INPLASY

INPLASY202570031

doi: 10.37766/inplasy2025.7.0031

Received: 8 July 2025

Published: 8 July 2025

Corresponding author:
Xingyuan Liu

365979543@qq.com

Author Affiliation:
Departments of Radiology, Second Affiliated Hospital, Harbin Medical University.

Diagnostic Accuracy of Artificial Intelligence (AI) and Radiomics for Axillary Lymph Node Metastasis in Breast Cancer: A Systematic Review and Meta-Analysis Protocol

Liu, XY; Ruan, Y; Gao, B.

ADMINISTRATIVE INFORMATION

Support - This work was supported by the National Natural Science Foundation of China (62172129).

Review Stage at time of this submission - Preliminary searches.

Conflicts of interest - None declared.

INPLASY registration number: INPLASY202570031

Amendments - This protocol was registered with the International Platform of Registered Systematic Review and Meta-Analysis Protocols (INPLASY) on 8 July 2025 and was last updated on 8 July 2025.

INTRODUCTION

Review question / Objective 1. Quantify pooled diagnostic accuracy (sensitivity, specificity, DOR, AUC) of preoperative AI/radiomics models for predicting ALNM, validated against SLNB/ALND pathology.
2. Assess heterogeneity from five technical covariates:
• Segmentation method: Automatic, semi-automatic, manual
• Classifier type: Statistical models, machine learning, deep learning
• Feature integration: Imaging-only, imaging+clinical
• Contrast enhancement: Enhanced (DCE-MRI/CE-CT/CEM/CEUS) vs. non-enhanced
• ROI target: Primary tumor, axillary lymph node, combined
3. Develop evidence-based selection criteria for optimal AI configurations in axillary staging.

Rationale The axillary lymph node status is a critical determinant of staging and treatment strategies in breast cancer. Current standard diagnostic methods—including ultrasound-guided biopsy and surgical sentinel lymph node biopsy (SLNB)—are either limited by high false-negative rates (up to 25% for ultrasound) or invasiveness (surgical complications in 10-15% of SLNB cases). Emerging artificial intelligence (AI) and radiomics approaches offer non-invasive alternatives for preoperative metastasis prediction, yet their clinical adoption is hampered by unresolved technical heterogeneity. Key variables such as segmentation methods (manual vs. automated), classifier architectures (statistical vs. machine learning vs. deep learning), and feature integration strategies directly impact diagnostic reliability but remain systematically unassessed. This meta-analysis will be the first to quantify how these technical covariates influence accuracy, addressing a pivotal gap in translating AI tools to clinical axillary staging workflows. By establishing evidence-based criteria for optimal model configuration, we aim to accelerate standardized implementation, reduce unnecessary invasive

procedures, and personalize surgical planning for 2.3 million new breast cancer patients annually.

Condition being studied Axillary lymph node metastasis (ALNM) refers to the spread of breast cancer cells from the primary tumor to lymph nodes in the ipsilateral axilla. As a key component of breast cancer staging (AJCC 8th edition), ALNM status directly determines:
Prognosis: 5-year survival drops from 99% (node-negative) to 86% (1–3 nodes+) and 57% (≥4 nodes+).
Treatment decisions:
Node-negative: Sentinel lymph node biopsy (SLNB) alone
Node-positive: Axillary lymph node dissection (ALND) ± radiotherapy
Systemic therapy escalation (e.g., chemotherapy for >3 nodes+)
Current diagnostic challenges include:
① Limited sensitivity of imaging: Ultrasound misses 20–30% of micrometastases (<2 mm).
② Invasive gold standard: SLNB causes lymphedema (10–15%), sensory loss, and mobility impairment.
③ Overtreatment risks: Up to 40% of SLNB-positive patients receive unnecessary ALND.
This meta-analysis focuses on non-invasive AI/radiomics tools to address these gaps by enabling accurate preoperative ALNM prediction.

