METIMMOX: COLORECTAL CANCER METASTASIS – SHAPING ANTI-TUMOR IMMUNITY BY OXALIPLATIN

Protocol Acronym: METIMMOX

ClinicalTrials ID: NCT03388190

STATISTICAL ANALYSIS PLAN

Version: 1

Date: 06-May-2024

Authors:

Anne Hansen Ree
Akershus University Hospital
Coordinating Investigator /
Principal Investigator

Jūratė Šaltytė Benth
Akershus University Hospital
Study Statistician

Contents:

1. Introduction

1.1. Background and rationale

Owing to an aging population, colorectal cancer (CRC) is a common malignancy with a sharp rise in incidence from the age of 60. Immune checkpoint blockade (ICB) is efficacious in the small CRC subgroup of patients with highly immunogenic disease, the microsatellite-instable/mismatch repair (MMR)-deficient entity. ICB is, however, considered inefficacious for the majority of patients presenting microsatellite-stable (MSS)/MMR-proficient CRC. Unresectable abdominal metastases commonly reflect an aggressive phenotype.

Our previous findings for short-course oxaliplatin-containing chemotherapy in locally advanced or early metastatic CRC support a notion that oxaliplatin may invoke tumor-defeating immunity. Specifically, patients who presented unresectable single-organ liver metastases as the first metastatic event, given oxaliplatin as hepatic arterial infusion chemotherapy and responding with a rapid rise in a circulating anti-tumor immune factor, were alive 8-12 years later.

In the METIMMOX trial, patients with previously untreated, unresectable abdominal metastases from MSS-CRC will be randomly assigned to short-course oxaliplatin-based chemotherapy (the Nordic FLOX regimen) alternating with ICB (nivolumab) or standard FLOX chemotherapy.

1.2. Objectives

Primary objective: To determine progression-free survival (PFS), in terms of failure of treatment strategy, on sequential treatment with the Nordic FLOX regimen and nivolumab compared with the standard-of-care Nordic FLOX regimen in previously untreated metastatic MSS-CRC.

Secondary objectives:

i) To determine safety and tolerability of sequential treatment with the Nordic FLOX regimen and nivolumab compared with the standard-of-care Nordic FLOX regimen.
ii) To determine objective response rate (ORR), duration of response (DOR), secondary surgical curative resection rate (SSCRR), and overall survival (OS).
iii) To monitor and compare quality-of-life (QoL) alterations during therapy courses.

Tertiary objective: To compare costs for the resource use (in diagnostic work-up, treatment, and any adverse events) for the sequential therapy with that of the standard-of-care, applying a model specifically developed for CRC.

Exploratory objectives:

i) To develop circulating biomarkers of cytotoxic T lymphocyte activity.
ii) To develop circulating biomarkers of tumor response.
iii) To develop circulating biomarkers of immunogenic cell death.
iv) To develop functional MR imaging biomarkers.
v) To develop histologic and molecular tumor biomarkers.

2. Study Methods

2.1. Trial design

The METIMMOX study is a multicenter open-label randomized phase 2 trial in first-line treatment of metastatic MSS-CRC using the standard-of-care Nordic FLOX regimen (control arm) or sequential therapy with the Nordic FLOX regimen and nivolumab (experimental arm), to investigate whether the experimental arm shows superiority in PFS, safety, tolerability, and QoL.

Study period

- Estimated date of first patient enrolled: 1 March 2018
- Anticipated recruitment period: 12 months
- Estimated date of last patient completed: 1 September 2020

Treatment duration

- Expected treatment duration per patient:
  - Experimental arm: 18 months (until failure of treatment strategy)
  - Control arm: 9 months (until failure of treatment strategy)

Follow-up:

- Expected follow-up period per patient:
  - Experimental arm: 30 months (12 months following failure of treatment strategy)
  - Control arm: 21 months (12 months following failure of treatment strategy)

2.2. Randomization

The patients will be block-randomized into treatment arms with ratio 1:1 following a computer-randomized allocation sequence. The blocks will be defined according to sidedness and (K)RAS/BRAF mutation status of the primary tumor (right-sided versus left-sided and (K)RAS/BRAF wild-type versus mutation). The allocation sequence will be generated by the web-based Viedoc system.

