STASTISTICAL ANALYSIS PLAN

A randomised, open-label, Phase II, dose/schedule optimisation study of NUC-3373/leucovorin/irinotecan plus bevacizumab (NUFIRI-bev) versus 5-FU/leucovorin/irinotecan plus bevacizumab (FOLFIRI-bev) for the treatment of patients with previously treated unresectable metastatic colorectal cancer

IMP	NUC-3373
Protocol Number	NuTide:323
SAP Version, Date	Version 2.0, 17 April 2024
Development Phase	II
IND Number	135275
EudraCT Number	2022-001459-17
ClinicalTrials.gov ID	NCT05678257
Sponsor:	**NuCana plc** 3 Lochside Way Edinburgh, EH12 9DT, UK Tel: +44 (0)131 357 1111 Email: NuTide323@nucana.com

Statistical Analysis Plan

Protocol No:	NuTide:323
EudraCT No:	2022-001459-17
Title:	A randomised, open-label, Phase II, dose/schedule optimisation study of NUC-3373/leucovorin/irinotecan plus bevacizumab (NUFIRI-bev) versus 5-FU/leucovorin/irinotecan plus bevacizumab (FOLFIRI-bev) for the treatment of patients with previously treated unresectable metastatic colorectal cancer
Development Phase:	II
Investigational Product:	NUC-3373
Sponsor:	NuCana plc. 3 Lochside Way, Edinburgh, EH12 9DT, UK

Author:	[redacted], KJC Statistics Ltd.
SAP Version	Version 2.0, FINAL
SAP Date	17 Apr 2024
Apr-24-2024 | 5:23 PM BST	
Based on Protocol	Version 2.0, 26 Sep 2023

The undersigned certify that they have read, reviewed and approved this document:

Author: [redacted] Signature: [redacted]

Date:

NuCana plc: [redacted] Signature: [redacted]

Senior VP Medical & Clinical

Development Date:

Table of Contents

Contents

List of Tables

List of Figures

Version History

Version	Description of Changes	Version Date	Timing of Changes
1	N/A	09 Oct 2023	N/A
2	The end of study definition has been clarified Derived visit windows have been updated to align with updates made in the study protocol Colorectal cancer history and current disease status and prior/subsequent colorectal cancer therapy parameters have been updated to align with updates made in the CRF and study protocol Details of the calculation for 'time from the end of the last prior line of therapy to the start of NuTide:323' have been added Treatment status definitions have been added to the adverse event section Lists of laboratory paramaters with and without CTCAE grading have been updated The ECG analyses have been updated to align with the data collected and streamline the presentation The study schedule of events has been updated to align with the updated study protocol	17 Apr 2024	During study conduct, prior to database lock

Abbreviations

Abbreviation	Term
5-FU	5-fluorouracil
AE	Adverse Event
AUC	Area under the plasma concentration-time curve
Bev	Bevacizumab
BMI	Body Mass Index
BOR	Best Overall Response
BRAF	v-Raf murine sarcoma viral oncogene homolog B
BSA	Body surface area
CI	Confidence Interval
Cinf	Concentration at the end of infusion
CL	Apparent clearance
C_{max}	Maximum plasma concentration; peak plasma concentration
COVID-19	Coronavirus Disease 2019
CR	Complete response
CRC	Colorectal cancer
CRF	Case report form
CRO	Clinical Research Organisation
CSR	Clinical study report
CT	Computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
DCR	Disease control rate
DLT	Dose-limiting toxicity
DOR	Duration of Response
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic Case Report Form
EDC	Electronic data capture
EDOR	Expected DOR
ÊDOR	Estimation of Expected DOR
EMA	European Medicines Agency
EudraCT	European Union Drug Regulating Authorities Clinical Trials Database
FAS	Full analysis set
FDA	(US) Food and Drug Administration
FOLFIRI	5-FU + LV + irinotecan
FOLFOX	5-FU + LV + oxaliplatin
FOLFOXIRI	5-FU + LV + oxaliplatin + irinotecan
HR	Hazard Ratio
ICF	Informed consent form
IV	Intravenous(ly)
KRAS	Kirsten rat sarcoma virus
LV	Leucovorin
MedDRA	Medical Dictionary for Regulatory Activities
mFAS	Modified full analysis set
MI	Multiple Imputation
MRI	Magnetic resonance imaging
MTD	Maximum tolerated dose
NCI-CTCAE	National Cancer Institute-Common Terminology Criteria for Adverse Event

Abbreviation	Term
NE	Not evaluable
NRAS	Neuroblastoma RAS viral oncogene homolog
NUFIRI	NUC-3373 + LV + irinotecan
NUFOX	NUC-3373 + LV + oxaliplatin
ORR	Objective Response Rate
OS	Overall survival
PD	Progressive Disease
PFS	Progression-Free Survival
PH	Proportional Hazards
PK	Pharmacokinetics
PR	Partial Response
PT	Preferred Term
Q1W	Weekly
Q2W	Alternate weekly (fortnightly)
QoL	Quality of Life
QT/QTc	QT interval / Corrected QT interval
RE	Response Evaluable
RECIST	Response Evaluation Criteria in Solid Tumours (version 1.1)
SAE	Serious adverse event
SAP	Statistical analysis plan
SARS-CoV-2	Severe acute respiratory syndrome-related coronavirus 2
SD	Stable Disease
SI units	International System of Units
SOC	System Organ Class
SS	Safety set
t1/2	Terminal half-life
TEAE	Treatment-emergent adverse event
UGT1A1	UDP glucuronosyltransferase 1A1
ULN	Upper limit of normal
Vd	Volume of distribution
WHO DD	World Health Organisation Drug Dictionary

1. Introduction

This document presents the Statistical Analysis Plan (SAP) for protocol NuTide:323: A randomised, open-label, Phase II, dose/schedule optimisation study of NUC-3373/leucovorin/irinotecan plus bevacizumab (NUFIRI-bev) versus 5-FU/leucovorin/irinotecan plus bevacizumab (FOLFIRI-bev) for the treatment of patients with previously treated unresectable metastatic colorectal cancer.

The SAP provides the description of the final analysis for the primary and secondary endpoints. The analysis for the single exploratory endpoint will be described separately. The details pertaining to the potential interim analysis will be captured in a separate addendum to this SAP.

2. Protocol Details

2.1 Study Objectives

Primary Objective

The primary objectives are to:

- Compare progression-free survival (PFS) of NUC-3373 in combination with leucovorin (LV), irinotecan and bevacizumab (NUFIRI-bev) (two dosing schedules) with 5-fluorouracil (5-FU) in combination with LV, irinotecan and bevacizumab (FOLFIRI-bev)
- Determine the optimal NUFIRI-bev dosing schedule

Secondary Objectives

The secondary objectives are to:

- Compare the efficacy of NUFIRI-bev to FOLFIRI-bev in terms of:
  - Objective response rate (ORR)
  - Duration of response (DoR)
  - Disease control rate (DCR)
  - Maximum percentage change in tumour size
  - Overall survival (OS)
- Assess the safety and tolerability of NUFIRI-bev compared to FOLFIRI-bev
- Assess the pharmacokinetics (PK) of NUFIRI-bev

Exploratory Objective

The exploratory objective is to determine if there are tumour cell characteristics that may further elucidate the mechanisms through which the clinical activity of NUC-3373 is achieved.

2.2 Overall Study Design

NuTide:323 is a randomised, open-label, dose/schedule optimisation study of NUFIRI-bev versus FOLFIRI-bev for the treatment of patients with unresectable metastatic colorectal cancer (CRC). Two NUFIRI-bev dosing schedules will be

assessed, namely Q1W NUC-3373 + LV (abbreviated as NUFIRI-bev Q1W) and Q2W NUC-3373 + LV (abbreviated as NUFIRI-bev 2QW). Irinotecan and bevacizumab will be given Q2W as per SoC as stated in the protocol.

The overall aim of the study is to estimate the efficacy and safety of the two NUFIRI arms as compared to the FOLFIRI control arm.

A total of 171 patients will be randomised 1:1:1 (57 patients per arm) to either NUFIRI-bev Q1W, NUFIRI-bev Q2W, or FOLFIRI-bev Q2W. Randomisation is stratified by RAS status (wild-type vs KRAS mutant vs NRAS mutant), prior bevacizumab treatment (yes vs no) and duration of prior line of therapy (<6 months vs ≥6 months).

Bevacizumab may be substituted with any approved biosimilar consistent with institutional practice and the specific agent administered will be recorded.

Patients may continue to receive treatment in the absence of disease progression or unacceptable toxicity that is not ameliorated by optimal medical or non-medical supportive or prophylactic care. All patients will be followed up on an intent-to-treat basis until withdrawal of consent, lost to follow-up, death, or the overall end of study has been reached, whichever occurs first.

The primary efficacy analysis will be performed when a total of 139 progression or death (i.e. PFS) events have occurred. Following accrual of 139 PFS events, the study will be considered complete once the final patient has completed their End of Treatment visit.

An evaluation of efficacy may be performed 3 months after the last patient has been randomised, at which time a total of 70 PFS events are expected to have occurred.

Based on the hypothesised median PFS time for each of the three study arms (see Section 2.3, Sample Size Determination), the estimated inclusion period is 13 months; the estimated time from last patient randomised to 70 PFS events is 3 months and from last patient randomised to 139 PFS events is 18 months.

The study design is summarised in Figure 1 and the study schedule is provided in Appendix 1.

