# STATISTICAL ANALYSIS PLAN

**Celsion Corporation**
997 Lenox Drive, Suite 100
Lawrenceville, NJ 08648 USA

## A Phase III, Randomized, Double Blind, Dummy-Controlled Study of ThermoDox® (Lyso-Thermosensitive Liposomal Doxorubicin-LTLD) in Hepatocellular Carcinoma (HCC) using standardized Radiofrequency Ablation (RFA) treatment time ≥ 45 minutes for solitary lesions ≥ 3 cm to ≤ 7 cm.

**Clinical Study Protocol No. 104-13-302**
**NCT No. 02112656**

**Prepared by**

[redacted]
inVentiv Health Clinical
500 Atrium Dr.
Somerset, NJ 08873 USA



Version 3.0 Date: 16 June 2020

# STATISTICAL ANALYSIS PLAN

## SIGNATURE PAGE

**Celsion Corporation**
997 Lenox Drive, Suite 100
Lawrenceville, NJ 08648 USA

**A Phase III, Randomized, Double Blind, Dummy-Controlled Study of ThermoDox® (Lyso-Thermosensitive Liposomal Doxorubicin-LTLD) in Hepatocellular Carcinoma (HCC) using standardized Radiofrequency Ablation (RFA) treatment time ≥ 45 minutes for solitary lesions ≥ 3 cm to ≤ 7 cm.**

**Clinical Study Protocol No. 104-13-302**

| [redacted] **Celsion Corporation** | **Approval Date** |
|---|---|
| • | • |
| [redacted] **Celsion Corporation** | **Approval Date** |
| • | • |
| [redacted] **Syneos Health** | **Approval Date** |
| • | • |
| [redacted] **Syneos Health** | **Approval Date** |
| • | • |

## STATISTICAL ANALYSIS PLAN

## REVISION HISTORY

| Version | Version Date | Author | Description of Modification |
|---|---|---|---|
| 0.1 | 27 March 2014 | [redacted] | First Draft based on protocol dated 14 March 2014 |
| 0.2 | 13 May 2014 | [redacted] | Updated due to FDA comments |
| 1.0 | 05 February 2016 | [redacted] | 1. Updated footnote on the cover page from Version Date: 13 May 2014 to Version Date: 02 February 2016.<br>2. Updated footnote in the body of the document from Draft 0.2 - 13May2014 to Final 1.0 02Feb2016.<br>3. Updated Signature page with Elizabeth Zobre's title and replaced Lori Davis with Daniel Strieter.<br>4. Updated to be based on Amended Protocol Version 2.1e, 04-August-2014: Section 1.0 Introduction change from Protocol Appendix IV to Protocol Appendix 20.4 |
| 1.1 | 04 April 2019 | [redacted] | Updated An additional 72-hour CBC with differential is required for eligibility criteria from Version 2.0, 14-March-2014 to Version 2.1.e, 01-Aug-2014 |
| 1.2 | 22 May 2019 | [redacted] | Updated with amendments to tables, based on client comments |
| 2.0 | 05 June 2019 | [redacted] | Included last comments and provided for approval |
| 3.0 | 06 May 2020 | [redacted] | Revise erroneous wording for PFS censoring rule (section 3.0) and include detailed classification of Protocol Deviations in Section 6.2.2 |

## TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

| Abbreviation | Definition |
|---|---|
| AASLD | American Association for the Study of Liver Disease |
| AE | Adverse Event |
| AT | As-Treated |
| CBC | Complete Blood Count |
| CI | Confidence Interval |
| CPH | Cox Proportional Hazard |
| CRF | Case Report Form |
| CR | Complete Response |
| CT | Computerized Tomography |
| CTCAE | Common Terminology Criteria for Adverse Events |
| DMC | Data Monitoring Committee |
| ECG | Electrocardiogram |
| HCC | Hepatocellular Carcinoma |
| ITT | Intent-to-treat |
| LVEF | Left Ventricular Ejection Fraction |
| MedDRA | Medical Dictionary for Regulatory Activities |
| MRI | Magnetic Resonance Imaging |
| MUGA | Multiple gated acquisition |
| NCI | National Cancer Institute |
| OS | Overall Survival |
| PFS | Progression-Free Survival |
| PT/INR | Prothrombin Time/International Normalized Ratio |
| RFA | Radiofrequency Ablation |
| SAE | Serious Adverse Event |
| SOC | System Organ Class |
| sRFA | Standardized Radiofrequency Ablation with treatment dwell time ≥45 min. |
| TEAE | Treatment-emergent adverse events |
| WHO | World Health Organization |

# 1. INTRODUCTION

Hepatocellular carcinoma (HCC) is the sixth most common neoplasm worldwide, but its very poor prognosis makes it the third leading cause of cancer-related mortality, responsible for about 600,000 deaths annually (Parkin, et al. 2005). In the US, 18,000 new cases of HCC are diagnosed each year and the incidence is steadily increasing, almost doubling since 1998. HCC is commonly diagnosed in patients with longstanding hepatic disease and cirrhosis (primarily due to hepatitis C in US and Europe and to hepatitis B in Asia). Mortality and hospitalization due to HCC, as well as hospital-related costs (inflation adjusted), increased approximately two-fold from 1988 to 2000 (Kim et al. 2005, Thomas and Abbruzzese 2005). The incidence of extrahepatic metastases in patients with HCC is generally around 15%, which includes such metastases both at the initial diagnosis of HCC or during follow-up regardless of treatment (Uka, et al. 2007; Yang, et al. 2007).

Surgical resection is the mainstay of curative treatment. However, no more than 30% of HCC patients are considered suitable for surgical treatment because of tumor size, multifocal tumors, vascular invasion, presence of extrahepatic metastases, and/or extensive liver impairment. Liver transplantation is an alternative curative treatment, but its application is limited by a severe shortage of liver graft donors (Llovet, et al. 2008). Thermal ablation modalities such as radiofrequency ablation (RFA), microwave ablation, and high-intensity focused ultrasound have emerged as important treatment options for such patients in recent years.

