes:

- SBP (< 90, [90, 140[, ≥140 mmHg).
- DBP (< 60, [60, 90[, ≥ 90 mmHg).
- HR (< 60, [60, 100[, ≥ 100 mmHg).
- ECOG (0, 1, 2, 3, 4).

- BSA (< 1.07, [1.07,1.23[, [1.23, 1.38[, [1.38, 1.53[, [1.53, 1.69[, [1.69, 1.84[, [1.84, 1.99[, [1.99, 2.15[, [2.15, 2.30[, ≥ 2.30).

3.2.4.7. ECG at baseline

The frequency distribution (n, %) of patients with an ECG performed as well as heart rate (b.p.m) (< 60, [60, 100[,≥ 100), uncorrected QT, Fridericia's corrected QT, Bazett's corrected QT and RR intervals will be described in terms of quantitative values and in classes, considering thresholds defined in ICH E14 (*i.e.*, ≤ 450,] 450; 480],] 480; 500] and > 500 ms) in the FAS. Both Fridericia's correction ($QTcF = QT/\sqrt[2]{RR}$) and Bazett's correction ($QTcB = QT/\sqrt[3]{RR}$) will be used for QTc interval.

3.2.4.8. Clinical laboratory at baseline

All laboratory data will be analysed using NCI-CTCAE grade criteria.
Clinical laboratory parameters will be summarized according to the following classes:

- For highest and/or lowest gradable parameters at baseline: (Grade 0, Grade 1, Grade 2, Grade 3, Grade 4).
- For non-gradable parameters and gradable parameters without baseline, except for creatinine clearance: (< LLN, [LLN, ULN], > ULN).
- For urinalysis parameter: proteinuria (Absence, Trace, Presence (+, More than one +)).
- Neutrophils-lymphocyte ratio (NLR) (< 3, ≥ 3).
- Neutrophils-lymphocyte ratio (NLR) (< 5, ≥ 5).

Creatinine clearance will be categorised in 4 levels of impairment and analysed by classes (severe, moderate, mild, normal):

- Normal renal function (CLcr ≥ 90 mL/min).
- Mild renal impairment (CLcr 60-89 mL/min).
- Moderate renal impairment (CLcr 30-59 mL/min).
- Severe (< 30mL/mn).

3.3. Treatments

3.3.1. Extent of exposure and treatment compliance

3.3.1.1. Trifluridine/tipiracil

Extent of exposure and treatment compliance of trifluridine/tipiracil will be described in the SS by patient for each treatment arm including:

- Treatment duration (months).
- Treatment duration in classes (≤ 2,]2; 4],]4-6],]6-8],]8; 10], > 10 months).
- Follow-up duration (months) with median based on reverse Kaplan Meier method.
- Number of cycles per patient.
- Number of cycles per patient (1, 2, …, > 10 cycles).
- Number of cycles per patient (< 2, [2-4],]4-6],]6-8],]8-10], > 10 cycles).
- Cumulative dose (mg/m^2).
- Dose intensity (DI) (mg/m^2/week).
- Relative Dose Intensity (RDI) (%).
- RDI in classes (≤ 60,] 60-80],] 80-100],] 100- 110], > 110%).
- Patients with at least one dose reduction (Yes, No).

- Number of dose reduction per patient.
- Number of dose reduction per patient (1, 2, 3, > 3 reductions).
- Number of patients by actual dose level on day 1 of each cycle.
- Patients with at least one dispensation postponed (Yes, No).
- Number of dispensation postponed per patient.
- Number of dispensation postponed per patient (1, 2, 3, > 3 dispensations postponed).
- Patients with at least one unplanned treatment interruption (Yes, No).
- Number of unplanned treatment interruption per patient.
- Number of unplanned treatment interruption in classes (1, 2, 3, > 3 unplanned treatment interruptions).

Individual listings by cycle will be provided including: cumulative dose, planned dose intensity, DI, RDI, total number of tablets given, estimated and returned (15 or 20 mg for trifluridine/tipiracil), cycle duration (days), duration of unplanned treatment interruption (days), reason for interruption, dispensation postponed (Yes, No), reason for dispensation postponed (medical reason, non-medical reason) and dose reduction (Yes, No).

Individual listing per patient will be provided including: number of cycles, cumulative dose, planned dose intensity, DI, RDI, treatment duration, number of cycles with unplanned treatment interruption, number of cycles with dispensation postponed and number of dose reductions.

3.3.1.2. Associated agent: bevacizumab

Number of infusion per patients, patients without partial dose administered, patients with at least one missed intake, number of cycles with at least one missed intake, cumulative dose (mg/kg), DI (mg/kg/week), RDI (%) and RDI in classes (≤ 60,] 60-80],] 80-100],] 100- 110], > 110%) will be described by patient in the SS.

Individual listing by cycle will be provided including: number of intake, reason for miss-intake, cumulative dose, planned dose intensity, DI, RDI, cycle duration (days), cycle with full dose administered (Yes, No) and reason for incomplete volume infusion (medical reason, non-medical reason).

Individual listing per patient will be provided including: number of cycles, number of intakes, cumulative dose, planned dose intensity, DI, RDI, treatment duration, number of cycles with at least one partial dose administered and number of cycles with at least one missed intake.

3.3.2. Concomitant treatments and new anti-cancer therapies

Drug treatments will be described by pharmacological class, pharmacological sub-class, therapeutic class and preferred name (WHO-DD classification, latest available version at the analysis date). Post-study surgeries and radiotherapy procedures related to colorectal cancer will be summarized by SOC and PT according to the MedDRA dictionary (latest available version at the analysis date). All concomitant treatments and new anti-cancer therapies will be summarized according to the following variables:

- Concomitant treatment (Yes, No).
- Received new anti-cancer therapies (Yes, No).
- Received new surgery (Yes, No).
- Received new radiotherapy (Yes, No).

Concomitant treatments taken at inclusion, those taken before, during and after the treatment period, as well as new anticancer therapies (as defined in efficacy part) will be described in the SS and in the FAS.

For concomitant treatments, individual listing will be provided including pharmacological class, pharmacological sub-class, therapeutic class and preferred name, medications starting and ending date, administration schedule and dose unit administered. The listing will include medications 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.

The number and percentage of patients receiving:
- At least one concomitant G-CSF treatment.
- At least one concomitant G-CSF treatment for primary prophylaxis.
- At least one concomitant G-CSF treatment for secondary prophylaxis.
- At least one concomitant G-CSF treatment for therapeutic use.

will be described in the SS for each treatment arm during the treatment period.

3.4. Efficacy analysis

Efficacy analyses will be carried out in the FAS by treatment arm.
Tumour response is evaluated according to the "New Response Evaluation Criteria in Solid Tumours: Revised RECIST guideline" (version 1.1) (Eisenhauer et al, 2009)
Strata will be based on IWRS data, unless otherwise specified.
Brookmeyer and Crowley confidence interval methodology will be applied for median and Kalbfleisch and Prentice confidence interval methodology for survival probabilities.

Primary objective and associated estimands of interest

The primary objective is to demonstrate the superiority of trifluridine/tipiracil in combination with bevacizumab over trifluridine/tipiracil monotherapy in terms of OS in patients with refractory mCRC.

Table (3.4) 1 - Summary of estimands of interest associated to primary objective

Estimand	Variable	Population level summary	Handling IEs	Treatment condition	Population
Primary estimand: effect of the randomised treatments on the survival duration in all subjects regardless of whether or not intercurrent events occur	OS	HR	Treatment policy strategy	trifluridine/tipiracil + bevacizumab potentially followed by further anti-cancer therapy vs trifluridine/tipiracil monotherapy potentially followed by further anti-cancer therapy	Full Analysis Set (FAS)
Additional estimand: effect of the randomised treatments on the survival duration in all subjects before patients receive further anti-cancer therapy	OS	HR	While on treatment strategy	trifluridine/tipiracil + bevacizumab vs trifluridine/tipiracil monotherapy	Full Analysis Set (FAS)

Intercurrent event

During the study the following intercurrent events could occur:

- Administration of further anti-cancer therapy.
- Treatment discontinuation.
- Treatment switch (from trifluridine/tipiracil in combination with bevacizumab to trifluridine/tipiracil monotherapy and from trifluridine/tipiracil monotherapy to trifluridine/tipiracil in combination with bevacizumab (i.e. planned arm different from actual arm)).

Secondary objectives and associated estimands of interest

Secondary objectives are to estimate the effect of trifluridine/tipiracil in combination with bevacizumab versus trifluridine/tipiracil monotherapy in terms of PFS, ORR, and DCR in patients with refractory mCRC.

Table (3.4) 2 - Summary of estimands of interest associated to secondary objective

Estimand	Variable	Population level summary	Handling IEs	Treatment condition	Population
Secondary estimand: effect of the randomised treatments on progression-free survival in all subjects regardless of whether or not intercurrent events occur	PFS	HR	Treatment policy strategy	trifluridine/tipiracil + bevacizumab potentially followed by further anti-cancer therapy vs trifluridine/tipiracil monotherapy potentially followed by further anti-cancer therapy	Full Analysis Set (FAS)
Secondary estimand: effect of the randomised treatments on response in all subjects before modification of randomised treatment	ORR DCR	Risk difference	While on treatment strategy	trifluridine/tipiracil + bevacizumab vs trifluridine/tipiracil monotherapy	Full Analysis Set (FAS)

Another secondary objective is to compare the impact on QoL of trifluridine/tipiracil in combination with bevacizumab to trifluridine/tipiracil monotherapy in patients with refractory mCRC.

3.4.1. Primary estimand based on the OS

The primary estimand of interest is the effect of the randomised treatments on the survival duration in all subjects regardless of whether or not intercurrent events occur (treatment policy strategy). The motivation for this choice is to assess the efficacy of the trifluridine/tipiracil in combination with bevacizumab compared to trifluridine/tipiracil monotherapy under the ITT principle assuming further anti-cancer therapies represent clinical practice (intercurrent events are considered to be part of the treatments being compared).

All data collected during the trial regardless of occurrence of an IE will be used. This is aligned with the treatment policy strategy.

3.4.1.1. Primary analysis

Overall survival is defined as the time elapsed between the date of randomisation and the date of death due to any cause. For patients without documentation of death, the OS will be censored at the last contact date the patient was known to be alive.

Table (3.4.1.1) 1 - Event and Censoring Rules for primary OS

Event/Censor	Decision	Date of event or censor to consider for analysis
Death	Event	date of death
Patient alive without documented death	Censor	date of last contact*

** Maximum date among completed dates relative to patient's information*

The distribution of OS will be compared between the two treatment groups using a stratified log-rank test at one-sided 2.5% level of significance (stratification factors based on IWRS data).

OS for each arm will be summarized using Kaplan Meier curves and further characterized in terms of the median and survival probabilities at 6, 12 and 18 months along with the corresponding 2-sided 95% CI for the estimates.

The hazard ratio of OS with its 95% confidence interval will be estimated with a stratified Cox proportional hazard model (stratification factors based on IWRS data).

For missing data (not linked to intercurrent events), *i.e.* patients without documentation of death (lost to follow-up, withdrawal of consent, administrative end of study), the OS will be censored on the last contact date the patient was known to be alive or the survival cut-off date, whichever is earlier.

Underlying assumptions of proportional hazards will be checked using Schoenfeld Residuals test and graphical methods (Log cumulative hazard curve, etc…). If proportionality is not observed, sensitivity analyses other than those already planned in the SAP could be carried out.

3.4.1.2. Sensitivity analyses

As sensitivity analyses, the following analyses will be done:

- The distribution of OS will be compared between the treatment groups using an unstratified log-rank test and the hazard ratio (together with associated 95% CI) resulting from an unstratified Cox model will be presented.
- The distribution of OS excluding patients not fulfilling one of the following criteria:
  - Has histologically confirmed unresectable adenocarcinoma of the colon or rectum.
  - Has received a maximum of 2 prior chemotherapy regimens for the treatment of advanced colorectal cancer and had demonstrated progressive disease or intolerance to their last regimen.
  - Has measurable or non-measurable metastatic lesion(s) as defined by Response Evaluation Criteria in Solid Tumours (RECIST) version 1.1.
  - Has an Eastern Cooperative Oncology Group (ECOG) performance status ≤ 1.
  - More than 2 prior chemotherapy regimens for the treatment of advanced colorectal cancer.
  - Has previously received trifluridine/tipiracil.
  - Does not currently received or has not received anticancer therapies within 4 weeks prior to randomisation

Will be compared between the treatment groups using a stratified log-rank test (stratification factors based on IWRS data) and the hazard ratio (together with associated 95% CI) resulting from a stratified Cox proportional hazard model (stratification factors based on IWRS data) will be presented.

