Statistical Analysis Plan

Version AA

A Single-arm Trial of Transcatheter Arterial Chemoembolization with Tandem Microspheres in the Treatment of Localized Hepatocellular Carcinoma

A Study of Tandem Microspheres in Localized Hepatocellular Carcinoma

STOPPER

Study Reference Number **S2382**

CONFIDENTIAL

DO NOT COPY OR DISTRIBUTE WITHOUT WRITTEN PERMISSION

TABLE OF CONTENTS

1 PROTOCOL SYNOPSIS

<table>
<tr><td colspan="2">A Single-arm Trial of Transcatheter Arterial Chemoembolization with Tandem Microspheres in the Treatment of Localized Hepatocellular Carcinoma</td></tr>
<tr><td>Study Objective</td><td>The primary objective of this study is to evaluate the safety and efficacy of transcatheter arterial chemoembolization with Tandem Microspheres loaded with Epirubicin in the treatment of subjects with localized hepatocellular carcinoma (HCC)</td></tr>
<tr><td>Study Device</td><td>Tandem Microspheres are spherical, tightly calibrated, biocompatible, non-resorbable, hydrogel microspheres coated with Polyzene®-F, an inorganic polymer.</td></tr>
<tr><td>Planned Indication(s) for Use</td><td>Tandem Microspheres are indicated for the embolization of arteriovenous malformations and hypervascular tumors</td></tr>
<tr><td>Device Specifications</td><td>Tandem Microspheres are stored in 2 ml and 3 ml sized syringes. They are available in a range of sizes (40, 75, and 100 μm). Microspheres used will be at the discretion of the treating physician(s) based on the tumor size and/or vascular structure.
<table>
<tr><th>Microspheres Diameter（μl）</th><th>Size（ml）</th></tr>
<tr><td>40</td><td>2</td></tr>
<tr><td>40</td><td>3</td></tr>
<tr><td>75</td><td>2</td></tr>
<tr><td>75</td><td>3</td></tr>
<tr><td>100</td><td>2</td></tr>
<tr><td>100</td><td>3</td></tr>
</table></td></tr>
<tr><td>Study Design</td><td>This is a prospective, single-arm, multicenter study to evaluate the safety and efficacy of transcatheter arterial chemoembolization with Tandem Microspheres loaded with Epirubicin in the treatment of subjects with localized hepatocellular carcinoma. All qualified subjects will be enrolled in the study and receive transcatheter arterial chemoembolization treatment with Tandem Microspheres loaded with Epirubicin</td></tr>
<tr><td>Sample Size</td><td>109 subjects (including 10% loss to follow-up)</td></tr>
<tr><td>Number of Study Sites</td><td>10 clinical centers will be participating in the study</td></tr>
<tr><td>Primary Effectiveness Endpoint</td><td>The primary effectiveness endpoint for this clinical trial is 6-month objective tumor response rate (ORR).</td></tr>
</table>

Secondary Endpoints	• ORR at 30-day and 3-month • Time to Progression (TTP) • Time to Extrahepatic Spread • Proportion Progression-Free (PPF) at one year • Overall survival at one year • The frequency of Tandem treatment associated adverse events
Study Enrollment	Subject will be qualified to be enrolled in the clinical trial once signed the IRB/EC approved informed consent, meet all inclusion criteria and does not meet any of the exclusion criteria. All enrolled subjects will receive the transcatheter arterial chemoembolization with Tandem Microspheres. Enrollment timepoint will be based on subject receiving transcatheter arterial chemoembolization treatment with Tandem Microspheres loaded with Epirubicin
Follow-up Plan	Follow-up will be conducted at 30 days and 3, 6, and 12 months post initial treatment. The tests during the follow-ups include: • Laboratory (include Complete Blood Count, blood biochemistry, blood coagulation function, renal and liver function, electrolyte, AFP) and imaging tests (include plain scan plus enhanced MRI, Chest plain scan CT) at 30 days and 3, 6 and 12 months post procedure • ECG at 3 months post procedure
Study Duration	The study will end when all subjects completed the 12-month follow-up post initial treatment, or withdraw (death loss to follow-up, or withdraw consent).

