RANDOMIZED PHASE II TRIAL IN METASTATIC PANCREATIC CANCER EVALUATING FOLFIRINOX +/- LV5FU2 IN MAINTENANCE AND FIRGEM IN 1ière LINE

Phase II randomized - multicenter

PRODIGE 35

Statistical Analysis Plan

Final analysis

Version: 2.0

Date: March 25, 2016

Editor: Karine Le Malicot

Review committee: Laetitia Dahan, Jérémie Bez , Emilie Maillard, Paolo Carni, Fadil Masskouri.

Developer

Fédération Francophone de Cancérologie Digestive (FFCD),

Faculty of Medicine,

7, Boulevard Jeanne d'Arc, BP 87900,

21079 Dijon Cedex

Coordinator

Pr Laetitia DAHAN

Digestive Oncology Department

Hôpital de la Timone, 264 rue Saint Pierre

13005 MARSEILLE

Tel. 04 91 38 60 23

Email : Laetitia.DAHAN@ap-hm.fr

Project **Manager**

Mr Jérémie Bez

FFCD

Faculty of Medicine

7 boulevard Jeanne d'Arc

BP 87900, 21079, Dijon Cedex

Tel.: 03 80 39 34 83

Email : jérémie.bez@u-bourgogne.fr

1 Table of contents

2 Abbreviations and definitions

ACE	embryonic carcino antigen
ALAT	Alanine aminotransferase (or SGPT: serum glutamic pyruvic transaminase)
ANSM	National Agency for the Safety of Medicines and Health Products
ASAT	Aspartate aminotransferase (or SGOT: serum glutamic oxaloacetic transaminase)
PPC	Comité de Protection des Personnes
EI	Undesirable event
EIG	Serious adverse event
5FU	5-fluorouracil
FFCD	French-speaking Federation of Digestive Oncology
FOLFIRI	Folinic acid - Fluorouracil - IRInotecan
GDS	Geriatric depression scale
GGT	Gamma glutamyl transpeptidase
Hb	Hemoglobin
HR	Hazard ratio
ITT	Intention to treat
KM	Kaplan Meier
LDH	Lactate dehydrogenase
LSN	Upper limit of normal
MMSE	Mini mental state examination
N	Normal
NCI-CTCAE	National Cancer Institute - Common Terminology Criteria for Adverse Events
NFS	Blood count
WHO	World Health Organization
PAL	Alkaline phosphatases
PNN	Neutrophil Polynuclear
Q1-Q3	Quartiles
RECIST	Response Evaluation Criteria In Solid Tumors
RC	Complete Answer
RP	Partial response
RO	Objective answer
PNN	Neutrophils
SG	Overall survival
SSP	Progression-free survival
SD	Stable disease
TAP	Thoraco-Abdomino-Pelvic
CT SCAN	Computed tomography
TP	Prothrombin levels

3 Introduction

3.1 Study objectives

3.1.1 Main objective

The main aim of this Phase II study is to assess the rate of patients alive and free of radiological (RECIST V1.1) and/or clinical progression at 6 months in each arm, **and to** select the best therapeutic strategy for a future Phase III trial.

3.1.2 Secondary objectives

The secondary objectives of the study are to evaluate:

- Duration of disease control
- progression-free survival
- Progression time during maintenance treatment,
- Median time on maintenance treatment (arm B)
- Objective response rate
- Overall survival
- Toxicity according to NCI CTC v4.0
- Overall survival
- The rate of patients receiving 2nd-line therapy
- Quality of life (EORTC QLQ-C30 questionnaire)
- Geriatric frailty factors predictive of treatment failure in patients over 65.

In arm A (Folfirinox), duration of disease control is defined as the time between the date of randomization and the date of treatment progression (clinical and/or radiological) or death (from any cause and at any time) or the date of line change.

In arm B with maintenance (arm B), the duration of disease control is defined as follows:

- After progression during LV5FU2 maintenance (SSP1), if the disease is not controlled after reintroduction of FOLFIRINOX (SSP2), then the duration of disease control will be equal to the time between randomization and progression before reintroduction of FOLFIRINOX (SSP1).
- After progression during maintenance on LV5FU2 (PFS1), if the disease is again controlled after reintroduction of FOLFIRINOX, then the duration of disease control (stability, OR) will be equal to the time between randomization and progression (PFS2) on FOLFIRINOX (PFS1 + PFS2).