METHODS

Search strategy The search strategy employs terms covering: (1) Disease/Population: "Axillary Lymph Node Metastasis" OR "ALNM" OR "SLNB" OR "Sentinel Lymph Node Biopsy" OR "ALND" OR "Axillary Lymph Node Dissection"; (2) Index Test: "Artificial Intelligence" OR "Deep Learning" OR "Radiomics" OR "Radiomic"; (3) Imaging: "MRI" OR "Ultrasound" OR "Mammography" OR "CT" OR "PET-CT"; (4) Outcome: "Sensitivity" OR "Specificity" OR "Diagnostic Accuracy". PubMed syntax: ("axillary lymph node metastasis"[Title/Abstract] OR ALNM[Title/Abstract] OR SLNB[Title/Abstract] OR "axillary lymph node dissection"[Title/Abstract]) AND ("breast cancer"[Title/Abstract]) AND ("Radiomics"[Title/Abstract] OR "Radiomic"[Title/Abstract]) AND ("PET-CT"[Title/Abstract]) AND ("Sensitivity"[Title/Abstract]).检索策略包含：(1) 疾病/人群：腋窝淋巴结转移/ALNM/前哨淋巴结活检/SLNB/腋窝淋巴结清扫/ALND；(2) 待评价技术：人工智能/深度学习/影像组学/Radiomics/Radiomic；(3) 影像模态：MRI/超声/钼靶/CT/PET-CT；(4) 结局指标：敏感度/特异度/诊断准确性。PubMed检索式：("axillary lymph node metastasis"[Title/Abstract] OR ALNM[Title/Abstract] OR SLNB[Title/Abstract] OR "axillary lymph node dissection"[Title/Abstract]) AND ("breast cancer"[Title/Abstract]) AND ("Radiomics"[Title/Abstract] OR "Radiomic"[Title/Abstract]) AND ("PET-CT"[Title/Abstract]) AND ("Sensitivity"[Title/Abstract]).

Participant or population
Inclusion Criteria:
1. Breast cancer patients with histopathologically confirmed ALNM status via SLNB/ALND.
2. Development/validation of AI or radiomics models using preoperative imaging (MRI, US, mammography, CT, or PET-CT) for ALNM diagnosis.
3. Report of sufficient data (TP, FP, FN, TN) to calculate sensitivity/specificity.
Exclusion Criteria:
1. Duplicate patient cohorts.
2. Incomplete 2×2 contingency table data.
3. Non-preoperative imaging-based models.
4. Studies exclusively predicting ALN metastatic burden without diagnostic accuracy metrics.
5. Non-eligible publication types (reviews, case reports, etc.).

Intervention The index test under evaluation is the application of artificial intelligence (AI) and radiomics methodologies to preoperative medical imaging (including MRI, ultrasound, mammography, CT, or PET-CT) for diagnosing axillary lymph node metastasis (ALNM). This encompasses the entire technical workflow from image acquisition and region-of-interest (ROI) segmentation (via automatic, semi-automatic, or manual methods) to feature extraction—where radiomics quantifies traditional features (e.g., texture, shape, intensity) while deep learning autonomously derives hierarchical features—and subsequent model development using statistical, machine learning, or deep learning classifiers to generate binary ALNM classification (positive/negative). Critical covariates for subgroup analysis include segmentation methods, classifier architectures, feature integration strategies (imaging-only vs. imaging+clinical), imaging modalities, and ROI targets (primary tumor, lymph node, or combined).

Comparator In this diagnostic accuracy meta-analysis, the reference standard is histopathological examination of axillary lymph nodes obtained through surgical procedures:
Sentinel lymph node biopsy (SLNB) for clinically node-negative (cN0) patients

Axillary lymph node dissection (ALND) for patients with confirmed SLNB metastasis or clinically node-positive (cN+) status
Histopathology must adhere to international diagnostic criteria (e.g., AJCC 8th edition), with metastasis defined as: Macrometastasis (>2 mm)
Micrometastasis (0.2–2 mm)
Isolated tumor cells (ITCs, 95% specificity and >99% sensitivity for detecting ALNM when comprehensive nodal sampling is performed (Ann Surg Oncol 2021;28:2237). This standard aligns with NCCN Guidelines (v.3.2024) recommendations for axillary staging.

Study designs to be included This review will include: Diagnostic cohort studies (prospective or retrospective) Cross-sectional studies with consecutive or random patient sampling Diagnostic test accuracy studies reporting AI/radiomics models validated against histopathology Excluded designs: Case-control studies (due to high risk of spectrum bias)Prognostic/predictive modeling studies without diagnostic accuracy metricsCase reports, reviews, conference abstracts without full data Rationale: Cohort and cross-sectional designs align with STARD-AI guidelines, providing unbiased estimates of se.

