- Randomized, double-blind, phase II/III trial evaluating combination chemoembolization with Sunitinib or placebo in patients with hepatocellular carcinoma

Randomized Phase II/III

PRODIGE 16 (FFCD 0905) - SATURNE

Statistical Analysis Plan

Phase II

Version: Final

Date: February 2015

Editor: Karine Le Malicot

Review committee: Pr Mohammed Hebbar, Dr Thierry de Baere, Dr Jean-Didier Grangé, Carole Montérymard, Emilie Maillard.

Developer

FFCD

Faculty of Medicine, 7 boulevard Jeanne d'Arc, BP87900, 21079, Dijon Cedex

Tel: 03 80 39 33 87- Fax: 03 80 38 18 41

Email : cecile.girault@u-bourgogne.fr

Coordinator Pr Mohammed Hebbar

Medical Oncology Department, CHU Lille

Tel: 03 20 44 54 61

Email : m-hebbar@chru-lille.fr

Scientific co-leaderDr Thierry de Baere

Interventional Radiology Department - Institut Gustave Roussy - Villejuif

Phone: 01 42 11 54 28

E-mail : debaere@igr.fr

Scientific co-leaderDr Jean-Didier Grangé

Hepato-gastroenterology department - Hôpital Tenon - Paris

Tel: 01 56 01 70 09

E-mail : jean-didier.grange@tnn.ap-hop-paris.fr

Project Manager Ms. Marie Moreau

FFCD

Faculty of Medicine, 7 boulevard Jeanne d'Arc, BP87900, 21079, Dijon Cedex

Tel: 03 80 73 77 84

Email : marie.moreau@u-bourgogne.fr

Statistician : Ms. Karine Le Malicot

FFCD

Faculty of Medicine, 7 boulevard Jeanne d'Arc, BP87900, 21079, Dijon Cedex

Tel: 03 80 39 34 71

Email : karine.le-malicot@u-bourgogne.fr

1 Table of contents

1 Table of contents 4
2 Abbreviations and definitions 6
3 Introduction 7
3.1 Rationale for the study 7
3.2 Objectives 7
3.2.1 Pilot phase 7
3.2.2 Phase II 7
3.2.3 Phase III 7
4 Study population 8
4.1 Inclusion criteria 8
4.2 Criteria for non-inclusion 8
5 Experimental design 9
5.1 Scheme of the study 9
5.2 Study arms 9
5.3 Randomization and blinding 9
5.4 Chronological sequence 9
5.5 Justification of the number of subjects needed 11
5.5.1 Pilot phase 11
5.5.2 Phase II 11
5.5.3 Phase III 11
5.6 Steps of the test 12
5.6.1 Pilot phase 12
5.6.2 Phase II 12
5.6.3 Phase III 12
5.6.4 Transition between steps 12
5.7 Judgement criteria for phase II 13
5.7.1 Main criterion 13
5.7.2 Secondary efficacy criteria 14
5.7.3 Secondary safety criteria 15
5.7.4 Quality of life 16
5.7.5 Other criteria 16
5.7.6 Exploratory analyses 16
5.7.7 Management of criteria between Phase II and Phase III 16
6 Study populations 17
6.1 Definition of Analysis Populations 17
6.1.1 Intent-to-treat (ITT) population 17
6.1.2 Modified intention-to-treat population (mITT) 17
6.1.3 Population for Tolerance Analysis (SP) 17
6.2 Definition of analysis subgroups 17
7 Recruitment and analysis planning 17
7.1 Recruitment 17

7.2 Planning of analyses 18
7.2.1 Pilot phase 18
7.2.2 Phase II 18
7.2.3 Phase III 18
7.2.4 Transition from Phase II to Phase III 18
7.3 Adjustments 18
8 Statistical methods 18
8.1 General information on statistical analysis methods 18
8.1.1 Software 18
8.1.2 Agreements concerning dates and durations 18
8.1.3 Outlier Conventions 19
8.1.4 Missing Data Conventions 19
8.1.5 Definition of the baseline 19
8.1.6 Statistics 19
8.2 Patient characteristics at inclusion 20
8.2.1 Patient eligibility 20
8.2.2 Demographic characteristics 20
8.2.3 Clinical characteristics 20
8.2.4 Biological characteristics 20
8.2.5 Disease-related characteristics 21
8.3 Follow-up characteristics (if Phase III not done) 21
8.4 Phase II primary endpoint assessment 22
8.5 Evaluation of effectiveness 22
8.5.1 Description of chemo-embolization 22
8.6 Evaluation of tolerance 23
8.6.2 Quality of life 23
8.6.3 Other criteria 23
8.7 Common criteria for Phase II and Phase III 23
8.7.1 Progression-free survival 23
8.7.2 Overall survival 24
8.7.3 Tumor control rate at 3, 6, 9, 12 months 24
8.7.4 Remedial treatments 24
8.8 Exploratory analyses 24
9 Quality Assurance 24
9.1 Input 24
9.2 Monitoring of the base 24
9.2.1 Test analyses 24
9.2.2 Forecast on the analyses 25
9.2.3 Dual programming 25
10 Bibliographic references 26

2 Abbreviations and definitions

ALAT : Alanine transaminases

ASAT : Aspartate transaminases

HCC : HepatoCellular Carninoma

VEF : Ventricular Ejection Fraction

LSN : Upper Normal Limit

WHO : World Health Organization

RECISTm : RECIST Modified

TACE: Transarterial chemoembolization

TSH Thyroid Stimulin

AFP Alpha Foeto Protein

3 Introduction

3.1 Rationale for the study

Transarterial chemoembolization (TACE) is a commonly used technique in the treatment of patients with isolated or small hepatocellular carcinoma not amenable to surgical resection (intermediate stage), with an advantage in terms of survival time over symptomatic treatments. This approach is justified by the fact that HCC is a highly vascularized tumor, with a major role of angiogenesis.

