ClinicalTrials.gov Document Cover Page

Statistical Analysis Plan

Study CS1003-305

Study Title: A Multi-Center, Double-Blind, Randomized, Phase III Study to Investigate the Efficacy and Safety of CS1003 in Combination with Lenvatinib Compared to Placebo in Combination with Lenvatinib as First-Line Therapy in Subjects with Advanced Hepatocellular Carcinoma (HCC)

Document Date: 02-Feb-2024

NCT Number: NCT04194775

STATISTICAL ANALYSIS PLAN

TITLE:	**A Multi-Center, Double-Blind, Randomized, Phase III Study to Investigate the Efficacy and Safety of CS1003 in Combination with Lenvatinib Compared to Placebo in Combination with Lenvatinib as First-Line Therapy in Subjects with Advanced Hepatocellular Carcinoma (HCC)**
PROTOCOL NO.:	**CS1003-305**
INVESTIGATIONAL PRODUCT:	**CS1003**
CURRENT VERSION NO.:	**Final v1.0**
SPONSOR:	**CStone Pharmaceuticals (Suzhou) Co., Ltd.**
PLAN PREPARED BY:	**Hangzhou Tigermed Consulting Co., Ltd.**
DATE FINAL:	**02Feb2024**
DATE AMENDED:	**NA**

SPONSOR APPROVAL PAGE

A Multi-Center, Double-Blind, Randomized, Phase III Study to Investigate the Efficacy and Safety of CS1003 in Combination with Lenvatinib Compared to Placebo in Combination with Lenvatinib as First-Line Therapy in Subjects with Advanced Hepatocellular Carcinoma (HCC)
CS1003-305

Statistical Analysis Plan

Version: Final v1.0

Prepared by:

[redacted] Date: [redacted]

Reviewed by:

[redacted] Date: [redacted]

Approved by:

[redacted] Date: [redacted]

Table of Contents

1. BACKGROUND 8
2. STUDY DESIGN 8
2.1 STUDY OBJECTIVES, ESTIMANDS AND ENDPOINTS 8
2.1.1 Primary Objective, Estimand, and Endpoint 8
2.1.2 Secondary Objectives, Estimands, and Endpoints 8
2.1.3 Exploratory Objectives and Endpoints 10
2.2 OVERALL STUDY DESIGN 10
2.3 DETERMINATION OF SAMPLE SIZE 11
2.4 PLANNED ANALYSES 12
3. STUDY CONDUCT 13
3.1 RANDOMIZATION 13
3.2 DATA MONITORING 13
4. STATISTICAL METHODS 13
4.1 ANALYSIS SETS 13
4.2 GENERAL STATISTICAL CONSIDERATIONS 14
4.2.1 Data Analysis General Principles and Definition 14
4.2.2 Methods for Handling Missing Data 15
4.3 ANALYSIS OF STUDY CONDUCT 15
4.3.1 Patient Disposition 15
4.3.2 Protocol Deviations 16
4.3.3 Demographics and Baseline Characteristics 16
4.3.4 Prior and Concomitant Medications 16
4.3.5 On-Study or Follow-up Anti-Cancer Treatment 17
4.3.6 Duration of follow-up time 17
4.4 EFFICACY ANALYSES 17
4.4.1 Primary Efficacy Endpoint 17
4.4.2 Key Secondary Efficacy Endpoints 19
4.4.3 Multiplicity 20
4.4.4 Other Secondary Efficacy Endpoints 20
4.4.5 Exploratory Efficacy Endpoints 21
4.4.6 Subgroup Analyses 22

4.5 SAFETY ANALYSES 22
4.5.1 Exposure of Study Medication 22
4.5.2 Adverse Events 23
4.5.3 Adverse Events of Special Interest 23
4.5.4 ADeath 24
4.5.5 Clinical Laboratory Evaluations 24
4.5.6 Vital Signs 25
4.5.7 Electrocardiograms (ECGs) 26
4.5.8 ECOG Performance Status 26
4.6 PHARMACOKINETIC ANALYSES 26
4.7 IMMUNOGENICITY ANALYSES 26
4.8 PATIENT-REPORTED OUTCOMES 27
Appendix 1 PROTOCOL SYNOPSIS 29
Appendix 2 SCHEDULE OF ASSESSMENT 42
Appendix 3 TIME POINTS OF BLOOD SAMPLING FOR CS1003 PK AND ADA 49
Appendix 4 HANDLING OF MISSING DATES 50
Appendix 5 CTCAE GRADES VERSION 5.0 FOR LABORATORY PHRAMETERS TO BE ANALYZED 52
Appendix 6 ADVERSE EVENTS OF SPECIAL INTEREST 56

List of abbreviations

Abbreviation	Definition
ADA	Anti-drug antibody
AE	Adverse events
AFP	Alpha fetoprotein
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
BCLC	Barcelona Clinic Liver Cancer
BICR	Blinded Independent Central Review Committee
CI	Confidence interval
CR	Complete response
CRF	Case report form
CRO	Contract research organization
CTCAE	Common Terminology Criteria for Adverse Events
DCR	Disease control rate
DoR	Duration of response
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic case report form
EDC	Electronic data capture
EGD	Esophagogastroduodenscopy
EHS	Extra-hepatic spread
EOI	End of infusion
EORTC	European Organization for Research and Treatment of Cancer
EOT	End of treatment visit
EQ-5D-5L	European quality of life 5-dimensions using 5-level scale
GHS	Global health status
HBcAb	Hepatitis B core antibody
HBsAg	Hepatitis B surface antigen
HBV	Hepatitis B virus
HCC	Hepatocellular carcinoma
HCV	Hepatitis C virus
HR	Hazard ratio
HRQoL	Health-related quality of life

Abbreviation	Definition
IC	Tumor Infiltrating Immune Cells
IDMC	Independent Data Monitoring Committee
IgG	Immunoglobulin G, for example IgG1, IgG2, IgG3 and IgG4; other immunoglobulins include IgM, IgD, etc.
IMAS	Immunogenicity analysis set
INR	International Normalized Ratio
IO	Immuno-oncology
IP	Investigational product
ITT	Intention-to-treat
IVRS	Interactive Voice Response System
IWRS	Integrated Web Response System
LLQ	Lower level of quantification
MedDRA	Medical Dictionary for Regulatory Activities
mg	milligram
mL	milliliter
ms	millisecond
msec	millisecond
MRI	Magnetic resonance imaging
MVI	Macroscopic vascular invasion
Nab	Neutralizing antibodies
NCI-CTCAE	National Cancer Institute- Common Terminology Criteria for Adverse Events
ORR	Objective response rate
OS	Overall survival
PD	Disease progression
PD-L1	Programmed death ligand-1
PK	Pharmacokinetics
PKAS	Pharmacokinetic analysis set
PPK	Population pharmacokinetics
PR	Partial response
PRO	Patient-reported outcome
PT	Preferred Term
Q3W	Every 3 weeks
QLQ-C30	Quality of Life Questionnaire Core 30
QLQ-HCC18	Quality of Life Questionnaire for Hepatocellular Carcinoma 18

Abbreviation	**Definition**
QTcF	Fridericia corrected QT interval
RECIST	Response Evaluation Criteria in Solid Tumors
ROW	Rest of the world
SAE	Serious adverse event
SAS	Safety analysis set
SD	Stable disease
SI	Standard international
SOC	System organ class
TBIL	Total bilirubin
TC	Tumor Cells
TEAE	Treatment emergent adverse event
TTD	Time to deterioration
ULN	Upper limit of normal
ULQ	Upper level of quantification
WHO	Word Health Organization
WHO-Drug	World Health Organization Drug Dictionary

1. BACKGROUND

This statistical analysis plan (SAP) describes the analysis for Protocol CS1003-305, a Multi-Center, Double-Blind, Randomized, Phase III Study to Investigate the Efficacy and Safety of CS1003 in Combination with Lenvatinib Compared to Placebo in Combination with Lenvatinib as First-Line Therapy in Subjects with Advanced Hepatocellular Carcinoma (HCC), version 5.0, which was issued on 24Nov2022.

This SAP provides a comprehensive and detailed description of the strategy, rationale, and statistical techniques to evaluate the efficacy and safety endpoints of the study. The purpose of this SAP is to ensure that the study findings are obtained based on pre-specified statistical approaches to the analysis of study data prior to database lock.

2. STUDY DESIGN

2.1 STUDY OBJECTIVES, ESTIMANDS AND ENDPOINTS

2.1.1 Primary Objective, Estimand, and Endpoint

Primary Objective:	Primary Endpoint/Estimand:
• To compare the efficacy of CS1003 in combination with lenvatinib vs. placebo in combination with lenvatinib	• Primary endpoint: Overall survival (OS) • Estimand: The primary objective will be evaluated under the treatment policy strategy to reflect the effect of the initially assigned randomized study treatment, irrespective of intercurrent events such as discontinuing treatment, withdrawing from or dropping out of the study, and starting non-protocol anti-cancer treatment. The population-level summary for OS will be estimated by the HR from stratified Cox regression model together with its corresponding 95% CI and stratified log-rank p-value for the ITT set.

2.1.2 Secondary Objectives, Estimands, and Endpoints

Secondary Objectives:	Secondary Endpoints:
• To compare the efficacy of CS1003 in combination with lenvatinib vs. placebo in combination with lenvatinib	Key secondary endpoints: • Objective response rate (ORR) assessed by Blinded Independent Central Review Committee (BICR) • Estimand: ORR assessed by BICR will be evaluated under the modified while-on-treatment strategy using tumor assessments prior to initiating non-protocol anti-cancer treatment to reflect the effect of the initially

	assigned randomized study treatment. The population-level summary for ORR evaluated by BICR will be presented by the rate difference between the two groups and the corresponding 95% CI using normal approximation to the binomial distribution, and p-value calculated by Stratified Mantel-Haenszel test for the ITT set. • Progression-free survival (PFS) assessed by BICR based on Response Evaluation Criteria in Solid Tumors (RECIST) v1.1 • Estimand: PFS assessed by BICR will be evaluated under the treatment policy strategy to reflect the effect of the initially assigned randomized study treatment, irrespective of intercurrent events such as discontinuing treatment before BICR assessed disease progression and starting of non-protocol anti-cancer treatment before disease progression. The population-level summary for PFS evaluated by BICR will be estimated by the hazard ratio (HR) from stratified Cox regression model together with its corresponding 95% CI and stratified log-rank p-value for the ITT set. Other secondary endpoints • PFS evaluated by investigators based on RECIST v1.1 • ORR evaluated by investigators based on RECIST v1.1 • Duration of response (DoR) and disease control rate (DCR) evaluated by BICR and investigators based on RECIST v1.1
• To compare the safety of CS1003 in combination with lenvatinib vs. placebo in combination with lenvatinib	• Adverse events (AEs), serious adverse events (SAEs), death, vital signs (including blood pressure and pulse, etc.), electrocardiogram (ECG), laboratory tests (including serum chemistry, hematology, and urinalysis, etc.) and ECOG performance status, etc.
• To characterize the population pharmacokinetics (PopPK) and immunogenicity of CS1003 when dosed in combination with lenvatinib	• Peak and trough serum concentrations of CS1003 • Number and percentage of subjects who develop anti-CS1003 antibody (ADA)
• To assess the disease/treatment-related patient-reported outcome (PRO), Health related Quality of Life (HRQoL)/Global health status (GHS) and function of	• Time to deterioration (TTD), defined as the time from randomization to the first deterioration of European Organization for the Research and Treatment of Cancer (EORTC) Quality-of-Life Questionnaire-Core 30 (QLQ-C30) scale

subjects treated with CS1003 in combination with lenvatinib compared to placebo in combination with lenvatinib	

2.1.3 Exploratory Objectives and Endpoints

Exploratory objectives:	Exploratory endpoints:
• To explore the relationship between biomarker and the efficacy of CS1003 in combination with lenvatinib	• Association of baseline PD-L1 expression with the efficacy of CS1003 in combination with lenvatinib
• To evaluate the disease/treatment-related PRO, HRQoL/GHS and function of subjects treated with CS1003 in combination with lenvatinib compared to placebo in combination with lenvatinib	• The scores of EORTC QLQ-C30 and EORTC QLQ-HCC18 subscales and the changes from baseline by cycle
• To evaluate the health status of subjects treated with CS1003 in combination with lenvatinib compared to placebo in combination with lenvatinib	• The health utilities index of EQ-5D-5L scale

2.2 OVERALL STUDY DESIGN

This study is a multi-center, double-blind, randomized, phase III study to investigate the efficacy and safety of CS1003 in combination with lenvatinib compared to placebo in combination with lenvatinib in the treatment for subjects with advanced HCC, who has no prior systematic treatment for HCC and is not amenable for surgery or locoregional therapy or have progressed after surgery or locoregional therapy.

