Research Title: Development and Prospective Validation of an AI-Based Diagnostic Model for Hepato-Pancreato-Biliary Diseases

NCT Number: Under Review

Document date: 2026/2/26

Study Protocol

1. Research Background

With the rapid advancement of artificial intelligence (AI) technology, its applications in healthcare have expanded significantly, particularly in diagnostic assistance, intelligent triage, and patient interaction. Hepatobiliary and pancreatic diseases (e.g., liver cancer, pancreatic cancer, cholangiocarcinoma, cirrhosis) are characterized by insidious onset, rapid progression, low early diagnosis rates, and poor prognosis, necessitating high demands for early identification and precise clinical evaluation. However, in China's primary healthcare institutions, challenges such as insufficient specialist resources, variable patient health literacy, and incomplete initial data collection contribute to high risks of missed or incorrect diagnoses, delaying critical intervention opportunities.

Recent breakthroughs in large language models (LLMs) in natural language understanding and generation provide a technical foundation for developing intelligent, personalized medical dialogue systems. Multi-agent systems (MAS) simulate collaborative roles to enable complex, structured decision-making, offering advantages in task decomposition, knowledge integration, and dynamic reasoning. Integrating MAS with medical LLMs could create a specialized intelligent pre-consultation system for hepatobiliary and pancreatic diseases, systematically collecting and analyzing key patient information (e.g., symptoms, risk factors, family history, lifestyle) to improve consultation efficiency and completeness.

This study aims to develop a hepatobiliary-pancreatic disease intelligent pre-consultation system based on a MAS architecture. By designing specialized agents (e.g., "Guidance Agent," "Medical History Agent," "Risk Assessment Agent," "Physician Summary Agent"), the system will simulate clinical reasoning to guide patients in structured symptom reporting and generate standardized preliminary reports for physician reference. The system seeks to enhance early disease recognition by primary care physicians, optimize outpatient resource allocation, and improve patient experiences.

2. Research Objectives

The study aims to design, develop, and preliminarily validate a MAS-based intelligent pre-consultation system for hepatobiliary-pancreatic diseases, exploring its feasibility and potential clinical value. Specific objectives include:

Developing a Specialized MAS Framework: Design a collaborative MAS with 分工 (division of labor) capabilities (guidance, history-taking, risk screening, report generation) to simulate clinical reasoning and systematically collect hepatobiliary-pancreatic symptom data.

Improving Pre-Consultation Completeness and Accuracy: Use AI-driven dialogue to guide patients in recalling critical risk factors (e.g., jaundice, abdominal pain, weight loss, alcohol history, viral hepatitis, family cancer history), reducing information omissions due to memory bias or poor articulation.

Generating Structured Preliminary Reports: Automatically synthesize patient complaints and medical history into standardized summaries highlighting key signs and risk alerts, aiding physicians in rapid case comprehension and shortening consultation time.

Laying Foundations for Clinical Validation: Conduct small-scale simulations and expert evaluations to verify usability, gather feedback, optimize workflows, and provide evidence for future prospective studies and ethical compliance.

The system is not intended for direct diagnosis/treatment but as a pre-consultation tool to advance safe, equitable AI applications in early hepatobiliary-pancreatic disease screening and tiered healthcare delivery.

3. Study Design and Methods

Design: Prospective, multi-arm, non-interventional study comparing AI-assisted and traditional consultation pathways, with a gold-standard reference.

Study Duration: 9 months (start date: upon IRB approval; end: September 30, 2026).

Inclusion Criteria:

Age ≥18 years;

Scheduled inpatient admission (non-emergency);

Willingness to participate and operate AI system;

Proficiency in Mandarin/English.

Exclusion Criteria:

Critical conditions; cognitive/communication barriers; concurrent trials; pregnancy; refusal of consent.

Withdrawal Criteria: Patient withdrawal, adverse events, protocol deviations, or unrecognized exclusions.

Group Allocation:

Experimental Group 1: AI consultation followed by blinded physician consultation (AI output not shared).

Experimental Group 2: AI consultation followed by physician consultation with AI-generated summary (excluding diagnostic advice).

Control Group: Traditional physician consultation without AI.

Safety Mechanisms: Multi-tiered risk monitoring (e.g., alerts for "hematemesis" or "severe allergy") with mandatory physician review.

Sample Size: 400 patients (120 per group + 10% dropout buffer) to detect 8% improvement in accuracy (α=0.05, 80% power).

Gold Standard: Expert committee-generated reference records for unbiased evaluation.

Outcome Measures:

Algorithm Performance: Accuracy, omission rate, diagnostic Kappa.

Clinical Utility: Consultation time, physician workload (NASA-TLX), patient satisfaction.

System Stability: Availability (≥99%), latency (<2s), crash rate (<0.5%).

Statistical Analysis: Multivariable logistic regression, ANOVA, Bland-Altman plots.

3.2 AI System Deployment and Intervention

The AI consultation system is a hepatobiliary-pancreatic MAS built on LangChain, featuring a state-machine-driven architecture with five specialized agents:

Nexus Agent: Patient interaction with clinically relevant questions.

Cortex Agent: Dynamic generation of structured "Case Characteristics" (CC) following clinical guidelines.

Sentinel Agent: Validates CC completeness against predefined checklists.

Arbiter Agent: Resolves patient feedback on CC accuracy.

Oracle Agent: Generates evidence-based diagnostic recommendations using hybrid retrieval (Chroma + BM25).

Deployment: Integrated into hospital mobile apps and terminals. AI accesses EMR demographics (readonly) but does not write outputs. All interactions are encrypted and stored on isolated research servers.

Safety Protocol: Nexus Agent triggers risk alerts (e.g., "Severe allergy detected—review with physician") and logs critical events for physician review, ensuring a "AI screening–human verification" safety loop.