Inclusion Criteria	Subjects may be included in the study only if they meet all of the following criteria: 1. Subject is able to provide informed consent and must sign the Institutional Review Board/Ethics Committee (IRB/EC) approved Informed Consent Form. 2. Male or female of age ≥18 and ≤75 years. 3. Confirmed diagnosis of HCC according to the diagnostic criteria included in the management guideline issued by China's Ministry of Health in 2017. 4. HCC is diagnosed for the first time or recurrence of tumor after surgical or ablation treatment. 5. Single tumor less than 7cm in diameter or multiple tumors with maximum 3 lesions with >1cm in diameter, individual diameter <7cm and less than 10cm in total diameter. 6. no previous chemotherapy, radiotherapy or transarterial embolization (with or without chemotherapy) for HCC 7. Preserved liver function (Child-Pugh A or B7). 8. ECOG Performance Status 0 or 1.

Exclusion Criteria	Subjects will be excluded for the following reasons: 1. Presence of vascular invasion or extra-hepatic spread of disease, or diffuse HCC, defined as >50% liver involvement , or arteriovenuous fistula 2. Macrovascular invasion of main or primary branches of portal vein at entry into the study 3. Any contraindication for TACE treatment 4. Any contraindication for Epirubicin administration 5. Advanced liver disease (bilirubin levels >2 mg/dl, AST or ALT >5 times upper limit of normal) 6. Renal failure or insufficient renal function 7. Subject unable to receive MRI examination 8. Pregnant or breast feeding woman, or plan to become pregnant during treatment or within 12 months of treatment 9. Couldn't commit reliable birth control measures during treatment or within 12 months of treatment. 10. Subject is participating other investigational drug or device clinical trial within 30 days of signing the informed consent. 11. Subject is not suitable to participate in the study as judged by investigator
Statistical Methods	
Primary Study Hypotheses	The primary hypothesis is that the 6-month ORR in the Tandem microspheres embolization and drug delivery arm is higher than the performance goal (PG). The null (H_0) and alternative (H_1) hypotheses for this primary endpoint are presented below. H_0: $P_e \leqslant PG$ H_1: $P_e > PG$ where P_e is the expected 6-month ORR (%).
Statistical Test	A normal approximation test will be used to assess the one-sided hypotheses of the test device superior to the PG

Success Criteria for the PG	The test device will be concluded to be superior to the PG in the primary analysis if the one-sided lower 97.5% confidence bound for the test device performance in 6-month ORR is greater than PG.
Sample Size Parameters	The overall sample size (N=109) is driven by the primary effectiveness endpoint. • Power ⩾ 85% • One-sided significance level = 2.5% (alpha) • PG for 6-month ORR = 45% • Expected 6-month ORR = 60% • Total N=98 subjects are required for 6-month assessment • Attrition rate = 10% (loss to follow-up in 6 months) • Approximately N=109 subjects to enroll

2 INTRODUCTION

This Statistical Analysis Plan (SAP) has been designed to document the planned analyses to be consistent with the objectives of the STOPPER China protocol. The specified analyses may be provided in reports to competent authorities and/or for scientific presentations and/or manuscripts. The primary analyses will be based on the 6-month assessment for the tumor objective response rate (ORR) after the treatment initiation (i.e. the 1st treatment). The study success will be based on the primary effectiveness testing hypotheses (i.e. the only statistical hypotheses).

3 PRIMARY ENDPOINTS ANALYSES

The sample size is justified by hypotheses parameters and driven by the 6-month primary effectiveness endpoint to preserve adequate statistical testing power.

The primary effectiveness hypothesis is planned for being tested after the initial treatment and through 6 months at the pre-specified significance level of one-sided 2.5%.

