



Artificial Intelligence In The Diagnosis And Prediction Of Breast Cancer-Related Lymphedema: A Scoping Review

Shely Kagan¹ · Lyndsey Huynh² · David Chen² · Caro Strickland³ · Christina Yang⁴ · Jennifer Y. Y. Kwan^{5,6} · Shing Fung Lee⁷ · Adrian Wai Chan⁸ · Srinivas Raman⁸ · Edward Chow⁹ · Henry C. Y. Wong¹⁰

Received: 29 July 2025 / Accepted: 7 April 2026
© Crown 2026

Abstract

Introduction Lymphedema is a chronic complication of breast cancer treatments that can significantly impact the well-being of survivors. This scoping review aims to evaluate the role of artificial intelligence (AI) in enhancing diagnostic precision and supporting timely interventions.

Methods PubMed, Scopus, EMBASE, Web of Science, and the Cochrane Central Register of Controlled Trials were searched from database inception until September 2024. Studies were included if they examined AI-based techniques for detecting lymphedema, assessing severity, and/or facilitating early diagnosis. Inclusion criteria required studies to be published in English with full-text availability. Editorials, review papers, and inaccessible full-text studies were excluded.

Results From 300 studies identified from the database search, 13 studies met the inclusion criteria. The investigated AI-based models used input data such as Electronic Health Record (EHR) and clinical data (5 studies, 38.5%), patient-reported symptoms and demographics (4 studies, 30.8%), imaging (3 studies, 23.1%), and clinical factors like BMI and hypertension (2 studies, 15.4%) for outcome prediction. The most commonly used AI model was built using the Support Vector Machine (SVM) algorithm, which appeared in 8 studies (61.5%) and was often combined with other supervised learning techniques. Risk prediction models achieved accuracy rates of 81% to 93.75%, with sensitivity of 95.65%, specificity of 91.03%, and Area Under the Curve (AUC) of 0.751. Models classifying lymphedema severity demonstrated accuracy rates between 81 and 91%, with the best-performing models achieving a balanced accuracy of 99.4% and AUC ranging from 0.889 to 0.931.

Conclusions AI demonstrates potential in improving the diagnosis and prediction of breast cancer-related lymphedema, offering enhanced diagnostic capabilities and personalized interventions. Further research is required to address data standardization, model validations, and the development of frameworks for clinical implementation.

Keywords Lymphoedema · Breast cancer · Artificial intelligence · Machine learning · Prediction model

✉ Henry C. Y. Wong
henrywong3011@gmail.com

¹ University of Western Ontario, London, Canada

² Temerty Faculty of Medicine, University of Toronto, Toronto, Canada

³ Faculty of Computer Science, University of Western Ontario, London, Canada

⁴ University of Waterloo, Kitchener, Canada

⁵ Radiation Medicine Program, Princess Margaret Cancer Centre, Toronto, Canada

⁶ Department of Radiation Oncology, University of Toronto, Toronto, Canada

⁷ Department of Radiation Oncology, National University Cancer Institute, Singapore, Singapore

⁸ Department of Radiation Oncology, British Columbia Cancer Agency - Vancouver, Vancouver, BC, Canada

⁹ Department of Radiation Oncology, Sunnybrook Health Sciences Centre, Odette Cancer Centre, University of Toronto, Toronto, ON, Canada

¹⁰ Department of Oncology, Princess Margaret Hospital, Kowloon West Cluster, Hospital Authority, 6/F, Block H, Princess Margaret Hospital, 2-10 Princess Margaret Hospital Road, Lai Chi Kok, Kowloon, Hong Kong S.A.R., China

Introduction

Lymphedema is a chronic and often debilitating complication of breast cancer treatment, particularly among survivors undergoing axillary lymph node dissection. [1, 2] It is estimated that up to 1 in 5 breast cancer survivors experience lymphedema, underscoring its significance in survivorship care. [1] This condition impacts physical health and imposes substantial psychological and social burdens. [2] Symptoms such as swelling, pain, reduced mobility, and recurrent infections negatively impact the quality of life and functional independence of survivors. [3] Despite advancements in cancer care, early detection and effective management of lymphedema remain challenging. [4]

Traditional methods for detecting lymphedema rely on subjective evaluations including patient-reported outcomes, or gross changes in tape measurements, which may be insufficient for the detection of early lymphoedema. This gap has fueled interest in innovative approaches leveraging Artificial Intelligence (AI) that can more sensitively predict and detect early lymphedema before definitive clinical symptoms are present. [5] By processing complex, multimodal data, AI techniques capture useful patterns that can improve diagnostic precision and accurately predict disease progression to augment clinical decision-making.

The integration of AI into healthcare, while promising, requires careful evaluation. [5] Its potential to revolutionize diagnostic and treatment paradigms must be weighed against challenges in implementation, such as data limitations, algorithm biases, and the need for clinical validation in real-world settings. This scoping review aims to address these issues by assessing the current state of AI applications in the prevention, diagnosis, and treatment of breast cancer-related lymphedema. The objectives of this study include evaluating the performance of AI models, identifying knowledge gaps, and proposing recommendations to optimize the integration of AI into clinical practice.

Methods

Search Protocol and Databases

This review was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses—Scoping Reviews (PRISMA-ScR) guidelines. A comprehensive literature search was performed across PubMed (including MEDLINE), Scopus, EMBASE, Web of Science, and the Cochrane Library, covering all available studies until September 2024. The search strategy employed a

combination of controlled vocabulary and free-text terms related to lymphedema, artificial intelligence, and clinical outcomes. Key search terms included “lymphedema,” “artificial intelligence,” “machine learning,” “deep learning,” “diagnosis,” and “treatment.” Boolean operators were used to refine results, ensuring inclusivity across prevention, diagnosis, and treatment domains.