- The Restricted Mean Survival Time (RMST) for OS will be reported with its 95% confidence interval in each treatment group. The difference in RMST between treatment groups will be tested.

3.4.1.3. Supplementary analysis

3.4.1.3.1. Subgroup analysis

OS subgroup analyses are planned in the FAS to further explore the homogeneity of the treatment effect across patient subsets. Depending on the sample size, the subgroups, as defined in Section 2.3 will be examined.

An unstratified Cox-regression model with treatment arm as predictor variable will be fitted separately for each subgroup category. The hazard ratio for treatment along with the associated 95% confidence interval will be provided.

Forest Plot of Hazard Ratios for treatment effect on primary estimand based on OS by selected subgroup will be provided.

3.4.2. Additional estimand based on OS

The other estimand based on OS is defined in order to assess the effect of the randomised treatments on the survival time in all subjects before patients receive further anti-cancer therapy (while on treatment strategy). The aim of this estimand is to evaluate the effect of the study treatment without taking into account any potential effects of another anti-cancer therapy.

The use of data post intercurrent events will be different according to the IE:

- For IE “administration of further anti-cancer therapy” (5 Appendix B) and “treatment switch”: data obtained post IE will not be used for the analysis. OS will be censored at the time of administration of further anti-cancer therapy.
- For IE “treatment discontinuation”: time of death for any cause after treatment discontinuation will be taken into account.

Table (3.4.2) 1 - Event and Censoring Rules for OS considering intercurrent events

Event/Censor	Decision	Date of event or censor to consider for analysis
Death before or at administration of further anti-cancer therapy, treatment switch and treatment discontinuation	Event	Date of death
Death after treatment discontinuation and before or at administration of further anti-cancer therapy and treatment switch	Event	Date of death
Death after administration of further anti-cancer therapy or treatment switch	Censor	Min (date of administration of further anti-cancer therapy;date of treatment switch)
Patient alive without documented death and with administration of further anti-cancer therapy or treatment switch	Censor	Min (date of administration of further anti-cancer therapy;date of treatment switch)
Patient alive without documented death and without administration of further anti-cancer therapy and treatment switch	Censor	Date of last contact*

** Maximum date among completed dates relative to patient's information*

The distribution of OS will be compared between the two treatment groups using a stratified log-rank test at one-sided 2.5% level of significance (stratification factors based on IWRS data).

OS for each arm will be summarized using Kaplan Meier curves and further characterized in terms of the median and survival probabilities at 6, 12 and 18 months along with the corresponding 2-sided 95% CI for the estimates.

The hazard ratio of OS with its 95% confidence interval will be estimated with a stratified Cox proportional hazard model (stratification factors based on IWRS data).

For missing data (not linked to intercurrent events), i.e. patients without documentation of death (lost to follow-up, withdrawal of consent, administrative end of study), the OS will be censored on the last contact date the patient was known to be alive or the survival cut-off date, whichever is earlier.

3.4.3. Secondary estimands based on PFS, ORR and DCR

3.4.3.1. Secondary estimand based on PFS

Important secondary estimands of interest is the effect of the randomised treatments on progression-free survival in all subjects regardless of whether or not intercurrent events occur (treatment policy strategy). The PFS is based on investigator judgement and is defined as the time elapsed between the randomisation and the date of radiologic tumour progression (according to RECIST version 1.1 (Eisenhauer, 2009) or death from any cause.

All data collected during the trial regardless of occurrence of an IE will be used. This is aligned with the estimand and the treatment policy strategy.

Table (3.4.3.1) 1 - Event and Censoring Rules for primary PFS

Event/Censor	Decision	Date of event or censor to consider for analysis
Documented first radiological progression	Event	Date of 1st radiological PD assessment
Death without radiological progression	Event	Date of Death
Patient alive without documented radiological progression	Censor	Date of last evaluable tumour assessment (non-NE)
Patient lost to follow-up without documented radiological progression	Censor	Date of last evaluable tumour assessment (non-NE)
No baseline or post-baseline tumour assessment	Censor	Date of randomisation

For missing data (not linked to intercurrent events), i.e. patients who were lost to follow-up or who have withdrawn their consent without radiological progression or reached the time point of analysis without a known record of death or radiological progression, the PFS will be censored at the date of last evaluable tumour assessment or the survival cut-off date, whichever is earlier.

The following analyses will be done as primary analyses:
- The distribution of PFS will be compared between the two treatment groups using a stratified log-rank test at one-sided 2.5% level of significance (stratification factors based on IWRS data).
- PFS for each arm will be summarized using Kaplan Meier curves and further characterized in terms of the median and survival probabilities at 3, 6, 9 and 12 months along with the corresponding 2-sided 95% CI for the estimates.
- The hazard ratio of PFS with its 95% confidence interval will be estimated with a stratified Cox proportional hazard model (stratification factors based on IWRS data).

PFS is identified as the key secondary variable. A hierarchical testing strategy, where PFS is to be statistically evaluated and interpreted only if the primary efficacy estimand OS is significantly different between the 2 treatment groups, will be used to control the overall type-I error rate.

Sensitivity analyses

The following analyses will be done as sensitivity analyses:
- PFS sensitivity analysis to address the impact of stratification (supportive 1): the distribution of PFS will be also compared between the treatment groups using an unstratified log-rank test and the hazard ratio (together with associated 95% CI) resulting from an unstratified Cox model will be presented.
- PFS sensitivity analysis to address the impact of further therapies and clinical progression (supportive 2): an analysis that considers clinical progression and administration of further anti-cancer therapy as PFS events in addition to the radiological progression or death.

- PFS sensitivity analysis to address the impact of further therapies and missing tumour assessments (supportive 3): an analysis censoring the PFS duration at the date of last evaluable tumour assessment before or the day of administration of further anti-cancer therapy and censoring patients with radiological progression or who die after ≥ 2 consecutive missed radiological assessments at the latest evaluable assessment prior to the missing assessments. Based on the time interval (distance) between the last evaluable tumor assessment date and the event date, if the distance is greater than (2*8) + 2 = 18 weeks then the analysis will assume two missing assessments. The threshold is formed based on the protocol specified interval between the tumor assessments plus 2 weeks.

CL3-95005-007 - Statistical Analysis Plan | VV-TMF-103206 | 1.0

Table (3.4.3.1) 2 - Event and Censoring Rules for PFS in supportive analysis 2

Event/Censor	Decision	Date of event or censor to consider for analysis
Documented first radiological progression without previous clinical progression nor new anti-cancer therapy or documented first radiological progression with new anti-cancer therapy at the same date without previous clinical progression or documented first radiological progression with clinical progression at the same date without new anti-cancer therapy or documented first radiological progression with clinical progression and with new anti-cancer therapy at the same date	Event	Date of 1st radiological PD assessment
Clinical progression without previous radiological progression nor new anti-cancer therapy or clinical progression with new anti-cancer therapy at the same date without previous radiological progression	Event	Minimum of: • Date of withdrawal if reason for withdrawal = progressive disease • Date of progression in AE e-CRF page (PT "Malignant neoplasm progression" with code MedDRA 10051398) • Date of first progression at the follow-up visit
Death without radiological progression and without clinical progression and (without previous new anti-cancer therapy or with previous new anti-cancer therapy at the same date)	Event	Date of Death
Patient receiving new anti-cancer therapy before any radiological progression, any death and any clinical progression	Event	Date of start of new therapy
Patient alive without documented radiological progression, any clinical progression and any new anti-cancer therapy	Censor	Date of last evaluable tumour assessment (non-NE)
Patient lost to follow-up without documented radiological progression, any clinical progression and any new anti-cancer therapy	Censor	Date of last evaluable tumour assessment (non-NE)
No baseline or post-baseline tumour assessment	Censor	Date of randomisation

Table (3.4.3.1) 3 - Event and Censoring Rules for PFS in supportive analysis 3

Event/Censor	Decision	Date of event or censor to consider for analysis
Documented first radiological progression not preceded by ≥2 consecutive missed tumour assessments* and without new anti-cancer therapy previously reported	Event	Date of 1st radiological PD assessment
Death not preceded by ≥2 consecutive missed tumour assessments* without radiological progression and without new anti-cancer therapy previously reported	Event	Date of Death
Patient alive without documented radiological progression and without new anti-cancer therapy previously reported	Censor	Date of last evaluable tumour assessment (non-NE)
Patient receiving new anti-cancer therapy before any radiological progression or death not preceded by ≥2 consecutive missed tumour assessments*	Censor	Date of last evaluable tumour assessment (non-NE) prior to or the day of new anti-cancer therapy
Patient receiving new anti-cancer therapy before any radiological progression or death preceded by ≥2 consecutive missed tumour assessments*	Censor	Min(Date of last evaluable tumour assessment (non-NE) prior to or the day of new anti-cancer therapy; Date of last evaluable tumour assessment (non-NE) prior to missing visits)
Document first radiological progression or death preceded by ≥2 consecutive missed tumour assessments* without new anti-cancer therapy previously reported	Censor	Date of last evaluable tumour assessment (non-NE) prior to missing visits
Patient lost to follow-up without documented radiological progression and without new anti-cancer therapy previously reported	Censor	Date of last evaluable tumour assessment (non-NE)
No baseline or post-baseline tumour assessment	Censor	Date of randomisation

* Based on the time interval (distance) between the last evaluable tumor assessment date and the event date, if the distance is greater than (2*8) + 2 = 18 weeks then the analysis will assume two missing assessments.

For missing data (not linked to intercurrent events), i.e. patients who were lost to follow-up or who have withdrawn their consent without radiological progression or reached the time point of analysis without a known record of death or radiological progression, the PFS will be censored at the date of last evaluable tumour assessment or the survival cut-off date, whichever is earlier.

The same subgroup analyses as used for the primary estimand will be done for PFS.

3.4.3.2. Secondary estimands based on ORR and DCR

Other secondary estimand of interest are the effect of the randomised treatments on response in all subjects before modification of randomised treatment.

Responses recorded after intercurrent event will be excluded. This is aligned with the estimand and the while on treatment strategy.

The frequency distribution of Best Overall Response (BOR) will be provided on the FAS according to the following classes: CR, PR, SD, Non CR/ Non PD, PD, NE.

The following analyses will be done:

- Overall Response Rate (ORR) is based on the investigator's tumour assessment and is defined as the proportion of patients with objective evidence of complete response (CR) or partial response (PR) according to RECIST version 1.1 criteria and using investigator's tumour assessment.
  ORR will be described using 2-sided 95% Clopper-Pearson CIs in each treatment arm. A Fisher's exact test and a 2-sided 95% CI for the difference in ORR between the two treatment arms will also be provided based on the normal approximation. If a stratified analysis is required a Cochran-Mantel-Haenszel (CMH) test will be used.
- Disease Control Rate (DCR) is based on the investigator's tumour assessment and is defined as the proportion of patients with objective evidence of CR or PR or stable disease (SD) according to RECIST version 1.1 criteria and using investigator's tumour assessment.
  DCR will be described using 2-sided 95% Clopper-Pearson CIs in each treatment arm. A Fisher's exact test and a 2-sided 95% CI for the difference in DCR between the two treatment arms will also be provided based on the normal approximation. If a stratified analysis is required a Cochran-Mantel-Haenszel (CMH) test will be used.

As supplementary analysis, BOR, ORR and DCR will be evaluated on population evaluable for tumour response,i.e., all patients of the FAS with measurable disease at baseline (at least one target lesion) and with at least one tumour evaluation while on treatment (with the same method of measurement as baseline).

For missing data (not linked to intercurrent events), an unfavourable outcome will be taken into account.

3.5. Safety analysis

All safety analyses will be performed in the SS.

Calculation rules for safety endpoints and other specific definitions are provided in 6 Appendix C.

Adverse events (AEs) will be graded using the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE version 5.0) with the most recently available version.

Subgroup analyses of adverse events and Grade 3 or 4 clinical laboratory abnormalities will be performed based on the following patient characteristics at baseline:

- ECOG performance status (0, 1).
- Age (< 65, ≥ 65 years).
- Gender.

- Baseline creatinine clearance (< 60ml/min, ≥ 60ml/min).

3.5.1. Adverse events

Number of events, number and percentage of patients reporting at least one event, presented by SOC and/or PT (depending on the analysis) for each treatment arm, will be provided for serious adverse events (SAE) and emergent adverse events (EAE) over the study and treatment periods, respectively.
EAE and SEAE will be described for each treatment arm according to:

- Worst grade.
- Severity.
- Relationship (related to trifluridine/tipiracil, related to bevacizumab, related to the combination (trifluridine/tipiracil or bevacizumab) (7 Appendix D).
- Outcome (recovered/resolved, recovered/resolved with sequelae, recovering/resolving, not recovered/not resolved, fatal, unkown).
- Leading to drug withdrawal.

