Evaluation of the Effectiveness and Safety of CascadeDiagnose Prostate MRI: A Multi-Center Clinical Study

NCT number: N/A

Lead Institution:

The First Affiliated Hospital of Guangxi Medical University

Participating Institutions:

Nanning Second People's Hospital

Guilin People's Hospital

Liuzhou People's Hospital

Guigang People's Hospital

Wuzhou Red Cross Hospital

Principal Investigator: Jiwen Cheng

July 1, 2026

1. Research Background

Prostate cancer (PCa) ranks as the second most prevalent malignant tumor in men worldwide, with its incidence continuously rising. Magnetic resonance imaging (MRI) has become a core tool in the diagnostic workflow for prostate cancer, and clinical practice widely adopts the Prostate Imaging-Reporting and Data System version 2 (PI-RADS v2) for standardized interpretation. Nevertheless, MRI interpretation suffers from substantial inter-observer variability. Besides, the PI-RADS scoring strategy is oversimplified and lacks flexibility, resulting in sub-optimal diagnostic performance.

In recent years, artificial intelligence (AI) technologies have achieved remarkable progress in MRI-based prostate cancer diagnosis. Deep-learning models have evolved from radiomics-driven approaches to end-to-end architectures, and multimodal fusion strategies have further improved prostate cancer detection performance. Recent studies demonstrated that Prost-LM, a multimodal large-language-model-based algorithm, achieved an AUC of 0.954 for differentiating prostate cancer from benign lesions when validated on a multi-center cohort of 3,940 subjects, outperforming conventional single-modal models.

Our research group has developed the intelligent MRI-based prostate lesion detection system (CascadeDiagnose Prostate MRI), an intelligent prostate-lesion detection system built upon a cascaded deep-learning architecture. Adopting a cascade detection-then-classification framework, the system first performs lesion detection and segmentation on prostate MRI images, followed by malignant-benign discrimination and risk stratification for detected lesions. It integrates three MRI sequences: T2-weighted imaging (T2WI), diffusion-weighted imaging (DWI), and apparent diffusion coefficient (ADC), and is trained and validated using multi-center datasets. The system features credibility review of imaging data, standardization assessment, lesion identification and auxiliary diagnosis, alongside interpretable outputs. Traceable evidence is provided for lesion detection outputs, including MRI sequences, slice locations, spatial coordinates, confidence scores, heatmaps, and segmentation masks. For low-quality, non-standard, suspected tampered, cross-patient mixed, or

non-interpretable data, the system returns indeterminate or rejection outputs rather than forcing diagnostic conclusions. Full-link traceability is implemented for data versions, model versions, threshold versions, inference logs, and manual review records. To comprehensively evaluate the clinical effectiveness and safety of this system in Chinese populations, this multi-center clinical study is designed.

Currently, advancing multi-center clinical validation faces dual bottlenecks in regulatory compliance and engineering implementation. Prostate MRI images, patient PSA levels, medical histories, and biopsy pathological results from each site represent highly sensitive health-related data regulated by the Data Security Law of the People's Republic of China and the Personal Information Protection Law of the People's Republic of China. Raw imaging data, personal identifiers, and pathological materials cannot be copied or centrally uploaded for cross-hospital aggregation. Traditional centralized dataset paradigms carry substantial risks of privacy leakage, violate hospital data sovereignty, and struggle to pass multi-center ethical reviews, severely hindering large-scale joint training and synchronous performance validation.

This study leverages the Guangxi-wide medical-sharing distributed trusted computing network for all multi-center data collaboration. The core concept is "Don't move data; move computing power". Edge computing nodes are deployed locally in each hospital, connected via encrypted dedicated lines to an aggregation node, establishing a distributed trusted medical computing network where data remains in-hospital and models collaborate across institutions. A U-shaped federated-learning architecture is implemented. Raw clinical and imaging data from each hospital are processed locally within on-premises servers. Only desensitized and SM4-encrypted intermediate features and gradient parameters are transmitted to the regional aggregation node for model aggregation and multi-center statistical summarization. This architecture fundamentally realizes "data available but not visible", satisfying requirements for multi-center joint modeling and synchronous clinical assessment while safeguarding patient privacy and institutional data sovereignty, and resolving compliance barriers for multi-center research data sharing.