Selection Process

Results from bibliographic databases and trial registries were imported into Covidence, a web-based software tool designed to facilitate the production of systematic reviews by removing duplicates and assisting with screening. [6] Covidence automatically eliminated duplicate records, and the remaining references were screened by two independent reviewers (S.K, L.H) based on the eligibility criteria. In cases of discrepancies, a third reviewer (H.CY.W) was consulted to reach a resolution.

After manually removing any missed duplicates and completing the title and abstract screening, full-text screening was carried out by two independent reviewers for studies that met the eligibility criteria for this phase. Any disagreements were resolved through discussion or by referring to a third reviewer (H.CY.W) when necessary.

Inclusion and Exclusion Criteria

Studies were included only if they evaluated breast cancer-related arm lymphedema, either in patients diagnosed with the condition or those at risk of developing it due to treatments such as surgery or radiotherapy for breast cancer. Only studies exploring the use of AI were considered, including machine learning (ML), a method where computers learn from data to make predictions or decisions without being explicitly programmed; fuzzy models, which handle uncertain or imprecise information in a way that mimics human reasoning; and natural language processing. The focus was on applications of these AI methods in the prevention, diagnosis, or severity of lymphedema. Comparisons between AI-based approaches and traditional methods of managing lymphedema were eligible, as well as studies offering insights into AI’s role without a direct comparison. Included studies contained outcomes such as clinical improvements (such as better diagnosis, early detection, or enhanced management), technical measures (e.g., accuracy, sensitivity, specificity), and practical aspects of AI implementation in clinical settings. Studies were excluded if they did not focus on lymphedema or AI applications related to lymphedema, as well as animal studies, in-vitro research, non-peer-reviewed articles, case studies, conference abstracts, reviews

articles, and studies published in languages other than English.

Data Extraction

For this review, data extraction was conducted by two independent reviewers, with a third reviewer involved if necessary to resolve discrepancies. The extracted information included general study characteristics such as the article name, authors, publication year, journal, and country/region of study. Additionally, the focus of the AI application, specific aims, and the type of AI model were recorded. Specific AI techniques used, and data type inputs were noted, along with model training methodology and performance metrics. Additional information such as sample size, age group, inclusion/exclusion criteria, underlying health conditions, AI prediction targets, and the clinical outcomes predicted were also extracted. The prediction timeline, clinical validation (if available), model accuracy, limitations, and data challenges were documented.

Results

Study Selection

The initial database search yielded 436 studies. After excluding 136 duplicate records, 300 articles underwent title and abstract screening with 284 excluded based on inclusion and exclusion criteria. A total of 16 studies progressed to full-text review, with 3 excluded due to a focus on non-breast cancer patient populations. This left 13 studies eligible for data extraction and synthesis (Fig. 1).

Geographic Scope and Focus of Studies

The 13 studies included in this review represent a global effort to apply AI in predicting, diagnosing, and detecting breast cancer-related lymphedema, with research spanning North America ($n=2$, 15.4%), South America ($n=1$, 7.7%) Europe ($n=3$, 23.1%), and Asia ($n=7$, 53.8%). Four studies aimed to enhance diagnostic accuracy (30.8%), five aimed facilitate early prediction of lymphedema (38.5%), and three aimed personalize treatment plans for affected patients (23.1%). The median publication year of 2022 and a range of 13 (2011–2024) indicates a growing interest in using AI for addressing this complex medical issue, with an increasing number of researchers exploring its potential.

Factors Used in the Models

The models utilized various patient-related factors, cited in 11 (84.6%) studies, as well as radiological and

treatment-related factors, reported in 5 (38.5%) and 6 (46.2%) studies respectively, to enhance prediction accuracy. Patient-related factors included demographics (age, BMI), medical history (comorbidities such as diabetes, cardiovascular diseases), and surgical history. Radiological factors encompassed imaging reports from modalities such as MRI, CT scans, and mammograms, with extracted features including tumor size, lymph node involvement, and tissue density. Treatment-related factors comprised details on chemotherapy regimens, exposure to radiation therapy, type of surgical intervention (lumpectomy, mastectomy, axillary dissection), and the presence of post-treatment complications.

The studies reviewed showed a wide range of sample sizes for developing and validating AI models for breast cancer-related lymphedema. Large datasets included Carleton's analysis of 580,000 clinical documents (with the number of unique patients unspecified), [18] Notash (2022) with 1,117 patients, [13] Peng's retrospective study of 3,201 participants, [14] and Trinh's 2,137 patients. [15] Medium-sized datasets consisted of Du's 670 breast cancer patients, [7] Wei's 533 participants, [16] Wu's 370 patients, [17] Fu's 355 participants, [8] and Nascimben's multicentric dataset of 294 patients. [11] Smaller datasets involved Kistenev's 76 individuals, [9] Notash (2021) with data from 30 individuals (60 arms), [12] Vicentini's fuzzy logic system tested on 7 patients, [19] and Moreira's Kinect-based assessment, for which the sample size was not reported. [10]

AI Techniques and Applications

Recent studies have highlighted the role of AI-driven tools, such as ML algorithms, in enhancing the accuracy, scalability, and timeliness of lymphedema management (Table 1). The AI techniques used in these studies varied widely, reflecting the specific objectives of each model (Table 2). For example, Natural Language Understanding (NLU) (1 study) was employed [18] to analyze unstructured clinical notes from electronic health records (EHRs). This method successfully identified lymphedema incidence with an accuracy rate of 81%, demonstrating the potential of AI in extracting valuable insights from the vast amounts of unstructured data routinely recorded in clinical settings. In another study, fuzzy logic systems were used [19] to classify the severity of lymphedema. These systems exhibited high reliability in clinical practice, categorizing severity levels based on various inputs to ensure timely, appropriate care. Additionally, ML techniques such as LASSO regression (1 study) and decision trees (6 studies) were applied for risk stratification and predictive modeling. [7, 8] These methods helped identify patients at high risk for developing lymphedema, enabling earlier interventions that could prevent or mitigate the condition.