EAE will be also described for each treatment arm according to:

- Action taken regarding trifluridine/tipiracil (drug interrupted, dose reduced, dose delayed, dose reduced and delayed, dose not changed), action taken regarding bevacizumab (drug interrupted, dose delayed).
- EAE related to disease progression.
- EAE leading to death.
- Requirement of added therapy.

Non SEAE (EudraCT analyses) and Non SEAE related (related to trifluridine/tipiracil, related to bevacizumab, related to the combination (trifluridine/tipiracil or bevacizumab)) will be described by SOC and PT for each treatment arm.

Of note, the seriousness and the relationship with trifluridine/tipiracil and/or bevacizumab are based on investigator opinion but also include sponsor decision to upgrade seriousness and/or relationship.

3.5.2. Death

Number of deaths will be described on each period (during the treatment and follow-up period / during the treatment period / during the follow-up period) by treatment arm.

Reason of death on each period will be provided. TEAE leading to death and TEAE related to disease progression and leading to death will be presented by SOC and PT during the treatment period / during the treatment or follow-up periods. During the follow-up period, reason of death (PD, other) will be summarized.

Listing of patients who died during the treatment period or during the follow-up period will be provided.

3.5.3. Clinical laboratory evaluation

For haematological, biochemistry, coagulation and urinary parameters, descriptive statistics on value at baseline, on value at each cycle under treatment (worst value at each cycle), on worst value under treatment, on last post-baseline value under treatment and, on change from baseline to each cycle (worst value at each cycle) under treatment, to worst value under treatment and to last post-baseline value under treatment will be provided.

Number and percentage of patients for each laboratory parameter classified according to: reference ranges (< LLN/≥ LLN or ≤ ULN/> ULN) for non-gradable parameters except creatinine clearance, grade (Grade 0, Grade 1, Grade 2, Grade 3, Grade 4) for gradable parameters, following classes for proteinuria (Absence, Trace, Presence (+, More than one +)) and following classes for creatinine clearance (severe, mild, moderate, normal) and using shift tables from baseline to the worst (high and/or low) class or value under treatment will be described by treatment arm.
Some parameters with grade dependent on baseline value (no grade cannot be computed at baseline).

3.5.4. Vital signs, clinical examination and other observations related to safety

Vital signs and clinical examination will be described, in terms of value at baseline, value at each cycle (worst value at each cycle) under treatment, worst value under treatment and last post-baseline value under treatment; as well as in terms of change from baseline to each cycle (worst value at each cycle) under treatment, to worst value under treatment, to last post-baseline value under treatment by treatment arm. Vital signs and clinical examination will be also described by classes.
Time from randomisation to worsening ECOG performance status of ≥ 2 will be analysed in the FAS using Kaplan-Meier methodology and compared between treatment groups with a stratified log-rank test (strata based on IWRS data). To be considered as a worsening, the ECOG PS must be increased by at least one category and to at least a minimal PS value of 2. Death will be considered as worsening of ECOG PS ≥ 2**.** Patients without ECOG PS worsening as of the survival cut-off date will be censored at the last recorded ECOG PS assessment.

Table (3.5.4) 1 - Event and Censoring Rules for time to ECOG performance status worsening ECOG of ≥ 2

Event/Censor	Decision	Date of event or censor to consider for analysis
ECOG from 0-1 at baseline to ≥ 2 post baseline	Event	Date of first worsening ECOG assessment
ECOG from 2 at baseline to ≥ 3 post baseline	Event	Date of first worsening ECOG assessment
Death without previous ECOG ≥ 2 deterioration	Event	Date of Death
Patient without worsening ECOG with at least a minimum ECOG value of 2 (ECOG from 0 at baseline to 1 post baseline, same ECOG value at baseline and post baseline)	Censor	Date of last ECOG assessment
No baseline or/and no post baseline ECOG evaluation	Censor	Date of randomisation

3.5.5. ECG

ECG parameters will be described, in terms of value at baseline and last post-baseline value under treatment; as well as, for quantitative endpoints, in terms of change from baseline to last post-baseline value under treatment Moreover values and changes from baseline of uncorrected and corrected (Fridericia's correction and Bazett's correction) QT interval and RR will also be described in classes, considering the thresholds defined in ICH E14:

- ≤ 450,] 450; 480],] 480; 500] and > 500 ms for values.
- ≤ 30,] 30; 60] and > 60 msec for changes.

3.6. QoL

3.6.1. EORTC QLQ-C30

The QLQ-C30 is a cancer specific instrument that contains 30 questions and provides a multi-dimensional assessment of health related QoL. The QLQ-C30 is composed of five functional scales, three symptom scales, six single items and a global health status (GHS).
The raw scores are linearly transformed to give standard scores in the range of 0-100. Higher scores in the global and functioning scales and lower scores in the symptom/single scales indicate better quality of life.

The number of patients from the FAS who have completed at least one questionnaire item will be described by treatment group at each visit as well as the reasons of non-completion among patients still on treatment.

For the QLQ-C30 analyses, patients from the FAS who have completed at least one questionnaire item at baseline and during the study period will be analysed in the arm they were assigned to at randomisation, unless otherwise specified.

The raw QoL data will be scored according to the EORTC scoring manual and the change in score from baseline in the GHS scale is identified as the primary QoL variable of interest.

For each QLQ-C30 scale, the completion and compliance rates will be summarised by treatment group for each scheduled assessment timepoints.

The completion rate is defined as the rate of patients with a non-missing scale score among patients of the FAS.

The compliance rate is defined as the rate of patients with a non-missing scale score among patients of the FAS still on treatment at each visit.

Absolute difference less than 5 points in the scores represent 'no change' (Osoba D, 1998), absolute difference of 5-10 points in the scores represents a small change, 10-20 points a moderate change and greater than 20 points a large change from the patient's perspective. Any difference greater than 10 points will be considered as clinically significant.

Descriptive statistics will be used to summarize baseline values, post-baseline values and changes from baseline in the GHS and sub-scale scores at each scheduled assessment time point prior to any procedure (at each cycle and at withdrawal visit) and by classes (large benefit, moderate benefit, small benefit, no change, small deterioration, moderate deterioration, and large deterioration) for each treatment arm.

Figures will be displayed in order to show the evolution of each sub-scale score from baseline values to post-baseline values at each scheduled assessment time point prior to any procedure (at each cycle and at withdrawal visit).

A repeated-measures mixed-effects model (SAS PROC MIXED) that includes terms for treatment arms, baseline stratification factors, baseline sub-scale score and time to visit (at each scheduled assessment time point) will be used to compare the two treatment groups with respect to changes from baseline in the sub-scales scores longitudinally over time (Global health status, functional scales and symptoms scales). Treatment arms by time to visit interaction will be explored as well. The most suitable covariance structure will be chosen according to Akaike's Information Criterion and Schwartz' Bayesian Criterion, between unstructured, compound symmetric and autoregressive of order 1.

Time Until Definitive Deterioration (TUDD) of 10 points will be compared for each treatment arm. TUDD is defined as the interval between baseline date and the first worsening in QoL score ≥ 10 points compared to baseline QoL score with no later improvement above this threshold observed during the study. Death is considered as deterioration of QoL. Patients receiving any further anti-tumoral treatment before definitive worsening will be censored at the date of their last QoL assessment before starting this therapy. Patients that have not definitively worsened as of the cut-off date for the analysis will be censored at the date of their last assessment before the cut-off. Patients with a missing scale score at baseline will be censored at their randomisation date + 1 day.

TUDD of 10 points in each sub-scale will be compared between the two treatment arms using the stratified log-rank test at a 2-sided 5% level of significance. The distributions will be described using Kaplan-Meier curves including the median time to definitive 10 points deterioration along with the 95% confidence intervals.

3.6.2. EQ-5D-5L

The EQ-5D 5L is a standardized health utility measure developed by the EuroQOL Group in order to provide a simple, generic measure of health for clinical and economic appraisal. It is composed of 5 questions focused on mobility, self-care, usual activities, pain/discomfort and anxiety/depression, and a Visual analogue scale (VAS). Each of the 5 questions has 5 levels: no problem, slight problems, moderate problems, severe problems and extreme problems.

The VAS records the respondent's self-rated health on a vertical visual analogue scale. The VAS 'thermometer' has endpoints of 100 (Best imaginable health state) at the top and 0 (Worst imaginable health state) at the bottom. This information can be used as a quantitative measure of health outcome as judged by the individual respondents.

For the EQ-5D 5L analyses, patients from the FAS with an evaluable EQ-5D-5L assessment at baseline and at least one evaluable assessment post baseline will be analysed in the arm they were assigned to at randomisation, unless otherwise specified.

The dimensional 5-level system will be converted into a single index utility score: values for the 3125 theoretically possible health states defined by the EuroQol classification are calculated using a regression model (Busschbach et al., 2003).

The index utility score will be derived according to Euroqol specific country value set. For all countries, the French Value Set will be used to generate health utility scores.

The number of patients from the FAS who have completed at least one questionnaire item will be described by treatment group at each visit as well as the reasons of non-completion among patients still on treatment.

The completion and compliance rates will be summarized by treatment group for each scheduled assessments timepoints.

The completion rate is defined as the rate of patients with evaluable EQ-5D-5L assessment among patients of FAS.

The compliance rate is defined as the rate of patients with evaluable EQ-5D-5L assessment among patients still on treatment at each visit among patients of FAS.

A breakdown table to include the proportions of reported problems at baseline versus each scheduled assessment timepoint (at each cycle and at withdrawal visit) will be presented in each dimension of EQ-5D-5L by treatment arm.

Descriptive statistics will be used to summarize baseline values and changes from baseline in EQ-5D-5L VAS and EQ-5D-5L health utility index by treatment arm at each scheduled assessment time point prior to any procedure (at each cycle and at withdrawal visit).

Figures will be displayed in order to show the evolution of EQ-5D-5L VAS and EQ-5D-5L health utility index from baseline values to post-baseline values at each scheduled assessment time point prior to any procedure (at each cycle and at withdrawal visit).

A repeated-measures mixed-effects model (SAS PROC MIXED) that includes terms for treatment, baseline stratification factors, baseline score and time to visit prior to any procedure (at each scheduled assessment time point) will be used to compare the two treatment groups with respect to changes from baseline in the EQ-5D-5L VAS and EQ-5D-5L health utility index. Treatment arms by time to visit interaction will be explored as well. The most suitable covariance structure will be chosen according to Akaike's Information Criterion and Schwartz' Bayesian Criterion, between unstructured, compound symmetric and autoregressive of order 1.

4. APPENDIX A: DATA HANDLING AND PROGRAMMING SPECIFICATIONS

4.1. General analytic definitions

4.1.1. Expressions

Reliable value for biological and coagulation samplings is identified in datasets with flag BIOFG_*parameter name* ("Non analysable value") different from 1 and non-missing result. For other panels, **reliable value** is a non-missing result.

Value at baseline is defined as the last reliable value prior to treatment intake
Note: In case of patient included and/or randomised but not treated (*i.e.* patients with treatment duration equal to 0): value at baseline is defined as the last reliable value.

Post-baseline value is defined as a reliable value after the first study treatments intake.

Last value is defined as the last reliable post-baseline value of the criteria of interest during treatment period.

Withdrawal value is defined as the reliable value of the criteria of interest during the withdrawal visit.

Lowest value is defined as the lowest value during the treatment period.

Highest value is defined as the highest value during the treatment period.

Worst (lowest/highest) value at cycle i is defined as the worst (lowest/highest) value at $cycle_i$.

Change from baseline to worst (lowest/highest) value at cycle i is calculated as:
Worst value at the $Cycle_i$ - Value at baseline.

Change from baseline to last value is calculated as:
Last value - Value at baseline.

Change from baseline to worst (highest/lowest) value is calculated as:
Worst value - Value at baseline.

Change from baseline to withdrawal value is calculated as:
Withdrawal value - Value at baseline.