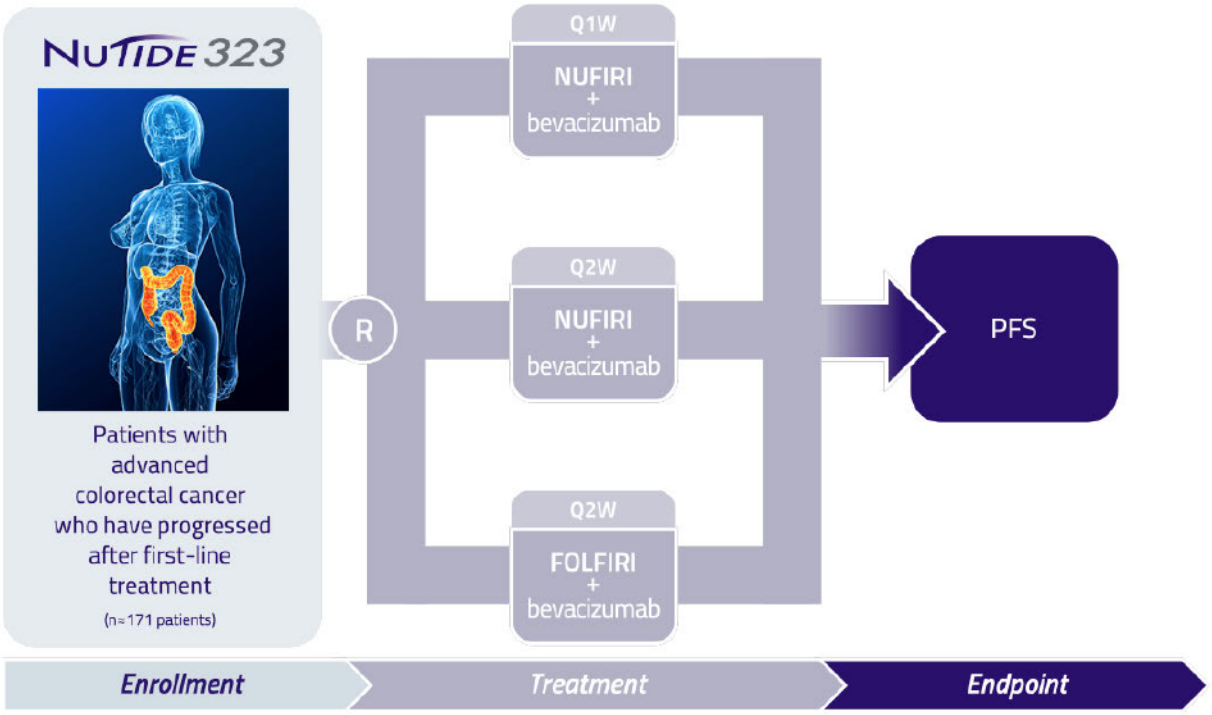
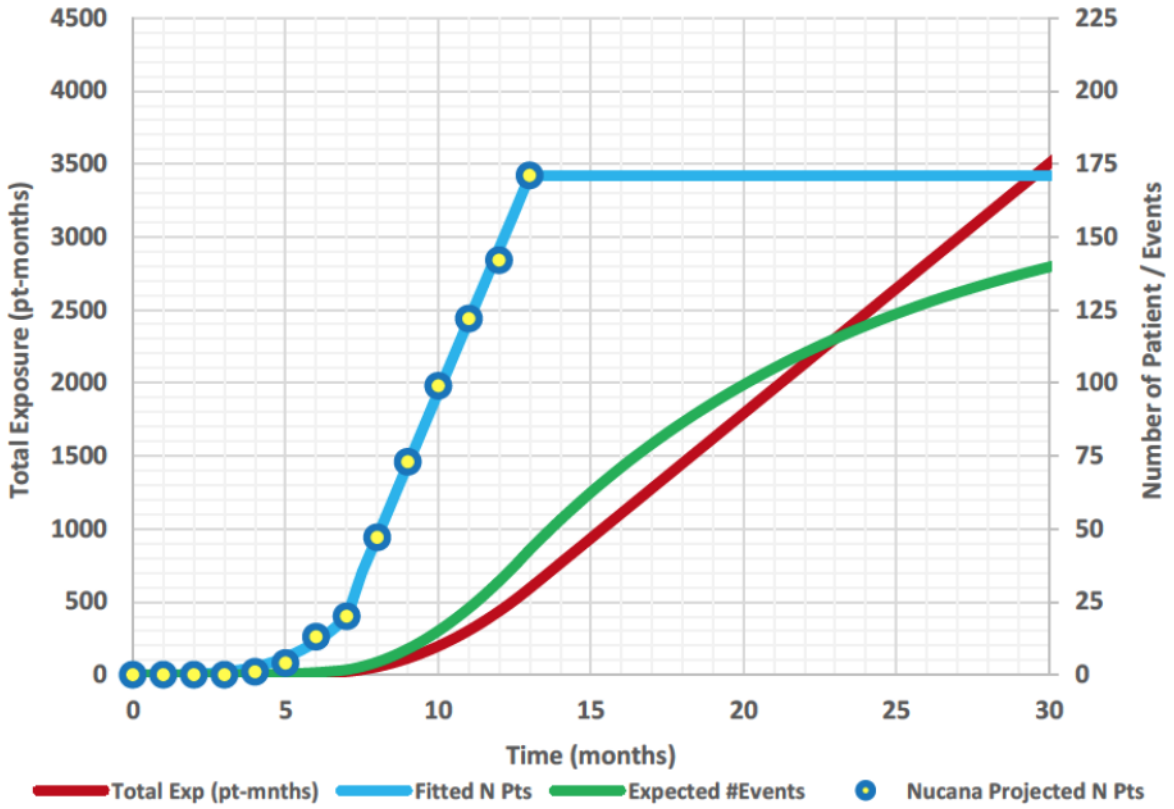


Figure 1: NuTide:323 Study Design

2.3 Sample Size

In total, 171 patients will be randomised on a 1:1:1 basis (57 patients per arm) to either NUFIRI-bev Q1W, NUFIRI-bev Q2W or FOLFIRI-bev Q2W.

The principal statistical objective of this study is to estimate the likely efficacy of the two NUFIRI arms as compared to the FOLFIRI control arm to support decision making regarding the further development of NUFIRI. Median PFS is expected to be 7 months on the FOLFIRI control arm and at least 9.9 months on each of the two NUFIRI arms. As displayed in Figure 2, assuming a non-linear recruitment profile over the planned 13-month accrual period and with a minimum of 17 months follow-up post-accrual, a total of 139 PFS events are expected across the three randomised arms. With this amount of information, this study will provide an 80% probability of correctly concluding superiority when NUFIRI is truly better than FOLFIRI in terms of PFS and, similarly, an 80% probability of correctly concluding non-superiority when NUFIRI is truly the same as FOLFIRI in terms of PFS. The smallest observed improvement in median PFS for either NUFIRI arm relative to FOLFIRI to conclude NUFIRI is truly better than FOLFIRI is 1.3 months.



1. The cumulative distribution of subject accrual time, $F(t)$ is modelled as:
   $F(t) = 0\ if\ t = 0;\ (t/13)^{3.5}$ if $t \le 7;\ 1.88(t/13) - 0.88$ if $7 < t \le 13;\ 1\ if\ t > 13$
2. The probability of an event at a given mean follow-up time is as per Carroll (2009).

Figure 2: Expected Patient and PFS Event Accrual Over Time[1,2]

Further, an evaluation of efficacy may be performed 3 months after the last patient has been randomised. At this time a total of 70 PFS events are expected across the three randomised arms. With this amount of information, the smallest observed improvement in median PFS for either NUFIRI arm relative to FOLFIRI to conclude NUFIRI is truly better than FOLFIRI is 1.9 months.

2.4 Timing of Analyses

It should be noted that this study is event driven, and not calendar-based, so that the timing of the principal analysis, while expected to occur 18 months after the last patient is randomised, is dependent upon the attainment of 139 PFS events.

Similarly, while an evaluation of the data may be performed 3 months after the last patient is randomised, the timing of this analysis will not be calendar driven but rather based on the accrual of 70 PFS events. The details pertaining to this analysis will be captured in a separate addendum to this SAP.

3. Analysis Populations

The analysis populations for this study are detailed below.

3.1 Full Analysis Set (FAS)

The Full Analysis Set (FAS) is based on the intention-to-treat principle and includes all randomised patients regardless of any early cessation of randomised treatment or any intercurrent events interfering with the determination of response, progression and/or death. The FAS will be the principal population for the assessment of efficacy endpoint data. Demographic and baseline characteristics will also be summarised in the FAS. Patients will be analysed based on randomised treatment.

3.2 Modified Full Analysis Set (mFAS)

The Modified Full Analysis Set (mFAS) is defined as a subset of the FAS who had at least one dose of randomised study treatment and had a baseline tumour assessment as well as at least one follow-up tumour assessment. The mFAS will be used to perform a supportive analysis of efficacy endpoint data. Patients will be analysed based on randomised treatment.

3.3 Safety Set (SS)

Safety will be assessed in the Safety Set (SS), defined as all randomised patients who received at least one dose (or partial dose) of study therapy (any administered treatment from either the NUFIRI arms or the FOLFIRI control arm). Patients will be analysed by treatment actually received, based on their first dose of randomised study drug.

4. Efficacy and Safety Endpoints

4.1 Primary Efficacy Endpoint

The primary efficacy endpoint is PFS as determined by RECIST v1.1 and defined as the time from randomisation to the first observation of objective tumour progression or death from any cause. PFS will be assessed in the FAS.

PFS time will be calculated in months as follows:

PFS = (date of event/date censored – date of randomisation + 1) / 30.4375

Patients who have not experienced disease progression or died at the time of analysis will be censored at the time of the latest date of assessment from their last evaluable RECIST v1.1 assessment.

Patients who start another anti-cancer therapy prior to progression will be censored at the date of the last evaluable RECIST v1.1 assessment.

The PFS time will always be derived based on imaging/assessment dates rather than cycle / study day or visit dates and the following rules will be applied:

- Date of disease progression will be determined based on the earliest of the dates of the component that triggered the disease progression, *i.e.,* if both the target lesions and the non-target lesions indicate disease progression but

were scanned on different days, the earlier of the 2 dates would be applied

- When censoring a patient for PFS the patient will be censored at the latest of the imaging/assessment dates contributing to a particular overall visit assessment

The definition of date of PFS or censoring refers to Food and Drug Administration (FDA) and European Medicines Agency (EMA) guidelines for PFS [2][2][3] and will be defined as shown in Table 1.

Table 1: Event and Censoring Times for PFS

Situation	Date of Event or Censoring	Outcome
Incomplete or no baseline tumour assessments	Randomisation	Censored
Complete baseline tumour assessment but no further tumour assessments	Randomisation	Censored
Progression documented based on RECIST v1.1 between scheduled visits	Date of progression based on RECIST v1.1	Event
Progression after 1 missing RECIST v1.1 assessment	Date of progression based on RECIST v1.1	Event
Progression after ≥2 consecutive missing RECIST v1.1 assessments	Date of last prior evaluable radiological assessment/date of progression based on RECIST v1.1	Event
Radiation therapy or surgical removal of non-target lesions and progression is documented based on RECIST v1.1	Date of progression based on RECIST v1.1	Event
Radiation therapy or surgical removal of target lesions and progression is documented based on RECIST v1.1	After the radiation therapy, earliest date between date of 'not evaluable' assessment and date of progression based on RECIST v1.1.	Event
Radiation therapy or surgical removal of target lesions and 'not evaluable' is documented based on RECIST v1.1	Date of last 'not evaluable' assessment	Censored
New systemic anticancer therapy prior to progression disease per RECIST v1.1	Date of last assessment before the start of the new anticancer therapy	Censored

Patients without any progression at 139 PFS endpoints	Date of last assessment without progression	Censored
Death at any time prior to progression disease per RECIST v1.1	Date of death	Event

4.2 Secondary Efficacy Endpoints

Secondary efficacy endpoints are listed below. All secondary efficacy endpoints will be assessed in the FAS.

