Official Title:

Evaluation of the Effectiveness and Safety of a Locally Deployed AI Model in Multidisciplinary Team (MDT) Decision-Making for Non-Small Cell Lung Cancer: A Prospective, Controlled Clinical Trial

NCT Number: Not yet assigned

Document Date:

April 05, 2026

# Evaluation of the Effectiveness and Safety of a Locally Deployed AI Model in Multidisciplinary Team (MDT) Decision-Making for Non-Small Cell Lung Cancer: A Prospective, Controlled Clinical Trial

Team Leader: Guangdong Provincial People's Hospital

Project Leader: Wenzhao Zhong

Department: Pulmonary Surgery

Contact Number: 13609777314

Participating Institution: Guangdong Provincial People's Hospital

Research Period: October 2025 – October 2027

Version Number and Date: V3.0 2026-04-05

PROTOCOL SUMMARY

| | |
|---|---|
| Project Title | Evaluation of the Effectiveness and Safety of a Locally Deployed AI Model in Multidisciplinary Team (MDT) Decision-Making for Non-Small Cell Lung Cancer: A Prospective, Controlled Clinical Trial |
| Study Purpose | This is a prospective, single-center clinical trial aimed at evaluating the effectiveness and safety of a locally deployed AI decision-support model developed by our department in the multidisciplinary team (MDT) process for non-small cell lung cancer (NSCLC). |
| Study Design | Patient clinical data will be submitted concurrently to a traditional MDT team and the locally deployed AI model for analysis. The traditional MDT will provide Treatment Plan 1, while the AI model will independently generate Treatment Plan 2. After reviewing the AI model's recommendations, clinicians will make the final treatment decision, designated as Treatment Plan 3, which will be implemented for the patient. The primary endpoints of this study are to evaluate the consistency between the AI model's recommendations and the traditional MDT's decisions, as well as the rate of clinician adoption of the AI model's recommendations (i.e., decision modification rate). Secondary endpoints include improvements in MDT discussion efficiency, physician satisfaction, and long-term patient survival |

| | |
|---|---|
| | outcomes (DFS, PFS, OS). This study aims to validate that the locally deployed AI model can serve as an efficient and accurate adjunct tool in MDT, contributing to improved standardization and homogenization of lung cancer care. |
| Total Sample Size | 300 |
| Patient Selection | Inclusion Criteria: Patients with stage II-IV NSCLC undergoing MDT discussion; availability of detailed clinical data. Exclusion |
| | Exclusion Criteria: Stage I NSCLC; non-NSCLC thoracic tumors; lack of detailed clinical data. |
| Treatment Plan | Treatment is based on the conclusions of the AI model and the real-world MDT team. If the conclusions are consistent, treatment proceeds accordingly. If discrepancies exist, the real-world MDT team re-discusses the case considering the AI model's output to formulate a new plan, which the patient receives. |
| Efficacy Assessment<br>Efficacy Assessment | Effectiveness Evaluation Indicators (Primary and Secondary Endpoints): The primary focus is to investigate the consistency between the locally deployed AI model's conclusions and the real-world MDT team's conclusions, as well as the clinician's decision modification rate based on the AI model's recommendations. Subgroup analyses will be performed to assess the AI model's sensitivity and accuracy based on factors such as disease stage and gene expression status. Safety Evaluation Indicators: This study is led by clinicians, and the final treatment plan is determined by clinicians. |
| **Statistical** | Kaplan-Meier method, log-rank test, McNemar's test, Kappa |

| | |
|---|---|
| **Methods** | coefficient, etc. |
| **Study Duration** | 2 years |

## I. BACKGROUND

Lung cancer has the highest morbidity and mortality rates of all malignancies worldwide, with non-small cell lung cancer (NSCLC) accounting for approximately 85% of all lung cancer cases [1]. For patients with stage II-IV NSCLC, treatment strategies are complex and diverse, involving the single or combined use of surgery, chemotherapy, radiotherapy, targeted therapy, immunotherapy, and other modalities. Traditional single-department care models are often insufficient to meet the clinical needs of these patients.

The multidisciplinary team (MDT) model integrates specialists from departments such as thoracic surgery, medical oncology, radiation oncology, radiology, pathology, and respiratory medicine to collaboratively formulate individualized, optimized, and continuous diagnostic and treatment plans for patients. Numerous studies have demonstrated that the MDT model can significantly improve diagnostic accuracy, treatment adherence, survival rates, and quality of life for cancer patients [1]. Consequently, leading international and domestic guidelines (e.g., NCCN, CSCO) strongly recommend MDT discussion for NSCLC patients.

