Title: An Open-Label, Multi-Center, Phase II Study of Anti-PD-1 Antibody SHR-1210 Plus Capecitabine and Oxaliplatin Sequenced by SHR-1210 with Apatinib Mesylate or Apatinib in Patients with Previously Untreated Advanced or Metastatic Gastric (GC) or Gastroesophageal Junction (GEJ) Cancer

NCT number: NCT03472365

Date: 22 Dec., 2020

AN OPEN-LABEL, MULTI-CENTER, PHASE II STUDY OF ANTI-PD-1 ANTIBODY SHR-1210 PLUS CAPECITABINE AND OXALIPLATIN SEQUENCED BY SHR-1210 WITH APATINIB MESYLATE OR APATINIB IN PATIENTS WITH PREVIOUSLY UNTREATED ADVANCED OR METASTATIC GASTRIC (GC) OR GASTROESOPHAGEAL JUNCTION (GEJ) CANCER

Statistical Analysis Plan

(SAP)

Author: Yanfei Tai

Company: Jiangsu Hengrui Pharmaceuticals Co., Ltd.

Version: 1.0

Date: 22 Dec., 2020

This SAP has been reviewed by the following personnel before being approved and effective.

Functional Role	**Reviewer**
Medical Science	[redacted]
Statistics	[redacted]
Clinical Pharmacology	[redacted]

TABLE OF CONTENTS

LIST OF TABLES

ABBREVIATIONS

AE	Adverse event
BOR	Best overall response
CR	Complete response
DCR	Disease control rate
DoR	Duration of response
FAS	Full analysis set
ORR	Objective response rate
PD	Pharmacodynamics
PFS	Progression-free survival
PR	Partial response
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SD	Stable disease
SS	Safety set
TTOR	Time to objective response

1 REVISION

None.

2 PROTOCOL AMENDMENT

Added overall survival (OS), time to objective response (TTOR), and subgroup analysis as exploratory efficacy analysis to provide guidance for subsequent phase III clinical studies

3 INTRODUCTION

This statistical analysis plan is formulated to provide specific statistical analysis and reporting methods or strategies for a phase II study on the efficacy and safety of SHR-1210 plus capecitabine and oxaliplatin sequential treatment (cohort 1) or SHR-1210 plus apatinib (cohort 2) in previously untreated advanced or metastatic gastric or gastroesophageal junction cancer. The final draft of this plan will be completed before database locking and will be signed by various functional departments for confirmation.

3.1 Study Design

This is a randomized, open-label, multi-center, phase II clinical study.

A total of 98 patients with previously untreated advanced or metastatic gastric cancer or gastroesophageal junction cancer will be enrolled. Subjects will be randomized to receive sequential therapy of SHR-1210 plus capecitabine and oxaliplatin (cohort 1, n = 43) or SHR-1210 plus apatinib treatment (cohort 2, n = 55).

Cohort 1 (sequential therapy of SHR-1210 combined with capecitabine and oxaliplatin): SHR-1210 + capecitabine + oxaliplatin for 4-6 cycles, followed by SHR-1210 + apatinib for subjects without progressive disease (PD);

Cohort 2 (SHR-1210 combined with apatinib): SHR-1210 + apatinib.

A two-stage design is adopted to minimize the exposure of subject to ineffective treatment. Nineteen subjects will be enrolled into cohort 1 and cohort 2 at the first stage, respectively. If at least 4 subjects achieve objective response (complete response or partial response), the second stage will be initiated. And more subjects will be enrolled in cohort 1(to 43 subjects) and in cohort 2 (to 55 subjects), respectively.

In this study, the screening period should be no more than 28 days. After completing screening examinations and assessments, eligible subjects will enter the treatment period (21 days/cycle) and begin the study treatment and study visits according to the protocol. Tumor imaging assessment will be performed once in every 2 cycles (6 weeks ±7 days) during the first 12 months (the first 16 cycles) of the study treatment period and once every 3 cycles (9 weeks ± 7 days) thereafter. The subjects will complete safety examinations and imaging assessments upon discontinuation of the study treatment.

Thereafter, the subjects will enter the follow-up period. The safety follow-up period will start after the last administration, and the subjects are followed up every 30 days (±7 days) until 90 days after the last administration. The first safety follow-up will be carried out at the study site; the second and the third follow-up visits will be made via telephone calls. The survival follow-up period starts after the end of the safety follow-up period. The survival follow-up period ends upon the subject's death, lost to follow-up, withdrawal of informed consent, or study termination by sponsor. During this period, a follow-up will be conducted every month via telephone or other effective methods to collect information on subject survival and subsequent treatments. For subjects who show no evidence of radiographic progression, an imaging evaluation should continue at the frequency of response evaluations specified in the protocol, until progressive disease, death, lost to follow-up, withdrawal of informed consent, start of other anti-cancer treatments, or trial termination by the sponsor.