2. Research Objectives

2.1 Primary Objectives

(1) To assess the diagnostic accuracy of the intelligent MRI-based prostate lesion detection system for prostate cancer detection, with AUC, sensitivity, and specificity as primary endpoints.

(2) To evaluate system-related clinical safety, focusing on false-negative rate, false-positive rate, risks associated with rejection or indeterminate outputs, system operational failures, and clinical-decision bias potentially induced by erroneous AI outputs. Adverse events linked to investigational medical-device usage will be documented.

2.2 Secondary Objectives

(1) To compare the diagnostic performance of the intelligent MRI-based prostate lesion detection system with standalone interpretations by senior radiologists for prostate-cancer MRI reading.

(2) To evaluate the generalization performance of the system across different MRI vendors (GE, Siemens, etc.), magnetic-field strengths (1.5 T vs. 3.0 T), and clinical centers.

(3) To quantify changes in diagnostic accuracy and reading efficiency for junior radiologists after receiving auxiliary support from the system.

(4) To assess the potential clinical value of the system in reducing unnecessary prostate biopsy procedures.

3. Study Design

3.1 Study Type

This is a multi-center diagnostic clinical study combining retrospective validation and prospective clinical validation, complying with the Guidelines for Clinical Evaluation and Registration Review of Artificial-Intelligence-Assisted Detection Medical Devices (Software).

Cross-center synchronous AI inference and multi-center aggregated statistics are implemented on the distributed trusted medical-sharing computing network following the principle of "Don’t move data; move computing power". The network adopts a topology of 1 aggregation node + 6 hospital edge nodes. Raw MRI and pathological clinical data from each site are encrypted and stored locally, with image preprocessing and low-level AI feature extraction performed on-site. Only encrypted intermediate representations are transmitted across nodes; no raw patient-identifiable private data flows cross hospitals, satisfying medical-data compliance and ethical-review requirements.

3.2 Study Phases

The study consists of two phases:

Phase 1 (Retrospective Validation): Consecutive historical subjects with prior prostate MRI examinations and eligible reference-standard results are enrolled from participating centers. Enrolled cases shall not have been used for model training, parameter tuning, model selection, or decision-threshold determination of the intelligent MRI-based prostate lesion detection system, and serve for independent cross-center performance assessment of the locked-version model.

Phase 2 (Prospective Clinical Validation): Each center consecutively enrolls eligible subjects according to unified inclusion/exclusion criteria. AI analysis and human reading are performed under a fixed system version to evaluate real-world diagnostic performance, operational stability, safety, and clinical benefits of AI-assisted reading. The protocol pre-specifies whether AI outputs are accessible to treating clinicians and defines permitted usage scope.

3.3 Study Population

A minimum of 2 000 eligible patients are planned for enrollment: approximately 1 500 for the retrospective phase and 500 for the prospective phase.

4. Participating Centers

This multi-center study is conducted across six tertiary-level hospitals covering diverse MRI platforms, case-mix profiles, and radiologist experience levels to evaluate cross-center stability and device adaptability of the system.

Table 1 Participating Institutions

No	Center Name	Role
1	The First Affiliated Hospital of Guangxi Medical University	Lead Center
2	Nanning Second People's Hospital	Participating Center
3	Guilin People's Hospital	Participating Center
4	Liuzhou People's Hospital	Participating Center
5	Guigang People's Hospital	Participating Center
6	Wuzhou Red Cross Hospital	Participating Center

Each of the six hospitals deploys an independent edge-computing node connected via three 10 Gb/s national-cryptography VPN encrypted dedicated lines to the computing aggregation node at Wuxiang Unicom Computer Room. The core "Don't move data; move computing power" mechanism is fully implemented. Raw prostate MRI, medical record, and pathological data remain closed-loop stored on local hospital servers; only encrypted intermediate computational results participate in cross-center model collaboration and data aggregation.

