Statistical Analysis Plan (SAP)

A Single-Arm, Prospective, Multicenter Study to Evaluate the Safety and Effectiveness of the NeuWave Certus Microwave Ablation System in Chinese Patients with Hepatocellular Carcinoma

Protocol Version: NEU_2017_07 v4.0 (dated 06August2020)

SAP Version: 1.0
Date: 28Apr2021

A Single-Arm, Prospective, Multicenter Study to Evaluate the Safety and Effectiveness of the NeuWave Certus Microwave Ablation System in Chinese Patients with Hepatocellular Carcinoma

Protocol Version: NEU_2017_07 v4.0 (dated 06August2020)

The following individuals have reviewed this version of the Statistical Analysis Plan and are in agreement with the content:

Signature Page

[redacted]

(Print)

Revision History

Revision Number	Revision Date (DD/MM/YYYY)	Reasons for Revision
1.0	28/04/2021	Not Applicable

Table of Contents

1 Study Design

This is a single-arm, prospective, multicenter study. Individuals who are assessed for NeuWave Certus Microwave (MW) ablation of hepatocellular carcinoma (HCC) in accordance with their institution's standard of care, who meet study entry criteria, and sign the informed consent will be enrolled. The subjects will be treated with MW ablation and will be followed for up to 36 months after the ablation procedure to assess safety and ablation outcomes.

The total sample size will be 137 subjects. To provide sites with an opportunity to get equal experience in the use of the Certus system, there will be 3 patients per site treated as part of a run-in phase. There will be 4 sites participating in the study. Therefore, 12 of the 137 subjects will be included in the run-in phase.

2 Treatment Assignment

All subjects enrolled will undergo microwave ablation of HCC using only the NeuWave Certus Microwave Ablation System.

3 Randomization and Blinding Procedures

This is a single-arm study. Hence, there will be no control or comparison group. Therefore, there is no need for randomization or blinding.

4 Visit Schedule / Interval Windows

Table 1: Study Timepoint

Visit Number	**Visit Label for Statistical Outputs**	**Visit Name**	**Study Phase**	**Visit Name (Visit Window)**
1	Visit 1	Screening	Screening	Screening (within 14 days before ablation)
2.1	Visit 2A	Day 0	Ablation	Ablation (day 0)
2.2	Visit 2B	Post-Ablation	Post-Ablation	Post-Ablation (0 to 7 days)
3	Visit 3	Month 1	Follow-up	1 months (± 7 days) post-ablation
4	Visit 4	Month 3		3 months (± 7 days) post-ablation
5	Visit 5	Month 6		6 months (± 14 days) post-ablation
6	Visit 6	Month 9		9 months (± 14 days) post-ablation
7	Visit 7	Month 12		12 months (± 14 days) post-ablation
8	Visit 8	Month 18		18 months (± 28 days) post-ablation
9	Visit 9	Month 24		24 months (± 28 days) post-ablation
10	Visit 10	Month 30		30 months (± 28 days) post-ablation
11	Visit 11	Month 36		36 months (± 28 days) post-ablation
XX.X	USV	Unsheduled		Unscheduled Visit

The visit in the outputs should be derived based on the above visit schedule.

If a visit is performed “out of window” (i.e.between two visit windows), the visit should be analyzed as part of the the visit closest in terms of time or number of days as outlined by visit schedule, above. More details can be found in Table 1 of the protocol.

In case any retest performed within one visit, the visit in the outputs will be derived based on the above time windows. The visit with worst result will be used as the planned visit.

5 Levels of Significance

Given that no hypotheses are formally being tested in this study, no levels of significance are specified. Estimation of specified study endpoints will be performed using two-sided 95% confidence intervals (CI).

6 Analysis Sets

6.1 ***All Subjects Screened Set (SCR)***

The All Subjects Screened Set (SCR) contains all subjects who provide informed consent for this study. The SCR will be used for disposition of study subjects.

6.2 ***Full Analysis Set (FAS)***

The Full Analysis Set (FAS) contains all subjects who are enrolled in the study, are not part of the run-in phase, and have the NeuWave Microwave Ablation System used for ablation. The FAS will be used for effectiveness endpoints and safety analysis.

Enrolled subjects means subjects who meet study entry criteria and sign the informed consent.

6.3 ***Per Protocol Analysis Set (PPS)***

The Per Protocol Analysis Set (PPS) contains all subjects in the Full Analysis Set who have no major protocol deviations that impacting the effectiveness endpoint. PPS will be used for effectiveness analyses.