For HCC tumors ≥ 3 cm, overlapping ablations are required, and this overlapping can miss some micro metastases (Chen, et al. 2004; Curley, 2001; Dodd, et al. 2001; Chen, et al. 2006). If the efficacy of RFA for HCC tumors ≥ 3.0 cm could be increased, as by an adjuvant, it would be a formidable curative modality.

In this study, RFA treatments will be administered according to the RFA Treatment Procedure and device instructions (Protocol Appendix 20.4). Investigators at each site that meet the experience qualifications to perform RFA in this study will complete formal training with an accompanying RFA training manual developed by the Sponsor. All subjects will undergo an RFA treatment dwell time lasting ≥ 45 minutes.

To sum up, the worldwide HCC population is large and growing. The vast majority, including all those with tumors ≥ 3 cm, are incurable. These patients have an unmet need for more effective therapies.

## 1.1 Study Objectives

The primary objective is to compare overall survival (OS) between subjects receiving RFA plus ThermoDox versus RFA alone, using a standardized Radiofrequency Ablation (sRFA) treatment dwell time ≥ 45 minutes.

The secondary objectives are to compare progression-free survival (PFS) and safety between subjects receiving sRFA plus ThermoDox versus sRFA alone, using a standardized treatment dwell time ≥ 45 minutes.

## 1.2 Study Design

This is a 1:1 randomized, double blind, dummy controlled multicenter trial to evaluate the efficacy and safety of ThermoDox plus sRFA compared to sRFA alone using standardized treatment dwell time ≥ 45 min. for solitary HCC lesions ≥ 3.0 cm to ≤7.0 cm for 550 subjects at up to 100 study centers.

An sRFA treatment for this protocol is defined as the dwell time of ≥ 45 minutes measured from the first activation of the RFA probe to produce coagulative necrosis of target tissue through removal of the RFA probe

after the final ablation cycle or deployment. This includes the multiple ablation cycles and repositioning time between cycles for an individual subject.

Eligible HCC subjects will have a solitary lesion not amenable to curative resection consistent with clinical diagnosis of AASLD classification guidelines and will be candidates for RFA. Subjects will be randomly assigned to receive either standardized sRFA plus ThermoDox at 50 mg/m$^2$ or standardized sRFA plus a dummy infusion using a standardized RFA dwell time ≥ 45 minutes. Randomization and analysis will be stratified by maximum lesion diameter (3-5 cm versus > 5-7 cm) and RFA route (laparoscopic, open surgical, percutaneous).

Screening Period

Subjects will be evaluated up to 21 days prior to the RFA procedure date to establish eligibility for study treatment. Subjects must meet all the study inclusion/exclusion criteria prior to randomization to a treatment arm.

Treatment Period

Subjects who meet the eligibility criteria will be randomly assigned to either sRFA plus ThermoDox or sRFA plus dummy infusion using a web-based randomization system. The RFA procedure day will be Day 0 and subjects will return to the clinic Day 14 and Day 28. Subjects with a complete ablation by imaging will continue in the follow up period described below.

A subject who has an incomplete ablation is eligible for 1 retreatment procedure within 21 days after radiological imaging exam showing residual disease at Day 28. Subjects will be retreated only once with the same RFA equipment and treatment assigned at randomization. Baseline safety evaluations must meet the eligibility parameters prior to a retreatment. Subjects with a complete ablation after retreatment will be followed for both OS and PFS. If after 2 ablations the subject has local, distant intrahepatic, or extrahepatic HCC, then the subject will be considered a treatment failure and will have met the PFS endpoint. The subject is still followed for OS every 3 months after progression.

Among subjects who are not treatment failures, up to five repeat treatments are permitted to treat a recurrent lesion or to treat newly identified local or distant intrahepatic lesions at the Investigator's discretion after the PFS endpoint is reported and with agreement from the Sponsor. The subject must be eligible for retreatment consistent with the safety eligibility criteria and will be retreated with the same randomized treatment.

Subjects who develop extrahepatic lesions will have met the PFS endpoint and are no longer eligible for further protocol treatment; they will be followed for OS.

Follow up visits are performed Day 14 and Day 28 (+/- 3 days) following the treatment. Subjects with bilirubin levels > 2.0 mg/dL and ≤ 3.0 mg/dL will return to the clinic on Day 7 for additional safety assessments. Exclusion concomitant medications are to be restricted through Day 28 following study treatment.

Follow-Up Period

Following study treatment, subjects will undergo CT or MRI imaging scans (chest, abdomen, and pelvis) at months 1, 5, 9, 13, 17, 21, and 25 (+/- 2 weeks), then at 6-month intervals (+/- 2 weeks) until radiological progression is seen. The same imaging modality and measurement of assessment should be used to characterize disease at baseline and during follow up for an individual subject. Investigator determined radiological progression must be observed and recorded prior to beginning alternate treatments for HCC. Post-progression treatments will be reported and the subject will continue to be followed for OS.

To provide surveillance for any late hematologic, cardiac, or liver toxicity, the following additional safety assessments will be done:

- Physical examination, assessment of vital signs, ECGs, CBC with differential, serum chemistry, PT/INR and urinalysis at Month 5, Month 9, and Month 13 (+/- 2 weeks) or at disease progression, whichever occurs first.
- LVEF monitoring at Month 13 or at disease progression, whichever occurs first.
- Serum chemistry and PT/INR assessments at Month 17, Month 21, and Month 25 (+/- 2 weeks) or at disease progression, whichever occurs first.

The site will contact all subjects every 3 months after radiological progression to document vital status until the subject expires or withdraws consent from the study. The follow-up contacts across all clinical sites may be coordinated using the same 3-month interval.

### 1.3 Study Procedures and Timepoints

The study procedures and time points are described in Section 10 of the Protocol.