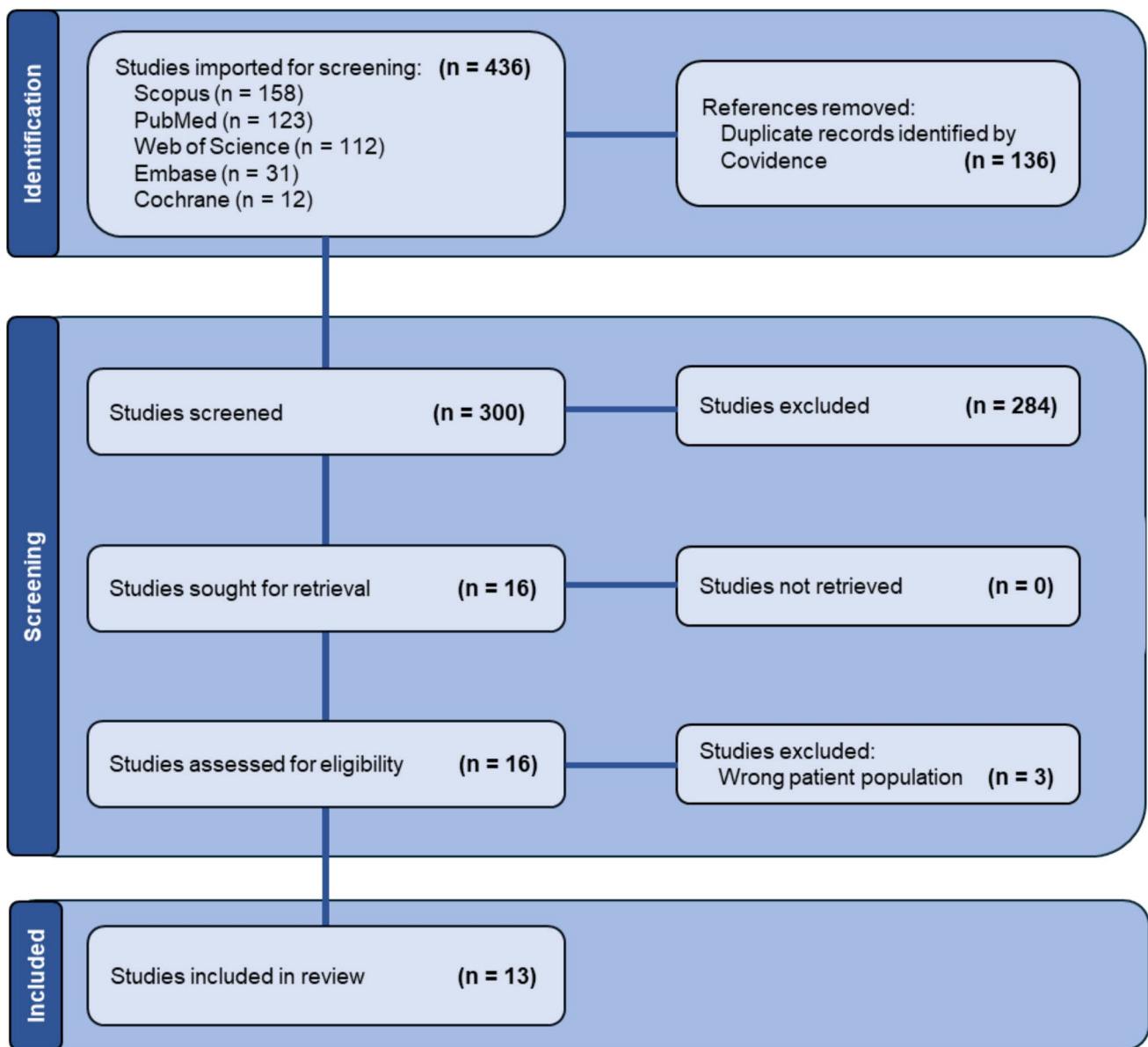


Fig. 1 PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) Chart

AI Models and Performance

Performance outcomes from these AI models were promising. For instance, NLU was utilized to process clinical and demographic data, achieving 81% accuracy in identifying lymphedema. [18] Studies employing ML techniques, such as gradient boosting [8] and decision trees [7, 8] by reached cross-validation accuracies of up to 93.75%, [8] highlighting the potential of AI to detect early signs of lymphedema. Furthermore, fuzzy logic systems demonstrated success in classifying the severity of lymphedema using clinical parameters like skin changes and pitting. [19] Input data for these models included EHRs, clinical variables, imaging data, and

patient-reported outcomes (PROs), allowing for the creation of more personalized models to address the multifaceted nature of lymphedema. A summary of these diverse AI approaches, their specific aims, and reported performance metrics is presented in Table 3.

Most studies (8 total, 61.5%) employed long-term prediction horizons ranging from one year to over 20 years post-treatment. Validation methods varied across studies, with many employing internal validations through cross-validation, with resulting Area Under the Curve (AUC) performance ranging from 0.81 to 0.93. [14, 16] These AUC values indicated strong predictive performance, particularly in early detection and severity classification. Few studies

Table 1 AI Models in Lymphedema Prediction and Diagnosis

AI Models	Description	Studies	Domains Examined
Machine Learning Algorithms	Algorithms used for prediction, classification, or regression	Du et al ⁷ , Fu et al ⁸ , Kitsenev et al ⁹ , Moreira et al ¹⁰ , Nascimben et al ¹¹ , Notash et al. (2021) ¹² , Notash et al. (2022) ¹³ , Peng & Lu ¹⁴ , Trinh et al ¹⁵ , Wei et al ¹⁶ , Wu et al ¹⁷	Lymphedema detection, risk prediction, severity prediction
Deep Learning Models	Advanced neural networks used to analyze large datasets and images	Fu et al ⁸ , Notash et al. (2021) ¹² , Peng & Lu ¹⁴ , Trinh et al ¹⁵ , Wei et al ¹⁶	Lymphedema prediction, image analysis, outcome prediction
Natural Language Understanding	Natural Language Understanding used to process clinical data	Carleton et al ¹⁸	Identification of Lymphedema, pathologically positive lymph nodes (pN1) prediction
Imaging Studies	Imaging data analyzed using AI models for diagnosis and prediction	Vicentini et al ¹⁹ , Notash et al. (2021) ¹²	Lymphedema tissue analysis, arm volume measurement

presented data on the external validation of the developed models.