6.4 ***Safety Analysis Set (SAF)***

The Safety Analysis Set (SAF) contains all subjects who provide informed consent and have MW ablation attempted with the NeuWave system and includes run-in phase subjects. The Safety Set will be used for summarization of safety endpoints only.

6.5 Run-in Analysis Set

The Run-in analysis set contains all subjects participating in the run-in phase.

7 Sample Size Justification

The total sample size will be 137 subjects. This includes 3 patients at each site (12 patients across 4 sites) who will be part of the run-in phase. There are no stated hypotheses in this study from which a power calculation and sample size determination will be performed; rather, a minimum of 100 subjects providing data to at least 1 year is determined to be sufficient to provide an evaluation of safety and effectiveness of the NeuWave Certus Microwave Ablation System. In anticipation of up to 20% dropout, a total of 125 subjects will be enrolled after completion of the run-in phase for a total of 137 subjects.

8 Analyses to be Conducted

8.1 ***General Conventions***

8.1.1 Reference Start Dates

Reference start date is defined as the day of subjects undergoing first microwave ablation, considered as Day 0 (Day 1 is the next day of the first microwave ablation) and appear in every listing where an assessment date or event date appears.

8.1.2 Study day

The onset day of AE is calculated from the date of first microwave ablation administrated to the date of the AE start date, as follows:

Onset Day = (date of event – date of first microwave ablation + 1) if date of event happens on or later than date of ablation.

Onset Day = (date of event – date of first microwave ablation) if date of event happens earlier than date of ablation.

Study Day will be calculated from the reference start date (date of first microwave ablation) and will be used to show start/stop day of assessments and events.

Study Day = (date of event – date of first microwave ablation + 1) if date of event happens on or later than date of ablation

Study Day = (date of event – date of first microwave ablation) if date of event happens earlier than date of ablation.

In the situation where the event date is partial or missing, the Onset Day/Study Day and any corresponding durations will appear missing.

If a date is partial, no imputation will be done except for adverse events (AEs). If the day of AE start is missing, the first day of the month is considered as the start date, unless the month is the same month as the month of ablation, in which case day will be set equal to the day of ablation. If month part is missing, the AE start date is imputed as the first day of the year, unless the same year of ablation, in which case day will be set equal to the day of ablation. If the day or month of AE end is missing, the day will be imputed as the end day of the month or the year, unless the same month or year as the day of ablation. If the imputed date happens later than death date, the date is replaced with death date.

Based on the protocol, there will be no imputation of missing data for any parameters or for early terminated patients.

8.1.3 Baseline

Unless otherwise specified, baseline is defined as the last non-missing measurement taken prior to the date of initial ablation (Day 0) (including unscheduled assessments). Ultrasound and/or CT scan on Day 0 is taken prior to ablation, and is considered baseline (including unscheduled assessments).

Change from baseline will be defined as the difference between the post-baseline value at the point of interest and the baseline value.

8.1.4 Visit 2 Summaries, Retests, Unscheduled Visits

Visit 2 is split into 2 distinct phases: ablation (visit 2A) and post-ablation (visit 2B). Listings and summaries will be presented by phase of the visit, as appropriate.

In general, for by-visit summaries, data recorded at the scheduled visit will be presented. Unscheduled and reduced follow-up measurements will not be included in by-visit summaries.

Listing will include scheduled, unscheduled, retest.

8.1.5 Last Contact Date

The last contact date will be derived for subjects not known to have died at the analysis cut-off date using the latest date among the following applicable or available data:

- All visit dates, including examination/scan dates (e.g. laboratory, vital signs, NPRS pain, EORTC QLQ- HCC 18/C 30, ECG, ECOG, tumor imaging scan, tumor assessment)
- Medication/procedure dates (including ablation, concomitant medications/surgeries, therapies/procedures administered after study treatment discontinuation, etc.)
- Adverse event start and end dates

- Dates of protocol deviations
- Date of trial completion/discontinuation in [CRF form Subject Completion/Discontinuation]

8.1.6 Cut-off conventions

All available data will be included at the time of the cut-off date. Missing or partial dates (unable to compare) will also be presented, but no imputation for any date will occur except for partial adverse event dates as described in section *8.1.2*.

Adverse events and concomitant medications with no end date available by the cut-off date will be considered as ongoing events. If the start date is prior to / on the cutoff date, the end dates will be presented regardless if the end dates are post cut-off date or not. Ongoing events will be summarized for analysis using the cut-off date as the date of completion, with an indication that event is ongoing.

8.1.7 Conversion factors

The following conversion factors will be used to convert days into months or years: 1 month = 30.4375 days, 1 year = 365.25 days.