3.1 Primary Effectiveness Endpoint

The 6-month ORR is selected to be assessed for the primary effectiveness endpoint. The goal is designed to demonstrate that Tandem meets the PG in terms of the ORR through 6 months post treatment initiation.

3.1.1 Definition of the ORR

Objective tumor response will be assessed by two response evaluation systems for solid tumors:

- European Association for the Study of the Liver (EASL); and
- Modified Response Evaluation Criteria in Solid Tumors (mRECIST).

The EASL and mRECIST responses measure on MRI in HCC subjects treated with chemoembolization to predict long-term survival. The mRECIST response demonstrated higher survival correlation than the EASL (cf. Prajapati et al, 2013 Annuals of Oncology).

EASL Response Assessment of Target Lesion for HCC:

Response Category	EASL Criteria
Complete Response (CR)	Disappearance of all known disease and no new lesions determined by two observations not less than 4 weeks apart
Partial Response (PR)	At least 50% reduction in total tumor load of all measurable lesions determined by two observations not less than 4 weeks apart
Stable Disease (SD)	Any cases that do not qualify for either PR or PD
Disease Progression (PD)	At least 25% increase in size of one or more measurable lesions or the appearance of new lesions

mRECIST Response Assessment of Target Lesion for HCC:

Response Category	mRECIST Criteria
Complete Response (CR)	Disappearance of any intratumoral arterial enhancement in all target lesions
Partial Response (PR)	At least a 30% decrease in the sum of the diameters of viable (enhancement in the arterial phase) target lesions, taking as reference the baseline sum of the diameters of target lesions.
Stable Disease (SD)	Any cases that do not qualify for either PR or PD
Disease Progression (PD)	An increase of at least 20% in the sum of the diameters of viable (enhancing) target lesions, taking as a reference the smallest sum of the diameters of viable (enhancing) target lesions recorded since treatment started.

A responder is for a subject being assessed as CR or PR as opposed to a non-responder is for being assessed as SD or PD (i.e. numerator). Subjects with valid assessment (i.e. CR, PR, SD, or PD) will be considered as evaluable subjects (i.e. denominator).

A subject without 6 month tumor assessment after the treatment initiation (i.e. the 1st DEB-TACE) will be regarded as a missing data and will be excluded from the primary effectiveness endpoint analysis.

3.1.2 Effectiveness Hypotheses

The primary effectiveness hypothesis to be tested is that the 6-month ORR in subjects treated with Tandem meets the PG at a one-sided significance level of 2.5%.

The null hypothesis (H_0) states that the PG is not met as opposed to the alternative hypothesis (H_1) which states that the PG is met. The hypotheses inequalities are shown below:

H_0: Pt ≤ PG (not met)

H_1: Pt > PG (met)

where Pt is the 6-month ORR for the subjects treated with Tandem and the PG is 45%. The PG development is based on the PRECISION V, an international multicenter phase II randomized clinical trial (RCT).

3.1.3 Effectiveness Sample Size

The overall sample size is driven by the primary effectiveness endpoint. Approximately 109 subjects are planned to be enrolled in the single-arm study. The sample size justification is based on the following assumptions.

- Power ≥ 85%
- One-sided significance level (alpha) = 2.5%
- PG for 6-month ORR = 45%

- Expected 6-month ORR = 60%
- A minimum of 98 evaluable subjects to be required at 6 months
- Attrition rate in 6 months ≤ 10%
- A maximum of 109 subjects to be enrolled in the study at baseline

3.1.4 Effectiveness Statistical Methods

A normal approximation test (e.g. Chi-Square Test) for comparing the observed 6-month ORR with the PG will be used to assess the effectiveness hypotheses for a minimum of 98 evaluable subjects.

3.1.5 Worst Case Assessment for Effectiveness

The worst case below reflects the analysis for 98 evaluable subjects. If the actual attrition is better than expected, the worst case scenario will be reassessed.

The PG of 45% will require a minimum of 52 responders at 6 months out of 98 evaluable subjects treated with Tandem. That is, the observed 6-month ORR in Tandem will need to be at least 53.1% (52/98) in order to claim the study effectiveness.