Discussion

The studies reviewed demonstrate the significant potential of AI in the prediction, detection, and management of breast cancer-related lymphedema (Table 3). AI models have shown high predictive accuracy, with some studies reporting values as high as 99.9% for balanced accuracy, AUC = 0.999 for training and 0.931 for external validation, [15] enabling earlier identification of patients at risk for lymphedema. The ability of AI to analyze diverse data sources, including EHRs, [17, 18] imaging, [7, 12, 18] and patient-reported outcomes, [8, 19] underscores its transformative role in improving patient care. This review is particularly important as lymphedema remains a common yet underdiagnosed condition among breast cancer survivors. [1, 4] By highlighting AI's role in early detection and personalized intervention, this review contributes to the ongoing efforts to enhance lymphedema management and improve patient outcomes.

AI models offer significant advantages in identifying and diagnosing breast cancer-related lymphedema (Fig. 2). One of the most notable benefits is the potential to enhance early detection, a crucial factor for effective intervention. [15] Traditional clinical methods such as subjective measurements often fail to identify lymphedema in its early stages, which can lead to worsening symptoms and long-term disability. [1, 4] AI-driven tools, particularly those using ML algorithms like gradient boosting and decision trees, have demonstrated the ability to predict the risk of lymphedema development much earlier than conventional methods. [8] These findings suggest that AI has the potential to provide healthcare professionals with crucial early warnings, enabling

them to initiate preventive measures sooner. Incorporating AI into prospective surveillance programs may be able aid in selection of high-risk patients to be enrolled, improve the efficiency and cost-effectiveness of these programs, and thereby reducing the workload of healthcare professionals.

Another advantage of AI in lymphedema care is its ability to personalize treatment. The use of diverse inputs allows AI models to create individualized care plans based on a patient's unique medical history and current condition. For example, the fuzzy logic system used for severity classification of lymphedema allows for more nuanced decision-making by providing healthcare professionals with precise recommendations on the appropriate level of intervention for each patient. [19] This approach contrasts with traditional methods that often apply one-size-fits-all guidelines, which may not fully account for the individual variation in lymphedema development and progression.

Despite the promising results, the studies also revealed several key limitations that need to be addressed in future research. One major challenge was the quality of the data used to train the models. Missing data, particularly in relation to self-reported symptoms, and variability in clinical parameters posed significant hurdles for building robust, generalizable models. [8, 9, 13, 14, 16, 17, 19] Self-reported outcomes were prone to inconsistencies, which affected the reliability of the models that relied on them. Another limitation was the lack of generalizability of the findings, [14, 16, 17] as many of the datasets used in the studies were small, (with the smallest number of patients being 7) and homogenous, [9, 10, 16, 17] limiting the applicability of the models to broader, more diverse patient populations. This lack of diversity in the datasets may have impacted the models' ability to accurately predict outcomes for patients from different demographic backgrounds, highlighting the need for more inclusive research. [20] Additionally, some studies

Table 2 Study Characteristics

First Author	Year	Country	Sample Size	Prediction Timeline	Exact Time Horizon	Application (AI Used)
Carleton et al ¹⁸	2024	United States	925	Long-term	5.5 years (median)	Large Language Model—NLU
Du et al ⁷	2024	China	670	Long-term	Post-surgery	Machine Learning—LASSO (for variable selection), comparison of models (including techniques like RF, SVM, and DT)
Fu et al ⁸	2018	United States	355	Long-term	1–5 years (up to 20 post-cancer treatment)	Machine Learning—DT C4.5, DT C5.0, GBM, ANN (using feed forward multi-layer network), SVM [using radial basis function (RBF) kernel function]
Vicentini et al ¹⁹	2011	Brazil	7	Long-term	N/A	Fuzzy Logic System—Mamdani Fuzzy Inference
Kistenev et al ⁹	2019	Russia	34	Short-term	N/A	Machine Learning – SVM
Moreira et al ¹⁰	2015	Portugal	Not specified	Short-term	Immediately after cancer treatment(s)	Machine Learning – Fisher LDA, Naïve Bayes, SVM
Nascimben et al ¹¹	2023	Italy	294	Long-term	854.85 days (2 years, 4 months)	Machine Learning—UMAP
Notash et al ¹²	2021	Iran	60	Short-term	Immediate prediction after measurement	Machine Learning—EEFSL, SVM, KNN, ANFIS, DT, Naïve Bayes
Notash et al ¹³	2022	Iran	1,117	Long-term	N/A: Post-surgery and treatment (months or years after treatment)	Machine Learning – DT C5.0, SVM (with different kernels Linear, RBF, and Polynomial), KNN, MLP (shallow)
Peng & Lu ¹⁴	2024	China	3,201	Long-term	6 + months	Machine Learning—GLRM, RF model, and ANN model
Trinh et al ¹⁵	2023	Korea	2,137	Short-term	Immediately or prior to observable symptoms	Machine Learning—RF, GBM, DT, LR, ANN
Wei et al ¹⁶	2021	China	533	Short-term	Within the first stages post-surgery, identified by a 1 cm increase in arm circumference compared to the healthy side	Machine Learning – LR, RF, ANN, SVM, CART, DT C5.0
Wu et al ¹⁷	2022	China	370	Long-term	28.73 months on average after surgery	Machine Learning—Naïve Bayes Model, K-KNN model, SVM, LR, MLP