Although conceptually advanced, the MDT model faces several limitations in its practical implementation and dissemination, hindering its full potential. Firstly, traditional MDT processes often suffer from inefficiency and high resource consumption. Coordinating senior specialists from multiple disciplines for simultaneous meetings is logistically challenging and time-consuming. In-depth discussion of a single case typically requires 10-20 minutes, limiting the number of patients that can be discussed and making MDT a scarce medical resource [3]. Secondly, there is variability in the quality and standardization of care. The quality of MDT decisions heavily depends on the individual knowledge, experience, and clinical practice of participating experts. Significant discrepancies in the understanding and application of guidelines can exist between hospitals in different regions or at different levels, and even among specialists within the same institution. This leads to "heterogeneity" or "inequity" in lung cancer care, where patients may not receive standardized, high-quality treatment across different healthcare settings [4]. Thirdly, there is information overload and

escalating decision-making complexity. In the era of "precision medicine," treatment decisions for NSCLC require integrating an increasing amount of complex information, including detailed imaging features, pathological subtypes, PD-L1 expression levels, mutation status of numerous driver genes (e.g., EGFR, ALK, ROS1, MET, RET), and patient performance status (PS score) and comorbidities. Faced with such vast information, there is a risk of overlooking key details or failing to adhere to the latest guideline recommendations.

Artificial intelligence, particularly machine learning and natural language processing, is increasingly being applied in medicine, demonstrating significant potential in image recognition, data integration, and clinical decision support [5]. AI models can rapidly process massive amounts of data, integrating and analyzing both structured and unstructured patient information within seconds. They can perform logical reasoning based on the latest versions of international and domestic authoritative clinical guidelines, ensuring the standardization and timeliness of decisions. Additionally, they provide an objective, standardized decision-making reference, potentially mitigating inconsistencies arising from variations in expert experience or fatigue.

Based on these clinical needs and technological possibilities, we have developed a locally deployed AI-assisted decision-support model. This model, trained through deep learning on a large corpus of high-quality clinical guidelines, literature, and de-identified real-world data, aims to emulate a highly skilled, standardized, and tireless MDT expert team, generating treatment recommendations for NSCLC patients that align with guideline recommendations.

Currently, there is a lack of high-quality studies evaluating the role of AI models within the overall MDT decision-making process. Therefore, this study aims to conduct a rigorously designed prospective trial to conduct a head-to-head comparison between the locally deployed AI model and traditional human MDT decision-making, scientifically addressing the following key questions: Can the locally deployed AI model achieve a level of decision-making comparable to or even higher than that of a human expert team? Can it be

efficiently integrated into clinical workflows as a reliable "intelligent assistant" for physicians, ultimately improving the standardization, homogenization, and efficiency of lung cancer care.

## II. STUDY OBJECTIVES

Primary Objectives: To evaluate the consistency between the plan generated by the locally deployed AI model (Treatment Plan 2) and the plan from the traditional MDT discussion (Treatment Plan 1). To evaluate the consistency between the clinician's final decision after reviewing the AI model's opinion (Treatment Plan 3) and the initial traditional MDT plan (Treatment Plan 1), i.e., the clinician's decision modification rate.

Secondary Objectives: To evaluate the impact of using the locally deployed AI model on MDT discussion efficiency (time from data collection to the end of discussion). To survey MDT physician satisfaction with the discussion process and outcomes after integrating the AI model's opinion. To preliminarily observe and compare long-term survival outcomes (Disease-Free Survival, DFS; Progression-Free Survival, PFS; Overall Survival, OS) in patients treated based on the final MDT plan (Treatment Plan 3).

## III. STUDY DESIGN TYPE, PRINCIPLES, AND PROCEDURES

Study Design: Prospective, single-center, self-controlled clinical trial.

Study Procedures:

Retrospective Phase: Clinical data and treatment plans from patients who underwent MDT discussion at our hospital between January 1, 2023, and October 1, 2025, will be retrospectively collected. These data and the original MDT treatment plan will serve as the retrospective gold standard (Treatment Plan 1). Patient details will be input into the locally deployed AI model by independent researchers, and the model will generate Treatment Plan 2. Clinicians (2-3 senior physicians) will simultaneously review Treatment Plan 1 and Treatment Plan 2 to assess consistency and calculate the agreement rate.