3.1.6 Additional Justification of Sample Size Parameters

The PG of 6-month ORR for Tandem was based on the primary endpoint results (6-month tumor response by EASL) in the PRECISION V (Lammer et. al. 2010), a phase II RCT, published by CardioVascular and Interventional Radiology (Springer).

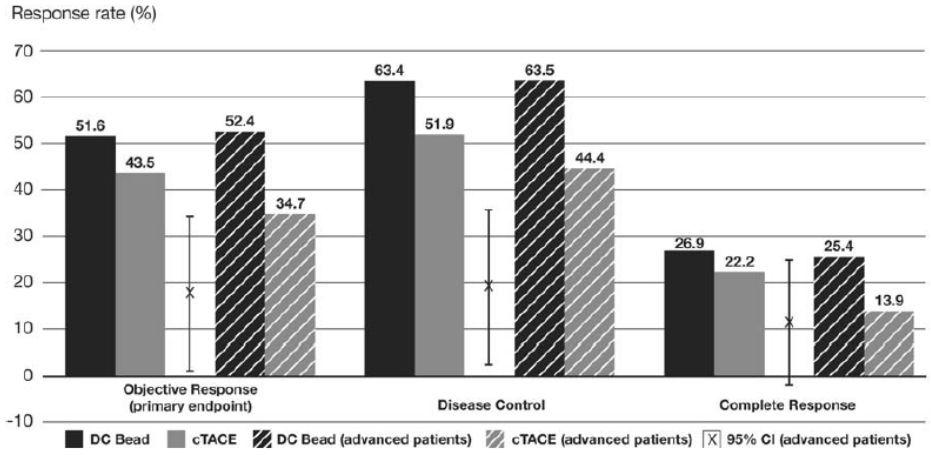

Fig. 2 Tumor response at 6 months (LOCF) (MITT population and advanced patient group*, **). * More advanced disease was at least one of Child-Pugh B, ECOG 1, undergone prior curative treatment (i.e., recurrent disease), and presence of bilobar disease. In accordance with the EASL criteria: complete response (CR)—complete disappearance of all known viable tumor (assessed via uptake of contrast in the arterial phase of the MRI scan) and no new lesions; partial response (PR)—50% reduction in viable tumor area of all measurable lesions; stable disease (SD)—all other cases; progressive disease (PD)—25% increase in size of one or more measurable lesions or the appearance of new lesions. Objective response (OR) was defined as CR + PR, and disease control (DC) as CR + PR + SD. ** Analysis of advanced patient subgroup: OR rate, $P = 0.038$; DC rate, $P = 0.026$; CR rate, $P = 0.091$ (chi-square analysis)

The 6-month ORRs were reported as 51.6% (48/93) vs. 43.5% (47/108) for DC Bead and cTACE respectively. The 95% exact confidence intervals were estimated by SAS: 51.6% with 95% CI: 41.0%, 62.1% and 43.5% with 95% CI: 34.0%, 53.4% for DC Bead and cTACE

respectively. Therefore in order to demonstrate the significant treatment effect over the cTACE, the PG of 45% based on the performance of cTACE was chosen.

3.2 Primary Safety Endpoint

The primary safety hypotheses are not specified. However the safety endpoints are pre-specified in the Safety Plan and will be monitored against the reference rates on a regular basis.

The safety endpoints to be observed in 6 weeks post treatment initiation are:

- All-cause death
- Liver failure
- Hepatic abscess

3.3 Success Criteria

The study success is based on the PG assessment for the primary effectiveness endpoint in at least the minimum of 6-month evaluable subjects. The Tandem will be concluded as meeting the endpoint for device effectiveness if the one-sided lower 97.5% confidence bound of the observed 6-month ORR is greater than zero.

4 GENERAL STATISTICAL CONSIDERATION

4.1 Analysis Sets

The primary and pre-specified additional endpoints will be analyzed on a per-protocol basis. Only enrolled subjects who receive the first treatment will be included in the analysis sample.