## 2. STUDY POPULATIONS

- Intent-to-Treat (ITT) population: This population includes all randomized subjects and is the primary population for subject characteristics and all efficacy parameters. All analyses using this population will be based on the treatment to which each subject was randomized.
- As-Treated (AT) population: This population includes all subjects who actually received at least one application of RFA plus ThermoDox (TR) or RFA + dummy infusion (PR). This population will be used for the safety analyses. All analyses using this population will be based on the treatment actually received.

## 3. DEFINITIONS AND DERIVED VARIABLES

Age (years): integer [(Date of Informed Consent –Date of Birth +1)/365.25]

Baseline results are those reported at the last visit on or before the Procedure Day 0 pre-dose visit.

Study Day: Date in question – Date of Randomization

Overall Survival is defined as the time (in months) from the date of randomization to the death date. In the absence of death confirmation or for subjects alive as of the OS cut-off date, survival time will be censored at the date of last study follow-up, or the cut-off date, whichever is earlier.

Progression-free survival is defined as the time (in months) from the date of randomization until the date of the Investigator-assessed radiological disease progression (PD) or death due to any cause. In the absence of PD confirmation or for subjects without PD as of the cut-off date, survival time will be censored at the date of last tumor follow-up.

Censoring Rules for the Secondary Efficacy Parameter: PFS

| Description | Outcome | Date PD/Censored |
|---|---|---|
| No baseline radiological assessments | Censored | Randomization |
| Treatment failures (subjects with incomplete ablations after two RFA procedures) | PD event | Date of radiological assessment after 2nd RFA procedure that shows residual disease |
| PD documented between scheduled visits | PD event | Date of earliest radiological assessment with evidence of PD |
| PD documented after one missed scheduled visit | PD event | Date of earliest radiological assessment with evidence of PD |
| Death before first radiological assessment | PD event | Date of death |
| Death between adequate assessment visits | PD event | Date of death |
| Death or PD after two or more missed scheduled visits | Censored | Date of latest radiological assessment prior to missing assessments |
| Alive with no PD at the analysis cut-off date | Censored | Date of last radiological assessment |
| Receive non-study cancer treatment before progression | Censored | Date of last radiological assessment before non-study cancer treatment initiated |
| | | |
| Liver transplant or liver resection before PD | Censored | Date of last tumor assessment for progression before the procedure |
| Discontinues/Withdraws from study without progression; no other follow-up information | Censored | Date of last radiological assessment for progression before date of discontinuation |

Treatment-emergent adverse events (TEAEs) are defined as those adverse events which start or worsen on or after the first dose.

## 4. EFFICACY ENDPOINTS

Overall Survival is the primary efficacy endpoint of this study. All patients will be monitored for survival by recording their visits during routine follow up for response to treatment. The visits are scheduled to occur every four months from the first imaging study confirming complete ablation until month 25 or radiological progression, whichever comes first. If patients have not demonstrated radiological progression at month 25 then the imaging visit schedule is reduced to every six months until progression. Survival is confirmed at every imaging visit. Once radiological progression is confirmed then follow up for overall survival will be confirmed every three months. It is expected that subject follow up will be about five years.

Progression-free Survival is the secondary efficacy endpoint of this study. The protocol incorporates modified RECIST (mRECIST) developed for HCC clinical research as a basis to evaluate tumor response. CT or MRI scans (chest, abdomen, and pelvis) will be done at baseline and post-treatment at months 1 (Day 28), 5, 9, 13, 17, 21, 25, then every 6 months until radiological progression is seen.

## 5. SAFETY ENDPOINTS

- Adverse events
- Laboratory assessments: hematology including CBC with differential, clinical chemistry, PT/INR, and urinalysis
- Physical Examinations
- Vital signs
- 12-Lead Electrocardiograms (ECGs)
- Echocardiograms/MUGA scans
- LVEF monitoring

Adverse events (AEs) assessment will begin at time of signing informed consent through Day 28 visit following the last study treatment. Adverse events which are assessed as possibly, probably, or definitely related to study treatment will be recorded at any point during the trial and must be followed until resolution or the subject is clinically stable.

Other safety data including physical examinations, vital signs, 12-Lead ECGs, echocardiograms/MUGA scans, LVEF monitoring and hematology, clinical chemistry, and urinalysis will be assessed through Day 28 following study treatment.

To provide surveillance for any late hematologic, cardiac, or liver toxicity, the following additional safety assessments will be done:

- Physical examination, assessment of vital signs, ECGs, CBC with differential, serum chemistry, PT/INR and urinalysis at Month 5, Month 9, and Month 13 or at disease progression, whichever occurs first.
- LVEF monitoring at Month 13 or at disease progression, whichever occurs first
- Serum chemistry and PT/INR assessments at Month 17, Month 21, and Month 25 (+/- 2 weeks) or at disease progression, whichever occurs first.

Subjects with elevated bilirubin levels > 2.0 mg/dL and ≤ 3.0 mg/dL at baseline will return to the clinic on Day 7 for additional safety assessments.

## 6. STATISTICAL METHODOLOGY

### 6.1 Statistical and Analytical Issues

#### 6.1.1 Statistical Methods

Continuous variables will be described in terms of mean, standard deviation, median, minimum, and maximum for each treatment arm. Qualitative variables will be presented in terms of frequency and percent for each treatment arm.

In summary and analysis tables of continuous variables, the minimum and maximum statistics will be presented to the same number of decimal places as the original data. The mean, median, quartiles, and 2-sided 95% confidence interval (CI) will be presented to 1 more decimal place than the original data. The standard deviation and standard error will be presented to 2 more decimal places than the original data.

In summary tables of categorical variables, the number of non-missing observations and percentages will be presented. The denominator for each percentage will be the number of subjects within the population of the treatment arm (unless otherwise specified).

Baseline characteristics and the efficacy parameters will be summarized for the ITT population while the safety parameters will be summarized for the As-Treated population.

All statistical analyses will be performed using SAS® v9.4 or higher.