Approximately 534 subjects are planned to be enrolled in this study. Eligible subjects will be randomized in a 2:1 ratio into the following two groups according to stratified block randomization method by Interactive Voice Response System (IVRS) and/or Interactive Web Response System (IWRS):

Group A: CS1003 in combination with lenvatinib

Or

Group B: Placebo in combination with lenvatinib

Randomization will be based on the following stratification factors:

- [redacted]

[redacted]

Subjects randomized to group A/B will receive CS1003 or placebo in combination with lenvatinib:

- CS1003 (200 mg) or placebo is administered by i.v. infusion, Q3W.
- Lenvatinib is orally administered at 12 mg or 8 mg, with or without food, once daily, at approximately the same time every day (refer to the prescription information of lenvatinib).

All subjects will be treated until disease progression (PD) or unacceptable toxicity or the withdrawal from the study due to other causes, whichever occurs first (refer to Protocol section 5.7.1).

Tumor assessments will be performed every 6 weeks (±7 days) after the randomization for the first 48 weeks, and then every 12 weeks (±7 days) thereafter, regardless of treatment schedule, until PD confirmed by radiological examination per RECISIT v1.1, subject withdraws informed consent, lost to follow-up, end of study, or death, whichever occurs first. Subjects who discontinue study treatment due to causes other than PD will continue to receive radiological assessments at pre-determined time points until PD, subject withdraws informed consent, lost to follow-up, end of study, or death, whichever occurs first.

Dosing adjustment for CS1003/placebo and lenvatinib may occur according to the guidance described in Protocol sections 6.6.1 and 6.6.2.

2.3 DETERMINATION OF SAMPLE SIZE

The primary endpoint of this study is OS.

Number of events and sample size are estimated based on the following hypotheses:

- Randomization ratio is 2:1.
- Type I error is controlled for two-sided test at 0.05.
- OS will be tested with log-rank test.
- [redacted]
- One interim analysis of OS will be conducted. Type I error will be controlled using Lan-DeMets approximation to the O'Brien-Fleming boundary.
- Survival times are exponentially distributed.
- [redacted]
- [redacted]

Based on the above hypotheses, [redacted]

[redacted]

ORR and PFS will be sequentially tested if OS is statistically significant. [redacted]

2.4 PLANNED ANALYSES

[redacted]

The OS will be tested using log-rank test with the following boundaries:

[redacted]

IDMC

An independent data monitoring committee (IDMC) will review the safety data collected in the trial regularly about every 3 months for the first two meetings from the first subject enrollment, and the following frequency will be adjusted based on the available safety data, e.g. every 6 months. The IDMC will also monitor the data for OS interim analysis and make recommendations to the Sponsor according to the pre-defined boundaries. Members of the IDMC are from outside the Sponsor and must follow the charter about duties and responsibilities. Any safety monitoring result that may affect the conduct of this trial must be notified to the investigator so that Ethics Committee (EC) will be informed. Membership, roles and responsibilities, and implementation details of IDMC will be described in the IDMC charter.

3. STUDY CONDUCT

3.1 RANDOMIZATION

Eligible subjects will be randomized (at 2:1) using the IVRS/IWRS into one of the two treatment groups.

Randomization will be conducted based on the following stratification factors:

3.2 DATA MONITORING

The IDMC will be established for the purpose of monitoring the safety of investigational treatment and OS interim analysis. IDMC consists of independent experts (including at least one oncologist and one statistician) who are not involved in the study, will carry out evaluation based on acceptable benefit/risk ratio, and make recommendations to the Sponsor on issues such as whether the study should continue, unchanged, be stopped, or be modified in any way. Once the IDMC has reached a recommendation, a report will be provided to the Sponsor. The report will include the recommendation and any potential protocol amendments and will not contain any unblinding information.

The final decision to modify or stop the study will sit with the Sponsor. The Sponsor or IDMC may call additional meetings if at any time there is concern about the safety of the study. All data provided to IDMC for official safety and/or efficacy analyses will be kept blinded to the Sponsor. Refer to the IDMC charter for the constitution, roles and responsibilities, and execution of IDMC.

4. STATISTICAL METHODS

4.1 ANALYSIS SETS

Intent-to-treat (ITT) set: All randomized subjects; these subjects will be grouped by the assigned treatment at randomization. Subject disposition, demographics, and other baseline characteristics, OS, ORR, PFS, and DCR analyses will be based on the ITT set. DoR analysis will be based on patients in the ITT set who achieve an objective response. TTD analysis is for ITT subjects with both a non-missing baseline PRO assessment and at least one post-baseline PRO assessment.

Safety analysis set (SAS): All subjects who have received at least one dose of either study treatment. Randomized subjects who haven't received any dose of study treatment will not be included in the SAS. For safety analysis, subjects will be grouped according to the actually received treatment (patients who receive any dose of CS1003 incorrectly will be counted in CS1003-treated group). The safety analyses will be based on the SAS.

Pharmacokinetic analysis set (PKAS): All subjects who receive at least one dose of CS1003 per the protocol, for whom there is at least one available PK concentration data, will be included in the PK analysis set.

Immunogenicity analysis set (IMAS): All subjects who have received at least one dose of CS1003 administration and have at least one reportable ADA result.

4.2 GENERAL STATISTICAL CONSIDERATIONS

4.2.1 Data Analysis General Principles and Definition

The below-mentioned general principles will be followed throughout the study.

Descriptive statistics will be used for all variables, as appropriate. Continuous variables will be summarized by the number of observations, mean, standard deviation (SD), median, minimum, and maximum. Categorical variables will be summarized by frequency counts and percentages for each category.

Reporting conventions will be as follows unless otherwise specified:

- Number of patients in the treatment group (N) and number of patients with non-missing values (n) will always be displayed to 0 decimal places.
- Mean and median will be displayed to one more decimal place than the raw data.
- SD will be displayed in two more decimal places than the raw data.
- The number of decimal places of minimum and maximum will be the same as the raw data.
- Percentages will be displayed to 1 decimal place. However, the percentage will not be displayed when the count equals 0.
- The maximum number of decimal places will be 4.
- In statistical test, if P value is equal to or greater than 0.0001 then 4 decimal places will be retained; and if P value is less than 0.0001, it will be expressed as "< 0.0001".

dThe data will be summarized by treatment group.

SAS® version 9.4 (as a minimum) will be used for all analyses.

4.2.1.1 Investigational Products/Study treatments

In this study, CS1003 (or placebo) and lenvatinib are both considered as the study treatments while CS1003 (or placebo) is the investigational product, and lenvatinib is selected as the basic treatment for HCC. Eligible subjects are randomized to one of the 2 groups:

- Experimental group: CS1003 in combination with lenvatinib
- Control group: placebo in combination with lenvatinib

CS1003 is a monoclonal antibody drug which is liquid for single use and diluted with normal saline (0.9% sodium chloride injection) or 5% dextrose prior to i.v. administration.

4.2.1.2 Baseline Values

For efficacy evaluations, the last adequate assessment on or before the date of randomization will be used as the baseline.

For summaries of baseline characteristics based on the ITT, baseline will be defined as follows:

- For subjects randomized and not dosed: the last assessment on or before the date of randomization
- For subjects randomized and dosed: the last assessment on or before the first dose of the study treatment.

For safety evaluations, the last assessment on or before the first dose of the study treatment will be used as the baseline.

For other evaluations(if appropriate), baseline will be the last non-missing observation prior to the first dose of the study treatment.

4.2.2 Methods for Handling Missing Data

Dates in the listings will be presented as stated on the eCRF.

Missing data on efficacy will be handled and censored as follows:

Subjects without death reports or missing death dates will be censored at the date of last survival confirmation in the OS analysis. Subjects without any data after baseline will be censored at the date of randomization.

For ORR analysis, subjects who are included in the analysis without any measurable disease at baseline or any tumor assessment after baseline will be considered as non-responders.

For TTD analysis, patients without events will be censored at the date of the last PRO assessment.

The approach regarding handling missing response assessments and censoring are presented in the *table 4.4.2*.

Detailed date imputation guidelines of concomitant medications, adverse events and death dates are presented in Appendix 4.

4.3 ANALYSIS OF STUDY CONDUCT

4.3.1 Patient Disposition

The number of patients who sign the informed consent form, the number of patients who fail screening and the number of patients who succeed in screening but are not randomized will be calculated. The number of randomized patients and the number of patients who have used the study treatments will be summarized by treatment group.

Based on all randomized patients (ITT set), the number and percentage of patients who discontinue treatment, the number and percentage of patients who discontinue treatment for any reason, the number and percentage of patients who withdraw from the study, and the number and percentage of patients who withdraw from the study for any reason will be calculated, with corresponding reasons listed.

The number and percentage of patients who are included in the ITT, SAS, PKAS, and IMAS will be calculated, respectively.

4.3.2 Protocol Deviations

The important protocol deviations will be summarized and listed for ITT set. Protocol deviations due to COVID-19 will also be summarized and listed.

4.3.3 Demographics and Baseline Characteristics

The following will be summarized for all patients in the ITT set (unless otherwise specified) by treatment group:

Demographics, including age, sex, race, geographic region, ethnicity, height, weight, baseline ECOG, etc.

Baseline disease characteristics, including AFP at baseline (< 400 ng/mL vs. ⩾ 400 ng/mL), Hepatitis virus (HBV vs. HCV vs. no hepatitis virus infection), Child-Pugh score (5, 6), BCLC stage(B, C), Tumor diagnosis at screening, HCC diagnosis criteria, Histological grade, Macroscopic vascular invasion (MVI) status, Extra-hepatic spread (EHS), MVI and/or EHS, Location of Metastatic Disease.

Prior anti-cancer treatment, including systemic anti-cancer therapy, cancer-related surgery/procedure, radiotherapy, etc.

General medical history other than cancer will be coded by Medical Dictionary for Regulatory Activities (MedDRA) and summarized by system organ class (SOC) and preferred term (PT).

4.3.4 Prior and Concomitant Medications

Medications received prior to or concomitantly with study drug will be coded using WHO Drug Dictionary.

Prior medications and concomitant medications will be summarized separately following below definitions:

- Prior medications are those taken prior to study treatment with a stop date prior to the first dose of study treatment.
- Concomitant medications are those with a stop date on or after the first dose date of study treatment (and could have started prior to or during treatment).

Missing coding terms should be listed and summarized as "Not coded".

4.3.5 On-Study or Follow-up Anti-Cancer Treatment

On-study or follow-up anti-cancer treatment data includes follow-up systemic anti-cancer therapy, on-study or follow-up radiotherapy and cancer-related surgery/procedure.

On-study or follow-up anti-cancer treatment will be summarized based on the ITT set, presence of systemic anti-cancer therapy, radiotherapy and cancer-related surgery/procedure, if any systemic anti-cancer agents, then they will be summarized based on level 2 of ATC, level 4 of ATC and PT.

On-study or follow-up anti-cancer treatment will be listed by treatment group based on the ITT.

4.3.6 Duration of follow-up time

Duration of follow-up refers to the time interval between the start of randomization and the last known alive date or death date, whichever occurs later. Patients with death reports before data cutoff will be censored on the day of death. For patients without any data after baseline, the date of randomization will be considered as event date. The median duration and quantile will be estimated using Kaplan-Meier method.

4.4 EFFICACY ANALYSES

Unless otherwise specified, the analyses of efficacy endpoints will be based on ITT set.

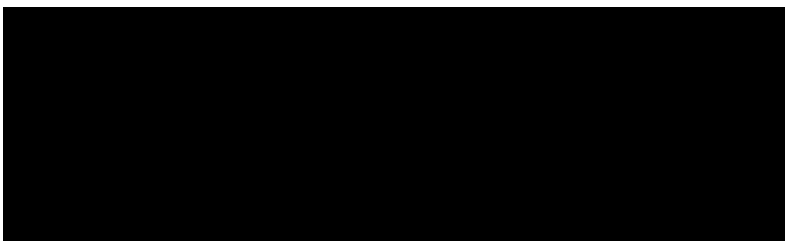
4.4.1 Primary Efficacy Endpoint

The primary efficacy endpoint of this study is OS.

Overall survival (OS): defined as the time interval between the date of randomization to the date of death from any cause. Patients who are alive at the time of analysis will be censored at the last known date of alive. Patients without any data after baseline will be censored on the randomization day.

The primary objective will be evaluated under the treatment policy strategy to reflect the effect of the initially assigned randomized study treatment, irrespective of intercurrent events such as discontinuing treatment, withdrawing further participation in the study but not the informed consent to further information disclosure, and starting of new anti-cancer treatment.

