INPLASY

INPLASY202520053

doi: 10.37766/inplasy2025.2.0053

Received: 10 February 2025

Published: 10 February 2025

Deep Learning-Based Breast MRI for Predicting Axillary Lymph Node Metastasis: A Systematic Review and Meta-Analysis

Wu, WP; Lee, CF.

Corresponding author:
Chia-Fen Lee

mini124719@gmail.com

Author Affiliation:
Changhua Christian Hospital.

ADMINISTRATIVE INFORMATION

Support - None.

Review Stage at time of this submission - Data analysis.

Conflicts of interest - The authors declare no conflicts of interest.The authors have no relevant financial or non-financial interests to disclose.

INPLASY registration number: INPLASY202520053

Amendments - This protocol was registered with the International Platform of Registered Systematic Review and Meta-Analysis Protocols (INPLASY) on 10 February 2025 and was last updated on 10 February 2025.

INTRODUCTION

Review question / Objective The diagnostic performance of deep learning algorithms applied to breast MRI in predicting axillary lymph node metastases in breast cancer patients.

Rationale Breast cancer is a leading cause of morbidity and mortality among women, highlighting the need for effective early detection and staging methods. Accurate assessment of axillary lymph node status is crucial for determining treatment strategies. Although sentinel lymph node biopsy (SLNB) and axillary lymph node dissection (ALND) are the gold standards, their invasiveness and associated risks emphasize the need for non-invasive alternatives.

Recent advances in deep learning applied to breast MRI offer a promising non-invasive solution for assessing lymph node metastasis. These algorithms can improve diagnostic accuracy by analyzing complex imaging data, potentially identifying lymph node involvement earlier. This approach reduces reliance on invasive procedures, improving patient outcomes, shortening recovery times, and lowering healthcare costs, particularly for early-stage breast cancer patients.

This systematic review and meta-analysis aim to evaluate the diagnostic accuracy of deep learning-based MRI in predicting lymph node metastases in breast cancer, providing a comprehensive assessment of its clinical efficacy.

Condition being studied Breast cancer is one of the most common malignant tumors in women. According to the results of the Z11 and NSABP B-32 studies, sentinel lymph node biopsy (SLNB) has been proven to have comparable prognostic outcomes to traditional axillary lymph node dissection (ALND) while significantly reducing the incidence of postoperative complications. However, traditional lymph node metastasis detection methods are invasive and carry certain risks. Therefore, developing and establishing a non-invasive method to assess the metastasis status of sentinel and axillary lymph nodes would help improve diagnostic efficiency, reduce patient discomfort, and minimize surgical complications, which is especially important for patients requiring early diagnosis.

METHODS

Search strategy

1. Pubmed(Advanced search):
(artificial intelligence) AND(lymph) AND (breast),
(deep learning) AND (lymph)AND (breast cancer),
(neural network) AND (breast)AND (lymph) AND (MRI).
2. Medline(Multi-field search):
(deep learning) AND (lymph) AND (breast),
(lymph) AND (mri) AND (neural network),
(breast) AND (lymph) AND (mri) AND (neural network).
3. Embase(Advanced search):
(artificial intelligence) AND (lymph) AND (breast),
(deep learning) AND (lymph) AND (breast cancer),
(neural network) AND (breast) AND (lymph) AND (MRI).

Participant or population Inclusion: Patients diagnosed with early stage breast cancers and received breast MRI.
Exlucsion: Patients received neoadjuvant chemotherapy.

Intervention (I) inclusion of patients diagnosed with breast cancers; (II) deep learning using breast MRI images were applied to predict axillary lymph nodes metastases; (III) sufficient data were present in terms of predictive performance of the deep learning algorithms; (IV) original research articles. Studies without enough information to calculate true positive (TP), true negative (TN), false positive (FP), and false negative (FN) values were excluded.

Comparator Not applicable.

Study designs to be included We will include original research studies that apply deep learning algorithms to breast MRI for predicting axillary lymph node metastases in breast cancer patients. Studies must report performance metrics like sensitivity, specificity, accuracy, and AUC.