### 6.1.2 Handling of Dropouts and Missing Data

#### 6.1.2.1 Dropouts

Subjects who complete the Day 0 procedures but withdraw prior to the Follow-up Period will not be replaced

#### 6.1.2.2 Missing Data

All analysis will be based on available data to summarize safety and efficacy results. However, a conservative approach to partial dates and missing adverse event's relationships will be as follows:

- For the OS analyses, if a subject's death month and year are provided but the day is missing, the day will set to the first day of the month, unless other qualifying study data support survival until a later date in the same month. The same concept will be applied if only the year is known. For a completely missing date of death, the fatality will be imputed to have been occurred on the last available contact of the subject where we still know he/she was alive.
- Adverse events with partial dates with the same month and year as the Procedure Day 0 will be considered as treatment-emergent adverse events.
- If an adverse event's relationship is missing it will be assumed to be related.

### 6.1.3 Pooling of Investigator Sites

Data will be pooled from centers for these analyses. The justification for pooling is made on a clinical basis (Meinert, 1986). The basis for pooling comes from three critical factors: 1) The study sites must implement one common protocol. 2) The sponsor must provide very close monitoring of study site compliance. 3) The study sites must use common data collection procedures.

However, region (EU and US, China, Korea, other) will be included as a factor in supportive and exploratory CPH analyses for OS and PFS, as well as safety analyses.

### 6.1.4 Determination of Sample Size

The study is designed to detect with 80% power a hazard ratio for OS of 0.67 (33% risk reduction) in the ThermoDox (TR) arm compared with the control (PR) arm with an overall 1-sided type 1 error of 0.025. An OS hazard ratio of 0.63 was observed among Celsion's initial phase III trial, the HEAT study (Protocol 104-06-301) subjects with a solitary 3-7 cm lesion treated with ≥ 45 minutes of RFA. A 3%/year loss to survival follow-up rate has been assumed and using a 1:1 treatment allocation (TR:PR) of 550 subjects, a target of 197 events (deaths) will be required for the primary analysis.

Based on these design operating characteristics and assuming a median survival time of approximately 4.5 years in the control arm, the primary analysis target events milestone will be reached approximately 30 months after the last subject is randomized in the study. The median OS in the control arm was estimated based on the interim OS results in the HEAT study.

## 6.2 Subject Disposition, Baseline and Treatment Characteristics

### 6.2.1 Subject Disposition

The number of subjects in each study population and the reasons for exclusion, along with any randomization and/or stratification errors will be summarized by treatment arm. In addition, subjects that discontinue study treatment or study follow-up will also be summarized, along with reasons for study discontinuation.

The number of screening failures, and the reason, will be tabulated.

### 6.2.2 Protocol Deviations

CSR Reportable

A CSR reportable Protocol Deviation is related to inclusion/exclusion criteria, conduct of the trial, patient management or patient assessments that impact the safety of the subjects or jeopardize the quality of the study data.

The following classifications will be used to report CSR reportable Protocol Deviations:

RD1 - A subject that did not meet entry criteria

RD2 - A subject that developed withdrawal criteria but was not withdrawn

RD3 - A subject that received the wrong treatment or incorrect dose

RD4 - A subject that received an excluded medication

RD5 - Critical ICF, GCP and other Protocol Deviations

Major CSR Reportable Protocol Deviations

In addition, among all CSR Reportable Protocol Deviations, a set of “major protocol deviations” is defined as a means to measure adherence to key aspects of the protocol using prespecified sensitivity analyses.

A summary of all major protocol deviations (at study entrance and during the study period) by type of deviation will be provided.

A patient will be classified as having a major protocol CSR reportable deviation if she/he meets at least one of the criteria presented below.

MRD1 - No HCC lesions

MRD2 - single HCC lesion < 3.0 or > 7.0 cm or multiple lesions

MRD3 - ECOG performance status at baseline not 0

MRD4 - Not a Child-Pugh A classification

MRD5 - baseline platelets < 75,000/mm3

MRD6 - RFA start time > 20 minutes from start of infusion

MRD7 - RFA dwell time < 45 minutes

CSR Non-Reportable

A CSR non-reportable Protocol Deviation may be important to address and document as part of site management and oversight, but will not be considered reportable in the CSR.

The following classifications will be used to capture non-reportable PDs.

NRD1 - SAE reporting

NRD2 - Informed consent (other than those captured above)

NRD3 - Study procedures

NRD4 - Investigational product (other than incorrect dose or wrong treatment)

### 6.2.3 Subject Baseline Characteristics

Demographic data (age, race, gender) will be summarized by treatment arm for the ITT and AT population.

Subject characteristics such as height (in cm), weight (in kg), ECOG performance and disease history reported at the Screening/Baseline visit will be summarized by treatment arm for the ITT population.

Data listings will present the demographic and baseline characteristic data.

### 6.2.4 Treatment Exposure and Compliance

The study treatment administration and compliance profile will be summarized descriptively by treatment arm for the AT population and displayed in data listings.

### 6.2.5 Prior and Concomitant Medications and Therapies

Prior and concomitant medications will be coded according to the World Health Organization (WHO) Medication Dictionary for Concomitant Medication.

Prior and concomitant medications will be summarized descriptively by treatment arm for the ITT population and displayed in data listings.

### 6.2.6 Medical Histories

Medical history including cancer and non-cancer history data listing will be presented by treatment arm for the ITT population.

### 6.2.7 Follow-up Period

Duration of Follow-up for Survival:

A summary table showing the duration of follow-up will be presented by treatment arm for the ITT population. The follow-up period is defined as the time starting on the date randomized until the last day of contact, or the day of withdrawal from study or the day of death, whichever occurs last.

