informatics

MDPI

Systematic Review

Feature-Centric AI for Breast Cancer Metastasis Analytics: A Systematic Review and Evidence-Based Framework for Feature Engineering and Multimodal Integration

Wellington Kanyongo [1,*] and Bester Chimbo [2]

[1] Department of Computer Science, School of Computing, College of Science, Engineering & Technology, University of South Africa (UNISA), 28 Pioneer Avenue, Florida Park, Roodepoort 1709, South Africa
[2] Department of Information Systems, School of Computing, College of Science, Engineering & Technology, University of South Africa (UNISA), 28 Pioneer Avenue, Florida Park, Roodepoort 1709, South Africa; chimbb@unisa.ac.za
* Correspondence: kanyow@unisa.ac.za; Tel.: +27-73-245-4283

Abstract

The importance of understanding how feature engineering contributes to breast cancer metastasis analytics is growing as artificial intelligence (AI) technologies rapidly transform the field of precision oncology. This systematic review collated, thematically synthesised, and analysed the evidence related to feature engineering processes, multimodal integration techniques and explainability mechanisms used in AI-driven breast cancer metastasis analytics. Following PRISMA 2020, studies from 2020 to 2026 were retrieved from MEDLINE, Scopus, Web of Science, Embase and IEEE Xplore. A total of 50 empirical studies investigating AI, feature engineering, radiomics and multimodal approaches for breast cancer metastasis prediction were included. The research findings reveal that contemporary AI-based breast cancer metastasis prediction is predominantly based on imaging-derived features, while clinical, pathological, biomarker and molecular variables increasingly enhance multimodal model development. Frequently applied techniques were observed across feature extraction, preprocessing, reproducibility assessment, dimensionality reduction, feature integration and explainability, all of which helped to create more interpretable, robust and clinically meaningful predictive systems. The study crafts an evidence-informed conceptual 5Fs framework comprising feature extraction or generation, feature preparation and transformation, feature selection and dimensionality reduction, feature construction and integration, and feature interpretability and explainability. The framework highlights the feature-centric nature of AI-driven breast cancer metastasis analytics and emphasises the importance of interpretable, multimodal and potentially clinically translatable predictive systems.

Keywords: artificial intelligence (AI); breast cancer; feature engineering; feature extraction; machine learning (ML); metastasis; radiomics; multimodal integration

Check for updates

Academic Editor: Jiang Bian

Received: 26 May 2026
Revised: 22 August 2026
Accepted: 24 August 2026
Published: 27 August 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Breast cancer remains the most commonly diagnosed cancer among women worldwide and continues to pose a significant public health challenge due to its high rates of incidence, recurrence and metastatic progression. Breast cancer caused about 694,000 deaths worldwide in 2024 and it was also the most frequently diagnosed cancer among women in 164 of 186 countries [1]. In 2024, approximately 2.4 million women were diagnosed

with breast cancer, which highlights the global burden and the need for effective new solutions [1]. Breast cancer is a condition in which abnormal cells in the breast grow out of control and can form tumors that can be deadly if left untreated. It typically starts in the milk ducts or lobules and the early stage (in situ) is generally not life-threatening and often detectable early if screening is sought [2]. Cancer cells that spread into nearby breast tissue form a lump or result in thickening of the breast tissue. In more advanced cases, invasive cancers can invade other organs or lymph nodes by metastasis, which is life-threatening and can lead to death [1]. Lymphadenopathy, bone, liver, lung and brain metastasis significantly worsen prognosis and treatment. It is thus crucial to identify the risk of metastasis at an early stage to improve prognosis, to better inform personalised treatment and to better inform clinical decisions.

Currently, the most common methods for predicting breast cancer metastasis are based on a histopathological examination, clinical staging, radiological analysis and molecular typing. While these strategies remain clinically relevant, they continue to have several limitations in predictive reliability and translational efficiency. Traditional imaging interpretation is often radiologist-dependent and can miss subtle quantitative signatures of intratumoral heterogeneity as metastatic progression unfolds [3–5]. Furthermore, the traditional statistical prediction models lack the capacity to incorporate multimodal data from heterogeneous imaging, biomarkers, pathological and molecular evidence [6]. Histopathological evaluation and clinicopathological scoring systems also have low sensitivity for early detection of metastatic transformation and are often difficult to interpret due to inter-observer variability [7]. In addition, traditional statistical prediction models are limited by their low-dimensional data-handling capacity and lack of integration between heterogeneous multimodal evidence from imaging, biomarkers, pathology and molecular datasets [6]. The shortcomings of these approaches have underscored the need for new computational systems capable of providing actionable clinical insights from large, complex oncology datasets.

The use of artificial intelligence (AI), machine learning (ML), deep learning (DL), and radiomics has revolutionised the analysis of breast cancer metastasis. The ability of AI-driven systems to extract high-dimensional quantitative imaging features, identify nonlinear relationships and integrate heterogeneous data sources for metastasis prediction and prognosis stratification is gaining in capability. Random Forest, support vector machine, XGBoost and logistic regression models are ML methods that have proven to be highly useful for predicting metastasis from clinical and radiomics data [8]. In a similar vein, DL setups including convolutional neural networks (CNNs), DenseNet, EfficientNet, and transformer-based systems have enhanced automated feature extraction, lesion characterisation and metastasis classification using imaging as well as pathology data [9–11]. Radiomics has also proven to be a translational direction that enables the extraction of non-invasive quantitative markers from routine medical images, which characterise tumor heterogeneity, microenvironmental dynamics and metastatic potential [12]. Artificial intelligence (AI) is also an emerging transformative technology and WHO has acknowledged its potential to enhance healthcare diagnostics, precision oncology and evidence-based decision-making in the healthcare sector, if used appropriately and clinically validated [13].

While there has been increasing interest in the scope of AI in breast cancer metastasis analytics, the currently available review studies remain fragmented and methodologically limited. Some recent reviews have largely centered on radiomics applications in breast imaging, overlooking the important discussion of the feature engineering processes underlying the development of metastasis prediction systems [14,15]. Other reviews focused predominantly on DL algorithms, imaging modalities, or specific metastatic endpoints, such as lymph node metastasis, while providing limited synthesis of multimodal feature

integration and explainability frameworks [10,16,17]. Additional reviews have covered AI use in cancer, but none have specifically explored the feature-engineering pipelines used to analyse breast cancer metastasis [18,19]. Existing literature also demonstrates limited conceptual integration of feature extraction, preprocessing, dimensionality reduction, multimodal fusion and explainability into a unified analytical framework. Moreover, there is relatively limited research work to synthesise how imaging radiomics can be integrated with clinicopathological, molecular, biomarker and pathology-derived evidence inside metastasis prediction ecosystems.

Recent reviews by [20–23] similarly emphasised ongoing challenges related to reproducibility, standardisation, external validation, data harmonisation and explainability in AI-driven oncology analytics. These methodological challenges suggest the need for a comprehensive synthesis specifically on feature engineering processes in breast cancer metastasis analytics. Consequently, the present study has been designed to systematically synthesise the feature processing techniques, feature engineering strategies, feature types, multimodal integration approaches and explainability mechanisms used in AI-based breast cancer metastasis analytics. This study also aims to provide a conceptual understanding of how diverse imaging, molecular, biomarker, pathological and clinical characteristics are integrated into clinically relevant evidence for predicting metastasis. The following research questions were addressed in this study:

1. What feature engineering techniques are employed in AI-based breast cancer metastasis analytics?
2. Which feature types are most commonly used in AI models for breast cancer metastasis prediction?
3. What evidence-informed conceptual feature-centric framework can be derived from feature engineering pipelines used in AI-driven breast cancer metastasis analytics?

Technical Contributions

This study makes the following technical contributions:

- It develops a feature-centered synthesis of AI-driven breast cancer metastasis analytics by showing how raw imaging, clinical, biomarker, molecular and pathology data are transformed into metastasis-relevant predictive features.
- It proposes the 5Fs Framework as a structured feature engineering pipeline comprising feature extraction or generation, feature preparation and transformation, feature selection and dimensionality reduction, feature construction and integration and feature interpretability and explainability.
- It identifies the key feature engineering techniques employed in breast cancer metastasis prediction across diverse methodological contexts.
- It maps the major feature categories employed in AI-based metastasis analytics, encompassing imaging modalities, clinicopathological variables, molecular and biomarker indicators, pathology-derived features and site-specific labels.
- It demonstrates that multimodal feature fusion is a central technical pathway for improving metastasis prediction by integrating radiomics, clinical, pathological, molecular and biomarker evidence.
- It provides a technical foundation for future AI-based breast cancer metastasis prediction systems by emphasising reproducibility, data harmonisation, external validation, multimodal integration and interpretable model deployment.

2. Materials and Methods

2.1. Review Design

The study design was based on the systematic review method to provide a synthesis of evidence related to feature engineering processes, feature types, multimodal integration strategies and explainability mechanisms used in AI-driven breast cancer metastasis analytics. Studies related to AI, ML, DL, radiomics and multimodal predictive systems for metastasis analytics and prediction in breast cancer were investigated in the review. The review process followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines, which promote methodological transparency, reproducibility and rigor in reporting. A structured review protocol was therefore adhered to in a manner that guided the identification of literature, screening, eligibility assessment, quality appraisal, data extraction and thematic synthesis to ensure that it was aligned with the research questions of this study. The PRISMA 2020 study selection process is presented in Figure 1, while the PRISMA 2020 for Abstracts Checklist and the PRISMA 2020 Checklist are provided in Appendices A.1 and A.2, respectively.

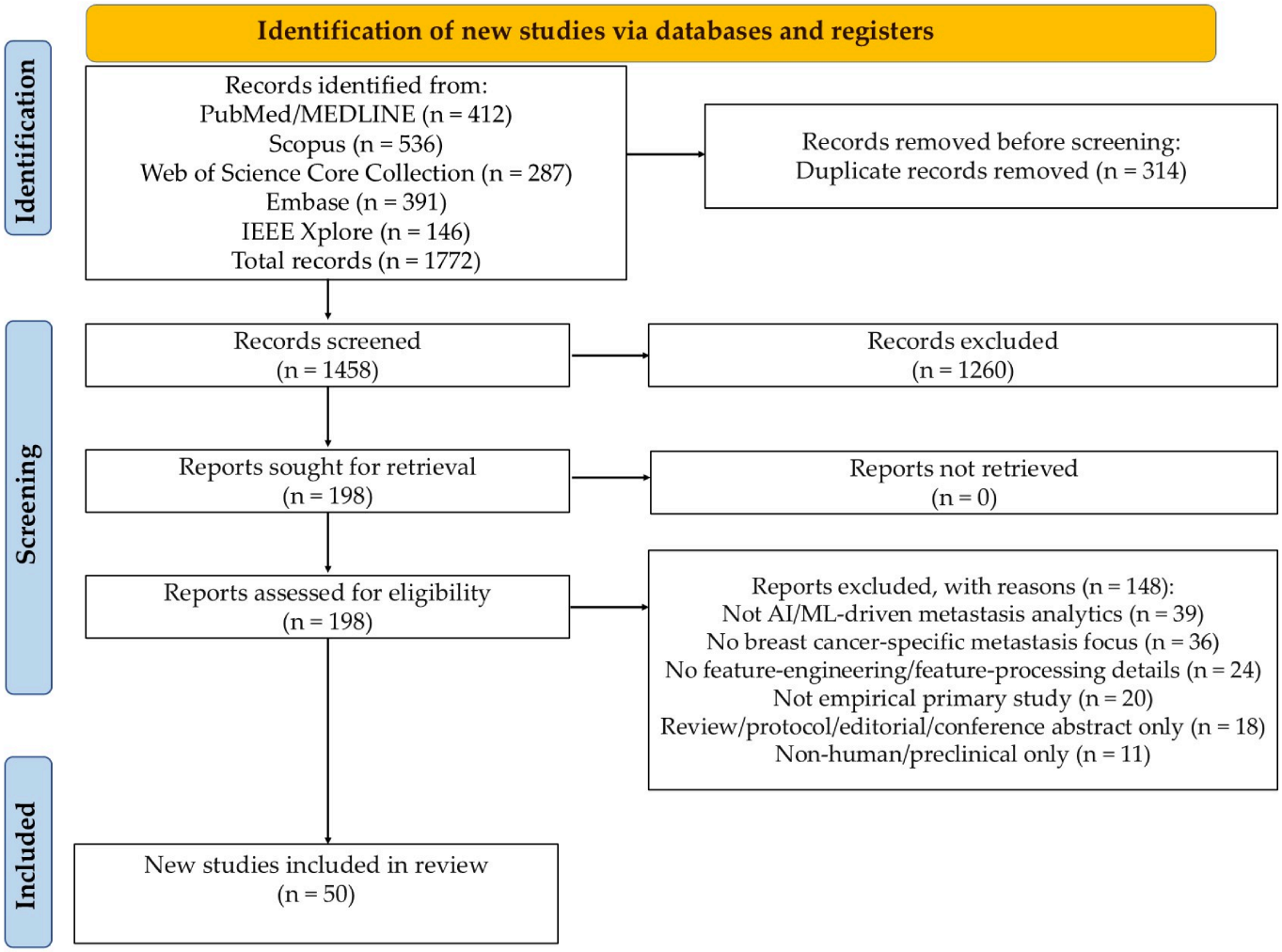


Figure 1. PRISMA study identification and selection flow diagram.

2.2. Information Sources and Search Strategy

A comprehensive literature search was conducted between 1 May 2026 and 15 May 2026 across multiple multidisciplinary and healthcare-related electronic databases including MEDLINE, Scopus, Web of Science, Embase and IEEE Xplore. These databases were

chosen as they cover a wide range of research in oncology, radiology, radiomics, AI, medical imaging and healthcare analytics. The reference lists of the eligible articles and relevant review studies were manually screened to identify additional studies. The search strategy was developed using keywords derived from the research questions and the thematic concepts central to breast cancer metastasis analytics. The search incorporated combinations of Boolean operators and keywords such as "breast cancer", "metastasis", "lymph node metastasis", "distant metastasis", "artificial intelligence", "machine learning", "deep learning", "radiomics", "feature engineering" and "multimodal integration". An example of the search string used was: ("breast cancer" OR "breast neoplasm") AND ("metastasis" OR "metastatic progression") AND ("artificial intelligence" OR "machine learning" OR "deep learning" OR "radiomics") AND ("feature engineering" OR "feature extraction" OR "feature processing" OR "feature selection" OR "prediction analytics").

2.3. Eligibility Criteria

Studies that examined the use of AI, ML, or DL in breast cancer metastasis analytics and prediction or multimodal analytical or feature engineering techniques were included. Studies were eligible for inclusion if they incorporated feature engineering processes, imaging-derived features, clinicopathological variables, molecular integration, biomarker integration and/or explainability processes related to metastasis analytics. Inclusion of articles was limited to full-text peer-reviewed articles published in English in the 2020–2026 period. Articles focusing on cancers other than breast cancer, non-metastatic outcomes, purely laboratory experiments without AI analytics, editorials, conference abstracts, protocols, dissertations and review articles were excluded. Additionally, studies that did not provide adequate methodological detail on feature engineering and/or predictive analytics were excluded from the synthesis.

2.4. Study Selection Process

Selection of studies was based on a sequential screening procedure that was based on predefined criteria. All studies retrieved from the various electronic literature databases were exported to a reference management system, where duplicate records were identified and removed electronically and manually before screening commenced. Subsequently, the titles and abstracts were screened independently to identify potentially relevant studies related to AI-driven breast cancer metastasis analytics. Studies that passed the initial screening were then assessed in full text to determine eligibility for final use in the synthesis. The screening process focused on relevance to feature engineering techniques, type of features, multimodal integration, explainability methods and outcomes of metastasis prediction. Any disagreements that arose during the screening and eligibility process were discussed, and a consensus was reached to ensure a consistent and reliable methodological approach in the selection process.