8.1.8 Reporting Conventions

Quantitative variables will be summarized using descriptive statistics (by visit, when appropriate) and will consist of values for number of evaluable subjects, mean, standard deviation (SD), median,Q1, Q3, minimum, and maximum. In calculated statistics, mean, median, Q1, and Q3 will be displayed with one more digit than original data while the SD will be displayed with two more digits. Minimum and maximum will keep the same number of decimal digits as the original data.

Table 2 Presentation of Numbers

Statistics	Format
N	xx
n	xx
Mean	x.x
SD	x.xx
Median	x.x
Q1, Q3	x.x, x.x
Min, Max	x, x
SE	x.xx
95% CI	[x.x, x.x]
P-Value	0.XXX
Missing	xx
Percentage	x.x%

N = total number of subjects in the defined analysis set; n = total number of subjects in the specific category. If N is specified in the column heading, then any reference to the number of subjects in the body should be no larger than N.

Qualitative variables are summarized by counts and percentages and displayed by decreasing order of modalities if ordinal, and by alphabetical order otherwise. The number of subjects with missing data should be displayed when there is missing data (number of total evaluable subjects is less than number of total subjects in the defined analysis population set) . The numerator is the total number of evaluable subjects in the specific category (n).

Unless otherwise stated, for percentages:

- One decimal digit should be given
- No percentage should be given for cells with 0 subjects
- Based on the number of subjects in the analysis set
- Counts of missing observations will not be included in the denominator, but will be presented.

8.1.9 Presentation of Tables/Listings/Figures:

The treatment label for all Tables, Listings, and Figures will be: NeuWave MW Ablation.

All data recorded during the trial will be presented in individual data listings on the SAF, unless otherwise specified (including scheduled and unscheduled data). All listings will be sorted by phase (main or run-in), site, subject, parameter, date, and time point (where applicable), if not otherwise stated. Further details are provided in the appropriate section for the analysis of the specific parameter.

Tables and figures will be presented by scheduled time point (where applicable), unless otherwise specified. By-visit tables and figures will be sorted by chronological scheduled time point (where applicable).

8.1.10 Software

All analyses will be conducted using SAS version 9.4 or higher.

8.2 ***Disposition of Study Subjects***

Subject disposition will be summarized for the SCR. The number and percentage of subjects screened, screen failure, treated, completed and discontinued will be tabulated along with the specific reasons for discontinuation. Analysis population (ENR, Run-in, FAS, PPS and SAF) will be presented. Listing of subject disposition and eligibility criteria will be prepared.

Counts and percentages of protocol deviations (PD) recorded during the study will be provided for the type of deviation, the rationale for the deviation, and classification (minor or major). Listing of protocol deviations will be presented for SAF population. Sort by phase (main or run-in), site, subject ID, and date of deviation.

8.3 ***Demographic and Baseline Characteristics***

8.3.1 Demographics

Demographic data and other baseline characteristics will be presented for the FAS and Run-in analysis set.

The following demographic and other baseline characteristics will be reported for this study:

- Sites
- Age (years) at informed consent date
- Age category: <= 65 years, and > 65 years
- Sex (including childbearing potential of female subjects)
- Ethnicity
- Race
- Pregnancy Test Result
- Alcohol Use
- # of target tumor(s) (1,2,3)
- Baseline ECOG performance status (0-2)
- Baseline Child-Pugh score
- Baseline ASA score
- Baseline NPRS Pain score

8.3.2 Baseline Target Tumor Assessement (Analyzed on tumor level)

Baseline target tumor assessment will be summarized for FAS and Run-in

In “[Target Tumor Assessments]” CRF page, maximum tumor dimension, tumor type and vertical dimension be summarized, same as other continuous measurement for each target lesion. Counts for liver segment and major vessel assessment will be tabulated at baseline.

Other details on Target tumor assessments will only be listed, sorted by phase, site, subject, target tumor.

8.3.3 Medical History and Surgical History

General medical history will be presented for the FAS and Run-in. Number and percentage of subjects with at least one medical history will be tabulated by System Organ Class (SOC) and Preferred Term (PT). Medical history will be coded using the MedDRA codes (MedDRA version 22.0 or above). A listing of all subject medical histories will be prepared.

Surgical histories and radiotherapy hstories will also be presented in a listing for the SAF.

8.3.4 Prior and Concomitant Medications and Procedures

Medications and procedures will be presented for the SAF. Medications will be presented by preferred term. Medications will be coded using the World Health Organization Drug Dictionary (WHO-DD) latest version.