In the FIRGEM arm (arm C), the duration of disease control is defined as follows:

- If the discontinuation of alternating treatments (FOLFIRI.3/Gemcitabine) is due to progression (SSP1), then:
- If the disease is not controlled after treatment change, then the duration of disease control will be equal to the time between randomization and progression before treatment change (PFS1).
- If the disease is again controlled after a change of treatment, then the duration of control (stability, RO) of the disease will be equal to the time between randomization and progression (SSP2) under treatment (SSP1+ SSP2).

If the discontinuation of alternating treatments is due to a cause other than progression (toxicity, investigator's decision, etc.), then the duration of disease control will be the time to first progression (PFS1).

3.1.3 Exploratory studies

- Blood sampling study of constitutional genetic polymorphisms that can influence the efficacy and tolerance of chemotherapy molecules (UGT1A1, ERCC1, GSTT1, MTHFR, DPD, TS ...).
- Study of kerosene sections for immunohistochemical biomarkers predictive of response to treatment.

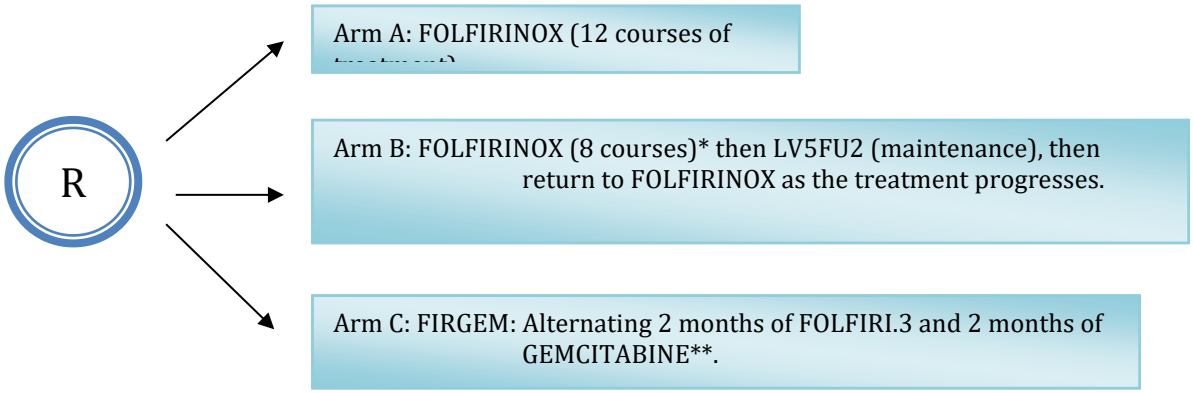
4 Experimental design

4.1 Study diagram

This is an open-label, Phase II, parallel-group (3-arm), multicenter trial.

4.2 Treatment arms

- Arm A: FOLFIRINOX (12 courses)
- Arm B: FOLFIRINOX (8 courses) followed by LV5FU2 (maintenance), then resumption of FOLFIRINOX as the treatment progresses.
- Arm C: FIRGEM: Alternating 2 months of FOLFIRI.3 and 2 months of GEMCITABINE