2.5. Quality Appraisal

The Mixed Methods Appraisal Tool (MMAT) was used to evaluate the methodological quality of the included studies. The MMAT was selected because the evidence comprised heterogeneous empirical quantitative studies, including retrospective diagnostic, prognostic, radiomics, machine-learning, deep-learning, computational pathology and biomarker-based studies. Thus, the appraisal was confined to the applicable MMAT criteria, within which relevant aspects of AI study design, data processing, measurement, analysis and validation were considered integral components of methodological quality rather than separate appraisal criteria. Since majority of the studies were quantitative non-randomised, the tool was used to ensure consistent appraisal of various designs within one framework and operationalised through MMAT screening questions and criteria 3.1–3.5: clarity of research

questions, appropriateness of data, participant representativeness, measurement validity, completeness of outcome data, control of confounders and suitability of statistical analysis.

Each MMAT criterion was rated "Yes", "No", or "Can't tell". A "Yes" rating indicated that the criterion was satisfied, a "No" rating indicated that it was not satisfied, and "Can't tell" indicated that the reported information was insufficient to determine whether the criterion was satisfied. In accordance with MMAT guidance, no overall numerical score or percentage was calculated. Instead, author-defined narrative categories were used to summarise the pattern of MMAT responses. Studies were classified as high quality when both screening questions were rated "Yes", and three to five of the criteria 3.1–3.5 were rated "Yes." Studies were classified as moderate quality when both screening questions were rated "Yes," but no more than two of the criteria 3.1–3.5 were rated "Yes." Only criteria rated "Yes" were regarded as satisfied when assigning the narrative quality category. These classifications are narrative summaries developed for this review and should not be interpreted as official MMAT categories. The detailed quality appraisal results are outlined in Appendix B.

2.6. Data Extraction and Synthesis

A structured extraction table was developed based on the study's research questions. Extracted data included author details, publication year, study aim, study methods, feature engineering techniques, feature types and AI approaches, as well as information regarding the results such as model performance indicators. The summary table of data extracted from the articles included is presented in Table 1 and Appendix C. Thematic synthesis was used to analyse the extracted evidence and find common methodological and analytical patterns found in the studies. The thematic synthesis was guided by thematic codes that were progressively refined to basic themes, organising themes and grand themes, following the method developed by [24] called thematic network analysis. Appendix C shows the data extraction summary table. The thematic interpretation also guided the development of the conceptual framework representing the feature-engineering pipeline in AI-driven breast cancer metastasis analytics.

Table 1. Feature-engineering techniques employed in AI-based breast cancer metastasis analytics and the feature types.

Grand Theme	Organising Theme	Basic Theme	Thematic Code	References
Feature-processing and engineering	Feature extraction or generation	Imaging radiomics extraction	PyRadiomics	[25–32]
			Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) radiomics	[25,31,33–38]
			Ultrasound radiomics	[27,28,39–41]
		Texture-feature extraction	Gray level co-occurrence matrix (GLCM) texture features	[25,29–33,37,38,42,43]
			Gray level run length matrix (GLRLM) texture features	[25,29–33,38,42]
			Gray level size zone matrix (GLSZM) texture features	[25,29–33,38,42]
			Gray level dependence matrix (GLDM) texture features	[25,29,31–33,35,38,42,43]
			Neighbourhood gray tone difference matrix (NGTDM) texture features	[25,29,30,32,38,43]
		Deep-feature extraction	CNN-derived deep features	[29,36,39,41,44–49]
			VGG/DenseNet/MobileNet/EfficientNet deep-feature extraction	[36,41,44,47,50]
		Pathology feature extraction	WSI patch extraction	[45,48–52]

Table 1. *Cont.*

Grand Theme	Organising Theme	Basic Theme	Thematic Code	References
Feature-processing and engineering	Feature preparation and transformation	Region of Interest (ROI) and image preparation	Manual ROI/Volume of Interest (VOI) segmentation	[25,27,28,30,31,33,36–38,42,43,51]
		Image/data standardisation	Z-score normalisation	[27,31,37,39,40,53–58]
		Reproducibility testing	Intra-class correlation coefficient (ICC) feature reproducibility filtering	[25,27,28,31–33,37,40,41,46]
		Data balancing	SMOTE oversampling	[26,36,42,59,60]
	Feature selection and dimensionality reduction	Penalized regression	Least absolute shrinkage and selection operator (LASSO) feature selection	[25,27,28,30,31,33,37,38,40,46,51,57,58,61–65]
		Redundancy reduction	Pearson correlation filtering	[30,33,37,43,51]
			Spearman correlation filtering	[40,51,55,57,58]
			MRMR feature selection	[27,28,40,61,62]
		ML-based feature ranking	Random Forest feature ranking or mean decrease accuracy/Gini	[29,54,66–68]
	Feature construction and integration	Radiomics-score construction	Rad-score/radiomics signature construction	[25,27–30,33,36–38,43]
		Nomogram construction	Clinical-radiomics or multimodal nomogram	[25,28,30,36,37,43,44,51,54,69]
		Clinical-imaging fusion	Imaging radiomics integrated with clinical/pathological variables	[25,27,28,30,35–38,44,46,60,70]
	Feature interpretability and explainability	SHAP explainability	SHAP feature attribution	[27,31,41,56,64,68,71]
		Visual explainability	Gradient-weighted Class Activation Mapping (Grad-CAM), class activation map (CAM), attention heatmaps or prototype visualisation	[29,39,41,45,47,49,52]
		Clinical utility testing	Decision curve analysis	[25,27–30,33,43,51,54,55,57,58,68,70]
Feature types	Imaging-derived feature types	MRI feature types	DCE-MRI features	[25,29,31,33–38,44,53]
			Diffusion-weighted imaging (DWI)/apparent diffusion coefficient (ADC) features	[29,31,37,42,53,60]
			Contrast-enhanced T1-weighted imaging (T1 + C)/T2-weighted imaging (T2WI) multiparametric MRI features	[29,46,60,62,64,65,71]
		Ultrasound feature types	B-mode ultrasound features	[27,28,39–41]
	Clinical and clinicopathological feature types	Tumor and nodal variables	Tumour size or lesion diameter	[28,38,51,58,68]
		Pathology variables	ER, PR, human epidermal growth factor receptor (HER2) or Ki-67 proliferation index (Ki-67) status	[29,38,40,43,48,51,58,64,65,68]
	Digital pathology and pathomics feature types	WSI pathology features	Haematoxylin and eosin (H&E) whole-slide image features	[45,48–52]
	Metastasis-site and outcome-related feature types	Lymph node metastasis outcomes	Axillary lymph node metastasis (ALNM) prediction	[25,27,28,30,31,33,39,40,48,51,53,62,65,70]
		Distant metastasis outcomes	Distant metastasis prediction	[49,55–58,61,69,72,73]
		Organ-specific metastasis outcomes	Bone metastasis prediction	[43,57,58,63,67,68]
			Bone/liver/lung/brain metastasis site labels	[32,43,57,58,67,68,74]

3. Results

3.1. Study Selection

A search of the electronic literature databases identified 1772 records from PubMed/MEDLINE, Scopus, Web of Science Core Collection, Embase and IEEE Xplore. A total of 314 duplicate records were removed, and 1458 studies were screened by title and abstract, of which 1260 were excluded. A total of 198 reports were retrieved and evaluated for eligibility. After full-text evaluation, 148 studies were excluded for the following reasons: no information on metastasis, no breast cancer-specific focus on metastasis, insufficient information on the features being engineered, non-empirical designs, review papers only, or preclinical studies. In the end, 50 studies were included in this systematic review.

3.2. Study Characteristics

All 50 included studies were related to breast cancer metastasis analytics, prediction, classification, or feature identification. The majority of the studies focused on lymph-node metastasis outcomes (30/50, 60%), such as axillary lymph node metastasis (ALNM), sentinel lymph node metastasis and high axillary lymph node burden. An additional 18 studies (36%) examined distant, organ-specific, recurrence or metastasis or survival-related outcomes, such as bone, liver, lung, brain, distant metastasis, metabolic response, recurrence/metastasis risk and 5-year survival. Of these, bone metastasis was the most common organ-specific distant metastasis focus, reported in 6 studies (12%), followed by broader distant metastasis or prediction of recurrence/metastasis seen in 7 studies (14%). A small number of studies used comparator or mixed metastatic settings (2/50; 4%): one differentiated breast-cancer bone metastases from non-breast-cancer bone lesions and one classified the primary tumor origin of brain metastases while including breast cancer brain metastases.

Across publication years, the included studies were published between 2021 and 2025: 2 studies (4%) in 2021, 12 studies (24%) in 2022, 10 studies (20%) in 2023, 11 studies (22%) in 2024 and 15 studies (30%) in 2025. Thus, the overall evidence is heavily concentrated in the most recent years, with 37 of 50 studies (74%) published from 2023 onward. By study design, the evidence base was mostly retrospective empirical work (43/50; 86%).

A variety of AI algorithms were used across the included studies. The most frequently used predictive algorithms were support vector machine [SVM] (23 studies, 46%), logistic regression [LR] (21 studies, 42%), random forest [RF] (20 studies, 40%) and gradient-boosting algorithms ([XGBoost, LightGBM, GBM, CatBoost, AdaBoost, gradient boosting] 18 studies, 36%). Decision tree (DT) models were used in 24% of the studies and k-nearest neighbor (KNN) was employed in 22% of the studies. Deep-learning architectures, including CNN-based models, ResNet, DenseNet, VGG, MobileNet, EfficientNet, vision transformer and attention-based multiple instance learning (MIL) frameworks, were used in 13 studies (26%), mainly in imaging and digital pathology-based metastasis prediction.

3.3. Quality Appraisal Results

The quality appraisal process showed that 43 of 50 studies (86%) were rated high quality, while 7 studies (14%) were rated moderate quality. The high-quality studies showed strong methodological features, such as strong preprocessing, feature selection, cross-validation, external/multicenter validation, explainable AI, calibration analysis and clinically relevant prediction performance. The moderate-quality studies remained methodologically useful but were limited by retrospective single-center design, small sample size, manual segmentation, moderate performance, or limited external validation. The appraisal overall demonstrated that the included studies provided sufficiently rigorous and relevant methodological evidence to support synthesis in this review.

3.4. Feature-Engineering Techniques and Feature Types

The evidence retrieved demonstrates a rapidly expanding application of AI, radiomics and ML, as well as multimodal data analytics in breast cancer metastasis analytics and metastasis-related clinical decision support. The findings reveal themes around extraction, preparation, transformation, integration and interpretation of heterogeneous features derived from imaging, clinicopathological, molecular and biomarker datasets to improve metastasis detection, prediction, classification and prognostic stratification in breast cancer. The findings are grouped into two broad themes: feature-processing and engineering techniques used in AI-based breast cancer metastasis analytics and the types of features used in AI-based breast cancer metastasis analytics.

3.4.1. Feature-Processing and Engineering Techniques in AI-Based Breast Cancer Metastasis Analytics

Feature extraction and generation emerged as the most dominant analytical process across the extracted evidence. The evidence revealed that imaging radiomics extraction is a key component of most of the pipelines for metastasis prediction, notably those used in magnetic resonance imaging (MRI), computed tomography (CT), ultrasound imaging, PET-CT and diffusion-weighted imaging modalities. Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) radiomics, contrast-enhanced magnetic resonance imaging (CE-MRI) radiomics, CT radiomics, DWI radiomics, ADC feature extraction and multimodal MRI radiomics were commonly used to characterise tumor heterogeneity, microenvironmental alterations and metastatic behavior. High-dimensional quantitative descriptors were extracted repeatedly from regions of interest (ROIs) for the tumor, vertebral bone marrow regions, lymph nodes and metastatic lesions on radiomics platforms such as PyRadiomics, Darwin Scientific Research Platform, 3D Slicer, ITK-SNAP medical image segmentation software and PyRadiomics libraries. Typical feature-generation processes were the first-order statistical descriptors, shape-based, texture-based and matrix-based radiomics features, such as gray level co-occurrence matrix (GLCM) features, gray level run length matrix (GLRLM) features, gray level dependence matrix (GLDM) features, gray level size zone matrix (GLSZM) features and neighborhood gray tone difference matrix (NGTDM) features. The primary aim of radiomics extraction was to predict ALN metastasis, sentinel lymph node metastasis, distant metastasis, bone metastasis, liver metastasis, lung metastasis and brain metastasis.

The results also underscored the importance of feature preparation and transformation as a key methodological step in the analysis of breast cancer metastasis. Most research applied extensive preprocessing algorithms in order to optimise the reproducibility, standardisation and robustness of the features prior to the subsequent analytical modelling. Commonly used preprocessing methods are normalisation techniques such as z-score normalisation, maximum–minimum normalisation, pixel intensity standardisation, standard normal distribution transformation and unit standardisation. Voxel resampling, image reconstruction standardisation, ROI contour review and segmentation refinement, inter-observer reproducibility analysis (using intraclass correlation coefficients (ICC) or Cohen's kappa statistics) were often included in imaging research. Additional techniques employed also encompass redundancy reduction and correlation filtering, notably Pearson correlation filtering, Spearman correlation analysis, variance thresholding and redundancy mitigation. The widespread use of dimensionality preprocessing techniques such as principal component analysis (PCA), missing-value imputation, outlier detection and harmonisation of imaging parameters was also evident. These preprocessing techniques demonstrate a focus on maintaining the stability of the analytical process, enabling robust and reliable prediction of metastasis, as well as minimising noise in AI-driven metastasis prediction pipelines.

Feature selection and dimensionality reduction further emerged as an increasingly important methodological consideration. The use of feature reduction methods based on statistical and ML approaches to identify metastasis-relevant predictors from high-dimensional data was reported. The most commonly mentioned dimensionality reduction method is the Least Absolute Shrinkage and Selection Operator (LASSO) regression. LASSO has been widely applied with sparse modelling, L1 regularisation, coefficient shrinkage and cross-validation optimisation to identify the most informative radiomics, biomarker and clinicopathological features associated with metastatic progression. Other studies employed SelectKBest, minimum redundancy maximum relevance (mRMR), variance filtering, sample variance F-value filtering, logistic regression-based selection, Cox regression analysis and statistical significance thresholding to reduce dimensionality and keep metastasis-related predictors. The feature selection processes frequently yield compact feature subsets that contain highly discriminative imaging signatures, biomarker combinations, or clinicopathological variables predictive of lymphatic or distant metastatic spread.

Another important finding reveals that feature integration and fusion represent a more conceptually relevant analytical component in metastasis-oriented feature engineering. The results indicate that integrating heterogeneous data sources enhances the predictive performance of metastasis prediction models and, in turn, improves their clinical applicability. In related results, radiomics signatures are often integrated with clinical variables, clinicopathological data, molecular biomarkers, blood-based biomarkers and imaging descriptors in multimodal fusion strategies, such as forming nomograms, comprehensive prediction models and multimodal analytical frameworks. Multiple studies reveal integration of combined DCE-MRI and diffusion-weighted imaging (DWI) radiomics features and other studies combined CT radiomics with DL segmentation output, or ultrasound with pathology and blood biomarker profiles. Multiple clinical-radiomic nomograms and multifactorial integration strategies, as well as multimodal feature-fusion frameworks, have been repeatedly reported to improve metastasis prediction performance relative to single-modality approaches. The results indicate that integration of heterogeneous feature domains is an important aspect of contemporary AI-based breast cancer metastasis analytics.

Feature interpretability and explainability are also more prominent methodological techniques in AI-based breast cancer metastasis analytics. The results explicitly show that explainable AI (XAI) techniques improve transparency of prediction models, feature attribution and clinical interpretability of metastasis prediction systems. Shapley Additive Explanations (SHAP) is the most commonly used explainability framework, reporting and depicting SHAP values, beeswarm plots, feature importance and variable contribution analyses to identify metastasis-associated predictors. Other interpretability tools included correlation heatmaps, feature coefficient analysis, decision curve analysis, odds ratio interpretation, confusion matrices, and radiomics score comparisons between metastasis-positive and metastasis-negative groups. By and large, the findings highlight the need for interpretable AI systems to translate complex metastasis prediction analytics into clinically actionable decision support tools in breast oncology practice.