Abbreviations: NLU, natural language understanding; LASSO, least absolute shrinkage and selection operator; RF, random forest; SVM, support vector machine; DT, decision trees; GBM, gradient boosting model; ANN, artificial neural network; LDA, linear discriminant analysis; UMAP, uniform manifold approximation and projection; EEFSL, evolutionary ensemble feature selection learning; KNN, K-nearest neighbour; MLP, multilayer perceptron; GLRM, Generalized Logistic Regression Model; LR, Logistic Regression; CART, Classification and Regression Tree; ANFIS, Adaptive Neuro-Fuzzy Inference System

lacked external validation, which is crucial for ensuring that AI models can perform effectively in real-world clinical settings. [8, 9, 14, 16, 17] Without external validation, the confidence in the model's ability to generalize to new populations and healthcare environments remains uncertain. [20] Future research should place greater emphasis on rigorous external validation through concrete strategies such as the use of multi-center datasets and cross-institutional collaborations to evaluate model performance across diverse populations and care settings.

Future research should focus on expanding the datasets used to train AI models. This includes using data from

diverse patient populations to ensure the generalizability of models across various demographic groups. Additionally, incorporating more objective data into AI models, such as measurements from wearable sensors and devices such as perometry and bioimpedance spectroscopy may improve model reliability and accuracy. [5] Beyond model development, greater attention is needed to the real-world integration of AI into clinical practice. Practical challenges such as clinician trust, workflow integration, model transparency, and the cost of adoption remain significant barriers to implementation and warrant further exploration. Another important area for future research is the development of hybrid

Table 3 AI Model Effectiveness

First Author	Specific Aim of AI	Effectiveness	Key Insights
Carleton et al ¹⁸	Determine rates of pathologically positive lymph nodes (pN1) and lymphedema	15% false-positive rate, 0% false-negative rate, 100% sensitivity, 95% specificity, 95% overall accuracy for detecting breast/arm lymphedema, and 93% per-patient accuracy	Supports omitting SLNB in select ER +/cN0 breast cancer patients and highlights NLU's potential for better clinical data extraction
Du et al ⁷	Prediction of lymphedema development	Logistic Regression (LR) outperformed nine models with 0.81 accuracy, 0.79 sensitivity and recall, 0.82 specificity, 0.60 precision, 0.68 F-score, and 0.87 AUC in predicting BCRL	LASSO regression identified BMI as key in predicting BCRL, with the LR model effectively identifying high-risk patients
Fu et al ⁸	Early detection of lymphedema in breast cancer survivors	The model achieved 93.75% average cross-validation accuracy, 95.65% sensitivity, 91.05% specificity, and an AUC of 0.751	A well-trained ANN classifier using real-time symptom reports offers more accurate lymphedema detection than traditional methods
Vicentini et al ¹⁹	Prediction of lymphedema severity and classification	The fuzzy classification method produces results that are similar to the SBL approach, suggesting good accuracy in differentiating severity levels, particularly when dealing with patients exhibiting overlapping clinical characteristics	Approach enhances precision in diagnosis by allowing both qualitative and quantitative evaluations, making it more adaptable for clinical decision-making
Kistenev et al ⁹	Detection of lymphedema tissue remodeling	Sobel Operator: 0.59, Canny: 0.571, Morphology: 0.62, LoG: 0.571; Best results with sensitivity: 0.79 ± 0.11 , specificity: 0.77 ± 0.10	SVM model using machine learning and collagen analysis show promise for automated, accurate lymphedema diagnosis
Moreira et al ¹⁰	Prediction of lymphedema development	Accuracy of 81% (SVM with polynomial/RBF kernel, 0.19 error rate) with high precision but low recall	Uses Kinect and machine learning to assess upper-body function and lymphedema in breast cancer patients, showing promising classification accuracy
Nascimben et al ¹¹	Predicting risk of lymphedema development	Balanced accuracy of 99.4% (RF); 99.0% (ET, RF for two labels)	Developed a prognostic map using clinical features to categorize patients into clusters, helping assess BCRL risk
Notash et al ¹²	Measurement of arm volume	The EEFSL model's hyperparameters iteratively adjusted until convergence, achieving a high correlation of $R = 0.994$, indicating a global best solution	The HVIS-EEFSL method, optimized using genetic algorithms, offers a reliable and clinically suitable alternative to traditional arm volume measurement techniques for patients with lymphedema
Notash et al ¹³	Prediction of lymphedema development	91% accuracy	The study highlights the use of ensemble learning to predict lymphedema risk factors, supporting digital technologies in management
Peng & Lu ¹⁴	Predicting risk of lymphedema development	The RF model achieves AUCs of 0.894 (training) and 0.889 (verification) with 95% CI of 0.839–0.949 and 0.834–0.944, while the ANN model shows AUCs of 0.819 (training) and 0.807 (verification) with 95% CI of 0.762–0.876 and 0.750–0.864	The interpretable RFM model improves limb lymphedema risk prediction, enabling early identification and personalized treatment to reduce complications
Trinh et al ¹⁵	Early detection of lymphedema	The RF model achieves 99.9% balanced accuracy, sensitivity, specificity, precision, F1 score, and an AUC of 0.999 on the training set, while on the external validation dataset, it shows 87.0% balanced accuracy, 84.3% sensitivity, 89.1% specificity, 97.4% precision, 90.4% F1 score, and an AUC of 0.931	Developed machine learning models using blood test and therapy data for early lymphedema detection and stage prediction, with an open-access web app for doctors

Table 3 (continued)