All subjects should continue to receive the study treatment until progressive disease, intolerable toxicity, voluntary treatment discontinuation or study withdrawal by the subject, or discontinuation determined by the investigator.

SHR-1210 is an immune checkpoint inhibitor and according to the experience of drugs of same class, some subjects may experience delayed or early pseudoprogression after receiving immunotherapy drugs. Therefore, if subjects in the treatment group who have experienced progressive disease (PD) for the first time meet the criteria for continuing treatment after progressive disease, they can continue to be treated according to the original treatment regimen. Subjects who do not have progressive disease or intolerable toxicities may continue the SHR-1210 treatment for no more than 24 cumulative months.

Subjects who do not have progressive disease after 24-month treatment may continue the apatinib monotherapy according to the prescribing information of apatinib.

3.2 Study Objectives

3.2.1 Primary objective

To evaluate the efficacy of the sequential therapy of SHR-1210 combined with capecitabine and oxaliplatin or SHR-1210 combined with apatinib mesylate as the first-line therapy for advanced or metastatic gastric (GC) or gastroesophageal junction (GEJ) cancer.

3.2.2 Secondary objective

To evaluate the safety of the sequential therapy of SHR-1210 combined with capecitabine and oxaliplatin or SHR-1210 combined with apatinib mesylate for the treatment of advanced or metastatic gastric (GC) or gastroesophageal junction (GEJ) cancer.

4 INTERIM ANALYSIS, FINAL ANALYSIS, AND UNBLINDING

4.1 Interim Analysis

At the end of Stage I of this study, it will be judged whether to continue the trial in a certain cohort according to the rules in Section 5.

4.2 Final Analysis

The final analysis will be performed after the database is locked and released.

4.3 Unblinding

This is an open-label clinical study and does not involve unblinding.

5 STATISTICAL HYPOTHESES AND DECISION RULES

Cohort	H0	H1	N1	R1	N	R	Power	Alpha
Cohort 1	35% ORR	55% ORR	19	4	43	21	80%	0.05
Cohort 2	15% ORR	30% ORR	19	4	55	13	80%	0.05

N1: Number of subjects at the first stage;

R1: The minimum effective number of subjects required in Stage I; if the effective number of subjects in Stage I is less than R1, the trial will be discontinued because of invalidity. If ≥ 4 subjects in cohort 1 show disease response (CR or PR), the study will enter Stage II and 24 subjects will be added. If ≥ 4 subjects in cohort 2 show disease response, the study will enter Stage II and 36 subjects will be added (55 subjects in cohort 2).

N: Final number of subjects;

R: The minimum effective number of subjects eventually required. If the total effective number of subjects is less than R after the end of Stage II (cohort 1 fails to show at least 21 subjects of disease responses, and cohort 2 fails to show at least 13 subjects of disease responses), subsequent drug study will be not considered necessary.

6 ANALYSIS SETS

6.1 Full Analysis Set (FAS)

It includes all subjects who are randomized and have received study drugs at least once. This analysis set will be used for the efficacy analysis.

6.2 Per-Protocol Set (PPS)

It is a subset of the full analysis set, and subjects with important protocol deviations judged to have a significant impact on treatment efficacy will be excluded from this set. The list of subjects included into or excluded from the PPS should be reviewed and determined by the sponsor and the investigator before the database is locked.

6.3 Safety Analysis Set

It includes all enrolled subjects who have received the study drugs at least once and have post-administration safety evaluation data.

6.4 Evaluable Set

It is a subset of FAS and includes all enrolled subjects who have received the study drugs at least once and have at least one post-baseline tumor assessment.

6.5 PK Analysis Set (PKS)

It includes all enrolled subjects who have received the study drugs and have post-administration PK evaluation data (concentration and/or parameter).

6.6 Immunogenicity Analysis Set (ADA)

It includes all enrolled subjects who have received the study drugs at least once and have baseline and at least one set of post-baseline ADA evaluation data.

7 MISALLOCATION OF TREATMENT GROUPS

Not applicable.

8 PROTOCOL DEVIATION

The protocol deviation analysis is based on FAS. Protocol deviations may include but are not limited to the following types:

- Not meeting the inclusion/exclusion criteria;
- Meeting the withdrawal criteria but the subject has not withdrawn from the study;
- Not taking study drugs according to the protocol, such as taking wrong doses and missed doses;
- Using combined medications prohibited by the protocol, such as other anti-tumor drugs and adjuvant drugs related to tumor treatment (including anti-cancer Chinese medicine);
- Non-compliance to the procedures stipulated in the protocol, such as omission of blood sample collection, and blood sample collection time exceeding the time window.