Prior medications are defined as the medications/therapies that were taken and ended prior to the procedure; while concomitant medications are defined as the medications/therapies with end date after the procedure (Day 0 but after the procedure, or after Day 0) or ongoing medication/therapies.

Summary of indications for concomitant medications and procedures(indication, related to AE or not) will also be tabulated.

Prior and concomitant medications and procedures will be listed.

8.4 ***Effectiveness and Safety Endpoint(s) and Associated Hypotheses***

In this study, all analyses are for descriptive purpose only. No hypothese testing will be conducted on the effectiveness and safety endpoints.

8.4.1 Effectiveness Endpoint(s)

All effectiveness endpoints will be summarized for FAS and PPS. Subjects in run-in analysis set will be also summarized for technical success and technical efficacy endpoints.

The effectiveness endpoints are defined, as follows:

Technical Success

Technical Success, defined as a combination of A0 ablations (complete tumor ablation with a surrounding 5 mm margin) and A1 ablations (complete tumor ablation with inadequate margins) based on contrast-enhanced MRI performed at Visit 2B (up to 7 days following the original ablation procedure).

The number and percentage of tumors achieving Technical Success will be summarized. Percentage will be calculated relative to the total count of evaluable target tumor lesions at Visit 2B. A 95% confidence interval will be estimated using wald method. The number and percentage of tumors with A0, A1 and A2 margins will be presented as a sub-categorization :

> Note: Ablation evaluation by Investigator is collected in “Ablation Zone Assessment” CRF page.
> A0: complete tumor ablation with adequate margin (surrounding 5 mm margin)
> A1: complete tumor ablation with inadequate margins
> A2: Incomplete tumor ablation

A subject archieving technical success is defined as a subject with all tumors reaching A0 or A1 at Visit 2B. The number and percentage of subjects archieving Technical Success will be summaried as well. Percentage will be calculated relative to the total number of subjects in relavent analysis population (FAS or PPS). A 95% confidence interval will be estimated using wald method.

Number and percentages of reasons for A1 and A2 ablation categorization will be also summarized.

Technical Efficacy

Technical Efficacy, defined as a combination of A0 ablations (complete tumor ablation with a surrounding 5 mm margin) and A1 ablations (complete tumor ablation with inadequate margins), based on contrast-enhanced MRI scans at Visit 3 (1 month +/- 7 days) after the original ablation procedure.

Tumor-level analysis method is same as that of primary endpoint.The number and percentage of tumors achieving Technical Efficacy will be summarized and a 95% confidence interval will be estimated using Wald method. The number and percentage of tumors with A0 and A1 margins will be presented as a sub-categorization within each presentation of Technical Efficacy, as well as a summarization for the number and reasons for repeat ablations through Visit 3.

A subject archieving technical efficacy is defined as a patient with all tumors reaching A0 or A1 at Visit 3. The number and percentage of subjects archieving Technical Efficacy will be summaried as well. Percentage will be calculated relative to the total number of subjects in relavent analysis population (FAS or PPS). A 95% confidence interval will be estimated using wald method. But the subjects who have any A2 margin since the visit 2B scan would not be counted as technical efficacy regardless if the tumor was corrected with repeat ablation.

Local tumor progression (LTP) (the 2nd interim, final)

LTP is evaluated after the ablation of the target tumor(s). LTP describes the appearance of tumor foci at the edge of the ablation zone, after at least one contrast-enhanced follow-up

MRI scan has documented adequate ablation, and an absence of viable tissue in the target tumor and surrounding ablation margin, by using imaging criteria.

Tumor-level analysis method is same as that of Technical Success endpoint. The number and percentage of tumors with LTP will be summarized and a 95% confidence interval will be estimated using Wald method.

A subject with LTP is defined as a patient with any target tumor LTP. The number and percentage of subjects with LTP. Percentage will be calculated relative to the total number of subjects in relevant analysis population (FAS or PPS).

Subgroup analysis is defined in SAP section 8.11. Two subgroup columns (subjects/tumors with only initial ablation and subjects/tumors with repeat ablation) and 1 total column will be used as analysis columns in LTP tables.

Tumor type (primary HCC and recurrent HCC and other) will be another subgroup variable for LTP (analyzed for patient level and tumor level).

