International Platform of Registered Systematic Review and Meta-analysis Protocols

INPLASY

Artificial Intelligence-Assisted Lung Tumour Detection: A Systematic Review and Meta-Analysis

INPLASY202590109
doi: 10.37766/inplasy2025.9.0109
Received: 26 September 2025
Published: 27 September 2025

Zheng, YM; Zhang, SB; Geng, L; Li, JJ; Shi, KM; Yao, D; Li, QZ; Zhang, T; Zhao, YS; Zhang, L; Sun, Y; Wan, JX.

Corresponding author:
Jinxin Wan

jjjjiiiinnnn_3054@163.com

Author Affiliation:
The Second People's Hospital of Lianyungang; Cancer hospital of Lianyungang; The Second People's Hospital of Lianyungang Affiliated with Kangda College of Nanjing Medical University.

ADMINISTRATIVE INFORMATION

Support - This research was supported by the Health Science and Technology Project of the Lianyungang Health Commission (202325); the Scientific Research Project of the Lianyungang Anti-Cancer Association (MS202312); Science and Technology Plan Project of Lianyungang (No. SF2311); Aging Health Research Project of Lianyungang (No. L202318); Health Science and Technology Project of Lianyungang (No. 202219); Medical Education Collaborative Innovation Fund of Jiangsu University (No. JDYY2023087).

Review Stage at time of this submission - Completed but not published.

Conflicts of interest - None declared.

INPLASY registration number: INPLASY202590109

Amendments - This protocol was registered with the International Platform of Registered Systematic Review and Meta-Analysis Protocols (INPLASY) on 27 September 2025 and was last updated on 27 September 2025.

INTRODUCTION

Review question / Objective This study aimed to quantify the pooled diagnostic performance of AI algorithms for lung tumour detection through systematic review and meta-analysis, addressing (1) comparative performance versus control methods; (2) algorithm family differences; and (3) sources of heterogeneity.

Condition being studied Lung cancer remains the leading cause of cancer-related mortality worldwide.Artificial intelligence (AI) has emerged as a promising technology to enhance lung cancer detection accuracy. In this study, we address three questions: (1) the pooled sensitivity and specificity of AI; (2) the comparative performance of algorithm families; and (3) study-level factors driving heterogeneity.

METHODS

Participant or population This study aimed to quantify the pooled diagnostic performance of AI algorithms for lung tumour detection.

Intervention N/A.

Comparator N/A.

Study designs to be included Studies evaluating AI diagnostic accuracy for lung nodule/tumour detection with sufficient data for 2×2 contingency tables, comparison with histopathology or multidisciplinary consensus reference standards and full-text availability were included. Reviews, case reports, conference abstracts and studies with insufficient diagnostic data were excluded.

Eligibility criteria (1) studies evaluating AI algorithms for lung nodule or tumour detection and classification; (2) studies reporting sufficient data on diagnostic performance measures (at least one 2×2 table derivable); (3) studies including a comparison with a reference standard; and (4) full-text availability for thorough assessment.
Studies were excluded under the following conditions: (1) review articles, case reports, conference abstracts or letters without primary data; (2) studies lacking sufficient information to calculate diagnostic performance; (3) duplicate publications or studies with overlapping datasets; (4) studies focusing solely on technical aspects of AI development without reporting diagnostic performance; and (5) preprints.

Information sources A comprehensive systematic search was conducted across seven databases :PubMed, Embase, Cochrane Library, Web of Science, SinoMed, China National Knowledge Infrastructure and Wanfang.

Main outcome(s) 1. Specificity of Artificial Intelligence for Lung Tumour Detection 2. Sensitivity of Artificial Intelligence for Lung Tumour Detection.

Quality assessment / Risk of bias analysis Study quality was assessed using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool, which evaluates risk of bias and applicability concerns across four domains: patient selection, index test, reference standard, and flow and timing. Each domain was classified as having low, unclear or high risk of bias based on the information provided in the study.

Strategy of data synthesis Meta-analysis was performed using Review Manager 5.4 software (The Cochrane Collaboration, Copenhagen, Denmark). The primary outcome measures were the comparative diagnostic performance between AI algorithms and control methods for lung tumour detection and classification. Risk ratios (RRs) with 95%CIs were calculated to compare the sensitivity, specificity and accuracy between AI groups (experimental) and control groups, with RR>1 denoting AI outperforming the control method. Pooled RRs were estimated using random-effects models.
Heterogeneity among studies was assessed using the chi-square test and the I^2 statistic. The I^2 statistic quantifies the proportion of inter-study variance relative to total variance; I^2 values of 25%, 50% and 75% were considered to represent low, moderate and high heterogeneity, respectively. A fixed-effects model was employed when heterogeneity was low to moderate (I^2<50%), and a random-effects model was applied when heterogeneity was significant (I^2≥50%). Pre-specified subgroups were algorithm family and training database; leave-one-out sensitivity analysis was used to assess each study's influence. Meta-regression was conducted to explore sources of heterogeneity. Publication bias was assessed using Egger's regression and Begg's rank-correlation tests (two-tailed, α=0.05).
Statistical significance was set at P<0.05 for all analyses. The results are presented as forest plots showing individual and pooled effect estimates with corresponding 95%CIs.

Subgroup analysis Pre-specified subgroups were algorithm family and training database; leave-one-out sensitivity analysis was used to assess each study's influence:Subgroup analysis by AI algorithm type and Subgroup analysis by learning database.

Sensitivity analysis leave-one-out sensitivity analysis was used to assess each study's influence.

Country(ies) involved Multiple countries.

Keywords deep learning; computer-aided detection; diagnostic accuracy; Quality Assessment of Diagnostic Accuracy Studies-2; systematic review.

Contributions of each author
Author 1 - Yongming Zheng.
Author 2 - SuBo Zhang.
Author 3 - Lei Geng.
Author 4 - Jiajing Li.
Author 5 - Kaimin Shi.
Author 6 - Di Yao.
Author 7 - Qiuzi Li.
Author 8 - Tao Zhang.
Author 9 - Yusheng Zhao.
Author 10 - Lei Zhang.
Author 11 - Yi Sun.
Author 12 - Jinxin Wan.