In addition, the list of subjects experiencing protocol deviations should be reviewed and confirmed before the database is locked.

The corresponding listing of subjects excluded from PPS due to important protocol deviations will be reported.

9 EVALUATION ENDPOINTS

9.1 Efficacy Endpoints

Primary endpoint:

- **Objective response rate (ORR):** the percentage of participants in the analysis population who have a complete response (CR) or a partial response (PR).

Secondary endpoints:

- **Progression free survival (PFS):** the time from the start of study treatment (the time of the first administration provided by the investigator) to the time of objective progressive disease or death due to any cause, whichever occurs first.
- **Duration of response (DoR)**: the time from first evidence of CR or PR until first evidence of disease progression or death due to any cause. The date of response is the date of first documented response rather than the date of confirmation.
- **Disease control rate (DCR)**: the percentage of participants in the analysis population who have CR, PR, or stable disease (SD).

9.2 Safety Endpoints

The following safety data will be collected according to the study protocol.

- Adverse event
- Laboratory test data
- Vital signs data
- 12-Lead ECG
- ECOG PS
- Physical examination

9.3 Other Endpoints

Not applicable.

9.4 Covariates

Not applicable.

9.5 Analysis Window

Data obtained from post-baseline visits will be summarized by protocol visits shown in eCRF, with no need to consider whether the visit window specified by the protocol has been exceeded.

When a planned visit is missing, or one of the test items in the planned visit is missing or its results are invalid, the unplanned visit closest to the planned visit will be considered as the protocol visit. If the time of two unplanned visits is equally close to the time of planned visit, the later unplanned visit will be selected as the protocol visit. All visits in the protocol will be sorted in chronological order. In addition, once defined as a protocol visit, an unplanned visit will no longer be defined as other protocol visits. In the analysis carried out by visits, the statistical analysis will be performed according to the planned time points in the protocol, i.e., the time points of unplanned visits do not need to be shown.

10 HANDLING OF MISSING VALUES

10.1 Handling of Missing Data of Tumor Assessment

When missing, the tumor assessment data will be treated according to RECIST (v1.1, refer to the table in appendix for details).

10.2 Handling of Missing Data of Safety Evaluation

Unless otherwise specified, missing safety endpoint values will not be imputed. Please refer to Appendix 1.4 for the imputing of date/time of adverse events and concomitant medications.

10.3 Handling of Missing Data of Tumor Biomarkers

Missing tumor biomarker data will not be imputed.

11 STATISTICAL ANALYSIS

The statistical analysis involved in this study will be carried out in accordance with cohorts. The statistical analysis plan must be finalized before database locking. Statistical tables, listings, and figures will be generated using SAS (v9.4).

In this study, unless otherwise stated, all data will be summarized using descriptive statistics in accordance with the following general principles.

Measurement data will be summarized by mean, standard deviation, median, maximum, and minimum. The decimal places of minimum and maximum should be consistent with those recorded on the CRF, mean and median have one more decimal place than those of the raw data, and the standard deviation has 2 more decimal places than that of the raw data. Count data will be summarized by frequency and percentage; categorical variables will be statistically described using the frequency and percentage of each category, with one decimal place rounded for percentages.

Time-to-event data will adopt the Kaplan-Meier method to estimate survival rates and plot survival curves; serum concentration data will be summarized using geometric mean, geometric standard deviation, geometric coefficient of variation, mean, standard deviation, coefficient of variation, median, maximum, and minimum.

11.1 Subject Distribution

The reasons for screening failure will be summarized by frequency and percentage, and the percentage will be calculated using the number of subjects subjected to screening as the denominator.

All enrolled subjects will be divided into cohorts and summarized by frequency and percentage. In addition, as for subject distribution, the following information will be summarized by frequency and percentage for cohorts:

- Number of subjects enrolled;
- Number of subjects withdrawn from the study;
- Number of subjects whose study is underway;
- Reason for treatment discontinuation;
- Reason for subject withdrawal after the end of study;

The percentages will be calculated using the number of all enrolled subjects as the denominator. In addition, the corresponding listing of subjects will be provided.

11.2 Demographics and Other Disease Characteristics

This analysis will be carried out based on FAS.

11.2.1 Demographics

Demographics (age, gender, ethnicity, height, body weight, BMI [kg/m^2]) and ECOG PS will be summarized and listed by cohort. The calculation of age (years) is based on the date of signing of the informed consent form and the date of birth.

The BMI calculation formula is: Baseline body weight (kg)/baseline height (m^2).

11.2.2 Tumor diagnosis

The descriptive statistics of baseline disease characteristics will be provided by cohort.

11.2.3 History of drug allergy

The descriptive statistics of allergy history data will be provided and listed by cohort.

11.2.4 Medical history

The descriptive statistics of medical history data will be provided and listed by cohort.