3.4.2. Feature Types Used in AI-Based Breast Cancer Metastasis Analytics

The second grand theme reveals various types of features utilised in AI-based breast cancer metastasis analytics. The most prevalent data type comprises imaging-derived features. Extensive usage of MRI-derived radiomics, CT-derived radiomics, ultrasound imaging descriptors, PET-CT imaging variables, DWI features, ADC maps and vertebral bone marrow imaging signatures is evident. The imaging-based features that are mainly captured include tumor morphology, texture, lesion intensity distributions, volumetric features and microenvironmental changes that can occur during metastatic dissemination.

Texture-based radiomics matrices such as GLCM, GLRLM, GLDM, GLSZM and NGTDM are common matrices reflected in analytics focused on tumor heterogeneity and metastasis risk stratification.

Clinical and clinicopathological feature types also appeared extensively. The most commonly used variables are tumor size, histological subtype, tumor–node–metastasis (TNM) stage, lymph node status, menopausal status, age, molecular subtype, Breast Imaging Reporting and Data System (BI-RADS) classification, treatment variables and surgical status. Such clinicopathological characteristics are often integrated with imaging features to improve the accuracy of predictions of lymph node metastasis, distant metastasis and survival. Despite the growing prominence of radiomic and AI-based analysis methods, clinicopathological variables remain foundational predictors within multimodal metastasis prediction systems.

Another important feature type is blood-based biomarkers, especially those used to study distant metastasis and predict bone metastasis. Many of the biomarkers that are commonly reported are carcinoembryonic antigen (CEA), cancer antigen 15-3 (CA153), cancer antigen 125 (CA125), cancer antigen 19-9 (CA199), alpha-fetoprotein (AFP), creatine kinase-MB isoenzyme (CK-MB), alpha-hydroxybutyrate dehydrogenase (α-HBDH), apolipoprotein B, albumin–globulin ratio, alkaline phosphatase, lipoprotein-a and gamma-glutamyl transferase. These biomarkers are often integrated with clinical or image-derived markers to further enhance predictive discrimination between metastasis-positive and non-metastatic patient groups. The integration of systemic biochemical indicators therefore reflects growing interest in combining tumor imaging phenotypes with circulating physiological signatures associated with metastatic progression.

Molecular and omics-derived feature types were also commonly represented. Estrogen receptor/progesterone receptor (ER/PR), human epidermal growth factor receptor 2 (HER2), Ki-67 proliferation index (Ki-67), molecular subtypes, luminal subtypes, HER2-overexpression profiles and triple-negative breast cancer markers are repeatedly reported as metastasis-associated markers. These molecular descriptors are often integrated into predictive models and linked to metastatic risk stratification, tumor aggressiveness and survival outcomes. The results indicate that molecular subtyping also remains an integral part of AI-driven metastasis analytics in breast cancer.

In contrast, digital pathology and pathomics feature types appear less frequently in breast cancer metastasis analytics work. Only isolated instances included biopsy pathology descriptors, histopathological examination variables, or pathology-derived features in the metastasis prediction systems. The results show that digital pathology integration remains comparatively underdeveloped relative to imaging radiomics and clinical-radiomic analytics in current breast cancer AI-based metastasis modelling.

Finally, the evidence consistently showed feature types associated with metastatic sites and outcomes. Multiple metastatic endpoints emerging from the results include metastasis to ALNs, metastasis to sentinel lymph nodes, bone metastasis, liver metastasis, lung metastasis, brain metastasis, distant metastasis, metastatic recurrence and metastasis-related survival outcomes. Various studies specifically differentiated metastasis-positive and metastasis-negative patient groups, while others focused on predicting future metastatic occurrence within predefined follow-up periods. By and large, the current evidence highlights the progressive maturation of AI-driven breast cancer metastasis analytics, transforming it into multimodal, interpretable, and comprehensive predictive systems that can aid metastasis detection and diagnostics, prognostic stratification, and personalised oncology decision-making across diverse metastatic contexts.

3.5. Evidence-Informed Conceptual 5Fs Framework for Feature Engineering in AI-Driven Breast Cancer Metastasis Analytics

The evidence extracted in this study provided a strong basis for developing a layered, feature-centric framework (Figure 2), referred to as the 5Fs Framework for the feature-engineering pipeline in AI-driven breast cancer metastasis analytics. It is worth noting that the acronym 5Fs is used to refer to five related feature-engineering steps: feature extraction or generation (FE/G), feature preparation and transformation (FPT), feature selection and dimensionality reduction (FSD), feature construction and integration (FCI) and feature interpretability and explainability (FIE). These stages collectively illustrate the 5Fs approach for feature engineering the breast cancer metastasis data pipeline from raw data to clinically useful, interpretable and metastasis-relevant predictive evidence.

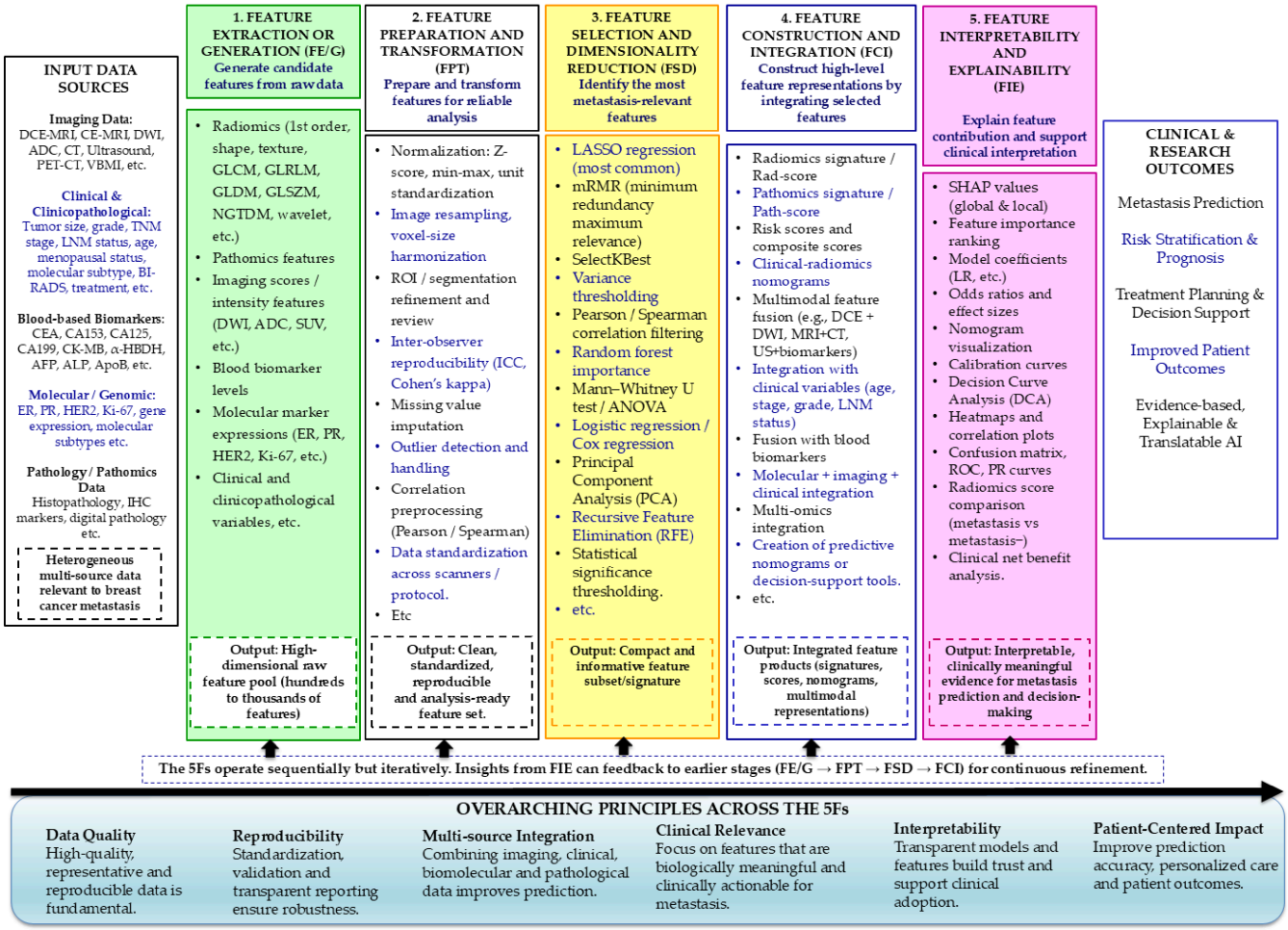


Figure 2. The evidence-informed conceptual 5Fs Framework for feature engineering pipeline in AI-driven breast cancer metastasis analytics (The 5Fs framework (FE/G–FPT–FSD–FCI–FIE) provides a systematic, evidence-based roadmap for engineering high-quality, interpretable and clinically relevant features for AI-driven breast cancer metastasis analytics).

The framework outlines the five stages of evidence-based metastasis prediction, which are sequential but not mutually exclusive: FE/G, where metastasis-relevant data are extracted or generated; FSD—high-dimensional features and transformed into features suitable for reliable analysis; FSD-high-dimensional features are reduced to the most informative metastasis predictors; FCI—features are assembled into an integrated predictive signature; and FIE—feature contributions are explained for clinical interpretation. This framework elucidates a model-based approach that is fundamentally feature-driven, inte-

grating imaging, clinical, biomarker and pathology-derived evidence to enhance metastasis prediction, diagnostic accuracy and AI-driven decision making.

Layer 1: Feature Extraction or Generation (FE/G)

The base level of the framework is FE/G. It illustrates how to turn raw imaging, clinical, pathological, blood-based, molecular and omics data into measurable candidate predictors of metastasis modelling. This layer captures tumor characteristics, biological behavior, patient clinical status and metastatic risk patterns through radiomics, pathomics, biomarker extraction and molecular profiling.

Layer 2: Feature Preparation and Transformation (FPT)

This layer preprocesses extracted features to make them amenable for reliable analysis by improving quality, standardisation, reproducibility and comparability. This layer converts raw extracted features into analysis-ready variables through normalisation, resampling, segmentation refinement, missing-value handling, outlier management, reproducibility checks and correlation preprocessing, among other techniques.

Layer 3: Feature Selection and Dimensionality Reduction (FSD)

This layer detects the most metastasis-relevant features from the high-dimensional feature space. At this stage, FSD converts high-dimensional feature pools into low-dimensional feature signatures that represent features associated with metastases. It filters out the noise of features to obtain short, metastasis-relevant feature signatures for downstream integration and interpretation, while keeping only features which are most relevant to metastatic outcomes.

Layer 4: Feature Construction and Integration (FCI)

This layer synthesises the selected features together to create higher-order predictive signatures, scores, nomograms and multimodal feature combinations. It presents the process of how refined predictors are combined into clinically relevant feature products that integrate imaging, clinical, biomarker, molecular, pathological and omics evidence for breast cancer metastasis analytics. The role of FCI in the framework is to convert the selected predictors into integrated representations of the metastasis risk that can be used for diagnosis, prediction, stratification or prognosis.

Layer 5: Feature Interpretability and Explainability (FIE)

This is the final translational layer of the framework. It describes how engineered features can aid breast cancer metastasis prediction and how their results can inform clinical decisions. These include various feature-importance rankings, SHAP values, coefficient interpretation, odds ratios, radiomics score interpretation, nomogram visualisation, calibration assessment, decision curve analysis, heatmaps and clinical net benefit evaluation. Thus, the role of FIE is to translate engineered features into clinically meaningful evidence.

4. Discussion

The results revealed that AI-powered breast cancer metastasis analytics evolve from stand-alone prediction models to feature engineering ecosystems that effectively integrate imaging, clinical, pathological, biomarker and molecular evidence to build clinically interpretable metastasis prediction systems. The dominance of radiomics-based feature extraction reflects the growing recognition that tumor heterogeneity contains clinically meaningful signatures associated with metastatic progression. Evidence from previous studies also demonstrated that radiomics derived from MRI, PET-CT and ultrasound can capture intratumoral heterogeneity and microenvironmental variations linked to metastatic dissemination and survival outcomes in breast cancer [75,76]. The extensive use of texture matrices such as GLCM, GLRLM, GLSZM and GLDM further supports the premise

that spatial heterogeneity within tumor regions represents an important biomarker for metastatic behavior.

The findings also establish that the preparation and transformation of features remain foundational to reliable AI-based metastasis analytics. Normalisation, segmentation refinement, ICC reproducibility testing and correlation filtering are widely adopted, pointing towards a growing awareness that poorly standardised radiomics pipelines are likely to limit reproducibility and external validity. This observation has been echoed in recent radiomics standardisation research, which has pinpointed segmentation, image acquisition and radiomics preprocessing as major barriers to clinical translation [77–79]. The emphasis on preprocessing therefore reflects an emerging methodological shift toward reproducible and clinically robust AI workflows.

LASSO, mRMR and correlation filtering are the most frequently used methods in the feature selection process, which highlights the difficult situation of dealing with high-dimensional metastasis datasets. The results suggest that AI-powered metastasis analytics is not just an algorithmic phenomenon, but a feature-based one as well. Recent evidence similarly suggests that compact and biologically meaningful feature subsets often outperform excessively large feature pools in oncological prediction systems due to improved generalisability and reduced overfitting [80,81]. The finding is consistent with other recommendations for AI governance by the World Health Organization, which highlight the importance of transparency, robustness, explainability and generation of clinically meaningful evidence in the context of AI-supported health systems [13].

A key finding is the growing use of clinical radiomics fusion, nomograms and multimodal predictive frameworks to integrate heterogeneous multimodal data. This indicates that metastatic progression is increasingly conceptualised as a multidimensional biological process requiring simultaneous interpretation of imaging phenotypes, molecular markers and clinicopathological characteristics. Empirical evidence from past studies has also shown that multimodal AI systems integrating both radiomics and molecular and clinical data achieve better performance in metastasis prediction or detection and survival stratification than single-modality models [82,83]. The crafted evidence-based 5Fs Framework for feature-engineering pipeline in AI-driven breast cancer metastasis analytics therefore provides an important conceptual contribution by systematically positioning feature engineering as the central translational mechanism linking raw cancer data to clinically interpretable metastasis prediction.

The increasing relevance of explainable AI approaches, Grad-CAM and SHAP further demonstrates the rising appreciation of the importance of interpretability and clinician trust for clinical use. Recent studies have found that explainable AI can enhance the transparency, clinician trust and clinical acceptance of oncology decision support systems [22,84,85]. The results thus indicate that future breast cancer metastasis analytics is likely to evolve towards an explainable, multimodal and clinically integrated AI ecosystem that will enable personalised oncology decision making, metastasis risk stratification and precision medicine.

5. Conclusions

The application of AI tools for breast cancer metastasis analysis relies on complex feature engineering workflows that convert complex imaging, clinical, molecular, biomarker and pathological data into meaningful clinical evidence in the form of predictions. The results show that radiomics-based imaging features such as MRI, DWI, Apparent Diffusion Coefficient (ADC), ultrasound and multimodal imaging signatures are the most prominent features for the identification of metastatic risk patterns and tumor heterogeneity associated with disease progression. The consistent usage of preprocessing, normalisation,

reproducibility assessment, dimensionality reduction and feature selection methods further highlights the importance of analytical stability and feature robustness in metastasis prediction systems. The integration of heterogeneous feature domains through multimodal fusion frameworks, radiomics signatures and clinical nomograms reflects a growing transition toward precision-oriented oncology analytics capable of supporting individualised metastasis prediction and AI-driven decision-making. The increasing incorporation of explainable AI approaches such as SHAP, Grad-CAM and feature attribution analysis further demonstrates the need for transparency and interpretability in clinically deployable AI systems.