4.1.2. Value prior to treatment / under treatment

Table (4.1.2) 1 - Time frame

Criteria	Value prior to treatment if measured between (D1=) days before and (D2=) days after the first treatment intake (D1 and D2 included)		Post-baseline value on studied/interest period if measured between (D1=) days after the first treatment intake and until (D2=) days after the last treatment intake as defined in Section 4.1.5 (D1 and D2 included)	
Tumour assessment (all lesions)	D1 = -∞	D2 = Xsup	D1 = Xinf	D2 = 30
Biochemistry/haematology/ coagulation/urinanalysis	D1 = -∞*	D2 = Xsup*	D1 = Xinf*	D2 = 30
Vital signs and physical examinations	D1 = -∞*	D2 = Xsup*	D1 = Xinf*	D2 = 30
Pregnancy test	D1 = -∞*	D2 = Xsup	D1 = NA	D2 = NA
ECG	D1 = -∞	D2 = Xsup	D1 = Xinf	D2 = 30
QoL (score, completion, compliance)	D1 = -∞	D2 = Xsup**	D1 = Xinf**	D2 = +∞
QoL (reason of non-completion)	D1 = -∞	D2 = Xsup***	D1 = Xinf***	D2 = +∞
AE expected	D1 = -∞	D2 = Xsup	D1 = Xinf	D2 = 30

-∞ : the last evaluable assessment before the first treatment intake date where visit< "C001"
-∞: the last evaluable assessment before the first treatment intake date where visit≤"C001D001"*
Xsup: 0 and" Visit< "C001"
Xsup: 0 and" Visit≤ "C001D001"*
*Xsup**: 0*
*Xsup***: Visit ≤ "A000"*
Xinf: 0 and" Visit ≥ "C001"
Xinf: 1 and" Visit ≥ "C001",(Note: C001UNS is considered as a post-dose assessment when reported the same date as the first treatment intake)*
*Xinf**: 1*
*Xinf***: Visit > "A000"*

Table (4.1.2) 2 - Time frame of cycle n

Notations	Definitions
Cycle n	Start date: First trifluridine/tipiracil intake date of cycle n End date: The day before the 1st trifluridine/tipiracil intake of cycle (n+1) ➢ It means that assessments planned the same day but before the 1st trifluridine/tipiracil drug intake in cycle (n+1) are part of cycle n. For the last cycle, End date = min ((*), date of death). (*) first intake date of last cycle +27 days for trifluridine/tipiracil monotherapy and max (first intake date of trifluridine/tipiracil of last cycle +27 days, first intake date of bevacizumab of last cycle +27 days) for trifluridine/tipiracil + bevacizumab arm The Data Management variable VISIT in exposure dataset will be used to determine the start date of each cycle.

4.1.3. General duration derivation and conversion

In instances where duration or times-to-event are calculated, the convention to be used unless otherwise specified is [later date] – [earlier date] + 1 day.

When converting a number of days to other units, the following conversion factors will be used: 1 year = 365.25 days; 1 month = 30.44 days

4.1.4. Analysable value

Table (4.1.4) 1 - Definition of analysable value

General definition	
Non missing value	
Specific definitions	
Laboratory parameters	Only reliable values are considered for analyses. Unreliable values are flagged into the database.
ECG parameters	For the analysis of QT criteria per measurement time, only scheduled values will be analyzed.

4.1.5. First and last IMP intake dates

For patients having taken at least one dose of IMP, the dates of first and last IMP intake on the analysis period will be defined as follows:

- The date of the first IMP intake = min (first trifluridine/tipiracil intake date, first bevacizumab infusion date) within the analysis period.
- The date of the last IMP intake = max (last trifluridine/tipiracil intake date, last bevacizumab infusion date) within the analysis period.

After selection of the dates of first and last IMP intake as defined above, if these dates are missing or incomplete, the following substitution rules will be applied:

Table (4.1.5) 1 - Substitution rules of IMP intake dates

Date to substitute		**Substituted date**
First IMP intake	../mmm/yyyy	Inclusion date* if complete with same month and year *Otherwise:* **01/mmm/yyyy**
	../…/yyyy	Inclusion date* if complete with same year *Otherwise:* **01/JAN/yyyy**
	../.../….	Inclusion date*if totally incomplete
Last IMP intake	../mmm/yyyy	**Last available date**** if same month and year *Otherwise:* **last day of the month/mmm/yyyy**
	../…/yyyy	**Last available date**** if same year *Otherwise:* **31/DEC/yyyy**
	../.../….	**Last available date****

Notes:

- *Missing dates will be substituted only for patients having taken at least one dose of study treatment*
  ../mmm/yyyy = missing day
  ../.../ yyyy = missing day and month
  ../.../.... = totally missing date
- ** inclusion date (or dispensation date if inclusion date is missing, or randomization date if dispensation date is missing)*
- *** Last available date (only for patients included) is defined as date of death if patient died, and as maximum date among completed dates relative to patient's information otherwise.*
-

Note:

Cycles with both missing first and last IMP intake dates and with number of returned tablets equal to number of tablets delivered at the previous cycle (or with estimated number of tablets taken equal to 0) will not be taken into account.

4.1.6. Other dates

The rules for substitution of missing or incomplete death date are as follows:

Table (4.1.6) 1 - Substitution rules of death date

Date to substitute		Substituted date
Death date	../mmm/yyyy	**Last available date*** if same month and year *Otherwise:* **01/mmm/yyyy**
	../…/yyyy	**Last available date*** if same year *Otherwise:* **01/JAN/yyyy**
	../.../….	**No substitution**

Notes:

../mmm/yyyy = missing day
../.../ yyyy = missing day and month
../.../.... = totally missing date

- ** Last available date is defined as maximum date among completed dates relative to patient's information otherwise.*

If no specific management of dates is defined, missing information is substituted as defined below:

Table (4.1.6) 2 - Substitution rules of dates if no specific management is defined

Date to substitute		Substituted date
Date	../mmm/yyyy	01/mmm/yyyy
	../…/yyyy	01/JAN/yyyy
	../.../….	No substitution

Note:
../mmm/yyyy = missing day
../.../ yyyy = missing day and month
../.../.... = missing date

4.2. Specific analytic definitions and data handling conventions

4.2.1. Study patients: Disposition, baseline characteristics and follow-up

4.2.1.1. Disposition of patients

All treatment withdrawal reasons occurred in the study will be taken into account (adverse event, protocol deviation, progressive disease (radiological progressive disease, clinical progressive disease, radiological and clinical progressive disease), non-medical reason (consent withdrawal from study treatment, consent withdrawal including survival follow-up), lost to follow-up, physician decision or other medical reason).

4.2.1.2. Protocol deviations

Deviation categories have been defined in order to gather all the deviations relative to the same topic.

For the description of protocol deviations, the 6 following categories are considered in accordance with ICH E3 guideline and ICH E3 Q&A:

- Selection/inclusion criteria not fulfilled.
- Patient having withdrawal criteria but not withdrawn.
- Incorrect treatment or dose received.
- Forbidden concomitant treatment.
- Endpoint assessment possibly affected.
- Safety possibly affected.

4.2.1.3. Demographic data and other baseline characteristics

4.2.1.3.1. Demographic data

Age is calculated as difference between year of informed consent and year of date of birth.

Geographic region is defined from Enrollment e-CRF page:
- North America: USA.
- European Union: Austria, Belgium, Denmark, France, Germany, Hungary, Italy, Poland and Spain.
- Rest of the World: Brazil, Ukraine and Russia.

4.2.1.3.2. History of Colorectal cancer

Primary diagnosis in classes, colon (including colorectal) versus rectal versus other, will be defined according to the codelists 4012.0 (colon and colorectal cancer) and 4013.0 (rectal cancer). If a patient is not associated with any of these lists, he will be considered as other.

Primary tumour localisation:
- **Right** if primary tumour site = right colon(1) or transverse or other*
- **Left** if primary tumour site = left colon(2) or rectum

(1) As per protocol, right colon should include ascending colon, cecum, appendix and hepatic flexure.
(2) As per protocol, left colon should include descending colon, sigmoid colon, splenic flexure and rectosigmoid segment.
*Other: *in the case the patient present multiple tumours in both sides, the disease will be considered as "right-sided".*

Site of metastasis is defined from History of colorectal cancer e-CRF page.

Number of metastatic site is defined as the number of distinct site of metastasis (whether the site is considered for target or non-target lesion) from History of colorectal cancer e-CRF page. Sites filled in 'Other' will be counted as 1 site.

Disease duration
Disease duration (years) is defined as (Randomisation date – date of the first diagnosis)/365.25

Time from first metastasis diagnosis to randomization
Time from first metastasis diagnosis to randomization (months) is defined as (Randomization date – date of the first metastasis diagnosis)/30.44

Progression free interval
Progression free interval (months) is defined as (Date of progression prior to the randomisation date – date of end of the last prior anti-cancer therapy*)/30.44.
The progression free interval of patients progressing during the last regimen is estimated to last one day.
*Including drug treatment, radiotherapy and surgery (whatever the intent).

Treatment free interval

Treatment free interval (months) is defined as (Date of first trifluridine/tipiracil intake – date of end of the last prior anti-cancer therapy*)/30.44.

*Including drug treatment, radiotherapy and surgery.

Of note, in case of partial or incomplete date:

Table (4.2.1.3.2) 1 - Substitution rules of History of colorectal cancer

Missing diagnosis date		**Substituted diagnosis date**
../mm/yyyy	⇨	Diagnosis date = **01**/mm/yyyy
../…/yyyy	⇨	Diagnosis date = **01/01**/yyyy
Missing metastasis date		**Substituted metastasis date**
../mm/yyyy	⇨	Metastasis date = **01**/mm/yyyy
../…/yyyy	⇨	Metastasis date =**01/01**/yyyy
Missing start date of the last prior therapy		**Substituted start date of the last prior therapy**
../mm/yyyy	⇨	start date of the last prior therapy = **01**/mm/yyyy
../…/yyyy	⇨	start date of the last prior therapy = **01/01**/yyyy
Missing end date of the last prior therapy		**Substituted end date of the last prior therapy**
../mm/yyyy	⇨	end date of the last prior therapy = start date of the last prior therapy + 1day if same month and year *otherwise*: end date of the last prior therapy = 01/mm/yyyy
../…/yyyy	⇨	end date of the last prior therapy = start date of the last prior therapy + 1day if same year *otherwise*: end date of the last prior therapy = 01/01/yyyy
../…/…	⇨	end date of the last prior therapy = start date of the last prior therapy + 1day
Missing start date of a previous treatment	⇨	**Substituted start date of a previous treatment**
../mmm/yyyy	⇨	01/mmm/yyyy
../…/yyyy	⇨	01/JAN/yyyy
../.../….	⇨	Inclusion date

Note: ../mm/yyyy = missing day
../.../ yyyy = missing day and month
../.../.... = missing date

RAS status

Table (4.2.1.3.2) 2 - RAS status

KRAS	**NRAS**	**RAS**
MUTANT	whatever (mutant, wild, missing)	MUTANT
whatever (mutant, wild, missing)	MUTANT	MUTANT
WILD	WILD	WILD
MISSING	WILD	NE
WILD	MISSING	NE
MISSING	MISSING	NE

Stage at diagnosis

Table (4.2.1.3.2) 3 - Staging based on classification

Stage in classes	AJCC stage	Dukes' classification	Astler-Coller modified Dukes' classification
	Stage 0	-	-
I-II	**Stage I**	A	A
I-II	**Stage I**	A	B1
I-II	**Stage II-A**	B	B2
I-II	**Stage II-B**	B	B2
I-II	**Stage II-C**	B	B3
III	**Stage III-A**	C	C1
	Stage III-B	C	C2
		C	C1/C2
		C	C1
	Stage III-C	C	C2
		C	C2/C3
		C	C3
IV	**Stage IV-A**	D	D
	Stage IV-B		
	Stage IV-C		

4.2.1.3.3. Medical history and surgical or medical procedures history other than colorectal cancer

The existence of a history (Yes/No) is defined from the presence, or not, of a Primary system organ class and/or Preferred term.

4.2.1.3.4. Previous therapies for colorectal cancer

The anatomical therapeutic chemical classification (ATC code = 5 digits) is composed of 4 levels:

- The first (1 digit) represents the anatomo-physiological class.
- The second (2 digits) represents the pharmacological class.
- The third (1 digit) represents the pharmacological sub-class.
- The last (1 digit) represents the therapeutic class.

The existence of a previous surgery (Yes, No) is defined from the presence, or not of a SOC and/or a "preferred name" for previous surgery at inclusion visit.

The existence of a previous radiotherapy (Yes, No) is defined from the presence or not of a verbatim for the previous radiotherapy involved site at inclusion visit.

The existence of a previous drug treatment (Yes, No) is defined from the presence, or not of an "anatomical therapeutic chemical classification" and/or a "preferred name" for the previous drug treatment at inclusion visit.