11.2.5 History of concomitant diseases

The descriptive statistics of medical history data will be provided and listed by cohort.

11.2.6 Tumor treatment history

The frequency and percentage of subjects with history of tumor surgery, history of radiotherapy, history of neoadjuvant chemotherapy, history of adjuvant chemotherapy, and other tumor treatment history will be summarized.

11.2.7 Medication history and concomitant medication

Prior medications are defined as drugs whose use has been completed before the first dose of the investigational drug during the treatment period. Concomitant medications are defined as drugs that are used after the first dose of the investigational drug, or drugs that are used before the first dose of the investigational drug but are continued to be used during the treatment period.

Medication history and concomitant medications will be descriptively summarized and listed by cohort.

11.3 Drug Compliance

The descriptive statistics of the exposure time, amount of exposure, number of treatment cycle, frequency of dose interruption, frequency and reason of dose discontinuation, frequency and reason of dose delay, final dosage, and reason for dose modification of study drugs (SHR-1210, oxaliplatin, apatinib, and capecitabine) will be summarized by cohort. In this study, SHR-1210 and oxaliplatin will be administered via intravenous drip infusion. The amount of apatinib and capecitabine dispensed and returned at each visit will be counted by tablet. The strength of apatinib is 375 mg/tablet and 250 mg/tablet, and that of capecitabine is 500 mg/tablet and 150 mg/tablet.

1) Actual cumulative drug exposure (mg)

SHR-1210:

Dose per cycle in study = 200 mg ×Actual dosing percentage (%)

Oxaliplatin:

Dose per cycle in study = Total dose of this dosing (mg) ×Actual dosing percentage (%)

Apatinib:

Dose per cycle in study = Total daily actual dose ×Number of days the corresponding dose is used (days)

Capecitabine:

Actual dose per cycle in study = Total daily actual dose ×Number of days the corresponding dose is used (days) = Total daily dose ×(Start date - End date + 1)

2) Scheduled dose (mg)

SHR-1210:

200 mg per cycle.

Oxaliplatin:

Scheduled dose per cycle = Total dose of this dosing (mg)

Apatinib:

Scheduled dose per cycle = Total dose of this dosing (mg)

Capecitabine:

Scheduled dose per cycle = Total daily scheduled dose ×Number of da ys the corresponding dose is used (days)

3) Subject compliance (%) = (Actual dose/Scheduled dose) ×100%.

All study drugs will be listed and analyzed by cohort, subject code, treatment stage, dosing date, and actual dose.

4) Duration of exposure to study drugs (weeks) = (Date of last dosing - Date of first dose + 1)/7.

Actual dose, scheduled dose, and compliance will be summarized by drug and cohort.

11.4 Safety Analysis

The safety analysis will be carried out based on the safety set by cohort.

11.4.1 Adverse event

An adverse event (AE) refers to any untoward medical occurrence in a study subject administered a medicinal product and which does not necessarily have a causal relationship with this treatment. In this study, AE should be collected from the signing of informed consent form (ICF) to 90 days after the last dosing of study drugs. AE includes any adverse symptom or sign, laboratory test abnormality, or disease. AE will be summarized separately in different stages of cohort 1 (SHR-1210 + capecitabine + oxaliplatin, sequence by SHR-1210 + apatinib mesylate, and total) and cohort 2.

AE will be analyzed based on treatment-emergent adverse event (TEAE). TEAE is defined as an AE that occurs or becomes aggravated on the same day or after the first dosing of study drugs but not later than 90 days after the last dosing of study drugs. Only TEAE will be summarized and analyzed. A separate data listing will be provided only for AE that occurs before treatment.

Each AE will be coded by MedDRA (v23.0) and graded by NCI-CTCAE v4.03. For the same SOC and/or PT, multiple cases of the same events that occur in one subject will be counted only once. For the same AE reported in one subject multiple times but varying in CTCAE grade, the greatest grade episode will be enumerated.

A summary description report will be provided for the number and percentage of subjects experiencing at least one TEAE, treatment-related TEAE (SHR-1210, oxaliplatin, capecitabine, and apatinib), severe TEAE (CTCAE Grade ≥ 3), TEAE leading to study drug discontinuation/interruption, treatment-related TEAE leading to drug discontinuation/interruption, TEAE leading to dose modification of study drugs, treatment-related TEAE leading to dose modification of the study drugs, TEAE leading to subject withdrawal from treatment, serious adverse event (SAE), and death.