The proposed evidence-informed conceptual 5Fs framework provides a structured representation of how metastasis-relevant data can be systematically extracted, refined, selected, integrated and interpreted within AI-driven analytical environments. The future of AI-enabled metastasis analytics promises to be greatly shaped by advancements in data harmonisation, external validation across diverse patient populations, more standardised feature engineering workflows and stronger integration of multimodal biological evidence. These advances may contribute toward more accurate metastasis prediction, earlier detection of disease progression and more personalised therapeutic decision making in breast cancer management.

6. Strengths and Limitations

One of the strengths of this study is its thorough synthesis of recent empirical research on feature-engineering pipelines for breast cancer metastasis analytics across 50 included studies. Imaging, pathology, clinical, genomic, proteomic, metabolomic and multimodal AI studies added richness and depth to evidence synthesis and brought a more holistic picture to the emerging AI applications in breast cancer metastasis prediction. In addition, the review incorporated a multidisciplinary approach, including evidence from radiomics, ML and DL, genomics, pathomics, biomarker analytics and multimodal AI frameworks, which informed the development of the evidence-informed conceptual 5Fs Framework. However, there are limitations to point out. Some studies included in this review applied different imaging modalities, feature-processing workflows, validation methods, outcome definitions and AI algorithms, which makes direct comparison challenging. Comparability and generalisability of findings may be influenced by variability in reporting and limited external validation.

7. Future Directions

Future studies should be directed towards the development of more standardised, externally validated and clinically implementable pipelines for the creation of AI-driven breast cancer metastasis analytics. Research should prioritise standardised radiomics and pathomics extraction procedures, transparent reporting of feature-selection methods, and reproducible validation across different imaging systems, populations and clinical contexts. Future studies should also improve multimodal integration by integrating imaging, clinicopathological, blood-based, molecular, genomic and digital pathology features in the same prediction frameworks. In addition, explainable AI methods should be routinely incorporated to clarify feature contribution, improve clinical trust and support decision making. The proposed 5Fs Framework may provide a conceptual guide for future feature-centric and interpretable AI research; however, its structure, practical utility, and clinical relevance should be evaluated through expert consultation, prospective studies, external validation, and implementation research before clinical application.

Author Contributions: Conceptualisation, W.K. and B.C.; Methodology, W.K. and B.C.; Validation, W.K. and B.C.; Formal Analysis, W.K. and B.C.; Investigation, W.K. and B.C.; Writing—Original Draft, W.K. and B.C.; Writing—Review and Editing, W.K. and B.C.; Visualisation, W.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: During the preparation of this manuscript/study, the author(s) used QuillBot and Grammarly for the purposes of grammar and spelling checks and improving clarity. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AB-MIL/ABMIL	Attention-Based Multiple Instance Learning
ADC	Apparent Diffusion Coefficient
AI	Artificial Intelligence
ALN	Axillary Lymph Node
ALNM	Axillary Lymph Node Metastasis
AUC	Area Under the Curve
AUROC	Area Under the Receiver Operating Characteristic Curve
CNN	Convolutional Neural Network
CT	Computed Tomography
DCE	Dynamic Contrast-Enhanced
DCE-MRI	Dynamic Contrast-Enhanced Magnetic Resonance Imaging
DFNN	Deep Feedforward Neural Network
DL	Deep Learning
DT	Decision Trees
DWI	Diffusion-Weighted Imaging
XAI	Explainable AI
GLCM	Gray Level Co-occurrence Matrix
GLDM	Gray Level Dependence Matrix
GLM	Generalised Linear Model
GLRLM	Gray Level Run Length Matrix
GLSZM	Gray Level Size Zone Matrix
H&E	Haematoxylin and Eosin
LASSO	Least Absolute Shrinkage and Selection Operator
LGBM	Light Gradient Boosting Machine
LIME	Local Interpretable Model-Agnostic Explanations
ML	Machine Learning
MRI	Magnetic Resonance Imaging
mRMR	Minimum Redundancy Maximum Relevance
NGTDM	Neighbourhood Gray Tone Difference Matrix
NLP	Natural Language Processing
NPV	Negative Predictive Value
PCA	Principal Component Analysis
PET	Positron Emission Tomography
PET-CT/PET/CT	Positron Emission Tomography-Computed Tomography
PET/MRI	Positron Emission Tomography-Magnetic Resonance Imaging
RF	Random Forest

RFE	Recursive Feature Elimination
ROI	Region of Interest
RSD	Relative Standard Deviation
SHAP	Shapley Additive Explanations
SVM	Support Vector Machine
WSI	Whole-Slide Image
XGBoost	Extreme Gradient Boosting

Appendix A

Appendix A.1

Table A1. PRISMA abstract checklist.

Section and Topic	Item #	Checklist Item	Reported (Yes/No)
TITLE			
Title	1	Identify the report as a systematic review.	Yes
BACKGROUND			
Objectives	2	Provide an explicit statement of the main objective(s) or question(s) the review addresses.	Yes
METHODS			
Eligibility criteria	3	Specify the inclusion and exclusion criteria for the review.	Yes
Information sources	4	Specify the information sources (e.g., databases, registers) used to identify studies and the date when each was last searched.	Yes
Risk of bias	5	Specify the methods used to assess risk of bias in the included studies.	Yes
Synthesis of results	6	Specify the methods used to present and synthesise results.	Yes
RESULTS			
Included studies	7	Give the total number of included studies and participants and summarise relevant characteristics of studies.	Yes
Synthesis of results	8	Present results for main outcomes, preferably indicating the number of studies included and participants for each. If meta-analysis was done, report the summary estimate and confidence/credible interval. If comparing groups, indicate the direction of the effect (i.e., which group is favored).	Yes
DISCUSSION			
Limitations of evidence	9	Provide a brief summary of the limitations of the evidence included in the review (e.g., study risk of bias, inconsistency and imprecision).	Yes
Interpretation	10	Provide a general interpretation of the results and important implications.	Yes
OTHER			
Funding	11	Specify the primary source of funding for the review.	No
Registration	12	Provide the register name and registration number.	No

Appendix A.2 The Summarised PRISMA 2020 Checklist

Table A2. PRISMA 2020 checklist. Topic: Feature-Centric AI for Breast Cancer Metastasis Analytics: A Systematic Review and Evidence-Informed Conceptual Framework for Feature Engineering and Multimodal Integration.

Section and Topic	Item #	Checklist Item	Location Where Item Is Reported
TITLE			
Title	1	Identify the report as a systematic review.	Title block, page 1: the article title identifies the report as "A Systematic Review" and specifies the feature-centric breast cancer metastasis analytics focus.
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	Abstract, page 1: reports PRISMA 2020 guidance, databases searched, publication period, 50 included empirical studies, synthesis focus, main findings and the 5Fs framework. Section 2.1, paragraph 1 also points to Appendix A.1 for the PRISMA 2020 for Abstracts Checklist.
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Section 1, paragraphs 1–5: explains breast cancer burden, metastatic dissemination, limitations of conventional staging/imaging, growth of AI/radiomics, fragmented prior reviews and the need for feature-engineering synthesis in breast cancer metastasis analytics.
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Section 1, Research Questions 1–3: states the three review questions on feature-engineering techniques, feature types and the evidence-based feature-centric framework. Section Technical Contributions further specifies the technical contributions, including the 5Fs framework.
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Section 2.3, paragraph 1: states the inclusion and exclusion criteria. Section 2.6, paragraph 1 and Section 3.4, paragraph 1 explain how studies were grouped for thematic synthesis by feature-processing and engineering techniques and feature types, with Table 1 presenting the grouped synthesis.
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Section 2.2, paragraph 1: identifies MEDLINE, Scopus, Web of Science, Embase and IEEE Xplore as the electronic databases searched between 1 May 2026 and 15 May 2026 and also reports manual screening of reference lists from eligible articles and relevant review studies.
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Section 2.2, paragraph 1: reports the main keywords, Boolean operators and an example search string.
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	Section 2.4, paragraph 1: describes export to a reference-management system, electronic and manual duplicate removal, title/abstract screening, full-text eligibility assessment, independent screening and consensus resolution of disagreements.
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Section 2.6, paragraph 1: states that a structured extraction table aligned with the research questions was used to collect author details, publication year, study aim, methods, feature-engineering techniques, feature types, AI approaches and model-performance results.
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results compatible with each outcome domain in each study were sought and, if not, the methods used to decide which results to collect.	Section 2.3, paragraph 1; Section 3.2, paragraph 1; and Section 3.4.2, paragraphs 1 and 6: metastasis-related outcomes included axillary lymph node metastasis, sentinel lymph node metastasis, high axillary lymph-node burden, distant metastasis, bone/liver/lung/brain metastasis, recurrence/metastasis risk, metabolic response and survival-related endpoints. Section 2.6, paragraph 1 states that model performance indicators were extracted when reported.

Table A2. *Cont.*

Section and Topic	Item #	Checklist Item	Location Where Item Is Reported
Data items	10b	List and define all other variables for which data were sought (e.g., participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Section 2.6, paragraph 1 and Appendix C: extracted variables included author details, publication year, study aim, study methods, feature-engineering techniques, feature types, AI approaches and model-performance results.
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	Section 2.5, paragraphs 1–3; Section 3.3, paragraph 1; and Appendix B: methodological quality was assessed using the Mixed Methods Appraisal Tool (MMAT), operationalised through screening questions and criteria 3.1–3.5, with additional appraisal of processing pipelines, feature selection, validation, explainable AI, calibration and multimodal integration.
Effect measures	12	Specify for each outcome the effect measure(s) used in the synthesis or presentation of results.	Section 2.6, paragraph 1 states that model-performance indicators were extracted.
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis.	Section 2.3, paragraph 1; Section 2.6, paragraph 1; Section 3.4, paragraph 1; and Table 1: studies were eligible for synthesis if they met the empirical AI-based breast cancer metastasis analytics criteria and contributed evidence to the planned synthesis domains of feature-processing and engineering techniques, feature types, multimodal integration, explainability and framework development.
Synthesis methods	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Section 2.6, paragraph 1 and Section 3.4, paragraph 1: extracted evidence was organised into thematic codes, basic themes, organising themes and grand themes.
Synthesis methods	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Section 3.1, paragraph 1 and Figure 1 display the study-selection process; Section 3.4, paragraph 1 and Table 1 display the thematic synthesis of feature-engineering techniques and feature types; Section 3.5, paragraphs 1–2 and Figure 2 display the 5Fs framework; Appendix B presents quality appraisal; and Appendix C presents the extracted study summary table.
Synthesis methods	13d	Describe any methods used to synthesise results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify heterogeneity, and software package(s) used.	Section 2.6, paragraph 1: thematic synthesis was guided by Attride-Stirling's thematic network analysis, with thematic codes refined into basic themes, organising themes and grand themes.
Synthesis methods	13e		
Synthesis methods	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesised results.	Sensitivity analyses were not conducted because no pooled quantitative synthesis was undertaken.
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis arising from reporting biases.	Reporting bias due to missing results was not statistically assessed because the review used heterogeneous narrative/thematic synthesis; note that MMAT quality appraisal and reporting transparency were considered when interpreting the evidence.
Certainty assessment	15	Describe any methods used to assess certainty or confidence in the body of evidence for an outcome.	Section 2.5, paragraphs 1–3; Section 3.3, paragraph 1; Section 6, paragraph 1; and Appendix B support narrative confidence assessment using MMAT-informed methodological quality, validation, reproducibility, reporting completeness and convergence across feature domains.
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Section 3.1, paragraph 1 and Figure 1: database searches identified 1772 records from PubMed/MEDLINE (n = 412), Scopus (n = 536), Web of Science Core Collection (n = 287), Embase (n = 391) and IEEE Xplore (n = 146); 314 duplicates were removed; 1458 records were screened; 1260 records were excluded; and 50 studies were included.

Table A2. *Cont.*

Section and Topic	Item #	Checklist Item	Location Where Item Is Reported
Study selection	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Section 3.1, paragraph 1 and Figure 1: full-text exclusion reasons included no information on metastasis/not AI/ML-driven metastasis analytics, no breast cancer-specific metastasis focus, insufficient/no feature-engineering details, non-empirical designs, review-only publications and non-human/preclinical studies.
Study characteristics	17	Cite each included study and present its characteristics.	Section 3.2, paragraphs 1–3 summarises the 50 included studies by metastasis endpoint, publication year, study design and AI algorithms. Appendix C contains the data extraction summary of included articles, including author/year, aim, methods, feature-engineering techniques, feature types, AI approaches and results.
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Section 2.5, paragraphs 1–3; Section 3.3, paragraph 1; and Appendix B present the study-level methodological quality/risk-of-bias appraisal using MMAT. Section 3.3, paragraph 1 summarises 43 high-quality studies (86%) and 7 moderate-quality studies (14%).
Results of individual studies	19	For all outcomes, present, for each study: summary statistics for each group where appropriate and an effect estimate and its precision, ideally using structured tables or plots.	Section 2.6, paragraph 1 and Appendix C present extracted study-level data, including model-performance indicators where reported. Section 3.2, paragraph 3 summarises AI algorithm distributions, while Section 3.4 and Table 1 synthesise feature-engineering techniques and feature types.
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Section 3.2, paragraphs 1–3; Section 3.3, paragraph 1; Section 3.4, paragraph 1; Table 1; and Appendix B summarise the characteristics and quality appraisal of contributing studies.
Results of syntheses	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present the summary estimate and precision and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	No meta-analysis or pooled statistical synthesis was conducted. Narrative/thematic synthesis results are presented in Section 3.4.1, paragraphs 1–5; Section 3.4.2, paragraphs 1–6; and Section 3.5, paragraphs 1–2.
Results of syntheses	20c	Present results of all investigations of possible causes of heterogeneity among study results.	Section 3.2, paragraphs 1–3; Section 3.4.1, paragraphs 1–5; Section 3.4.2, paragraphs 1–6; and Section 4, paragraphs 1–5 report heterogeneity narratively across metastasis outcome types, feature domains, feature-engineering workflows, imaging modalities, multimodal integration approaches, explainability methods and AI algorithms.
Results of syntheses	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesised results.	No sensitivity analyses were conducted because there was no pooled quantitative synthesis.
Reporting biases	21	Present assessments of risk of bias due to missing results arising from reporting biases for each synthesis assessed.	Reporting bias due to missing results was not statistically assessed because evidence was heterogeneous and synthesised narratively/thematically.
Certainty of evidence	22	Present assessments of certainty or confidence in the body of evidence for each outcome assessed.	Section 3.3, paragraph 1; Section 2.5, paragraphs 1–3; Appendix B; and Section 6, paragraph 1 can be used to report narrative confidence in the evidence.
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Section 4, paragraphs 1–5: interprets the dominance of radiomics extraction, preprocessing, LASSO/mRMR selection, multimodal integration and explainable AI in relation to prior oncology, radiomics, explainability and AI governance evidence.
Discussion	23b	Discuss any limitations of the evidence included in the review.	Section 6, paragraph 1 discusses methodological heterogeneity, variability in reporting and limited external validation across included studies. Section 4, paragraph 2 also discusses reproducibility and standardisation barriers in radiomics pipelines.
Discussion	23c	Discuss any limitations of the review processes used.	Section 2.3, paragraph 1 identifies review-process restrictions to English-language, full-text, peer-reviewed empirical studies from 2020–2026. Section 6, paragraph 1 discusses comparability and generalisability limitations arising from heterogeneity and limited external validation.