2.3. Sample size

Calculation of sample size was performed for PFS – the primary end point – and based on PFS data from recent publications. A median PFS of 9 months was assumed for the standard-of-care (control arm) [Cartwright, 2012; Tveit et al., 2012; Cremolini et al., 2015]. Regarding the therapy concept of ICB, there is no PFS data published as yet for the setting of interest. In the study of patients with treatment-refractory metastatic CRC, median PFS was not reached in the cohort with mismatch repair-deficient disease at the time of reporting [Le et al., 2015]. In the study of previously untreated patients with advanced non-small cell lung cancer, which is a tumor entity with considerably poorer prognosis than advanced CRC, patients with PD-L1-positive tumors that were randomized to ICB showed median PFS of 10.3 months compared to 6.0 months for standard chemotherapy [Reck et al., 2016].

Extrapolating to the setting of interest, with a median PFS of 9 months for standard-of-care, a median PFS twice as long (18 months) was entered for the sequential therapy. Assuming the exponential distribution of survival functions, the median PFS estimates were converted to hazard ratio (HR) of 0.5. Allowing for 10% censoring rate of subjects, the required sample size was estimated to be 40 patients in each arm to be able to show that HR of 0.5, implying the risk of a PFS event in the experimental group is 50% lower than in the control group, is significantly different from 1 with the power of 80% at significance level of 5%, given that log-rank test will be used.

However, it was necessary with several adjustments of the above calculations to maintain sufficient power. First, the sample size was adjusted to enable removal of a maximum of 5% enrolled cases with non-MSS tumor before data analysis. Next, keeping in mind the condition of patients, it is not unlikely that some will be lost from the study before a valid PFS event has been reached. Also, a number of analyses stratified by primary tumor sidedness and (K)RAS/BRAF mutation status will be performed. Hence, to enable high enough power in the statistical analysis, the sample size was increased by an additional 20%. After

incorporating the adjustments, the final sample size was estimated to be 100 patients, 50 in each arm, with 1:1 randomization.

With the total catchment populations of the participating centers we envisage an accrual duration period of approximately 12 months to procure the sufficient number of study cases. As there is no PFS data published as of yet for the novel therapy concept of repeat sequential immunogenic chemotherapy and ICB, and further, the optimum duration of immune-modulating therapy is currently unknown, the study duration of the expected METIMMOX longer-lasting arm (experimental arm) was estimated to 30 months (12 months following failure of treatment strategy). Because these estimates were pure conjecture, they were not entered into the sample size calculation.

2.4. Framework

The analyses comparing experimental and control groups will be performed within the framework of superiority hypotheses testing.

2.5. Statistical interim analysis and stopping guidance

Interim analyses will particularly apply to data on safety, tolerability, possible detrimental effects, and curative outcomes, and will be based on purely clinical judgment by the External Data and Safety Monitoring Committee. In particular, CTCAE grade 3–5, PFS measures considerably shorter than 9 months, complete response (CR), and secondary surgical curative resection events will be counted.

2.6. Timing of final analysis

Analyses of the various end point data will not be undertaken simultaneously, as the various end points will not mature simultaneously. For example, some biomarker analyses may be done before the final outcome data are mature.

2.7. Timing of outcome assessments

See 5.1 for end point definitions and time points at which end points will be measured.

3. Statistical Principals

3.1. Confidence intervals and P values

Level of statistical significance is set to 5% in all analyses. Whenever relevant, 95% confidence intervals (CIs) will be reported. No adjustment for multiple testing will be introduced. However, all performed statistical analyses will be reported in the publications leading to transparency and enabling reader to judge the validity of the reported p-values.

3.2. Adherence and Protocol Deviations

Investigators ascertain they will apply due diligence to avoid protocol deviations. All significant protocol deviations will be recorded and reported in the clinical study report.