Prospective Phase: Eligible patients will be enrolled, and their complete clinical data will be collected. The data will be submitted concurrently, in a blinded manner, to:

Traditional MDT Group: Conducts routine discussion and derives Treatment Plan 1.

Locally Deployed AI Model: Independent researchers input data, and the model generates Treatment Plan 2.

The clinician (MDT chair or primary decision-maker) will simultaneously review Treatment Plan 1 and Treatment Plan 2. If Plan 1 and Plan 2 are consistent, the discussion concludes, and the patient proceeds with this agreed-upon plan. If Plan 1 and Plan 2 are inconsistent, Plan 2 is disclosed. The clinician, considering the patient's specific situation along with the conclusions and rationale of Plan 2, re-discusses the case to formulate Treatment Plan 3, documenting the reasons for the final choice. The patient receives Treatment Plan 3.

Study Endpoints and Evaluation Metrics:

Primary Endpoints:

Consistency rate between Plan 1 and Plan 2 (calculated using Kappa coefficient).

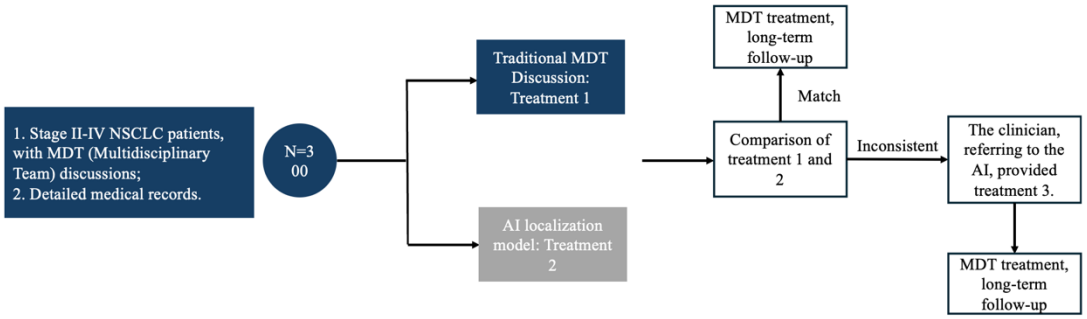
Consistency rate between Plan 1 and Plan 3 (decision modification rate).

Secondary Endpoints:

Duration of MDT discussion process (minutes).

MDT physician satisfaction questionnaire score (using a 5-point Likert scale).

Long-term patient survival outcomes (DFS, PFS, OS).




IV. PATIENT SELECTION

Inclusion Criteria:

Study Population: Patients with pathologically confirmed stage II-IV NSCLC.

Prospective Validation Sample Size: 300 cases.

Rationale: The primary objective is to verify the non-inferiority of the locally deployed AI

model compared to traditional MDT in NSCLC treatment decision-making. Based on preliminary retrospective data, the expected consistency rate between the AI model and human MDT is 95%, with a non-inferiority margin of 5% ($\Delta = 0.05$). Under a significance level $\alpha = 0.025$ (one-sided) and power $1-\beta = 0.8$, the calculated sample size is 300 cases. This sample size is sufficient to scientifically address the primary research question.

Inclusion Criteria:

1. Age $\geqslant$ 18 years.
2. Deemed by MDT discussion to require formulation of a systemic treatment plan.
3. Possess complete clinical, imaging, and molecular pathology data.

Exclusion Criteria:

1. Stage I disease.
2. Diagnosis of thoracic tumors other than NSCLC.
3. Lack of detailed clinical data or missing essential information.

V. RESEARCH METHODS AND TECHNICAL ROUTE

Enrolled eligible patients and collect all clinical data. The data will be submitted concurrently, in a blinded manner, to: the Traditional MDT Group, which conducts routine discussion to derive Treatment Plan 1; and the Locally Deployed AI Model, where independent researchers input data for the model to generate Treatment Plan 2. The clinician (MDT chair or primary decision-maker) reviews both plans. If Plan 1 and Plan 2 are consistent, the discussion concludes, and the patient follows this plan. If inconsistent, Plan 2 is disclosed, and the clinician, considering the patient's specifics and Plan 2's rationale, re-discusses to formulate Treatment Plan 3, documenting the choice rationale. The patient receives Treatment Plan 3.