Sunitinib is a novel multi-targeted tyrosine kinase inhibitor with both angiogenesis inhibition and direct anti-tumor properties. The addition of sunitinib may therefore reduce the risk of tumor recurrence in HCC patients treated with TACE. Nevertheless, some angiogenesis inhibitors are associated with an increased risk of bleeding complications, which may raise concerns about their use in patients with HCC. Thus, it is important to evaluate the safety of the sunitinib-chimioembolization combination, and compare it to chemoembolization alone (chemoembolization plus placebo).

3.2 Objectives

3.2.1 Pilot phase

To assess the rate of bleeding and/or severe liver failure within one week of chemoembolization sessions in patients receiving sunitinib.

3.2.2 Phase II

3.2.2.1 Main objective

Assess the rate of bleeding and/or severe liver failure within one week of chemoembolization sessions.

3.2.2.2 Secondary objectives

The secondary objectives of the phase II study are to evaluate:

- tumor control rate* at 3, 6, 9 months after the last TACE ;
- Progression-free survival
- Overall survival*;
- Tolerance*;
- Quality of life*.

*These objectives are common objectives with la phase III

3.2.3 Phase III

3.2.3.1 Main objective

Compare overall survival in the sunitinib and placebo groups.

3.2.3.2 Secondary objectives

- The common criteria between la Phase IIand la Phase IIIare described in Section **3.2.2.2 (marked with an *).**
- Disease-free survival

4 Study population

4.1 Inclusion criteria

- Histologically proven hepatocellular carcinoma (HCC) or liver tumor meeting the Barcelona criteria.
- Child-Pugh A score.
- Tumor not accessible to surgical resection or radiofrequency.
- Tumor undergoing transarterial chemoembolization (TACE).
- Age ≥ 18 years.
- Performance Index (WHO) ≤ 2.
- A previous radiofrequency treatment is allowed if it took place more than 3 months before.
- Satisfactory hematologic, renal, and hepatic function (neutrophils ≥ 1.5 x109/L, platelets ≥ 70 x $^{109/L}$, hemoglobin ≥ 10 g/dL, prothrombin time ≥ 50%, creatinine level ≤ 120 µmol/L, bilirubin level (≤ 15 mg/L), alanine and aspartate transaminases (ALAT and ASAT) ≤4 times the upper limit of normal (ULN), alkaline phosphatases ≤ 5 times ULN, and fibrinogen level ≥1.5 g/L.
- Signed written consent.

4.2 Non-inclusion criteria

- History of chemoembolization.
- Portal thrombosis.
- Extra-hepatic metastases including brain metastases.
- Uncontrolled hypertension or requiring at least 2 classes of antihypertensive agents (uncontrolled hypertension (systolic pressure > 160 160 mmdiastolic pressure > 100 mm100 mmwell-conducted antihypertensive treatment)
- Uncontrolled heart failure (left LVEF < 50%) or coronary artery disease, recent myocardial infarction (less than one year), QT interval prolongation > 450 ms in men and > 470 ms in women.
- Concomitant disease or severe uncontrolled clinical situation.
- Uncontrolled severe infection.
- Patient treated with a CYP3A4 inhibitor within 7 days prior to treatment.
- Patient treated with a CYP3A4 potentiator within 12 days.
- Patients requiring long-term anticoagulant treatment.
- Pregnancy or breastfeeding.
- Lack of effective contraception (for men or women of childbearing age).
- Prior treatment with sunitinib, sorafenib or other angiogenesis inhibitors.
- History of other cancers excluding cancers known to be cured for more than 5 years (in this case histological evidence of HCC is required), or basal cell skin tumors or cervical cancer in situ adequately treated with curative intent.
- Patient who for psychological, social, family or geographical reasons could not be followed regularly.
- Concurrent participation of the patient in another experiment.
- Patient with a contraindication to vascular occlusion procedures.
    - Vascular anatomy that precludes catheterization or injection of emboli.
    - Presence or likely onset of vasospasm.

- Presence or probable occurrence of a hemorrhage.
- Presence of severe atheromatous damage.
- Presence of feeder arteries smaller than the distal branches from which they emerge.
- Presence of extra to intracranial anastomoses or patent shunts.
- Presence of collateral vascular pathways potentially endangering normal territories during embolization.
- Presence of terminal arteries leading directly to the cranial nerves.
- Presence of arteries irrigating the lesion not wide enough to receive the microbeads.
- Peripheral vascular resistance of the feeding arteries excluding the passage of microbeads into the lesion.
- Do not use microbeads in the following applications:
  - embolization of non-malignant tumors;
  - embolization of large-diameter arteriovenous shunts (i.e., where the blood does not pass through the arterial/capillary/venous transition but goes directly from the artery to the vein);
  - any vascular system where the microbead embolic agent is CN00103.2 likely to pass directly into the internal carotid artery or other non-target territory.