The existence of a previous metastatic drug treatment (Yes, No) is defined from the presence, or not of an "anatomical therapeutic chemical classification" and/or a "preferred name" for the previous drug treatment:

- Given in palliative indication.
- Given in adjuvant or neoadjuvant indication and with a progression before or within 6 months after the drug treatment completion

The existence of a previous therapy (Yes, No) is defined from the presence or not of a surgery, radiotherapy and/or a drug treatment.

The existence of a previous curative surgery (Yes, No) is defined from the presence or not of a surgery with a curative indication.

The existence of a previous palliative surgery (Yes, No) is defined from the presence or not of a surgery with a palliative indication.

The existence of a previous curative radiotherapy (Yes, No) is defined from the presence or not of a previous radiotherapy with a curative, adjuvant or neoadjuvant indication.

The existence of a previous palliative radiotherapy (Yes, No) is defined from the presence or not of a previous radiotherapy with a palliative indication.

The existence of a previous fluoropyrimidine drug treatment (Yes, No) is defined from the presence, or not of a previous drug treatment belonging to the list 3678.0.

The existence of a previous irinotecan drug treatment (Yes, No) is defined from the presence, or not of a previous drug treatment belonging to the list 3674.0.

The existence of a previous oxaliplatin drug treatment (Yes, No) is defined from the presence, or not of a previous drug treatment belonging to the list 3675.0.

The existence of a previous anti-VEGF monoclonal antibody drug treatment (Yes, No) is defined from the presence, or not of a previous drug treatment belonging to the list 4281.0.

The existence of a previous anti-EGFR monoclonal antibody drug treatment (Yes, No) is defined from the presence, or not of a previous drug treatment belonging to the list 4280.0.

Number of prior drug regimens is calculated as the sum of previous drug treatments whatever the intent.

Number of prior metastatic drug regimens is calculated as the sum of previous metastatic drug treatments.

4.2.1.3.5. Initial tumour assessment

Not applicable.

4.2.1.4. Extent of exposure and treatment compliance

Treatment duration for trifluridine/tipiracil

Treatment duration (months) for trifluridine/tipiracil is defined as ([min (first intake date of the last cycle + 27days**,** death date) - first trifluridine/tipiracil intake date] +1)/30.44.

Treatment duration for bevacizumab in trifluridine/tipiracil + bevacizumab arm
Treatment duration (months) for bevacizumab in trifluridine/tipiracil is defined as [(min (first intake date of the last cycle + 27days, death date) - first intake date) +1] / 30.44.

In case of missing or incomplete first intake date of the last cycle, the last intake date of the last cycle will be considered.

Cycle duration for trifluridine/tipiracil
Duration of cycle *i* (weeks) for trifluridine/tipiracil is defined as [[first trifluridine/tipiracil intake of cycle (*i+1*)] – [first trifluridine/tipiracil intake of cycle *i*]] / 7.
Of note, the duration of the last cycle for trifluridine/tipiracil will be estimated to be 28 days for the calculation of Dose Intensity and Relative Dose Intensity at the last cycle.

Cycle duration for bevacizumab
Duration of cycle *i* (weeks) for bevacizumab is defined as [[first trifluridine/tipiracil intake of cycle (*i+1*)] – [first trifluridine/tipiracil intake of cycle *i*]] / 7.
Of note, the duration of the last cycle for bevacizumab will be estimated to be 28 days (trifluridine/tipiracil + bevacizumab arm) for the calculation of Dose Intensity and Relative Dose Intensity at the last cycle.

Follow-up duration
The follow-up duration (months) is calculated as the duration between the last available date and the randomisation date [last available date* – randomisation date +1]/30.44. The events of interest are being alive or lost to follow-up and death is censored.
* Last available date is defined as date of death if patient died, and as maximum date among completed dates relative to patient's information otherwise.

Number of cycles for trifluridine/tipiracil (resp. bevacizumab)
The number of cycles with trifluridine/tipiracil will be defined based on exposure data. A patient is considered to enter in a cycle if there is at least one intake date.

Real administrated dose (mg) for trifluridine/tipiracil
On a time period, real administered dose for trifluridine/tipiracil (mg) = Number of capsules 15 mg taken*15 + Number of capsules 20 mg taken*20
With:

- Number of capsules taken = Number of capsules dispensed - Number of capsules returned

Else, if Number of capsules returned is missing, Number of capsules taken = Estimated number of capsules taken.

Real administrated dose (mg) for bevacizumab
On a time period, real administered dose for bevacizumab (mg) = sum of total dose administrated (full + incomplete doses).
Note: Avastin Summary of Product Characteristics (SmPC): Each ml of concentrate contains 25 mg of bevacizumab so 5mg of bevacizumab corresponds to 0.2 mL and 7.5 mg corresponds to 0.3 mL.

Planned dose intensity (PDI) for trifluridine/tipiracil (mg/m²/week)

	10 days of intake per cycle
35 mg/m² **(bid)** (**70** mg/m²/day)	$\frac{\mathbf{70}(\text{mg/m}^2/\text{day}) * \mathbf{10}}{4\ (\text{wk})}$

Planned dose intensity (PDI) for bevacizumab (mg/kg/week)

	2 days of intake per cycle
5 mg/kg **(qd)** (**5** mg/kg/day)	$\frac{\mathbf{5}(\text{mg/kg/day}) * \mathbf{2}}{4\ (\text{wk})}$

Cumulative dose

The cumulative dose (mg/m²*)* for trifluridine/tipiracil per patient in a time period (during the treatment period or per cycle) is the sum of the total dose that the patient received within that period according to the compliance.

$$Cumulative\ dose\ (mg/m^2) = \sum_{timeperiod} \left(\frac{Real\ Administrated\ dose\ (mg)}{BSA\ (m^2)} \right)$$

Where BSA is calculated in the ClinTrial database by the data management.

The cumulative dose (mg/kg*)* for the bevacizumab per patient in a time period (during the treatment period or per cycle) is the sum of the total dose that the patient received within that period according to the compliance.

$$Cumulative\ dose\ (mg/kg) = \sum_{timeperiod} \left(\frac{Real\ Administrated\ dose\ (mg)}{weight\ (kg)} \right)$$

Dose intensity (DI) per patient

The dose intensity (mg/m²/week) for trifluridine/tipiracil per patient is defined as the cumulative dose (mg/m²*)* received during the treatment period divided by the total treatment duration in weeks.

$$DI\ (mg/m^2/wk)\ per\ patient\ = \frac{Cumulative\ dose\ (mg/m^2)}{Treatment\ duration\ (weeks)}$$

The dose intensity (mg/kg/week) for the bevacizumab per patient is defined as the cumulative dose (mg/kg*)* received during the treatment period divided by the total treatment duration in weeks.

$$DI\ (mg/kg/wk)\ per\ patient\ = \frac{Cumulative\ dose\ (mg/kg)}{Treatment\ duration\ (weeks)}$$

Dose intensity (DI) per cycle

The dose intensity (mg/m²/week) for trifluridine/tipiracil per cycle is defined as the cumulative dose (mg/m²) received during the cycle divided by the cycle duration in weeks.

$$DI\ (mg/m^2/wk)\ at\ cycle\ i\ = \frac{Cumulative\ dose\ (mg/m^2)\ received\ at\ cycle\ i}{Duration\ of\ cycle\ i\ (weeks)}$$

The dose intensity (mg/kg/week) for the bevacizumab per cycle is defined as the cumulative dose (mg/kg) received during the cycle divided by the cycle duration in weeks.

$$DI\ (mg/kg/wk)\ at\ cycle\ i\ = \frac{Cumulative\ dose\ (mg/kg)\ received\ at\ cycle\ i}{Duration\ of\ cycle\ i\ (weeks)}$$

RDI per patient
The RDI (%) per patient is defined as the ratio of the DI to the initial PDI.

$$RDI\ (\%)\ per\ patient\ = \frac{DI}{Planned\ dose\ intensity} * 100$$

RDI per cycle
The relative dose intensity (%) per cycle is defined as the ratio of the dose intensity at cycle i to the planned dose intensity.

$$RDI\ (\%)\ at\ cycle\ i\ = \frac{DI\ at\ cycle\ i}{Planned\ dose\ intensity} * 100$$

Dose reduction for trifluridine/tipiracil
A cycle with a dose reduction is defined according to the trifluridine/tipiracil dispensing e-CRF page (dose level modified since the last cycle (Yes, No)).

Partial dose administered for bevacizumab
Partial dose administered for the bevacizumab is defined by an answer "No" at the eCRF interrogation "Was full dose administered?".

Missed intake for bevacizumab
A missed intake for the bevacizumab is defined by an answer "No" at the e-CRF interrogation "Did the patient receive the associated agent infusion?".

Dispensation postponed for trifluridine/tipiracil
A cycle with a dispensation postponed is defined according to the trifluridine/tipiracil dispensing e-CRF page (Dispensation postponed (Yes, No)).

Unplanned treatment interruption for trifluridine/tipiracil
A treatment interruption is defined according to the number of days between two consecutive end of treatment intake and start of treatment intake when this duration is one day or longer.
Treatment interruption (days) = (date of treatment restarted - date of last dose intake before interruption) -1.
Note: only intra-cycle interruptions will be considered
If at least one of the end or restart intake date is missing then the number of interruption days will be defined according to the e-CRF page (Number of days of interruption).

4.2.1.5. Concomitant treatments

The anatomical therapeutic chemical classification (ATC code = 5 digits) is composed of 4 levels:

- The first (1 digit) represents the anatomo-physiological class.
- The second (2 digits) represents the pharmacological class.

- The third (1 digit) represents the pharmacological sub-class.
- The last (1 digit) represents the therapeutic class.

The existence of a concomitant treatment (Yes, No) is defined from the presence, or not, of an Anatomical therapeutic chemical classification and/or Preferred name.

The **periods** considered **for the analysis** are:
At inclusion for which treatments:
- With start date ≤ inclusion date and stop date ≥ inclusion date or missing are taken into account. Inclusion date is equal to the date of inclusion visit, *i.e.* A000/D000 visit date.

Before the treatment period for which treatments:
- With start date < first trifluridine / tipiracil intake date are taken into account.

During the treatment period for which treatments:
- With start date ≥ first trifluridine/tipiracil intake date and < (*), or
- With start date ≤ first trifluridine/tipiracil intake date and stop date ≥ first trifluridine / tipiracil intake date or missing

After the treatment period for which treatments:
- With start date ≥ (*), or
- With start date ≤ (*) and stop date > (*) or missing.

(*) first intake date of last cycle +28 days for trifluridine/tipiracil monotherapy and max (first intake date of last cycle +28 days of trifluridine/tipiracil, first intake date of last cycle +28 days of bevacizumab) for trifluridine/tipiracil + bevacizumab arm

The following **rules for substitution** of missing or incomplete start and stop dates are so that the concomitance period is maximised:

Table (4.2.1.5) 1 - Substitution rules of concomitant treatments intake dates

Date to substitute		**Substituted date**
Start date	../mmm/yyyy	If the year and the month are the same as the year and the month of inclusion then Start date= **Inclusion date,** otherwise Start date=**01/mm/yyyy**
	../…/yyyy	If the year is the same as the year of inclusion then Start date= **Inclusion date,** otherwise Start date=**01/01/yyyy**
	../.../….	If stop date is non-missing and inferior to inclusion date then: **Stop date** Else: **Inclusion date**

Stop date	../mmm/yyyy	If patient died same month and year then
		Date of death
		Else
		Last day of the month/mmm/yyyy

	../…/yyyy	If patient died same year then
		Date of death
		Else
		31/DEC/yyyy

	../.../….	If patient died then
		Date of death
		Else
		No substitution
		(*i.e.,* treatment considered as still ongoing)

Note: ../mm/yyyy = missing day
../.../ yyyy = missing day and month
../.../.... = missing date

The following rules for substitution of totally or partially missing start and stop dates are used for concomitant treatments (only further anti-tumour therapy):

Table (4.2.1.5) 2 - Substitution rules of further anti-tumour therapy intake dates

Date to substitute		Substituted date
Start date	../mmm/yyyy	If the year and the month are the same as the year and the month of the last trifluridine/tipiracil intake date then **Last trifluridine/tipiracil intake date** otherwise **01/mm/yyyy**
	../…/yyyy	If the year is the same as the year of the last trifluridine/tipiracil intake date then **Last trifluridine/tipiracil intake date**, otherwise **01/01/yyyy**
	../.../….	**Last trifluridine/tipiracil intake date**

Stop date	../mmm/yyyy	If patient died same month and year then
		Date of death
		Else
		Last day of the month/mmm/yyyy ,

	../…/yyyy	If patient died same year then
		Date of death
		Else
		31/DEC/yyyy

	../.../….	If patient died then
		Date of death
		Else
		No substitution
		(*i.e.,* treatment considered as still ongoing)

Note: ../mm/yyyy = missing day
../.../ yyyy = missing day and month
../.../.... = completely missing date
The last trifluridine/tipiracil intake corresponds to the real last trifluridine/tipiracil intake within the analysis period.