5. Inclusion and Exclusion Criteria

5.1 Inclusion Criteria

(1) Male patients aged ≥ 18 years undergoing prostate MRI examination.

(2) Complete MRI sequences including at minimum T2WI, DWI, and ADC sequences.

(3) Definitive pathological diagnosis (prostate biopsy or post-surgical pathology) with complete Gleason-score information.

(4) Complete clinical data including serum PSA level, age, and prior medical history.

(5) Patient informed consent (written informed consent required for the prospective phase).

5.2 Exclusion Criteria

a) Substantial MRI image degradation caused by motion artifacts or metal artifacts that severely impair interpretation.

b) Prior prostate biopsy, prostate surgery, radiotherapy, or endocrine therapy.

c) Medical history of other malignant neoplasms.

d) Missing key MRI data or reference-standard materials required for primary-endpoint evaluation, precluding assessment of primary study outcomes.

e) Severe systemic diseases (e.g., heart failure, end-stage renal disease) interfering with imaging assessment or prognostic evaluation.

6. Study Methods

6.1 Overview of the Intelligent MRI-Based Prostate Lesion Detection System

The intelligent MRI-based prostate lesion detection system is an auxiliary analytical system for MRI-based prostate-lesion detection research and engineering demonstration. Built on cascaded deep-learning architectures, it automatically analyzes eligible prostate MRI scans to assist radiologists in detecting and localizing suspicious

lesions and assessing prostate-cancer risk, and delivers corresponding visualized auxiliary outputs.

6.2 Image Acquisition and Preprocessing

Image-quality review and rejection module: Image-quality assessment algorithms automatically evaluate artifact severity, sequence completeness, signal-to-noise ratio, and resolution. For images failing quality criteria, the system outputs "indeterminate" or "rejected" results with documented root causes such as severe motion artifacts, insufficient signal-to-noise ratio, or incomplete sequences. This module also performs data-integrity checks to identify anomalies including suspected tampering and cross-patient data mixing.

Prostate-region segmentation module: Built upon the 3D U-Net architecture, it automatically segments the whole prostate region to generate regions-of-interest (ROIs) for subsequent lesion detection. Automatic quality control is performed on prostate-segmentation outputs. Cases with algorithm-detected processing failures are recorded as technical failures, rejected, or non-evaluable per pre-defined rules and included in performance and safety analyses.

Lesion detection and segmentation module: Automatically detects suspicious lesions within the prostate region and performs precise lesion segmentation.

Lesion classification module: Based on residual-network architectures, it conducts binary malignant-benign classification and Gleason-score-based risk stratification for detected lesions.

Interpretable-output module: For each case, the system automatically generates a structured evidence-chain report containing:

a. Sequence and slice information: specific MRI sequences and key slices supporting the diagnostic inference.

b. Spatial location: precise intra-prostatic zonal localization (peripheral zone / transition zone / central zone) and three-dimensional coordinates of lesions.

c. Confidence score: model-derived confidence (0-100 %) for each diagnostic conclusion.

d. Grad-CAM heatmap: attention heatmap overlaid on raw MRI illustrating image regions driving model decisions.

e. Segmentation mask: precise contour delineation of lesions.

Input data consist of T2WI, DWI, and ADC MRI sequences. Outputs contain annotated lesion locations, malignant probability scores, evidence-chain reports, and risk-stratification recommendations.

6.3 Rejection and Indeterminate-Result Mechanism

The gold-standard reference for this study is histopathological diagnosis:

Malignant lesion: pathologically confirmed prostate cancer.

Benign lesion: pathologically confirmed benign prostatic hyperplasia (BPH), prostatitis, or other benign conditions.

Histopathology serves as the primary reference standard. For subjects undergoing radical prostatectomy, radical-specimen pathology is prioritized. For patients without radical surgery, systematic prostate-biopsy pathology is adopted. When multiple pathological reports exist for one subject, patient-level final pathological classification is determined by the highest ISUP grade.

All pathological reports, PI-RADS scores, and other sensitive clinical texts are encrypted and stored locally on each hospital's edge node. Only desensitized categorical statistical labels are encrypted and uploaded to the aggregation node for joint modeling.