5 Experimental design

5.1 Scheme of the study

The study is a phase II-III, parallel-group, multicenter study.

5.2 Study arms

- Arm A: chemoembolization plus sunitinib.
- Arm B: chemoembolization plus placebo.

5.3 Randomization and blinding

The study is double-blind for sunitinib or placebo (neither the doctor nor the patient knows which treatment is being given).

Stratification by minimization is performed, starting in la phase II, on:

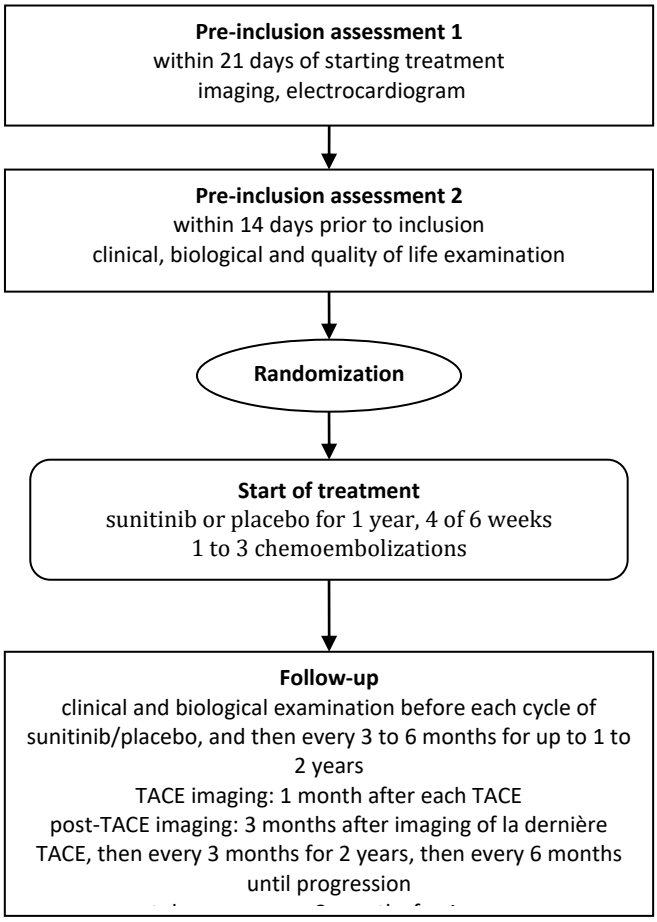
- diameter of the primary tumor (< 5 versus ≥ 5 5 cm
- the presence of one nodule versus the presence of more than one nodule;
- the investigating center.

When a stratified analysis will be performed, it will be performed only according to the diameter of the main tumor and the number of nodules, according to 4 groups:

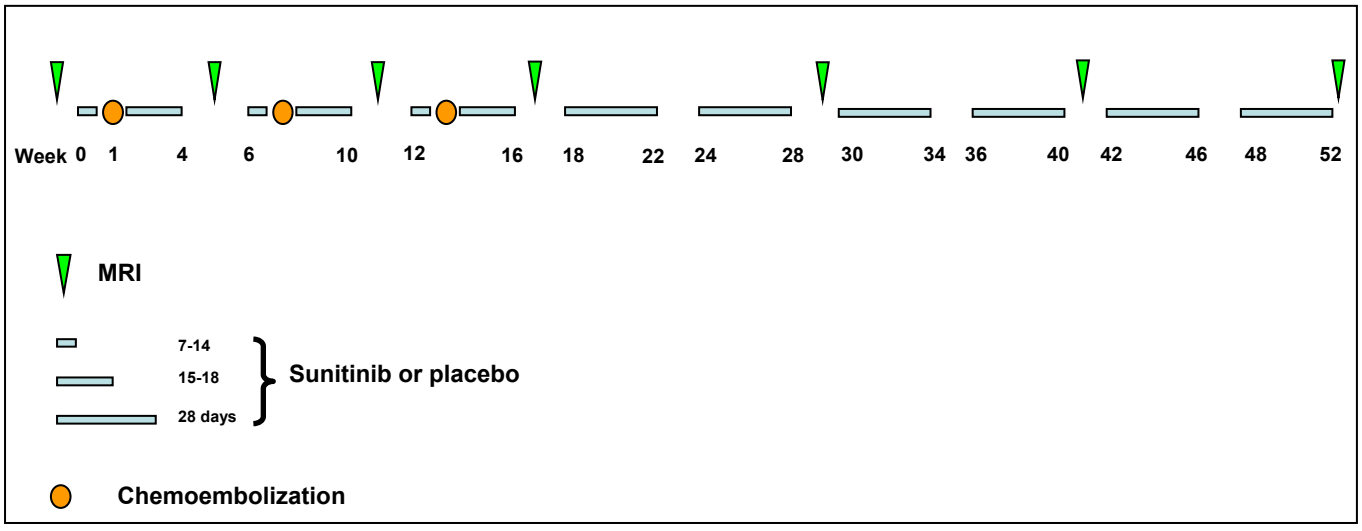
- diameter < 5 cm and a nodule,
- diameter ≥ 5 cmand a nodule,
- diameter < 5 cm and more than one nodule,
- diameter ≥ 5 cmand more than one nodule.