The existence of a concomitant G-CSF treatment (Yes, No) is defined from the presence, or not, of a CMCLAS= "COLONY STIMULATING FACTORS".

Primary prophylaxis is defined as:
- A G-CSF intake start date at cycle 1 and no neutropenia at cycle 1, or
- A G-CSF intake start date at cycle 1 and neutropenia at cycle 1 and neutropenia onset date > G-CSF intake start date.

Only the first G-CSF intake start date is considered.

Secondary prophylaxis is defined as:
- Neutropenia at any cycle (cycle i) and a G-CSF intake start date at further cycles (cycle i+1, i+2, …) and neutropenia onset date < G-CSF intake start date.

Therapeutic use is defined as:
- Neutropenia onset date = G-CSF intake start date.
- Neutropenia onset date + 1 day = G-CSF intake start date.

Of note:

- If a G-CSF intake is defined both as a secondary prophylaxis and as a therapeutic use, then it will be defined as a therapeutic use only.
- A patient may have more than one indication related to G-CSF use.
- The following MedDRA codes/preferred terms (PTs) reported are considered as neutropenia:

MedDRA PTs codes	MedDRA PTs
10029354	Neutropenia
10049151	Neutropenic sepsis
10016288	Febrile neutropenia
10029366	Neutrophil count decreased

5. APPENDIX B: EFFICACY

5.1. Non-confirmed Best overall response

Best overall response is defined as the best overall response across all time points excluding "Follow-up" e-CRF information. If applicable, responses recorded after intercurrent events will be excluded.

When SD is believed to be best overall response, it needs to be assessed a minimum of 6 weeks after study randomisation. Otherwise, the best overall response will be NE, unless any PD was further documented, in which case BOR will be PD.

A patient dead because of progression* before the first assessment planned per protocol (8 weeks) will be considered as 'early death'. If a patient progressed* before this first assessment, he will be considered as 'early progressive'. If any of these two events occur, the overall response of the patient will be resumed as progression (PD).

*: Progression/Death due to progression before the end of cycle 2:

- According to tumoral evaluation.
- Withdrawal reason status.
- Or before 8 weeks after the treatment start on AE e-CRF page *(PT « Malignant neoplasm progression » with code MedDRA 10051398.)*

In case of best overall response missing under the studied period, this one will be considered as Non Evaluable (NE).
Complete response (CR) and partial response (PR) will not be confirmed following initial documentation of overall response.

Table (5.1) 1 - BOR when confirmation of CR and PR not required

Overall response First time point (i)	Overall response Next time point (i+1)	BOR
CR	Whatever the overall response	CR
PR	Whatever the overall response other than CR	PR
SD	PD	SD (if the time length between the study randomisation and the SD is ≥ 6 weeks), otherwise, PD
SD	SD or NE or missing	SD (if the time length between the study randomisation and one of the SD is ≥ 6 weeks), otherwise, NE
PD	Whatever the overall response	PD
NE or missing	PD	PD
NE or missing	NE or missing	NE

5.2. Lesions diameters

Relative change of the sum of the lesions diameters (%)
The Relative change and the best relative change of the sum of lesions diameters will be only calculated if the overall response for the target lesions is not equal to 'Non Evaluable'.
Baseline value will be the sum of all target lesions measurements done prior to randomisation.

Relative change from baseline of the sum of the lesions diameters (%) is defined at cycle i as follow:

$$\left(\frac{\text{sum of the lesions diameters at the cycle i} - \text{sum of the lesions diameters at baseline}}{\text{sum of the lesions diameters at baseline}}\right) * 100$$

Relative change from NADIR (SMALLEST) of the sum of the lesions diameters (%) is defined at cycle i as follow:

$$\left(\frac{\text{sum of the lesions diameters at the cycle i} - \text{sum of the lesions diameters (SMALLEST) at baseline or cycle} < \text{i}}{\text{sum of the lesions diameters (SMALLEST) at baseline or cycle} < \text{i}}\right) * 100$$

Best relative change of the sum of the lesions diameters (%)
The best relative change from baseline of the sum of the lesions diameters (%) is defined as the greatest decrease of the sum of the lesions diameters recorded. It can be calculated as follow:

$$\min_{\substack{i \\ i>0}} \left(\frac{\text{sum of the lesions diameters at the cycle i} - \text{sum of the lesions diameters at baseline}}{\text{sum of the lesions diameters at baseline}}\right) * 100$$

5.3. Events and censoring dates

Date of treatment switch
Date of switch from trifluridine/tipiracil in combination with bevacizumab to trifluridine/tipiracil monotherapy or from trifluridine/tipiracil monotherapy to trifluridine/tipiracil in combination with bevacizumab.

1st radiological PD assessment:
Date of the first assessment of the series of the tests that determined PD.
The date of first progression is defined as the first date of PD where the overall response is PD.
The date of PD is defined as the earliest date among the following dates:

- Date of PD for target lesions: If the response for target lesions is PD at the cycle i, the date of PD for target lesions corresponds to the latest date of examination among the target lesions at cycle i.
- Date of PD for non-target lesions: If the response for non-target lesions is PD at the cycle i, the date of PD for non-target lesions corresponds to the earliest date of examination among the non-target lesions showing a PD at the cycle i. A non-target lesion showing a PD is defined as a lesion where the status is "Unequivocal progression".
- Date of PD for new lesion: If there is at least one new lesion at cycle i, the date of PD for new lesions is the earliest date of examination among the new lesions at cycle i.

1st clinical PD assessment
min (
Date of withdrawal if reason for withdrawal = "Clinical Progressive disease" or "Radiological and clinical progressive disease".
Date of progression in AE e-CRF page (Is this event related to disease progression).
Date of PD according to the patient'status filled in the follow-up e-CRF page.
)

Death date

min(
date of death from AE e-CRF page.
date of death from follow-up/status of the patient e-CRF page.
)

New anti-cancer therapy considered for efficacy analyses will be defined by:

- A concomitant medication with a pharmacological class, including antineoplastic agents, immunotherapy and endocrine therapy from codelist number = 3997.0.
- A concomitant medication based on all investigational drugs from codelist number = 3594.2.
- A further therapy reported in Post-study e-CRF pages.

Lost to follow-up

Patient lost to follow-up is defined by an answer "Lost to follow-up" and/or "Non-medical reason" related to consent withdrawal* at the eCRF interrogation "Reason for withdrawal from the study".

*related to consent withdrawal is defined as one of the two following subcategories:

- "Consent withdrawal from study treatment period and survival follow-up".

5.4. Proportional hazard assumption

The proportional hazard assumption across treatment arms will be checked.
The plot of the « log of the negative log of the estimated survivor function » (LLS) against log time will be done to provide an empirical check of the appropriateness of the proportional hazards model assumption. Under this assumption, the LLS curves should be approximately parallel across treatment groups. In addition, a Schoenfeld Residuals test will be performed to test whether the slope of scaled Schoenfeld residuals on time is zero or not. If the slope is not zero then the proportional hazard assumption has been violated. As there are certain types on non-proportionality that will not be detected by the tests of non-zero slopes alone but that might become obvious when looking at the graphs of the residuals, the graph of the scaled Schoenfeld residuals will be drawn. If necessary, sensitivity analyses other than those planned in the SAP could be carried out, in the framework of the validation of the assumptions underlying the model.

6. APPENDIX C: SAFETY

The **periods** considered **for the analysis** of adverse events, deaths and clinical laboratory evaluation are:

- During the treatment period:
  - With date ≥ first IMP intake date and <=last IMP intake date as defined in the Section 4.1.5 + 30 days.
- After the treatment period:
  - With date > last IMP intake date as defined in the Section 4.1.5 + 30 days.

6.1. Adverse events

Each **medical concept of adverse events coded according to the internal "multiple medical concept" process** is taken into account as a single adverse event in the statistical analysis.

The modalities of the adverse event (onset and end dates, severity, seriousness, action taken, additional therapy, relationship, outcome...) replicated by default to each medical concept are also taken into account in the statistical analyses.

Emergent adverse events on treatment are defined as all adverse events:

- Which occur between the first IMP intake date (included) and the last IMP intake date "+ 30 days″ (included), considering the definition of last IMP intake date defined in Section 4.1.5

or

- Which occur before the first IMP intake date and which worsen (in terms of severity) or become serious according to the investigator opinion between the first IMP intake date (included) and the last IMP intake date "+ 30 days″ (included), considering the definition of last IMP intake date defined in Section 4.1.5.

Note: Adverse events occurring or worsening or becoming serious on the day of the first IMP intake (if any) is considered as emergent.

Serious adverse events are defined as all adverse events upgraded by the sponsor during the IME or PharmacoVigilance (PV) process (upgrade of seriousness) or considered as "serious" from investigator assessment.
Serious adverse events from investigator assessment are defined as all adverse events fulfilling at least one of the following seriousness criteria for immediate notification: death, hospitalisation or prolongation of hospitalisation, medically important, life-threatening, disability/incapacity or congenital anomaly.

A fatal adverse event corresponds to an adverse event with “Fatal” as outcome.

Adverse events related to trifluridine/tipiracil correspond to adverse events with relationship forced by the sponsor during the PV process (upgrade of relationship for adverse events assessed as serious according to investigator or sponsor opinion) or considered as "related" from investigator assessment. In other cases, the adverse events are considered as "not related to trifluridine/tipiracil".
Adverse events related to study drug from investigator assessment correspond to adverse events associated with the answer "Related" to the "Is this event related to test drug (trifluridine/tipiracil)" question.

Adverse events related to Bevacizumab correspond to adverse events associated with the answer "Related" to the "Is this event related to the associated agent (Bevacizumab)" question.

Adverse events related to the combination (trifluridine/tipiracil + bevacizumab or trifluridine/tipiracil monotherapy) correspond to adverse events associated with the answer "Related" to the "Is this event related to test drug (trifluridine/tipiracil)" question and/or with the aswer “Related” to the "Is this event related to the associated agent (Bevacizumab)" question.

Adverse events related to disease progression correspond to adverse events associated with the answer "Related" to the "Is this event related to disease progression" question.

The following information will be taken into account:

- For the analyses where the severity of the adverse event is considered, the worst severity from the day of emergence and during the studied period (i.e. between the first IMPintake date (included) and the last IMP intake date (as defined in Section 4.1.5) “+ 30 days” (included) will be taken into account.
- The number of patient by worst grade: for each patient, system organ class and preferred term, we analyze the worst grade of events.
- The number of event by worst grade: for each patient, system organ class and preferred term, we analyse the worst grade of each event. The percentage will be calculated as the number of events of the PT concerned at grade X divided by the total number of events of the PT concerned.
- For the analyses where the action taken regarding the study drug and the additional therapy requirement are considered, all the actions taken and additional therapy requirements recorded from the day of emergence and on the studied period will be taken into account.
- However, in case of an episode of an emergent adverse event leading to studied treatment withdrawal reported after the last IMP intake date (as defined in Section 4.1.5) “+ 30 days”, the adverse event will be considered as leading to studied treatment withdrawal during the studied period.
- Seriousness is judged by event: if one episode is serious (whatever the time it occurs), the whole event will be considered as serious.
- A severe adverse event corresponds to an adverse event with NCI CTCAE grade 3, 4 or 5.
- Any grade corresponds to an adverse event with NCI CTCAE grade 1, 2, 3, 4 or 5.
- All multicoded events will be taken into account in the analyses.
- For the analysis of recovered emergent adverse event during / after treatment period, an EAE is considered as recovered "during treatment period" ("after end of treatment period", respectively) if the associated outcome is “recovered" or "recovered with sequelae" and occurs between the first IMP intake date and the last IMP intake date (as defined in Section 4.1.5) “+ 30 days” (included) (strictly after the last IMP intake date (as defined in Section 4.1.5) “+ 30 days”, respectively).