Comparison between treatment groups of OS will be performed using a two-sided stratified log-rank test at a significance level of 5%. Stratified Cox regression model will be used to estimate the HR (ties will be handled with Efron's method), and its 95% confidence interval (CI). Stratification factors will be those used during randomization from IVRS/IWRS, the sequence of stratification factors is ordered below：



There will be a total of 24 strata based on the four stratification factors, and if the number of patients in any stratum doesn't meet the requirement of analytical method, e.g. less than 15 patients, the level of this stratum will be collapsed in the inverse order until the minimum 15 patients criteria is achieved. The detailed process is as follows:

Kaplan-Meier method will be used to estimate the median OS, OS rates at different time points, and their 95% CIs (95% CIs of median OS by Brookmeyer Crowley method, 95% CIs of OS rates at different time points by Greenwood method). A Kaplan-Meier curve will be provided to visually describe OS changes over time.

Sensitivity Analysis

The following sensitivity analyses for OS will be performed to evaluate the robustness of primary analysis:

- OS will be analyzed using the log-rank test and Cox regression stratified by the stratification factors collected by EDC.
- OS will be analyzed using the unstratified method.
- OS will be tested using the same statistical method as the primary analysis except for patients who used PD-1/PD-L1 checkpoint inhibitors subsequent anti-cancer treatment prior to death, who will be censored on the day they receive PD-1/PD-L1 checkpoint inhibitors as subsequent anti-cancer treatments.
- The impact of COVID-19 on OS will be evaluated as follows: if the patient had not taken the investigational product as planned due to COVID-19 (and related quarantine measures) and had not restarted the treatment till 12 weeks afterward, then the patient would be regarded as being impacted by COVID-19, and the date of the last dose before being impacted plus 3 weeks would be considered as the start date of the investigational product administration being affected by the pandemic (refer to as "impact start date" hereinafter). Similar rule will be used for lenvatinib, and if the patient had not taken lenvatinib as planned due to COVID-19 (and related quarantine measures) and had not

restarted the treatment till 30 days afterward, then the patient would be regarded as being impacted by COVID-19 and the date of the last dose of lenvatinib before being impacted plus 1 day would be considered as the "impact start date". If the administration of investigational product and lenvatinib were both impacted by COVID-19, then the earlier "impact start date" would be regarded as the actual "impact start date" for the patient. For this OS sensitivity analysis, patients with the above situation will be censored on the "impact start date".

- OS will be analyzed using stratified methods based on a smaller pre-specified subset of the randomization stratification factors.

4.4.2 Key Secondary Efficacy Endpoints

The key secondary efficacy endpoints of this study include ORR and PFS assessed by BICR based on RECIST v1.1.

- **Objective response rate (ORR) assessed by BICR based on RECIST v1.1:** the number and proportion of patients who achieve objective response (CR or PR) assessed by BICR will be summarized. Patients without any measurable disease at baseline or any tumor assessment after baseline will be considered as non-responders. The analysis will be based on ITT set. Stratified Mantel-Haenszel test will be used to compare the ORR between groups. Normal approximation to binomial distribution will be used to calculate 95% CI for ORR difference between groups. Clopper-Pearson method will be used to calculate the ORR and its 95% CI of each group.

- **PFS assessed by BICR based on RECIST v1.1:** defined as the time from the date of randomization to PD assessed by BICR or death, whichever occurs first. The outcome (event or censor), and the date of event or censoring to be considered are presented in the below table. The statistical analysis method will be the same as that for the primary efficacy endpoint.

Table 4.4.2 Evaluation and Censoring of PFS and DoR

Situation	Outcome	Date of Event or Censoring
PD documented between/on scheduled response assessments	Progressed	Date of first overall response of PD
Death between/on adequate response assessments	Progressed	Date of death
Death or documented PD after missing 2 or more consecutive scheduled tumor assessments (for patients with at least one post-baseline tumor assessment)	Progressed	Date of first overall response of PD or death

No documented PD or death before data cutoff date	Censored	Date of last adequate radiologic assessment (not NE or missing)
Study discontinuation for undocumented progression	Censored	Date of last adequate radiologic assessment (not NE or missing)
Alternate anticancer treatment started before documented PD or death	Progressed	Date of first overall response of PD or death
No adequate post-baseline response assessment and no death prior to first scheduled tumor assessment	Censored	Date of randomization
No baseline assessment and no death	Censored	Date of randomization

4.4.3 Multiplicity

The primary and key secondary endpoints will be tested at an overall 2-sided Type I error rate at 5% level based on the fixed sequence testing procedure at the time of primary analysis. These endpoints will be tested in the following order:

- OS
- ORR assessed by BICR
- PFS assessed by BICR

Two analyses are planned for OS – one interim analysis and one final analysis (refer to SAP section 2.4 for more details). The predicted analysis timings of OS interim analysis and final analysis will be around 47 and 67 months respectively after the first subject is enrolled, and the data cutoff date of the key secondary endpoints will be the date when 387 PFS events are observed which is predicted to be around 38 months after the first subject is enrolled. Since the key secondary endpoints will be sequentially tested when OS is statistically significant, and the PFS events will for sure exceed 387 at either OS IA or FA timepoint, the data will be cut back to the time when the 387th PFS event occurs, and the cutback data will be used for the ORR and PFS analyses.

4.4.4 Other Secondary Efficacy Endpoints

The other secondary efficacy endpoints of this study include ORR and PFS assessed by investigators based on RECIST v1.1, DCR, and DoR assessed by BICR and investigators, and TTD.

- **ORR assessed by investigators based on RECIST v1.1:** the number and proportion of subjects who achieve objective response (CR or PR) assessed by investigators will be summarized. Patients without any measurable disease at baseline or any tumor assessment after baseline will be considered non-responders. The analysis will be based on ITT set. Stratified Mantel-Haenszel test will be used to compare the ORR between groups. Normal approximation to binomial distribution will be used to calculate 95% CI for ORR difference between groups. Clopper-Pearson method will be used to calculate the ORR and its 95% CI of each group.

- **PFS assessed by investigators based on RECIST v1.1** is defined as the time from the date of randomization to the first documented disease progression assessed by investigators or death, whichever occurs first. The outcome (event or censor), and the date of event or censoring to be considered are presented in the table 4.4.2.

- **Duration of response (DoR)** is defined as, among responders (CR or PR), the time interval between the date of the first occurrence of complete or partial response and the date of PD or death due to any cause, whichever occurs first. The censoring rule will be the same as that for the PFS analysis. The analysis method will be the same as that for the primary efficacy endpoint. Duration of response will be analyzed in ITT subjects with CR or PR response. All statistical analyses are descriptive.

- **Disease Control Rate (DCR)** is defined as the proportion of patients who achieve CR, PR, and SD based on RECIST v1.1. It will be analyzed similarly to ORR.

- **Time to deterioration (TTD)** is defined as the time from randomization to the first confirmed clinically meaningful deterioration of quality of life, physical functioning, and role functioning, respectively, as reported by the patient. Confirmed clinically meaningful deterioration is defined as a decrease from baseline of 10 points or more on the EORTC QLQ-C30 maintained for at least two consecutive assessments or a decrease from baseline of 10 points or more in one assessment followed by death from any cause within 3 weeks. TTD analysis will be conducted on all patients with both a non-missing baseline PRO assessment and at least one post-baseline PRO assessment and will include all data collected through disease progression and survival follow-up. The statistical analysis method will be the same as that for the primary efficacy endpoint. Patients without deterioration at the time of analysis will be censored at the last time they were known to have not deteriorated. There will be no imputation for missing baseline or post-baseline data for the TTD analysis.

PFS, DoR, and TTD will be analyzed according to the above methods in section 4.4.1, and DCR will be analyzed using the same method as ORR. The other secondary efficacy endpoints except TTD will be analyzed using the same data cutoff date as the key secondary endpoints.

4.4.5 Exploratory Efficacy Endpoints

To explore the relationship between biomarker and the efficacy of CS1003 in combination with lenvatinib, PD-L1 expression will be evaluated by percentage of tumor cells and tumor-infiltrating immune cells with positive staining of PD-L1 (Tumor Cells(TC)% or Tumor Infiltrating Immune Cells (IC)%, respectively). PD-L1 expression will be categorized as TC<1% or TC≥1% and IC<1% or IC≥1%. The correlation of PD-L1 expression with the efficacy endpoints will be compared

between patients with TC<1% vs. TC≥1% and patients with IC<1% vs. IC≥1%. According to data availability, all biomarker data will be listed, and analysis results will be presented by descriptive statistics in tables.

To evaluate the disease/treatment-related PRO, HRQoL/GHS, and function of patients treated with CS1003 in combination with lenvatinib compared to placebo in combination with lenvatinib, the descriptive statistics (e.g., mean, median, standard deviation, and range) of the absolute scores and change from baseline by cycle will be calculated for EORTC QLQ-C30 and EORTC QLQ-HCC18 and by treatment group. Refer to detailed description in section 4.8.

To assess the health status of patients treated with CS1003 in combination with lenvatinib compared to placebo in combination with lenvatinib, the health utilities index of EQ-5D-5L scale is calculated with descriptive statistics by treatment group.

4.4.6 Subgroup Analyses

Subgroup analyses will be performed to compare the primary endpoint between the treatment groups in the following subgroups based on the ITT set. The unstratified HR with its corresponding 95% CI will be estimated using the unstratified Cox regression model and depicted in the forest plot. The subgroups will include:

- Sex (Female, Male)
- Age group (<65 years, ≥65 years)
- Barcelona Clinic Liver Cancer (BCLC) stage at screening (B, C)
- Child-Pugh (5, 6)
- ECOG performance status (0, 1)
- Geographic region (Asia, ROW)
- Vascular invasion (yes, no)
- Extrahepatic metastasis (yes, no)
- Vascular invasion and/or extrahepatic metastasis (yes vs. no)
- AFP at baseline (< 400 ng/mL, ≥ 400 ng/mL)
- Hepatitis virus (HBV, HCV, no hepatitis virus infection)
- Hepatitis virus (yes, no)
- PD-L1 expression level (TC<1% vs. TC≥1% and IC<1% vs. IC≥1%)

The subgroup analyses will be based on values recorded on the eCRF. The results will be purely exploratory and are intended to evaluate the consistency of treatment effect.

4.5 SAFETY ANALYSES

Safety analyses will be performed on the safety analysis set unless otherwise specified.

4.5.1 Exposure of Study Medication

The following will be summarized by treatment group:

- Number of cycles for CS1003/Placebo;

- Duration of treatment exposure in months, cumulative actual dose, and relative dose intensity for CS1003/Placebo and lenvatinib.

4.5.2 Adverse Events

All AEs will be coded using MedDRA dictionary. The severity will be graded using the National Cancer Institute (NCI) Common Terminology Criteria for Adverse Events (CTCAE), version 5.0.

Number of patients experiencing treatment-emergent adverse events (TEAEs) will be summarized. TEAE was defined as (regardless of intensity):

- AE occurred on or after study treatment initiation
- AE occurred before study treatment initiation but worsened on or after study treatment initiation

In NCI CTCAE version 5.0, the death caused by PD should be considered AE and the preferred term is "Disease progression". It will also be reported as SAE in this study. In this study, these AEs will be excluded from summary tables but reported in the listings only.

All TEAEs will be summarized by treatment group, including any TEAEs, treatment-related TEAEs, CS1003/Placebo related TEAEs, lenvatinib-related TEAEs, Grade ≥ 3 TEAEs, treatment-related Grade ≥ 3 TEAEs, CS1003/Placebo related Grade ≥ 3 TEAEs, lenvatinib related Grade ≥ 3 TEAEs, serious TEAEs, treatment-related serious TEAEs, CS1003/Placebo related serious TEAEs, lenvatinib related serious TEAEs, immune-related TEAEs assessed by investigator, infusion-related reaction, TEAEs leading to CS1003/Placebo dose rate reduction/interruption/withdrawal, TEAEs leading to lenvatinib dose reduction/interruption/withdrawal, TEAEs leading to death, treatment-related TEAEs leading to death, CS1003/Placebo related TEAEs leading to death, lenvatinib related TEAEs leading to death.

In addition, serious TEAEs, Grade ⩾ 3 TEAEs, treatment-related TEAEs, treatment-related Grade ≥ 3 TEAEs, treatment-related serious TEAEs, immune-related TEAEs assessed by the investigator, infusion-related reactions, TEAEs leading to study treatment interruption/reduction (dose rate reduction for CS1003/Placebo)/withdrawal, TEAEs leading to death, treatment-related TEAEs leading to death, and treatment-related TEAEs leading to study treatment withdrawal will be summarized by SOC and PT. All TEAEs and treatment-related TEAEs will also be summarized by SOC, PT, and highest NCI-CTCAE grade. A patient experiencing multiple AEs under the same SOC or PT will be counted only once for that SOC or PT by maximum severity.