5.4 Chronological sequence

The chronological sequence for a patient is described in the following graph:



The study design is as follows:



5.5 Justification of the number of subjects needed

5.5.1 Pilot phase

A non-randomized pilot phase was performed on 10 chemoembolizations performed with sunitinib (regardless of the number of patients included). An independent committee ruled on the safety data for continuation or not in phase II and a specific report was sent to the ANSM for information.

5.5.2 Phase II

The assumptions for calculating the number of subjects needed are:

- H0: A rate of bleeding and/or severe liver failure occurring within one week of the chemoembolization session of 30% is considered unacceptable;
- $_{H1}$: A rate of bleeding and/or severe liver failure occurring in the week following the chemoembolization session of less than 30% is necessary to consider the treatment non-toxic; a rate of 15% is expected.

To achieve 85% power (assuming a 15% rate of severe bleeding and/or liver failure), with a two-sided first-step risk of 13.3%, 35 patients per arm are required (in a one-step Fleming design).

5.5.3 Phase III

The assumptions for calculating the number of subjects needed are:

- H0: Overall survival in the sunitinib arm is not different from that in the placebo arm; expected 2-year overall survival in the placebo arm is 40%;
- $_{H1}$: Overall survival in the sunitinib arm is different from that in the placebo arm; a 20% difference at 2 years in favor of sunitinib is expected between the two arms (corresponding to a hazard ratio of 0.56).

With a two-sided 5% risk of first-species error, 130 events will be required to achieve 90% power under the previously stated assumption $_{H1.}$ Assuming uniform patient inclusion over 2 years, and allowing for a 3-year follow-up after the last inclusion, the number of patients will be 180. To account for a 5% rate of patient attrition, a total of 190 patients will need to be included.

The 70 patients already included in la phase IIwill be retained for la Phase III, leaving 120 patients to be included.

An interim analysis of the primary endpoint is planned at death 65. This may provide early evidence of superiority of the sunitinib arm. A method of expenditure of the first kind risk (DeMets, 1994), will be used to calculate the nominal values of "p"; this method will make it possible to adjust these threshold values if the intermediate and final analyses are not carried out exactly at the right numbers of events. The expenditure function chosen is an O'Brien function.

Calculations were performed using the R software and the gsDesign library.

Analysis	Number of events	Nominal value of p	Power to the analysis under hypothesis H1
Inter 1 analysis	65	0.0054	0.3156
Final analysis	130	0.0439	0.5811

These values will be adjusted according to the number of actual events at the time of the analyses.

5.6 Stages of the test

5.6.1 Pilot phase

During the pilot phase, 10 chemoembolizations (8 patients included) were performed with sunitinib. Of these 10 chemoembolizations, less than 4 toxicities were retained for the primary endpoint (see protocol, Amendment 3), the trial continued in randomized phase II. **Patients included in the pilot phase are not included in la phase II**.
The pilot phase was the subject of a specific report which was sent to the ANSM for information

5.6.2 Phase II

At the end of the 35 patients/arm inclusions of the phase II:

- if the number of patients with severe toxicity (liver failure, bleeding) in the sunitinib arm is less than or equal to 7, the trial will be continued in phase III;
- if the number of patients with severe toxicity (liver failure, bleeding) in the sunitinib arm is greater than or equal to 8, the number of complications (liver failure, bleeding) in the placebo arm should also be considered:
    - if the complication rate (liver failure, bleeding) in the placebo arm is greater than or equal to 8, an independent committee of experts will have to judge the appropriateness of continuing the trial in phase III since the toxicity in the sunitinib arm may be related to the "background noise" of the toxicity of chemoembolization alone;
    - if the complication rate (liver failure, bleeding) in the placebo arm is less than or equal to 7, the independent expert committee will have to judge whether to stop the trial.

The decision rules will apply only to the sunitinib/chimio/embolization arm; the placebo/chimio/embolization arm will be maintained in all cases for la phase III, if it occurs, with this arm serving as the control arm.

5.6.3 Phase III

This step will only be performed if the trial has not been stopped at the end of la Phase II.

5.6.3.1 Intermediate Analysis

An interim analysis at 65 deaths is planned. This may provide early evidence of superiority of the sunitinib arm.

5.6.3.2 Final Analysis

The final analysis will be done at approximately the 130th event (death), after inclusion of all 180 patients and a minimum follow-up for the last included patient of 3 years.

5.6.4 Transition between steps

A suspension of inclusions is planned between la phase IIand la phase IIIresults of la phase II.analyzed.
Inclusions between the 2 stages of la phase IIIwill not be suspended for the time of the interim statistical analysis.

5.7 Judgement criteria for la phase II

5.7.1 Main criterion

The primary endpoint of la phase IIis the rate of bleeding and/or severe liver failure within one week of chemoembolization sessions (TACE), over all chemoembolization sessions for a patient. The precise definition of bleeding and liver failure to be retained is given in the protocol with amendment 3 :

- Severe liver failure is defined by the occurrence of at least one of the following complications:
  - encephalopathy;
  - clinical ascites; in case of clinical ascites already present at inclusion and persistence of ascites at the time of evaluation for liver failure, an independent committee will have to judge from the case report whether ascites has evolved since inclusion and should be counted as an event in the definition of the primary endpoint or not; if clinical ascites is present at inclusion, it will be graded according to NCI-CTC v4 and reported through a CID; during follow-up, clinical encephalopathy and ascites will be graded according to NCI-CTC v4 and will appear on form 3.1 of the case report form in the other toxicities tab
  - bilirubin levels greater than or equal to 50 mg/L (85.5 µmol/L);
  - TP ≤ 40%;
- bleeding : (definition p14 of the protocol)

These events count as primary endpoints if they occur within 1 week of chemoembolization, i.e., between D0 and D7 inclusive.
In the case of an event counting for the primary endpoint after D7, it will be recorded in the case report form 3.1 in the column after D5 only if the investigator considers that this event is related to the TACE, and the day must then be filled in. For events recorded after D7 in the case report form, an independent committee will meet to decide whether or not this event counts in the definition of the primary endpoint.