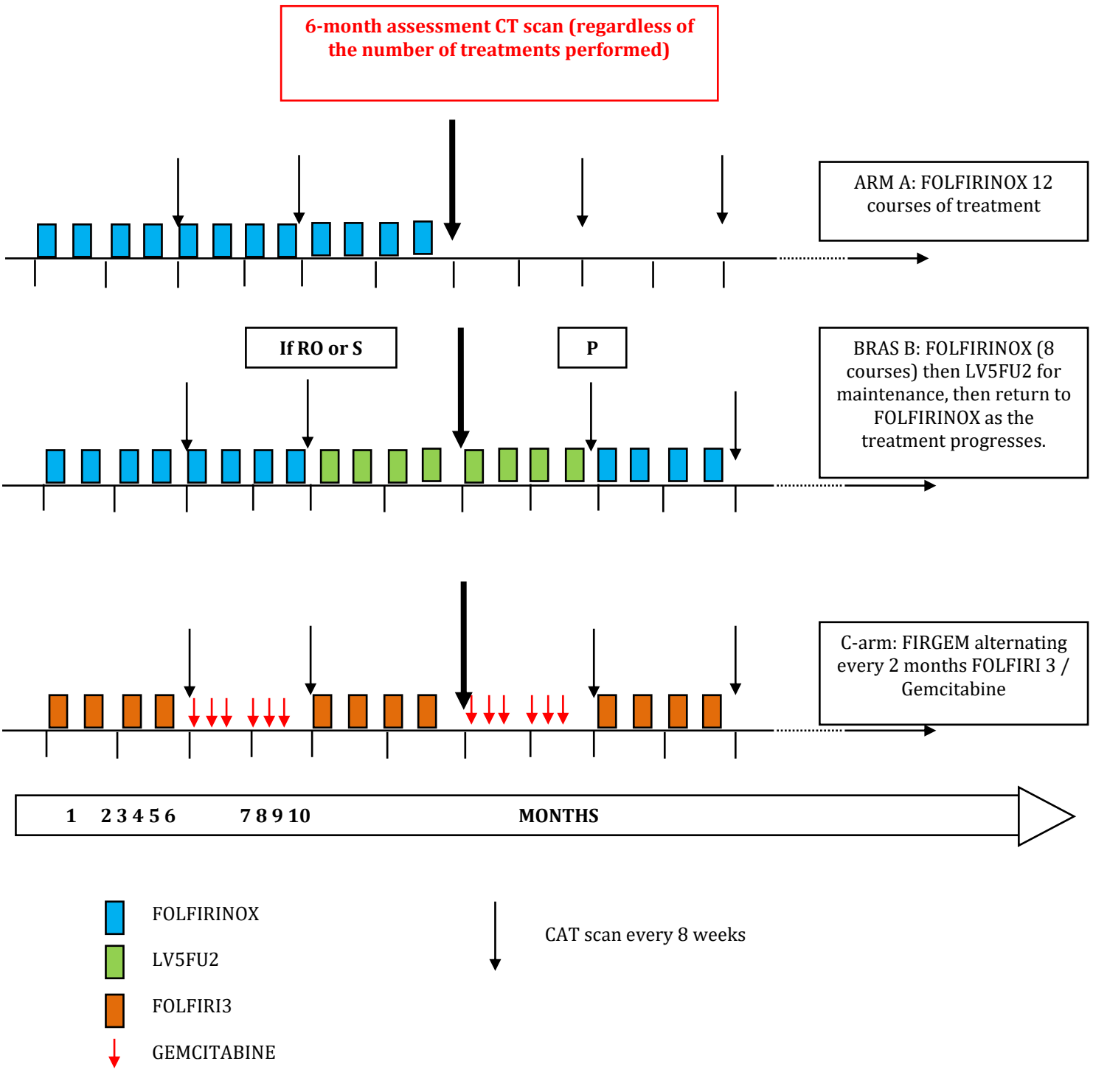


4.3 Randomization and blinding

Patients will be randomized (1:1:1) using a minimization technique based on the following **stratification factors:**

- Center
- Biliary prosthesis: Yes vs.
- Age: ≤ 65 years vs > 65 years

4.4 Chronological sequence



4.5 Justification of the number of subjects required

In this trial, two new strategies are being evaluated in the treatment of metastatic pancreatic adenocarcinoma.

The clinical assumptions used are as follows:
H_0 : A rate of patients alive and progression-free at 6 months ≤ 30% is unacceptable
H_1 : A rate of patients alive and progression-free at 6 months > 30% would be interesting. A rate of 45% is expected.

According to Fleming's one-stage design (α = 5%, power = 90%) 87 patients per arm will be randomized. Taking into account a 5% rate of lost to follow-up and non-evaluables, 92 patients will be randomized per arm, for a total of 276 patients.

The decision rules are as follows: Of the 87 evaluable patients :

- If 33 or fewer patients are alive and progression-free at 6 months, treatment is considered ineffective.
- If 34 or more patients are alive and progression-free at 6 months, the treatment is considered effective.

The decision rules will only be applied to the FOLFIRINOX arm (arm B) with maintenance and to the FIRGEM arm (arm C), with the FOLFIRINOX arm (arm A) being the reference arm.

4.6 Planning test analyses

A final analysis is planned after randomization of the 87 patients required for the arm-by-arm analysis.

4.7 Adjustments

Adjustments may be made to this analysis plan in the event of amendments to the protocol, or if phenomena not initially foreseen require statistical adaptations. In all cases, these modifications must be made before the database is frozen.

5 Study populations for analysis

5.1 Definition of analysis populations

5.1.1 Intention-to-treat (ITT) population

The intention-to-treat population is defined as all patients included in the study, regardless of eligibility criteria.