In the summary of AEs, the incidence of AEs will be the number of patients reporting AEs and not the number of AEs reported.

All AEs will be listed by treatment group and patient.

4.5.3 Adverse Events of Special Interest

Adverse Events of Special Interest (AESI) evaluated by the Sponsor will be summarized by the following:

a) By treatment group and AESI categories, including:
   - Patients with at least one AESI
   - Patients with at least one NCI-CTCAE Grade ≥ 3 AESI
   - Patients with at least one serious AESI
   - Patients with at least one immune-related TEAE assessed by the investigator
   - Patients with at least one AESI leading to investigational drug discontinuation/interruption
   - Patients with at least one AESI leading to death
   - Patients with at least one AESI treated with any Concomitant Medicine
   - Patients with at least one AESI treated with Systemic Corticosteroid
   - Patients with at least one AESI treated with High Dose of Systemic Corticosteroid

b) By treatment group, AESI categories, worst NCI-CTCAE Grade, and PT
c) By PT, outcome and treatment group
d) By treatment group
   - Other Concomitant Medications for AESI excluding Systemic Corticosteroid

In addition, time to onset and duration of AESIs will also be analyzed.

Please refer to Appendix 6 for predefined AESI categories with PTs.

4.5.4 ADeath

All deaths and the reason for deaths will be summarized by treatment based on safety analysis set.

In addition, all deaths will be listed by treatment group and patient.

4.5.5 Clinical Laboratory Evaluations

The laboratory data will be converted to Standard International (SI) units.

The absolute value and change from baseline of each lab parameter will be summarized by treatment group and scheduled visits. If there are multiple recorded for a specific visit, then the last reported value will be used for that visit.

Patient listing of abnormal laboratory data will be provided.

Grading of laboratory data

Laboratory data will be derived programmatically according to the NCI CTCAE version 5.0. The calculation of CTC grades will be based on the observed laboratory values only, clinical assessments will not be considered. The derivation rule of all laboratory parameters for NCI CTCAE version 5.0 is listed in Appendix 5. A shift table will be produced to reflect the grade change at post-baseline visits from baseline visits. Patients with baseline and post-baseline assessments available will be included in the corresponding parameters. The shift table in maximum CTCAE grades of worst (high) and worst (low) assessment will be presented separately.

Liver function parameters

The number and percentage of patients with Liver Function Laboratory Findings will be summarized based on the event criteria given in *Table 4.5.5*

Table 4.5.5: Liver-lab test findings

Parameter	Criterion
ALT	>=3xULN; >=5xULN; >=10xULN, >=20xULN
AST	>=3xULN; >=5xULN; >=10xULN; >=20xULN
ALT or AST	>=3xULN; >=5xULN; >=10xULN; >=20xULN
total bilirubin (TBIL)	>1.5xULN; >2xULN
ALP	>1.5xULN
ALT and/or AST & TBIL	ALT and/or AST>=3xULN & TBIL >1.5xULN ALT and/or AST>=3xULN & TBIL >2xULN
ALT and/or AST & TBIL & ALP	ALT and/or AST>=3xULN & TBIL >2xULN & ALP <2xULN

For a combined criterion to be fulfilled, all conditions must be fulfilled on the same visit. The criteria are not mutually exclusive, e.g., a subject with ALT = 6.42xULN is counted for ALT >=3xULN and ALT>=5xULN.

Handing of the laboratory values with reported sign ("<" or ">")

For laboratory tests values below Lower Level of Quantification (LLQ) or above Upper Level of Quantification (ULQ) will be imputed as LLQ or ULQ values, respectively. The numerical part of the reported result will be treated as the actual LLQ or ULQ. These laboratory values will be displayed in listings using the standard unit with the reported sign ("<" or ">").

4.5.6 Vital Signs

Vital signs (weight, systolic blood pressure, diastolic blood pressure, heart rate, and temperature) will be summarized at each scheduled visit for the safety analysis set. The change and percentage change from baseline will also be summarized for post-baseline visits.

4.5.7 Electrocardiograms (ECGs)

Overall Interpretation of ECG results will be summarized by visit.

In addition, number and percentage of patients with notable ECG values which are defined as those in the following categories will be presented.

- QTcF(msec) interval increase from baseline >30 ms, >60ms
- QTcF(msec) interval >450 ms, >480 ms, >500 ms

A separate listing of patients with clinically significant abnormal ECG results at any visit will be provided.

4.5.8 ECOG Performance Status

Shift table will be produced for ECOG performance status at post-baseline visits and baseline for patients in safety analysis set.

4.6 PHARMACOKINETIC ANALYSES

The PK data in this study will be pooled with data from other CS1003 clinical trials to develop a population PK model. This model will be used to evaluate internal and external variables and their influence on the PK of CS1003. Furthermore, exposure-response analysis will be performed on specific pharmacodynamic and safety endpoints. Results of the above population PK and exposure-response analyses will be recorded in separate reports.

4.7 IMMUNOGENICITY ANALYSES

The evaluation will be summarized as below definition based on immunogenicity analysis set.

Treatment-induced ADA positive: patients who had a baseline-negative ADA result and at least one positive ADA result at any time after first study drug administration.

Treatment-enhanced ADA positive: patients who had a baseline-positive ADA result and at least one enhanced titer result (by a scientifically reasonable margin ninefold greater than baseline ADA titer) at any time after first study drug administration.

Treatment-unaffected ADA: patients with positive ADA results at baseline and all post-baseline titer results are not greater than the baseline titer result (by a scientifically reasonable margin ninefold greater than baseline ADA titer); or patients with positive ADA results at baseline and all post-baseline results are negative or missing.

Unevaluable ADA: patients missing baseline results and at least one positive ADA result at any time after first study drug administration.

Overall ADA prevalence: the sum of all treatment-induce, treatment-enhanced, treatment-unaffected, and unevaluable ADA-positive patients as a proportion of the evaluable subject population (IMAS).

Overall ADA incidence: both treatment-induce and treatment-enhanced ADA-positive patients as a proportion of the evaluable subject population (IMAS).

For ADA positive patients, neutralizing antibodies (Nab) will be tested, and the number and percentage of patients will be summarized by:

- Nab results at baseline
- Nab results post-baseline

Patients with at least one post-baseline positive Nab result will be defined as post-baseline Nab positive, otherwise, defined as post-baseline Nab negative.

ADA and Nab results will also be listed.

4.8 PATIENT-REPORTED OUTCOMES

Patient-reported outcomes (PROs) will be evaluated in this study through the following 3 instruments: EORTC QLQ-C30, EORTC QLQ-HCC18, and the EQ-5D-5L. The EORTC QLQC30, EORTC QLQ-HCC18, and the EQ-5D-5L will be scored based on the instrument developer guidelines. No formal power calculation was done for the PRO analyses.

Analysis of PRO data will be performed on patients who have both a non-missing baseline assessment and at least one post-baseline assessment.

EORTC QLQ-C30

The EORTC QLQ-C30 is a valid and reliable self-reported questionnaire, it includes 30 items resulting in 5 functional scales (physical functioning, role functioning, emotional functioning, cognitive functioning, and social functioning), 1 global health status scale, 3 symptom scales (fatigue, nausea and vomiting, and pain), and 6 single items (dyspnea, insomnia, appetite loss, constipation, diarrhea, and financial difficulties).

EORTC QLQ-HCC18

The EORTC QLQ-HCC18 is a valid and reliable self-reported questionnaire used to assess hepatocellular carcinoma-specific symptoms (fatigue, body image, jaundice, nutrition, fever, pain, abdominal swelling, and sex life).

The EORTC QLQ-C30 and EORTC QLQ-HCC18 will be scored according to the EORTC scoring manual. In the event of incomplete data for any scales, if 50% or more of the constituent items are completed, a prorated score will be computed. For scales with less than 50% of the items completed, the score will be considered to be missing. On all scale scores, a within-patient change of 10 points or greater from baseline denotes a clinically meaningful change.

All of the EORTC scales and single-item measures will be linearly transformed so that each score has a range of 0−100. A high score for a functional scale represents a high or healthy level of functioning, and a high score for the global health status and HRQoL represents a high HRQoL; however, a high score for a symptom scale or item represents a high level of symptomatology or problems.

For EORTC QLQ-C30 and EORTC QLQ-HCC18, below summary analysis is conducted:

- Completion rates will be calculated as the number of patients who completed all the questions on a given questionnaire divided by the number of patients expected to complete the questionnaire (i.e., those still on study) at each scheduled visit. Completion rates will be calculated separately for the EORTC QLQ-C30 and the EORTC QLQ-HCC18.
- Proportion of patients with clinically meaningful changes (refer to SAP section 4.4.4) from baseline in selected scales of the EORTC QLQ-C30 (quality of life, physical functioning, role functioning, appetite loss, diarrhea, fatigue, and pain) and EORTC QLQ-HCC18 (fatigue, jaundice, and pain); Patients will be categorised as improved (if their scale score improved by ≥10 points), deteriorated (if their scale score deteriorated by ≥10 points), or stable (if their scale score changed by any other amount). The results will be summarised by treatment group and treatment cycle.
- Changes from baseline by treatment cycle in all scales of the EORTC QLQ-C30 and EORTC QLQ-HCC18 will be presented. Summary statistics will be calculated, including the number of patients, mean, SD, median, min, and max.

The Kaplan-Meier method will be used to estimate deterioration probabilities for each treatment group. The HRs and respective 95% CIs comparing CS1003 in combination with lenvatinib vs. placebo in combination with lenvatinib will be estimated using a stratified Cox regression model.

EQ-5D-5L

The EQ-5D-5L is a generic measure of health status. The EQ-5D-5L is a 5-item questionnaire that assesses 5 domains including mobility, self-care, usual activities, pain/discomfort, and anxiety/depression plus a visual analog scale rating "health today" with anchors ranging from 0 (worst imaginable health state) to 100 (best imaginable health state). The scores for the 5 separate questions are categorical and cannot be analyzed as cardinal numbers. However, the scores for the 5 dimensions are used to compute a single utility score ranging from zero (0.0) to 1 (1.0) representing the general health status of the individual.

Descriptive statistics will be calculated for each scheduled visit/time point in the study, for each treatment group. These will report the number of patients, the number of EQ-5D questionnaires completed at each visit, the number and proportion responding to each dimension of the EQ-5D-5L. Additionally, summary statistics (e.g. n, mean, median, SD, min, max) will be reported for the EQ-5D index score and the change from baseline for the EQ-5D index score.

Line plots of the mean EQ-5D index score, including change from baseline, and associated 95% CI by scheduled visits/time points in the study will be produced.

APPENDIX 1 PROTOCOL SYNOPSIS

<table>
<tr><td colspan="2">Sponsor: CStone Pharmaceuticals (Suzhou) Co. Ltd.</td></tr>
<tr><td colspan="2">Final Product: N/A</td></tr>
<tr><td colspan="2">Investigational product: CS1003 (or placebo)</td></tr>
<tr><td colspan="2">Active Substance: a humanized, recombinant IgG4 anti-programmed death-1 (anti-PD-1) monoclonal antibody</td></tr>
<tr><td>Title of Study:</td><td>A Multi-Center, Double-Blind, Randomized, Phase III Study to Investigate the Efficacy and Safety of CS1003 in Combination with Lenvatinib Compared to Placebo in Combination with Lenvatinib as First-Line Therapy in Subjects with Advanced Hepatocellular Carcinoma (HCC)</td></tr>
<tr><td>Protocol No.:</td><td>CS1003-305</td></tr>
<tr><td>Sites:</td><td>Approximately 74 sites globally</td></tr>
<tr><td>Duration of the Study:</td><td>Screening (up to 28 days), Treatment and Follow-up (including Safety Follow-up, up to 90 days, and Survival Follow-up, every 12 weeks until end of study)</td></tr>
<tr><td>Study Phase:</td><td>III</td></tr>
<tr><td colspan="2">Study Objectives, Estimands and Endpoints:
Primary objective, estimand and primary endpoint:
<table>
<tr><td>Primary Objective:</td><td>Primary Endpoint/Estimand:</td></tr>
<tr><td>• To compare the efficacy of CS1003 in combination with lenvatinib vs. placebo in combination with lenvatinib</td><td>• Primary endpoint: Overall survival (OS)
• Estimand: The primary objective will be evaluated under the treatment policy strategy to reflect the effect of the initially assigned randomized study treatment, irrespective of intercurrent events such as discontinuing treatment, withdrawing from or dropping out of the study, and starting of non-protocol anti-cancer treatment. The population-level summary for OS will be estimated by the HR from stratified Cox regression model together with its corresponding 95% CI and stratified logrank p-value for the ITT set.</td></tr>
</table>
Secondary objectives and secondary endpoints:
<table>
<tr><td>Secondary Objectives:</td><td>Secondary Endpoints:</td></tr>
<tr><td>• To compare the efficacy of CS1003 in combination with lenvatinib vs. placebo in combination with lenvatinib</td><td>Key secondary endpoints:
• Objective response rate (ORR) assessed by Blinded Independent Central Review Committee (BICR)
• Estimand: ORR assessed by BICR will be evaluated under the modified while-on-treatment strategy using tumor assessments prior to initiating non-protocol anti-cancer</td></tr>
</table>
</td></tr>
</table>