VI. OBSERVATION ITEMS AND ASSESSMENT SCHEDULE

| Observation Category | Specific Observation Items | Screening | MDT Discussion Period | Treatment Period | Follow-up Period |
|---|---|---|---|---|---|
| **General Data** | Demographics, Medical History, Smoking History | ✓ | | | |
| Tumor Characteristics | TNM Stage, Histopathology, Grade | ✓ | | | |
| **Molecular Markers** | Driver Genes (EGFR, ALK, ROS1, etc.), PD-L1 Expression | ✓ | | | |
| **safety Indicators** | Physical Exam, ECOG PS Score, Vital Signs | ✓ | | Each Visit | Each Visit |
| | CBC, Liver/Kidney Function, Electrolytes, ECG | ✓ | | Per Protocol | |
| | Adverse Events (AE) / Serious Adverse Events (SAE) | | | | |

| Observation Category | Specific Observation Items | Screening | MDT Discussion Period | Treatment Period | Follow-up Period |
|---|---|---|---|---|---|
| **Efficacy Indicators** | Recording & Comparison of Treatment Plans 1, 2, 3 | | ✓ | | |
| | MDT Discussion Duration | | ✓ | | |
| | Physician Satisfaction Survey | | | (Immediately post-MDT) | |
| | Imaging Assessment (CT/MRI/PET-CT) | (Baseline) | | Every 6-12 weeks | Every 3-6 months |
| **Survival Indicators** | Progression-Free Survival (PFS), Overall Survival (OS) | | | | Every 3 months |
| | Disease-Free Survival (DFS) | | | | (For surgical patients) |

VII. EFFICACY EVALUATION CRITERIA

Decision Consistency Evaluation:

Definition: Assessing the degree of agreement between different treatment plans.

Method: An adjudication committee, consisting of two senior oncology specialists unaware of group allocation, will review the three treatment plans for each case. Plans will be decomposed into core decision elements (e.g., primary treatment modality: surgery/radiotherapy/systemic therapy; specific drug regimen; treatment intent: neoadjuvant/adjuvant/palliative, etc.) for comparison.

Criteria:

Complete Agreement: All core decision elements are identical.

Partial Agreement: Core principles are consistent, with differences only in minor details (e.g., treatment cycles, specific radiotherapy dose) that do not affect the overall treatment strategy or expected efficacy.

Disagreement: Principled differences in one or more core decision elements (e.g., targeted therapy vs. immunotherapy; surgery vs. palliative treatment).

Tumor Response Evaluation:

Definition: Evaluating objective tumor response after implementing the treatment plan.

Criteria: Using the internationally recognized Response Evaluation Criteria in Solid Tumors version 1.1 (RECIST 1.1).

Complete Response (CR): Disappearance of all target lesions.

Partial Response (PR): At least a 30% decrease in the sum of diameters of target lesions.

Stable Disease (SD): Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD.

Progressive Disease (PD): At least a 20% increase in the sum of diameters of target lesions, or appearance of new lesions.

Objective Response Rate (ORR) = CR + PR

Disease Control Rate (DCR) = CR + PR + SD

Survival Evaluation:

Progression-Free Survival (PFS): Time from the start of treatment plan implementation to the first documented disease progression (per RECIST 1.1) or death from any cause.

Overall Survival (OS): Time from the start of treatment plan implementation to death from any cause.

Disease-Free Survival (DFS): (Applicable to patients receiving curative-intent surgery) Time from the end of surgery to the first documented disease recurrence or death from any cause.

VIII. ADVERSE EVENT MONITORING

Definition of Adverse Event (AE): Any untoward medical occurrence occurring in a patient during treatment with Treatment Plan 3, regardless of its causal relationship to the treatment.

Definition of Serious Adverse Event (SAE): An adverse event occurring at any dose that:

Results in death.

Is life-threatening.

Requires inpatient hospitalization or prolongation of existing hospitalization.

Results in persistent or significant disability/incapacity.

Is a congenital anomaly/birth defect.

Is an important medical event (may not be immediately life-threatening but jeopardizes the patient or requires intervention to prevent one of the above outcomes).