Table A2. *Cont.*

Section and Topic	Item #	Checklist Item	Location Where Item Is Reported
Discussion	23d	Discuss implications of the results for practice, policy, and future research.	Section 4, paragraph 5; Section 5, paragraphs 1–2; and Section 7, paragraph 1 discuss implications for standardised feature-engineering pipelines, prospective multicenter validation, radiomics/pathomics harmonisation, multimodal integration, explainable AI, clinical trust and decision-support deployment.
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	This systematic review was not registered.
Registration and protocol	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Section 2.1, paragraph 1 states that a structured review protocol was followed, but no public protocol location is reported.
Registration and protocol	24c	Describe and explain any amendments to information provided at registration or in the protocol.	The review was not registered and no public protocol was prepared.
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Funding statement after Section 7: states that the research did not receive external financial support.
Competing interests	26	Declare any competing interests of review authors.	Conflicts of Interest statement after Section 7: states that the authors declare no conflicts of interest.
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	Data Availability Statement after Section 7: states that no datasets were generated or analysed during the current study. Appendices B and C are identified as locations for quality appraisal and for the extraction of study data.

Appendix B

Table A3. MMAT-based quality appraisal summary for included studies.

Reference	Method Category	MMAT Criteria (Yes [Y]/No [N]/Can't Tell [CT])	Assessment	Overall Notes
[61]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: Y; 3.3: CT; 3.4: CT; 3.5: Y	High Quality	Compared to LASSO, Boruta, mRMR and tree-based classifiers after reducing 54,675 microarray features to 235; the Histogram Difference pathway gave the clearest metastasis signal, with XGBoost leading recall and F1.
[74]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: N; 3.2: CT; 3.3: CT; 3.4: N; 3.5: Y	Moderate Quality	EMR text mining, natural language processing (NLP) preprocessing, k-means outlier handling and nine ML classifiers supported a practical non-invasive blood-profile model; moderate accuracy, reduced post-cleaning sample size and single-center EMR variability temper generalisation.
[70]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: N; 3.2: Y; 3.3: Y; 3.4: Y; 3.5: CT	High Quality	Six ML algorithms, RFE, 10-fold cross-validation and multimodal breast-specific gamma imaging (BSGI)-ultrasound features produced a clinically usable ALN calculator, with SVM providing balanced specificity, negative predictive value (NPV) and improved ROC performance over ultrasound alone.

Table A3. *Cont.*

Reference	Method Category	MMAT Criteria (Yes [Y]/No [N]/Can't Tell [CT])	Assessment	Overall Notes
[53]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: Y; 3.3: Y; 3.4: Y; 3.5: CT	High Quality	Radiogenomic modelling drew strength from the cancer genome atlas (TCGA)/the cancer imaging archive (TCIA) integration, DCE-MRI radiomics, Boruta selection, ICC checks and external hospital validation; the combined model outperformed standalone radiomics or genomics for Axillary Lymph Node Metastasis (ALNM) prediction.
[33]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: N; 3.2: CT; 3.3: Y; 3.4: Y; 3.5: CT	Moderate Quality	DCE-MRI radiomics were carefully denoised, discretised and reduced using Pearson filtering and LASSO, with LR showing strong area under the curves (AUCs); however, single-center design, manual ROI work and limited positive-node representation constrain transferability.
[39]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: N; 3.2: Y; 3.3: Y; 3.4: Y; 3.5: CT	High Quality	Ultrasound radiomics, peritumoral expansion, CNN deep features and Grad-CAM interpretation gave this multimodal framework clear methodological depth, with the CRD_RF_Ext3 model achieving strong independent-test discrimination.
[44]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: N; 3.2: Y; 3.3: Y; 3.4: Y; 3.5: CT	High Quality	DenseNet121 extraction from DCE-MRI and DWI-ADC, mRMR ranking, LASSO selection and synthetic minority over-sampling technique (SMOTE) balancing supported a useful clinicopathological-radiomic nomogram, improving ALNM prediction beyond imaging signatures alone.
[62]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: Y; 3.3: Y; 3.4: N; 3.5: Y	High Quality	Domain adaptation was a major strength: BDA alignment addressed multicenter MRI distribution differences after mRMR-LASSO reduction, and external validation showed better performance than traditional RF modelling.
[45]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: Y; 3.3: Y; 3.4: N; 3.5: Y	High Quality	Whole-slide preprocessing, stain normalisation, Co-Mixup augmentation and attention-based MIL enabled direct metastasis prediction from primary tumor slides, with MEAI outperforming pathologists and supporting AI-assisted diagnostic collaboration.
[59]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: N; 3.2: Y; 3.3: Y; 3.4: Y; 3.5: CT	High Quality	The PET/CT radiomics workflow tested 49 feature-selection-classifier combinations, used SMOTE and independent validation, and identified Lasso-based pipelines as particularly effective for metabolic response prediction.
[46]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: Y; 3.3: CT; 3.4: N; 3.5: Y	High Quality	Multicenter DCE-MRI modelling combined bias correction, resampling, ICC filtering, LASSO radiomics and GBM-selected CNN features; the DL-clinical fusion model showed excellent internal and external AUCs.
[50]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: Y; 3.3: CT; 3.4: N; 3.5: Y	High Quality	Transfer learning, hard-negative mining and local retraining gave this pathology model strong real-world relevance, reducing false positives substantially while testing generalisation across sentinel and axillary-node slides.
[72]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: Y; 3.3: Y; 3.4: N; 3.5: CT	High Quality	Large clinical datasets, grid-search tuning, regularisation and 5-fold validation allowed fair comparison of Deep Feedforward Neural Network (DFNN) and conventional ML, showing XGB/RF advantages in imbalance and SVM strength for long-term prediction.
[66]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: Y; 3.3: Y; 3.4: N; 3.5: Y	High Quality	Proteomics, mutual-information ranking, network propagation, SVM classification, molecular taxonomy of breast cancer international consortium (METABRIC) survival validation and IHC/functional assays made the biomarker work methodologically rich and biologically grounded.

Table A3. *Cont.*

Reference	Method Category	MMAT Criteria (Yes [Y]/No [N]/Can't Tell [CT])	Assessment	Overall Notes
[64]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: CT; 3.3: Y; 3.4: Y; 3.5: Y	High Quality	MRI radiomics, whole-transcriptome sequencing, support vector machine-recursive feature elimination (SVM-RFE), SHAP explanations and international multicohort validation strengthened the radiogenomic nomogram, while immune and drug-response analyses added translational value.
[34]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: N; 3.2: Y; 3.3: Y; 3.4: N; 3.5: Y	High Quality	Deep-learning-assisted DCE-MRI segmentation, LASSO feature reduction and ten-fold validation showed that combining radiomics with clinicopathological features improved distant-metastasis prediction over radiomics alone.
[25]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: N; 3.2: Y; 3.3: Y; 3.4: Y; 3.5: Y	High Quality	Intratumoral and 5 mm peritumoral DCE-MRI radiomics, ICC filtering, LASSO and SVM modelling produced a clinically useful nomogram that outperformed radiologist assessment for ALNM prediction.
[35]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: N; 3.2: Y; 3.3: Y; 3.4: Y; 3.5: Y	High Quality	Peritumoral edema was incorporated thoughtfully with DCE-MRI radiomics and neural-network modelling; decision curve analysis (DCA), net reclassification improvement (NRI) and integrated discrimination improvement (IDI) analyses confirmed that edema information meaningfully improved ALNM-burden prediction.
[69]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: Y; 3.3: Y; 3.4: N; 3.5: CT	High Quality	MissForest imputation, edited nearest neighbor (ENN) balancing, genetic-algorithm selection and cost-sensitive CatBoost created an interpretable metastasis model, with local interpretable model-agnostic explanations (LIME) highlighting clinically plausible treatment and invasion-related contributors.
[47]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: N; 3.2: CT; 3.3: Y; 3.4: N; 3.5: CT	Moderate Quality	Conventional ultrasound and contrast-enhanced ultrasound (CEUS) were fused with EfficientNet and LightGBM, producing strong area under the receiver operating characteristic curve (AUROC) and feature-importance outputs; nonetheless, single-center retrospective data and physician-dependent annotations limit broader certainty.
[48]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: Y; 3.3: Y; 3.4: N; 3.5: Y	High Quality	Large multicenter WSI cohorts, stain normalisation, CTransPath embeddings, ABMIL and clinicopathological RF modelling supported METACANS, whose external validation and explainability strengthened preoperative ALN assessment.
[54]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: N; 3.2: Y; 3.3: Y; 3.4: Y; 3.5: CT	High Quality	Boruta, LASSO, spline modelling, Variance Inflation Factor (VIF) checks and molecular-subtype subgrouping gave the biomarker analysis an accessible but structured ML workflow for identifying inflammatory and nutritional metastasis indicators.
[26]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: N; 3.2: Y; 3.3: Y; 3.4: N; 3.5: CT	Moderate Quality	Prospective PET/MRI acquisition, multimodal lesion segmentation, PyRadiomics filtering, SMOTE and stratified cross-validation produced high-specificity ALN prediction; smaller single-center sampling and absent external validation remain important constraints.
[40]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: N; 3.2: Y; 3.3: Y; 3.4: Y; 3.5: CT	High Quality	Ultrasound radiomics combined with transcription polymerase chain reaction (qRT-PCR) and Beta-1,3-galactosyltransferase 4 (B3GALT4) measurement, mRMR-LASSO selection and nomogram validation were integrated well, producing very high AUCs and interpretable imaging-genomic ALNM prediction.

Table A3. *Cont.*

Reference	Method Category	MMAT Criteria (Yes [Y]/No [N]/Can't Tell [CT])	Assessment	Overall Notes
[65]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: Y; 3.3: Y; 3.4: Y; 3.5: Y	High Quality	Triple-negative breast cancer (TNBC)-specific modelling used multicenter MRI data, LASSO-selected T2WI radiomics and clinical predictors in an interpretable nomogram, supported by calibration and decision-curve analyses.
[27]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: N; 3.2: Y; 3.3: Y; 3.4: Y; 3.5: CT	High Quality	Intratumoral plus 3 mm peritumoral ultrasound radiomics, mRMR-LASSO reduction, XGBoost comparison and SHAP interpretation created a clinically transparent ALNM model with strong validation performance.
[55]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: Y; 3.3: Y; 3.4: Y; 3.5: Y	High Quality	Accessible blood markers were strengthened by ultrasound diameter, LASSO filtering, LightGBM tuning, external validation and feature-importance analysis, producing a practical distant-metastasis prediction approach.
[42]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: N; 3.2: Y; 3.3: Y; 3.4: N; 3.5: Y	High Quality	The DWI-MRI study benchmarked 24 feature selectors and 11 classifiers with repeated 10-fold validation, identifying MRMR-RF as both accurate and stable for non-invasive sentinel lymph node (SLN) prediction.
[36]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: N; 3.2: Y; 3.3: Y; 3.4: Y; 3.5: CT	High Quality	Hybrid MRI modelling combined VGG-16 transfer learning, radiomics, *t*-test/LASSO selection and multiple ML classifiers; the final nomogram performed well and exceeded either signature alone.
[51]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: CT; 3.3: Y; 3.4: Y; 3.5: Y	High Quality	Pathomics features from H&E tiles, LASSO Path-score construction, ML comparison, external validation and DCA supported the ALNET nomogram, improving beyond clinical-only ALNM prediction.
[28]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: N; 3.2: Y; 3.3: Y; 3.4: Y; 3.5: Y	High Quality	Intratumoral/peritumoral ultrasound radiomics, nested cross-validation, bootstrap calibration and a clinical-radiomics nomogram made the high-ALN-burden model interpretable and clinically relevant.
[56]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: Y; 3.3: Y; 3.4: N; 3.5: Y	High Quality	Elastic-net reduction of microarray genes, boosting-model comparison and SHAP explanations produced a highly interpretable genomic metastasis model, highlighting biologically plausible markers such as TSPYL5 and carbonic anhydrase 9 (CA9).
[71]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: N; 3.2: Y; 3.3: Y; 3.4: N; 3.5: Y	High Quality	Multiparametric MRI radiomics, SMOTE balancing, RFE/Kruskal–Wallis selection and SHAP interpretation supported accurate tumor-origin classification for brain metastases, especially when clinical features were fused.
[41]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: CT; 3.3: Y; 3.4: N; 3.5: CT	Moderate Quality	B-mode Ultrasound-Color Doppler Ultrasound (BMUS-CDUS) fusion, MobileNetV2 deep features, handcrafted radiomics, CAM visualisation, and multicenter validation strengthened SLN prediction; moderate cohort size and semi-manual ROI delineation temper generalisability.
[49]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: CT; 3.3: Y; 3.4: Y; 3.5: Y	High Quality	Weakly supervised WSI processing, ResNet18 encoding, attention pooling and multimodal fusion of clinical, sequencing and pathology data gave ICSDA prognostic value, with Kaplan–Meier separation supporting risk stratification.

Table A3. *Cont.*

Reference	Method Category	MMAT Criteria (Yes [Y]/No [N]/Can't Tell [CT])	Assessment	Overall Notes
[67]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: N; 3.2: Y; 3.3: Y; 3.4: N; 3.5: CT	Moderate Quality	Proteomic-metabolomic integration, Orthogonal Partial Least Squares Discriminant Analysis (OPLS-DA), LASSO, Random Forest ranking and ELISA validation generated biologically interpretable site-specific biomarker panels; small retrospective samples keep the findings preliminary despite strong AUCs.
[29]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: CT; 3.3: CT; 3.4: Y; 3.5: Y	High Quality	A very large multicenter MRI radiomics cohort, external plus prospective-retrospective validation, radiologist comparison and tumor-microenvironment analyses made this ALNM model especially robust and clinically persuasive.
[52]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: CT; 3.3: Y; 3.4: CT; 3.5: Y	High Quality	Vocabulary-based MIL, prototype discovery and CNN feature extraction from whole-slide images provided interpretable histology patterns for metastasis prediction, linking visual prototypes to survival-relevant tumor morphology.
[30]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: N; 3.2: Y; 3.3: Y; 3.4: Y; 3.5: CT	High Quality	Contrast-Enhanced Computed Tomography (CECT) lymph-node radiomics used ICC filtering, LASSO and logistic modelling to support nodal metastasis prediction, with imaging texture features adding value beyond conventional CT assessment.
[37]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: N; 3.2: Y; 3.3: Y; 3.4: Y; 3.5: Y	High Quality	The DCE-MRI study compared 2D and 3D radiomics, used PyRadiomics feature extraction and LASSO reduction, and showed that combined clinicoradiomic modelling improved prediction of ALNM.
[31]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: N; 3.2: CT; 3.3: Y; 3.4: Y; 3.5: CT	Moderate Quality	Dual-sequence MRI radiomics, ICC testing, Random Forest modelling and SHAP interpretation were used to predict ALNM; single-center design and moderate sample size moderate confidence despite transparent modelling.
[57]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: Y; 3.3: Y; 3.4: N; 3.5: CT	High Quality	Clinical blood biomarkers were filtered through z-score normalisation, Spearman correlation, LASSO and LightGBM, yielding a practical metastasis model based on routinely available laboratory measures.
[58]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: Y; 3.3: CT; 3.4: N; 3.5: Y	High Quality	Blood biomarkers, ultrasound and biopsy pathology were fused using LASSO and LightGBM, with internal and external validation showing that clinically accessible variables can meaningfully support distant-metastasis prediction.
[32]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: Y; 3.3: Y; 3.4: N; 3.5: CT	High Quality	Automated CNN-transformer CT segmentation, radiomics extraction and external validation gave this framework methodological strength, reducing manual segmentation dependence while supporting metastasis-related prediction.
[68]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: CT; 3.2: Y; 3.3: Y; 3.4: Y; 3.5: Y	High Quality	Surveillance, epidemiology and end results (SEER)-scale modelling, regression-based feature selection, 5-fold cross-validation and external Chinese validation gave this clinicopathological ML model strong population-level support and interpretable metastasis predictors.
[43]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: CT; 3.3: Y; 3.4: Y; 3.5: CT	High Quality	Thoracic vertebral marrow radiomics, multiple ML algorithms and SHAP interpretation offered a novel CT-based route for bone-metastasis prediction, linking marrow texture patterns with clinically relevant risk stratification.