LTP start date is the date of initial ablation of a subject. The event date of LTP is collected in “Ablation Zone Assessment” CRF page as following table:

Table 3: Censoring Rules for LTP

	Situation	**Event/Censoring Date**	**Outcome**
No post-baseline or no baseline Ablation Zone Assessment			
A1	No post-baseline local tumor evaluation	Date of initial ablation	Censored
A2	No baseline local tumor evaluation	Date of initial ablation	Censored
With baseline and post-baseline tumor assessments			
B1	Local Tumor Progression and zero or one missed tumor assessment prior to confirmed LTP	Date of first LTP date	Event
B2	Local Tumor Progression and over one missed tumor assessment prior to confirmed LTP	Date of the latest evaluable local tumor assessment prior to the missed assessment	Event
C	No progression, either alive or death	Date of last local tumor assessment with no documented LTP	Censored

D	Dead	Date of last tumor assessment with no documented progression	Censored
E	Withdrawal by Subject	Date of last tumor assessment with no documented progression	Censored

Local tumor progression rates at 3, 6, 9, 12, 18, 24, 30, and 36 months will be estimated using the Kaplan-Meier (K-M) method and 95% confidence intervals (Greenwood's formula) will be provided.

The K-M curve of time to LTP will be provided.

Secondary efficacy rate (the 2nd interim, final)

Secondary efficacy rate is defined as the percentage of subjects whose tumors have undergone successful repeat ablation following identification of local tumor progression at any time during study follow up.

It could be found in CRF page, Ablation Zone Assessment, "Re-ablate with Neuwave". For subgroup that has LTP and used Certus to re-ablate, the number and percentage of subjects achieving secondary efficacy, will be summarized and a 95% confidence interval will be estimated using wald method.

Progression-free survival (PFS) (the 2nd interim, final)

PFS is defined as length of time from initial ablation until disease progression.

Table 4: Censoring Rules for PFS

	Situation	**Event Date**	**Outcome**
No post-baseline or no baseline Ablation Zone Assessment			
A1	No post-baseline tumor evaluation and alive	Date Ablation performed	Censored
A2	No post-baseline tumor evaluation but died on or before the second scheduled tumor assessment	Date Ablation performed	Censored
With baseline and post-baseline tumor assessments			

B	Any Tumor Progression and zero or one missed tumor assessment prior to progression	Date of first progression	Event
C	PD happens after >=2 consecutively missing tumor assessment	Date of last progression assessment before missing assessments	Censored
D	No progression	Date of last progression assessment with no documented progression	Censored

Progression-Free Survival rate will be estimated using the Kaplan-Meier method and 95% confidence intervals (Greenwood's formula) will be provided for 6, 12, 18, 24, 30, 36 months.

The K-M curve of time to progression will be provided.

Progression-free survival (PFS) sensitivity analysis (the 2nd interim, final)

PFS is defined as length of time from initial ablation until disease progression or death, whichever comes first.

Overall survival (OS)

36-month OS is measured from the time of the first ablation procedure to the time of death or last follow-up, if death has not occurred.

For OS, death from any cause will be considered as an event. If no death is reported, then use the cut-off date as the censored date of OS.

Overall Survival rate will be estimated using the Kaplan-Meier method and 95% confidence intervals (Greenwood's formula) will be provided.

The K-M curve of time to death will be provided.

Quality of Life (QoL)

EORTC QLQ-C30

The EORTC QLQ-C30 (version 3.0) comprises 30 questions and provides a multi-dimensional assessment of QoL.

The QLQ-C30 is composed of both multi-item scales and single-item measures. EORTC QLQ-C30 can be combined to produce 5 functional scales (Physical, Role, Cognitive, Emotional, and Social), 3 symptom scales (Fatigue, Pain, and Nausea/vomiting), 5 individual items (dyspnea, insomnia, appetite loss, constipation, and diarrhea) and a global measure of health status. Each of the multi-item scales includes a different set of items - no item occurs in more than one scale. All of the scales and single-item measures range in score from 0 to 100. A high scale score represents a higher response level.

There are four items for each question (from Q1 to Q28): from not at all, a little, quite a bit, to very much, each will be considered as from 1 to 4. For Q29 and Q30, will be scored from 1 to 7.

No response (NR) will not be considered as missing and not included in the analysis.

Table 5: Scoring the EORTC QLQ-C30 version 3.0

Scale	Number of items	Item range*	Item numbers
Global health status / QoL			
Global health status/QoL	2	6	29, 30
Functional scales			
Physical functioning	5	3	1 to 5
Role functioning	2	3	6, 7
Emotional functioning	4	3	21 to 24
Cognitive functioning	2	3	20, 25
Social functioning	2	3	26, 27
Symptom scales / items			
Fatigue	3	3	10, 12, 18
Nausea and vomiting	2	3	14, 15
Pain	2	3	9, 19
Dyspnoea	1	3	8
Insomnia	1	3	11
Appetite loss	1	3	13
Constipation	1	3	16
Diarrhoea	1	3	17
Financial difficulties	1	3	28

* *Item range* is the difference between the possible maximum and the minimum response to individual items; most items take values from 1 to 4, giving *range* = 3.