6.4 Reading Protocol

AI standalone interpretation: The intelligent MRI-based prostate lesion detection system independently analyzes MRI images and generates diagnostic outputs.

Radiologist standalone interpretation: Radiologists of varying seniority (senior: $\geq$ 10 years of experience; mid-level: 5-10 years; junior: $<$ 5 years) from each center read scans independently and record PI-RADS scores and diagnostic opinions.

AI-assisted interpretation: Radiologists perform secondary reading with access to AI outputs. Diagnostic conclusions and reading time are recorded.

Radiology reading terminals at each hospital connect exclusively to the local institutional edge node. AI inference, lesion segmentation, and evidence-map generation are completed within the local closed-loop workflow. Reading terminals only access local institutional cases and cannot retrieve raw images from other hospitals. Multi-center performance comparisons are realized exclusively through encrypted aggregated metrics at the aggregation node.

6.5 System Traceability Management

Full-link traceability is implemented within the intelligent MRI-based prostate lesion detection system.

Data-version traceability: Records source, acquisition time, de-identified version number, and cryptographic hash value for each input dataset.

Model-version traceability: Records model version, training-dataset version, and model-parameter hash used for every inference run.

Threshold-version traceability: Documents decision-threshold versions and historical modification logs.

Inference-log traceability: Preserves complete inference logs including input-data summaries, outputs of each module, timestamps, and anomalous events.

Manual-review-record traceability: Logs personnel, timestamps, review comments, and modification records for all manual-review operations.

An independent cross-node collaborative audit ledger is built within the medical-sharing network to comprehensively log edge-node computing-resource invocation, encrypted intermediate-data transmission, global-model aggregation-update, and gradient back-propagation workflows. Each collaborative task is assigned a unique tamper-resistant storage-evidence identifier, enabling full-chain trace-back inspection by ethics committees and regulatory authorities.

7. Evaluation Endpoints

7.1 Effectiveness Endpoints

Primary Effectiveness Endpoints

a. Area under the receiver-operating-characteristic curve (AUC): overall diagnostic performance for discriminating prostate cancer from benign lesions.

b. Sensitivity: capacity of the system to correctly identify prostate-cancer cases.

c. Specificity: capacity of the system to correctly rule out benign lesions.

d. Positive predictive value (PPV) and negative predictive value (NPV).

Secondary Effectiveness Endpoints

a. Diagnostic-performance comparison against radiologists: differences in AUC, sensitivity, and specificity between the AI system and radiologists of varying seniority.

b. Changes in diagnostic performance under AI assistance: comparison of diagnostic accuracy and reading efficiency for radiologists before and after AI support.

c. Potential for reduction of unnecessary biopsies: proportion of avoidable unnecessary biopsies while maintaining high sensitivity ($\geq 95\,\%$) relative to pathological gold standards.

d. Decision-curve analysis (DCA): clinical net benefit of the system across varying decision thresholds.

e. Subgroup analyses: stratified by PSA level (grey zone 4-10 ng/mL vs. > 10 ng/mL), PI-RADS score (3 vs. 4-5), age, study center, and MRI-device vendor.

f. Rejection rate and indeterminate-output rate: proportion of cases rejected or flagged indeterminate due to poor image quality, data anomalies, or insufficient model confidence for assessment of the image-quality-control and rejection mechanism.

g. Evidence-chain completeness: completeness, accuracy, and clinical acceptability of system-generated evidence-chain reports (sequence/slice/spatial location/confidence score/heatmap/segmentation mask) in clinical validation.

7.2 Safety Endpoints

a. Adverse Events (AEs): all adverse events associated with system usage are documented.

b. False-negative rate: proportion of missed prostate-cancer cases, which may lead to delayed treatment.

c. False-positive rate: proportion of false alarms that may trigger unnecessary biopsies and related complications.

d. System failure rate: frequency of technical malfunctions (analysis failure, timeout, crash, etc.) during system operation.

e. Assessment of impact on clinical decisions: degree of influence of system outputs on clinical workflows and risk of misleading clinical judgment.

f. Safety assessment related to rejection outputs: clinical impact of rejection/indeterminate outputs; clinical outcomes of rejected cases to verify that rejection mechanisms do not result in missed clinically significant lesions.