4.2 Control of Systematic Error/Bias

Selection of subjects will be made from the Investigators' general or professional referral population. All subjects meeting the inclusion/exclusion criteria and that have signed the protocol-specific ICF will be eligible for enrollment in the trial. Consecutively eligible subjects should be enrolled into the trial to minimize selection bias. In determining subject eligibility for the trial, the investigator's assessment of imaging will be used. However, the MRI core laboratory will independently analyze the images and the data obtained from the core laboratory will be utilized for analyses.

4.3 Enrollment for Each Investigative Site

Each site will be allowed to enroll up to 25% of the total enrollment, with a maximum enrollment of 27 subjects.

4.4 Baseline Data Analyses

Baseline demographics, procedure and assessments will be summarized for the study. For continuous and/or ordinal variables, the descriptive statistics will include mean, standard deviation, number evaluated, minimum and maximum. Some specific variables may also include additional statistics such as median and confidence intervals. For binary or categorical variables,

the descriptive statistics will include percentage, numerator, denominator, and number evaluated. Some variables may include confidence intervals as needed.

The following assessments will be performed during the follow up visits as well as the baseline visit and change from baseline analysis may be performed on clinically meaningful variables.

- Physical examination
- Child-Pugh score
- ECG
- Laboratory tests
- Imaging and tumor measurement
- ECOG performance status
- Concomitant medication

5 ADDITIONAL STATISTICAL ANALYSES

5.1 Secondary Endpoints Assessments

The study is not powered for the secondary endpoints therefore there are no formal hypotheses. The secondary endpoints and analyses are shown below.

- ORR at 30 days and 3-month
- Time to progression (TTP)
- Time to extrahepatic spread (EHS)
- Proportion progression-free (PPF) at one year
- Overall survival (OS) at one year
- The frequency of treatment associated adverse events

Kaplan-Meier analyses will be used for time-to-event type of endpoints such as overall survival (OS), time to progression (TTP), and time to extrahepatic spread (EHS) through 12 months. A binary proportion will be presented for ORR at 30-day and 3-month and progression-free (PPF) rate at 12 months. A frequency will be provided for treatment associated adverse events.

All additional assessments are observational.

5.2 Interim Analyses

No formal interim analyses are planned for the purpose of stopping this study early declaring for efficacy or futility.

5.3 Subgroup Analyses

Primary endpoints and/or additional assessments will be summarized by the following categories (but not limit to):

- Gender (male vs. female)
- ECOG Score (0 vs. 1)
- BCLC Stage (A vs. B7)

All subgroup analyses are observational. No formal tests of hypotheses are proposed for subgroups and therefore alpha-adjustment for multiple comparisons will not be used.

5.4 Missing Data, Drop-Outs, and Protocol Deviations Handling

Boston Scientific will employ robust oversight in order to minimize the loss of subjects throughout any trial follow-up. Additionally, the case report forms are easy-to-follow and maximize the data collection required at each follow-up visit without placing undue burden on the subject. Strategies that are planned to be utilized in the study include:

- Ensure that site personnel are properly trained on the data that is required to be collected and the importance of planning for the follow-up visits.
- Tools in the site's Manual of Operations to assist with follow-up visit planning (e.g. follow-up wheels or similar tools).
- The use of trial newsletters to remind sites of upcoming visits and other project-related milestones to ensure data is being entered promptly and is complete.

5.5 Sensitivity Analysis for Missing Outcome Data

Sensitivity analyses for the primary effectiveness endpoint assessment will be conducted to assess the impact of missing data on the result's robustness. In addition to the use of the worst-case analysis, the tipping point analysis will be performed as post-hoc analysis to consider all combinations of present/absent for all subjects with missing primary outcome.

5.6 Multivariable Analyses

Univariate and multivariable analyses may be performed as post-hoc analyses to assess the effect of potential predictors for the primary effectiveness endpoint in a logistic regression model.