4.2.1 Objective Response Rate

ORR is defined as the number of patients achieving a response (CR [complete response] or PR [partial response]), defined in accordance with RECIST v1.1 [5].

The ORR, the CR or PR as Best Overall Response (BOR), of each patient will be derived according to RECIST v1.1 criteria. The BOR is the best response recorded during the study period and up to the earliest of disease progression, initiation of subsequent anti-cancer therapy or death. No confirmed response is required.

At each tumour assessment (imaging), the Overall Response will be assessed from the evaluation of target lesions, non-target lesions, and appearance of new lesions at the given visit (Table 2).

Table 2: Overall Response Assessment for ORR

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

CR=Complete Response, PR=Partial Response, SD=Stable Disease, PD=Progressive Disease and NE=Not Evaluable

Assessment of Overall Response will be either CR, PR, Stable Disease (SD), Progressive Disease (PD) or Not Evaluable (NE).

When SD is believed to be the BOR, it needs to be assessed a minimum of 8 weeks (+/- 7 days) after randomisation. Otherwise, the BOR will be NE, unless any PD was further documented, in which case BOR will be PD.

4.2.2 Duration of Response (DoR)

DoR is defined for the subset of the FAS categorised as responders for the assessment of ORR. DoR is defined as the time, in months, from the time measurement criteria are first met for CR or PR (whichever is first recorded) until the first date that progressive disease is objectively documented (taking as reference the nadir recorded since the treatment started) or death.

DOR time in months will be calculated as:

DOR (months) = ([date of progression or death – date of first CR or PR] + 1) /30.4375

For patients who were lost to follow-up without progression or reached the time point of analysis without a known record of death or progression, the DoR will be censored at the date of last RECIST v1.1 tumour assessment.

4.2.3 Disease Control Rate (DCR)

DCR is defined as the number of patients achieving a response (CR and PR) or SD as a best overall response.

4.2.4 Duration of Stable Disease

Stable disease is defined for the subset of the FAS population categorised as having neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for progressive disease. The duration of SD is defined as the time, in months, from the time measurement criteria are first met for SD until the first date that progressive disease is objectively documented (taking as reference the nadir recorded since the treatment started) or death. For patients who were lost to follow-up without progression or reached the time point of analysis without a known record of death or progression, the duration of SD will be censored at the date of last tumour assessment.

4.2.5 Overall Survival (OS)

OS is defined as the time from randomisation to the time of death due to any cause. For patients who are alive at the time of analysis, or are permanently lost to follow-up, duration of OS will be censored at the date at which they were last known to be alive. OS time in months will be calculated as:

OS (months) = ([date of death or date patient last known to be alive – date of randomisation] + 1)/30.4375

The date at which the patient is last known to be alive is defined as the latest date of: (i) last site visit; (ii) last date at which the patient had a radiographic scan; and (iii) last date at which the patient, the study investigator, their other physicians, or a family member confirmed that the patient was alive.

4.2.6 Change from Baseline in Tumour Size

The percentage change from baseline in tumour size (%ΔTS_T) at 8-week intervals will be defined as follows:

- Baseline tumour size (TS_B): sum of longest diameters of target lesions at baseline;
- On-study TS (TS_T): sum of longest diameters of target lesions at a post-treatment disease assessment timepoint; then:

$$\%\Delta TS_T = \frac{TS_T - TS_B}{TS_B} \times 100$$

The best percentage change from baseline in tumour size (%ΔTS_X) across all timepoints is be defined as follows:

- Best TS (TS_X): smallest sum of longest diameters of target lesions observed at any timepoint, regardless of whether the assessment was scheduled or unscheduled, after first dose and prior to disease progression; then:

$$\%\Delta TS_X = \frac{TS_X - TS_B}{TS_B} \times 100$$

4.3 Safety Endpoints

Safety and tolerability will be assessed by evaluation of:

- TEAEs and serious adverse events (SAEs; per Common Terminology Criteria for Adverse Events [CTCAE] v5.0)
- Deaths due to TEAEs
- Treatment modifications due to TEAEs
- Clinically-significant laboratory changes (per CTCAE v5.0)
- Electrocardiograms

4.4 PK Endpoints

The PK of the NUFIRI-bev regimen will be assessed, including:

- Concentration at end of infusion (C_{inf})
- C_{max}
- AUC
- $t_{1/2}$
- Volume of distribution (Vd)
- Clearance (CL)

The analytes measured in plasma are:

- NUC-3373, CPF-1027, FBAL, irinotecan, SN-38, SN-38G and APC

4.5 Exploratory Endpoints

Exploratory analyses regarding population or tumour characteristic subtypes that may determine benefit to NUC-3373 treatment will be captured in a separate SAP.

5. Data Handling

5.1 Time Points and Visit Windows

Cycle 1 Day 1 is defined as the day of first dose of randomised treatment.

Relative days after Day 1 are calculated as (assessment date – Day 1 date) + 1.
Relative days prior to Day 1 are calculated as (assessment date – Day 1 date).

The day prior to Day 1 is Day 0.

For statistical summary by cycle/visit, data will be assigned to a derived visit window as described in Table 3. If more than one eligible result is reported within the same window, the mean result will be used for the descriptive statistics and the worst value (most conservative) result will be used in shift or summary tables for abnormalities.

Table 3: Derived Visit Windows

Planned Visit	Target Visit	Target Day	Starting Hour/Day for Visit Window	Ending Hour/Day for Visit Window
Screening	Screening	0	Day -28 to 0	Day 0†
Cycle 1 Day 1	Cycle 1 Day 1	1	> 0 hr of Cycle 1 Day 1	<0 hr of Cycle 1 Day 2
Cycle 1 Day 8	Cycle 1 Day 8	8	> 0 hr of Cycle 1 Day 6	<0 hr of Cycle 1 Day 10
Cycle 1 Day 15	Cycle 1 Day 15	15	> 0 hr of Cycle 1 Day 13	<0 hr of Cycle 1 Day 17
Cycle 1 Day 22	Cycle 1 Day 22	22	> 0 hr of Cycle 1 Day 20	<0 hr of Cycle 1 Day 24
Cycle 2 Day 1	Cycle 2 Day 1	29	27	31
Cycle 2 Day 8	Cycle 2 Day 8	36	34	38
Cycle x Day y	Cycle x Day y		> 0 hr of Cycle x Day $y-2$	<0 hr of Cycle x Day $y+2$
End Of Treatment	30 Days Post Last Dose		≥ 30 Days Post Last Dose	≤ 37 Days Post Last Dose
Follow-up	Q4 Weeks (France only)		≥ 25 Days Post Last Dose	≤ 31 Days Post Last Dose
Follow-Up	Q8 Weeks		≥ 49 Days Post Last Dose	≤ 63 Days Post Last Dose
Follow-Up	Q12 Weeks		≥ 70 Days Post Last Dose	≤ 98 Days Post Last Dose

Assessment and visits performed after EoT will be capured in the Follow Up period.

† Day 0 24-hour clock ends at first administration of study drug (i.e. 1 hour prior to first dose in the same calendar day is still study day 0).

Tumour measurements and disease response assessments will be performed every 8 weeks (±7 days) from Cycle 1 Day 1. If the patient stops study treatment for reasons other than radiologically-confirmed progressive disease, tumour measurements and disease response assessments will continue every 8 weeks (±7 days) from Cycle 1 Day 1 until progressive disease is radiologically confirmed.

No visit windows will be applied for summaries and analyses of tumour imaging data, these data will be summarised based on the actual date of data collection.

5.2 Unscheduled Assessments

Unscheduled radiographic assessments, if performed during the study, will be listed and included in the determination of response and disease progression endpoints.

5.3 Handling of Dropouts and Missing Data

Missing data

In general, missing data will remain missing and will not be included in data summaries. Exceptions are described below.

Missing tumour assessment data:

Patients who withdraw consent to follow-up may do so prior to having experienced progressive disease. The impact of such patients on the efficacy analysis will be explored via multiple imputation and tipping point analyses.

Patients who have missing tumour assessment data leading to missing objective response data will be included in the ORR analysis as non-responders.

Missing baseline data:

If a baseline value is not available and a screening value is available for the same parameter, then the last screening value will be used as baseline. This value will also be used for calculations of changes from baseline. Unless otherwise defined, baseline will be defined as Cycle 1 Day 1.

Incomplete dates:

In all listings, incomplete dates will be presented as they have been recorded. However, for calculation / sorting / assignation based on dates, the following methods will be used:

- As general rule:
  - Incomplete start dates will be imputed as first day of the month/year
  - Incomplete end dates will be imputed as last day of the month/year
- The most conservative approach will be systematically considered (*i.e.,* if the onset date of an AE/concomitant medication is missing/incomplete, it is assumed to have occurred during the study treatment phase except if the partial onset date or other data [stop date, etc] indicates differently).
- A missing/incomplete date of medical history or disease diagnosis will be assumed to have occurred before any study treatment.
- If a partial date and the associated information do not allow a statement about the assignment to a group / category, all the possible groups / categories will be considered (*i.e.,* an AE could be assigned to several possible doses at event onset according to its partial onset date and stop date. Particularly an AE with missing start date will be assigned to each dose received before its end date).

Where possible, the derivations based on a partial date will be presented as superior inequalities (*i.e.,* for an AE started in MAY2021 after the administration performed on 30APR2021, the days since last administration will be "≥ 2". Similarly, the duration of ongoing AEs or medication will be "≥ xx" according to the start and last visit dates).

6. Statistical Methods

6.1 General Principles

All data collected in this trial will be documented using summary tables, figures, and patient data listings.