Observation, Recording, and Reporting Procedures:

Observation and Recording: Investigators will closely observe and record all AEs from the time of signed informed consent until the final follow-up visit. Recorded details will include AE name, severity (CTCAE grade), start/end dates, relationship to study treatment, actions taken, and outcome.

Severity Grading: Grading will be performed using the National Cancer Institute Common Terminology Criteria for Adverse Events version 5.0 (CTCAE v5.0, grades 1-5).

SAE Reporting: Any SAE must be reported by the investigator to the Institutional Ethics Committee and the Principal Investigator within 24 hours. A written report detailing patient information, a description of the SAE, treatment provided, and causality analysis is required.

Risk Control Measures:

The final treatment plan in this study is determined by the clinician, based on treatments with known safety profiles. Risks are comparable to standard clinical care.

Investigators will fully inform patients of potential risks.

The research team has established emergency procedures to manage potential serious adverse events (e.g., allergic reactions, myelosuppression, immune-related pneumonitis).

All AEs will be followed until resolution, stabilization, or confirmation of no relationship to the study.

## IX. DATA SAFETY MONITORING

The clinical study will establish a data safety monitoring plan commensurate with the level of risk. All AEs will be documented, managed appropriately, and followed until resolution or stabilization. SAEs and unexpected events will be reported promptly to the Ethics Committee, regulatory authorities, sponsor, and drug administration as required. The Principal Investigator will periodically review all AEs cumulatively and convene investigator meetings as needed to assess the risk-benefit profile. For studies involving greater than minimal risk, an independent data monitor will be appointed; for high-risk studies, an independent Data Safety Monitoring Board (DSMB) will be established to monitor cumulative safety and efficacy data and provide recommendations on study continuation.

## XI. STATISTICAL ANALYSIS

Statistical analysis will be performed using SPSS version 27.0. Categorical data will be presented as frequencies and percentages, with comparisons between groups using the chi-square test or Fisher's exact test. Consistency analysis will employ the Kappa coefficient.

Survival analysis will utilize the Kaplan-Meier method and log-rank test. All statistical tests will be two-sided, with a p-value < 0.05 considered statistically significant.

XI. ETHICS OF CLINICAL RESEARCH

The clinical study will adhere to the World Medical Association's Declaration of Helsinki and relevant regulations. The study will only commence after the protocol has been approved by the Ethics Committee. Prior to enrollment, investigators are responsible for providing each potential subject (or their legally authorized representative) with comprehensive information regarding the study's purpose, procedures, and potential risks, and obtaining written informed consent. Subjects must be informed of their right to withdraw from the study at any time. The informed consent form will be retained as a clinical study document. Patient privacy and data confidentiality will be protected throughout the study.

XIII. **STUDY PERSONNEL**

| Name | Title/Specialty | Responsibilities |
|---|---|---|
| Zhong Wenzhao | Chief Physician/Oncology | Guidance, supervision, review |
| Liang Qing | Residence/Clinical Medicine | Patient enrollment, data collection, manuscript writing |
| Li Yongqian | CRC | Data collection, patient follow-up |
| Chen Jiayun | CRO | Data collection, AE monitoring, quality control management |

XIV. **REFERENCES**

[1]. WHO. Lung cancer. World Health Organization. 2023. https://www.who.int/news-room/fact-sheets/detail/lung-cancer (accessed July 15, 2024).

[2]. Popat S, Navani N, Kerr KM, et al. Navigating Diagnostic and Treatment

Decisions in Non-Small Cell Lung Cancer: Expert Commentary on the Multidisciplinary Team Approach. *Oncologist*. 2021;26(2):e306-e315. doi:10.1002/onco.13586

[3]. Liu H, Song Y. MDT is still important in the treatment of early stage lung cancer. *J Thorac Dis*. 2018;10(Suppl 33):S3984-S3985. doi:10.21037/jtd.2018.09.121

[4]. Chen S, Wang Z, Sun B. Chinese Society of Clinical Oncology Non-small Cell Lung Cancer (CSCO NSCLC) guidelines in 2024: key update on the management of early and locally advanced NSCLC. *Cancer Biol Med*. 2025;22(3):191-196. doi:10.20892/j.issn.2095-3941.2024.0592

[5]. Bi WL, Hosny A, Schabath MB, et al. Artificial intelligence in cancer imaging: Clinical challenges and applications. *CA Cancer J Clin*. 2019;69(2):127-157. doi:10.3322/caac.21552