As requested, clinically and/or statistical meaningful baseline covariates will be selected in the regression model. For the 6-month ORR, the predictors will be listed in ascending order of p-value. Univariate analyses will be performed overall as well as separately for each treatment group. For the multivariable analyses, only coefficients in the final model, i.e., with p-values less than 0.1 will be listed.

No formal conclusion will be made by this secondary post-hoc analysis.

5.7 Time-To-Event Kaplan-Meier Analysis

The Kaplan-Meier product-limit method will be used to estimate event or event-free rates for time-to-event outcomes as post-hoc analyses, such as OS, TTP, EHS, and PPF through 12 months.

5.7.1 Kaplan-Meier for OS

OS is defined as the length of time from the treatment initiation, that treated subjects are still alive. The Kaplan-Meier analysis is aimed to capture the all-cause death for each subject. Subjects who are lost-to-follow will be considered being censored.

There are three critical data items to be captured:

- Death date
- The very last date in the database (e.g. follow-up visit date and/or site reported AE date)
- The desired cut-off days (e.g. 6-month OS at 182 days, 12-month OS at 365 days)

In addition to the Kaplan-Meier survival curve, the median survival time (mOS) will also be used to represent the study cohort.

5.7.2 Kaplan-Meier for TTP

TTP is defined as the length of time from the treatment initiation to either the date of the first disease progression occurred, as assessed by the investigators, or the date of the subject died due to any cause, whichever comes earlier.

TTP (also referred as “time to treatment failure”) will be measured by a few data items list below:

- The date of the disease progression in the Overall Response form
- The date of lost-to-follow due to:
  - Adverse events
  - Progressive disease/insufficient therapeutic response
  - Death
  - Failure to return
  - Refusing treatment/being unwilling to cooperate/withdrawing consent
- The very last date by scanning all available dates in the database
- The desired cut-off days

In addition to the Kaplan-Meier estimates, the median will also be used to represent the study cohort.

5.7.3 Kaplan-Meier for EHS

The time to EHS is defined as the length of time from the treatment initiation to the development of extrahepatic spread of the disease via imaging assessment.

The data items to be considered will include:

- The date of the extrahepatic spread in the Overall Response form
- The very last date by scanning all available dates in the database
- The desired cut-off days

5.7.4 Kaplan-Meier for PPF

PPF is defined as the length of time from the treatment initiation, that treated subjects are still progression-free.

The data items to be captured will include:

- The date of the disease progression in the Overall Response form
- The very last date by scanning all available dates in the database
- The desired cut-off days

The proportion of any time point will be determined by the Kaplan-Meier estimates.

5.8 Time to Events and Time to Adequate Follow-Up

The safety binary rates (specified in the Safety Plan), as opposed to Kaplan-Meier rates, will be calculated based on the subjects who have adequate follow-up and/or have experienced pre-specified events.

All reported AEs will be summarized as the safety outcome. The three protocol-defined safety events listed below are only monitored for safety trigger:

- All-cause death
- Liver failure
- Hepatic abscess

5.8.1 Event Rates Presented Using "Month" System

The denominator will be based on number of subjects who reach the protocol-defined lower window (i.e. adequate follow-up days) and/or subjects who experience the event. The numerator will be based on number of subjects-level events within the protocol-defined upper window. Subject-level events beyond the upper window will be counted as next visit.

Follow-up Visit	Protocol Defined Lower Window	Protocol Defined Upper Window
1 Month	16	44
3 Months	77	105
6 Months	169	197
12 Months	351	379

5.8.2 Event Rates Presented Using Exact Days Cut-Off System

The denominator will be based on number of subjects who reach the protocol-defined lower window (i.e. adequate follow-up days) and/or subjects who experience the event. The numerator will be based on number of subjects-level events within the exact desired cut-off days. Subject-level events beyond the exact cut-off days will be counted as next cut-off days.