Non-study anti-cancer treatment during follow-up period:

Patients receiving non-study anti-cancer treatments initiated during the follow-up period will be summarized in a table by treatment arm. The table will display the number, type and time to starting a new non-study anti-cancer treatment. The summary table will also show the number of patients starting a new anti-cancer treatment without having experienced documented radiological progression by the investigator. The time to start of the new anti-cancer treatment will be defined as the start date of the new anti-cancer therapy minus the date of last dose of study medication + 1. If the start date of the new anti-cancer treatment is missing, the time to start of anti-cancer therapy will be missing for that patient.

## 6.3 Efficacy Analysis

### 6.3.1 Primary Efficacy Variable(s)

Overall survival will be measured by time (in months) from randomization to death from any cause or the end of the study. Subjects who have not died will be censored at the date of last study follow-up, or the cut-off date, whichever is earlier as mentioned in Section 3.

The OS cut-off date used for the primary analysis will be based on the observations of the 197th death in the study. All patients dying on the calendar date of the 197$^{th}$ death will be included in the analysis should more than one patient die on the calendar date of the 197$^{th}$ death. Patients having a documented survival status (alive or dead) after this date are censored at the cut-off date.

With the OS cut-off date being event driven, for operational efficiency, the cut-off date for all other study endpoints (e.g., PFS) will be fixed at close proximity of the OS cut-off date, when the milestone is nearing completion. This especially applies to the interim analyses.

OS in the ITT population will be compared between the 2 treatment arms using the stratified log-rank test (Score statistic from PHREG and ties=Breslow (Kalbfleisch and Prentice 1980)). The estimate of the hazard ratio and corresponding 95% CI will be provided using a Cox proportional hazards (CPH) model including treatment and the stratification factors (maximum lesion diameter [3-5 cm versus >5–7 cm] and RFA route [laparoscopic, open surgical, percutaneous]) in the model. The survival curves will be estimated using Kaplan-Meier estimates. The stratification factors will be populated as per the randomization assignment.

Survival for each arm will be summarized using Kaplan Meier curves and is further characterized in terms of the median and survival probability at 12, 24, 36, 48, 60 and 72 months, along with the corresponding 2-sided 95% confidence intervals for the estimates. Confidence intervals for median survival are based upon the methods of (Brookmeyer and Crowley 1982). In addition to Kaplan-Meier estimates, corresponding 2-Sided 95%

confidence intervals constructed using the Clopper-Pearson approximation to the exact binomial proportion (Hollander and Wolfe 1973).

For the first interim analysis, if there are more than 129 events (118+10%), an additional table will be presented for sensitivity purposes, based only on the first 118 events. The same principle applies for the second interim analysis, accordingly.

### 6.3.2 OS Supportive Analyses

Supportive analyses for OS, conducted in the ITT population (unless otherwise noted), will include:

a. The unstratified log-rank test and a CPH model (only treatment effect in the model).

b. Multivariate analysis using the CPH model, including the stratification factors and the following set of potential prognostic/predictive factors: age (< 65 v. ≥ 65), race (Caucasian, Asian, Other), region (EU and America, China All, Korea, Other), RFA start time, RFA dwell time (< 90 min, ≥ 90 min), device, disease etiology (Hep B, Other), ECOG Performance Status (0, 1, 2, 3, 4, 5), Alfa-feto protein >200 ng/mL (yes/no).

c. Factors included in the model will be assessed for co-linearity and a stepwise selection process will be applied to identify a final subset of prognostic/predictive factors in the model. Once the subset has been established, treatment will be added to the final model to assess its effect. An exploratory analysis of treatment by factor interactions using the CPH model will be conducted, using the factors identified in the final model above.

d. Subgroup analyses will also be conducted for the stratification factors and the potential prognostic/predictive factors identified in Section 6.33.b above. The HR and associated 95% CI will be presented for each subgroup.

e. The primary efficacy analysis, as outlined in Section 6.3.1, will also be run excluding any subjects that met the Major CSR Reportable Protocol Deviation criteria MRD1 to MRD7 as outlined in Section 6.2.2.

f. A stratified log-rank test using the final strata as recorded on the CRF, in the event there are differences from the primary IVRS assignments. (AT population)

g. Additional exploratory analyses may be performed.

In addition, a study follow-up Kaplan-Meier analysis (as for OS) will be presented for all subjects. Subjects will be censored if they have died, on their date of death. Only median, 25$^{th}$ and 75$^{th}$ percentiles will be presented.

### 6.3.3 Secondary Efficacy Variable(s)

PFS will be measured from the date of randomization to the first date on which one of the following occurs, as determined by CT or MRI scan:

- Death of any cause
- Treatment failure (inability to achieve Complete Response (CR) after two RFA ± ThermoDox treatment sessions)
- Progression due to local tumor recurrence after initial CR

- Progression due to distant intrahepatic tumor recurrence
- Progression due to extrahepatic tumor recurrence

All secondary endpoints comparisons will be made at the 2-sided 0.05 significance level for the ITT population. Since PFS is the only secondary endpoint for regulatory registration purposes, no further multiplicity adjustments will be made. Assuming that OS demonstrates significance at the 1-sided 0.025 level, PFS can subsequently be tested at the 1-sided 0.025 level.

PFS will be analyzed with methodology applied to the OS endpoint. Specifically, PFS will be analyzed with the methodology specified in Section 6.3.1 and the supportive analyses in Section 6.3.2.

In line with FDA guidance (Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics, 2007), several sensitivity analyses for Progression-Free Survival are also conducted on the ITT Population as follows:

- Analysis that includes clinical progression as a PFS event in addition to the presence of radiological evidence of progression.

- Analysis including clinical progression as a PFS event that also counts initiation of non-study antitumor therapy as an event date rather than as date used to censor subsequent response assessments. (ITT population)

- An analysis that includes all death and response assessments (without censoring missed visits) and also counting as an event, clinical progression, the initiation of anti-tumor therapy or death through the date of cut-off for survival (ITT population)

- Analysis using the next scheduled visit as the event date for radiological progression when image assessment falls more than 1 month after a scheduled assessment. If reassignment of the progression date to the next scheduled visit results in a progression date after death, the death date is used as the event date. (ITT population)

- Kaplan-Meier curves of time to first, second, third, fourth and fifth radiological tumor assessments from date of randomization are done and compared between groups using the log-rank test. (ITT population)

## 6.4 Safety Analysis

The safety evaluations will focus primarily on AEs and laboratory assessments, but will also include physical examinations, vital signs, ECGs, echocardiograms/MUGA scans and LVEF testing. All subjects included in the As-Treated population will be summarized by treatment arms in the safety analysis.