First Author	Specific Aim of AI	Effectiveness	Key Insights				
Wei et al ¹⁶	Prediction of lymphedema development	For early detection: accuracy of 0.819–0.838 For late detection: accuracy of 0.863–0.894. Best model: AUC: 0.889; Sensitivity: 77.1%; Specificity: 88.3%	The logistic regression model enables early lymphedema detection and real-time monitoring through an open-access web app				
Wu et al ¹⁷	Prediction of lymphedema development	Five ML models could accurately predict the occurrence of BCRL (ranging from 0.783 to 0.916) 91.6% (LR), 0.975 (MLP, AUC)	Found that multilayer perceptron and logistic regression models best predicted BCRL				
First Author	Year	Sample Size	AI Focus	Prediction	Timeline	Exact Time Horizon	Application (AI Used)
Carleton et al ¹⁸	2024	925	Diagnosis	Long-term		5.5 years (median)	Large Language Model—NLU
Du et al ⁷	2024	670	Prediction	Long-term		Post-surgery	Machine Learning—LASSO (for variable selection), comparison of models (including techniques like RF, SVM, and DT)
Fu et al ⁸	2018	355	Prediction	Long-term		1–5 years (up to 20 post-cancer treatment)	Machine Learning—DT C4.5, DT C5.0, GBM, ANN (using feed forward multi-layer network), SVM [using radial basis function (RBF) kernel function]
Vicentini et al ¹⁹	2011	7	Prediction	Long-term		N/A	Fuzzy Logic System—Mamdani Fuzzy Inference
Kistenev et al ⁹	2019	34	Diagnosis	Short-term		N/A	Machine Learning—SVM
Moreira et al ¹⁰	2015	Not specified	Prediction	Short-term		Immediately after cancer treatment(s)	Machine Learning—Fisher LDA, Naïve Bayes, SVM
Nascimben et al ¹¹	2023	294	Prediction	Long-term		854.85 days (2 years, 4 months)	Machine Learning—UMAP
Notash et al ¹²	2021	60	Prediction	Short-term		Immediate prediction after measurement	Machine Learning—EFSL, SVM, KNN, ANFIS, DT, Naïve Bayes
Notash et al ¹³	2022	1,117	Prediction	Long-term		N/A: Post-surgery and treatment (months or years after treatment)	Machine Learning—DT C5.0, SVM (with different kernels Linear, RBF, and Polynomial), KNN, MLP (shallow)
Peng & Lu ¹⁴	2024	3,201	Prediction	Long-term		6+ months	Machine Learning—GLRM, RF model, and ANN model
Trinh et al ¹⁵	2023	2,137	Prediction	Short-term		Immediately or prior to observable symptoms	Machine Learning—RF, GBM, DT, LR, ANN
Wei et al ¹⁶	2021	533	Prediction	Short-term		Within the first stages post-surgery	Machine Learning—LR, RF, ANN, SVM, CART, DT C5.0
Wu et al ¹⁷	2022	370	Prediction	Long-term		28.73 months on average after surgery	Machine Learning—Naïve Bayes Model, K-KNN model, SVM, LR, MLP

Abbreviations: AI, artificial intelligence; ML, machine learning; AUC, area under the curve; SVM, standard vector model; LoG, Laplacian of Gaussian; RBF, radial basis function; RF, random forest; ET, envelope tracking; ANN, artificial neural network; MLP, multilayer perceptron; BCRL, breast cancer-related lymphedema; SLNB, sentinel lymph node biopsy