3.3. Analysis populations

The following study populations will be defined:

- Intention-to-treat population: Includes all randomized participants, regardless of protocol adherence. We expect about 5% of randomized patients to have non-MSS tumor. These patients will be excluded from the analysis post-randomization. According to Fergusson and co-workers [Fergusson et al., 2002], such exclusion does not introduce bias and may be legitimate if an independent, blinded adjudication committee makes the decision on exclusion after evaluating all randomized patients. Sample size calculations include necessary adjustments to preserve the same power after exclusion of misdiagnosed patients. The sample size calculation has included necessary adjustments to maintain the same power after exclusion of patients with non-MSS tumor.
- Safety population: Includes all randomized participants who have received at least one dose of study medication and had at least one sequential safety-related observation. Subjects who withdraw from the study will be included in the safety analysis given the safety population criteria are otherwise met. A list of withdrawn subjects, preferably with the reasons for withdrawal, will be made.
- Per-protocol population: Includes all subjects who have adhered to all protocol-defined treatments and outcome assessments.

We will account for all of the patients registered in the study. The number of patients who are not evaluable, or who die or withdraw before treatment has begun, will be specified. The distribution of follow-up time will be described and the number of patients lost to follow-up will be given. Start of follow-up will be defined as the date of randomization.

4. Trial Populations

4.1. Screening data

Data on subjects who do not meet the eligibility criteria will not be recorded.

4.2. Eligibility

Eligible patients, with no upper age limit to recruit subjects reflecting population-based incidence rates, have previously untreated, unresectable metastatic colorectal MSS adenocarcinoma. Essential study inclusion criteria are age ≥18 years, measurable infradiaphragmatic (liver, peritoneal and/or nodal) metastatic manifestation(s) according to Response Evaluation Criteria in Solid Tumors version 1.1 (RECIST 1.1), and Eastern Cooperative Oncology Group performance status 0-1. In addition, C-reactive protein (CRP) <60 mg/L is required at study entry. A period <6 months since discontinuation of neoadjuvant or adjuvant oxaliplatin-containing chemotherapy and a history of autoimmune disease are main exclusion criteria. The complete list of inclusion and exclusion criteria can be found with the clinical trial registration (ClinicalTrials.gov Identifier: NCT03388190) and in the trial protocol.

4.3. Recruitment

This information will be reported in a CONSORT flow diagram of the study cases.

4.4. Withdrawal / Follow-up

The number of subjects that withdraw consent or are lost to follow-up, with timing of such events, will be reported.

4.5. Baseline patient characteristics

Continuous baseline characteristics will be presented as means and standard deviations or as medians and minimum and maximum dependent on the distribution. Categorical data will be presented by frequencies and percentages. Baseline characteristics will be presented by intervention and control groups. Following characteristics will be presented: age, sex, ECOG performance status, primary tumor sidedness, tumor (K)RAS/BRAF mutation status, number of metastatic sites, and if the liver is involved or not.

5. Analysis

5.1. End point definitions

Primary end point: PFS; radiologic assessment every 8 weeks, according to the Response Evaluation Criteria in Solid Tumors (RECIST) and the RECIST consensus guideline for assessment of response to immune-modulating therapies, iRECIST.

Secondary end points:

Safety: incidence of adverse events (AE), as reported according to Common Terminology Criteria for Adverse Events (CTCAE) version 4.0 recorded on ongoing basis.

Tolerability: AE grading, as assessed by CTCAE v4.0 recorded on ongoing basis.

ORR: the percentage of patients with a confirmed complete response (CR)or partial response (PR).

DOR: the time from the first documentation of a CR or PR to disease progression.

SSCRR: the percentage of patients with a confirmed resection of metastatic disease with microscopically free margin (R0).

OS: the time from randomization to death of any cause.

QoL: assessment undertaken at study enrolment, completion of the active therapy of the first treatment sequence (after 16 weeks), and disease progression, using the consensus modules EORTC QLQ-C30, EORTC QLQ-CIPN20, and EQ-5D-5L.

Tertiary end point: Cost estimate; evaluation of the healthcare costs (relating to the sequencing procedures and administered medication) and health effects (quality-adjusted life years), expressed as incremental cost-effectiveness ratio, of the sequential therapy relative to the control group as comparator.

Exploratory end points: Various analyses will be undertaken for exploratory end points listed in 1.2.