### 6.4.1 Treatment-Emergent Adverse Events

Treatment-emergent adverse events will be coded according to the Medical Dictionary for Regulatory Activities (MedDRA) terminology and the severity of the toxicities will be graded according to the NCI CTCAE criteria, v4.03, where applicable.

All TEAEs will be summarized (incidence) and listed by the System Organ Class (SOC), preferred term, toxicity/severity grade, and causal relationship to study medication by treatment arm. In addition, separate

summaries of serious treatment-emergent adverse events (SAEs) and Grade 3 and 4 TEAEs will be presented by treatment arm.

Further, AE tabulations will be repeated using the following subgroups:

- Age Category at Screening (<65 years/>=65 Years)
- Gender (male/female),
- Ethnicity (Asian/Caucasian/Other),
- Region (America and Europe/China All/Asia Other), and
- Lesion Size (<5 cm/>=5cm)
- RFA Dwell Duration (<90, >=90 min)

### 6.4.2 Laboratory Parameters

Hematological and chemistry laboratory parameters will be graded according to the NCI CTCAE v.4.03 criteria (see appendix), where applicable. Absolute values and changes from baseline will be summarized by treatment arm at each visit. Shift tables will be provided for the shift from Baseline to the highest post Baseline value. The last on-study value will also be presented (called 'End of Study'), with change from baseline. In addition, worst severity grade, time to event, and time to resolution will also be summarized.

### 6.4.3 Physical Examination

The physical exam data will be listed by treatment arm at the screening visit, at post RFA on Day 0, Day 7 for subjects with elevated bilirubin, Day 14, Day 28, Month 5, Month 9, Month 13 or disease progression, whichever occurs first.

### 6.4.4 Vital Signs

Vital signs, (temperature (in Celsius), blood pressure, pulse rate, and respiration) will be collected at the screening visit, Day 0: pre-dose assessment the morning of the RFA procedure, post-RFA within 15 minutes after completion of study drug infusion, within 30 minutes and 1 hour after RFA procedure, then every 2 hours until the subject is stable or discharged, Day 7 for subjects with elevated bilirubin, Day 14, Day 28, Month 5, Month 9, Month 13 or disease progression, whichever occurs first.

Vital signs will be summarized at each visit by treatment arm for actual values and change from baseline. The last on-study value will also be presented (called 'End of Study'), with change from baseline.

### 6.4.5 12-Lead ECGs

12-Lead ECGs will be acquired at screening, post Day 0: post RFA, Day 28, Month 5, Month 9, and Month 13 or disease progression, whichever occurs first.

For all12-lead ECG abnormal findings will be presented for the actual values and change from baseline by visit and treatment arm.

Frequency tables will be provided by visit for QTcB and QTcF for the following categories:

- QTc <= 450 ms
- 450 ms < QTc <= 480 ms

- 480 ms < QTc <= 500 ms
- QTc > 500 ms

Frequency tables will be provided by visit for the changes from baseline of QTcB and QTcF for the following categories:

- QTc change from baseline <= 30 ms
- 30 ms < QTc change from baseline <= 60 ms
- QTc change from baseline > 60 ms

### 6.4.6 Echocardiograms/MUGA scans for LVEF monitoring

A baseline echocardiogram (ECHO) will be carried out at the screening visit. Measurements with a MUGA scan are allowed if an echocardiogram cannot be performed; however, the same modality used at baseline must be used on Day 28 and at Month 13 or at time progression of disease, whichever occurs first.

Counts and percentages of abnormal echocardiograms/MUGA scan will summarized for each treatment arm at each visit.

LVEF monitoring will be assessed at screening, Day 0: post RFA, Day 28, Month 5, Month 9, and Month 13 or disease progression, whichever occurs first.

LVEF % will be summarized by treatment arm at each visit for actual values and change from baseline.

## 6.5 Data Monitoring Committee

An independent Data Monitoring Committee (DMC) will periodically assess the safety data as well as the efficacy as part of the interim analyses described in Section 6.6. A description of the roles and responsibilities and details of the review processes are provided in a separate DMC charter.

## 6.6 Interim Analysis

Two interim analyses reviewed by the DMC, both for efficacy and futility, are planned for the study. The first is planned after 60% of the target events is reached (118 deaths) and the second after 80% of the events has been reached (158 deaths). The Lan-DeMets alpha-spending approach will be used with O'Brien-Fleming stopping boundaries to evaluate efficacy. Fixed HR boundaries will be used to assess futility. This approach will account for multiple testing and preserve the overall 1-sided study significance level of 0.025. Additional details are provided in the DMC charter.

No futility analyses are planned during the accrual period. Considering the relative short time of the accrual period compared to the slower accumulation rate of the events, there will not be sufficient events to make such assessment in the accrual study stage.

# 7. TABLES, LISTINGS, AND FIGURES LAYOUT

The default tables, listings and figures (TLF) layout will be as follows.

| | |
|---|---|
| **Orientation** | Landscape |
| **Paper Size** | Letter |
| **Margins** | Top: 3.2 cm<br>Bottom: 2.5 cm<br>Left: 2.5 cm<br>Right: 2.5 cm |
| **Font** | Courier New 9pt |
| **Headers** | Sponsor name and Protocol number (Left); Page X of Y (Right)<br>TLF Number and Title |
| **Footers** | SAS program path and file name<br>Date, Time TLF generated |

The font size may be reduced as necessary to allow additional columns to be presented, but not at the expense of clarity. Also, the orientation may be changed to portrait if appropriate.