5.1.2 Modified intention-to-treat population (mITT)

The modified intention-to-treat population is defined as all patients included in the study, irrespective of eligibility criteria and having received at least one dose of treatment (whatever the molecule).
A list of patients excluded from the ITTm population will be provided with the reasons for their exclusion.

6 General information on statistical methods

Statistical analyses will be carried out by CRGA.

6.1 Software

Statistical analyses will be carried out using SAS software version 9.4. Some graphs may be produced using R software version 2.11 or later.

6.2 Agreements concerning dates and durations

Time since randomization will be defined as the time elapsed since the day of randomization, with the day of randomization taken as day 1.
As a result, durations are calculated using the following rule, for example for the time elapsed between death and inclusion: day of death - day of inclusion **+ 1**.
The day preceding the day of inclusion (resp. the day preceding the day of treatment) will be considered as day -1 (day 0 does not exist).

The date of last news will be the date of the last examination or treatment.
The following conversion rules will be used to convert numbers of days into numbers of months or years: 1 month = 30.4375 days; 1 year = 365.25 days.

6.3 Missing data conventions

Except in the cases specified, missing data will not be replaced.

6.4 Baseline definition

Baseline measurements are the last measurements taken before randomization. In the event of missing data, the last measurement taken before the first treatment is used.

6.5 Statistics

The confidence intervals provided will be two-sided 95% confidence intervals, except for the main criterion where the confidence interval calculated will be a one-sided 5% confidence interval.

Quantitative data will be described using the following descriptive statistics: headcount, mean, standard deviation, median, first and third quartile, minimum and maximum. These statistics will be considered as the usual statistics for the analysis of quantitative variables. Quantitative variables can be categorized using their median or a cut-off known from the medical literature.

Categorical variables will be summarized, using the following descriptive statistics: number, frequencies and percentages for each level of the variable. These statistics will be considered as the usual statistics for the analysis of qualitative variables.
Where necessary, confidence intervals for the proportions will be calculated from the exact binomial distribution.

For survival data :
After describing the number of events according to treatment arm, time to progression, progression-free survival and overall survival will be estimated using the Kaplan Meier method, then described by curves and rates at different time points with their 95% confidence intervals.
The median follow-up time is estimated using the inverse Kaplan-Meier method (Shemper, 1996).

Descriptive baseline analyses will be presented by treatment arm and on the total population. Efficacy and safety results will be presented by treatment arm.

No statistical comparison between arms is planned.

7 Statistical analysis

	Population ITT	Population ITTm
Description of Baseline		
Inclusion eligibility criteria	X	
Demographic characteristics	X	
Clinical features at inclusion	X	
Biological characteristics at inclusion	X	
Disease-related characteristics	X	
Median follow-up time	X	
Efficiency		
Main criterion	X	X
Duration of disease control	X	X
Overall survival, progression-free survival	X	X
Time to progress during interview	X	
Duration of maintenance treatment	X	
Tumor response rate	X	
Tolerance		
Treatment administration		X
Toxicities and frequency of use of hematopoietic growth factors		X
Quality of life		X

7.1 Patient characteristics at randomization

7.1.1 Characteristics at randomization - Eligibility

Population: ITT

Patient eligibility for randomization will be described by :

- Number and percentage of patients who met all inclusion criteria
- Number and percentage of patients who met all non-inclusion criteria
- Number and percentage of patients who met all inclusion and non-inclusion criteria
- Stratification criteria
    - Biliary prosthesis: Yes vs.
    - Age: ≤ 65 years vs > 65 years

7.1.2 Demographic characteristics

Population: ITT

The following randomization characteristics will be described:

- Age (years) ;
- Inclusion center (number of patients included per center) ;
- Gender (Male *vs* Female).

7.1.3 Clinical features

Population: ITT

The following randomization characteristics will be described:

- Weight (kg) ;
- WHO general condition (0 *vs* 1 *vs* 2).
- G8 score
- Charlson Index
- MMSE score
- GDS 15 score

7.1.4 Biological characteristics

Population: ITT

The following characteristics will be described at randomization:

- Hemoglobin (g/dL) ;
- Platelets (10^9 /L);
- Leukocytes (/mm^3) ;
- PNN (/mm^3) ;
- Lymphocytes (/mm)^3
- TP (%) ;
- Natremia (mmol/L) ;
- Kalemia (mmol/L) ;
- Potassium (mmol/L)
- Albumin levels (g/L) ;
- Glucose (g/L)
- Protein level (mmol/L)
- Total and conjugated bilirubin (µmol/L) ;
- LDH (mmol/L)
- ASAT, ALAT, PAL, GGT (by class: <Normal; 1xN < ≤3xN; > 3N or even 3xN/5xN).