In the TEAE summary table, AE occurring after the study drug treatment will be analyzed by MedDRA system organ class (SOC) and preferred term (PT), and descriptive statistics will be provided by frequency and percentage. The summary table contains the number of subjects experiencing at least one TEAE, TEAEs leading to treatment withdrawal, SAEs, TEAEs related to each study drug, and CTCAE Grade ≥ 3 treatment-related TEAE, along with CTCAE grades. After study drug treatment, each TEAE, SAE, TEAE of special interest, immune-related TEAE, and possible liver injury event, along with CTCAE grade, in each cohort will be summarized by SOC and PT, and each treatment-related TEAE and SAE related to study drugs will be summarized by SOC and PT. Immune-related AE will also be summarized by immune-related category and PT. The incidence of AE is calculated based on the number of subjects having an AE, instead of the number of AE episodes. For TEAE occurring after the start of study treatment, causality with study drugs consists of definitely related, possibly related, and indeterminable; if causality is missing, the TEAE will be considered drug-related for analysis. TEAE with CTCAE grade missing will be counted as missing.

TEAE of special interest includes the following:

- Grade ≥ 3 infusion reaction;
- Grade ≥ 2 diarrhea/colitis, uveitis, and interstitial pneumonia;
- Other Grade ≥ 3 immune-related AE;
- Grade 4 amylase or lipase increased.

Each AE, TEAE leading to dose interruption or modification, TEAE leading to treatment withdrawal, severe TEAE, and death will be listed.

For immune-related TEAE (classified by immune-related TEAE), the descriptive summary of the number and percentage of subjects have received corticosteroids or immunomodulators, received immunomodulators, received corticosteroids (further divided into systemic and topical administration), and received high-dose corticosteroids (defined as at least one dose or daily dose of ≥ 40 mg prednisone or equivalent; further divided into systemic and topical administration), and the starting dose and duration of high-dose corticosteroids will also be provided. The dosage of different hormonal drugs will be converted into the equivalent dose of prednisone according to the following formula: 0.75mg of dexamethasone = 5 mg of prednisone = 4mg of methylprednisolone = 20mg of hydrocortisone = 0.8mg of betamethasone.

If one subject experiences one AE and receives two or more hormones during the presence of the AE, the starting dose of the hormone used first will be involved in the statistics of starting doses (i.e., the starting dose(s) of subsequent hormone therapy/therapies will not be involved in the statistics of starting doses). Then, the cumulative duration of the use of multiple hormones (time overlapping of hormone therapies should be avoided in the calculation) will be summarized as the duration of dosing. If one AE category (such as the category of immune-related AE) contains more than one AE (by PT), this AE category will be treated as one AE when being summarized. When the initial dose is summarized, if the dosing frequency of hormones is more than once a day, it will be converted into a single daily dose before being summarized. When the duration of dosing is summarized, if the end date of hormone drug treatment is unknown, it will be imputed with the start date of the last dose of the hormone drug treatment.

Corrective treatment, site, maximum diameter, incidence of hemorrhage, and incidence of infection will also be summarized for reactive capillary endothelial proliferation.

11.4.2 Laboratory test

The baseline laboratory test is defined as the most recent test before the first administration.

The quantitative parameter of hematology, blood biochemistry, and urinalysis in each cohort at baseline, the laboratory test result after treatment, the change from the baseline, and the normal/abnormal change from the baseline will be summarized by visit. The hematology, blood biochemistry, and urinalysis of each cohort will be summarized by the normal/abnormal changes before and after treatment.

All laboratory test results will be listed by subject ID, among which, abnormal values will be marked and indicated for clinical significance (if any).

11.4.3 Vital signs

The vital sign baseline is defined as the last sign result tested before the first administration.

The vital signs of each cohort will be summarized by visit using descriptive statistics. At the same time, the changes from baseline will also be summarized and analyzed by descriptive statistics.

All vital signs will be listed.

11.4.4 Physical examination

The physical examination baseline is defined as the last examination result before the first administration.

The physical examination will be summarized by the normal/abnormal changes before and after treatment.

The abnormal results of physical examination will be listed.

11.4.5 12-Lead ECG

Baseline is defined as the most recent test data before the first dose.

Items concerning ECG include: HR (times/min), PR interval (ms), QT interval (ms), and QTc (ms). The measured value and change from baseline of each ECG variable at each time point will be summarized using descriptive statistics. At the same time, ECG will be summarized by the normal/abnormal changes before and after treatment.

If QTcF is not directly provided on the ECG report, it will be obtained using the following formula (Fridericia formula):

$$QTcF = QT/(RR^{1/3})$$

RR = 60/heart rate (s)

Related data will also be reported in the form of listing.

11.4.6 ECOG PS

Baseline is defined as the most recent test data before the first dose.

The ECOG PS of each cohort will be analyzed by descriptive statistics.

All ECOG PS will be reported in the form of listing.

11.4.7 Other safety data

Examinations such as echocardiography, virology, fecal occult blood, pregnancy, and coagulation function will be listed.