Time to event data (PFS, OS, DOR, DOSD) will be summarised in the appropriate populations using Kaplan-Meier estimates. The Cox proportional hazards regression model will provide the main analysis and treatment effect estimates in terms of hazard ratios (HRs), confidence intervals (CIs) and p-values. The proportional hazards (PH) assumption will be assessed graphically by means of a plot of log(-log(survival)) versus log of survival time. If the data suggest PH assumption may not hold, the HR and CI can still be estimated and interpreted as the ratio of average hazards over the trial follow-up period.

The Log rank test will be performed where specified as a sensitivity analysis.

2-sided p-values will be presented along with 2-sided 95% CI.

All data processing, summarisation and analyses will be performed using SAS Version 9.4 (or later).

Sample SAS code is listed in Appendix 2.

The following principles will be applied to all tables, figures and listings unless otherwise stated:

Principle	Value
Significant tests	2-sided p-values will be presented along with 2-sided 95% CI.
Treatment group labels and order presented	NUFIRI-bev Q1W NUFIRI-bev 2QW FOLFIRI-bev
Tables	Data in summary tables presented by treatment group, and visit (where applicable).
Listings	All data collected presented by treatment group, patient and visit (where applicable), unless otherwise specified.
Descriptive summary statistics for continuous variables	Number of patients/observations (n), mean, standard deviation, median, first and third quartiles (Q1;Q3), minimum, and maximum.
Descriptive summary statistics for categorical variables	Frequency counts and percentages [n (%)]
Denominator for percentages	Number of patients in the analysis population, unless stated otherwise in table shell(s)
Include "Missing" as category	Demographics and Other baseline characteristics only, when the number missing is greater than zero for at least one treatment group.
Display for 0 percentages	0

Principle	Value
Display to one more decimal place than collected value	Mean Median Quartiles (Q1, Q3) Geometric Mean Confidence interval
Display to two more decimal places than collected value	Standard deviation % Coefficient Variation
Limit of precision for Hazard ratio and corresponding confidence interval, Kaplan-Meier estimates, and corresponding median and confidence interval, Odds ratio and corresponding confidence interval	2 decimal places
Limit of precision for KM displays	3 decimal places
Date format	DDMMMYYYY

6.2 Patient Disposition

Patient disposition will be listed and summarised by treatment group and overall for the FAS population and will include the number and percentage of patients:

- Screened (only number of patients in overall group);
- Randomised;
- Randomised and not treated;
- Treated;
- Included in each study population (FAS, mFAS, SS).

In addition, the number and percentage of patients who complete the study and who discontinue the study early, including a breakdown of the primary reasons for study discontinuation, will be presented for the FAS population. Similarly, patients who complete or prematurely discontinue randomised treatment will be summarised, including a breakdown of the primary reasons for discontinuation.

Corresponding listings will be provided with appropriate level of detail.

A listing of patients in each analysis set and those incorrectly randomised will be displayed.

6.3 Protocol Deviations

As per ICON process, important protocol deviations data will be entered into our system of record (PSO). The study team and the Sponsor will conduct on-going reviews of the deviation data from PSO and the resulting set of evaluable patients throughout the study, adjusting the deviation criteria as seems appropriate. The evaluable patients set must be finalised at the post-freeze data review meeting (or earlier), prior to database lock.

The number and percentage of patients with at least one important protocol deviation will be summarised overall and by category of important protocol deviation by treatment group. For each category and deviation patients will only be counted once even if they had multiple deviations in the category or deviation.

The important and non-important protocol deviations will be listed for the FAS population.

Protocol deviations will be classified as "important" or "non-important". Important protocol deviations may include:

- Incorrect study drug dose, frequency, timing or method of drug delivery;
- Excessive number of assessments
- Non-compliance with study procedures;
- Unauthorised medication;
- Other, e.g., consent not signed.

6.4 Demographics and Baseline Characteristics

Demographic and baseline characteristics will be listed and summarised by treatment group and overall for the FAS population.

As a general rule, the baseline value of study assessment (*e.g.,* weight) will be defined as last available value collected prior to the first dose of randomised treatment.

Standard descriptive statistics will be presented for the continuous variables of:

- Age (years);
- Weight (kg);
- Height (cm);
- Body mass index (BMI) (kg/(m)^2) [calculated as (weight/(height)^2) where weight is in kg and height is in m].

The total counts and percentages of patients will be presented for the randomisation stratification factors as well as the categorical variables of:

- Sex (Female, Male);
- Age Group (>= 18 - < 65; >= 65 years);
- Race;

- Ethnicity;
- Country;
- Childbearing potential (yes, no);

All demographics and baseline characteristics will be presented in listings, by patient.

No formal tests of statistical significance will be performed on the demographic and baseline data.

6.4.1 Medical History

Medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) [Current Version]. All medical history will be listed, and the number and percentage of patients with any medical history as well as number of events will be summarised for the FAS population by system organ class (SOC) and preferred term (PT) for each treatment group and overall.

6.4.2 Colorectal Cancer History and Current Disease Status

All CRC history will be listed and summarised by treatment group and overall for the FAS population. Standard descriptive statistics will be presented for the continuous variables of:

- Time since initial diagnosis of CRC (months) [calculated as randomisation date – date of initial diagnosis of CRC)/365.25)*12 and reported to 1 decimal place(s)]. No imputation will be applicable for incomplete or missing initial disgnosis date.
- Primary tumour location (colon – right side; colon – left side; colon – unknown; rectum)

The total counts and percentages of patients will be presented for the categorical variables of:

- Initial diagnosis (metastatic, locally advanced or locally non-advanced)*
- Stage at initial diagnosis (Stage IA-C, Stage IIA-C, Stage IIIA-C, Stage IVA-C);
- Number of metastatic sites (1, 2, 3, ≥4). This is the number of non-target lesions plus the number of target lesions identified for each patient (not including colon or rectum);
- Liver metastases (yes, no). This count includes patients for whom liver has been selected as the location of a target and/or non-target lesion;
- BRAF mutation status (non-V600E mutant, V600E mutant, wild-type);
- KRAS mutation status (mutant, wild-type);
- UGT1A1 status (UGT1A1 *1/*1, UGT1A1 *1/*28, UGT1A1 *28/*28, other);
- NRAS status (mutant, wild-type);
- MSI status (MSI-L, MSI-H, MSS);
- MMR status (normal/negative, positive);
- DPD/TYMP status (mutation associated with known toxicity to fluoropyrimidines, mutation not associated with known toxicity to fluoropyrimidines, wild-type);

- ECOG performance status (0, 1)

* Metastatic = Stage IVA to IVC; Locally advanced = Stage IIC and Stage IIIA to IIIC; Locally non-advanced = Stage IA to IC and Stage IIA to IIB.

6.4.3 Prior and Concomitant Medications

Medications received prior to or concomitantly with randomised treatment will be coded using the current WHO Drug Dictionary [Anatomical Therapeutic Chemical (ATC) Classification codes].

Prior medications and concomitant medications are defined as follows:

- Prior medications are those taken within 28 days prior to randomisation with a stop date prior to the date of randomisation.
- Concomitant medications are those with a start date before, on or after the date of randomisation and a stop date on or after the date of randomisation or ongoing at the end of study

If a medication cannot be classified as "prior" or "concomitant" after applying imputation rules for missing/incomplete dates (see Section 5.3), it will be classified as concomitant.

Prior medications and concomitant medications will be listed together using anatomic therapeutic chemical (ATC) Level 4 and preferred term, and sorted by medication start date and medication name and summarised separately for the FAS population.

The number and percentage of patients using each medication will be displayed together with the number and percentage of patients using at least one medication within each ATC Level 4 and preferred term.

6.4.4 Prior Systemic Colorectal Cancer Therapy

All prior systemic CRC therapies will be listed and summarised by treatment group and overall for the FAS population. Standard descriptive statistics will be presented for the continuous variables of:

- Number of lines of systemic therapies administered;
- Time to most recent progression during or after line of treatment (months) [calculated as ((date of progression during or after line of treatment - start date of prior systemic CRC therapy administered)/365.25)*12 and reported to 1 decimal place(s)]. In case of missing or incomplete last stop date of prior systemic CRC therapy administered, time to progression during or after line of treatment will not be calculated;
- Time from the end of the last prior line of therapy to the start of NuTide:323 (months) [calculated as ((date of randomisation - stop date of last prior line of systemic CRC therapy administered)/365.25)*12 and reported to 1 decimal place(s)]. In case of missing or incomplete stop date of last prior line of systemic CRC therapy, time from the end of the last prior line of therapy to the start of NuTide:323 will not be calculated;

The total counts and percentages of patients will be presented for the categorical variables of:

- Prior neoadjuvant treatment (yes, no)
- Prior adjuvant treatment (yes, no)
- Prior treatment for metastatic disease (yes, no)
- Prior chemotherapy (yes, no)
    - 5-FU
    - Capecitabine
    - Oxaliplatin
- Prior targeted/biologic therapy (yes, no)
    - Bevacizumab
    - Ramucirumab
    - Cetuximab
    - Panitumumab
    - Pembrolizumab
    - Nivolumab
- Duration of prior line of therapy (<6 months, ≥6 months)
- Best overall response on prior line of therapy (CR, PR, SD, PD, NE, unknown)
- Prior radiotherapy (yes, no)
- Prior cancer surgery (yes, no)

6.4.5 Subsequent Anticancer Treatment

All systemic CRC therapy and other cancer therapies will be listed and summarised by treatment group and overall for the FAS population.

Standard descriptive statistics will be presented for the continuous variables of:

- Number of subsequent systemic CRC therapies administered.

The total counts and percentages of patients will be presented for the categorical variables of:

- At least one subsequent systemic CRC therapy administered after randomisation;
- Type of therapy (description of other therapy type will be listed);
- Therapy setting;
- Best overall response on subsequent line of therapy (CR, PR, SD, PD, NE, unknown)

Palliative radiotherapy, if applied, post-randomisation will be listed and summarised by treatment group and overall for the FAS population.

The total counts and percentages of patients will be presented for the categorical variables of:

- At least one radiotherapy administered after randomisation;
- Radiotherapy type (description of other radiotherapy type will be listed);
- Site location (description of other site location will be listed);
- Intent (description of other intent will be listed);
- Therapy setting;
- Modality type;
- Best overall response on subsequent radiotherapy (CR, PR, SD, PD, MR, SR, NE, unknown)

6.5 Extent of Exposure

NUFIRI-bev Q1W, NUFIRI-bev 2QW and FOLFIRI-bev will be summarised using categorical variables and descriptive statistics for the Safety Set.