	treatment to reflect the effect of the initially assigned randomized study treatment. The population-level summary for ORR evaluated by BICR will be presented by the rate difference between the two groups and the corresponding 95% CI using normal approximation to the binomial distribution, and p-value calculated by Stratified Mantel-Haenszel test for the ITT set. • Progression-free survival (PFS) assessed by BICR based on Response Evaluation Criteria in Solid Tumors (RECIST) v1.1 • Estimand: PFS assessed by BICR will be evaluated under the treatment policy strategy to reflect the effect of the initially assigned randomized study treatment, irrespective of intercurrent events such as discontinuing treatment before BICR assessed disease progression and starting of non-protocol anti-cancer treatment before disease progression. The population-level summary for PFS evaluated by BICR will be estimated by the hazard ratio (HR) from stratified Cox regression model together with its corresponding 95% CI and stratified log-rank p-value for the ITT set. Other secondary endpoints • PFS evaluated by investigator based on RECIST v1.1 • ORR evaluated by investigators based on RECIST v1.1 • Duration of response (DoR) and disease control rate (DCR) evaluated by BICR and investigators based on RECIST v1.1
• To compare the safety of CS1003 in combination with lenvatinib vs. placebo in combination with lenvatinib	• Adverse events (AEs), serious adverse events (SAEs), death, vital signs (including blood pressure and pulse, etc.), electrocardiogram (ECG), laboratory tests (including serum chemistry, hematology, and urinalysis, etc.) and ECOG performance status, etc.
• To characterize the population pharmacokinetics (PopPK) and immunogenicity of	• Peak and trough serum concentrations of CS1003 • Number and percentage of subjects who develop anti-CS1003 antibody (ADA)

CS1003 when dosed in combination with lenvatinib	
• To assess the disease/treatment-related patient-reported outcome (PRO), Health-related Quality of Life (HRQoL)/Global health status (GHS), and function of subjects treated with CS1003 in combination with lenvatinib compared to placebo in combination with lenvatinib	• Time to deterioration (TTD), defined as the time from randomization to the first deterioration of European Organization for the Research and Treatment of Cancer (EORTC) Quality-of-Life Questionnaire-Core 30 (QLQ-C30) scale

Exploratory objectives and exploratory endpoints:

Exploratory Objectives:	Exploratory Endpoints:
• To explore the relationship between biomarker and the efficacy of CS1003 in combination with lenvatinib	• Association of baseline PD-L1 expression with the efficacy of CS1003 in combination with lenvatinib
• To evaluate the disease/treatment-related PRO, HRQoL/GHS, and function of subjects treated with CS1003 in combination with lenvatinib compared to placebo in combination with lenvatinib	• The scores of EORTC QLQ-C30 and EORTC QLQ-HCC18 subscales and the changes from baseline by cycle
• To assess the health status of subjects treated with CS1003 in combination with lenvatinib compared to placebo in combination with lenvatinib	• The health utilities index of EQ-5D-5L scale

Study Design:

This study is a multi-center, double-blind, randomized, phase III study to investigate the efficacy and safety of CS1003 in combination with lenvatinib compared to placebo in combination with lenvatinib in the treatment of subjects with unresectable advanced HCC, not amenable for locoregional therapy and have no prior systematic treatment for HCC or have progressed after curative treatment and/or locoregional therapy.

Subjects who have completed all eligibility check and fulfilled all the inclusion criteria and none of the exclusion criteria will be randomized in a 2:1 ratio according to stratified block randomization method:

Group A: CS1003, intravenous (i.v.) administration every 21 days (3 weeks) (Q3W);

Lenvatinib 12 mg (if the subject's weight ⩾ 60 kg) or 8 mg (if the subject's weight < 60 kg), oral administration, once daily (refer to the prescription information of lenvatinib)

or

Group B: Placebo, i.v. administration every 21 days (3 weeks) (Q3W);

Lenvatinib 12 mg (if the subject's weight ⩾ 60 kg) or 8 mg (if the subject's weight < 60 kg), oral administration, once daily (refer to the prescription information of lenvatinib)

Subjects should take lenvatinib with or without food at approximately the same time each day.

See Section 6.6 for instructions on study treatment adjustment during the study.

Randomization will be based on the following stratification factors:

[redacted]

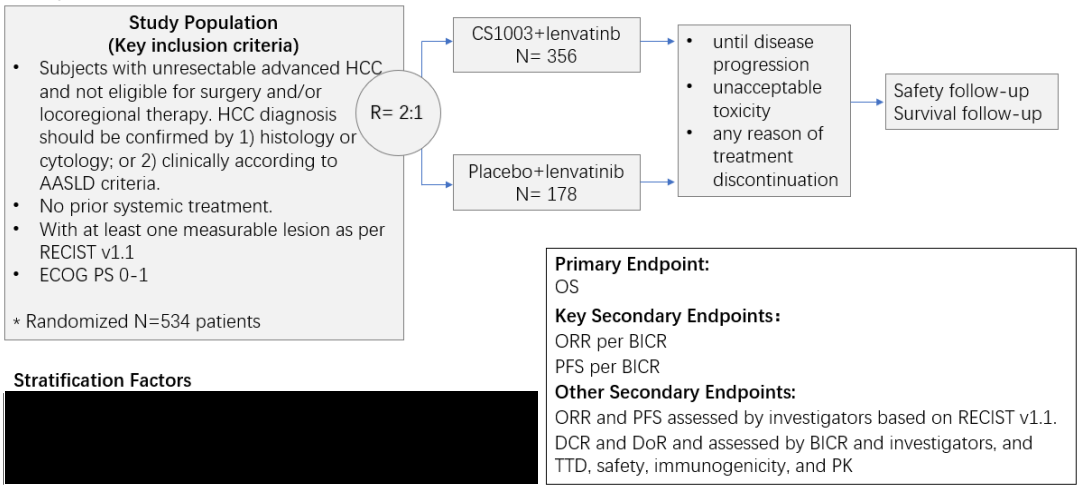
All subjects will be treated until disease progression (PD) or unacceptable toxicity or withdrawal from the study due to other causes, whichever occurs first. If the tumor assessment has indicated stable disease (SD) or greater benefit, and the study treatment has been tolerable, subjects who are still on treatment at the end of study may continue to receive study treatment with an expansion study or other forms after consultation with Sponsor, and after approval by the health authorities and Ethics Committees.

Anti-cancer response will be evaluated based on RECIST v1.1. Pseudo-progression may occur in subjects who receive treatment with immunotherapies such as CS1003. Therefore, if pseudo-progression is suspected by investigator, study treatment may continue until confirmation of PD with imaging repeated at least 4 weeks later or at the next scheduled imaging assessment, but not to exceed 9 weeks from the initial documentation of PD. Subject must be re-consented and study treatment continuation decision should be notified to the sponsor, and the following criteria must be met in order to continue the study treatment after initial PD.

- Absence of clinical symptoms and signs of disease progression (including worsening of laboratory results);
- The study treatment is tolerable to subject;
- Stable ECOG performance status (PS);

- Absence of rapid progression of disease or tumor at critical anatomical sites (e.g., spinal cord compression) that requires urgent medical intervention.

Study Schema:



Abbreviation:
AASLD: American Association for the Study of Liver Diseases; BICR: Blinded Independent Central Review Committee; PK: Pharmacokinetics; RECIST: Response Evaluation Criteria in Solid Tumors. All efficacy test are based on RECIST v1.1

Planned Number of Randomized Subjects:	A total of approximately 534 subjects (including approximately 356 in CS1003+lenvatinib group and 178 in placebo + lenvatinib group) are planned to be randomized.
Study Population:	**Subjects must meet all of the following inclusion criteria for this study:** **General inclusion criteria** 1. Age ⩾18 years on the day of signing informed consent (For Taiwan, the lower limit of age is 20 years). 2. Fully informed of the study, with good compliance and willing to provide written informed consent. The informed consent must be signed before performing any protocol-related procedure (that is not a part of subject's routine medical care). **Target Disease-Specific Inclusion Criteria** 3. Subjects with unresectable advanced HCC that is not eligible for surgery and/or locoregional therapy (Stage B or Stage C based on Barcelona Clinic Liver Cancer [BCLC] staging system indicated in Appendix 14.6), and meets either one of the following criteria: 1) histologically or cytologically confirmed diagnosis of HCC, 2) clinically confirmed diagnosis of HCC according to American Association for the Study of Liver Diseases (AASLD) criteria. Patients without cirrhosis require histological confirmation of diagnosis. The curative treatment and/or locoregional therapy must have been completed ⩾ 4 weeks prior to the baseline scan for

	subjects who experienced disease progression. 4. At least one measurable lesion can be assessed by the investigator based on Response Evaluation Criteria in Solid Tumors (RECIST) v1.1 (Appendix 14.2) within 14 days prior to randomization. Target lesions in the past radiation fields or that underwent locoregional therapy (e.g., transarterial chemoembolization [TACE] or ablation), if confirmed as radiographic progression, are considered as measurable lesions. 5. Eastern Cooperative Oncology Group performance status (ECOG PS) 0 or 1. 6. Life expectancy $\geqslant$ 3 months. 7. Child-Pugh $\leqslant$ 6 (Child-Pugh A). 8. No prior systemic treatment for advanced HCC, including target therapy, chemotherapy, immunotherapy (immune checkpoint inhibitors, interferons, interleukins), biological therapy (anti-cancer vaccines, cytokines, or growth factors), antibodies/drugs that target at any T cell co-regulatory protein (immune checkpoint), e.g., anti-PD-1, anti-PD-L1, anti-CTLA-4, anti-OX-40, anti-CD137, anti-TIM-3, anti-LAG-3 antibodies, etc. a) Interferon is only allowed for hepatitis treatment with a washout period of no less than 5 half-lives prior to the initiation of study treatment. 9. Subjects with hepatitis B virus (HBV) infection must have HBV DNA < 2000 IU/mL prior to randomization. Subjects with positive HBsAg and/or detectable HBV DNA must be treated with antiviral therapy for at least 2 weeks prior to the first dose of study treatment and must remain on antiviral therapy during the study. Subjects with hepatitis C virus (HCV) infection (active or past infection) are allowed to be enrolled and HCV RNA detectable subjects are allowed to receive approved anti-HCV treatment during the study. 10. After the approval from local health authorities is obtained (if applicable), at screening, subjects must agree to provide, if available, tumor tissue for biomarker analysis. When archival tumor tissue is not available, it is optional for the subject to undergo a fresh biopsy to collect tumor tissue if deemed medically safe by the investigator (details are in section 7.5). **Inclusion criteria related to laboratory tests at screening**

	Adequate organ and marrow function, as defined below. 11. Absolute neutrophil count (ANC) $\geqslant 1.5 \times 10^9$ /L 12. Platelet count $\geqslant 75 \times 10^9$ /L 13. Hemoglobin $\geqslant$ 90 g/L 14. Serum albumin $\geqslant$ 29 g/L 15. Serum creatinine $\leqslant$ 1.5 × upper limit of normal range (ULN) or creatinine clearance (CL) $\geqslant$ 60 mL/min (according to Cockcroft-Gault equation) For female, CrCl = $\frac{(140 - \text{Age}) \times \text{Weight (kg)} \times 0.85}{72 \times \text{Serum creatinine (mg/dL)}}$ For male, CrCl = $\frac{(140 - \text{Age}) \times \text{Weight (kg)} \times 1.00}{72 \times \text{Serum creatinine (mg/dL)}}$ 16. Serum total bilirubin $\leqslant$ 2 × ULN. 17. Aspartate transaminase (AST) and alanine transaminase (ALT): AST $\leqslant$ 5 × ULN ALT $\leqslant$ 5 × ULN 18. Coagulation: International normalized ratio (INR) or prothrombin time (PT) $\leqslant$ 1.5 × ULN. **Other inclusion criteria** 19. Female subjects with childbearing potential must have negative serum pregnancy test result at screening. Female subjects with childbearing potential, and male subjects and their female partners with childbearing potential must agree to use an contraceptive method(s) (detailed contraception guidance is provided in Section 14.7 Appendix 7) from the day of signing informed consent form (ICF), during the study and till at least 6 months after the last dose of study treatment. **Exclusion Criteria** **Subjects who meet any of the following criteria will be excluded from the study:** **Target disease-specific exclusion criteria** 1. Fibrolamellar HCC, sarcomatoid HCC, cholangiocellular carcinoma or mixed cholangiocarcinoma and HCC. 2. History of hepatic encephalopathy. 3. A prior bleeding event due to esophageal or gastric varices within