Patients who did not receive their oral treatment will still be counted in the primary endpoint; however, a very precise description of the reason for not taking the oral treatment will be made.

Patients who did not have chemoembolization will be excluded from this analysis (modified ITT).

5.7.2 Secondary efficacy criteria

5.7.2.1 RECISTm definition of treatment response

This definition will be used for progression-free survival, disease-free survival and tumor control rate.

Progression of target and non-target lesions will always be measured relative to the inclusion MRI.

- For evaluations following TACE: treatment response will be assessed only for lesions (target and non-target) treated during TACE according to RECISTm criteria. Progression is defined as either a treated lesion that progresses or the appearance of a new hepatic or extrahepatic lesion. Progression related to untreated lesions will not be taken into account with the exception of one case: progression of an untreated lesion that prevents further treatment with TACE (this information will be reported manually by the CRAs).

- For follow-up evaluations (after TACE is completed), response to treatment will be assessed in relation to all lesions, whether treated during TACE or not.

5.7.2.2 Tumor control rates at 3, 6, 9, 12 months

It is defined as the percentage of patients in Complete Response or Partial Response or in stability at the given time.
Patients for whom tumor assessments would not allow classification will be reviewed by a clinician.
Patients who died will be considered uncontrolled.

5.7.2.3 Progression-free survival (if Phase III not done)

It is defined as the time interval between the date of randomization and the date of first progression or death (regardless of cause). Progression is defined as progression of lesions according to RECISTm criteria, the appearance of a new liver lesion or an extra-hepatic lesion.
Since progress can be documented on the in-process assessments or the monitoring form, the definition in 5.7.2 applies. The dates to be considered are the dates of the CTs.
Live patients without progression will be censored at the date of the last informative examination or at the date of last news if a surveillance sheet is provided not documenting progression or at the point date.

Clinical progressions will not be taken into account, based on the principle that for this type of pathology, a progression in imaging is necessarily associated with a clinical progression.

5.7.2.4 Disease-free survival (if Phase III not done)

It is defined as the time interval between the date of randomization and the date of progression, or the date of a second cancer, or death (from any cause).
Since progression or second cancer may be documented on in-process assessments or the surveillance form, the definition in Section 5.7.2 applies. The dates to be considered are the dates of the CT scans.

Living patients without progression or second cancers will be censored at the date of the last informative review or at the date of last news if a surveillance form is provided that does not document progression or second cancers at the point date.

5.7.2.5 Overall survival (if Phase III not done)

Overall survival is defined as the time from the date of randomization to the date of death (from any cause).

Patients lost to follow-up or alive at the time of analysis will be censored at the last news date or the point date.

This time will also be used to calculate the median follow-up time.

5.7.3 Secondary tolerability criteria

Tolerance to treatment will be assessed by:

- chemo-embolization
- duration of sunitinib/placebo treatment, number of tablets administered, dose reductions, treatment discontinuations and their causes ;

The duration of the treatment (converted into months) will be calculated by adding the cycle durations: (cycle end date - cycle start date) + 1

- Toxicities, collected at the end of each treatment cycle, will be described according to the NCI-CTC version 4.0 criteria and their grade;
- the time to onset of grade 3-4 toxicity will be assessed;

It is defined as the time interval between the date of randomization and the date of first grade 3-4 toxicity estimated by the cycle start date + 2 weeks. Patients without Grade 3-4 toxicity at the earlier of the point or last date will be censored at the point or last date;

- The description of SAEs (analysis of pharmacovigilance)

5.7.4 Quality of life

It will be assessed using the QLQ-C30 (version 3) questionnaire for cancer and the FACT-HEP questionnaire specific to HCC.

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).

The FACT-HEP (version 4) is composed of the 27 items of the FACT-G evaluating 4 dimensions (physical, social and family, emotional, functional) and 18 specific items (HepCS) allowing to evaluate the specific symptoms of HCC and the side effects of its management. Three scores will be calculated: the FACT-Hep TOI, taking into account the physical, functional, and specific symptoms of HCC; the FACT-G Total score, taking into account the physical, functional, emotional, social, and family dimensions; and the FACT-Hep Total score, taking into account all dimensions.

These different scores will be described at inclusion and during follow-up.

For the QLQ-C30 Global Health, Physical Ability, Cognitive Ability, and Fatigue scores, and the FACT-Hep TOI, FACT-G Total score, and FACT-Hep Total score, longitudinal progression will be studied, with the percentage of symptomatic progression defined as the percentage of patients who had a score decrease of at least 5 points from inclusion without subsequent improvement (Osoba, 1998). Patients lost to follow-up without a decrease in score of at least 5 points from inclusion will be considered without symptomatic progression. For the overall health score, FACT-Hep TOI, FACT-G Total score, and FACT-Hep Total score, time to definite deterioration, defined as the time interval between the date of randomization and the date of a score decrease of more than 5 points (from the inclusion score) without subsequent improvement of more than 5 points, will also be studied. Patients alive or deceased without a score decrease of more than 5 points will be censored at the point date or the date of last news if earlier. For these last three criteria, patients who did not have a quality of life questionnaire at inclusion will be excluded (their number and percentage will be given).