7.1.5 Disease-related characteristics

Population: ITT

The following randomization characteristics will be described:

- Time since diagnosis of metastatic disease (months)
- Primary surgery
- Previous treatment and description of type of treatment
- Metastatic localization

7.2 Follow-up features

7.2.1 Definition of median follow-up time

Population: ITT

Median follow-up time is defined as the time interval between the date of randomization and the date of last news (patients alive or lost to follow-up) or, for deceased patients, the date of death (whatever the cause).

7.2.2 Evaluation of median follow-up time

The median follow-up time (for living patients) and its 95% confidence interval will be calculated in months per treatment. It will be estimated using the inverse Kaplan Meier method.

7.3 Evaluation of primary efficacy endpoint

Population: ITTm and ITT

7.3.1 Definition of primary endpoint

The primary endpoint is the rate of patients alive and progression-free at 6 months. Progression is defined as radiological progression according to RECIST v1.1 criteria and/or clinical progression according to the investigator. Progression or death (for any reason) will be considered if the event occurs within the first 6 months since randomization.

6-month scans will be 6-month scans with a +/- 1-month window

Patients with no evaluation at 6 months but later (8 months, for example) will be reviewed.

7.3.2 Evaluation of primary endpoint

The primary endpoint will be assessed by treatment arm in the ITTm and ITT populations. A one-sided 5% confidence interval will be calculated using the exact method for each treatment arm.

The decision rules will be applied to arms B and C, with arm A being the reference arm.

As a reminder, for arms B and C, on the first 87 evaluable patients (from the ITTm population):

- If 33 or fewer patients are alive and progression-free at 6 months, treatment is considered ineffective.
- If 34 or more patients are alive and progression-free at 6 months, the treatment is considered effective.

(Rules will be recalculated for the ITT population)

7.4 Evaluation of secondary efficacy endpoints

7.4.1 Duration of disease control

Population: ITTm and ITT

7.4.1.1 Definition

In arm A (FOLFIRINOX), the duration of disease control is defined as the time between the date of randomization and the date of treatment progression (clinical and/or radiological) or death (from any cause and at any time) or the date of line change.

In the maintenance arm (arm B), the duration of disease control is defined as follows:

- After progression during LV5FU2 maintenance (SSP1), if the disease is not controlled after reintroduction of FOLFIRINOX (SSP2), then the duration of disease control will be equal to the time between randomization and progression before reintroduction of FOLFIRINOX (SSP1).
- After progression during maintenance on LV5FU2 (PFS1), if the disease is again controlled after reintroduction of FOLFIRINOX, then the duration of disease control (stability, OR) will be equal to the time between randomization and progression (PFS2) on FOLFIRINOX (PFS1 + PFS2).

In the FIRGEM arm (arm C), the duration of disease control is defined as follows:

- If the discontinuation of alternating treatments (FOLFIRI.3/Gemcitabine) is due to progression (SSP1), then:
- If the disease is not controlled after treatment change, then the duration of disease control will be equal to the time between randomization and progression before treatment change (PFS1).
- If the disease is again controlled after a change of treatment, then the duration of control (stability, RO) of the disease will be equal to the time between randomization and progression (SSP2) under treatment (SSP1+ SSP2).

If the discontinuation of alternating treatments is due to a cause other than progression (toxicity, investigator's decision, etc.), then the duration of disease control will be the time to first progression (PFS1).
Event-free patients will be censored at the date of last radiological evaluation or the point date.

7.4.1.2 Evaluation

Duration of disease control will be described using standard descriptive statistics for each treatment arm.
The time scale considered is the month.

7.4.2 Progression-free survival

Population: ITT

7.4.2.1 Defining progression-free survival

It is defined as the time interval between the date of randomization and the date of 1ère radiological progression or the date of first clinical progression or the date of death (all causes combined). Patients alive and progression-free will be censored at the date of last news.

7.4.2.2 Assessment of progression-free survival

The time scale considered is the month.
Progression-free survival will be plotted using the Kaplan Meier estimator per treatment arm, and survival rates at different time points (6, 9, 12 and 18 months) will be calculated along with their 95% confidence intervals.