The following variables will be summarised for each arm:

- The number of doses adjusted, including reasons for adjustment
- The number of infusions interrupted, including a summary of duration of interruption (hours) and the reasons for interruption

The following parameters will be calculated for each arm:

- The number of infusions received
- Study treatment duration (days): date of last dose – date of first dose + 1.
- Number of cycles: Total number of complete or partial treatment cycles the patient received.
- Duration of infusion: Completion time of infusion – start time of infusion; where infusions have been paused, the duration of the infusion will be computed as the sum of the start/stop times as captured in the CRF.

The following parameters will be calculated for LV, irinotecan, 5-FU and NUC-3373, as appropriate, for each arm:

- Dose intensity (mg/m^2/cycle): [Sum of actual doses received (mg/m^2) across all cycles] / (number of cycles received [total number of complete or partial treatment cycles the patient received])
- Average dose per infusion per patient: mean of (the actual dose received/BSA at that cycle) across all cycles.

And for bevacizumab, for each arm:

- Dose intensity (mg/kg/cycle): [Sum of actual doses received (mg/kg) across all cycles] / (number of cycles received)
- Average dose per infusion: mean of (the actual dose received/weight at that cycle) across all cycles.

6.6 Efficacy

6.6.1 General

The principal population for the summary and analysis of efficacy endpoint data is the FAS. Supportive analyses will be conducted in the mFAS.

6.6.2 Primary Efficacy Analysis

6.6.2.1 Progression-Free Survival

The number of patients with a PFS event or censoring will be presented by treatment group, along with the reasons for the event or censoring. PFS will be analysed via Cox regression modelling stratified for the randomisation stratification factors (RAS status, first-line treatment, and duration of prior line of therapy) and including a fixed effect term for randomised treatment. The hazard ratio will be estimated for each NUFIRI arm versus the FOLFIRI control arm, along with the associated CIs and 2-sided p-values. The data will also be displayed using Kaplan-Meier curves and median PFS times will be estimated along with the associated CIs using the Brookmeyer and Crowley method. PFS rates at 12 months will also be estimated from the Kaplan-Meier curves along with the associated CIs. Considering the randomisation stratification factors, in the event that one or more cells are found to contain zero PFS events, a supportive non-stratified Cox analysis will be performed and the resulting HRs, CIs and p-value presented.

6.6.3 Secondary Efficacy Analyses

6.6.3.1 Change from Baseline in Tumour Size

The percentage change from baseline in tumour size at 8-week intervals ($\%\Delta TS_T$) will be calculated for each patient by treatment group. The best percentage change from baseline in tumour size ($\%\Delta TS_B$) across all timepoints will also be calculated for each patient.

$\%\Delta TS_T$ and $\%\Delta TS_B$ will be presented graphically by treatment group using waterfall plots for presenting each patient's percentage as a separate bar with the bars ordered from the largest increase to the largest decrease. Reference lines at the +20% and −30% change in tumour size levels will be added to the plots, which correspond with the definitions of disease progression and PR, respectively.

If a patient with measurable disease has no evaluable post-dose target lesion data, then they will be excluded from the waterfall plot of $\%\Delta TS_T$ and $\%\Delta TS_B$.

In addition, spider plots by treatment group will be also provided, showing percent change in tumour size by patient.

Scheduled assessments for tumour size evaluations will be summarised in a table and all assessments will be listed.

Objective Response

The number and percentage of patients in each RECIST v1.1 overall response category (CR, PR, SD, PD or NE) will be summarised by treatment group at fixed time intervals (*i.e.,* by assessment point) in the FAS. Similarly, the numbers and proportions of the patients who had a BOR of PD, SD, PR or CR will be summarised by treatment group.

The ORR will be analysed using exact logistic regression. The model will include a class effect for randomised treatment and will be stratified for the randomisation stratification factors. The exact odds ratio will be estimated for each NUFIRI arm versus the FOLFIRI control arm, along with the associated CI and 2-sided p-value. Clopper–Pearson exact 95% CIs for the ORR will be provided by treatment group.

Disease Control Rate

The number of patients with DCR (*i.e.,* achieving a BoR of CR, PR or SD) will be summarised by treatment group in the FAS. The exact odds ratio will be estimated for each NUFIRI arm versus the FOLFIRI control arm, along with the associated CI and 2-sided p-value. Clopper–Pearson exact 95% CIs for the DCR will be provided by treatment group.

Duration of Response

The DOR will be summarised by treatment group in the subset of FAS patients with a CR or PR using Kaplan-Meier display. Median DOR and the associated 95% CI will also be provided by treatment group.

The DOR will be analysed via Cox proportional hazards regression model in the same manner as that described for PFS.

Duration of Stable Disease (DOSD)

The DOSD will be summarised by treatment group in the FAS using Kaplan-Meier display. Median DOSD and the associated 95% CI will be provided by treatment group.

The DOSD will be analysed via Cox proportional hazards regression model in the same manner as that described for PFS.

Overall Survival (OS)

OS will be analysed in the FAS via Cox proportional hazards regression model in the same manner as that described for PFS. The number of patients that have died will be summarised; the number censored will also be presented along with reasons for censoring.

The OS will be displayed by treatment group using Kaplan-Meier curves. Median OS with the associated 95% 2-sided CI will be presented by treatment group.

6.6.4 Sensitivity Analyses

Progression Free Survival

Several sensitivity analyses will be performed on the primary endpoint in the FAS:

a) A log rank test on PFS. Each NUFIRI arm will be compared to FOLFIRI using the log rank test via PROC LIFETEST in SAS; the associated 2-sided p-value will be extracted and presented.

b) In order to assess the potential impact of missing PFS endpoint data, multiple imputation will be used. Patients who are non-administratively censored for PFS prior to a given database cutoff date will be imputed. To achieve this, the methodology described by Carroll (2003) [6] will be used whereby a parametric survival distribution is fitted, by randomised treatment group, to the PFS data at a given data cutoff date, T. Using PROC LIFEREG in SAS, Weibull, LogNormal and LogLogisitic distributions will be fitted to the PFS data by treatment group and the model with the lowest AIC for a given treatment group is chosen; covariates for time since initial diagnosis of CRC, number of prior cancer therapies administered, RAS status, first-line treatment, and duration of prior line of therapy will be included in the models. For a patient non-administratively censored at time j prior to time T, the fitted parametric distribution is then used to impute a PFS time to event or censoring in the period T-j. The same procedure is applied to all patients non-administratively censored before time T. A complete dataset of PFS times to event/censoring is thus constructed and the primary endpoint analysis model is then applied to this dataset to arrive at a new, imputed log HR estimates and associated SEs for each NUFIRI group verses FOLFIRI. This process is repeated 100 times to provide 100 imputed log HR estimates and associated SEs for each NUFIRI group verses FOLFIRI. These log hazard ratio estimates will then be combined using Rubin's rules for a point estimate and standard error, as implemented in SAS proc MIANALYZE. The final resulting log HR estimates and associated CIs will be back transformed and presented along with the associated 2-sided p-values.

   This entire process will be repeated but with imputation of non-administratively censored patients in both NUFIRI arms and the FOLFIRI arm under the parametric survival distribution fitted to the FOLFIRI arm. This sensitivity analysis is therefore akin to 'jump to placebo' imputation for data missing not at random.

c) An additional sensitivity analysis using a tipping-point approach will be conducted to assess the possible impact of informative censoring. Tipping point analysis explores the influence of potentially informative censoring on the primary endpoint analysis by progressively penalising patients non-administratively censored in the NUFIRI arms. The goal is to find the level of penalisation that results in loss of statistical significance for either or both NUFIRI groups vs FOLFIRI on the primary endpoint.

The steps to perform the tipping point analysis are as follows:

1. PFS in non-administratively censored patients will be imputed using the same parametric distribution and imputation method described in b) above.
2. The imputed PFS times in the NUFIRI arms will be penalised by adding a small penalty of δ days (e.g., δ = 1) to each imputed patients' PFS time.
3. The primary endpoint analysis model will then be applied to the resulting PFS dataset.
4. Steps 1 and 3 are repeated 100 times to arrive at 100 HR estimates and associated SEs for each NUFIRI group versus FOLFIRI where non-administratively censored patients have been imputed in both arms and, additionally, patients on the NUFIRI arms have been penalised by an amount of δ days.
5. These 100 hazard ratio estimates are then combined using Rubin's rules for a point estimate and standard error, as implemented in SAS proc MIANALZE.
6. Steps 1-5 are repeated for 2δ, 3δ, …, until statistical significance is lost.

Overall Survival

All following sensitivity analyses performed on the primary endpoint will be also performed for OS:

a) A log rank test on OS. Each NUFIRI arm will be compared to FOLFIRI and the associated 2-sided p-value will be extracted and presented.

b) Sensitivity analysis based on multiple imputation.

c) Sensitivity analysis using the tipping-point approach.

Duration of Response

A sensitivity analysis will be performed on the FAS population, based on the expected duration of response as per Ellis *et al* [7]. For this analysis, Weibull, LogNormal and LogLogistic probability distributions will be considered, and fitted by treatment group, with the lowest AIC retained.

To do so, the LIFEREG procedure will be run for each of the distribution above with the treatment arm as covariate (see Appendix 2).

Once the best model has been selected, the estimated EDOR will be calculated using the formulas (4), (5) and (6) from Ellis *et al* (2008) as follows:

1. First, calculate the proportion of responders in each treatment group, among the randomised patients from the FAS.
2. Consider first NUFIRI-bev Q1W versus FOLFIRI-bev; estimate p_E and p_C:

$$p_E = \frac{r_E}{N_E} \; ; \; p_C = \frac{r_C}{N_C}$$

   where r_E and r_C are the number of NUFIRI-bev Q1W responders and FOLFIRI-bev responders respectively, and N_E and N_C are the corresponding number of randomised patients
3. Then, calculate the logarithm of the ratio of EDOR, noted R and computed as $R = \frac{p_E M_E}{p_C M_C}$ so that $\ln(R) = \ln\left(\frac{p_E}{p_C}\right) + \ln\left(\frac{M_E}{M_C}\right)$ where M_E and M_C are the estimated

mean of DOR for the NUFIRI-bev Q1W and FOLFIRI-bev treatment groups respectively.