The date format for all presentations will be 'DDMMMYYYY'.

All TLF outputs will be generated using SAS® v9.4 or higher for Windows.

CRF data collected will be presented within data listings. The data listings will be sorted by treatment arm, site number, subject number, visit, and time point.

## 8. LIST OF TABLES, LISTINGS AND FIGURES

The list of tables, listings and figures will be provided in a separate document.

## 9. REFERENCES

1. Chen MH, Wei Y, Yan K, Gao W, Dai Y, Huo L, et al. Treatment of strategy to optimize radiofrequency ablation for liver malignancies. J Vasc Interv Radiol. 2006;17(4):671-83.
2. Chen MS, Li JQ, Zheng Y, Guo RP, Liang HH, Zhang YQ, et al. A prospective randomized trial comparing percutaneous local ablative therapy and partial hepatectomy for small hepatocellular carcinoma. Ann Surg. 2006;243(3):321-8.
3. Chen MH, Yang W, Yan K, Zou MW, Solbiati L, Liu JB, et al. Large liver tumors: protocol for radiofrequency ablation and its clinical application in 110 patients - mathematic model, overlapping mode, and electrode placement process. Radiology. 2004;232(1):260-71.
4. Curley SA. Radiofrequency ablation of malignant liver tumors. Oncologist 2001; 6:14-23.

5. Dodd GD 3rd, Frank MS, Aribandi M, Chopra S, Chintapalli KN. Radiofrequency thermal ablation: computer analysis of the size of the thermal injury created by overlapping ablations. AJR Am J Roentgenol. 2001;177(4):777-82.

6. Kim WR, Gores GJ, Benson JT, Therneau TM, Melton LJ 3rd. Mortality and hospital utilization for hepatocellular carcinoma in the United States. Gastroenterology. 2005;129(2):486-93.

7. Llovet JM, Di Bisceglie AM, Bruix J, Kramer BS, Lencioni R, Zhu AX, et al. Design and endpoints of clinical trials in hepatocellular carcinoma. J Natl Cancer Inst. 2008;100(10):698-711.

8. Meinert, C. (1986). *Clinical Trials: Design, Conduct, and Analysis*. Oxford University Press, New York.

9. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. CA Cancer J Clin. 2005;55:74-108.

10. Thomas MB, Abbruzzese JL. Opportunities for targeted therapies in hepatocellular carcinoma. J Clin Oncol. 2005;23:8093-108.

11. Uka K, Aikata H, Takaki S, Shirakawa H, Jeong SC, Yamashina K, et al. Clinical features and prognosis of patients with extrahepatic metastases from hepatocellular carcinoma. World J Gastroenterol. 2007;13(3): 414-20.

12. Yang Y, Nagano H, Ota H, Morimoto O, Nakamura M, Wada H, et al. Patterns and clinicopathologic features of extrahepatic recurrence of hepatocellular carcinoma after curative resection. Surgery. 2007;141(2):196-202

13. Kalbfleisch, J. D. and Prentice, R. L. (1980), The Statistical Analysis of Failure Time Data, New York: John Wiley & Sons, Inc.

14. Brookmeyer and Crowley, A Confidence Interval for the Median Survival Time, Biometrics 38, 29-41, March 1982.

15. Hollander and Wolfe, D. A.: Nonparametric Statistical Methods, Wiley 1973.

16. *Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics*. FDA, May 2007. Web. 27 Mar.2014.
http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm071590.pdf

# 10. APPENDICES

## 10.1 Standard Laboratory Ranges

### Table 10.1.1. Standard Normal Ranges for Hematological Tests

| Hematology Laboratory Test | Unit | Normal Range |
|---|---|---|
| Hemoglobin | g/L | 120 - 170 |
| Hematocrit | % | 37 - 50 |
| Platelets | x $10^3$/uL | 110 - 450 |
| RBC | x $10^6$/uL | 4 - 6.2 |
| WBC | x $10^3$/uL | 3.5 - 11 |
| Neutrophil, Absolute | x $10^3$/uL | 2.0 – 7.0 |
| Lymphocytes, Absolute | x $10^3$/uL | 0.9 - 3.5 |
| Monocytes, Absolute | x $10^3$/uL | 0.12 – 0.9 |
| Eosinophils, Absolute | x $10^3$/uL | 0 – 0.66 |
| Basophils, Absolute | x $10^3$/uL | 0 – 0.2 |

### Table 10.1.2. Standard Normal Ranges for Serum Chemistry Tests

| Serum Chemistry Laboratory Test | Units | Normal Range |
|---|---|---|
| Sodium | mEq/L | 132 - 145 |
| Potassium | mEq/L | 3.4 - 5.4 |
| Chloride | mEq/L | 94 - 112 |
| Albumin | g/L | 33 - 49 |
| Calcium | g/L | 0.083 – 0.106 |
| BUN | g/L | 0.04 – 0.24 |
| Bilirubin | g/L | 0.002 – 0.012 |
| Bilirubin, indirect | mg/dL | 0.1 – 1.0 |
| Alkaline Phosphatase | units/L | 35 - 127 |
| AST (SGOT) | units/L | 5.0 - 50 |
| ALT (SGPT) | units/L | 5.0 - 40 |
| Creatinine | g/L | 0.005 – 0.012 |
| Glucose | g/L | 0.7 – 1.00 |

