Statistical Analysis Plan

Study Title: Prospective Single-Arm Feasibility Clinical Study of Huangqi Yangyuan Formula for Adverse Reactions Induced by Capecitabine in Patients with Advanced Colorectal Cancer

Study Sponsor: Hebei Provincial Hospital of Traditional Chinese Medicine

Ethics Committee: Ethics Committee of Hebei Provincial Hospital of Traditional Chinese Medicine

SAP Version: V1.0

Version Date: 2026-08-31

Target Sample Size: 35 subjects

Study Design: Prospective, single - arm, single - center feasibility exploratory study

Document note: This statistical analysis plan shall be finalized and signed prior to database lock. Any modifications shall be implemented via formal SAP amendment. This study has no concurrent control group; all efficacy - related analyses are exploratory descriptive analyses rather than confirmatory hypothesis testing.

Table of Contents

1. Introduction

2. Study Objectives

3. Study Population

4. Analysis Datasets

5. Baseline Characteristics Analysis

6. Primary Feasibility Endpoint Analyses

7. Exploratory Anti-Tumour Outcome Analyses

8. Safety and Adverse Event Analyses

9. Exploratory Biomarker (T-lymphocyte Subsets) Analyses

10. Quality of Life and TCM Constitution Scale Analyses

11. Post-chemotherapy 3-Month Follow-up Analyses

12. Missing Data Management

13. Statistical Methods and Software

14. Statistical Inference Rule

15. Protocol Deviation and SAP Amendment Rules

16. Appendices

1. Introduction

This is a prospective single-arm feasibility exploratory study conducted in Hebei Provincial Hospital of Traditional Chinese Medicine, enrolling 35 patients with advanced colorectal cancer receiving capecitabine monotherapy. No concurrent blank control group is set in this trial.

The primary purpose of this trial is to assess the clinical implementation feasibility of capecitabine combined with Huangqi Yangyuan Formula. Secondary and exploratory objectives are to preliminarily observe potential signals of alleviating capecitabine-related adverse reactions, changes in patient quality of life, TCM constitution status and peripheral blood T-lymphocyte subsets, so as to provide evidence basis for subsequent larger-sample clinical trials.

This Statistical Analysis Plan (SAP) prespecifies all analytical contents before database lock. Post-hoc hypothesis testing is not permitted without formal SAP amendment.

2. Study Objectives

2.1 Primary Objective

To evaluate the clinical feasibility of capecitabine combined with Huangqi Yangyuan Formula for advanced colorectal cancer patients receiving capecitabine monotherapy.

Feasibility core indicators include chemotherapy completion rate,

Huangqi Yangyuan Formula intervention compliance rate and subject retention rate.

2.2 Secondary / Exploratory Objectives

(1) To describe the incidence and grade distribution of capecitabine-related adverse reactions under combined treatment regimen.

(2) To describe changes of EORTC-QLQ-C30 quality-of-life scale scores at baseline, the end of Cycle 4 and the end of Cycle 8.

(3) To describe changes of TCM BCQ constitution scale scores across study time points.

(4) To describe changes of peripheral blood T-lymphocyte subsets ($CD3^+$, $CD4^+$, $CD8^+$, $CD4^+/CD8^+$ ratio, etc.) at baseline, the end of Cycle 4 and the end of Cycle 8.

(5) To describe tumour-related events within 3-month follow-up after chemotherapy completion, including tumour progression, recurrence and metastasis.

Important note: This study is not designed to confirm definite anti-tumour or toxicity-mitigating efficacy of Huangqi Yangyuan Formula. All efficacy-related results are for exploratory signal observation only.

3. Study Population

3.1 Inclusion and Exclusion Criteria

Consistent with inclusion criteria and exclusion criteria specified in study

informed consent form and clinical research protocol.

3.2 Sample Size Justification

Thirty - five (35) subjects will be enrolled. Since this is a feasibility exploratory single-arm study, sample size is set for feasibility assessment, and no formal statistical power calculation is performed.

4. Analysis Datasets

4.1 Full Analysis Set (FAS)

All subjects who have signed informed consent and received at least one dose of capecitabine or Huangqi Yangyuan Formula. FAS is the primary dataset for feasibility assessment, safety evaluation and exploratory outcome analysis.

4.2 Per-Protocol Set (PPS)

Subjects from FAS satisfying all below conditions:

- Meet all inclusion criteria and have no major exclusion criteria after enrolment;
- Complete no less than 70% of the planned study interventions;
- Without major protocol violations.

PPS will be used for sensitivity analysis of exploratory endpoints.

4.3 Safety Analysis Set (SAS)

Identical to FAS. All subjects receiving at least one study - related treatment will be included for adverse event and safety analysis.

Pre-defined Major Protocol Violations

(1) Violation of key inclusion or exclusion criteria after subject enrolment;

(2) Failure to complete required 2 - week wash - out period for anti - tumour Chinese medicine before enrolment;

(3) Use of prohibited anti - tumour concomitant therapies during study treatment phase;

(4) Loss of both baseline and post-baseline core assessment data.