5.7.5 Other criteria

- The evolution of AFP rates over time.

5.7.6 Exploratory analyses

Not Applicable

5.7.7 Management of criteria between la phase II and la phase III

The common criteria for la phase IIand la phase III(Section 5.7.2 and 5.7.4.) will be reviewed at the end of la phase III, or at the end of la phase II if la phase IIIis not performed. If these criteria are only assessed in la phase II, there will be no comparison made between arms.

In order not to introduce operational bias, all criteria can be described in la phase II(common or not to la phase III), but the common criteria must remain confidential.

In all cases, a description of the TACE treatment and a description of the catch-up treatments will be made.

6 Study populations

The study focuses on patients with hepatocellular carcinoma.

6.1 Definition of Analysis Populations

6.1.1 Intent-to-treat (ITT) population

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

The description of the population at inclusion will be based on this population.

6.1.2 Modified intention-to-treat population (mITT)

The modified IIT population is defined as all patients included in the study, regardless of eligibility criteria and amount of treatment taken who had at least one chemoembolization.

The primary endpoint will be assessed in this population.

6.1.3 Population for Tolerance Analysis (SP)

It is defined as la population ITTthat received at least one day of treatment with sunitinib/placebo (no constraints on chemoembolization).

The analysis of secondary efficacy endpoints will be done in this population (depending on the randomization arm)

Patients who did not take treatment will be described by individual data.

The analyses for tolerability will be performed according to the actual treatment received. Thus, if despite randomization, a patient received placebo instead of sunitinib from baseline to the end of treatment, this patient will be analyzed for safety in the placebo arm (and vice versa).

However, if a change of treatment (sunitinib/placebo) occurs during treatment, the patient will be analyzed in the group corresponding to the first treatment he/she received.

6.2 Definition of analysis subgroups

Not applicable

7 Recruitment and analysis planning

7.1 Recruiting

The expected rate of inclusion is 7 to 8 patients per month.

7.2 Analysis planning

7.2.1 Pilot phase

The analysis of the pilot phase was conducted in April 2011.

7.2.2 Phase II

The cut-off date for the la phase IIanalysis will be one week after the completion of the last chemoembolization of the last included la phase II.patient.

7.2.3 Phase III

7.2.3.1 Intermediate Analysis

An interim analysis is planned during la phase III. Cetteanalysis is planned at the midpoint of the number of events, i.e. at 65 deaths. Based on the assumptions made, this analysis could take place 28 months after the inclusion of the first patient.

7.2.3.2 Final Analysis

The point date for the final analysis is the 130th event.

7.2.3.3 Late analysis after Phase III

A final analysis will be performed to account for overall survival at 4 years. The time point is 4 years after inclusion of the last patient.

7.2.4 Transition from Phase II to Phase III

In case of transition to phase III, only results concerning the primary endpoint, safety data, and control rates during each TACE will be reported for la phase II

7.3 Adjustments

Adjustments may be made to this analysis plan in case 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.

8 Statistical methods

Statistical analyses will be performed by the CRGA.

8.1 General information on statistical analysis methods

8.1.1 Software

Statistical analyses will be performed with SAS version 9 or later and STATA version 10 or later. Some graphs can be made with R software version 2.11or later.

8.1.2 Agreements concerning dates and durations

Time since randomization will be defined as the time since the day of randomization, with the day of randomization considered day 1.

Time since initiation of treatment will be defined as the time since the first day of sunitinib/placebo, which is considered Day 1.
Therefore, the durations will be calculated using the following rule, for example, for the time between death and randomization: day of death - day of randomization **+ 1**.
The day before the randomization day (resp. the day before the treatment day) will be considered as day -1 (day 0 does not exist).

The date of last news will be the date mentioned on the monitoring form, and in case of missing information, the date of the last examination performed or the last treatment.

The following conversion rules will be used to convert the number of days into months or years: 1 month = 30.4375 days; 1 year = 365.25 days.

8.1.3 Outlier Conventions

Outliers will be subject to a confirmation request to the investigating center. In case of confirmation, their value will not be modified and will be taken into account as it is during the analysis.

8.1.4 Missing Data Conventions

Except in the cases specified, missing data will not be replaced.
In case of partially missing data concerning dates, the following rule will be applied: when the day is missing, the day will be considered as the 15th of the month.

8.1.5 Definition of the baseline

Baseline measurements will be the last measurements taken before la randomisation. Encase of missing data, the last measurement taken before the first treatment will be used.

8.1.6 Statistics

The confidence intervals provided will be two-sided 95% confidence intervals.

Quantitative data will be described by treatment group and for the entire population, using the following descriptive statistics: number of patients, number of missing values, mean, standard deviation, median, first and third quartile, and minimum and maximum. These statistics will be considered as the usual statistics for the analysis of quantitative variables. Quantitative variables may be categorized using their median or a cut-off known from the medical literature.

Categorical variables will be summarized by treatment group and for the entire population, using the following descriptive statistics: frequencies and percentages for each level of the variable and missing values (missing values will be included in the denominator of the frequency calculation). These statistics will be considered the usual statistics for the analysis of categorical variables.
When necessary, confidence intervals for the proportions will be calculated from the exact binomial distribution.