Table A3. *Cont.*

Reference	Method Category	MMAT Criteria (Yes [Y]/No [N]/Can't Tell [CT])	Assessment	Overall Notes
[38]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: N; 3.2: Y; 3.3: Y; 3.4: Y; 3.5: CT	High Quality	CE-MRI radiomics used manual 3D segmentation, reproducibility filtering and model validation to generate a focused ALNM prediction tool, with imaging features complementing clinical nodal assessment.
[60]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: N; 3.2: Y; 3.3: Y; 3.4: Y; 3.5: CT	High Quality	Multimodal clinical-imaging modelling incorporated MRI preprocessing, LASSO reduction, classifier comparison and validation analyses, showing that combined imaging and clinicopathologic information improved metastasis prediction.
[63]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: Y; 3.3: Y; 3.4: N; 3.5: CT	High Quality	Triphasic contrast-enhanced CT radiomics, manual liver-metastasis ROI delineation and clinical integration supported prediction of treatment response, with portal-venous and delayed-phase features contributing useful texture information.
[73]	Quantitative Non-Randomised (Criteria 3.1–3.5)	S1: Y; S2: Y; 3.1: Y; 3.2: Y; 3.3: Y; 3.4: N; 3.5: CT	High Quality	Logistic regression-based recursive feature elimination (LR-RFE) reduced high-dimensional gene-expression data into interpretable markers, and multiple ML comparisons supported a compact metastasis-prediction model grounded in genomic feature selection.

Appendix C

Table A4. The data extraction summary table.

Reference	Aim of the Study	Research Type	Study Methods	Countries/Setting/Location	Participant Details	AI System/Algorithm/Application Involved	Feature-Processing System or Technique Used	Results	Strengths (+) & Limitations (−)
[61]	To predict breast cancer metastasis using ML algorithms and compare effectiveness of feature-selection methods for high-dimensional microarray data	Empirical quantitative computational study using retrospective secondary microarray dataset	Microarray breast cancer datasets analysed using statistical preprocessing, feature selection, and tree-based ML classification models. Compared LASSO, Boruta, and MRMR feature-selection methods with Decision Tree, Extremely Randomized Trees, and extreme gradient boosting (XGBoost) classifiers. Performance evaluated using precision, recall, and F1-score.	Public genomic datasets from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) Databank (GSE102484 and GSE20685); multicenter/international secondary dataset.	202 breast cancer patients total: 101 metastasis and 101 non-metastasis patients. Initial dataset contained 54,675 genomic features reduced to 235 features after preprocessing.	Machine learning (ML); decision tree, extremely randomized trees (Extra-Trees), gradient boosting tree (XGBoost).	Histogram Difference Method; LASSO (Least Absolute Shrinkage and Selection Operator); Boruta; MRMR (Maximum Relevance Minimum Redundancy); variable importance ranking using tree-based models; dimensionality reduction from 54,675 to 235 features.	Preprocessing reduced 54,675 features to 235. Extra-Trees achieved highest precision (0.7934), while XGBoost achieved highest recall (0.8140) and F1-score (0.7423). Histogram Difference Method showed best overall performance, and top 10 important arrays were identified.	(+) Compared multiple feature-selection methods for metastasis prediction; clear dimensionality-reduction and preprocessing pipeline. (–) Small sample size (n = 202); extensive preprocessing reduced usable data; missing values limited reliable imputation.
[74]	To develop a non-invasive ML-based system for classification and diagnostic prediction of breast cancer metastasis using blood profile and clinical text data.	Empirical quantitative retrospective computational study using EMRs (electronic medical records) and blood profile data.	Clinical text data from EMRs were processed using text mining and NLP techniques to identify metastatic breast cancer cases. Blood profile data were analysed using nine ML classification models including Logistic Regression, KNN, Decision Tree, Random Forest, SVM, Gradient Boosting, and XGBoost. Model performance was evaluated using five-fold cross-validation, accuracy, recall, precision, F1-score, AUC, and receiver operating characteristic (ROC) analysis.	Basavatarakam Indo-American Cancer Hospital and Research Institute (BIACH & RI), Hyderabad, Telangana, India; additional normal blood-profile data from hospitals/health check-up records.	Initial dataset included 25,652 EMR entries, reduced to 5176 breast cancer-related records after tokenisation. Final ML dataset included 1449 blood-profile records (1073 metastatic breast cancer patients and 376 normal subjects), reduced to 676 entries after outlier removal. Additional chemotherapy-response analysis included 40 patients.	Machine learning (ML); logistic regression; k-nearest neighbor (KNN); decision tree (DT); random forest (RF); support vector machine (SVM-linear, radial, polynomial); gradient boosting (GB); XGBoost; Flask-based web application deployment.	Text mining and NLP using tokenisation, split/enumerate functions, NLTK libraries, lemmatisation, word-cloud visualisation, k-means clustering for outlier removal, PCA dimensionality reduction, train-test split, ROC/AUC evaluation, and statistical analysis using Welch unpaired t-test.	Decision Tree classifier achieved the best performance with 83% accuracy and AUC of 0.87 after outlier removal. ML performance improved substantially after k-means-based outlier reduction. Monocyte counts showed significant differentiation between metastatic breast cancer and healthy subjects. A Flask-based web application was successfully deployed for metastasis prediction.	(+) Integrated EMR text mining with blood-profile ML analysis; compared multiple ML models; included web-based clinical deployment; non-invasive prediction approach. (–) Moderate prediction accuracy; small final dataset after preprocessing/outlier removal; retrospective single-center study; potential variability in EMR quality and chemotherapy status.
[70]	To develop a non-invasive ML model combining breast-specific gamma imaging (BSGI) and ultrasonographic parameters to predict axillary lymph node (ALN) metastasis in breast cancer patients.	Empirical retrospective diagnostic ML modelling study.	Retrospective imaging analysis using ultrasonographic and BSGI features. Six ML algorithms were trained with recursive feature elimination and 10-fold cross-validation. Model performance was evaluated using ROC, calibration analysis, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy. A web-based prediction calculator was developed.	Zhongshan Hospital, Fudan University, Shanghai, China.	334 breast cancer patients (2012–2021). Training set: 235 patients with 1104 ultrasound and 940 BSGI images; test set: 99 patients with 568 ultrasound and 396 BSGI images.	Machine learning (ML); support vector machine (SVM), random forest (RF), generalised linear model (GLM), neural network (NNET), gradient boosting machine (GBM), extreme gradient boosting (XGB); online prediction calculator.	Recursive feature elimination (RFE); multimodal imaging feature integration; BSGI and ultrasonographic feature extraction; restricted cubic splines (RCS); correlation matrix collinearity analysis; categorical transformation; 10-fold cross-validation. Final selected features included tumor and lymph-node dimensions, lymphatic echogenicity, lymphatic color doppler flow imaging (CDFI) grade, and BSGI axillary mass status.	The SVM model achieved the best predictive performance with AUC = 0.794, sensitivity = 0.641, specificity = 0.800, PPV = 0.676, NPV = 0.774, and accuracy = 0.737. Incorporating BSGI features improved ROC performance compared with ultrasonography alone. Final model included six ultrasonographic parameters and one BSGI feature.	(+) Integrated multimodal imaging features; non-invasive ALN metastasis prediction; compared multiple ML models; included interpretable imaging biomarkers and web-based calculator. (–) Single-center retrospective study; no external validation; moderate sensitivity; BSGI limited by cost and radiation exposure; possible selection bias.
[53]	To develop and validate a radiogenomics model integrating MRI radiomics and transcriptomic features for predicting axillary lymph node metastasis (ALNM) in breast cancer.	Empirical retrospective multicohort radiogenomics study.	Transcriptomic data from the cancer genome atlas (TCGA) and matched DCE-MRI data from the cancer imaging archive (TCIA) were integrated with external hospital validation data. Differential gene-expression analysis, MRI radiomics extraction, feature selection, dimensionality reduction, and logistic regression modelling was performed. Model performance was evaluated using ROC, calibration, decision-curve, and DeLong analyses.	TCGA/TCIA public datasets and Chongqing University Cancer Hospital, China.	126 breast cancer patients' total. Training/testing cohorts included 111 patients from TCGA/TCIA; external validation cohort included 15 patients from one hospital.	Radiogenomics AI framework; logistic regression machine learning (ML); dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) radiomics; transcriptomics integration.	Radiomics extraction of 1651 DCE-MRI features using Pyradiomics; differential gene-expression analysis (DEGs); Boruta feature selection; Spearman correlation filtering; Z-score normalisation; intraclass correlation coefficient (ICC) reproducibility analysis; stepwise logistic regression; multimodal fusion of radiomics and genomics features; fivefold cross-validation.	The radiogenomics model achieved the best performance with AUCs of 0.84 (training), 0.75 (testing), and 0.82 (external validation). It outperformed standalone radiomics and genomics models and significantly improved prediction compared with the radiomics model in the training group (p = 0.01). Final model incorporated five genomics and three radiomics features.	(+) Integrated MRI radiomics with transcriptomics; included external validation and multicohort design; robust feature-selection and preprocessing pipeline; improved ALNM prediction performance. (–) Small external validation cohort; retrospective design; only logistic regression used; no deep-learning models; only DCE-MRI modality evaluated.

Table A4. *Cont.*

Reference	Aim of the Study	Research Type	Study Methods	Countries/Setting/Location	Participant Details	AI System/Algorithm/Application Involved	Feature-Processing System or Technique Used	Results	Strengths (+) & Limitations (−)
[33]	To predict axillary lymph node metastasis using radiomics features extracted from DCE-MRI in breast cancer.	Retrospective single-center radiomics ML study.	DCE-MRI radiomics features were extracted and analysed using multiple ML algorithms. Feature selection, dimensionality reduction, model construction, ROC evaluation, calibration analysis, and decision-curve analysis were performed to predict axillary metastasis.	First Affiliated Hospital of Wenzhou Medical University, China.	243 patients with pathologically confirmed invasive ductal breast cancer. Training cohort: 160 patients; validation cohort: 83 patients.	Machine Learning (ML); logistic regression (LR), Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM); radiomics framework.	ROI manual segmentation; Gaussian denoising; grayscale discretisation and normalisation; extraction of 833 handcrafted radiomics features; ICC reproducibility testing; Pearson correlation filtering; LASSO dimensionality reduction; texture analysis using GLCM and GLRLM features.	LASSO reduced 833 radiomics features to three final features. Logistic regression showed the best performance with AUCs of 0.97 (training) and 0.85 (validation), outperforming DT, RF, and SVM models. Decision-curve and calibration analyses demonstrated good clinical utility and model agreement.	(+) Comprehensive radiomics preprocessing and feature-selection pipeline; compared four ML algorithms; reproducibility testing with ICC; calibration and decision-curve analysis supported clinical applicability. (–) Single-center retrospective study; no external validation; relatively small positive-node subgroup; manual ROI segmentation may introduce subjectivity.
[39]	To develop a multimodal AI framework integrating clinical, radiomics, and deep-learning features for non-invasive prediction of axillary lymph node (ALN) metastasis in breast cancer.	Retrospective multimodal AI imaging study.	Preoperative ultrasound images were analysed using radiomics extraction, CNN-based deep-feature extraction, multimodal feature fusion, and ML classification. Models were evaluated using ROC, DeLong test, decision-curve analysis, and gradient-weighted class activation mapping (Grad-CAM) explainability.	China: single-center hospital study affiliated with Zhejiang University.	431 breast cancer patients (2017–2022); training set: 301 patients; independent test set: 130 patients. Axillary lymph node dissection (ALND) pathology served as the reference standard.	ML; LR, DT, RF, SVM; radiomics framework.	Manual ROI segmentation; intratumoral and peritumoral region expansion (1–3 mm); radiomics extraction using Pyradiomics; ICC reproducibility analysis; LASSO feature selection; PCA reduction of deep-learning features; multimodal feature fusion; Grad-CAM visualisation; fivefold cross-validation.	The combined RF-based model integrating clinical, radiomics, and ResNet50 deep-learning features from the intratumoral +3 mm peritumoral region (CRD_RF_Ext3) achieved the best performance with test AUC = 0.869, sensitivity = 0.886, specificity = 0.856, PPV = 0.856, and NPV = 0.839. Incorporating peritumoral information improved predictive performance across models.	(+) Strong multimodal fusion framework; explainable AI (XAI) using Grad-CAM; systematic peritumoral-region analysis; large cohort; extensive ML/DL comparison; decision-curve clinical utility analysis. (–) Single-center retrospective design; no external validation; operator-dependent ultrasound segmentation; inter-scanner variability may affect generalisability.
[44]	To develop a deep learning radiomics signature and nomogram using preoperative breast MRI for predicting axillary lymph node metastasis (ALNM) in breast cancer.	Retrospective MRI radiomics and deep-learning study.	Preoperative multiparametric breast MRI was analysed using deep learning radiomics extraction, feature selection, and logistic regression modelling. A clinicopathological-radiomic nomogram was developed and evaluated using ROC, calibration, and decision-curve analyses.	Xinhua Hospital, Shanghai Jiao Tong University School of Medicine, China.	479 breast cancer patients with 488 lesions. Training cohort: 366 lesions (123 ALNM-positive, 243 negative); testing cohort: 122 lesions (44 positive, 78 negative).	DenseNet121 deep-learning radiomics; logistic regression nomogram.	Manual ROI segmentation using insight toolkit-snap medical image segmentation software (ITK-SNAP); deep-learning feature extraction from DCE-MRI and diffusion-weighted imaging/apparent diffusion coefficient (DWI-ADC) sequences using pretrained DenseNet121; mRMR feature ranking; LASSO feature selection; synthetic minority over-sampling technique (SMOTE) oversampling; multiparametric MRI fusion; tenfold cross-validation.	Twenty-three deep-learning radiomics features were selected to construct the radiomics signature. The combined nomogram achieved AUCs of 0.80 (training) and 0.71 (testing). In the testing cohort, sensitivity was 65%, specificity 80%, and accuracy 75%. Combining radiomics with clinicopathological factors improved predictive performance over radiomics alone.	(+) Integrated multiparametric MRI and deep-learning radiomics; robust feature-selection pipeline; relatively large cohort; combined clinicopathological and imaging features; calibration and decision-curve analyses supported clinical utility. (–) Single-center retrospective study; no external validation; moderate testing performance; manual segmentation dependency; limited scanner diversity may affect generalisability.
[62]	To develop a domain-adaptive radiomics pipeline using multiparametric MRI for predicting axillary lymph node metastasis (ALNM) in multicenter breast cancer studies.	Retrospective multicenter MRI radiomics study.	Multiparametric breast MRI radiomics features were extracted and analysed using a domain-adaptive ML pipeline. Feature selection, domain adaptation, classifier training, internal validation, and external validation were performed. Predictive performance was compared with six traditional ML models using ROC and DeLong analysis.	Four hospitals in China.	501 breast lesions from 496 breast cancer patients. Source domain: 396 lesions from 391 patients (training/internal validation); target domain: 105 lesions from 105 patients (external validation).	BDA classifier; SVM, KNN, RF, DT, XGBoost, linear discriminant analysis (LDA).	Semi-automatic ROI segmentation; multiparametric MRI radiomics extraction from t2-weighted imaging (T2WI), non-fat-saturated t1-weighted imaging (T1WI), and DCE-MRI sequences; Variance Threshold filtering; mRMR feature reduction; LASSO sparsity selection; maximum mean discrepancy (MMD)-based domain alignment; multicenter distribution harmonisation.	A total of 13,184 radiomics features were extracted per lesion and reduced to 12 final features after feature selection. The BDA domain-adaptive pipeline achieved external validation AUC = 0.781 (95% CI: 0.692–0.870), outperforming the best traditional RF model (AUC = 0.700; $p = 0.048$).	(+) Included multicenter external validation; innovative domain-adaptation framework; robust multistage feature-selection pipeline; addressed cross-center MRI distribution differences. (–) Retrospective design; semi-automatic segmentation required manual refinement; moderate sample size; no prospective validation; generalisability requires further testing.