5. Baseline Characteristics Analysis

Baseline demographic and clinical characteristics will be summarized based on FAS dataset.

- Continuous variables: sample size (n), mean ± standard deviation, median, interquartile range (Q1-Q3), minimum and maximum value.

- Categorical variables: count and percentage (%).

Baseline analysis items include but are not limited to:

(1) Demographics: age, gender;

(2) Tumour characteristics: tumour stage (ⅢB/Ⅳ), primary tumour site, history of radical surgery;

(3) ECOG performance status score;

(4) Baseline laboratory tests: liver function, renal function, bone marrow function, tumour markers;

(5) Baseline scales: EORTC-QLQ-C30 scores, TCM BCQ constitution scale scores;

(6) Baseline immune biomarkers: peripheral blood T - lymphocyte subsets.

6. Primary Feasibility Endpoint Analyses

All analyses are performed on FAS dataset.

(1) Chemotherapy completion rate: The proportion of subjects who finish all planned 8 cycles of capecitabine. Numerator = number of subjects completing 8 cycles; denominator = total enrolled subjects. Reasons for non-completion (adverse reaction, disease progression, subject voluntary withdrawal and other causes) shall be tabulated.

(2) Huangqi Yangyuan Formula compliance rate: Proportion of subjects achieving ≥ 80% of prescribed TCM decoction intake during treatment period, presented as count and percentage.

(3) Subject retention rate: Proportion of subjects finishing all prespecified study assessments (baseline, end of Cycle 4, end of Cycle 8 and 3-month follow-up). Drop-out and withdrawal reasons shall be recorded.

7. Exploratory Anti-Tumour Outcome Analyses

Only descriptive statistics will be applied; no formal statistical

comparison test will be conducted.

(1) Tumour responses assessed by contrast-enhanced CT at Cycle 4 and Cycle 8 will be classified as CR, PR, SD, PD according to RECIST 1.1 criteria. Count and percentage for each response category will be presented.

(2) Within 3-month follow-up after chemotherapy completion: count and percentage of tumour progression, recurrence and metastasis events will be summarized.

8. Safety and Adverse Event Analyses

All safety analyses are based on SAS dataset. Adverse events will be coded using CTCAE Version 5.0.

(1) Summarize all treatment-emergent adverse events (TEAEs): number and percentage of subjects with at least one TEAE; stratified by system-organ class, preferred term and maximum adverse event grade.

(2) Summarize capecitabine-specific adverse reactions, including hand-foot syndrome, diarrhoea, nausea and vomiting, myelosuppression, oral mucositis, liver-renal function abnormality; report incidence and grade distribution.

(3) Summarize adverse events potentially related to Huangqi Yangyuan Formula, including abdominal distension, diarrhoea, constipation

and liver-renal function abnormality.

(4) Summarize serious adverse events (SAEs): case - wise description, causality assessment (definitely related / probably related / possibly related / unrelated to study interventions).

(5) Summarize capecitabine modification circumstances: frequency of dose reduction, cycle delay and treatment discontinuation caused by adverse reactions.

9. Exploratory Biomarker (T-lymphocyte Subsets) Analyses

Peripheral blood T-lymphocyte subset indicators (percentage of $CD3^+$, $CD4^+$, $CD8^+$ cells, $CD4^+/CD8^+$ ratio) are measured at baseline, end of Cycle 4 and end of Cycle 8.

- Descriptive statistics (mean, median, interquartile range) will be provided for each time point.

- Within-subject changes from baseline at Cycle 4 and Cycle 8 will be described.

- No formal hypothesis testing will be performed; results are only for exploratory signal observation.

10. Quality of Life and TCM Constitution Scale Analyses

EORTC-QLQ-C30 scale and TCM BCQ constitution scale are collected at baseline, end of Cycle 4 and end of Cycle 8.

- Domain and item scores at each time point will be summarized with mean, median and interquartile range.

- Subject-level change-from-baseline values at Cycle 4 and Cycle 8 will be calculated.

- Missing scale items shall be handled following official scoring manuals of corresponding scales.

11. Post-chemotherapy 3-Month Follow-up Analyses

After completion of 8-cycle chemotherapy, 3-month follow-up data contain adverse events, physical condition, tumour recurrence and metastasis status.

- Categorical outcomes will be presented as count and percentage.

- Time-to-tumour-progression data will be described descriptively. Survival-related formal model analysis will not be implemented given small sample size and single-arm design.

12. Missing Data Management

(1) Baseline variables: missing baseline values will not be imputed; missing rate shall be reported.

(2) Repeated-measurement indicators (laboratory indices, scales, biomarkers): Multiple imputation is not adopted as primary analysis method. Missing proportion for each time-point will be reported.

Last Observation Carried Forward (LOCF) will only be used for supplementary sensitivity descriptive reference, not for primary result conclusion.