Eligibility criteria Types of study to be included: We will include original research studies, that involve the application of deep learning algorithms to breast MRI images for predicting axillary lymph node metastases in breast cancer patients. The studies must provide sufficient data to allow the calculation of performance metrics, including true positive (TP), true negative (TN), false positive (FP), and false negative (FN) values. We will focus on studies that report on the predictive performance of these algorithms, particularly in terms of sensitivity, specificity, accuracy, and area under the curve (AUC).
Exclusion criteria: Studies without enough information to calculate TP, TN, FP, and FN values will be excluded. Additionally, studies that are not original research (e.g., reviews, commentaries, or editorials) will not be included in this review.

Information sources PubMed, MEDLINE, Embase.

Main outcome(s) Sensitivity, specificity, accuracy, and AUC of deep learning-based breast MRI in predicting axillary lymph node metastasis in breast cancer patients.

Quality assessment / Risk of bias analysis To assess the risk of bias and applicability, we will use two primary tools:

QUADAS-AI Tool: This specialized tool evaluates the risk of bias and applicability in AI-based diagnostic accuracy studies. It will help assess study quality, focusing on patient selection, AI model implementation, and the reference standard used.

CLAIM Criteria: The CLAIM (Checklist for Artificial Intelligence in Medical Imaging) criteria will provide a more comprehensive quality assessment.

Two reviewers will independently evaluate the risk of bias and quality of each study, ensuring a rigorous process. Any disagreements will be resolved through discussion. We will assess internal validity based on the criteria of the QUADAS-AI tool and CLAIM, focusing on the robustness and reliability of study design, data collection, and result reporting.

Strategy of data synthesis Statistical analysis was performed using R version 4.4.0. The pooled proportion analysis of diagnostic accuracy estimates, including sensitivity and specificity, was conducted using a random-effects model with 95% confidence intervals (CIs). This model was selected to account for the expected variability between studies, both within each study (intra-study sampling errors) and between different studies (inter-study variances). The random-effects model generally produces wider confidence intervals than the fixed-effects model, reflecting the diverse nature of the included studies.

The predictive accuracy of the deep learning models was quantified by pooling sensitivity and specificity, both with 95% CIs. Heterogeneity among studies was assessed using the I^2 statistic, where values of 25%, 50%, and 75% were interpreted as low, moderate, and high heterogeneity, respectively. A p-value < 0.10 was considered indicative of significant heterogeneity across studies.

The analysis will also explore the effect of heterogeneity by using the I^2 index to quantify the variation across studies and further assess the robustness of the overall pooled estimates.

Subgroup analysis We plan to conduct subgroup analyses to investigate potential effect modifiers based on the target lesions of primary tumor solely orthe primary tumor and axillary lymph nodes, AI algorithm models, and reference standards used in the studies. The rationale for these subgroup analyses is to determine whether variations in diagnostic performance (e.g., sensitivity, specificity, AUC) are influenced by factors such as study settings or the specific type of deep learning algorithm used.

Sensitivity analysis Sensitivity analysis will be conducted to evaluate the robustness of the pooled estimates and the influence of study characteristics on the overall results. Heterogeneity between studies will be quantified using the inconsistency index (I^2), with values of 25%, 50%, and 75% representing low, moderate, and high heterogeneity, respectively. A p-value < 0.10 indicates the presence of significant heterogeneity. In addition, a random-effects model will be used to account for both intra-study sampling errors and inter-study variances, as it assumes significant diversity among studies. To further assess the variability in sensitivity and specificity estimates, a cross-hairs plot will be generated. Sensitivity analysis will also examine the impact of excluding studies with high risk of bias or small sample sizes on the pooled diagnostic accuracy metrics, including sensitivity, specificity, and diagnostic odds ratio (DOR).

Country(ies) involved Taiwan.

Keywords Deep Learning, Breast Magnetic Resonance Imaging, Lymph Node Metastasis.

Contributions of each author

Author 1 - Chia-Fen Lee.
Email: mini124719@gmail.com
Author 2 - Wen-Pei Wu.
Email: 139496@cch.org.tw