Follow-Up Cut-Off	Days for Adequate Follow-Up	Maximum Days to Event
30 Days	16	30
91 Days	77	91

Follow-Up Cut-Off	Days for Adequate Follow-Up	Maximum Days to Event
183 Days	169	183
365 Days	351	365

5.8.3 Missing Event Dates Considerations

All event rates will be calculated relative to the date of procedure (i.e. post-procedure).

When calculating rates of adverse events with missing event date (i.e. mm/dd/yyyy), the ideal is to work with safety and/or data management representatives to query sites for missing data. However missing and partial missing dates may be handled as using the worst case scenario as follows:

Partial Date Description	Action Taken
Entire onset date is missing	The procedure date will be used for the onset date.
The month and the day of the month are missing but the year is available	January 1st will be used for the month and day of the onset date. However, if the imputed date falls before the procedure date, then the procedure date will be used for the onset date.
Day is missing, but the month and year are available	The 1st will be used as the day of the onset date. However, if the imputed date falls before the procedure date, then the procedure date will be used for the onset date.

5.9 Analysis of Site-Reported Serious and Non-Serious Adverse Events

Subject-level event rates will be calculated at various time points (e.g. exact days) based on all events reported by the site regardless of whether or not they are ultimately adjudicated. These safety parameters will be summarized using descriptive statistics.

5.10 Scoring Systems of Target Lesion for HCC

There is a few scoring systems to measure a liver function in the study.

5.10.1 ECOG Performance Status

The Eastern Cooperative Oncology Group (ECOG) performance status describes a subject's level of functioning in terms of the ability to self-care, daily activity, and physical ability (e.g. walking, working, etc.).

Score	Description
0	Fully active, able to carry on all pre-disease performance without restriction
1	Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature (e.g. light housework, office work).

2	Ambulatory and capable of all self-care but unable to carry out any work activities. Up and about more than 50% of waking hours.
3	Capable of only limited self-care, confined to bed or chair more than 50% of waking hours.
4	Completely disabled. Cannot carry on any self-care. Totally confined to bed or chair.
5	Dead

5.10.2 Child-Pugh Score

The Child-Pugh score is a measure of liver function based on the following 5 items that tells how well the liver is working.

- Bilirubin levels in the blood
- Albumin levels in the blood
- How quickly the blood clots (prothrombin time)
- If fluid has collected in the abdomen (ascites)
- Brain function (encephalopathy)

Each one is given a number score as follows:

		1 point	2 points	3 points
Serum Bilirubin	mg/dL	<2	2-3	>3
	µmol/L	<34.2	34.2-51.3	>51.3
Serum Albumin	g/dL	>3.5	2.8-3.5	<2.8
	g/L	>35	28-35	<28
INR		<1.7	1.7-2.30	>2.30
Ascites		None	Mild	Severe
Hepatic encephalopathy		None	Grade I-II	Grade III-IV

If there are several test results for one test item, the lower point (the worst case) will result will be used to determine the Child-Pugh class. Based on that score, people fall into 1 of 3 classes:

- Class A (5-6 points): the liver is working normally.
- Class B (7-9 points): the liver is mild to moderate illness and the subject may be offered treatment such as surgery or chemoembolization.
- Class C (10-15 points): there is severe liver damage; unfortunately the outlook is then quite poor. The subject is often too sick to have treatment for the cancer.

Chronic liver disease is classified into Child-Pugh class A to C. The one-year and two-year survival rates are based on meta-analysis (Cholongitas, E. et al, 2005, Alimentary Pharmacology

and Therapeutics). The category may be classified further such as "B7" representing Child-Pugh class B with a score of 7. The interpretation for Child-Pugh scores are shown below:

Point	Class	One year survival	Two year survival
5-6	A	100%	85%
7-9	B	81%	57%
10-15	C	45%	35%

5.10.3 mRECIST Criteria

The mRECIST response assessment of target lesion for HCC is to measure on MRI in HCC subjects treated with chemoembolization to predict long-term survival. The EASL criteria describes below:

Response Category	mRECIST Criteria
Complete Response (CR)	Disappearance of any intratumoral arterial enhancement in all target lesions
Partial Response (PR)	At least a 30% decrease in the sum of the diameters of viable (enhancement in the arterial phase) target lesions, taking as reference the baseline sum of the diameters of target lesions.
Stable Disease (SD)	Any cases that do not qualify for either PR or PD
Disease Progression (PD)	An increase of at least 20% in the sum of the diameters of viable (enhancing) target lesions, taking as a reference the smallest sum of the diameters of viable (enhancing) target lesions recorded since treatment started.