Table A4. *Cont.*

Reference	Aim of the Study	Research Type	Study Methods	Countries/Setting/Location	Participant Details	AI System/Algorithm/Application Involved	Feature-Processing System or Technique Used	Results	Strengths (+) & Limitations (−)
[45]	To develop an AI platform for predicting lymph node and distant metastases directly from primary breast cancer pathology whole-slide images (WSIs).	Retrospective digital pathology deep-learning study.	H&E-stained WSIs from primary breast cancer were preprocessed and analysed using a deep-learning multiple-instance learning framework. CNN feature extraction, augmentation-bag generation, and attention-based multiple instance learning (MIL) were used for metastasis prediction. Model performance was compared with six board-certified pathologists and mixed AI–pathologist models.	The cancer genome atlas (TCGA) public dataset and Affiliated Hospital of Jiangnan University, China.	832 WSIs total: 520 non-metastatic and 312 metastatic breast cancer cases. Metastatic cohort included lymph node, bone, lung, liver, and multiple-organ metastases. Independent test cohort included 187 patients.	MEAI platform; metastasis evaluation convolutional neural network (MECNN) deep-learning framework; ResNet50 backbone CNN; attention-based multiple-instance learning (AB-MIL).	OTSU tissue localisation; non-overlapping patch extraction (224 × 224 pixels); structure-preserving color normalisation (SPCN) stain normalisation; Co-Mixup augmentation; random rotation/flip and stain perturbation; CNN feature extraction; augmentation-bag generation; AB-MIL aggregation.	The MEAI system achieved test AUROC = 0.934 (95% CI: 0.922–0.945), outperforming six pathologists (mean AUROC = 0.811). ResNet50 showed the best performance among evaluated CNN backbones. AI-assisted mixed diagnostic models improved pathologist accuracy, sensitivity, specificity, PPV, and NPV.	(+) Strong predictive performance; direct metastasis prediction from primary-tumor WSIs; comparison with pathologists; explainable AI–human collaborative evaluation; multicenter/public dataset integration. (–) Retrospective design; limited prospective and external validation; computationally intensive WSI processing; image heterogeneity across datasets; single-image representation per lesion may limit real-world generalisability.
[59]	To identify optimal combinations of feature-selection and ML classifiers for predicting metabolic response of metastatic breast cancer lesions using pretreatment [18F]F-FDG PET/CT radiomics and clinical variables.	Retrospective PET/CT radiomics ML study.	Pretreatment and follow-up [18F] F-FDG PET/CT scans were analysed using radiomics extraction, feature selection, and ML classification. Forty-nine cross-combinations of seven feature-selection methods and seven ML classifiers were evaluated using 10-fold cross-validation and independent test validation with ROC analysis.	Vienna General Hospital, Austria.	48 patients with pathologically confirmed metastatic breast cancer (2010–2015); 228 metastatic lesions analysed (127 responders, 101 nonresponders). Training cohort: 182 lesions; testing cohort: 46 lesions.	SVM, RF, neural network (NN), LR, Gaussian naïve Bayes (GNB), KNN and AdaBoost	PET/CT radiomics extraction (202 features per lesion); semiautomatic volume of interest (VOI) segmentation; z-score normalisation; voxel resampling to $1 \times 1 \times 1$ mm^3; discretisation (64-bin); Spearman correlation filtering; SMOTE oversampling; feature-selection methods including Lasso, Mutual Information (MI), PCA, ICA, Wilcoxon test, analysis of variance (ANOVA) F-test, and clustering.	Lasso + SVM and Lasso + RF achieved the highest cross-validation AUCs (0.93 ± 0.06 and 0.92 ± 0.03). In the validation cohort, LASSO + NN achieved the best AUC (0.90), while MI + RF achieved the highest accuracy (0.85). On average, Lasso and SVM methods showed the best overall predictive performance.	(+) Comprehensive comparison of 49 ML pipelines; lesion-level response prediction; integration of PET/CT radiomics with clinical variables; robust feature-selection and validation workflow. (–) Small retrospective single-center cohort; heterogeneous treatments and patient characteristics; no external validation; lesion-based rather than patient-level split may limit generalisability.
[46]	To develop and validate a DCE-MRI-based deep-learning model for preoperative prediction of axillary lymph node metastasis (ALNM) in breast invasive ductal carcinoma (IDC).	Retrospective multicenter AI imaging study.	Preoperative DCE-MRI images were analysed using radiomics and CNN-based deep-learning feature extraction. Five predictive models (Clinical, Radiomics, Radiomics-Clinical, DL, DL-Clinical) were developed and evaluated using internal and external validation cohorts with ROC analysis.	Two hospitals in Qinhuangdao, China.	520 patients with histologically confirmed breast IDC. Training set: 287 patients; internal validation set: 124 patients; external testing set: 109 patients.	Convolutional neural network (CNN); logistic regression; LASSO radiomics; gradient boosting machine (GBM)-based feature selection; DL-CLINICAL fusion model.	N4 bias-field correction; B-spline interpolation voxel resampling ($1 \times 1 \times 1$ mm^3); histogram standardisation; semi-automatic ROI segmentation using ITK-SNAP; Z-score normalisation; ICC reproducibility filtering; Pearson correlation filtering; LASSO feature selection; GBM selection of top DL features; 5-fold cross-validation.	The DL-Clinical model achieved the best predictive performance with AUCs of 0.935 (training), 0.946 (internal validation), and 0.951 (external testing), outperforming Clinical, Radiomics, Radiomics-Clinical, and DL-only models. The DL model alone also outperformed traditional radiomics models.	(+) Multicenter external validation; strong DL-clinical integration; comprehensive preprocessing and feature-selection pipeline; good external generalisability; non-invasive ALNM prediction. (–) Retrospective design; included only IDC and unifocal lesions; multifocal/bilateral lesions excluded; reliance on semi-automatic segmentation; only DCE-MRI analysed without multimodal imaging integration.
[50]	To evaluate the generalisation performance of a pretrained deep-learning model for breast cancer lymph-node metastasis detection across new clinical settings and surgical indications.	Empirical retrospective multicenter computational pathology study.	Deep-learning models were trained, retrained, and evaluated on digital whole-slide images (WSIs) from multiple cohorts. Transfer learning, slide-level and lesion-level evaluation, receiver operating characteristic-area under the curve (ROC-AUC) and free-response receiver operating characteristic (FROC) analyses, heatmap generation, and qualitative pathology review were performed.	Sweden and the Netherlands; CAMELYON16/17 datasets and AIDA BRLN clinical datasets.	Training/evaluation datasets included 961 WSIs from sentinel and axillary lymph nodes, plus 259 additional negative axillary-node slides. Datasets contained breast cancer lymph-node metastases from multiple clinical indications.	Modified DenseNet convolutional neural network for digital pathology metastasis detection.	Whole-slide preprocessing; tissue segmentation; patch extraction; stain normalisation; augmentation; hard-negative mining; transfer learning/retraining; fully convolutional inference; heatmap generation; lesion-threshold detection; sampling optimisation strategies.	The pretrained base model showed reduced performance on unseen local clinical data, particularly for axillary-node specimens. Retraining on local data substantially improved performance and reduced false-positive detections from 9088 to 327 on axillary-node slides. Retrained models demonstrated improved ROC-AUC and FROC performance across datasets.	(+) Strong real-world generalisation assessment; multicenter external datasets; clinically relevant pathology evaluation; extensive transfer-learning and preprocessing pipeline. (–) Limited invasive lobular carcinoma representation; false positives remained despite retraining; isolated tumor cells (ITCs) excluded; retrospective design without prospective clinical deployment.

Table A4. *Cont.*

Reference	Aim of the Study	Research Type	Study Methods	Countries/Setting/Location	Participant Details	AI System/Algorithm/Application Involved	Feature-Processing System or Technique Used	Results	Strengths (+) & Limitations (−)
[72]	To compare deep learning and conventional ML methods for predicting later breast cancer metastasis using non-image clinical data.	Retrospective empirical machine-learning study.	Clinical metastasis datasets were analysed using deep feedforward neural networks (DFNNs) and nine conventional ML algorithms. Hyperparameter tuning with grid search, 5-fold cross-validation, ROC-AUC evaluation, and Wilcoxon statistical testing were performed for 5-, 10-, and 15-year metastasis prediction.	USA; Lynn Sage Dataset for Metastasis (LSM), Northwestern Memorial Hospital.	Three clinical datasets: LSM-5 year (4189 patients; 437 metastatic cases), Lynn Sage dataset for metastasis (LSM)-10 year (1827 patients; 572 metastatic cases), and LSM-15 year (751 patients; 608 metastatic cases). Each dataset included 31 clinical predictor variables.	Deep feedforward neural network (DFNN), XGBoost (XGB), RF, SVM, LR, LASSO, KNN, AdaBoost, DT, Naïve Bayes (NB).	Nearest-neighbor imputation; extensive grid-search hyperparameter optimisation; dropout regularisation; L1/L2 regularisation; stochastic gradient descent (SGD) and AdaGrad optimisation; class-balance analysis; 5-fold cross-validation; Wilcoxon rank-sum statistical testing.	DFNN achieved mean test AUCs of 0.769 (5-year), 0.793 (10-year), and 0.842 (15-year), ranking 6th, 4th, and 3rd among models, respectively. XGB and RF performed best on imbalanced datasets, while SVM, LR, LASSO, and DFNN performed better on more balanced long-term prediction datasets. The best overall AUC was 0.845 using SVM for 15-year prediction.	(+) Large clinical dataset; extensive comparison of DL and ML methods; comprehensive grid-search optimisation; detailed overfitting-control strategy; robust statistical comparison framework. (–) Used only non-image clinical data; retrospective design; high computational cost for DFNN grid search; no independent external validation.
[66]	To identify protein biomarkers associated with early and late breast cancer metastasis and validate VWA5A as a metastasis-related biomarker.	Empirical translational proteomics and ML biomarker study.	High-throughput proteomics, ML-based feature selection, SVM classification, molecular taxonomy of breast cancer international consortium (METABRIC) survival validation, immunohistochemistry (IHC) validation, and in-vitro functional assays were performed.	Seoul National University Hospital, South Korea; external validation using METABRIC dataset.	Discovery cohort: 29 breast cancer patients (10 no metastasis, 9 late metastasis, 10 early metastasis). Validation cohort: 1002 consecutive breast cancer patients; 966 evaluable for IHC. METABRIC cohort: 1904 patients.	Machine learning-based protein prioritisation; SVM classifier.	High-throughput formalin-fixed paraffin-embedded (FFPE) proteomics; k-nearest neighbor imputation; differential protein expression analysis; mutual information (MI) ranking; network propagation (NP) scoring; recursive feature addition; leave-one-out cross-validation (LOOCV); t-SNE visualisation; support vector machine-radial basis function (SVM)-RBF classification; IHC H-score validation.	A total of 9455 proteins were quantified, with 6639 retained after filtering and imputation. VWA5A expression progressively decreased from non-metastatic to early metastatic groups. The SVM classifier achieved precision = 0.80, recall = 0.79, and F1-score = 0.79. Low VWA5A expression was associated with shorter disease-free survival in METABRIC (hazard ratio [HR], 0.71; $p = 0.0001$), advanced T stage, lymphatic invasion, and increased migration/invasion in breast cancer cell lines following VWA5A knockdown.	(+) Integrated proteomics, ML-based biomarker prioritisation, survival validation, IHC, and functional assays; identified biologically relevant metastasis-associated biomarker; included large external validation cohort. (–) Small proteomics discovery cohort; biomarker-focused rather than deployable predictive model; limited human epidermal growth factor receptor 2 (HER2+) representation; retrospective validation design; no prospective clinical validation.
[64]	To develop a radiogenomic multimodal ML model integrating MRI, transcriptomics, and clinicopathologic data for preoperative prediction of axillary lymph node metastasis (ALNM) and therapeutic response in breast cancer.	Retrospective international multicohort radiogenomic ML study.	MRI radiomics extraction, whole-transcriptome sequencing (WTS) differential-expression analysis, support vector machine-recursive feature elimination (SVM-RFE) feature selection, SHAP explainability, logistic regression nomogram modelling, GSEA (gene set enrichment analysis), immune-checkpoint analysis, and drug-sensitivity testing were performed.	China and international public datasets (TCGA/TCIA, Duke, Foshan cohorts).	1078 breast cancer patients across five cohorts: TCIA cohort ($n = 103$), TCIA validation cohort ($n = 51$), Duke cohort ($n = 138$), Foshan cohort ($n = 106$), and TCGA cohort ($n = 680$). WTS cohort included 8 patients (4 ALN−, 4 ALN+).	SVM, SVM-recursive feature elimination (SVM-RFE), LR, SHAP explainable AI, GridSearchCV optimisation.	MRI voxel resampling ($1 \times 1 \times 1$ mm); manual 3D ROI segmentation with ITK-SNAP; extraction of 1014 radiomics features (GLCM, GLRLM, GLSZM, NGTDM, first-order features); differential gene-expression analysis using limma; Pearson correlation filtering; PCA visualisation; recursive feature elimination; multimodal radiogenomic fusion; 20-fold cross-validation; SHAP feature attribution.	Four radiomics features and three genomic features (MESP1, RAB40B, ACADS) were retained for model construction. The radiogenomic multimodal nomogram achieved AUCs of 0.82 (95% CI, 0.74–0.91) in the TCIA cohort and 0.77 (95% CI, 0.63–0.91) in the TCIA validation cohort for ALNM prediction. High-risk groups showed distinct metabolic and immune signatures, with higher expression of programmed death-ligand 1 (PD-L1), CTLA4, CD8A, CXCL10, GZMB, and HAVCR2, and differential sensitivity to lapatinib, doxorubicin, and methotrexate.	(+) Strong multimodal integration of MRI, transcriptomics, and clinicopathology; large international multicohort design; explainable AI using SHAP; external validation; incorporated immune and drug-response analyses. (–) Retrospective design; manual ROI delineation; MRI heterogeneity across centers; limited paired radiogenomic sample size; no prognostic survival validation; prospective validation required.
[34]	To develop a DCE-MRI radiomics prediction model for distant breast cancer metastasis using deep-learning-assisted imaging analysis.	Retrospective empirical radiomics/deep-learning study.	DCE-MRI lesion segmentation, radiomics extraction, feature reduction, ensemble classification, and ten-fold cross-validation were performed. Predictive models using imaging features alone and those combining radiomic and clinicopathologic features were compared.	Baoji Central Hospital, China.	288 breast cancer patients with 297 lesions. Metastatic group: 96 patients with 100 lesions; nonmetastatic group: 192 patients with 197 lesions.	Deep matrix factorisation deep-learning model with fully connected architecture; ensemble learning classifier.	Deep-learning-based lesion segmentation; ROI delineation; extraction of morphological, grayscale statistical, texture, and wavelet radiomics features; rank-sum test screening; LASSO feature reduction; dimensionality reduction; ten-fold cross-validation.	Of the 341 extracted radiomics features, 87 were retained after LASSO reduction. The radiomics-only model achieved a sensitivity 66.7%, a specificity 84.2%, an accuracy 78.3%, and an AUC of 0.744. The combined radiomics + clinicopathologic model improved performance with sensitivity 67.7%, specificity 86.8%, accuracy 80.5%, and AUC 0.763. Bone, lung/pleura, and liver were the most common metastatic sites.	(+) Combined radiomics with clinicopathologic data; incorporated deep-learning-assisted MRI segmentation; multifeature radiomics workflow; demonstrated improved prediction with multimodal integration. (–) Small single-center retrospective dataset; moderate predictive performance; no independent external validation; heterogeneous MRI acquisition platforms; limited deep-learning explainability.