Eligibility criteria
Supplementary Inclusion Criteria:
Sample size: ≥30 patients (ensuring model stability)
Time interval: ≤8 weeks between preoperative imaging and surgery (minimizing disease progression bias)
Data accessibility: Sufficient data to reconstruct 2×2 contingency tables (TP/FP/FN/TN)
Model validation: Internal/external validation cohort reporting
Supplementary Exclusion Criteria:
Neoadjuvant therapy: Studies including patients receiving chemotherapy before imaging
Image quality: Non-diagnostic image quality (e.g., motion artifacts affecting feature extraction)
Gray literature: Conference abstracts without accessible full data
Non-English/Chinese publications: Untranslated articles in other languages
Quality Control:
Exclude studies where >20% of images failed AI segmentation/feature extraction.

Information sources Electronic databases: PubMed, EMBASE, Web of Science, Scopus, Cochrane Library, CNKI, Wanfang Data, CBM. Grey literature: RSNA/SABCS conference abstracts (2018-2023), ClinicalTrials.gov, WHO ICTRP, arXiv (cs.CV/q-bio), and bioRxiv. Supplementary methods include backward/forward citation tracking, contacting authors for missing 2×2 data (3 email attempts), and consulting RSNA AI Committee experts. Limits: English/Chinese publications (2000-2025).

Main outcome(s) The primary outcomes are pooled sensitivity and specificity estimated via bivariate random-effects model (95% CI), diagnostic odds ratio (DOR) calculated using DerSimonian-Laird random-effects model (stratified by imaging modality), and area under the hierarchical summary receiver operating characteristic curve (HSROC AUC) quantifying overall diagnostic accuracy, all measured within ≤8 weeks from preoperative imaging to surgical pathology. Secondary outcomes include positive/negative likelihood ratios (PLR/NLR) interpreted via Fagan nomogram thresholds (PLR>10 or NLR50% deemed significant) with subgroup analyses (segmentation methods, classifier types, ROI targets), and clinical utility metrics (false-negative rate, false-positive rate). Effect measures will be visualized through forest plots for sensitivity/specificity/DOR, SROC curves with confidence regions, and Fagan plots for pretest-posttest probability conversion.

Quality assessment / Risk of bias analysis Risk of bias will be assessed using: QUADAS-2.
Data will be assessed by one person (or a machine) and checked by at least one other person (or machine).
The assessment wiil be done at study level.
There more two reviewers will be involed in the quality assessment.
Look for third parties other than reviewers to evaluate disagreements.

Strategy of data synthesis Primary outcomes (pooled SEN/SPE) analyzed via bivariate random-effects model; summary AUC via HSROC model. Secondary: DOR (random-effects), PLR/NLR (Fagan nomogram application). Heterogeneity: I^2 statistics, subgroup analyses (segmentation/classifier/ROI/QUADAS-2 risk), meta-regression (sample size/year). Sensitivity: Leave-one-out, Bayesian models, clinical scenario restrictions. Publication bias: Deeks' funnel plot asymmetry test ($p<0.10$). Software: R (mada/metafor), Stata (midas).

Subgroup analysis Prespecified subgroups: (1) Segmentation methods (auto/semi-auto/manual), (2) Classifier types (statistical/ML/DL), (3) Feature integration (imaging-only/imaging+clinical), (4) Imaging modalities (MRI/US/mammography/CT/PET-CT), (5) ROI targets (primary tumor/LN/

combined). Analyzed via bivariate meta-regression with interaction tests (p10% difference in SEN/SPE between subgroups. Sensitivity: Merge small subgroups (<3 studies); exclude studies with unreported data.

Sensitivity analysis Three-pronged approach: (1) Methodological: Compare bivariate vs. HSROC models; Bayesian sensitivity analysis; (2) Exclusion-based: Remove high-bias studies (QUADAS-2 ≥2 high-risk domains), small samples (n2mm), single-modality subgroups. Robustness threshold: 80% 95% CI overlap. Technical covariate impact: ΔAUC >0.05 for auto-segmentation studies.

Language restriction Only English and Chinese publications will be included.

Country(ies) involved China.

Keywords Breast cancer; Axillary lymph node metastasis; Artificial intelligence; Radiomics; Deep learning; Diagnostic accuracy; Sensitivity; Specificity; Meta-analysis.

Contributions of each author

Author 1 - Xingyuan Liu.
Email: 365979543@qq.com
Author 2 - Xingyuan Ruan.
Email: ruanye2000@163.com
Author 3 - Bo Gao.
Email: gaobo72519@hrbmu.edu.cn