7.4.3 Overall survival

Population: ITT

7.4.3.1 Defining overall survival

It is defined as the time interval between the date of randomization and the date of death (whatever the cause).
Patients lost to follow-up or alive at the time of analysis will be censored at the last news date.

7.4.3.2 Assessment of overall survival

The time scale considered is the month.
Overall survival will be plotted using the Kaplan Meier estimator per treatment arm, and survival rates at different time points (6, 9, 12 and 18 months) will be calculated along with their 95% confidence intervals.

7.4.4 Time to progress during maintenance (Arm B only)

Population: ITTm

7.4.4.1 Definition

Time to progression during maintenance is defined as the time between the end of induction therapy (FOLFIRINOX or FOLFIRI3) and the date of radiological progression during the maintenance period.

7.4.4.2 Evaluation

Time to progression will be described using standard descriptive statistics for each treatment arm.

7.4.5 Duration of maintenance treatment (Arm B only)

7.4.5.1 Definition

The duration of maintenance treatment (arm B) will be calculated as the time between Day 1 of the last course of FOLFIRINOX and the date of progression leading to resumption of FOLFIRINOX.

7.4.5.2 Evaluation

The duration of maintenance treatment will be described using standard descriptive statistics for arm B only.

7.4.6 Objective response rate and best tumor response during the first 6 months of treatment

Population: ITT

7.4.6.1 Definition

The objective response rate (complete response or partial response) according to RECIST 1.1 criteria at 2, 4 and 6 months will be derived from the tumor response measured according to RECIST criteria (modified version 1.1) every 2 months.
Patients who die without an objective response (partial response or complete response = OR) or who are lost to follow-up will not be considered as responders.
The best response to treatment during the first 6 months will also be derived from radiological assessments at 2, 4 and 6 months.

7.4.6.2 Evaluation

Objective response at 2, 4 and 6 months will be described according to the categories Responder vs. Nonresponder by treatment arm. Patients with a NE response will be considered Nonresponders.

Tumor responses at 2, 4 and 6 months, as well as the best response at the first 6 months of treatment, will be described according to the following 4 categories: Complete response / Partial response / Stability / Progression/ NE by treatment arm.

7.5 Tolerance assessment

Population: ITTm

Tolerance will be assessed by treatment arm.
Tolerance to treatment will be assessed by :

- Duration of treatment, doses received, dose reductions and postponements;
- Toxicities described according to NCI-CTC criteria version 4.0 ;
- The frequency of use of hematopoietic growth factors for hematological toxicity
- Percentage of patients receiving 2ème line therapy

7.5.1 Treatment administration

7.5.1.1 Definition of treatment duration

The duration of treatment will be described by:

- Number of treatments per patient;
- The duration of treatment (converted into months) will be calculated using the formula :

  Start date of last treatment - start date of first treatment + 1 day

  Therapeutic breaks and any postponed cure days during this period will not be subtracted from this duration.

7.5.1.2 Defining administered doses

- Compliance is calculated using the formula :

  (Cumulative dose received (in mg) / theoretical protocol dose (in mg)) x 100

Protocol treatment doses depend on the treatment arm:

- FOLFIRINOX (Arms A and B)

  Oxaliplatin = 85 mg/m^2 (in mg/m^2)

  Irinotecan = 180 mg/m^2 (in French)

  5FU bolus = 400 mg/m^2 (1.5 mg/m^2)

  Continuous 5FU = 2400 mg/m^2.

- Simplified 5FU (Arm B)

  5FU bolus = 400 mg/m^2 (1.5 mg/m^2)

  Continuous 5FU = 2400 mg/m^2.

- FOLFIRI.3

  Irinotecan = 280 mg/m^2 (in mg/m^2)

  Continuous 5FU = 2000 mg/m^2

- GEMCITABINE

  1000 mg/m^2 (3 out of 4 weeks at D1, D8, D15, D29, D36 and D43)

7.5.1.3 Postponement of courses of treatment, change of administration and reason for change, discontinuation of treatment and next line

- The number and percentage of patients with at least one postponed treatment/cycle and the reasons for these postponements
- The causes of dose modification will be presented. A list of others will be provided
- Definitive discontinuation of protocol treatment and reasons for discontinuation will be described (number, %). A list of other causes of stoppage will be provided.
- The number of patients receiving a subsequent line and subsequent treatments will be described.