## 10.2 NCI CTCAE Laboratory Grades

### Table 10.2.1 NCI CTCAE V4.03 LAB GRADES

| Lab Test | Unit | Grade 1 | Grade 2 | Grade 3 | Grade 4 |
|---|---|---|---|---|---|
| **SERUM CHEMISTRIES + COAGULATION** | | | | | |
| Albumin (hypoalbuminemia) | g/dL | [3, LLN) | [2, 3) | [0, 2) | UNDEFINED |
| Alkaline Phosphatase | units/L | (ULN, 2.5*ULN] | (2.5*ULN, 5*ULN] | (5*ULN, 20*ULN] | >20*ULN |
| ALT | units/L | (ULN, 3.0*ULN] | (3.0*ULN, 5*ULN] | (5*ULN, 20*ULN] | >20*ULN |
| AST | units/L | (ULN, 3.0*ULN] | (3.0*ULN, 5*ULN] | (5*ULN, 20*ULN] | >20*ULN |
| Amylase | units/L | (ULN, 1.5*ULN] | (1.5*ULN, 2*ULN] | (2*ULN, 5*ULN] | >5*ULN |
| Bilirubin | mg/dL | (ULN, 1.5*ULN] | (1.5*ULN, 3*ULN] | (3*ULN, 10*ULN] | >10*ULN |
| Calcium max (Hypercalcemia) | mg/dL | (ULN, 11.5] | (11.5, 12.5] | (12.5, 13.5] | >13.5 |
| Calcium min (Hypocalcemia) | mg/dL | [8, LLN) | [7, 8) | [6, 7) | [0, 6) |
| Cholesterol (Hypercholesterolemia) | mg/dL | (ULN, 300] | (300, 400] | (400, 500] | >500 |
| CK | unit/L | (ULN, 2.5*ULN] | (2.5*ULN, 5*ULN] | (5*ULN, 10*ULN] | >10*ULN |
| Creatinine | mg/dL | (ULN, 1.5*ULN] | (1.5*ULN, 3*ULN] | (3*ULN, 6*ULN] | >6*ULN |
| GGT | units/L | (ULN, 2.5*ULN] | (2.5*ULN, 5*ULN] | (5*ULN, 20*ULN] | >20*ULN |
| Glucose max (hyperglycemia) | mg/dL | (ULN, 160] | (160, 250] | (250, 500] | >500 |
| Glucose min (hypoglycemia) | mg/dL | [55, LLN) | [40, 55) | [30, 40) | [0, 30) |
| Lipase | units/dL | (ULN, 1.5*ULN] | (1.5*ULN, 2*ULN] | (2*ULN, 5*ULN] | >5*ULN |
| Fibrinogen | mg/dL | [0.75*LLN, LLN) or decreased from baseline by >0 to 25% | [0.5*LLN, 0.75*LLN) or decreased from baseline by 25 to <50% | [0.25*LLN, 0.5*LLN) or decreased from baseline by 50 to <75% | [0, 0.25*LLN) or decreased from baseline by ≥75% or <50 mg/dL |
| Magnesium max (Hypermagnesemia) | mg/dL | (ULN, 3] | UNDEFINED | (3, 8] | >8 |
| Magnesium min (Hypomagnesemia) | mg/dL | [1.2, LLN) | [0.9, 1.2) | [0.7, 0.9) | [0, 0.7) |
| Phosphates (hypophosphatemia) | mg/dL | [2.5, LLN) | [2, 2.5) | [1, 2) | [0, 1) |
| Potassium max (hyperkalemia) | mEq/L | (ULN, 5.5] | (5.5, 6] | (6, 7] | >7 |
| Potassium min (hypokalemia) | mEq/L | [3, LLN) | UNDEFINED | [2.5, 3) | [0, 2.5) |
| PT | Seconds | (ULN, 1.5*ULN] | (1.5*ULN, 2.5*ULN] | >2.5*ULN | UNDEFINED |
| INR | | (ULN, 1.5*ULN], >1 - 1.5 times above baseline if on anticoagulation | (1.5*ULN, 2.5*ULN], >1.5 - 2.5 times above baseline if on anticoagulation | >2.5*ULN, >2.5 times above baseline if on anticoagulation | UNDEFINED |
| PTT | Seconds | (ULN, 1.5*ULN] | (1.5*ULN, 2.5*ULN] | >2.5*ULN | UNDEFINED |
| Sodium max (hypernatremia) | mEq/L | (ULN, 150] | (150, 155] | (155, 160] | >160 |
| Sodium min (hyponatremia) | mEq/L | [130, LLN) | UNDEFINED | [120, 130) | [0, 120) |
| Triglycerides (Hypertriglyceridemia) | mg/dL | (150, 300] | (300, 500] | (500, 1000) | >=1000 |
| Uric Acid (hyperuricemia) | mg/dL | (ULN, 10] | UNDEFINED | UNDEFINED | >10 |

| Lab Test | Unit | Grade 1 | Grade 2 | Grade 3 | Grade 4 |
|---|---|---|---|---|---|
| **HEMATOLOGIES** | | | | | |
| Hemoglobin (High) | g/dL | Increase in >0 - 2 gm/dL above ULN or above baseline if baseline is above ULN | Increase in >2 - 4 gm/dL above ULN or above baseline if baseline is above ULN | Increase in >4 gm/dL above ULN or above baseline if baseline is above ULN | UNDEFINED |
| Hemoglobin (Low) | g/dL | [10, LLN) | [8, 10) | [0, 8) | UNDEFINED |
| Platelet | 10^3/uL | [75, LLN) | [50, 75) | [25, 50) | [0, 25) |
| WBC (High) | 10^3/uL | UNDEFINED | UNDEFINED | >100 | Undefined |
| WBC (Low) | 10^3/uL | [3, LLN) | [2, 3) | [1, 2) | [0, 1) |
| Lymphocytes (High) | 10^3/uL | UNDEFINED | (4 – 20] | >20 | UNDEFINED |
| Lymphocytes (Low) | 10^3/uL | [.8, LLN) | [.5 - .8) | [.2 - .5) | [0,.2) |
| Neutrophils | 10^3/uL | [1.5, LLN) | [1, 1.5) | [.5, 1) | [0, .5) |

LLN=Lower Limit of Normal; ULN=Upper Limit of Normal;

[a] Grade 4 criteria do exist for AE reporting based on clinical manifestations of leukocytosis needing urgent intervention. Not included in the programmed grading of laboratory data but is applicable to AE reporting.