11.5 Efficacy Analysis

The tumor assessment at each time point will be conducted by the investigator as per RECIST (v1.1). The descriptive statistics of ORR will be summarized by stages (Stage I: 19 subjects respectively in cohort 1 and cohort 2 who have completed the response evaluation; Stage II: the remaining subjects who have completed the response evaluation) and cohorts, while PFS, DoR, and DCR will be only summarized by cohort. The efficacy analysis will be carried out based on FAS. If a subject's baseline or post-baseline tumor assessment is missing, the subject's is non-evaluable (NE) in response. For the maximum post-baseline change in the sum of the longest diameter of the target lesion, a waterfall plot will be drawn by cohort; for the post-baseline change in each visit regarding the sum of longest diameter of the target lesion relative to the baseline, a spider chart will be drawn by cohort.

11.5.1 Primary efficacy endpoint analysis

11.5.1.1 Best overall response (BOR)

BOR is defined as the best response recorded from the start of the treatment until objective recording of progressive disease (PD) by RECIST (v1.1) or start of subsequent anti-cancer therapy (whichever occurs first). For subjects without documented progression or new anti-cancer treatment, BOR will be determined based on all response evaluations. BOR recorded during this period will be based on the confirmed response. For the criteria for response confirmation, see Appendix 1.3.

- Confirmed CR: Response evaluated as CR and reconfirmed as CR at a later time point (at least 4 weeks later and before PD).
- Confirmed PR:
  - Response evaluated as CR but reconfirmed as PR (CR not reached) at a later time point (at least 4 weeks later and before PD).
  - Response evaluated as PR but reconfirmed as CR/PR at a later time point (at least 4 weeks later and before PD).
- Confirmed SD:
  - Response evaluated as CR/PR but reconfirmed as SD (CR/PR not reached).
  - Response evaluated as SD at least once within 5 weeks after the start of the study (visit window: 6 weeks - 7 days).
- PD: Progressive disease not meeting the requirements for confirmed CR, PR, and SD
- NE: Other conditions

A response will be judged as PR, if a patient has two non-consecutive results of PR, the interval between the two is more than four weeks, and no PD occurs during the interval. Similarly, a response will be judged as CR, if a patient has two non-consecutive results of CR, the interval between the two is more than four weeks, and no PD occurs during the interval.

CR or PR occurring after the start of subsequent anti-cancer treatment will not be included in the calculation of BOR.

BOR (CR, PR, SD, PD, and NE) will be analyzed by frequency and percentage.

11.5.1.2 Objective response rate (ORR)

ORR is defined as the proportion of subjects whose BORs have been confirmed as CR or PR by RECIST (v1.1).

When the first response evaluation data of a subject after treatment are missing or unknown, the subject is not evaluable (NE) and will only be included in the denominator for percentage calculation.

ORR of a single cohort and its two-sided 95% exact (Clopper-Pearson) confidence interval will be calculated.

The confirmed CR, PR, SD, and PD will be listed and summarized.

The analysis for ORR will be repeated based on FAS, PPS, and ES.

11.5.2 Secondary efficacy endpoint analysis

11.5.2.1 Disease control rate (DCR)

DCR is defined as the percentage of subjects in the analysis population who have BOR evaluation of CR, PR, or SD.

When the response evaluation data of a subject after treatment are missing or unknown, the subject is not evaluable (NE) and will be included in the denominator for percentage calculation.

DCR and its two-sided 95% exact (Clopper-Pearson) confidence interval will be calculated.

11.5.2.2 Duration of response (DoR)

DoR is defined as the time from first documented evidence of CR or PR until the date of first documented evidence of progressive disease or death due to any cause. The end date of response must be consistent with the date of progressive disease or death of PFS. Tumor response is based on the confirmation, and the response that has not been confirmed will not be included. The date of response is the date of first observation (the last date of imaging assessment of visit, not the date of visit), not the date of confirmation. Table 2 describes information on progressive disease, death, or censoring.

For DoR, survival curves will be plotted using the Kaplan-Meier method and the 95% CI for median time will be calculated using the Brookmeyer-Crowley method. In addition, a swimmer plot will be drawn for each cohort.

Way to interpret Table 2: From the first question until the occurrence of progressive disease, death or censoring, and the corresponding date.