5.10.4 EASL Criteria

The EASL response assessment of target lesion for HCC is to measure on MRI in HCC subjects treated with chemoembolization to predict long-term survival. The EASL criteria describes below:

Response Category	EASL Criteria
Complete Response (CR)	Disappearance of all known disease and no new lesions determined by two observations not less than 4 weeks apart
Partial Response (PR)	At least 50% reduction in total tumor load of all measurable lesions determined by two observations not less than 4 weeks apart
Stable Disease (SD)	Any cases that do not qualify for either PR or PD

Disease Progression (PD)	At least 25% increase in size of one or more measurable lesions or the appearance of new lesions

5.11 Analyses Software

All statistical analyses will be performed and validated by the independent CRO (e.g. IQVia in Bangalore) using the Statistical Analysis Software (SAS), version 9.2 or later (Copyright © 2002-2010 by SAS Institute Inc., Cary, North Carolina 27513, USA. All rights reserved). BSC will review statistical reports.

5.12 Changes to Planned Analyses

Any changes to the planned statistical analyses made prior to performing the primary endpoint analysis will be documented in an amended Statistical Analysis Plan approved prior to performing the analysis. Changes from the planned statistical methods after performing the analysis will be documented in the clinical study report along with a reason for the deviation.

6 VALIDATION

All clinical data reports generated per this plan will be validated per 90702587, Global WI: Clinical Data Reporting Validation. Statistical analyses and validation will be done by an independent CRO.

7 PROGRAMMING CONSIDERATIONS

All statistical programming tasks will be performed by the independent CRO.

7.1 Derivation of Variables

- The number of subjects included in the event rates (overall and individual components) will be based on subjects who have adequate follow-up and/or have experienced any component of events within the analysis interval.
- The last follow-up date will be the latest of the following dates for each subject: onset date of an event, DEB-TACE dates, treatment evaluation follow-up dates, end of study date, end of treatment date, and follow-up visit dates.
- Days to event (or last known status) = event (or status) date – index procedure date (e.g. DEB TACE1).

7.2 SAS Codes for Chi-Square Test

The confidence intervals and the Chi-Square test p-value can be produced by the following SAS code.

```
proc freq data=;
    tables xx/binomial(p=) alpha=;
```

```
run;
```

For example, the worst case scenario in the PG testing of the primary effectiveness hypotheses is used for the exercise. A dummy frequency table is coded as below.

```
%let total=98; %let yes=54; %let no=&total.-&yes.;

data dsn1;
    yn=1; wgt=&yes.; output;
    yn=0; wgt=&no.; output;
run;
```

A list of SAS codes for Binomial and Chi-Square method is used for PG=45% and 95% confidence limits to control the Type I error under one-sided 2.5% (i.e. two-sided 5%).

```
proc freq data=dsn1 order=data;
   tables yn/ binomial(p=.45) alpha=.05;
   weight wgt;
run;
```

The SAS output for the worst case scenario is presented below.

```
        Binomial Proportion
                yn = 1

Proportion                   0.5510
ASE                          0.0502
95% Lower Conf Limit         0.4525
95% Upper Conf Limit         0.6495

Exact Conf Limits
95% Lower Conf Limit         0.4472
95% Upper Conf Limit         0.6517

 Test of H0: Proportion = 0.45

ASE under H0                 0.0503
Z                            2.0102
One-sided Pr >  Z            0.0222
Two-sided Pr > |Z|           0.0444

        Sample Size = 98
```