The estimation of survival for the censored variables will be done by the Kaplan Meier method (Kaplan and Meier, 1958). This will be described by the median and rates calculated at different times. The 95% confidence intervals will

be provided. Confidence intervals for the rates will be constructed from the Greenwood variance calculated using the log-log transformation. For the median confidence interval, the upper bound will be defined as the smallest time for which the upper bound of the associated survival rate confidence interval, calculated from the Greenwood method, is less than or equal to 50%. Similarly, the lower bound will be the smallest time whose lower bound of the confidence interval of the associated survival rate is less than or equal to 50%.

The estimation of the treatment-related hazard ratios will be carried out using a stratified Cox model (Cox, 1984); the proportionality hypothesis of the rates will be tested using the graphical representation and the test based on Schöenfeld residuals (Grambsch, 1994); the linearity of the effect of the continuous variables on the risk will be assessed from the graphical representation of the martingale residuals.
Confidence intervals for the coefficient estimates of the Cox models will be calculated using the Wald method.

The median follow-up time will be calculated using the reverse Kaplan-Meier method (Shemper, 1996).

Unless otherwise stated, **results will be described separately for each treatment arm**; therefore, when "total number of patients" is specified, the total number of patients in the arm should be considered.

8.2 Patient characteristics at inclusion

Patient characteristics at inclusion will be described by treatment group and on the total population.

8.2.1 Patient Eligibility

A description of the stratification factors will be performed to ensure proper distribution of prognostic factors, but no statistical testing will be performed :

- diameter of the main tumor (< 5cm vs. ≥ 5cm)
- number of nodules (1 vs > 1)
- inclusion center (Number of patients included per center)

Similarly, a description of the inclusion and non-inclusion criteria will be provided

8.2.2 Demographic characteristics

The following characteristics at inclusion will be described using descriptive statistics (no statistical tests):

- Age (years);
- Gender

8.2.3 Clinical characteristics

The following characteristics at inclusion will be described using descriptive statistics (no statistical tests):

- weight (kg) ;
- blood pressure (mmHg); if more than one SBP or DBP is indicated then the values will be averaged
- General condition WHO (0, 1, or 2)

8.2.4 Biological characteristics

The following characteristics at inclusion will be described using descriptive statistics (no statistical tests):

- TSH (mU/l);
- AFP (µg/l).

8.2.5 Characteristics related to the disease

The following characteristics at inclusion will be described using descriptive statistics (no statistical tests):

- CHILD-PUGH score (5 or 6);
- HCC involvement: unilobar or bilobar;
- number of hepatic nodules;
- diagnostic method: histological or according to the Barcelona criteria;
- Associated cirrhosis: yes or no; if yes, etiology: alcohol, hemochromatosis, HBV, HCV or other (A listing of other etiologies will be edited).

8.3 Follow-up characteristics (if Phase III not done)

The median follow-up time and its 95% confidence interval will be calculated in months for each treatment arm using the Inverse Kaplan-Meier method.

8.4 Evaluation of the primary endpoint of la phase II

For the primary endpoint analysis, the number of patients with at least one bleed and/or liver failure, the percentage of patients, and the 95% confidence interval will be calculated by treatment group:

Similarly, a more precise analysis will be performed describing, across all TACEs and by treatment group:

- the number of TACEs during which severe bleeding was observed;
- the number of TACEs during which at least one liver failure was noted;
- the number of TACEs during which at least one clinical ascites was noted ;
- the number of TACEs during which at least one encephalopathy was reported;
- The number of TACEs during which at least one bilirubin assay ≥ 50 mg/L was reported;
- the number of TACEs during which at least one decrease in prothrombin time of more than 40% from baseline was reported.

8.5 Evaluation of effectiveness

(See Section 8.7)

8.5.1 Description of chemo-embolization

The description of chemoembolization will be done by treatment group.

8.5.1.1 Description of the treatment

For each TACE (1, 2, 3 or more) will be described:

- the number of patients who received this TACE;
- the percentage of patients (from TACE) with all target lesions treated; ;
- the percentage of non-target lesions treated ;
- the treated area: total liver, lobar, sectorial, segmental, sub-segmental.
- A listing of reasons why CEAT was not done will be provided per patient

8.5.1.2 Response to treatment during treatment

For each of the TACE numbers (1, 2, 3 or more) will be described:

- The overall response to TACE for treated target lesions and/or for treated non-target lesions;
- the number of TACE followed by at least one new liver injury
- The number of TACE followed by at least one new extrahepatic lesion;
- The number of TACEs with the development of a second cancer, and the number of TACEs requiring additional TACE),

8.5.1.3 Toxicities during the days following HEC

- A description of the toxicities reported on the HEC charts will be made according to grade and treatment arm.

8.6 Assessment of tolerance

The assessment of tolerability will be described by treatment group.

8.6.1.1 Treatment administration

- The duration of treatment with sunitinib/placebo described using standard statistics
- The percentage of patients with at least one dose reduction will be calculated.
- For all treatments administered (at least one treatment), the complete or incomplete administration will be described as well as the causes if applicable (no reduction, toxicity, therapeutic pause, personal convenience of the patient, other; a listing of other causes will be provided)
- In case of premature discontinuation of treatment (sunitinib or placebo), the causes of discontinuation will be described; a listing of other causes will be provided.