5.2. Analysis methods

Main statistical analyses. The main statistical analysis is planned after the last patient has fulfilled the pre-defined follow-up period. PFS will be calculated from the day of randomization to first recorded disease progression according to RECIST v1.1 or when applicable iRECIST, or death. Patients who have not had disease progression at the time of final analysis will be censored at the date last known to be alive and progression-free. The analyses will compare the treatment groups in total as well as groups stratified by primary tumor sidedness and (K)RAS/BRAF mutation status.

Interim analyses. Interim analyses will apply to data on safety, tolerability, possible detrimental effects, and curative outcomes, and will be based on purely clinical judgment by the External Data and Safety Monitoring Committee. These types of data will be reported as the number of events.

Descriptive statistics. Demographic and clinical characteristics will be presented as means and standard deviations, medians with minimum and maximum values, or frequencies and percentages, as appropriate. Further statistical analyses will be separated to primary analyses, including only comparison of primary end point between two treatment arms, and secondary analyses, including stratified analyses on primary and secondary end points as well as number of analyses on secondary end points.

Primary analysis. Median PFS time will be compared between study arms by drawing Kaplan-Meier survival curves and performing log-rank test. HR with the corresponding 95% CI and p-values will be estimated by Cox proportional hazards model.

Secondary analyses. Distribution of PFS times stratified by the primary tumor sidedness and (K)RAS/BRAF mutation status will be illustrated graphically by Kaplan-Meier curves and compared by log-rank test. To assess the hazard rates for disease progression within each strata, Cox proportional hazards model will be estimated and, if appropriate, adjusted for relevant covariates. The results will be presented as HR with the corresponding 95% CI and p-values.

Safety. The number of serious adverse events (SAE) will be summarized and presented by trial arm. In addition, information will be given on the number of SAE per participant, together with details on the causality, expectedness, and outcome of each SAE experienced. Summaries of suspected unexpected serious adverse reaction (SUSAR) events will also be presented by trial arm.

Tolerability. Summaries by trial arm will be produced to show the proportion of participants experiencing each grade of CTCAE toxicities overall and during each treatment cycle. For more detailed summaries, this information will also be broken down into the different types of toxicities.

QoL. This will be summarized for each treatment arm at study enrolment, completion of the active therapy of the first treatment sequence (after 16 weeks), and disease progression, using the consensus modules EORTC QLQ-C30, EORTC QLQ-CIPN20, and EQ-5D-5L. The data will be illustrated graphically and presented as mean scores with 95% CI at separate time points. The differences in QoL between two treatment arms throughout the study period will be assessed by estimating a linear mixed model with fixed effects for time and treatment arm and the interaction between these two. Random effects for patients will be included. A significant interaction will quantify differences between the groups and serve as an omnibus test.

Safety and tolerability in the treatment arms will be compared by χ^2 or Fisher's exact test, as appropriate.

ORR is defined as the percentage of patients with a confirmed CR or PR. The proportion of participants achieving at least PR will be summarized with corresponding 95% CI for each treatment arm. The difference between the trial arms in the proportion of participants achieving at least PR will also be presented with corresponding 95% CI. The difference between participants achieving at least PR and remaining participants will be assessed using logistic regression adjusting for treatment group and stratification factors. Odds ratios with corresponding 95% CI and p-values will be presented for all variables incorporated in the model.

SSCRR is defined as the percentage of patients with a confirmed resection of metastatic disease with microscopically free margin (R0). Corresponding analyses as those to be performed for ORR will be applied for SSCRR.

DOR is defined as the time from the first documentation of CR/iCR or PR/iPR to disease progression. DOR Kaplan-Meier curves will be drawn and compared by log-rank test. The median DOR with corresponding 95% CI will be presented by treatment group. Stratified analyses will be performed as well. The Cox proportional hazards model, adjusting for the stratification factors, will also be used to compare DOR

between the treatment groups. HR, 95%CI and p-values will be presented for all variables incorporated in the model.

OS is defined as the time from randomization to death of any cause. The same analyses as for DOR will be performed for OS.