The following rules are applied in case of missing severity:

Table (6.1) 1 - Emergence of adverse events in case of missing severity/grade (if any)

Severity/Grade			Adverse event considered as
Nearest before the first study drug/associated agent intake date	**During the study period**		
Missing	Missing	→	Emergent
Missing	Grade 1	→	Non emergent
Missing	Grade 2, 3, 4 or 5	→	Emergent
Grade 1, 2, 3	Missing	→	Emergent
Grade 4	Missing	→	Non emergent

The rules for substitution of missing or incomplete episode date (onset date, dates of the *six* seriousness criteria and dates of change of severity and action taken) are as follows:

Table (6.1) 2 - Substitution rules of AE dates

Date to substitute	Substituted date (Episode date)
../mmm/yyyy	If same month and year than first IMP intake date then: **First IMPintake date** Else: **01/mmm/yyyy**
• ../…/yyyy	If same year than first IMP intake date then: **First IMPintake date** Else: **01/JAN/yyyy**
• ../.../….	**First IMP intake date**

Note: ../mm/yyyy = missing day
../.../ yyyy = missing day and month
../.../.... = missing date

The rules for substitution of missing or incomplete recovery dates, in case of AE outcome "recovered" or "recovered with sequelae" are as follows:

Table (6.1) 3 - Substitution rules of recovery date

Date to substitute	Substituted date
../mmm/yyyy	If same month and year than date of AE last information (*) then: **Date of AE last information** Else: **Last day of the month /mmm/yyyy**
../…/yyyy	If same year than date of AE last information (*) then: **Date of AE last information** Else: **31/DEC/yyyy**
../.../….	**Date of AE last information** (*)

Notes:
- *../mmm/yyyy = missing day,*

../.../ yyyy = missing day and month,
../.../.... = missing date.
(*) Date of AE last information is defined as the maximum between onset date, dates of change of severity and action taken and dates of the six seriousness criteria for this adverse event.

6.2. Death

Death information are taken on 'Adverse Event' (serious criteria for immediate reporting = Death) and 'Status of the patient at follow-up n° X' (reason for not continuing/completed the follow-up period = Death) e-CRF pages.
In case of death reported on 'Adverse Event' page, the date of death is compared to the date of last intake (as defined in Section 4.1.5) "+ 30 days" to classify occurrence on-treatment or during follow-up.

6.3. Clinical laboratory evaluation

Values
Only reliable values are considered for analyses. Unreliable values are flagged into the database.

In case of **multiple samples**:
- For the description of the values at each planned post-baseline visit, only the first analysable one measured under treatment at the visit is taken into account.
- Otherwise, each post-baseline value (test, re-test, planned, unplanned) measured under treatment is taken into account for analyses.

Units
All parameters will be analysed in international units (IU).

Abnormal values
Abnormal values are described according to:
- Reference laboratory ranges for the non-gradable parameters.
- CTCAE grade for gradable parameters.

Table (6.3) 1 - Gradable and non-gradable laboratory parameters

Laboratory parameters		Parameter	Highest	Lowest
Biochemistry	Gradable	Sodium	Hypernatremia	Hyponatremia
Biochemistry	Gradable	Potassium	Hyperkalemia	Hypokalemia
Biochemistry	Gradable	Magnesium	Hypermagnesemia	Hypomagnesemia
Biochemistry	Gradable	Serum creatinine	High creatinine	NA
Biochemistry	Gradable	Albumin	NA	Low albumin
Biochemistry	Gradable	Glucose	NA	Hypoglycemia
Biochemistry	Gradable	GGT	High GGT*	NA
Biochemistry	Gradable	AST	High AST*	NA
Biochemistry	Gradable	ALT	High ALT*	NA
Biochemistry	Gradable	Alkaline phosphatase	High alkaline phosphatase*	NA
Biochemistry	Gradable	Total bilirubin	High total bilirubin*	NA
Haematology	Gradable	Haemoglobin	High haemoglobin*	anemia
Haematology	Gradable	White blood cells	NA	Low WBC
Haematology	Gradable	Neutrophils	NA	Low neutrophils
Haematology	Gradable	Lymphocytes	High lymphocytes	Low lymphocytes
Haematology	Gradable	Platelets	NA	Low platelets
Coagulation	Non-gradable	INR	High INR	NA
Biochemistry	Non-gradable	Creatinine clearance**	NA	Low creatinine clearance
Biochemistry	Non-gradable	Chloride	High chloride	Low chloride
Biochemistry	Non-gradable	Blood urea nitrogen	High urea nitrogen	NA
Biochemistry	Non-gradable	Total calcium	High total calcium	Low total calcium
Biochemistry	Non-gradable	LDH	High LDH	NA
Biochemistry	Non-gradable	Phosphate	NA	Hypophosphatemia
Haematology	Non-gradable	Red blood cells	High RBC	Low RBC

**: Grade cannot be computed at baseline (grading depends on normality/abnormality of baseline value)*
***: Creatinine clearance is calculated in the ClinTrial database (by the data management department) if the data was not reported in the e-CRF.*

Laboratory reference limits and CTCAE grades are reported in the database.
For the parameters which are gradable according to CTCAE v5.0, the grades will be derived using the local laboratory reference limits in analysis datasets.

CL3-95005-007 - Statistical Analysis Plan | VV-TMF-103206 | 1.0

Table (6.3) 2 - Definition of CTCAE grade - Version 5.0 - Gradable parameters

Parameter	Grade 1	Grade 2	Grade 3	Grade 4
Sodium (mmol/L) - Hypernatremia	]ULN;150]	]150;155]	]155;160]	>160
Sodium (mmol/L) - Hyponatremia	[130;LLN[		[120;130[	< 120
Potassium (mmol/L) - Hyperkalemia	]ULN;5.5]	]5.5;6]	]6;7]	>7
Potassium (mmol/L) - Hypokalemia		[3.0;LLN[	[2.5;3.0[	<2.5
Calcium (mmol/L) - Hypercalcemia				
Magnesium (mmol/L) - Hypermagnesemia	]ULN;1.23]		]1.23;3.30]	> 3.30
Magnesium (mmol/L) - Hypomagnesemia	[0.5;LLN[	[0.4;0.5[	[0.3;0.4[	<0.3
Albumin (g/L) - Low	[30; LLN[	[20; 30[	< 20	
Glucose (mmol/L) - Hypoglycemia	[3.0;LLN[	[2.2;3.0[	[1.7;2.2[	< 1.7
ALAT increased	>ULN – 3.0 x ULN if baseline is normal [≤ULN]; 1.5 x baseline – 3.0 x baseline if baseline is abnormal [> ULN]	>3.0 x ULN – 5.0 x ULN if baseline is normal[≤ULN]; >3.0 x baseline – 5.0 x baseline if baseline is abnormal [> ULN]	>5.0 x ULN – 20.0 x ULN if baseline is normal [≤ULN]; >5.0 x baseline – 20.0 x baseline if baseline is abnormal [> ULN]	>20.0 x ULN x ULN if baseline is normal [≤ULN]; >20.0 x baseline if baseline is abnormal [> ULN]
ALP increased	>ULN – 2.5 x ULN if baseline is normal [≤ULN]; 2.0 x baseline – 2.5 x baseline if baseline is abnormal [> ULN]	>2.5 x ULN – 5.0 x ULN if baseline is normal [≤ULN]; 2.5 x baseline – 5.0 x baseline if baseline is abnormal [> ULN]	>5.0 x ULN – 20.0 x ULN if baseline is normal [≤ULN]; >5.0 x baseline – 20.0 x baseline if baseline is abnormal [> ULN]	>20.0 x ULN if baseline is normal [≤ULN]; >20.0 x baseline if baseline is abnormal [> ULN]
ASAT increased	>ULN – 3.0 x ULN if baseline is normal [≤ULN]; 1.5 x baseline – 3.0 x baseline if baseline is abnormal [> ULN]	>3.0 x ULN – 5.0 x ULN if baseline is norma [≤ULN]l; >3.0 x baseline – 5.0 x baseline if baseline is abnormal [> ULN]	>5.0 x ULN – 20.0 x ULN if baseline is normal [≤ULN]; >5.0 x baseline – 20.0 x baseline if baseline is abnormal [> ULN]	>20.0 x ULN if baseline is normal [≤ULN]; >20.0 x baseline if baseline is abnormal [> ULN]
GGT increased	>ULN – 2.5 x ULN if baseline is normal[≤ULN]; 2.0 x baseline – 2.5 x baseline if baseline is abnormal [> ULN]	>2.5 x ULN – 5.0 x ULN if baseline is normal [≤ULN]; >2.5 x baseline – 5.0 x baseline if baseline is abnormal [> ULN]	>5.0 x ULN – 20.0 x ULN if baseline is normal [≤ULN]; >5.0 x baseline – 20.0 x baseline if baseline is abnormal [> ULN]	>20.0 x ULN if baseline is normal [≤ULN]; >20.0 x baseline if baseline is abnormal [> ULN]

Table (6.3) 2 (Cont'd) - Definition of CTCAE grade - Version 5.0 - Gradable parameters

Parameter	Grade 1	Grade 2	Grade 3	Grade 4
Serum creatinine increased	>ULN – 1.5 x ULN	>1.5 x baseline – 3.0 x baseline (if baseline available) OR >1.5 x ULN – 3.0 x ULN	>3.0 x baseline (if baseline available) OR >3.0 x ULN – 6.0 x ULN	>6.0 x ULN
Total bilirubin increased	>ULN – 1.5 x ULN if baseline is normal [≤ULN]; 1.0 x baseline – 1.5 x baseline if baseline is abnormal [>ULN]	>1.5 x ULN – 3.0 x ULN if baseline is normal [≤ULN]; >1.5 x baseline – 3.0 x baseline if baseline is abnormal [>ULN]	>3.0 x ULN – 10.0 x ULN if baseline is normal [≤ULN]; >3.0 – 10.0 x baseline if baseline is abnormal [>ULN]	>10.0 x ULN if baseline is normal [≤ULN]; >10.0 x baseline if baseline is abnormal [>ULN]
Haemoglobin increased	> ULN and increase from baseline]0 ;20] (g/L)	> ULN and increase from baseline]20 ;40] (g/L)	> ULN and increase from baseline > 40 (g/L)	
Haemoglobin (g/L) - Anemia	[100;LLN[	[80;100[	< 80	
White Blood Cells (G/L) - Low	[3.0;LLN[	[2.0;3.0[	[1.0;2.0[	< 1.0
Neutrophils (G/L) - Low	[1.5;LLN[	[1.0;1.5[	[0.5;1.0[	<0.5
Lymphocytes (G/L) - Low	[0.8;LLN[	[0.5;0.8[	[0.2;0.5[	<0.2
Lymphocytes (G/L) - High		]4.0;20.0]	>20	
Platelets (G/L) - Low	[75;LLN[	[50;75[	[25;50[	<25

Any grade corresponds to a laboratory value with NCI CTCAE grade 0, 1, 2, 3 or 4.

A grade 0 corresponds to a laboratory value within limit or reference range.

The LOWEST value for a patient is defined as the lowest absolute laboratory value during the treatment period.
A Lowest value for a cycle is defined as the lowest laboratory value in that cycle.

The HIGHEST value for a patient is defined as the highest laboratory value during the treatment period.
A highest value for a cycle is defined as the highest laboratory value in that cycle.

Urinary results
The category 'Present' corresponds to results '+', '++', '+++' and '++++'.
The category 'More than one +' corresponds to results '++', '+++' and '++++'.

Worst value
For the urinalysis parameters, the worst class will correspond to 'Presence' class, then 'Trace' and finally 'Absence' will be considered as normal class.
For creatinine clearance, the worst class will correspond to "severe" category, then "moderate", then "mild" and finally "normal".
Baseline and post baseline worst grade for High and Low:

- Low: If the result is below lower normal range limit, Low = grade (for gradable parameter), Else Low = 0.

- High: If the result is above upper normal range limit, High = grade (for gradable parameter), Else High = 0.

The neutrophils-lymphocyte ratio (NLR) is defined as the absolute neutrophil count/absolute lymphocyte count

6.4. Vital signs and clinical examination

Worst value by cycle and overall
For ECOG, weight, BSA, SBP, DBP and HR, the worst (highest) value will be derived.
For weight, BSA, SBP, DBP and HR, the worst (lowest) value will be derived.

Last value under treatment
For ECOG, weight, BSA, SBP, DBP and HR, the last value under treatment will be derived.

Change from baseline
For weight, BSA, SBP, DBP and HR, the change from baseline to worst value at each cycle under treatment, to last post-baseline value under treatment and to worst (high and/or low) value by treatment arm will be derived.

Body Surface Area
BSA is calculated in the ClinTrial database (by data management department) using the height at inclusion visit and the weight at a corresponding cycle. This derived BSA will be used for compliance/exposure calculations as well as for vital signs and clinical examination analyses.
The BSA will be calculated using the following DuBois formula (all BSA calculations are rounded to 2 decimal places) (Du Bois, 1916).
BSA (m^2) = ([Body Weight (kg)] $^{0.425}$ x [Height (cm)] $^{0.725}$) x 0.007184

ECOG performance status

Table (6.4) 1 - Patient performance status

GRADE	*STATUS*
	ECOG
0	Normal unrestricted activity
1	Arduous physical activity restricted, but patient able to walk unaided and perform light work
2	Able to walk unaided and independent but unable to work more than half-time
3	Much less independent. Spends more than half his / her time in bed or seated.
4	Incapable of looking after him / herself. Completely confined to bed or to a chair

6.5. Electrocardiogram

Change from baseline to each post baseline visit = Post baseline absolute prolongation - Absolute prolongation at baseline

Change from baseline to last visit = Last absolute prolongation - Absolute prolongation at baseline.