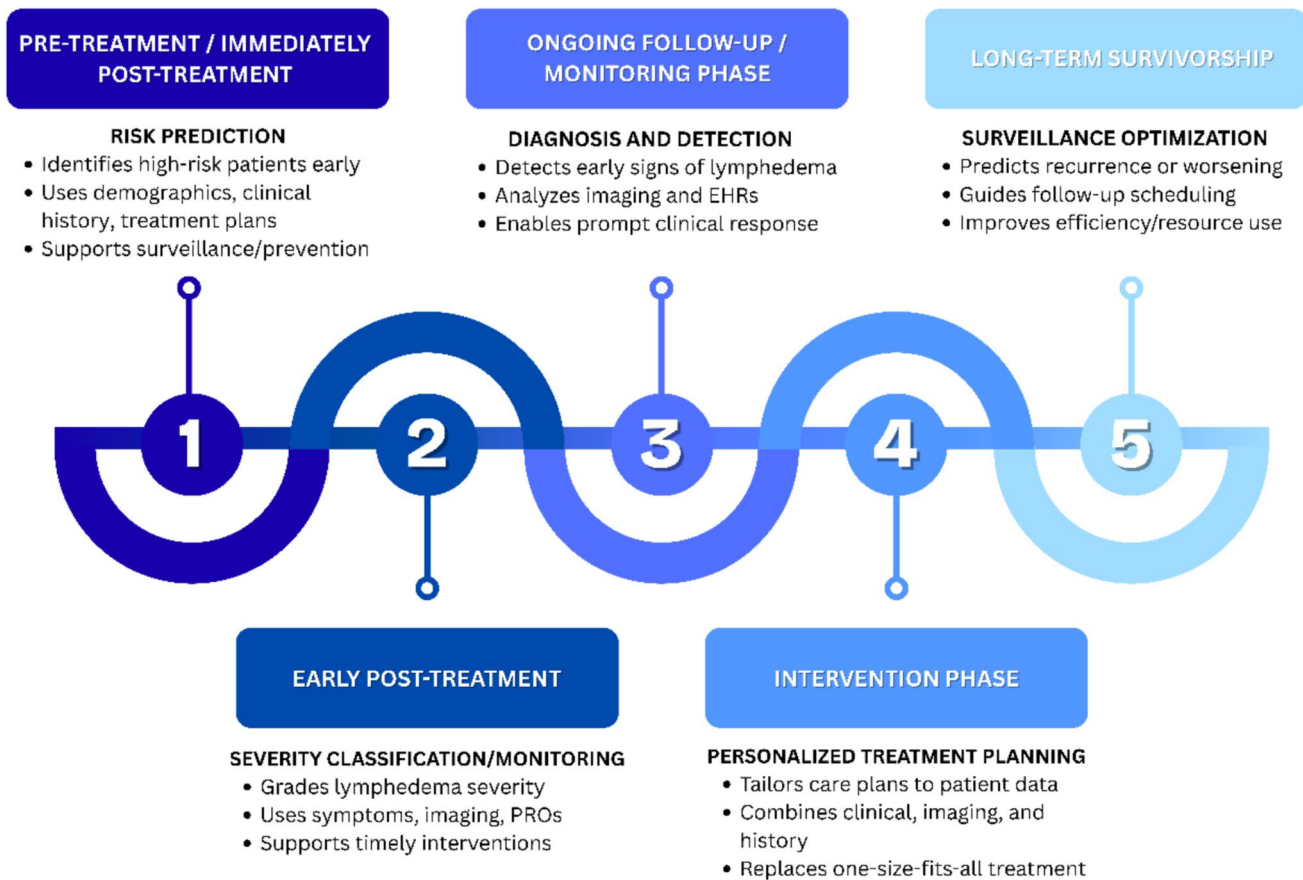


Fig. 2 Timeline of AI Applications in the Prediction, Diagnosis, and Management of Breast Cancer-Related Lymphedema

models that combine AI with traditional clinical expertise. While AI has the potential to enhance clinical decision-making, it should complement rather than replace the insights of experienced healthcare professionals. Implementation studies that embed AI tools within multidisciplinary care teams and assess usability, interpretability, and feasibility could generate robust evidence to guide how AI may be applied to deliver more holistic and clinically meaningful care for patients with lymphedema. This review has several potential limitations. First, many included studies relied on small, homogenous datasets, limiting the generalizability of AI models across diverse populations. Additionally, there was significant variability in AI methodologies, feature selection, and performance metrics, making direct comparisons difficult. The issue of missing or incomplete data was mentioned in 6 studies (46.2%), particularly in patient-reported outcomes, may also introduce bias in model training and predictions.

A key concern is the lack of external validation as many studies showed promising results using internal datasets but did not test their models on independent populations, raising concerns about real-world applicability. Furthermore, AI model transparency remains a significant challenge, as

many models function as black boxes with limited clinical interpretability. [21] These black box systems are often highly complex, particularly deep learning models such as neural networks, making it difficult to understand how specific inputs such as patient characteristics or imaging features lead to particular outputs such as diagnoses or risk predictions. This lack of interpretability poses a substantial barrier in clinical settings, where clinicians must be able to understand, justify, and communicate the rationale underlying diagnostic and treatment decisions. As a result, the black box nature of many AI models can undermine clinician trust and accountability, limiting their adoption in routine practice. Greater emphasis on the development and evaluation of explainable AI systems that provide transparent, interpretable, and clinically meaningful explanations for model recommendations is therefore essential. Incorporating explainable AI approaches may enhance clinician confidence, support informed decision making, and facilitate safer and more ethical integration of AI tools into clinical workflows. [22] Addressing these limitations (through diverse datasets, standardized evaluation, external validation, and improved model interpretability) will be essential for advancing AI-driven solutions in lymphedema care.

Ethical considerations are an important component of the clinical translation of AI and warrant greater attention, particularly with respect to algorithmic bias and fairness. AI models trained on non representative or imbalanced datasets may inadvertently perpetuate or amplify existing health disparities, leading to unequal performance across demographic groups and care settings [20, 21]. This risk is especially relevant in healthcare, where biased predictions may contribute to inequitable diagnosis, risk stratification, or treatment recommendations. Strategies to mitigate algorithmic bias include systematic data auditing to identify gaps or skewed representations within training datasets, as well as bias aware model design approaches that explicitly evaluate and adjust for differential performance across subpopulations [22]. Incorporating fairness assessments and transparency measures into model development and validation processes may help promote more equitable and accountable use of AI in clinical practice, complementing efforts to improve model generalizability, interpretability, and clinician trust [21, 22].

Conclusion

AI driven technologies hold significant potential to transform the prevention, diagnosis, and management of breast cancer related lymphedema by improving diagnostic accuracy, enabling earlier intervention, and supporting more personalized care. However, the successful translation of these tools into clinical practice requires addressing persistent challenges related to data standardization, external validation, ethical considerations, and real world implementation. Future research should prioritize the development of standardized and representative datasets, leverage multi modal data sources including imaging, clinical text, and objective biomarkers, and adopt rigorous multi center validation strategies. In parallel, greater emphasis on model transparency, fairness, and explainability will be essential to promote clinician trust, ensure equitable performance across patient populations, and support accountable clinical decision making. Through sustained interdisciplinary collaboration among clinicians, researchers, and data scientists, AI has the potential to meaningfully advance lymphedema care and improve quality of life for breast cancer survivors worldwide.

Author Contributions Conceptualization: Henry CY Wong, Shely Kagan; Methodology: Shely Kagan, Henry CY Wong; Formal analysis and investigation: Shely Kagan, Lyndsey Huynh, Henry CY Wong; Writing—original draft preparation: Shely Kagan; Writing—review and editing: Shely Kagan, Lyndsey Huynh, David Chen, Caro Strickland, Christina Yang, Jennifer Y Y Kwan, Shing Fung Lee, Adrian Chan, Srinivas Raman, Edward Chow, Henry C Y Wong; Funding acquisition: N/A; Resources: N/A; Supervision: Henry CY Wong.

Data Availability No datasets were generated or analysed during the current study.

Declarations

Conflicts of interest The authors declare no competing interests.

Financial support and sponsorship None.