The scoring procedure is as follows:

Step1: For all scales, the raw score (RS) is calculated as mean of the component items:

$$RawScore = RS = (I_1 + I_2 + \ldots + I_n)/n.$$

If at least half of the component items from the scale have been answered, the raw score is calculated as the average of the non-missing items, otherwise set to missing.

Step 2: Scale score (S) is calculated by a linear transformation to 0-100:

For Functional scales:

$$Score = S = \left(1 - \frac{RS-1}{Range}\right) \times 100$$

For Symptom scales / items and Global health status / QoL:

$$Score = S = \left(\frac{RS-1}{Range}\right) \times 100$$

where Range is the difference between the maximum possible value of RS and the minimum possible value. The EORTC QLQ-C30 has been designed so that all items in any scale have the same range. Therefore, the range of RS equals the range of the item values. Most items are scored 1 to 4, giving range = 3. The exceptions are the items contributing to the global health status / QoL, which are 7-point questions with range = 6.

Higher scores for the global health status/ QoL scale and functioning scales indicate a higher/healthier level of functioning and a higher/better QoL respectively, whereas higher scores in symptom scales represent a higher level of symptoms/problems.

Example:

Emotional functioning (EF), Raw score (RS) = (Q21+ Q22+ Q23+ Q24)/4
EF score = {1 – ((RS – 1)/3)} *100

If at least half of the items (i.e. 3 of 6 items, or 3 of 5 items) from the domain have been answered, assume that the missing items have values equal to the average of those items which are present for that respondent. This is equivalent to the raw score being taken as the mean of the non-missing item values. As a result none of the single-item measures can be imputed, i.e. if a single-item measure is missing, the subscale is missing.

Example:

Emotional Functioning if Q23 is missing (i.e. 3 items are not missing)
Raw Score (RS) = (Q21+Q22+Q24)/3
EF Score = {1 – ((RS – 1)/3)} *100

All of the scales and single-item measures are linearly transformed so that each score ranges from 0 to 100.

The EORTC QLQ-C30 Summary Score is calculated from the mean of 13 of the 15 QLQ-C30 scales (the Global Quality of Life scale and the Financial Impact scale are not included).

Prior to calculating the mean, the symptom scales need to be reversed to obtain a uniform direction of all scales. The summary score will only be calculated if all the required 13 scale scores are available (using scale scores based on the completed items, provided that at least 50% of the items in that scale have been completed.)

Observed values by visit and changes from baseline in EORTC QLQ-C30 scores in each scale/domain and totally are presented as n, mean, SD, median, Q1, Q3, minimum, and maximum. In addition, raw scores of EORTC QLQ-C30 are listed.

The changes from baseline in EORTC QLQ-C30 scores in each scale/domain and totally are presented as mean, 95%CI and p value compare with 0.

Liver-specific EORTC QLQ-HCC18

EORTC QLQ-HCC18 includes 18 multi-item scales. These items are grouped into 6 domains namely fatigue, body image, jaundice, nutrition, pain, and fever. Two remaining single items address abdominal swelling and sex life. All scales are grouped and converted to score 0 to 100 according to the scoring manual; a higher score represents a more severe symptom or problem.

Table 6 Scoring the EORTC QLQ-HCC18

Scale	Number of items	Item range*	Item numbers
Symptom scales / items			
Fatigue	3	3	15. 16, 17
Body Image	2	3	3, 5
Jaundice	2	3	6, 7
Nutrition	5	3	1, 2, 12, 13, 14
Pain	2	3	8, 9
Fever	2	3	10, 11
Abdominal Swelling	1	3	4
Sexual Interest	1	3	18

* *Item range* is the difference between the possible maximum and the minimum response to individual items; most items take values from 1 to 4, giving *range* = 3.

The scoring procedures as below:

For each of these domains, a symptom score is computed by taking the following steps:

(1) sum across the item responses on that domain,

(2) divide by the number of non-missing items to obtain the mean of item response (RS),

(3) subtract 1 from the mean of the item responses (i.e., item response mean – 1),

(4) divide by the number in Item Range

(5) multiply by 100.

This results in symptom scores that range from 0 to 100 with higher scores indicating greater symptom severity. These symptom scores can be computed as long as at least 50% of the items have responses (i.e., 1 item response needed for 1 item domains, 1 item response needed for 2 item domains, 2 item responses needed for 3 item domains).