	6 months or other gastrointestinal bleeding events within 28 days prior to screening. Untreated or incompletely treated esophageal or gastric varices that are considered by the investigator to be at high-risk for bleeding (Note: Patients must undergo an esophagogastroduodenoscopy [EGD], and all size of varices [small to large] must be assessed and treated per local standard of care prior to enrollment; patients who have undergone an EGD within 6 months prior to the initiation of study treatment do not need to repeat the procedure). Active gastric or duodenal ulcer. 4. Radiographic evidence of tumor thrombus in the main trunk of portal vein, vena cava or heart involvement on baseline imaging. 5. Malabsorption syndrome or inability to take oral medication due to other causes. 6. HBV and HCV co-infection. (Note: Subject has HCV infection history but HCV RNA is undetectable at screening can be considered as non-HCV infection in this study). 7. Surgery or locoregional therapy for palliative purpose (e.g., to treat bone metastases or metastases causing nerve impingement) within 4 weeks prior to study treatment. Palliative radiotherapy within 2 weeks prior to study treatment. Minor surgery (i.e, simple excision) within 7 days prior to study treatment. 8. Uncontrolled pleural effusion, pericardial effusion that are symptomatic and requiring repeated drainage or oral diuretics at screening. 9. Clinically meaningful ascites, defined as any ascites is detectable in the physical examination at screening, or is symptomatic, or requires special intervention, e.g., paracentesis or intraperitoneal perfusion (Note: Subjects with limited volume of ascites that is only identifiable by imaging can be enrolled). 10. Uncontrolled hypertension defined as systolic blood pressure > 150 mmHg and/or diastolic blood pressure > 90 mmHg and cannot be managed by medication. **Medical History and Concomitant Disease-Relevant Exclusion Criteria** 11. History of other malignancy(ies) in the past 5 years, except for malignant disease treated with curative intent and without active disease, e.g., basal cell carcinoma of skin, squamous cell carcinoma of skin, superficial bladder cancer or prostate cancer, breast cancer in situ or cervical cancer in situ.

	12. History of or current central nerve system (CNS) metastasis. 13. Active autoimmune disease or history of autoimmune disease that has possibility of relapse or at risk of having these conditions, e.g., history of organ transplantation and requiring immunosuppressants. Subjects with type I diabetes mellitus, hypothyroidism stable on hormone replacement, or dermatological condition that does not require systemic treatment (e.g., leukoderma, psoriasis or alopecia) are allowed to be enrolled. Any uncertainty, consultation with Sponsor's medical monitor is recommended before subjects signing informed consent. 14. Major cardiovascular diseases (e.g., congestive heart failure, unstable angina, atrial fibrillation, severe arrhythmia, etc.); any of acute myocardial infarction, unstable angina, stroke, or transient ischemic attack within 6 months before enrollment; New York Heart Association (NYHA) grade $\geqslant$ 2 congestive heart failure (detailed criteria is in appendix 5). 15. Presence of interstitial pneumonitis or non-infectious pneumonitis, which might confound the evaluation or management of pulmonary toxicity associated with study treatment-. 16. Any active infection (except for hepatitis virus infection) requires systemic treatment within 2 weeks prior to the first dose of study treatment. 17. Known history of human immunodeficiency virus (HIV) infection or acquired immunodeficiency syndrome (AIDS). 18. Active tuberculosis infection. **Previous Treatment-Relevant Exclusion Criteria:** 19. Any use of traditional Chinese medicine preparation with anti-tumor indication (e.g., Ban Mao capsule, elemene, Kanglaite, Cinobufacini, Xiaoaiping, Huai'er granule, Ganfule tablet, Jinlong capsule and Aidi) within 14 days prior to the first dose of study treatment. 20. Current or prior use of systemic corticosteroid (> 10 mg/day prednisone or equivalent) or other immunosuppressive medication within 14 days before the first dose of study treatment. Note: Inhaled or topical steroid, or adrenal replacement with $\leqslant$ 10 mg prednisone or equivalent is permitted if there is no active autoimmune disease. Short-term ($\leqslant$ 7 days) steroid for prophylactic treatment (e.g., for contrast allergy) or for non-autoimmune conditions (e.g., delayed hypersensitivity caused by contrast allergy) is allowed. Corticosteroid, if required by a subject,

	should be taken at least two days prior to or after CS1003 or placebo infusion. 21. History of bone marrow transplantation or organ transplantation. **Allergy and adverse reaction** 22. History of anaphylaxis or hypersensitivity to any ingredient of the investigational product. 23. Any contraindication of lenvatinib. **Other exclusion criteria** 24. Known history of drug abuse that would interfere with cooperation with the requirements of the trial. 25. Pregnant or lactating female subjects. 26. History of psychiatric disease that would interfere with cooperation with the requirements of the trial; lack of or with restricted physical capability. 27. QTc interval > 470 msec (as calculated with Fridericia's formula) at screening electrocardiogram (ECG). 28. Any condition that would in the investigator's judgement, prevent the subject from participating in this study. 29. Current participation in another clinical study or use of any investigational drug within 4 weeks prior the first dose of study treatment in this trial. 30. Major surgery within 4 weeks prior to the first dose of study treatment or planned major surgery during the study (mediastinoscopy, insertion of a central venous access device and insertion of a feeding tube are not considered major surgery). 31. Serious non-healing wound, ulcer, or bone fracture. 32. Urine protein ⩾1 g/24 hours (routine urine test, urine protein>1+ will have 24-hour urine collected to assess proteinuria).
Planned Study Duration:	[redacted]
End of Study:	[redacted]

Dosage and Administration of Study Treatment:	**Investigational product:** CS1003 (or placebo), 100 mg/4 mL/vial **Basic Treatment:** Lenvatinib, 4 mg/capsule **Dosage and administration:** Group A: CS1003 is administered at 200 mg by i.v. infusion Q3W. Lenvatinib is orally administered at 12 mg or 8 mg, with or without food, once daily, at approximately the same time every day (refer to the prescription information of lenvatinib). or Group B: Placebo is administered by i.v. infusion Q3W. Lenvatinib is orally administered at 12 mg or 8 mg, with or without food, once daily, at approximately the same time every day (refer to the prescription information of lenvatinib). When lenvatinib and CS1003/placebo are administrated on the same day, lenvatinib is to be taken first, and the i.v. infusion of CS1003 or placebo will be initiated within 30 minutes after lenvatinib administration. Refer to Section 6.6 for instructions on study treatment-related dose adjustment during the study.
Statistical Methods:	The primary endpoint is OS, defined as the time from randomization to the date of death of any cause. The key secondary efficacy endpoints include ORR and PFS assessed by BICR based on RECIST v1.1. The other secondary efficacy endpoints include: • PFS assessed by investigators based on RECIST v1.1 • ORR assessed by investigators based on RECIST v1.1 • DCR assessed by BICR and investigators based on RECIST v1.1 • DoR assessed by BICR and investigators based on RECIST v1.1 • TTD in EORTC QLQ-C30 Primary and secondary analyses (OS, PFS, ORR and DCR) are performed in all randomized subjects (intention-to-treat set). DoR analysis is performed in subjects with an objective response. TTD analysis is performed in all subjects with baseline PRO and at least once

	post-baseline PRO assessment. For efficacy analyses, subjects will be grouped according to treatment allocation at randomization. Safety analysis set is defined as all randomized subjects who received any study treatment. Randomized subjects who haven't received any dose of study treatment are not included in the safety analysis set. For safety analysis, subjects will be grouped by the actual received treatment group (patients who receive any dose of CS1003 incorrectly will be counted in the CS1003-treated group). Sample size determination: The primary endpoint of this study is OS. Number of events and sample size are estimated based on the following hypotheses: • Randomization ratio is 2:1. • Type I error is controlled for two-sided test at 0.05. • OS will be tested with log-rank test. • [redacted] • One interim analysis of OS will be conducted. Type I error will be controlled using Lan-DeMets approximation to the O'Brien-Fleming boundary. • Survival times are exponentially distributed • [redacted] • [redacted] Based on the above hypotheses, [redacted] **Interim analyses** [redacted] Log-rank test will be used for OS hypothesis, with its analysis boundaries shown in the table below:

	[redacted] **IDMC** An independent data monitoring committee (IDMC) is established to monitor safety data regularly and make recommendations to the Sponsor. The IDMC will also monitor the data for OS interim analysis and make recommendations to the Sponsor according to the pre-defined boundaries. Membership, roles and responsibilities, and implementation details of the IDMC will be described in the IDMC charter.

APPENDIX 2 SCHEDULE OF ASSESSMENT

Table 1 Study Assessments and Procedures Schedule

Study cycle	**Screening**	**Randomization**	**Treatment (1 cycle = 21 days)**								**Follow-up**			
			1	**2**	**3**	**4**	**5**	**6**	**Cn**	[x] **End of Treatment (EOT) visit**	**Safety follow-up visit 1 (FU1)**	**Safety follow-up visit 2 (FU2)**	**Safety follow-up visit 3 (FU3)**	[y] **Survival follow-up**
Date of visit (Day)	**-28 to -1**	**-3 to 1**	**1**	**22 (±3)**	**43 (±3)**	**64 (±3)**	**85 (±3)**	**106 (±3)**	**C(n-1)D1+21 (±3)**	**0-7 days after investigator's decision of EOT**	**30±7 (days) after the last dose**	**60±7 (days) after the last dose**	**90±7 (days) after the last dose**	**Every 12 weeks ±14 days**
[a] Informed Consent	X													
Inclusion/Exclusion Criteria	X													
Demographics	X													
Medical History	X													
Diagnosis and Confirmation of cancer	X													
Prior Anti-Cancer Treatment	X													
[b] Tumor Tissue	X													
Height	X													
Weight	X		X	X	X	X	X	X	X	X	X			
ECOG Performance Status	X		X	X	X	X	X	X	X	X	X			
[c] HBV, HCV and HIV tests	X						X		HBV-DNA or HCV-RNA is re-tested every	X	X			

									4 cycles for subjects with positive HBsAg/ HBcAb or positive HCV antibody					
[d] Physical Examination	X		X	X	X	X	X	X	X	X	(X)			
[e] Vital signs	X		X	X	X	X	X	X	X	X	X			
[f] 12-Lead Electrocardiogram (ECG)	X		X		X		X		Every 2 cycles	X	X			
[g] Echocardiogram	X		As clinically indicated											
[h] EGD	X													
[i] Laboratory Tests	X		X	X	X	X	X	X	X	X	X			
[j] AFP	X						X		Every 4 cycles	X	X			
[k] Coagulation Parameters	X		X	X	X	X	X	X	X	X	X			
[l] Thyroid Function	X		X		X		X		Every 2 cycles	X	X			
[m] Pregnancy Test	X		Every 2 cycles (every cycle in monthly intervals in EU countries)							X	X	X^{m}	X^{m}	
[n] Review Adverse Events	X	X	X	X	X	X	X	X	X	X	X	X	X	(X)
Review Concomitant Medications	X	X	X	X	X	X	X	X	X	X	X	X	X	
[o] Subsequent Anti-Cancer Treatment										X	X	X	X	X

[p] Tumor Imaging	X		**In the first 48 weeks:** every 6 weeks ± 7 days **After the first 48 weeks**: every 12 weeks ± 7 days; additional imaging is allowed as clinically indicated											
[q] PK Blood Sampling			X	X	X	X			(X)	X				
[r] ADA Sampling			X	X	X	X			(X)	X			X	
[s] Patient-Reported Outcome (PRO)	X		X	X	X	X	X	X	Every 2 cycles	X	X			X
[t] Randomization		X												
[u] CS1003 or Placebo Administration			X	X	X	X	X	X	X					
[v] Lenvatinib Administration			X	X	X	X	X	X	X					
[w] Review Treatment Compliance				X	X	X	X	X	X	X				
Survival status														X

Abbreviations: C, cycle; D, day; ADA, anti-drug antibody; ECOG, Eastern Cooperative Oncology Group Electrocardiogram; ECG, electrocardiogram; EGD, esophagogastroduodenscopy; HBV, hepatitis B virus; HCV, hepatitis C virus; HIV, human immunodeficiency virus; PK, pharmacokinetics; PRO, patient-reported outcomes using EORTC QLQ-C30, EORTC QLQ-HCC18 and EQ-5D-5L scales; SAE, serious adverse event.