The assumptions for the statistical method considered will be assessed by using standard techniques. In the case the assumptions are not met, the necessary adjustments for methods specified above will be introduced and reported.

Analysis of other end points. Cost modeling and validation will apply a model specifically developed for CRC by Investigator at Health Services Research Center, Akershus University Hospital [Joranger et al., 2015].

For *exploratory analyses*, appropriate statistical methods will be applied to analyze the various outcome measures.

5.3. Missing data

Number of missing data will be reported throughout all analyses. It is expected that the number of missing will be low and thus no missing values handling approach is planned. In the case the number of missing will show to be more substantial than expected however, an appropriate approach like for example weighting or multiple imputation will be considered.

5.4. Additional analyses

No other analyses than those specified above are planned.

5.5. Harms

The safety data will be recorded and reported according to severity, expectedness, and causality. Only CTCAE grade 3-5 will be recorded. CTCAE grade 2 events will be recorded only for immune-related hepatotoxicity (increased hepatic enzymes).

The following events will be individually recorded and reported according to prevailing national procedures:

Pregnancy (including pregnancy occurring in a female partner of a male study participant), overdose, and SUSARs. The events will be analyzed on incidence case basis. Each site investigator is responsible for assuring that there are procedures and expertise available to cope with emergencies during the study.

5.6. Statistical software

SPSS Statistics, SAS, STATA, and GraphPad Prism (the most recent versions for all) are relevant software for statistical analyses of the study data.

5.7. References

Cartwright TH. Treatment decisions after diagnosis of metastatic colorectal cancer. *Clin Colorectal Cancer* **2012**; 11, 155-66.

Cremolini C, Schirripa M, Antoniotti C, Moretto R, Salvatore L, Masi G, Falcone A, Loupakis F. First-line chemotherapy for mCRC–a review and evidence-based algorithm. *Nat Rev Clin Oncol* **2015**; 12, 607-19.

Fergusson D, Aaron SD, Guyatt G, Hebert P. Post-randomisation exclusions: the intention to treat principle and excluding patients from analysis. *BMJ* **2002**; 325, 652-4.

Joranger P, Nesbakken A, Hoff G, Sorbye H, Oshaug A, Aas E. Modeling and validating the cost and clinical pathway of colorectal cancer *Med Decis Making* **2015**; 35, 255-65.

Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, Skora AD, Luber BS, Azad NS, Laheru D, Biedrzycki B, Donehower RC, Zaheer A, Fisher GA, Crocenzi TS, Lee JJ, Duffy SM, Goldberg RM, de la Chapelle A, Koshiji M, Bhaijee F, Huebner T, Hruban RH, Wood LD, Cuka N, Pardoll DM, Papadopoulos N, Kinzler KW, Zhou S, Cornish TC, Taube JM, Anders RA, Eshleman JR, Vogelstein B, Diaz LA Jr. PD-1 blockade in tumors with mismatch-repair deficiency. *N Engl J Med* **2015**; 372, 2509- 20.

Reck M, Rodríguez-Abreu D, Robinson AG, Hui R, Csõszi T, Fülöp A, Gottfried M, Peled N, Tafreshi A, Cuffe S, O'Brien M, Rao S, Hotta K, Leiby MA, Lubiniecki GM, Shentu Y, Rangwala R, Brahmer JR; KEYNOTE-024 Investigators. Pembrolizumab versus chemotherapy for PD-L1-positive non-small-cell lung cancer. *N Engl J Med* **2016**; 375, 1823-33.

Tveit KM, Guren T, Glimelius B, Pfeiffer P, Sorbye H, Pyrhonen S, Sigurdsson F, Kure E, Ikdahl T, Skovlund E, Fokstuen T, Hansen F, Hofsli E, Birkemeyer E, Johnsson A, Starkhammar H, Yilmaz MK, Keldsen N, Erdal AB, Dajani O, Dahl O, Christoffersen T. Phase III trial of cetuximab with continuous or intermittent fluorouracil, leucovorin, and oxaliplatin (Nordic FLOX) versus FLOX alone in first-line treatment of metastatic colorectal cancer: the NORDIC-VII study. *J Clin Oncol* **2012**; 30, 1755-62