4. This gives the formula (4) from Ellis *et al* (2008) using substituted estimates obtained from the data: $\ln(\hat{R}) = \ln\left(\frac{\widehat{p_E}}{\widehat{p_C}}\right) + \ln\left(\frac{\widehat{M_E}}{\widehat{M_C}}\right)$
5. And thus, the formula (5) as follows:

$$Var\left[\ln(\hat{R})\right] = \frac{1-\widehat{p_E}}{\widehat{N_E}\widehat{p_E}} + \frac{1-\widehat{p_C}}{\widehat{N_C}\widehat{p_C}} + \frac{1}{\widehat{M_E^2}}Var\left[\widehat{M_E}\right] + \frac{1}{\widehat{M_C^2}}Var\left[\widehat{M_C}\right]$$

6. As stated in Ellis *et al* (2008), the hypothesis that the EDoR is equal for NUFIRI-bev Q1W and FOLFIRI-bev treatment groups can be tested by:

$$H_0 : R = \frac{EDOR_E}{EDOR_C} = 1 \quad versus \quad H_1 : R = \frac{EDOR_E}{EDOR_C} \neq 1$$

using formula (6):

$$z = \frac{\ln(\hat{R})}{\sqrt{Var\left[\ln(\hat{R})\right]}}$$

as the test statistic and comparing to a standard normal (0, 1) distribution.

DOR, number of patients who progress without response, number of patients who respond and are then censored, number of patients who respond and then progress, EDOR and variance of EDOR will be summarised by treatment group. The test statistic z will be estimated and corresponding p-value will be presented alongside in the summary table.

The same process is then followed to compare NUFIRI-bev Q2W versus FOLFIRI-bev.

6.6.5 Exploratory Subgroup Analysis

Descriptive subgroup analyses for key prognostic factors such as sex, age and race will be performed if the primary PFS endpoint is met. In particular, subgroup analyses may be performed in patients with and without liver metastases (where patients with liver metastases are those for whom liver has been selected as the location of a target and/or non-target lesion; and patients without liver metastases are those for whom liver has not been selected as the location of a target and/or non-target lesion).

As the study is not powered to document efficacy in subgroups, these exploratory analyses will be descriptive in nature.

In the event that the primary endpoint is not met, subgroups may still be presented whilst recognizing no alpha remains to support statistical interpretation.

6.7 Safety

Safety data will include AEs, deaths, laboratory values (haematology and chemistry), physical examinations, vital signs and ECOG.

All safety analyses will be presented by treatment group in patients from the SS population.

6.7.1 Adverse Events

AEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) Current Version, and will be classified by MedDRA Preferred Term (PT) and System Organ Class (SOC). The severity of AEs will be recorded and graded by investigators using the National Cancer Institute- Common Terminology Criteria for Adverse Events (CTCAE) current version.

AEs will be considered treatment-emergent if they start on or after the time of the first dose of study treatment and up to 30 days after the last dose of study treatment. Non-treatment-emergent AEs will be included in the patient listings and flagged as such but will not be included in the summary tables. Where an AE date is partial or missing, and it is unclear whether the AE is treatment-emergent, the AE will be assumed to be treatment-emergent. Any AEs with missing severity will be classified as severe (Grade 3). Any missing causality, or outcome will not be imputed, but will be classified as unknown. Deaths that occur within 90 days after the last dose of study drug are defined as on-study deaths.

A patient with more than one occurrence of the same adverse event in a particular system organ class will be counted only once in the total of those experiencing adverse events in that particular SOC.

If a patient experiences the same adverse event with more than one relationship to study drug, the stronger causal relationship to study drug will be given precedence.

If causality is assessed as related or possibly related, the event will be considered as related and if assessed as not related, the event will be considered as not related in the summary tables and listings. Causality will be defined globally (whatever the treatment) and for each treatment.

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

For tables by NCI-CTCAE grade, if a patient experiences the same AE at more than one grade, the worst grade will be summarised.

An overview table will summarise the number and percentage of patients with at least one of the following treatment-emergent adverse events (TEAEs) including number of events, presented by treatment group:

- TEAEs, treatment-related TEAEs, Grade 3+ TEAEs, serious TEAEs, TEAEs leading to dose interruption, TEAEs leading to treatment discontinuation, TEAEs with an outcome of death, treatment-related Grade 3+ TEAEs, treatment-related serious TEAEs, treatment-related TEAEs with an outcome of death.

The following classes of AEs will also summarised by treatment group with number and percentage of patients with at least one event and the number of events:

- Summary of TEAEs by SOC and PT
- Summary of treatment-related TEAEs by SOC and preferred term
- Summary of TEAEs occurring in at least 10% of patients, sorted by all grades and Grade 3+ and in descending order of frequency (*i.e.,* most frequent event shown first). The order of frequency will be determined by the most frequent preferred term across all cohorts.Summary of Grade 3+ TEAEs by SOC and preferred term
- Summary of treatment-related Grade 3+ TEAEs by SOC and preferred term
- Summary of TEAEs leading to treatment interruptions by SOC and preferred term
- Summary of TEAEs leading to treatment discontinuation by SOC and preferred term for each study treatment
- Summary of TEAEs by SOC, preferred term and worst grade
- Summary of TEAEs by SOC, preferred term and worst-case relationship attribution
- Summary of Hy's law cases
- Summary of non-serious TEAEs by SOC and preferred term
- Summary of SAEs by SOC and preferred term
- Summary of treatment-related SAEs by SOC and preferred term
- Summary of SAEs by preferred term sorted in descending order of frequency
- Summary of TEAEs leading to death by SOC and PT
- Summary of treatment-related TEAEs leading to death by SOC and PT
- Summary of all deaths

No statistical comparisons of AEs between treatment groups will be performed.

Only TEAEs will be included in the adverse event and SAE tables.

Listings will be presented and sorted by treatment group, patient and AE number for all adverse events recorded during the study.

The following listings will be provided:

- All AEs, for each patient and event, listings will state the date of onset, study day, dose at onset, treatment status at onset (pre-treatment, ongoing or post-treatment), investigator's assessment of severity and relationship to study drug, action taken with study treatment, and outcome.
- SAEs with outcome of death along with the date of onset, study day, dose at onset, treatment status at onset (pre-treatment, ongoing or post-treatment),

Investigator's assessment of severity and relationship to study drug, and action taken with study treatment,

- All SAEs along with the date of onset, study day, dose at onset, treatment status at onset (pre-treatment, ongoing or post-treatment), date of resolution (if SAE is resolved), Investigator's assessment of severity and relationship to study drug(s), action taken with study treatment, and outcome
- AEs leading to discontinuation of randomised treatment, listed along with the date of onset, study day, dose at onset, treatment status at onset (pre-treatment, ongoing or post-treatment), Investigator's assessment of severity and relationship to study drug, action taken with study treatment, and outcome
- AEs leading to treatment interruption, listed along with the date of onset, study day, dose at onset, treatment status at onset (pre-treatment, ongoing or post-treatment), Investigator's assessment of severity and relationship to study drug, action taken with study treatment, and outcome

Note: with regards to 'treatment status at onset', status is: pre-treatment if the AE start date is before the C1D1 date; ongoing if the AE start date is on or after the C1D1 date but before the end of treatment date; post-treatment if the AE start date is after the end of treatment date.

6.7.2 Laboratory Evaluations

Haematology, blood chemistry, coagulation, urinalysis and tumour marker variables recorded in the eCRF will be listed and summarised (in terms of actual and change from baseline values) by treatment group and cycle/study day.

All laboratory data will be reported in International System of Units (SI) units. Laboratory parameters will be assigned a grade using SI values and the NCI-CTCAE classification Version 5.0 or above. Grade 0 will be assigned for all non-missing values not graded as ≥1.

For analysis purposes, values preceded by a "<" or a ">" sign (i.e., those below or above the limits of quantification) will be considered equal to the lower or upper limit of quantification, respectively.

For each laboratory parameter, the baseline value will be defined as last scheduled or unscheduled value collected prior to the first dose of randomised treatment. Assessments carried out on Cycle 1 Day 1 are considered to have taken place before the first dose of randomised treatment.

Shift tables presenting worst post-baseline NCI-CTCAE grade according to baseline grade will be provided for each treatment group. All emergent (on treatment values with grade > baseline grade) values will be considered. Shift tables will be provided for the following parameters:

CRF variables (high/low values)	NCI-CTCAE variables
Haematology	
Eosinophils count (high)	Eosinophil count increased
Haemoglobin (Low values)	Anaemia
Haemoglobin (High values)	Haemoglobin increased
Lymphocytes count (Low values)	Lymphocyte count decreased
Lymphocytes count (High values)	Lymphocyte count increased
Neutrophils count (Low values)	Neutrophil count decreased
Platelets (Low values)	Platelet count decreased
White Blood Cell Count (Low values)	White blood cell decreased
White Blood Cell Count (High values)	Leukocytosis
Chemistry	
Alanine aminotransferase (High values)	Alanine aminotransferase increased
Albumin (Low values)	Hypoalbuminemia
Alkaline phosphatase (High values)	Alkaline phosphatase increased
Aspartate aminotransferase (High values)	Aspartate aminotransferase increased
Bicarbonate (Low values)	Bicarbonate decreased
Calcium (Low values)	Hypocalcaemia
Calcium (High values)	Hypercalcaemia
Creatinine (High values)	Creatinine increased
Glucose (Low values)	Hypoglycemia
Lactate dehydrogenase (high values)	Blood lactate dehydrogenase increased
Magnesium (Low values)	Hypomagnesaemia
Magnesium (High values)	Hypermagnesaemia
Potassium (Low values)	Hypokalemia
Potassium (High values)	Hyperkalemia
Sodium (Low values)	Hyponatremia
Sodium (High values)	Hypernatremia
Total bilirubin (High values)	Blood bilirubin increased
Coagulation	
Activated Partial Thromboplastin Time (High)	aPTT increased
Prothrombin International Normalised Ratio (High)	INR increased

In addition, for parameters that cannot be graded, shift tables of worst post baseline (normal/ (abnormal -Non clinical significant / Clinical Significant) categories according to baseline values will also be displayed. All values will be considered (repeat results within a visit or unscheduled visit). Those parameters are listed below

- Haematology
  - Basophils
  - Hematocrit
  - Monocytes
  - Red blood cell count
  - Lymphocytes/leukocytes (fraction of 1)
  - Neutrophils/leukocytes (fraction of 1)
  - Basophils/leukocytes (fraction of 1)
  - Eosinophils/leukocytes (fraction of 1)
  - Monocytes/leukocytes (fraction of 1)
- Chemistry
  - Urea
  - Chloride
  - Indirect bilirubin
  - Direct bilirubin
  - Phosphate
  - Protein
  - Urate
- Coagulation
  - Prothrombin time

All laboratory parameters (haematology, chemistry, coagulation, tumour markers, urinalysis, and pregnancy test) will be also listed. Out-of-reference-range values (haematology, chemistry) will be flagged as high (H) or low (L) in the listings.