Table 2. Progressive disease and censoring of the duration of response

		No	Yes
1.	Has the subject's response evaluation been confirmed as CR PR?	Not included in the analysis	Skip to 2
2.	Are new anti-cancer drugs used before progressive disease or death?	Skip to 3	Censoring The date of censoring is the date of the last objective tumor assessment before the use of the new anti-cancer drug
3.	Has the subject experienced progressive disease?	Skip to 5	Skip to 4
4.	Is the progressive disease of the subject recorded after an unacceptably long time interval (the response evaluation is missing by > 2 times, i.e., 98 days)?	The end date of response is the date of progressive disease	Censoring The date of censoring is the date of the last objective tumor assessment before the long-term lost to follow-up
5.	Is the subject dead?	Censoring The date of censoring is the date of the last objective tumor assessment	Skip to 6
6.	Is the death of the subject recorded after an unacceptably long time interval (the response evaluation is missing by > 2 times, i.e., 98 days)?	The end date of response is the date of death	Censoring The date of censoring is the date of the last objective tumor assessment before the long-term lost to follow-up

11.5.2.3 Progression free survival (PFS)

PFS refers to the time from the start of randomization to the date of objective progressive disease or death (caused by any reason), whichever occurs first. Table 3 describes progressive disease, death, or censoring.

Table 3. Date of progressive disease, death, or censoring

		No	Yes
1.	Does the subject have a baseline objective tumor assessment?	Censoring The date of censoring is the date of randomization	Skip to 2
2.	Are new anti-cancer drugs used before progressive disease or death?	Skip to 3	Censoring The date of censoring is the date of the last objective tumor assessment before the use of the new anti-cancer drug
3.	Has the subject experienced progressive disease?	Skip to 5	Skip to 4
4.	Is the progressive disease of the subject recorded after an unacceptably long time interval (the response evaluation is missing by > 2 times, i.e., 98 days)?	Progressive disease The date of progression is the date of earliest recorded progressive disease	Censoring The date of censoring is the date of the last objective tumor assessment before the long-term lost to follow-up
5.	Is the subject dead?	Censoring The date of censoring is the date of the last objective tumor assessment	Skip to 6
6.	Is the death of the subject recorded after an unacceptably long time interval (the response evaluation is missing by > 2 times, i.e., 98 days)?	Progressive disease The date of progression is the date of death	Censoring The date of censoring is the date of the last objective tumor assessment before the long-term lost to follow-up

The time of PFS is based on the time of the imaging examination rather than the time of the visit. The RECIST imaging examination time (targeted lesion, non-targeted lesion, and new lesion) for a special visit may be different, and the following rules apply:

- The date of progression will be based on the earliest date leading to progression.
- The date of censoring will be based on the last date of objective tumor assessment.

Survival curves will be plotted using the Kaplan-Meier method and the 95% CI for median time will be calculated using the Brookmeyer-Crowley method.

11.5.3 Exploratory efficacy analysis

11.5.3.1 Overall survival (OS)

OS is defined as the time from the start of randomization to the death of the subject caused by any reasons. Survival curves will be plotted using the Kaplan-Meier method and the 95% CI for median time will be calculated using the Brookmeyer-Crowley method. The 6-month, 9-month, and 12-month OS rates and their 95% CIs in the two cohorts will be estimated.

Censoring rules: For subjects who are still alive at the last follow-up, their OS will be censored at the date of last follow-up. For subjects who are lost to follow-up, their OS will be censored at the last confirmed survival time before lost to follow-up. The censored OS is defined as the time from randomization to censoring. OS will be calculated in months using the following formula: (Date of OS event or censored date - date of randomization + 1)/30.4375.

11.5.3.2 Time to objective response (TTOR)

TTOR is defined as the period of time from the start of randomization to the first documented BOR of objective response (as per RECIST 1.1).

TTOR will only be measured in subjects with a BOR after treatment. The date of response is the date of first documented response but not the date of confirmation. TTOR will be statistically described by the number of subjects, mean, standard deviation, median, minimum, and maximum.

11.5.3.3 Subgroup analysis of ORR, PFS, and OS

The subgroup analysis is exploratory and will only be performed based on FAS. This study includes the following subgroups:

- Length of chemotherapy cycles (≤ 4 cycles vs. > 4 cycles) (only in cohort 1)
- Age (< 60 vs. ≥ 60)
- Gender (male vs. female)
- Baseline ECOG PS score (0 vs. 1)
- Peritoneal metastasis (yes vs. no)
- Liver metastasis (yes vs. no)
- Metastasis sites (≤ 2 vs. > 2)
- History of radical gastrectomy (yes vs. no)
- Lesion site (stomach vs. gastroesophageal junction)

In order to explore the impact of subgroup factors on the ORR, the ORR and its two-sided 95% exact CI (Clopper-Pearson) in each subgroup of a single cohort will be calculated. The PFS and OS for each subgroup of a single cohort will be plotted using the Kaplan-Meier method and the 95% CI for median time will be calculated using the Brookmeyer-Crowley method.

11.6 Statistical Analysis of PK

Descriptive statistical analysis will performed for the PK drug concentrations of SHR-1210 and apatinib.