3.3 Bias Control and Fairness

Randomization: Stratified block randomization (block size 4/6) by department.

Blinding: Gold-standard experts, evaluators, and statisticians are blinded.

Equity Measures:

Provide devices and guidance for low-digital-literacy patients;

Include diverse demographics (age, education, dialects);

Monitor subgroup performance to prevent algorithmic bias.

3.4 Data Governance System

Collection: Automated (AI logs, EMR) and manual (audio transcripts) data entry.

Transmission: HTTPS + SM4 encryption.

Storage: Dual-factor authentication on hospital research servers.

Audit Trail: Full tracking of all data operations.

3.5 Adverse Events

Definitions: Adverse Event (AE): Any unfavorable medical occurrence (e.g., patient anxiety due to AI misunderstanding). Serious AE (SAE): Prolonged hospitalization,

life-threatening events, or medication errors (e.g., missed allergy history).

Reporting: SAEs reported to PI within 24 hours and IRB within 72 hours.

Management: Terminate AI use, initiate clinical intervention, and withdraw patient if necessary.

3.6 Statistical Analysis Plan

Analysis Populations:

Full Analysis Set (FAS): All randomized patients with gold-standard interviews.

Per-Protocol Set (PPS): FAS excluding major protocol violations.

Missing Data: Multiple imputation (<10%); sensitivity analysis (>10%).

Methods:

Logistic regression for accuracy; ANOVA/Kruskal-Wallis for time/satisfaction scores; Cohen's Kappa for diagnostic consistency.

3.7 Data Quality Control

Entry: Dual independent entry with arbitration.

Coding Standards: Predefined dictionaries (ICD-11, SNOMED CT).

Compliance:

Patient training and family assistance;

Physician SOPs and monthly training;

Centralized quality audits (10% monthly review; $\kappa \geq 0.8$ required for continuation).

Informed Consent Form for Human Research at The Second Affiliated Hospital of Zhejiang University School of Medicine

Dear Participant,

You are invited to participate in a clinical study titled "Development and Clinical Validation of an AI Consultation Model for Hepatobiliary and Pancreatic Diseases." Before deciding to join, please read this document carefully. It explains the purpose, procedures, potential benefits, risks, and your rights as a participant.

1. Study Background and Objectives

Artificial intelligence (AI) is increasingly applied in healthcare, particularly in diagnostic assistance and patient interaction. Hepatobiliary and pancreatic diseases (e.g., liver cancer, pancreatic cancer, cirrhosis) are challenging to diagnose early due to their silent onset, rapid progression, and poor prognosis. In primary care settings, shortages of specialists, variable patient health literacy, and incomplete initial data collection often lead to missed or delayed diagnoses.

Recent advances in large language models (LLMs) and multi-agent systems (MAS) enable the development of intelligent pre-consultation systems. This study aims to create an AI system that uses a MAS architecture to systematically collect patient symptoms, risk factors, and medical history, generating structured reports to aid physicians in early disease detection.

2. Study Procedures

Group Allocation:

Experimental Group 1:

Complete an AI-led consultation.

A blinded physician (unaware of AI results) conducts a separate consultation, generating a reference medical record.

AI-generated records are not shared with the physician.

Experimental Group 2:

Complete an AI-led consultation.

The physician reviews the AI-generated "Case Characteristics" (excluding diagnostic recommendations) before conducting a formal consultation.

Control Group:

Physicians conduct consultations without AI assistance, serving as the baseline for comparison.

Safety Measures:

The AI system includes a multi-tiered risk-monitoring mechanism. High-risk keywords (e.g., "hematemesis," "severe allergy") trigger alerts for physicians to manually review critical cases.

3. Your Role in the Study

Complete an AI-led consultation (Groups 1 and 2).

Participate in a standard physician consultation (all groups).

4. Potential Benefits

This study may not provide direct benefits to you but could improve AI-assisted diagnosis for future patients.

5. Risks and Discomforts

Prolonged Time: Groups 1 and 2 require two consultations, which may cause fatigue or scheduling inconvenience.

Psychological Discomfort: Repeated questions or unfamiliarity with AI may cause mild stress.

Privacy Risks: Data will be de-identified and stored on encrypted hospital servers. While risks are minimal, breaches cannot be entirely ruled out.

Our Safeguards:

Researchers will be present during consultations.

You may withdraw at any time without affecting your care.

Emergency protocols are in place for unexpected situations.

6. Costs and Compensation

No additional costs are associated with participation.

No compensation is provided, as the study does not involve experimental interventions.

7. Alternatives

You may decline participation without impacting your standard clinical care.

8. Confidentiality

Medical records will be stored securely.

Only authorized personnel (researchers, ethics committee) may access data.

Published results will not disclose personal identifiers.

9. Withdrawal from the Study

Participation is voluntary. You may withdraw at any time without penalty.

Reasons for termination include non-compliance, health risks, or researcher discretion.

10. Ethical Approval

This study was approved by the Ethics Committee of The Second Affiliated Hospital of Zhejiang University School of Medicine (Contact: Daytime: +86-571-87783759; Night (On-Call): +86-13757118366; Email: HREC2013@126.com).

Consent Statement

I confirm that I have read and understood this document, and I voluntarily agree to participate in the study. I consent to the use of my medical data for research publication.

Participant Signature: _____ Contact: _____

Date: _____

Legal Representative (if applicable):

Signature: _____ Relationship: _____

Contact: _____ Date: _____

Witness (if required):

Signature: _____ Contact: _____

Date: _____

Researcher Statement:

I confirm that I have explained the study details, including rights, risks, and benefits, to the participant and provided a signed copy of this consent form.

Researcher Signature: _____

Contact (Mobile): _____

Date: _____