Laboratory parameters with a grade ≥3 according to the NCI-CTCAE will be listed.

Cases of potential drug-induced liver injury are defined by the stopping criteria defined in the FDA DILI guidance document (FDA Guidance for Industry: Drug-induced liver injury: premarketing clinical evaluation, 2009) if they meet any of the following four criteria:

1. Treatment-emergent ALT or AST >8 x ULN
2. Treatment-emergent ALT or AST >5 x ULN for more than 2 weeks*
3. Treatment-emergent ALT or AST >3 x ULN with total bilirubin >2 x ULN or International Normalised Ratio (INR) >1.5
4. Treatment-emergent ALT or AST >3 x ULN with the appearance of fatigue, nausea, vomiting, right upper quadrant pain or tenderness, fever, rash, and/or esosinophilia

* Patients will be included in this category if the end date of the contributing event is more than 2 weeks after the start date of the contributing event.

MedDRA preferred terms for these criteria are ALT increased, AST increased, bilirubin increased and INR increased.

Patients, who meet any of the criteria above and any of the following criteria might potentially meet criteria for Hy's law:

- No initial finding of cholestasis (elevated serum ALP)
- No other reason can be found to explain the increases in ALT/AST and bilirubin or INR, such as viral hepatitis A, B, or C; pre-existing or acute liver disease; biliary duct obstruction or another drug capable of causing the observed injury

For any possible Hy's law cases, patient profiles for all liver function tests will also be produced.

For each of criteria 1-4, counts will be provided for patients meeting the criterion based on ALT but not AST, based on AST but not ALT, based on both ALT and AST, or based on either AST or ALT. Counts will also be provided for patients meeting any of criteria 1-4.

6.7.3 Vital Signs

The following vital signs will be listed and summarised using standard descriptive statistics by treatment group at baseline and at each cycle/study day post randomisation:

- Weight (kg);
- Systolic and diastolic blood pressure (mmHg);
- Heart rate (bpm);
- Body temperature (°C).
- Respiratory rate (breaths per minute)

The baseline value will be defined as last scheduled or unscheduled value collected prior to the first dose of randomised treatment. Assessments carried out on Cyle 1 Day 1 are considered to have taken place before randomised treatment was initiated.

Actual values at baseline and each scheduled visit and change from baseline at each post-baseline scheduled visit of vital signs (including heart, respiration, systolic and diastolic blood pressure, oral temperature, and weight) will be summarised with descriptive statistics by randomised treatment. All data will also be listed.

6.7.4 ECG

ECG parameters will be summarised descriptively by treatment group at each timepoint (pre- and post-dose on Day 1 and Day 15 of each cycle). The triplicate values taken at each timepoint for a patient will be averaged, and the average value will be used in the summaries. Actual values at baseline and each scheduled

timepoint and change from baseline at each post-baseline scheduled timepoint will be summarised.

A shift table of worst post-baseline corrected QTcF values according to baseline values and worst post-baseline change in corrected QTcF values according to baseline will be summarised using counts and percentages for the following categories:

- QTc ≤ 450 msec
- QTc > 450 msec to ≤ 480 msec
- QTc > 480 msec to ≤ 500 msec
- QTc > 500 msec
- QTc increase from baseline ≤ 30 msec
- QTc increase from baseline > 30 msec to ≤ 60 msec
- QTc increase from baseline > 60 msec

All ECG data will be listed.

6.7.5 ECOG Performance Status

ECOG performance status (Grade 0 – Grade 5) will be listed and the total counts and percentages of patients in each grade will be presented by treatment group at baseline and at each cycle/study day post randomisation.

Shift tables presenting worst post randomisation grade according to baseline grade will be provided for each treatment group. All emergent values will be considered (scheduled or unscheduled). In case of several values with the same worst grade, the first one will be taken into account.

6.7.6 PK Data

The PK concentration data will be provided to ICON for inclusion in the listings.

A listing of pharmacokinetic data will be presented.

The PK parameter analysis and reporting is not in scope of ICON biostatistics and programming team. The data analysis plan and population PK analysis and its results will be reported in separate documents.

6.8 Interim Analysis

The are no formal interim analyses planned for this study.

7. Changes in Planned Analysis

8. Data Issues

Not Applicable

9. References

1 ICH. *Statistical Principles for Clinical Trials*, Guideline E9, 1998. Available at http://www.emea.eu.int/pdfs/human/ich/036396en.pdf

2 US Food and Drug Administration, Guidance for Industry: Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics, December 2018.

3 US Food and Drug Administration, Guidance for Industry: Clinical Trial Endpoints for the Approval of Non- Small Cell Lung Cancer Drugs and Biologics. 2015

4 European Medicines Agency: Appendix 1 to the guideline on the evaluation of anticancer medicinal products in man, 13 December 2012.

5 Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). European journal of cancer (Oxford, England: 1990). 2009;45(2):228-247.

6 Carroll K: On the use and utility of the Weibull model in the analysis of survival data. Controlled Clinical Trials 24 (2003) 682-701

7 Ellis S, Carroll KJ, Pemberton K. Analysis of duration of response in oncology trials. Contemp Clin Trials. 2008;29(4):456-465.

8 U.S. Department of health and human services, National Institutes of Health, National Cancer Institute, November 27, 2017 (Version 5.0). Common Terminology Criteria for Adverse Events (CTCAE).

10.Appendices

Appendix 1: Study Schedule

Study Assessments[1]	Screening/ Baseline	Cycle 1				Additional Cycles				End of Treatment	Follow-up (France only)	Follow-up	Follow-up
	Days -28 to 0	D1	D8* (±2 days)	D15 (±2 days)	D22* (±2 days)	D1 (±2 days)	D8* (±2 days)	D15 (±2 days)	D22* (±2 days)	30 days post last dose (+7 days)	Q4 weeks (±3 days)	Q8 weeks (±7 days)	Q12 weeks (±14 days)
Informed consent	X												
Eligibility criteria	X												
Demographic data	X												
Medical history[2]	X												
Concomitant medications	X	X	X	X	X	X	X	X	X	X			
Full physical examination[3]	X	X[5]				X				X			
Directed physical examination[3]						X[3]							
Urinalysis[4]	X	X[5]		X		X		X		X			
ECOG status	X	X[5]				X				X			
Vital signs[6]	X	X	X	X	X	X	X	X	X	X			
ECG (pre-dose)[7]	X	X		X		X		X[7]		X			
ECG (post-dose)[8]		X		X		X[8]		X[8]					
Pregnancy test[9]	X	X				X				X	X[9]		
FBC and chemistry[10]	X	X[5]	X	X	X	X	X	X	X	X			
Coagulation profile	X	X				X				X			
Tumour markers[11]	X					X				X			
Blood uracil testing (DPD phenotyping)[12]	X												
BRAF/ KRAS/ NRAS/ MSI/ MMR status testing[13]	X												
UGT1A1 status testing[14]	X												
Radiologic tumour assessment (CT/MRI)[15]	X	*Every 8 weeks (±7 days) from C1D1 until disease progression*										X[16]	

Randomisation		X[17]											
NUC-3373 + LV administration (Q1W)		X	X	X	X	X	X	X	X				
NUC-3373 + LV administration (Q2W)		X		X		X		X					
5-FU + LV administration		X		X		X		X					
Irinotecan administration		X		X		X		X					
Bevacizumab administration		X		X		X		X					
AEs/SAEs[18]	X	X	X	X	X	X	X	X	X	X			
PK blood sample[19]						X[19]							
Archived sample[20]	X												
Dental examination[21]	X												
Follow-up		*Ongoing*											X[22]

* **Day 8 and Day 22 visits are only to be performed for patients receiving Q1W NUC-3373 (Arm A)**

1 Assessments scheduled on days of dosing should be done prior to administration of all investigation medicinal products (IMPs), unless otherwise specified. Lab assessments may be performed up to 72 hours prior to IMP administration.

2 Includes recording information on the sidedness of the patient's CRC (left-sided vs right-sided).

3 A full physical assessment should be completed at Cycle 1 Day 1, Cycle 2 Day 1 and at End of Treatment. From Cycle 3 Day 1 onwards, a directed physical assessment will be completed (in place of the full physical assessment), but only if clinically indicated.

4 Urinalysis testing for proteinuria should be performed prior to each bevacizumab administration.

5 Does not need to be repeated if Screening assessment was performed within 72 hours of C1D1.

6 Vital signs include respiration rate, pulse, temperature and blood pressure. Height should be recorded at baseline only. Weight should be recorded at baseline, Day 1 of every cycle and at end of study visit. If a patient's weight increases or decreases by ≥10% during the course of the study, the dose of study treatments should be recalculated.

7 **Pre-dose ECGs:** Standard 12-lead ECG measurements will be performed prior to administration of all IMPs, at the indicated visits (Day 1 of each cycle) for patients in all treatment arms.

All 12-lead ECG measurements should be performed in **triplicate** (keeping the leads in place and the patient supine during readings) and reviewed by the Investigator or qualified designee for safety and quality. The timing between the triplicate ECGs is recommended to be approximately 1 minute.

8 **Post-dose ECGs:** Additional standard 12-lead ECG measurements must be taken post administration of all IMPs on C1D1, C1D15, C2D1 and C2D15 (within 10 minutes of the end of infusion) for patients in the **NUFIRI treatment arms only** (Arms A and B).

9 Serum pregnancy assessment to be performed within 7 days of C1D1. Required only in women of childbearing potential. **In France only:** Pregnancy testing must continue every 4 weeks for 6 months after the last dose of study treatment. This post-study treatment pregnancy testing must not be performed with a home test; however, in order to avoid frequent hospital visits, it may be performed at an external laboratory that is local to the patient, with oversight by the study site. If this occurs, the study site must notify the Sponsor. All instances of positive pregnancy tests must be notified by the external laboratory to the study site, Sponsor and the patient's treating physician.