a. Subjects must provide written informed consent. Subjects must be assessed based on all the inclusion/exclusion criteria and eligibility is confirmed prior to randomization. Re-screening under limited conditions should be allowed only once after consultation with the Sponsor. Subject must undergo another informed consent process for re-screening and be allocated a new subject number. Assessments which are part of the patient's clinical standard of care may be performed before obtaining the written informed consent, if within the acceptable screening window.

b. During screening, subjects without cirrhosis require histological confirmation for HCC diagnosis (refer to Inclusion Criterion 3 for full details on HCC diagnosis). Under regulatory approval (if applicable), tumor tissue samples for exploratory biomarker assessment need to be collected. Tumor tissue samples can be archival formalin-fixed paraffin-embedded (FFPE) tissue block collected previously within 5 years before randomization (preferably within 12 months), or 3-5 freshly sectioned unstained slides (freshly sectioned within 6 months). When archival tissue is not available, it is optional for the subjects to

have a fresh biopsy to collect tumor tissue collection if deemed medicinally safe by the investigator. Refer to Protocol section 7.5 and Lab Manual for detailed requirements for tumor sample collection.

c. All subjects must receive hepatitis B surface antigen (HBsAg), hepatitis B core antibody (HBcAb) and hepatitis C virus (HCV) antibody tests. Those with positive HBsAg or positive HBcAb must have hepatitis B virus (HBV) DNA titer test. Subjects with positive HCV antibody must have HCV RNA test. Thereafter, HBV-DNA and/or HCV-RNA tests are repeated every 4 cycles started from Cycle 5 Day 1 (C5D1) for subjects with positive HBsAg or HBcAb or positive HCV antibody during study treatment phase (i.e. on C5D1, C9D1, C13D1 and so on), at the end-of-treatment (EOT) visit and safety follow-up visit 1 (no need to repeat the tests if there are available results obtained within 7 days prior to the EOT visit or safety follow-up visit 1).
d. A complete physical examination will be performed at screening and at every treatment cycle before administration of study treatment. Refer to Appendix 14.1 for detailed physical examination items. Physical examination will be performed at safety follow-up 1 when indicated by symptoms. All abnormal findings and any newly developed or worsened physical examination abnormality or related symptoms will be recorded.
e. Vital signs including temperature, pulse, respiratory rate and blood pressure must be taken at screening and at every treatment cycle before administration of drugs. If vital signs at screening are taken within 3 days prior to the first dose of study treatment, the measurements don't need to be repeated before the first dose. Tympanic approaches (if applicable) can be used for temperature monitoring. All abnormal findings and any newly developed or worsened vital sign abnormality or related symptoms will be recorded.
f. ECG should be performed before any planned vital sign measurement or blood sampling, if possible. The evaluations are performed at the following time points: at screening and every 2 cycles starting from Cycle 1 Day 1 during treatment, at the end-of-treatment visit and safety follow-up visit 1. Additional ECGs will be performed as clinically indicated. If ECG at screening is taken within 7 days prior to the first dose of study treatment, the measurements do not need to be repeated before the first dose. For ECG performed within 7 days prior to the end of treatment visit or safety follow-up visit 1, it is not necessary to repeat the measurements in these two visits. The ECG at each study treatment administration visit must be performed within 3 days prior to the administration of study treatment.
g. Echocardiogram evaluation is required at screening and will be performed as clinically indicated throughout treatment period.
h. EGD: All patients must undergo an EGD and all size of varices (small to large) must be assessed and treated per local standard of care prior to enrollment. The patients who have undergone an EGD within 6 months prior to the initiation of study treatment do not need to repeat the procedure.
i. Routine laboratory tests include hematology, serum chemistry and urinalysis are performed at screening, before administration of study treatment at each cycle, end-of-treatment visit and safety follow-up visit 1. If laboratory test at screening is performed within 7 days prior to the first dose of study treatment, the laboratory test of the first cycle before the first dose will not be required. When laboratory test and administration of CS1003 or lenvatinib are to be on the same day (e.g., the first day [D1] of each treatment cycle), the test result must be available before starting the study treatment. The laboratory tests at each study treatment administration visit must be performed within 3 days prior to the administration of study treatment. For laboratory tests with available results performed within 7 days prior to the end of treatment visit or safety follow-up visit 1, it's not necessary to repeat the tests in these two visits. Hematology test includes complete blood count and hemoglobin. Serum chemistry tests include blood urea or urea nitrogen (BUN), creatinine, sodium, potassium, magnesium, chloride, calcium, phosphate, blood glucose, total bilirubin, direct bilirubin, alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), lactate dehydrogenase (LDH), creatine kinase (CK), total cholesterol, total protein and albumin.

Urinalysis includes specific gravity, pH, urine glucose, protein, 24-hour urine protein (as clinical indicated), urine casts, ketones, occult blood and white blood cells. Refer to Appendix 14.1 for detailed test items.

j. AFP test is performed at screening and once every 4 cycles started from Cycle 5 Day 1 during treatment period until safety follow-up 1 (inclusive). For AFP test with available result performed within 7 days prior to the end of treatment visit or safety follow-up visit 1, it is not necessary to repeat the test in these two visits.

k. The coagulation tests are performed at screening, before administration of study treatment at each cycle, end-of-treatment visit and safety follow-up visit 1. If the test at screening is performed within 7 days prior to the first dose of study treatment, the laboratory test of the first cycle before the first dose will not be required. When laboratory test and administration of CS1003 or lenvatinib are to be on the same day (e.g., the first day [D1] of each cycle), test result must be available before starting the study treatment. Laboratory tests performed within 7 days prior to the end of treatment visit or safety follow-up visit 1 are not required to be repeated at these two visits. Refer to Appendix 14.1 for detailed test items.

l. Thyroid function tests are performed at screening, before first dose of study treatment and every two treatment cycles thereafter, at the end-of-treatment visit and safety follow-up visit 1. If the test at screening is performed within 7 days prior to the first dose of study treatment, the laboratory test of the first cycle before the first dose will not be required. For laboratory tests with available results performed within 7 days prior to the end of treatment visit or safety follow-up visit 1, they are not required to be repeated at these two visits. Refer to Appendix 14.1 for detailed test items. When the investigator encounters the difficulty in having all thyroid-function test results ready before the Day 1 dosing in certain cycle, a subject may continue with the dosing at the investigator's discretion only if the subject is not showing any symptom suspected of abnormal thyroid function. The investigator should continue to monitor the subject's condition and take immediate actions if necessary.

m. Female subjects of childbearing potential (except those with record of sterilization operation or after menopause) must receive serum pregnancy test at screening and have available results prior to the first study treatment. If the pregnancy test at screening is performed within 7 days prior to the first dose of study treatment, the test of the first cycle before the first dose will not be required. Urinary or serum pregnancy test will be performed every 2 cycles throughout treatment period, at the end of treatment visit, and at safety follow-up visit 1. If a subject has had pregnancy test done within 7 days prior to the end of treatment visit or safety follow-up visit 1, the tests at these two visits will not necessarily be repeated. Subjects with positive urine pregnancy test must receive serum pregnancy test for confirmation, and negative result must be obtained before any administration of study treatment at each treatment cycle. For EU countries, the pregnancy test should be done at every cycle in monthly intervals and in safety follow-up visit 2 and 3. Additional pregnancy tests may also be undertaken if requested by Institutional Review Board/Ethics Committees (IRB/ECs) or local regulations.

n. Adverse events (AEs): All AEs and serious adverse events (SAEs) will be recorded during the reporting period from subject signing the informed consent, throughout the treatment period and safety follow-up, till 90 days after the last dose of the study treatment, or the subject starts a new anti-cancer treatment, whichever occurs earlier, after which time, only related SAEs need to be recorded and reported. AEs that occur after safety follow-up visit 1 can be collected at safety follow-up visit 2 and 3 by phone calls.

o. Subsequent anti-cancer treatment: information of subsequent anti-cancer treatment (after the end of study treatment) will be collected at each subsequent scheduled visit starting from the end-of treatment visit, including generic drug names of the treatment, starting date and ending date of treatment.

p. Tumor imaging evaluation: Tumor response assessment will be performed based upon Response Evaluation Criteria in Solid Tumors (RECIST) v1.1. Refer to Appendix 14.2 for details. Tumor imaging will be conducted using CT or MRI according to the investigator's decision. MRI is preferred for brain

imaging, or CT is also permitted if a subject has any contraindication against MRI. For each subject, the same imaging method, equipment and technical parameters should be used throughout the study. The same investigator or radiologist should perform the imaging evaluation throughout the study, if possible. Contrast agent should be used if not contraindicated. Tumor assessment performed within 14 days prior to the randomization can be used as the baseline tumor evaluation if accepted by investigators. Baseline tumor evaluation should include brain, chest, abdomen, pelvis and any other site with suspected lesion. The tumor imaging will be performed every 6 weeks (±7 days) after the randomization in the first 48 weeks, and every 12 weeks (± 7 days) thereafter, until PD confirmed by radiological examination, withdrawal of informed consent, lost to follow-up, death, or end of study, whichever occurs first. Subjects who discontinue study treatment due to causes other than PD will continue to receive radiological assessments at pre-determined time points until PD, subject withdraws informed consent, lost to follow-up, end of study, or death, whichever occurs first. Subjects that are without brain metastasis as confirmed by brain imaging at screening are not required to receive regular brain imaging in the following visits. Additional brain imaging can be arranged by the investigator as clinically indicated. Other additional imaging is also allowed as clinically indicated. For any investigator-suspected pseudo-progression, prior to discontinuation of study treatment, a re-imaging will be performed at least 4 weeks later or at the next regularly scheduled imaging time point to confirm disease progression. (Note that the repeated radiologic examination must be performed no later than 9 weeks after the first recorded disease progression). For tumor radiologic evaluation performed within 28 days prior to the end of treatment visit or safety follow-up visit 1, they are not required to be repeated at these two visits.

All the images will be transferred to central radiology laboratory for blinded independent central review (BICR, refer to BICR Guidance for details) till the number of PFS events for PFS analysis has been reached.

q. Refer to Protocol Table 1.2-2.

r. Refer to Protocol Table 1.2-2.

s. Patient-reported outcomes (PRO) are collected through EORTC QLQ-C30, EORTC QLQ-HCC18 and EQ-5D-5L questionnaires completed by subjects on paper before any other evaluation at each visit. Before subject leaves the site, the investigator should review completeness of all questionnaires. These PROs should be assessed prior to any examinations scheduled on Day 1 of the first 6 treatment cycles, and then every 2 cycles from treatment Cycle 8 Day 1 on until the end-of-treatment visit (i.e. treatment C8D1, C10D1, C12D1, C14D1, and EOT), and safety follow-up visit 1. If PROs are assessed within 7 days prior to the end-of-treatment visit or safety follow-up visit, they are not required to be repeated in these two visits. During survival follow-up, an assessment is performed every 12 weeks unless the subject withdraws the ICF or the Sponsor ends this study. During survival follow-up, PRO could be collected at the site when subject comes for a visit or by phone calls. PRO assessment has a time window of -7 days (±7 days for EOT visit and visits thereafter), except the one on C1D1 which must be conducted on the same day and prior to any examinations scheduled on that day.

t. A subject can be randomized when eligibility is confirmed by the investigators. The first dose of study treatment should be administered within 3 days of randomization.

u. Subjects will receive CS1003 or placebo every 21 days (3 weeks), ± 3 days if necessary (but the interval between two doses must be ≥ 18 days) until end of treatment visit, disease progression, unaccepted toxicity, withdrawal of ICF, lost to follow-up, death or end of study, whichever occurs first.

v. Lenvatinib is administered once daily. When lenvatinib is given on the same day as CS1003 or placebo, please refer to Protocol Section 6.6.2 for details on administration and dose modification.

w. Treatment compliance will be monitored using drug dispense and return log, subject's medical record and electronic case report form (eCRF). Subjects should record the dose and time of lenvatinib administration every day and submit the record to the study coordinator at the next visit. At the end of treatment, subjects should return all used and unused drug containers for compliance evaluation.

x. End-of-treatment visit: Subjects who discontinue study treatment due to any reason should receive end of treatment visit within 7 days after the investigator determines discontinuation of treatment. If end of treatment visit is within the time window of safety follow-up visit 1, it will be done according to the safety follow-up requirements, and same test items will not be repeated.

y. Subjects should be contacted by telephone follow-up or site visit every 12 weeks (± 14 days) (or more frequently as needed) from the last dose of study treatment until lost to follow-up, death or end of study. Data on subject's survival, the subsequent anti-cancer therapy, PRO, and drug-related SAE will be collected in these telephone visits/site visits.