8. Data Collection and Pre-Processing

8.1 CT Image Acquisition

All participating centers adhere to standardized multi-phase renal-contrast-enhanced CT acquisition protocols adapted from Dai et al. Recommended scanning parameters:

a. Slice thickness: ≤ 5 mm; thin-slice reconstruction (1-3 mm) preferred.

b. Phases: non-contrast phase; corticomedullary phase (30-40 s post-contrast-media injection); nephrographic phase (90-120 s post-injection); excretory phase (3-5 min post-injection).

c. Reconstruction kernel: standard soft-tissue algorithm.

d. Matrix size: 512 × 512.

Distributed-computing-acquisition specification: all CT data are stored in DICOM format locally on hospital-based edge-node servers. No raw images are exported to portable media or transmitted cross-institutionally. Only locally-desensitized, quality-controlled, feature-extracted encrypted intermediate outputs participate in cross-site collaborative workflows.

8.2 Image Registration and Pre-Processing

To compensate for respiratory-motion-induced positional shifts across contrast phases, the corticomedullary phase serves as the reference volume. Non-contrast,

nephrographic, and excretory-phase volumes undergo rigid registration against the reference using 3D Slicer (version 5.10.0). Automatic image-quality assessment is performed by the built-in review module; scans flagged as poor quality receive secondary manual review by senior on-site radiologists.

Computational-resource allocation: all registration and pre-processing operations run locally on hospital edge nodes without consuming resources of the central convergence node, implementing the “local data processed locally” principle.

8.3 Histopathological Gold Standard

Surgical histopathology constitutes the reference gold standard. Pathological specimens from each site are interpreted by local-center pathologists following WHO-ISUP classification criteria. Challenging ambiguous cases are submitted for secondary review by the pathology department of the lead center.

Benign entities include fat-poor angiomyolipoma, oncocytoma, complex renal cysts, and other benign lesions.

Malignant entities include clear-cell renal-cell carcinoma, papillary renal-cell carcinoma, chromophobe renal-cell carcinoma, and other malignant renal neoplasms.

8.4 Clinical-Data Collection

Collected clinical variables: age, sex, clinical symptoms, prior medical history, laboratory tests (complete blood count, renal-and-hepatic biochemistry etc.), maximum tumor diameter measured on CT, tumor location (left/right kidney; upper/mid/lower pole), and CT imaging features (solid/cystic appearance, margin characteristics, enhancement patterns).

9. Observer Study

9.1 Panel of Readers

Table 2 Composition and Qualifications of Image Readers

Category	Number of Readers	Experience Requirement
Senior radiologists	2	≥ 10 years genitourinary-imaging experience

Mid-/junior-level radiologists	2	3-10 years radiological practice
Urologists	2	≥ 5 years renal-tumor clinical-management experience

9.2 Reading Workflow

a) Training phase: A genitourinary-imaging expert (lead center, > 5 000 renal-tumor-case reading experience) delivers standardized training for all readers, covering imaging-feature interpretation for 50 illustrative representative cases and standardized scoring frameworks (e.g., likelihood score for clear-cell carcinoma, Bosniak classification).
b) Independent-reading phase: Each reader interprets every CT dataset in isolation and blinded to clinical and pathological reference data, assigning a continuous 0-100% probability estimate of benign histology.
c) AI-assisted-reading phase: Readers re-evaluate identical cases with outputs from CascadeDiagnose Renal Tumor CT visible. Diagnostic conclusions and reading time are recorded.
d) Computing-workflow note: Reading terminals connect exclusively to the local hospital edge node. AI inference executes on-site. Under the "move computing power instead of raw data" architecture, readers never access raw case data originating from other hospitals. Only desensitized aggregated summary statistics are forwarded to the multi-center analysis platform.