7.5.1.4 Assessment of treatment administration parameters

All parameters are described using standard descriptive statistics for each treatment arm.

7.5.2 Toxicities

Toxicities will be described according to NCI-CTC version 4.0 criteria per treatment arm by :

- Number of patients and toxicities by grade and type;
- The number and percentage of patients with at least one toxicity over all courses of treatment according to grade;
- The number and percentage of patients who experienced at least one maximum grade 3-4-5 toxicity over all courses of treatment or at least one maximum grade 1-2 toxicity;

7.5.3 E.I.G

The pharmacovigilance department will provide a summary of all SAEs.

7.6 Quality of life assessment and Geriatric Assessment

Population: ITT

7.6.1 Quality of Life - QLQ-C30

Quality of life will be assessed using the QLQ-C30 questionnaire.

7.6.1.1 Definition

It will be assessed using the QLQ-C30 (version 3) cancer questionnaire. The QLQ-C30 will be completed before inclusion and during the study.

Internal validation of the 30-item QLQ-C30 questionnaire identified 15 dimensions and calculated 15 scores: 5 functional ability scores (physical ability, ability to work or perform any household task, cognitive ability, emotional state, social state), a global quality of life score, a financial problems score and 8 symptom scores (fatigue, nausea/vomiting, pain, dyspnea, sleep disturbance, loss of appetite, constipation, diarrhea).

For the QLQ-C30 scores of overall health, physical ability, cognitive ability and fatigue, longitudinal evolution will be studied, along with the percentage of symptomatic progression, defined as the percentage of patients with a score reduction of at least 5 points compared with inclusion, with no subsequent improvement.

For the global health score, time to definitive deterioration will also be studied, defined as the time interval between the date of randomization and the date of a score decrease of more than 5 points (compared with the score at inclusion) without subsequent improvement of more than 5 points, or death, or the date of last news. Living patients without a score decrease of more than 5 points will be censored at last news.

7.6.1.2 Evaluation

An estimate of the time to definitive deterioration of the global health score will be made using the Kaplan-Meier method, and compared, in an exploratory way, according to the treatment arm, by a Log Rank test.

An exploratory analysis will be carried out to account for the possible non-random nature of non-completed questionnaires, using simple imputation or pattern mixture methods. Similarly, a comparison between arms will be carried out using mixed models of analysis of variance. The methods used may change as the literature in this field develops.

7.6.2 Geriatric assessments

7.6.2.1 Definition

Geriatric parameters will be assessed every 2 months until progression, then 30 days after the last treatment.

- Charlson Score (/28)
- the *Mini-Geriatric Depression Scale* (GDS), a screening tool for depression,

  Score = sum of items coded yes=1 no=0 (score between 0 and 4)

  score ≥ 1: very high probability of depression

  score = 0: very high probability of no depression

- the *Mini-Mental State Examination* (MMSE) scale, which explores cognitive functions (orientation, learning, attention and calculation, free recall, language, constructive praxis),

  The MMSE score is calculated by summing the answers to the 30 questions, and dividing by the number of items completed. If more than 15 MMSE items have not been completed, the MMSE score is coded as missing.

  Category ≤ 27/30, > 27/30

- The G8 score: composed of 8 items:

  7 "Mini-Nutrional Assessment" items and chronological age

  Scores range from 0 to 17, and a score ≤ 14/17 indicates vulnerability or fragility.

  geriatric

7.6.2.2 Evaluation

On an exploratory basis, in order to study the temporal evolution of geriatric scores and quality of life (Spitzer QoL Index, IADL, ...), mixed analysis of variance models for repeated measurements of each score in each arm.

7.7 Exploratory analysis

Not applicable

8 Quality assurance

8.1 Input

Data entry and consistency checks are described in the data management plan.

8.2 Base monitoring

8.2.1 Test analysis

During the study, one or more test analyses may be carried out. All statistical analyses may be performed on the database extracted at that time. The purpose of these analyses is to :

- anticipate the drafting of statistical analysis programs;
- highlight inconsistencies in patient records not identified by data management rules;
- identify problems not considered when drafting the protocol and statistical analysis plan.

Problems encountered on patient files may be the subject of a request for additional information. Problems identified requiring modification or clarification of the protocol or analysis plan will be discussed with the coordinator.

8.2.2 Dual programming

Double programming of the main criterion will be carried out by two different statisticians. The results will be compared, and errors discussed and corrected.