APPENDIX 3 TIME POINTS OF BLOOD SAMPLING FOR CS1003 PK AND ADA

Study Day		**Time**	**Time window[a]**	**PK Blood Sampling[b]**	**ADA Blood Sampling[b]**
Cycle 1	D1	0 (before administration)	Within 60 minutes before administration	X	X
		EOI[c]	+ 30 minutes	X	
Cycle 2	D1	0 (before administration)	Within 60 minutes before administration	X	X
Cycle 3	D1	0 (before administration)	Within 60 minutes before administration	X	X
Cycle 4	D1	0 (before administration)	Within 60 minutes before administration	X	X
		EOI[c]	+ 30 minutes	X	
Cycle 8, and every 4 cycles thereafter	D1	0 (before administration)	Within 60 minutes before administration	X	X
End-of-treatment (EOT) visit	—	—	—	X	X
Safety follow-up visit 3	—	—	—		X

Abbreviations: EOI, end of infusion; ADA, anti-drug antibody; PK, pharmacokinetics; D1, day 1.
[a] PK and ADA sampling time window is calculated based on the administration of CS1003
[b] ADA and PK sample of 3.5 mL whole blood each.
[c] PK and ADA blood sampling follows the "rule of distal end" after the end of infusion; blood samples are taken from veins on the opposite limb of the infusion limb or on the limb without infusion site.

APPENDIX 4 HANDLING OF MISSING DATES

1. Missing dates of concomitant medications will be handled as below:

    1) For the start date, if only day is missing, then impute with the first day of the corresponding month; if both month and day are missing, then impute with 01 Jan. If the imputed start date is later than the end date, then it will be replaced by the end date.

    2) For the end date, if only day is missing, then impute with the last day of the corresponding month; if both month and day are missing, then impute with 31 Dec.

2. Missing dates of adverse events will be handled as below:

    1) For missing end date

        a) If only day is missing, then impute with the last day of the corresponding month;

        b) If both month and day are missing, then impute with 31 Dec.

        c) For AE leading to death, end date will be the earlier date of the death date and the imputed end date.

        d) Otherwise, set to missing.

    2) For missing start date

        a) If only day is missing, and year and month are earlier then the corresponding part of first dose date of study treatment, then impute with the last day of the corresponding month.

        b) If only day is missing, and year and month are the same as the corresponding part of first dose date of study treatment, then impute with the first dose date of study drug.

        c) If only day is missing, and year and month are later then the corresponding part of first dose date of study treatment, then impute with the first day of the corresponding month.

        d) If both month and day are missing, and year is earlier then the year of first dose date of study treatment, then impute with 31 Dec.

        e) If both month and day are missing, and year is the same as the year of first dose date of study treatment, then impute with the first dose date of study drug.

        f) If both month and day are missing, and year is later then the year of first dose date of study treatment, then impute with 01 Jan.

g) If day, month, and year are all missing, then impute with first dose date of study treatment.

h) If imputed start date is later than the end date, then it will be replaced by the end date.

i) Otherwise, set to missing.

3. Missing death dates will be handled as follows:

1) In safety analyses, all deaths are included, from all sources, regardless of completeness of death date. Patients who died with only a partial death date available will be included and will not be imputed.

2) In efficacy analyses, if only day is missing, then impute with the later date of the first day of the corresponding month and last known alive date+1. Missing death dates will not be imputed under other missing scenarios. A death is considered an event if the death date is complete after imputation.

APPENDIX 5 CTCAE GRADES VERSION 5.0 FOR LABORATORY PHRAMETERS TO BE ANALYZED

CTCAE v5.0 Term	Grade 1	Grade 2	Grade 3	Grade 4
Albumin decreased (Hypoalbuminemia)	<LLN - 30 g/L	<30 - 20 g/L	<20 g/L	
Calcium decreased (Hypocalcemia)	<LLN - 2.0 mmol/L	<2.0 - 1.75 mmol/L	<1.75 - 1.5 mmol/L	< 1.5 mmol/L
CD4 lymphocytes count decreased	<LLN - 0.5 x 10e9 /L	<0.5 - 0.2 x 10e9 /L	<0.2 x 0.05 - 10e9 /L	<0.05 x 10e9 /L
Glucose decreased (Hypoglycemia)	<LLN - 3.0 mmol/L	<3.0 - 2.2 mmol/L	<2.2 - 1.7 mmol/L	<1.7 mmol/L
Fibrinogen decreased	<1.0 - 0.75 x LLN; if abnormal, <25% decrease from baseline	<0.75 - 0.5 x LLN; if abnormal, 25 - <50% decrease from baseline	<0.5 - 0.25 x LLN or if abnormal, 50 - <75% decrease from baseline	<0.25 x LLN or if abnormal, 75% decrease from baseline or absolute value <50 mg/dL
HGB decreased (Anemia)	<LLN – 100 g/L	<100 – 80 g/L	<80 g/L	
Haptoglobin decreased	<LLN			
Lymphocyte count decreased	<LLN - 0.8 x 10e9/L	<0.8 - 0.5 x 10e9 /L	<0.5 - 0.2 x 10e9 /L	<0.2 x 10e9 /L
Magnesium decreased (Hypomagnesemia)	<LLN - 0.5 mmol/L	<0.5 - 0.4 mmol/L	<0.4 - 0.3 mmol/L	< 0.3 mmol/L
Neutrophil count decreased	<LLN - 1.5 x 10e9 /L	<1.5 - 1.0 x 10e9 /L	<1.0 - 0.5 x 10e9 /L	<0.5 x 10e9 /L
Platelet count decreased	<LLN – 75.0 x10e9 /L	<75.0 - 50.0 x10e9 /L	<50.0 – 25.0 x10e9 /L	<25.0 x 10e9 /L
Potassium decreased (Hypokalemia)		<LLN - 3.0 mmol/L	<3.0 - 2.5 mmol/L	< 2.5 mmol/L
*Sodium decreased (Hyponatremia)	<LLN - 130 mmol/L		<130 - 120 mmol/L	< 120 mmol/L
White blood cell decreased	<LLN - 3.0 x 10e9 /L	<3.0 - 2.0 x 10e9 /L	<2.0 - 1.0 x 10e9 /L	<1.0 x 10e9 /L

Alkaline phosphatase increased	When baseline: >ULN - 2.5 x ULN When post-baseline: >ULN - 2.5 x ULN if baseline was normal; 2.0 - 2.5 x baseline if baseline was abnormal	When baseline: >2.5 - 5.0 x ULN When post-baseline: >2.5 - 5.0 x ULN if baseline was normal; >2.5 - 5.0 x baseline if baseline was abnormal	When baseline: >5.0 - 20.0 x ULN When post-baseline: >5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	When baseline: >20.0 x ULN When post-baseline: >20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal
Alanine aminotransferase increased	When baseline: >ULN - 3.0 x ULN When post-baseline: >ULN - 3.0 x ULN if baseline was normal; 1.5 - 3.0 x baseline if baseline was abnormal	When baseline: >3.0 - 5.0 x ULN When post-baseline: >3.0 - 5.0 x ULN if baseline was normal; >3.0 - 5.0 x baseline if baseline was abnormal	When baseline: >5.0 - 20.0 x ULN When post-baseline: >5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	When baseline: >20.0 x ULN When post-baseline: >20.0 x ULN if baseline was normal;>20.0 x baseline if baseline was abnormal
Aspartate aminotransferase increased	When baseline: >ULN - 3.0 x ULN When post-baseline: >ULN - 3.0 x ULN if baseline was normal; 1.5 - 3.0 x baseline if baseline was abnormal	When baseline: >3.0 - 5.0 x ULN When post-baseline: >3.0 - 5.0 x ULN if baseline was normal; >3.0 - 5.0 x baseline if baseline was abnormal	When baseline: >5.0 - 20.0 x ULN When post-baseline: >5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	When baseline: >20.0 x ULN When post-baseline: >20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal
**Corrected Calcium increased (Hypercalcemia)	>ULN - 2.9 mmol/L	>2.9 - 3.1 mmol/L	>3.1 - 3.4mmol/L	>3.4 mmol/L
*Serum amylase increased	>ULN - 1.5 x ULN	>1.5 - 2.0 x ULN	>2.0 - 5.0 x ULN	>5.0 x ULN

Cholesterol high	>ULN - 7.75 mmol/L	>7.75 - 10.34 mmol/L	>10.34 - 12.92 mmol/L	>12.92 mmol/L
CPK increased	>ULN - 2.5 x ULN	>2.5 x ULN - 5 x ULN	>5 x ULN - 10 x ULN	>10 x ULN
Creatinine increased	>ULN - 1.5 x ULN	>1.5 - 3.0 x baseline; >1.5 - 3.0 x ULN	>3.0 x baseline; >3.0 - 6.0 x ULN	>6.0 x ULN
GGT increased	When baseline: >ULN - 2.5 x ULN When post-baseline: >ULN - 2.5 x ULN if baseline was normal; 2.0 - 2.5 x baseline if baseline was abnormal	When baseline: >2.5 - 5.0 x ULN When post-baseline: >2.5 - 5.0 x ULN if baseline was normal; >2.5 - 5.0 x baseline if baseline was abnormal	When baseline: >5.0 - 20.0 x ULN When post-baseline: >5.0 - 20.0 x ULN if baseline was normal; >5.0 - 20.0 x baseline if baseline was abnormal	When baseline: >20.0 x ULN When post-baseline: >20.0 x ULN if baseline was normal; >20.0 x baseline if baseline was abnormal
Hemoglobin increased	Increase in >0 - 2 g/dL above ULN	Increase in >2 - 4 g/dL above ULN	Increase in >4 g/dL above ULN	
INR Increased	>1.0 - 1.5 x ULN or >1.0 - 1.5 x baseline	>1.5 - 2.5 x ULN or >1.5 - 2.5 x baseline	>2.5 x ULN or >2.5 x baseline	
*Lipase increased	>ULN - 1.5 x ULN	>1.5 - 2.0 x ULN	>2.0 - 5.0 x ULN	>5.0 x ULN
Lymphocyte count increased		>4 - 20 x10e9 /L	>20 x10e9 /L	
Magnesium increased (Hypermagnesemia)	>ULN - 1.23 mmol/L		>1.23 - 3.3 mmol/L	>3.3 mmol/L
Potassium increased (Hyperkalemia)	>ULN - 5.5 mmol/L	>5.5 - 6.0 mmol/L	<6.0 - 7.0 mmol/L	>7.0 mmol/L
Activated partial thromboplastin time prolonged	>1.5 - ULN	>1.5 - 2.5 x ULN	>2.5 x ULN	
Sodium increased (Hypernatremia)	>ULN - 150 mmol/L	>150 - 155 mmol/L	>155 - 160 mmol/L	>160 mmol/L

Blood bilirubin increased	When baseline: >ULN - 1.5 x ULN When post-baseline: >ULN - 1.5 x ULN if baseline was normal; > 1.0 - 1.5 x baseline if baseline was abnormal	When baseline: >1.5 - 3.0 x ULN When post-baseline: >1.5 - 3.0 x ULN if baseline was normal; >1.5 - 3.0 x baseline if baseline was abnormal	When baseline: >3.0 - 10.0 x ULN When post-baseline: >3.0 - 10.0 x ULN if baseline was normal; >3.0 - 10.0 x baseline if baseline was abnormal	When baseline: >10.0 x ULN When post-baseline: >10.0 x ULN if baseline was normal; >10.0 x baseline if baseline was abnormal
Triglycerides increased (Hypertriglyceridemia)	1.71 - 3.42 mmol/L	>3.42 - 5.7 mmol/L	>5.7 - 11.4 mmol/L	>11.4 mmol/L
White blood cell increased (Leukocytosis)			>100 x10e9 /L	

Parameters of Hyperglycemia/Hypophosphatemia/Hyperuricemia are no quantified criteria and are removed from analysis in version 5.0.

*There are clinical judgements with quantified criteria for parameters of Hyponatremia/Serum amylase increased/Lipase increased in version 5.0, the rule of version 4.03 will be followed.

** Derivation for corrected calcium: If serum calcium in blood chemistry is expressed in mmol/L and albumin is expressed in g/L, corrected calcium can be derived using the reported total calcium value and albumin at the same assessment using the following formula: Corrected calcium (mmol/L) = serum calcium (mmol/L) + 0.0246 x (40 - serum albumin [g/L]) if serum albumin is less than 40 g/L. If serum albumin is ≥ 40 g/L, there is no need to correct albumin, use serum calcium in blood chemistry in the analysis.

APPENDIX 6 ADVERSE EVENTS OF SPECIAL INTEREST