10. Sample-Size Calculation

Sample-size assumptions are derived from published clinical-AI studies:

a. Expected system AUC = 0.85

b. Minimal clinically meaningful difference = 0.05

c. Two-sided $\alpha = 0.05$

d. Statistical power $(1-\beta) = 0.80$

e. Benign-to-malignant case ratio 1:1, reflecting real-world surgical-cohort distributions.

Calculations indicate approximately 1500 benign and 1500 malignant subjects are required. Accounting for multi-center heterogeneity and anticipated protocol violations/exclusions, the target total sample size is no fewer than 3000 participants (2600 retrospective, 400 prospective).

11. Statistical Analysis

11.1 Analysis Datasets

a) Full Analysis Set (FAS): All enrolled participants for baseline-characteristic descriptive statistics.

b) Per-Protocol Set (PPS): Subjects completing the full study workflow without major-protocol deviations; primary-efficacy analyses rely on this dataset.

c) Safety Set (SS): All patients receiving system assessment for safety-endpoint analysis.

d) Rejection Set (RS): Cases marked indeterminate or rejected by the system; clinical features and outcomes for these patients will be analyzed separately.

e) Distributed-computing-dataset specification: Desensitized feature subsets are generated locally on each edge node. Only encrypted aggregated summary statistics are transmitted to the convergence node for pooled analysis. Individual-level raw datasets are permanently retained within each participating hospital and are not centrally pooled.

11.2 Statistical Methods

a) Descriptive statistics: Continuous variables reported as mean ± standard-deviation or median (inter-quartile range). Categorical variables presented as frequency (percentage).

b) Diagnostic-performance comparisons: DeLong test for inter-group AUC comparisons. Analyses performed in R (version 4.5.1) and SPSS (version 26.0).

c) Sensitivity-and-specificity pairwise comparisons: McNemar's test.

d) Multi-center heterogeneity assessment: Random-effects modelling or Cochran's Q-statistic.

e) Subgroup analyses: Logistic-regression models incorporating interaction terms.

f) Rejected-case analysis: Descriptive characterization of rejection-reason distribution and associations with clinical outcomes.

g) All hypothesis tests are two-tailed. A P-value < 0.05 denotes statistical significance.

12. Quality Management and Traceability

12.1 Data-Quality Management

Standard operating procedures (SOPs) titled CT-Image-Acquisition and Data-Management Protocol for Renal-Tumor Research are enforced uniformly at all centers.

Local-center data managers oversee completeness, accuracy, and image-quality auditing.

De-identification is applied to all patient datasets prior to system ingestion in compliance with the Personal Information Protection Law of the People's Republic of China.

Specialized distributed-network quality control: Cross-node transmission integrity checks and anti-tampering receipts are implemented. Transfer of raw imaging or medical-record data across edge nodes is strictly prohibited. Periodic access-permission audits are performed on local edge-node storage to enforce the "move computing power instead of raw data" mandate.

12.2 System-Traceability Governance

Version control: Unique version identifiers, release timestamps, and change logs are maintained for datasets, AI-model iterations, and decision-threshold configurations.

Log management: The system automatically logs every inference workflow including inputs, outputs, timestamps, and exception events.

Review-audit trails: Manual-review events permanently store operator identity, timestamp, and review comments.

13. Ethical Considerations

a) Ethical-committee approval: The research protocol must receive ethical-review-committee clearance from each participating institution. The Ethics Committee of The First Affiliated Hospital of Guangxi Medical University serves as the primary reviewing body.

b) Informed consent: Waiver of written informed consent for retrospective data may be granted per local-ethical-committee determinations. Written informed consent must be obtained for all prospective participants.

c) Data protection: Patient datasets undergo anonymization complying with the Personal Information Protection Law and the Declaration of Helsinki. Built-in system safeguards mitigate data-leakage and unauthorized-access risks.

d) Clinical-trial registration: Trial registration shall be completed at the Chinese Clinical Trial Registry (ChiCTR) prior to subject enrollment.

e) Distributed-trustworthy-computing-network-specific ethical compliance: The multi-center collaborative framework adheres strictly to the “move computing power instead of raw data” principle:

① Data sovereignty: Original CT, electronic-medical-record, and pathology data remain stored and processed locally in-hospital. Each institution retains full sovereignty over its patient datasets without raw-data export or centralized aggregation.