(3) Feasibility endpoints: Subjects lost to follow-up are counted as non-completion / drop-out in primary analysis.

(4) Adverse event data: All available safety data will be analysed. Missing adverse event information will be documented in clinical study report.

This trial is feasibility-oriented exploratory research. Imputation shall not generate virtual observed data for core study conclusions.

13. Statistical Methods and Software

- Continuous variables: descriptive statistics including mean ± standard deviation, median, interquartile range, minimum and maximum. Within-subject changes over time will be described; no inter-group comparison as there is no control arm.

- Categorical variables: frequency count and percentage.

- Statistical software: SPSS 26.0 or above, or R software.

14. Statistical Inference Rule

This is a single-arm feasibility study focusing on descriptive statistics. No primary confirmatory hypothesis testing is planned. Any P - values

generated in exploratory analyses are for reference only, without pre-defined alpha level for confirmatory inference. Observed changes cannot establish causal relationship between combined intervention and study outcomes.

15. Protocol Deviation and SAP Amendment Rules

(1) Any deviation from this SAP must be documented via formal SAP amendment, with clear date and signature, completed before database lock.

(2) Exploratory analyses performed after database lock without formal SAP amendment shall be clearly labelled as post-hoc exploratory analyses in clinical study report.

(3) All protocol violations shall be listed in clinical study report; their impacts on each analysis dataset shall be assessed.

16. Appendices

- Appendix A: CTCAE Version 5.0 Adverse Event Coding Criteria

National Cancer Institute(NCI). Common Terminology Criteria for Adverse Events (CTCAE) Version 5.0. Published November 27, 2017.

https://dctd.cancer.gov/research/ctep-trials/for-sites/adverse-events

- Appendix B: RECIST 1.1 Tumour Response Evaluation Criteria

Citation: Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). European Journal of Cancer. 2009;45(2):228 - 247. DOI:10.1016/j.ejca.2008.10.026 .

Official PDF source: https://ctep.cancer.gov/protocoldevelopment/docs/recist_guideline.pdf

Core definitions for target lesions

- Complete Response (CR): Disappearance of all target lesions. Any pathological lymph-nodes (target or non-target) must have short-axis < 10 mm .

- Partial Response (PR): At least 30 % decrease in sum of longest diameters of target lesions, referenced against baseline sum of diameters .

- Stable Disease (SD): Neither sufficient shrinkage to meet PR criteria nor sufficient increase to meet PD criteria, referenced against the smallest sum recorded during study (nadir) .

- Progressive Disease (PD): $\geq$ 20 % increase in sum of target-lesion longest diameters compared with the nadir value; accompanied by an absolute increase $\geq$ 5 mm. Appearance of any new malignant lesion also counts as PD .

Note: This study uses contrast-enhanced chest-abdomen-pelvis CT for

tumour assessment, following RECIST 1.1 measurement rules. Maximum 5 target lesions in total, ≤ 2 target lesions per organ. Non-target lesions will be qualitatively documented.

- **Appendix C:** EORTC-QLQ-C30 Official Scoring Guideline

Citation: Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A, on behalf of the EORTC Quality of Life Group. The EORTC QLQ-C30 Scoring Manual (3rd Edition). Brussels: European Organisation for Research and Treatment of Cancer (EORTC); 2001.

Official manual source: https://www.eortc.org/app/uploads/sites/2/2018/02/SCmanual.pdf

Brief scale introduction & scoring rules

EORTC-QLQ-C30 contains total 30 items, including 5 functioning scales, 1 global health / quality-of-life scale, 3 symptom scales, and 6 single-item symptom measures .

- Raw scores are linearly transformed to 0-100 standardized scale scores following official scoring manual.

- For functioning scales and global health status: higher score = better function / better quality of life.

- For symptom scales and single-item symptoms: higher score = more severe symptom burden .

- Missing items are processed strictly in accordance with the scoring manual; scale will not be calculated if more than half items within that scale are missing.

- Assessments are performed at baseline, end of Cycle 4 and end of Cycle 8. Only descriptive statistics will be presented; no formal hypothesis testing.

- **Appendix D:** TCM BCQ (Body-Constitution Questionnaire) Constitution Scale Scoring Instruction

Citation: Su Y-C, et al. BCQ+: a body constitution questionnaire to assess Yang - Xu. Complementary Medicine Research, 2008,15(6):327 - 334. DOI:10.1159/000175938.

Brief note:

BCQ is a published, reliability-validated TCM self-report questionnaire covering three constitution sub-dimensions: Yang-Qi deficiency, Yin-Blood deficiency and Phlegm-stasis. It will be administered at baseline, end of Cycle 4 and end of Cycle 8 in this study. Only descriptive statistics will be performed, no hypothesis-testing will be conducted.

Scoring rule: Raw sub-scale score is calculated by summing item scores; higher score indicates more prominent corresponding constitution bias. If more than half items within one sub-scale are missing, that sub-dimension will not be calculated, and no data imputation will be

performed.