8.6.1.2 Serious adverse events

- A document will be produced by Pharmacovigilance.

8.6.1.3 Toxicity analysis

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

- by grade and type at the end of cycles during TACE treatment, over all cycles and during follow-up;
- Patients with at least one grade 3-4 toxicity of any type (during cycles including TACE and over all cycles), with 95% confidence interval;
- Time to Grade 3-4 toxicity will be (Toxicities collected during HEC and cycles) plotted with 95% confidence intervals and rates at different times (time to be adapted according to data) and their 95% confidence intervals.

8.6.2 Quality of life

- Each of the dimensions will be described at inclusion and at the end of follow-up; in addition, for each dimension, the difference between the score at inclusion and the score at the last visit will be described.
- The time without definite deterioration for each of these scores will be described by a survival curve with 95% confidence intervals.

8.6.3 Other criteria

- The evolution of AFP rates will be represented graphically.

8.7 Criteria common to la Phase II and la Phase III

These endpoints will only be studied in la Phase II if la Phase III not occur. Comparisons between arms will only be made in Phase III.

8.7.1 Progression-free survival

Progression-free survival will be compared between the 2 treatment groups (Phase III) or described (Phase II) using the stratified logrank test and using the simple logrank test.

The estimation of treatment-related hazard ratios will be performed using a stratified Cox model (Cox, 1984);

8.7.2 Overall survival

Overall survival will be compared between the 2 treatment groups (Phase III) or described (Phase II) using the stratified logrank test and using the simple logrank test.

The estimation of treatment-related hazard ratios will be performed using a stratified Cox model (Cox, 1984);

8.7.3 Tumor control rates at 3, 6, 9, 12 months

- The percentages will be from the Kaplan-Meier analysis as well as the confidence intervals.

8.7.4 Remedial treatments

- The number of patients who had at least one salvage treatment, and the type of salvage treatment: line 1 treatment, line 2 treatment, line 3 treatment, radiotherapy, radiofrequency, surgery, other.
- For each patient who had a 1st line, the average duration of the line will be calculated, and a listing of treatments will be provided. The description of the 2nd and 3rd line will be done in the same way.
- For each patient who had at least one radiation treatment, the number of radiation treatments will be described, and the average duration of radiation treatments will be calculated.
- For each of the patients who had radiofrequency treatment, the number of sessions will be described
- For patients who had surgery, the type of resection will be described. The number of patients who had a transplant will also be calculated.
- For each patient who has had other remedial treatment, a listing describing the treatments will be provided, and the average duration of treatment will be calculated.

8.8 Exploratory analyses

Not Applicable

9 Quality Assurance

9.1 Input

The input and consistency check are described in the data management plan.

9.2 Monitoring of the base

9.2.1 Test analyses

At several times during the study, test analyses will be performed. All statistical analyses will be performed on the database extracted during the test analysis. The purpose of these analyses is to:

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

Problems encountered on the patient charts will be discussed with a clinician; if no answer is provided by the clinician, they will eventually be the subject of a request for additional information.

Identified problems requiring modification or clarification of the protocol or analysis plan will be discussed with the coordinator.

9.2.2 Forecast on analysis

Phase	Period	Action	Objectives
Phase II	As soon as possible	Test analysis	▪ main criterion ▪ all study criteria (phase II and III, including quality of life)
	35 patients per arm + 1 week after the latest TACE of last patient on treatment	Pre-gel base Test analysis for base gel	▪ main criterion ▪ all the criteria of the study (phase II and III, including quality of life), ▪ basic freeze decision
		Base gel Phase II analysis	▪ main criterion ▪ tolerance criteria

9.2.3 Double programming

At least a double programming of the main criterion will be done by two different statisticians. Other criteria may be double-programmed (at the discretion of the statistician). The results will be compared, and errors discussed and corrected.

10 Bibliographic references

Bouvier AM, Jooste V, Bonnetain F, Cottet V, Bizollon MH, Bernard MP and Faivre J. Adjuvant treatments do not alter the quality of life in elderly patients with colorectal cancer. Cancer. 2008 ; 113 : 879-886.

Cox DR and Oakes D. Analysis of Survival Data. 1984 ; London Chapman & Hall.

DeMets DL and Lan KK. Interim analysis: the alpha spending function approach. Statistics in Medicine. 1994; 2:253-256.

Fairclough DL, Peterson HF, Cella D and Bonomi P. Comparison of several model-based methods for analyzing incomplete quality of life data in cancer clinical trials. Statistics in Medicine. 1998 ; 17 : 781-796.

Fleming TR. One-sample multiple testing procedure for phase II clinical trials. Biometrics. 1982; 38: 143-151.

Freedman L.S. Tables of the number of patients required in clinical trials using the logrank test. Statistics in Medicine. 1982 ; 1:121-129.

Grambsch P. and Therneau T. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika. 1994; 81, 515-526.

Kaplan EL and Meier P. Non-parametric estimation from incomplete observations. Journal of the American Statistical Association. 1958 ; 53 : 457-481.

Mantel N. Evaluation of survival data and two new rank order statistics arising and its consideration. Cancer Chemotherapy Reports. 1966 ; 50 : 163-170.

Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. Journal of Clinical Oncology. 1998 ; 16(1) :139-144.

Schemper M and Smith TL. A note on quantifying follow-up in studies of failure time. Control clinical trials. 1996 ; 17 : 343-346.