HCC18 index-score was defined as the sum of all 8 QLQ-HCC18 symptom/problem scales divided by 8 (the total number of QLQ-HCC18 scales). A higher HCC18 index-score reflects a worse overall HRQOL.

Observed values by visit and changes from baseline in EORTC QLQ-HCC18 scores in each domain and totally are presented as n, mean, SD, median, Q1, Q3, minimum, and maximum. In addition, raw scores of EORTC QLQ-HCC18 are listed.

The changes from baseline in EORTC QLQ-HCC18 scores in each scale/domain and totally will be estimated by mean, SE, 95%CI and p value.

Health Economics (the 1st interim)

Health economics will be summarized for the FAS and run-in. Procedure details, device information (as described in the “ABLATION” CRF page), and hospital discharge (as described in the “DISCHARGE FORM” CRF page) will be summarized by ablation (initial only vs repeat ablation)and listed.

- Complete Procedure Time
- Length of hospital stay.

- Ablation Time
- Number of Ablations
- Max Power
- Max Temperature
- Max Number of Probes of Each Type (PR, LK, and SR) and Overall
- Number of Probe RepositionNumber of ultrasound/CT scans performed for probe assessment
- Number of ultrasound/CT scans performed for margin assessment
- Other Ablation related information will be listed.

8.5 **Safety Analyses**

All outputs for safety outcomes will based on FAS and SAF.

8.5.1 Adverse Events

An AE is defined as any undesirable clinical occurrence in a patient. An AE does not necessarily have a causal relationship with the study medical device. AEs will be recorded since date of informed consent signature until 30 days after any ablation procedure except SAEs with 36 months follow-up period.

Adverse events will be summarized by System Organ Class (SOC) and Preferred Term (PT) and broken down further by maximum severity and relationship to study device or procedure. Number (n) and percentage (%) will be presented with the percentage calculated relative to the total number of subjects in the FAS/SAF as following:

- AE Overview (subgroup by initial only and repeat ablation, run-in analysis set and full analysis set, and total count of safety analysis set)
- AE by SOC and PT
- Procedure related AE by SOC and PT
- Study device related AE by SOC and PT
- Serious AE by SOC and PT
- Procedure related serious AE by SOC and PT
- Study device related serious AE by SOC and PT
- AE by SOC, PT, and maximum severity
- AE by SIR and PT
- AE listings will be prepared. Information including SIR, PT, relationship with device/procedure, seriousness, severity, AE start and end date, intervention/treatment taken, outcome of AEwill be presented.

AEs will be coded using MedDRA Version 19.0 or higher.

Procedure-related AE

A procedure-related AE is an adverse event related to the study procedure. The relationship with study procedure is recorded as “Unlikely related”, “Possibly related”, “Probably related”, and “Causal relationship”.

Device-related AE

A device-related AE is an adverse event related to the use of the study device. The relationship with study device is recorded as “Unlikely related”,“Possibly related”, “Probably related” and “Causal relationship”. Events marked with missing relationship with study device is considered as device-related AE.

Serious Adverse Event (SAE)

A Serious Adverse Event is any AE that:

- Resulted in death: An AE that resulted in the patient's death.
- Life-threatening illness or injury: The patient was at imminent risk of dying at the time of the AE.
- Permanent Impairment: An AE that resulted in permanent impairment of a body function or permanent damage to a body structure.
- Required in-patient or prolonged hospitalization. Note: If a patient has prolonged hospitalization due to having an additional ablative procedure to enhance the ablative margin and for no other adverse event, this should not be considered an SAE.
- Resulted in medical or surgical intervention to prevent life threatening illness or injury or permanent impairment to a body or body function
- Led to fetal distress, fetal death or congenital abnormality or birth defects.
- A persistent or significant disability or incapacity

SAE is collected in "Adverse Events Log" CRF page with answer "Yes" to question "Is this adverse event Serious".

Procedure related SAE

A procedure related SAE meets both procedure-related AE and SAE definitions.

Study device related SAE

A study device related SAE meets both device-related AE and SAE definitions.

AE by Maximum Severity

Based on investigator's assessment, the severity of adverse events is classed as mild, moderate, or severe. Only the highest reported severity of a given adverse event is counted for the individual subject.

AE leading to Death

AE leading to death is collected in "Adverse Events Log" CRF page with answer "Fatal" to question "Outcome".

Procedure related AE Leading to Death

A procedure related AE leading to death meets both procedure related AE and AE leading to death definitions.