CL3-95005-007 - Statistical Analysis Plan | VV-TMF-103206 | 1.0

7. APPENDIX D: QUALITY OF LIFE

7.1. QLQ C30 scores

The QLQ – C30 is a cancer specific instrument that contains 30 questions and provides a multi-dimensional assessment of health-related quality of life. The QLQ – C30 is composed of five functional scales, three symptom scales, six single items and a global health status.
The raw scores are linearly transformed to give standard scores in the range of 0-100. Score algorithm is provided below.

Higher scores in the global and functional scales and lower scores in the symptom/single scales indicate better quality of life.

Table (7.1) 1 - QLQ-C30 Score Algorithm

	Scale	Number of items	Item range*	**Version 3.0** Item numbers
Global health status / QoL				
Global health status/QoL (revised)†	QL2	2	6	29, 30
Functional scales				
Physical functioning (revised)†	PF2	5	3	1 to 5
Role functioning (revised)†	RF2	2	3	6, 7
Emotional functioning	EF	4	3	21 to 24
Cognitive functioning	CF	2	3	20, 25
Social functioning	SF	2	3	26, 27
Symptom scales / items				
Fatigue	FA	3	3	10, 12, 18
Nausea and vomiting	NV	2	3	14, 15
Pain	PA	2	3	9, 19
Dyspnoea	DY	1	3	8
Insomnia	SL	1	3	11
Appetite loss	AP	1	3	13
Constipation	CO	1	3	16
Diarrhoea	DI	1	3	17
Financial difficulties	FI	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.† (revised) scales are those that have been changed since version 1.0, and their short names are indicated in this manual by a suffix "2" – for example, PF2.*

QLQ C30 scores will be computed using the response to individual items noted I_x.

For the QLQ-C30, raw scores and scores will be calculated as follows using:

- For all scales, the Raw Score (RS), is the mean of the component items:
  $RS = (I_1 + I_2 + ... + I_n) / n$
- For Functional scales:
  $Score = \{1 - (RS - 1) / range\} \times 100$
- For Symptom scales, single items and global health status:
  $Score = \{(RS - 1) / range\} \times 100$

Global health status:
Score1: Global health status/QOL: $((I_{29}+I_{30})/2-1)/6 * 100$

Functional scales:

Score2: Physical functioning: $(1 - ((I_1+I_2+I_3+I_4+I_5)/5 -1)/3) * 100$
Score3: Role functioning: $(1 - ((I_6+I_7)/2-1)/3) * 100$
Score4: Emotional functioning: $(1 - ((I_{21}+I_{22}+I_{23}+I_{24})/4-1)/3) * 100$
Score5: Cognitive functioning: $(1 - ((I_{20}+I_{25})/2-1)/3) * 100$
Score6: Social functioning: $(1 - ((I_{26}+I_{27})/2-1)/3) * 100$

Symptom scales:

Score7: Fatigue: $((I_{10}+I_{12}+I_{18})/3-1)/3 * 100$
Score8: Nausea and vomiting: $((I_{14}+I_{15})/2-1)/3 * 100$
Score9: Pain: $((I_9+I_{19})/2-1)/3 * 100$

Single item scales:

Score10: Dyspnoea: $(I_8-1)/3 * 100$
Score11: Insomnia: $(I_{11}-1)/3 * 100$
Score12: Appetite loss: $(I_{13}-1)/3 * 100$
Score13: Constipation: $(I_{16}-1)/3 * 100$
Score14: Diarrhoea: $(I_{17}-1)/3 * 100$
Score15: Financial difficulties: $(I_{28}-1)/3 * 100$

For the QLQ-C30 analysis:

- QoL baseline questionnaire will be defined as the last questionnaire with at least one completed item prior to C1D1 included.
- QoL baseline score will be defined according to the QoL baseline questionnaire.

Scale scores in case of missing items

- Single item scales: in case of missing item, no scale score is computed.
- Multi-item scales: if at least half of the items from the scale have been answered, assume that the missing items have values equal to the average of those items which are present for that respondent. Then, the scale score is computed as described above.

Scheduled assessment time point

The quality of life of a patient at the time of a scheduled assessment will be based on the answers provided by the patient in the last questionnaire completed before or at the 1st administration of the following cycle according to the questionnaire completion date reported in the ePRO.
In case of unscheduled assessment, the QoL assessment will not be considered.

The description of reasons for non-completion at each scheduled time point will be based on the eCRF reported time point for the corresponding "Quality of life assessment" page.

Withdrawal questionnaire

The quality of life of a patient at withdrawal will be based on the answers provided by the patient in the first questionnaire completed at or after the withdrawal date.

7.2. EQ5D-5L

The EQ-5D 5L is a standardized health utility measure developed by the EuroQOL Group in order to provide a simple, generic measure of health for clinical and economic appraisal. It is composed of 5 questions focused on mobility, self-care, usual activities, pain/discomfort and anxiety/depression, and a Visual analogue scale (VAS). Each of the 5 questions has 5 levels: no problem, slight problems, moderate problems, severe problems and extreme problems.

The VAS records the respondent's self-rated health on a vertical visual analogue scale. The VAS 'thermometer' has endpoints of 100 (Best imaginable health state) at the top and 0 (Worst imaginable health state) at the bottom. This information can be used as a quantitative measure of health outcome as judged by the individual respondents.

The dimensional 5-level system will be converted into a single index utility score: values for the 3125 theoretically possible health states defined by the EuroQol classification are calculated using a regression model (Busschbach et al, 2003).

The index utility score will be derived according to Euroqol specific country value set. For all countries, the French Value Set will be used to generate health utility scores.

For the EQ-5D analysis:

- EQ-5D 5L utility score will be considered non evaluable when responses are missing for one or more of the dimensions.
- Baseline questionnaire will be defined as the last evaluable questionnaire prior to C1D1 included.
- Baseline score will be defined according to the baseline questionnaire.

EQ-5D 5L will be considered evaluable when questionnaire responses are completed for all the dimensions.

Scheduled assessment time point

The quality of life of a patient at the time of a scheduled assessment will be based on the answers provided by the patient in the last questionnaire completed before or at the 1st administration of the following cycle according to the questionnaire completion date reported in the ePRO.
In case of unscheduled assessment, the QoL assessment will not be considered.

The description of reasons for non-completion at each scheduled time point will be based on the eCRF reported time point for the corresponding "Quality of life assessment" page.

Withdrawal questionnaire

The quality of life of a patient at withdrawal will be based on the answers provided by the patient in the first questionnaire completed at or after the withdrawal date.

8. APPENDIX E: SOFTWARE AND PROGRAMMING CODES

Cox proportional hazard model stratifying for the stratification factors

```
PROC PHREG data = work.data;
        model TIME* CENSOR(censoredval) = ARM / RL TIES=efron;
        strata STRAT1 STRAT2 STRAT3 ;
RUN;
```

Stratified log-rank test

```
PROC LIFETEST data= work.data alpha=0.05 method=KM;
        time TIME *CENSOR (censoredval);
        strata STRAT1 STRAT2 STRAT3 / group=ARM;
RUN;
```

Cox proportional hazard model: stepwise procedure

```
PROC PHREG data = work.data;
        model TIME* CENSOR(censoredval)= STRAT1 STRAT2 STRAT3
        ALL_FACTORS /risklimits ties=efron selection=stepwise slstay=0.10 slentry=0.10
        include=3;
RUN;

PROC PHREG data = work.data;
        model TIME* CENSOR(censoredval)= ARM STRAT1 STRAT2 STRAT3
        SIG_FACTORS /risklimits ties=efron;
RUN;
```

Fisher's exact test

```
PROC FREQ data = work.data;
        tables ARM * VAR / fisher;
RUN;
```

Cochran-mantel-haenszel stratified analysis

```
PROC FREQ DATA= work.data;
        tables STRAT1*STRAT2*STRAT3 *ARM * VAR / CMH;
RUN;
```

Clopper Pearson CI

```
PROC FREQ data=work.data;
        tables RESPONSERATE / binomial (exact);
        by ARM;
RUN;
```

Kaplan Meier survival analysis

```
PROC LIFETEST data= work.data alpha=0.05 method=KM;
        time TIME *CENSOR (censoredval);
        strata ARM;
        id USUBJID;
RUN;
```

Mixed model repeated measures

```
PROC MIXED data= work.data;
        class ARM STRAT1 STRAT2 STRAT3 VISIT ;
        model CH_QoL_Score = BASE_QoL_Score ARM STRAT1 STRAT2 STRAT3
        VISIT VISIT*ARM;
        lsmeans ARM VISIT VISIT*ARM /cl pdiff adjust=tukey;
        repeated VISIT / subject=usubjid type=un or AR(1) or cs;
RUN;
```

9. REFERENCES

Guidelines

Guideline on Missing Data in Confirmatory Clinical Trials – Adopted by CHMP, June 2010, issued as EMA/CPMP/EWP/1776/99 Rev. 1.

ICH E3 - Structure and Content of Clinical Study Reports – Adopted by CPMP, December 1995, issued as CPMP/ICH/137/95/step 5.

ICH E3 Q&A - Structure and Content of Clinical Study Reports - Questions & Answers (R1) - July 2012, issued as EMA/CHMP/ICH/435606/2012.

ICH E9 - Statistical Principles for Clinical Trials - Adopted by CPMP, March 1998, issued as CPMP/ICH/363/96/step 5.

ICH E14 – The Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic drugs – Adopted by CHMP, May 2005, issued as CHMP/ICH/2/04.

CTCAE (Common Terminology Criteria for Adverse Events) version 5.0

Points to consider

Points to Consider on Adjustment for Baseline Covariates – Adopted by CPMP, May 2003, issued as CPMP/EWP/2863/99.

Points to Consider on Multiplicity Issues in Clinical Trials – Adopted by CPMP, September 2002, issued as CPMP/EWP/908/99.

Bibliography

Busschbach J. et al, "A comparison of EQ-5D time trade-off values obtained in Germany, the United Kingdom and Spain". In: Brooks R, Rabin R, de Charro F, editors. The measurement and valuation of health status using EQ-5D: a European perspective. Dordrecht: Kluwer Academic Publishers; 2003.

C. J. Clopper and E. S. Pearson, The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial *Biometrika* Vol. 26, No. 4 (Dec., 1934), pp. 404-413.

D. R. Cox, Regression Models and Life-Tables Journal of the Royal Statistical Society. Series B (Methodological), Vol. 34, No. 2. (1972), pp. 187-220.

Du Bois D, Du Bois EF. A formula to estimate the approximate surface area if height and weight be known. 1916. Nutr. Burbank Los Angel. City. Calif 1989;5:303-11.

Irwin, JO. The standard error of an estimate of expectation of life, with special reference to expectation of tumourless life in experiments with mice. Journal of Hygiene. 1949;47(2):188.

Kalbfleisch, J. D. and Prentice, R. L. (2002). The Statistical Analysis of Failure Time Data, 2nd edition. New York: John Wiley & Sons.

Kaplan E L, Meier P. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*.1958;53:457–481.

Mallinckrodt, C.H., Lin, Q. and Molenberghs M. A structured framework for assessing sensitivity to missing data assumptions in longitudinal clinical trials. Pharmaceutical Statistics, 2013, 12, 1-6. (reference associated with MMRM).

Mayer R., et al. "Randomised trial of TAS-102 for refractory metastatic colorectal cancer." The New England Journal of Medicine, 2015: 1901-1919.

New Response Evaluation Criteria in Solid Tumours: Revised RECIST guideline (version 1.1), Eisenhauer et al, 2009.

Osoba D, Rodrigues G, Myles J, ZEE B, PATER JInterpreting the significance of changes in health-related quality-of-life scores. J Clin Oncol 16: 139-144.

Tian, L., Zhao, L. & Wei, L. J. (2014). Predicting the restricted mean event time with the subject's baseline covariates in survival analysis. Biostatistics 15, 222-233.

CL3-95005-007 - Statistical Analysis Plan | VV-TMF-103206 | 1.0

Link to Avastin SmPC:
http://www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/000582/human_med_000663.jsp&mid=WC0b01ac058001d124.