10 Clinical chemistry (including hepatic panel) and haematology will be conducted weekly in the Q1W treatment arm and alternate weekly in the Q2W treatment arms throughout the study. In the event of neutropenia (ANC $<0.5\times10^9$/L), thrombocytopenia (platelet count $<50\times10^9$/L), or ≥Grade 2 clinical chemistry toxicity, these assessments will be conducted more frequently as clinically indicated until toxicity resolves to ≤Grade 1.

11 Collect a pre-dose blood sample for evaluation of carcinoembryonic antigen (CEA).

12 Blood uracil levels must be tested during screening to determine DPD phenotypic status. Patients with blood uracil ≥150 ng/mL are considered to be DPD deficient and are excluded from participation. Patients with blood uracil levels of ≥16 ng/mL to <150 ng/mL are considered to be partially DPD deficient and the 5-FU dose must be adapted in countries where this is standard practice as per national and/or local guidelines (refer to Section 9.2.1).

13 If KRAS/NRAS/BRAF/MSI/MMR status is not known, perform genetic testing and obtain results prior to dosing on C1D1.

14 If UGT1A1 status is not known, perform genetic testing. Results do not need to be obtained prior to dosing. If a patient's mutational status is known prior to dosing and they have a mutation that may affect their ability to metabolise irinotecan, an initial dose reduction of irinotecan may be implemented as per the irinotecan SmPC/Prescribing Information and site standard of care. This must be discussed on a case-by-case basis with the medical monitor.

15 Computed tomography (CT) / magnetic resonance imaging (MRI) disease assessments will be performed at Screening (within 28 days prior to randomisation) and every 8 weeks (±7 days) from C1D1.

Additional tests may be requested at the Investigator's discretion. The same modality should be used throughout.

16 Patients discontinuing study treatment with no radiological evidence of disease progression will remain in the study and receive scans every 8 weeks (±7 days) from C1D1 until disease progression, initiation of a subsequent line of therapy, or death in order to determine duration of overall response and PFS.

17 Randomisation takes place up to 3 working days prior to administration of study treatment on C1D1.

18 All adverse events (AEs) occurring from the time of informed consent up to and including 30 days after the last dose of study drug has been administered must be reported in detail on the AE case report form (CRF).

Note: Investigators must report all SAEs that they become aware of irrespective of the end of study treatment or the end of study, unless the patient has initiated a new therapy after which only SARs must be reported.

19 Collection of blood samples for PK analysis will be performed for patients in the NUFIRI treatment arms only (Arms A and B).

A total of 4 blood samples will be collected on **C2D1** only:

PK timepoint	Description
Pre-dose	Prior to administration of NUC-3373
End of infusion	Within 5 minutes before end of NUC-3373 infusion
2-4 hours post-infusion	2-4 hours after end of NUC-3373 infusion
6-24 hours post-infusion	6-24 hours after end of NUC-3373 infusion

The exact time that each PK sample is taken must be recorded.

20 Original diagnostic or other representative FFPE block containing tumour will be recalled (where available).

21 A dental examination and appropriate preventative dentistry should be considered prior to starting treatment with bevacizumab. In patients who have previously received or are receiving IV bisphosphonates, invasive dental procedures should be avoided if possible.

22 All patients, including those who discontinue study treatment, will be followed up for disease progression, initiation of new treatments and survival every 12 weeks (±14 days) from C1D1 until withdrawal of consent, lost to follow-up, death, or the overall end of study, whichever is earliest. A total of three attempts should be made before the patient is considered as lost to follow-up. Patients in follow-up who have not experienced disease progression should continue to attend the clinic for planned radiologic scans; however, other follow-up data can be collected remotely.

Collection of follow-up data can be performed via a telephone call with the patient where possible. Where not possible, follow-up with the patient's next of kin or physician should be performed, along with review of medical notes or national registries/databases if needed. The data collected and the follow-up schedule remain as per the schedule of events. Disease progression and survival will be censored at the end of study date.

Appendix 2: Sample SAS® code for analyses

- **Tables that need descriptive statistics – continuous variables:**

```
PROC UNIVARIATE DATA=dset NOPRINT;
    VAR var1 var2 var3 …varn;
    BY byvar; (optional)
    OUTPUT OUT=outname;
    N=n MEAN=mean MIN=min MAX=max MEDIAN=median STD=std;
RUN;
```

- **Tables that need frequency counts:**

```
PROC FREQ DATA=dset NOPRINT;
    BY byvar; (optional)
    TABLES var1*var2;
    OUTPUT OUT=outname;
RUN;
```

- **Tables that need exact (Clopper-Pearson) 95% CIs between groups for proportions:**

```
PROC FREQ DATA=dset;
    BY byvar; (optional)
    TABLES var1 * var2 / binomial(exact) MEASURES ALPHA=0.05;
RUN;
```

Notes: 1 Estimates are computed for 2x2 tables only

2 This code also gives exact 95% CIs within group for binomial proportions

3 Put *ALPHA=0.2* for giving exact 80% CIs within group for binomial proportions

- **Tables that need 95% CIs within group for binomial proportions:**

```
PROC FREQ DATA=dset;
    BY byvar; (optional)
    TABLES var1;
    EXACT BINOMIAL;
RUN;
```

- **Tables that require exact logistic regression:**

PROC LOGISTIC DATA=dset descending;

CLASS treatment;

STRATA byvar; (optional)

MODEL response = <treatment>;

EXACT treatment;

RUN;

- **Tables that need to use WILCOXON:**

PROC NPAR1WAY DATA=dset WILCOXON;

CLASS class variables;

VAR variable;

RUN;

- **Tables that need number of events/censored and probabilities of failure/survival at cut off times:**

PROC LIFETEST DATA=dset OUTSURV=LIFE METHOD=LT INTERVALS=12;

*TIME duration*censor (0 or 1);*

ID patient;

STRATA treatment;

RUN;

Notes: PROC ICLIFETEST will be used in case of interval censored method

- **Tables that need life table with estimates of survival, with CIs and log rank test:**

PROC LIFETEST DATA=dset OUTSURV=LIFE METHOD=KM;

*TIME duration*censor (0 or 1);*

ID patient;

STRATA treatment;

RUN;

- **Kaplan-Meier curves for treatment:**

PROC LIFETEST data=dataset plots=survival(strata=individual);

*TIME time*event(0);*

STRATA treatment;

RUN;

- **Tables that need Cox Proportional Hazards models:**

ODS OUTPUT PARAMETERESTIMATES=parms;

PROC PHREG data=dataset;

CLASS treatment covariable;

*MODEL time*event(0) = treatment covariable / ties=exact;*

CONTRAST 'a vs b' treatment 1 / estimate=exp;

RUN;

Notes: PROC ICPHREG will be used in case of interval censored method.

- **Tables that need Cox Proportional Hazards models (martingale residuals):**

ODS OUTPUT PARAMETERESTIMATES=parms;

PROC PHREG data=dataset;

CLASS treatment covariable;

*MODEL time*event(0) = treatment covariable / ties=exact;*

ASSESS var=(treatment) PH / resample;

CONTRAST 'a vs b' treatment 1 / estimate=exp;

RUN;

Notes: Baseline ECOG (0 versus 1) score used as a continuous variable in this analysis.

- **Multiple imputation**

Producing monotone missing data patterns

PROC MI DATA= dataset NIMPUTE=100 SEED=14823 OUT=mi1_dataset NOPRINT;

MCMC IMPUTE=MONOTONE;

VAR time;

BY treatment;

RUN;

Imputation step on monotone dataset

PROC MI DATA= mi1_dataset SEED=14823 nimpute=1 out= mi2_dataset NOPRINT;

MONOTONE REGRESSION;

VAR site time_1st_diag nb_prior_rad nb_prior_surg nb_prior_systh time;

BY _imputation_ treatment;

RUN;

With: site, the site number,

time_1st_diag, the time since initial diagnosis of CRC,

nb_prior_rad, the number of prior cancer radiotherapies administered,

nb_prior_surg, the number of prior surgeries/procedures for CRC

nb_prior_systh, the number of systemic therapies administered,

time, the event (PFS/OS).

Inference step

ODS OUTPUT PARAMETERESTIMATES=parms;

PROC PHREG data= mi2_dataset;

CLASS treatment covariable;

*MODEL time*event(0) = treatment covariable/ ties=exact;*

CONTRAST 'a vs b' treatment 1 / estimate=exp;

BY imputation;

RUN;

Combination step:

ODS OUTPUT PARAMETERESTIMATES= MI_parms;

PROC MIANALYZE DATA= parms;

MODELEFFECTS var;

BY TREATMENT TIME;

RUN;

- **Multiple imputation using the tipping-point approach**

Producing monotone missing data patterns

PROC MI DATA= dataset NIMPUTE=100 SEED=14823 OUT=mi1_dataset NOPRINT;

MCMC IMPUTE=MONOTONE;

VAR time;

BY treatment;

RUN;

Imputation step on monotone dataset

PROC MI DATA= mi1_dataset SEED=14823 nimpute=1 out= mi2_dataset NOPRINT;

MONOTONE REGRESSION;

VAR site time_1st_diag nb_prior_rad nb_prior_surg nb_prior_systh time;

MNAR ADJUST (time / SHIFT=&S ADJUSTOBS=(treatment ='Exp. Arm'));

BY _imputation_ treatment;

RUN;

With: site, the site number,

time_1st_diag, the time since initial diagnosis of CRC,

nb_prior_rad, the number of prior cancer radiotherapies administered,

nb_prior_surg, the number of prior surgeries/procedures for CRC,

nb_prior_systh, the number of systemic therapies administered,

time, the event (PFS/OS),

&S, shift parameters 1, 1.1, 1.2 …, 2.5. / This range assumes that the tipping point for the shift parameter that reverses the study conclusion is between 1 and 2.5. The following statement performs multiple imputation analysis for each of the shift parameters 1, 1.1, 1.2 …, 2.5.

Inference step

ODS OUTPUT PARAMETERESTIMATES=parms;

PROC PHREG data= mi2_dataset;

CLASS treatment covariable;

*MODEL time*event(0) = treatment covariable/ ties=exact;*

CONTRAST 'a vs b' treatment 1 / estimate=exp;

BY imputation;

```
RUN;
Combination step:
ODS OUTPUT PARAMETERESTIMATES= MI_parms;
PROC MIANALYZE DATA= parms;
    MODELEFFECTS var;
    BY TREATMENT TIME;
RUN;
Adding the shift parameter:
DATA MI_parms;
    SET MI_parms;
    SHIFT= &S;
RUN;
```