11.7 Statistical Analysis of PD

Not applicable

11.8 Statistical analysis of immunogenicity (ADA)

The number and percentage of subjects with positive ADA and negative ADA will be summarized using descriptive statistics.

For the population with positive ADA, the following items will be analyzed: the onset of ADA positive, the number and percentage of subjects with transient positive post-baseline results (transient ADA positive is defined as the subjects who are negative at baseline and have only one ADA-positive sample post baseline before the last sample and are not persistent ADA positive), the number and percentage of subjects with persistent positive post-baseline results (persistent ADA positive can be divided to two categories: 1) negative at baseline, there are at least two positive samples post baseline, where the first and last ADA-positive samples are separated by a period >16 weeks, or 2) negative at baseline, the last post-baseline sample is positive, or the last post-baseline sample is negative and the second last post-baseline sample is positive), and the number and percentage of subjects with other positive post-baseline results (other ADA positive is defined as the subjects who are negative at baseline, have at least two ADA positive samples post baseline, but not persistent ADA positive, and the last sample is negative).

In the meantime, the safety of ADA-positive and negative subjects will be summarized so as to investigate the effects of ADA on safety. ADA-positive subjects refer to subjects with post-baseline ADA-positive samples, mainly including: a: the subjects who have the baseline negative and post-baseline positive samples, or b. the subjects who have both baseline and post-baseline positive samples, and the titer of the post-baseline samples is ≥ 4-fold of the baseline titer. ADA-negative subjects are those who are not ADA positive.

The number and percentage of subjects with neutralizing antibodies (Nabs) positive will be summarized using descriptive statistics.

Immunogenicity raw data will be reported in the form of listing.

12 APPENDICES

Appendix 1.1 Time Point Response: Subjects with Target Lesions (Including or Excluding Non-Target Lesions)

Target Lesion	Non-Target Lesion	New lesion	Overall Response at the Time Point
CR	CR	Non	CR
CR	Non-CR/Non-PD	Non	PR
CR	Not evaluable	Non	PR
PR	Non-PD or not all evaluable	Non	PR
SD	Non-PD or not all evaluable	Non	SD
Not all evaluable	Non-PD	Non	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD

CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease, and NE = not evaluable

Appendix 1.2 Time Point Response - Subjects with Non-target Lesions Only

Non-Target Lesion	New lesion	Overall Response
CR	Non	CR
Non-CR/Non-PD	Non	Non-CR/Non-PD[a]
Not all evaluable	Non	Not evaluable
Equivocal PD	Yes or No	PD
Any	Yes	PD

a: "Non-CR/non-PD" is preferred over SD for non-target disease. Since SD is increasingly used as an endpoint for response evaluation, non-CR/non-PD response is developed to address the absence of lesion measurability.

Appendix 1.3 Best Overall Response with CR and PR to Be Confirmed

Evaluation at the Previous Time Point	Evaluation at the Subsequent Time Point	BOR
CR	CR	CR
CR	PD, missing	SD (> 6 weeks from the start of treatment (5 weeks after subtracting 1 week of visit window))
CR	PR or SD[a]	Note: This situation generally will not occur. ● If the last evaluation is CR, then a little bit of the lesion in the next evaluation will lead to the response evaluation of PD. The BOR depends on whether the minimum time of 6 weeks for SD is met. If met, it is SD.

Evaluation at the Previous Time Point	Evaluation at the Subsequent Time Point	BOR
PR	PR, unconfirmed CR	PR
PR	SD, missing	SD
PR	PD	SD (> 6 weeks from the start of treatment (5 weeks after subtracting 1 week of visit window))
SD	SD, unconfirmed PR, unconfirmed CR	SD
SD	NE –> SD	SD
PR	NE –> PR	PR
NE	NE	NE

CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease, and NE = not evaluable.

a: If a CR is truly met at first time point, then any disease seen at a subsequent time point, even the disease meeting PR criteria relative to baseline, makes the disease PD at that point (since disease must have reappeared after CR). Best overall response will depend on whether minimum duration for SD is met. However, sometimes CR may be claimed when subsequent scans suggest small lesions are likely still present and in fact the subject have PR, not CR at the first time point. Under these circumstances, the original CR should be changed to PR and the best overall response is PR.

Appendix 1.4 Rules for Imputing Incomplete Dates of AE and CM:

Date	Start Date	End Date
Missing "Day"	Impute with the start day of study treatment when the year and month of the occurrence are identical to those of study treatment; otherwise, impute with the first day of the corresponding month	Impute with the last day of the month
Missing "Month" and "Day"	Impute with the start month and day of study treatment when the year of the occurrence is identical to that of study treatment; otherwise, impute with "1 Jan."	31 Dec.
Missing "Year", "Month", and "Day"	No imputation	No imputation

Note: All imputed dates must be before the date of end/termination of study and date of death.