Study device related AE Leading to Death

A study device related AE leading to death meets both study device related AE and AE leading to death definitions.

SIR Classification System

A major complication is defined as an event that lead to substantial morbidity and disability that increases the level of care, or results in hospital admission, or substantially lengthens the hospital stay.

AE complications are classed by outcome by SIR Classification System as following:

- A. No therapy, no consequence
- B. Nominal therapy, no consequence; includes overnight admission for observation only
- C. Require therapy, minor hospitalization (< 48 hours)
- D. Require major therapy, unplanned increase in level of care prolonged hospitalization (> 48 hours)
- E. Permanent adverse sequelae
- F. Death.

The minor complications (A、B) and the major complcations (C-F) will be presented by preferred term.

8.5.2 Death

If any subjects that report death during the study as recorded on the “Death” CRF page, the information will be presented in a data listing

8.5.3 Laboratory Evaluations

Change from baseline at post ablation visit (V3) will be summarized.

Shift from baseline table of laboratory evaluations will be summarized descriptively by visit for each category in.

The laboratory tests are:

- Hematology (Complete Blood Count):
    - Red blood cell (RBC)

  - White blood cell (WBC)
  - Neutrophils (NEUT)
  - Lymphocytes (LYM)
  - Monocytes (MONO)
  - Eosinophils(EOS)
  - Basophils
  - Platelet count
- Chemistry (Liver Function and Renal Function)
  - Aspartate Aminotransferase (AST)
  - Alanine Aminotransferase (ALT)
  - Gamma-Glutamyl Transferase (GGT)
  - Albumin
  - Indirect bilirubin
  - Direct bilirubin
  - Total bilirubin
  - Total protein
  - Urea
  - Creatinine
  - Sodium
  - Potassium
  - Chloride
- Coagulation:
  - Prothrombin Time (PT)
  - Activated Partial Thromboplastin Time (APTT)
  - International Normalized Ratio (INR)

- Other:
    - Alpha-Fetoprotein (AFP)
    - Human Chorionic Gonadotropin (HCG), Quantitative, Serum

8.5.4 ECOG Performance Status

ECOG Performance Status will be summarized descriptively by visit for each score category. A shift table of ECOG from baseline to worst post-baseline ECOG score will be presented. A listing will also be prepared.

8.5.5 Numeric Pain Rating Scale (NPRS) Pain Score

NPRS score and change from baseline will be summarized descriptively by visit. A listing will also be prepared.

8.6 **Plans for Interim Analysis**

There will be 2 interim analyses.

The first interim analysis will occur after all enrolled subjects have completed the 1-month visit and is intended to provide an initial estimate of device effectiveness for Technical Success and Technical Efficacy as well as to summarize the peri-operative out to 1-month post-operative safety profile of subjects undergoing microwave ablation.

The second interim analysis will occur after all enrolled patients have completed the 12-month visit and will include a summary of local tumor progression rates as well as safety through one year.8.7 Handling of Missing Data

There will be no imputation of data for early terminated subjects or for enrolled subjects who do not provide a measurement at a given visit.

*8.8 **Adjustments for Multiplicity***

Not applicable.

*8.9 **Subgroup Analysis***

The subjects in Full Analysis Set are grouped by the number of ablation procedures:

- The initial only group contains all subjects who only receive the initial, planned ablation procedure at Visit 2A.

- The repeated group contains all subjects who receive a repeat ablation in the first 30 days

Subgroup analysis for Tumor type(primary HCC and recurrent HCC and other) for technical success and technical efficacy (analyzed for patient level and tumor level).

For subgroup analysis, the percentage and 95% CI will be provided by using forest plot.

8.10 ***Assessment of Site Homogeneity***

Technical Success and Technical Efficacy analysis will be performed by sites. The percentage and 95% CI will be provided by using forest plot. (by patient level and tumor level)

9 Data Monitoring Committee (DMC)

Not applicable.

End of Document

Document Approval Signature(s)

This document has been signed electronically in compliance with IQVIA and IQVIA Companies' electronic signature policies and procedures. This page is a manifestation of the electronic signatures, which may be confirmed within the IQVIA and IQVIA Companies' source system in accordance with IQVIA and IQVIA Companies' records retention policies. Note: the Title field in the signature block has been intentionally left blank.

UserName: Zheng, Boran (q829009)
Title:
Date: Wednesday, 28 April 2021, 06:08 AM Eastern Time
Meaning: I Approve the Document
==========

UserName: Zhang, Aiqin (q850404)
Title:
Date: Wednesday, 28 April 2021, 06:11 AM Eastern Time
Meaning: I Approve the Document
==========