References

- Gillespie TC, Sayegh HE, Brunelle CL, Daniell KM, Taghian AG (2018) Breast cancer-related lymphedema: risk factors, precautionary measures, and treatments. *Gland Surg* 7(4):379–403. <https://doi.org/10.21037/gs.2017.11.04>
- Ridner SH (2009) The psycho-social impact of lymphedema. *Lymphat Res Biol* 7(2):109–112. <https://doi.org/10.1089/lrb.2009.0004>
- Fu MR, Ridner SH, Hu SH, Stewart BR, Cormier JN, Armer JM (2012) Psychosocial impact of lymphedema: a systematic review of literature from 2004 to 2011. *Psychooncology* 22(7):1466–1484. <https://doi.org/10.1002/pon.3201>
- Tse SS, Wong C, Tane K, Kikawa Y, Rafn BS, Chan AW et al (2024) Prospective surveillance and early intervention to prevent chronic breast cancer-related arm lymphedema—what are the barriers? *Ann Palliat Med* 13(6):1513–1520. <https://doi.org/10.21037/apm-24-93>
- Eldaly AS, Avila FR, Torres-Guzman RA, Maita K, Garcia JP, Serrano LP, Forte AJ (2022) Artificial intelligence and lymphedema: State of the art. *J Clinical Translational Res* 8(3):234–242
- Covidence Systematic Review Software VHI [Software]. Melbourne, Australia: Veritas Health Innovation, 2023. www.covidence.org
- Du J, Yang J, Yang Q, Zhang X, Yuan L, Fu B (2024) Comparison of machine learning models to predict the risk of breast cancer-related lymphedema among breast cancer survivors: a cross-sectional study in China. *Front Oncol*. <https://doi.org/10.3389/fonc.2024.1334082>
- Fu MR, Wang Y, Li C, Qiu Z, Axelrod D, Guth AA et al (2018) Machine learning for detection of lymphedema among breast cancer survivors. *mHealth* 4:17–17. <https://doi.org/10.21037/mhealth.2018.04.02>
- Kistenev YV, Nikolaev VV, Kurochkina OS, Borisov AV, Vrazhnov DA, Sandykova EA (2019) Application of multiphoton imaging and machine learning to lymphedema tissue analysis. *Biomed Opt Express* 10(7):3353. <https://doi.org/10.1364/boe.10.003353>
- Moreira R, Magalhães A, Oliveira HP (2015) A Kinect-based system to assess lymphedema impairments in breast cancer patients. *Lecture Notes Computer Sci* 228–36. https://doi.org/10.1007/978-3-319-19390-8_26
- Nascimben M, Lippi L, de Sire A, Invernizzi M, Rimondini L (2023) Algorithm-based risk identification in patients with breast cancer-related lymphedema: a cross-sectional study. *Cancers* 15(2):336. <https://doi.org/10.3390/cancers15020336>
- Notash AY, Bayat P, Haghighat S, Notash AY (2021) Evolutionary ensemble feature selection learning for image-based assessment of lymphedema arm volume. *Concurrency Computat Pract Exp*. <https://doi.org/10.1002/cpe.6334>
- Notash AY, Notash AY, Omid Z, Haghighat S (2022) Prediction of lymphedema occurrence in patients with breast cancer using the optimized combination of ensemble learning algorithm and

- feature selection. *BMC Med Inform Decis Mak.* <https://doi.org/10.1186/s12911-022-01937-z>
14. Peng X, Lu Z (2024) Development and validation of upper limb lymphedema in patients after breast cancer surgery using a practicable machine learning model: a retrospective cohort study. *Int J Gen Med* 17:3799–3812. <https://doi.org/10.2147/ijgm.s478573>
 15. Trinh X-T, Chien PN, Long N-V, Van Anh LT, Giang NN, Nam S-Y et al (2023) Development of predictive models for lymphedema by using blood tests and therapy data. *Sci Rep.* <https://doi.org/10.1038/s41598-023-46567-1>
 16. Wei X, Lu Q, Jin S, Li F, Zhao Q, Cui Y et al (2021) Developing and validating a prediction model for lymphedema detection in breast cancer survivors. *Eur J Oncol Nurs* 54:102023. <https://doi.org/10.1016/j.ejon.2021.102023>
 17. Wu X, Guan Q, Cheng ASK, Guan C, Su Y, Jiang J et al (2022) Comparison of machine learning models for predicting the risk of breast cancer-related lymphedema in Chinese women. *Asia Pac J Oncol Nurs* 9(12):100101. <https://doi.org/10.1016/j.apjon.2022.100101>
 18. Carleton N, Saadawi G, McAuliffe PF, Soran A, Oesterreich S, Lee AV, et al (2024) Use of natural language understanding to facilitate surgical de-escalation of axillary staging in patients with breast cancer. <https://doi.org/10.1101/2024.02.03.24302095>
 19. Vicentini P, del Carmen M, Perez J, Araujo E (2011) Fuzzy lymphedema assessment based on clinical and functional criteria. 2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011). 2708–13. <https://doi.org/10.1109/fuzzy.2011.6007676>
 20. Hanna MG, Pantanowitz L, Jackson B, Palmer O, Visweswaran S, Pantanowitz J et al (2025) Ethical and bias considerations in Artificial Intelligence/Machine Learning. *Mod Pathol* 38(3):100686. <https://doi.org/10.1016/j.modpat.2024.100686>
 21. Salahuddin Z, Woodruff HC, Chatterjee A, Lambin P (2022) Transparency of deep neural networks for Medical Image Analysis: a review of Interpretability Methods. *Comput Biol Med* 140:105111. <https://doi.org/10.1016/j.combiomed.2021.105111>
 22. Hassija V, Chamola V, Mahapatra A, Singal A, Goel D, Huang K et al (2023) Interpreting black-box models: a review on Explainable Artificial Intelligence. *Cogn Comput* 16(1):45–74. <https://doi.org/10.1007/s12559-023-10179-8>
- Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.