② Privacy protection: Cross-institutional transmission is limited to desensitized, cryptographically-secured intermediate features and gradient parameters. The convergence node cannot reconstruct original patient images or personally-identifiable information, achieving the “data is usable but not observable” paradigm.

③ Transparent traceability: Tamper-resistant audit trails capture every cross-node computing-resource scheduling and encrypted-data-transfer event. Ethics committees may inspect these logs to verify data-use compliance.

④ Risk isolation: Physical and logical network isolation separates edge-node infrastructure from the convergence node. The convergence node cannot access local raw datasets, architecturally mitigating patient-privacy-breach risks inherent to conventional multi-center clinical research.

14. Study Timeline

Total study duration: August 2026 – July 2028.

Subject-recruitment windows: retrospective-phase recruitment October 2026 – December 2027; prospective-phase enrollment January 2027 – December 2027.

Table 3 Study-Phase Timeline

Phase	Time Window	Key Deliverables
Preparation	2026-08-01 – 2026-09-30	Protocol finalization, ethical approval, clinical-trial registration, system deployment, investigator training
Retrospective-data collection	2026-10-01 – 2027-12-30	Retrospective-case curation, quality-control, retrospective-phase system validation
Prospective enrollment	2027-01-01 – 2027-12-30	Prospective-subject enrolment, clinical-and-imaging-data capture, prospective-phase system validation
Statistical analysis	2028-01-01 – 2028-03-30	Data-lock, statistical computation, result interpretation
Study conclusion	2028-04-01 – 2028-07-30	Final-study-report composition, manuscript submission and publication, knowledge-translation and dissemination

15. Expected Outcomes

a) Comprehensive multi-center performance characterization of CascadeDiagnose Renal Tumor CT for differentiating benign-malignant small renal masses ≤ 3 cm, including sub-centimeter (≤ 1 cm) lesions.

b) Empirical validation of generalization and robustness across varying CT scanner hardware, acquisition parameters, and geographically-diverse clinical centers.

c) Empirical validation of image-quality screening, evidence-chain tracing, and case-rejection mechanisms.

d) Quantification of clinical-decision-support utility, especially potential reduction of unnecessary renal-surgical procedures.

e) Generation of high-quality clinical-evidence packages supporting medical-device-regulatory submission.

f) Publication of 2-3 high-impact peer-reviewed clinical-research articles.

g) Real-world demonstration of a "move computing power instead of raw data" distributed trustworthy-medical-computing network for multi-center medical-device clinical trials. A replicable compliance-oriented implementation specification for healthcare-AI multi-center collaborative research will be produced to serve as engineering-and-ethics reference for analogous medical-device-evaluation projects.

16. References

[1] Dai C, Xiong Y, Zhu P, et al. Deep Learning Assessment of Small Renal Masses at Contrast-enhanced Multiphase CT. Radiology. 2024;311(2):e232178.

[2] Center for Medical Device Evaluation, National Medical Products Administration. Guidelines for Clinical Evaluation and Registration Review of Artificial-Intelligence-Assisted Detection Medical Devices.

[3] Center for Medical Device Evaluation, National Medical Products Administration. Guidelines for Clinical Evaluation and Registration Review of Artificial-Intelligence-Assisted-Diagnosis Medical Devices.

[4] Çiçek Ö, Abdulkadir A, Lienkamp S S, et al. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2016: 424-432.

[5] Heller N, Isensee F, Maier-Hein KH, et al. The state of the art in kidney and kidney tumor segmentation in contrast-enhanced CT imaging: Results of the KiTS19 challenge. Med Image Anal. 2021;67:101821.

[6] SPIRIT-AI Extension: Guidelines for clinical trial protocols for interventions involving artificial intelligence. Equator Network, 2025.

[7] DECIDE-AI: Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence, 2025.

[8] Srigley JR, Delahunt B, Eble JN, et al. The International Society of Urological Pathology (ISUP) Vancouver Classification of Renal Neoplasia. Am J Surg Pathol. 2013;37(10):1469-1489.