Table A4. *Cont.*

Reference	Aim of the Study	Research Type	Study Methods	Countries/Setting/Location	Participant Details	AI System/Algorithm/Application Involved	Feature-Processing System or Technique Used	Results	Strengths (+) & Limitations (−)
[25]	To develop a DCE-MRI intratumoral and peritumoral radiomics nomogram for preoperative prediction of axillary lymph node metastasis (ALNM) in breast cancer.	Retrospective radiomics ML study.	DCE-MRI radiomics extraction, intratumoral/peritumoral VOI segmentation, feature selection, SVM modelling, logistic regression nomogram construction, ROC analysis, calibration analysis, and decision-curve analysis (DCA) were performed.	Affiliated Cancer Hospital of Xinjiang Medical University, China.	312 breast cancer patients (291 invasive ductal, 6 invasive lobular, 15 papillary carcinoma); 190 ALN-positive and 122 ALN-negative cases. Training cohort: 250 patients; validation cohort: 62 patients.	SVM; multivariable logistic regression radiomics nomogram.	Manual intratumoral ROI segmentation using 3D-Slicer; automatic 5 mm peritumoral VOI expansion; Pyradiomics extraction; ICC reproducibility filtering (>0.75); z-score normalisation; Spearman correlation filtering; LASSO feature selection with 10-fold cross-validation; radiomics signature construction.	A total of 3812 radiomics features were extracted, with 15 final ALN-related features retained in the combined model. The intratumoral and peritumoral radiomics model achieved AUCs of 0.867 (training) and 0.785 (validation), respectively. The final radiomics nomogram incorporating radiomics signature, MR-reported ALN status, and lesion diameter achieved AUCs of 0.872 (training) and 0.863 (validation), outperforming radiologist assessment (AUC 0.775).	(+) Integrated intratumoral and tumor-microenvironment radiomics; robust multistep feature-selection workflow; good validation and clinical utility performance; included calibration and DCA analyses; compared against radiologist assessment. (−) Single-center retrospective design; relatively small validation cohort; semiautomatic/manual segmentation dependency; no external multicenter validation; only second postcontrast DCE-MRI phase analysed.
[35]	To determine whether peritumoral edema (PE) improves MRI-based deep-learning radiomic prediction of axillary lymph node metastasis (ALNM) burden in breast cancer.	Retrospective multiparametric MRI deep learning radiomics study.	T2WI-based edema assessment, DCE-MRI radiomics extraction, neural-network modelling, LASSO feature selection, ROC analysis, net reclassification improvement (NRI)/integrated discrimination improvement (IDI) evaluation, calibration analysis, and decision-curve analysis (DCA) were performed.	Sichuan Cancer Hospital & Institute, China.	288 invasive breast cancer patients; 67 with high ALNM burden (≥2 positive lymph nodes). Training cohort: 177 patients; independent validation cohort: 111 patients.	Deep-learning neural network using TensorFlow/Keras; deep-learning radiomic (DLR) model; logistic regression MRI model.	Intra- and peri-tumoral ROI segmentation; extraction of 208 radiomics features (shape, first-order, texture); ICC reproducibility filtering (>0.75); z-score normalisation; LASSO feature selection with 10-fold cross-validation; GLCM texture analysis; PE classification from T2WI; dropout and L2 regularisation in neural-network optimisation.	The deep-learning-combined model integrating PE and radiomics achieved the best performance, with AUCs of 0.953 (training) and 0.963 (validation), outperforming MRI-only and DLR-only models. Addition of PE significantly improved prediction performance (validation NRI = 0.539, IDI = 0.387; $p < 0.001$). Tenfold cross-validation showed mean accuracy of 0.932.	(+) Strong integration of T2WI edema features with DCE-MRI radiomics; clinically relevant ALNM burden stratification; independent validation cohort; comprehensive DCA, NRI, and IDI analyses; robust deep-learning optimisation workflow. (−) Retrospective single-center design; subjective edema assessment despite good interobserver agreement; same MRI scanner used for all cohorts; possible biopsy-related edema bias; no external multicenter validation.
[69]	To develop an explainable AI system for predicting breast cancer metastasis using clinicopathological and treatment data.	Retrospective machine-learning prognostic study.	Data preprocessing, imbalance correction, feature selection, cost-sensitive boosting classification, comparative ML evaluation, and patient-level explainability modelling were performed.	Morocco–Algeria–France collaboration using a Moroccan public breast cancer dataset.	Original dataset: 716 Moroccan women with breast cancer. Final cleaned dataset: 413 patients, including 101 metastatic and 312 nonmetastatic cases; median follow-up 31 months.	Cost-sensitive CatBoost (CS-CatBoost), LIME explainer, XGBoost, LightGBM, LR, Decision Tree (DT), KNN, cost-sensitive extreme gradient boosting (CS-XGBoost), cost-sensitive Light Gradient Boosting Machine (CS-LightGBM).	MissForest imputation; edited nearest neighbor (ENN) undersampling; genetic algorithm (GA) feature selection; cost-sensitive learning with weighted loss function; patient-level LIME explanations; stratified train/test split (70/30).	CS-CatBoost achieved the best performance with precision 76.5%, recall 79.5%, and F1-score 77% for metastasis prediction, outperforming conventional and other boosting models. LIME explanations identified non-use of adjuvant chemotherapy and radiotherapy as the strongest metastasis-associated factors, followed by vascular invasion and high Scarff–Bloom–Richardson (SBR) grade.	(+) Explainable AI framework with patient-level interpretation; addressed class imbalance using cost-sensitive learning; robust preprocessing and feature-selection pipeline; clinically interpretable prognostic-factor analysis. (−) Single retrospective dataset; no external validation; relied only on clinicopathological/treatment data; LIME interface had limited clinical usability; causal interpretability not fully established.
[47]	To develop a bimodal AI system integrating contrast-enhanced ultrasound (CEUS) and imaging features for predicting axillary lymph node (ALN) metastasis in early breast cancer.	Retrospective multimodal AI diagnostic study.	Deep-learning analysis of conventional ultrasound and CEUS images, multimodal feature integration, LightGBM classification, feature-importance extraction, and five-fold cross-validation were performed.	Hokkaido University Hospital, Japan.	788 CEUS images from 788 early breast cancer patients (2013–2021); 190 patients (24%) had ALN metastases. Data were split into training and independent test cohorts (3:1 ratio).	EfficientNet B0/B4/B8, vision transformer (ViT Base/16), LightGBM.	Digital imaging and communications in medicine (DICOM)-to-raster conversion; ROI cropping and resizing; grayscale preprocessing; RGB image fusion of conventional US and CEUS; ImageNet-pretrained DL models; image augmentation (cropping, flipping, noise addition, color adjustment); tabular feature encoding; five-fold cross-validation; LightGBM feature-importance ranking.	EfficientNet-B8 achieved the best DL performance and was used for image-based prediction. The final bimodal model achieved AUROC 0.93 (95% CI 0.88–0.98), sensitivity 0.88, specificity 0.96, and accuracy 0.93, outperforming radiologist diagnosis (accuracy 0.82) and conventional US-only models (AUROC 0.76). The most important predictive features were heterogeneous enhancement, diffuse cortical thickening, and eccentric cortical thickening.	(+) Strong multimodal CEUS + imaging-feature integration; high diagnostic performance; interpretable feature-importance analysis; independent test cohort; clinically meaningful CEUS biomarkers. (−) Retrospective single-center design; no external validation; possible overfitting despite cross-validation; physician-dependent feature annotations; not fully end-to-end automated.

Table A4. *Cont.*

Reference	Aim of the Study	Research Type	Study Methods	Countries/Setting/Location	Participant Details	AI System/Algorithm/Application Involved	Feature-Processing System or Technique Used	Results	Strengths (+) & Limitations (−)
[48]	To develop a multimodal AI model integrating whole-slide images (WSIs) and clinicopathological features for preoperative prediction of axillary lymph node (ALN) metastasis in breast cancer.	Retrospective multicenter multimodal AI diagnostic study.	Deep-learning analysis of biopsy WSIs, machine-learning modelling of clinicopathological variables, weighted ensemble integration, multicenter external validation, ROC analysis, and explainable AI assessment were performed.	Six institutions in South Korea.	Training/internal dataset: 2491 patients from Sinchon Severance Hospital (1991 training; 500 internal validation). External validation: 2166 patients across five cohorts (1090, 486, 246, 197, and 147 patients).	METACANS multimodal AI framework combining pathology deep learning model (PathDL) and clinicopathological machine learning (ClinicML); CTransPath; attention-based multiple instance learning (ABMIL); Random Forest (RF).	Patch extraction (224 × 224 RGB patches); informative patch selection; Macenko stain normalisation; CTransPath embeddings (768-dimensional features); ABMIL attention weighting; weighted ensemble learning; clinicopathological feature integration; dropout regularisation; eight-fold cross-validation.	METACANS achieved an AUC of 0.743 in internal validation and a pooled external AUC of 0.733 (95% CI 0.711–0.755), outperforming PathDL and ClinicML alone. External validation sensitivity was 0.820, NPV 0.846, specificity 0.504, and balanced accuracy 0.662. Explainable AI identified micropapillary growth, infiltrative patterns, and necrosis as pathology patterns associated with ALN metastasis.	(+) Large multicenter external validation cohort; multimodal integration of WSIs and clinicopathological data; explainable AI analysis; robust stain-normalisation and feature-extraction pipeline; clinically relevant preoperative ALN prediction. (–) Retrospective design; relatively low specificity; performance variability across institutions; no prospective validation; PathDL contribution weaker than ClinicML; model not yet suitable for standalone clinical decision-making.
[54]	To identify optimal inflammatory and nutritional biomarkers associated with distant metastasis in breast cancer using ML-based feature-selection methods.	Retrospective cross-sectional ML-assisted biomarker study.	Nutritional/inflammatory biomarker analysis, Boruta feature selection, LASSO regression, multivariable logistic regression, restricted cubic spline analysis, ROC analysis, and subgroup analysis by molecular subtype were performed.	Guangxi Medical University Cancer Hospital, China.	348 newly diagnosed breast cancer patients: 185 nonmetastatic and 163 distant metastatic cases.	Boruta algorithm, LASSO regression, gradient boosting, ElasticNet, logistic regression modelling.	Spearman correlation analysis; Boruta random-forest feature screening; LASSO variable selection with 10-fold cross-validation and 1-SE rule; z-score standardisation; variance inflation factor (VIF) multicollinearity testing; restricted cubic spline modelling; ROC/Youden index analysis; Gradient Boosting feature importance; ElasticNet feature selection.	Boruta and LASSO identified five key biomarkers: ALI, systemic inflammation response index (SIRI), monocyte-to-lymphocyte ratio (MLR), albumin–globulin ratio (AGR), and geriatric nutritional risk index (GNRI). Higher SIRI and MLR were associated with increased distant metastasis risk, whereas higher ALI, AGR, and GNRI were protective. AGR showed the best individual predictive performance (AUC = 0.650). A nomogram incorporating molecular subtype achieved AUC = 0.744 (95% CI 0.693–0.796), outperforming several single biomarkers.	(+) Comprehensive comparison of 18 inflammatory/nutritional biomarkers; robust ML-based feature-selection workflow; interpretable and clinically accessible biomarkers; included spline and molecular subtype analyses. (–) Single-center retrospective design; relatively small sample size; moderate predictive performance only; no external validation; no prognostic follow-up outcomes included.
[26]	To predict axillary lymph node metastasis preoperatively using simultaneous ^18F-FDG PET/MRI radiomics and machine-learning analysis of primary breast tumors.	Prospective single-center imaging AI study.	Simultaneous PET/MRI radiomics extraction, quantitative imaging analysis, multistep feature selection, SVM classification, and cross-validation were performed to predict axillary lymph node status.	Medical University of Vienna, Austria.	117 female patients with 127 histologically confirmed breast cancer lesions; 85 lymph node-positive and 42 lymph node-negative lesions. Training set: 85 lesions; test set: 42 lesions.	SVM-based radiomics ML model using simultaneous ^18F-FDG PET/MRI.	Whole-lesion segmentation on DCE-MRI, T2WI, DWI/ADC, and PET; PyRadiomics preprocessing; voxel resampling ($1 \times 1 \times 1$ mm^3); gray-level normalisation; wavelet and Laplacian of Gaussian filtering; variance and Pearson-correlation filtering; SMOTE oversampling; information-gain feature selection; five-fold stratified cross-validation.	A total of 4424 radiomics features were extracted, with 13 radiomics features retained for model building. The SVM classifier achieved training accuracy 79.8% (AUC = 0.798) and test accuracy 78.6% (AUC = 0.839), with sensitivity 67.9% and specificity 100% in the test set. The model significantly outperformed baseline classifiers.	(+) Prospective PET/MRI radiomics design; multimodal functional and morphologic imaging integration; robust preprocessing and feature-selection pipeline; independent test-set validation; high specificity for ALN prediction. (–) Small single-center cohort; no external validation; heterogeneous tumor histology; no comparison with ultrasound; no deep-learning architecture; limited PET/MRI availability.
[40]	To predict axillary lymph node metastasis (ALNM) in primary breast cancer using ultrasound radiomics combined with beta-1,3-galactosyltransferase 4 (B3GALT4) expression.	Retrospective single-center imaging-genomic AI study.	Ultrasound radiomics extraction, quantitative reverse transcription polymerase chain reaction (qRT-PCR) analysis of B3GALT4, machine-learning model construction, nomogram development, ROC evaluation, calibration analysis, and decision-curve analysis were performed.	Xuzhou Central Hospital, China.	135 breast cancer patients undergoing ultrasound and axillary lymph node dissection. Training cohort: 94 patients; validation cohort: 41 patients.	SVM-based radiomics classifier; integrated radiomics-clinical nomogram; ML models including LR, KNN, RF, XGBoost, Extra Trees, multilayer perceptron (MLP), and NB.	Manual ROI segmentation using ITK-SNAP; ICC reproducibility filtering (>0.75); voxel resampling ($1 \times 1 \times 1$ mm^3); z-score normalisation; PyRadiomics extraction; Spearman correlation filtering; mRMR feature selection; LASSO regression with 10-fold cross-validation; quantitative reverse transcription polymerase chain reaction (qRT-PCR) quantification of B3GALT4 mRNA.	A total of 1562 radiomics features were extracted, with 12 final features retained after mRMR and LASSO selection. The SVM radiomics model achieved AUCs of 0.937 (training) and 0.932 (validation). B3GALT4 expression was significantly higher in the ALNM group ($p < 0.001$). The combined nomogram integrating radiomics and B3GALT4 achieved the best performance with AUCs of 0.991 (training) and 0.975 (validation), accuracy 0.947/0.927, and specificity 0.940/0.960 in the training/validation cohorts.	(+) Integrated ultrasound radiomics with genomic biomarker analysis; high predictive accuracy; multistep feature-selection workflow; calibration and decision-curve analyses; clinically interpretable nomogram. (–) Retrospective single-center design; relatively small cohort; manual ultrasound ROI segmentation; no external multicenter validation; possible overfitting due to very high model performance.

Table A4. *Cont.*

Reference	Aim of the Study	Research Type	Study Methods	Countries/Setting/Location	Participant Details	AI System/Algorithm/Application Involved	Feature-Processing System or Technique Used	Results	Strengths (+) & Limitations (−)
[65]	To develop and validate an MRI radiomics and clinical-feature model for predicting axillary lymph node (ALN) metastasis in triple-negative breast cancer (TNBC).	Retrospective multicenter radiomics ML study.	MRI T2WI radiomics extraction, clinical-feature analysis, LASSO feature selection, ML model construction, nomogram development, calibration analysis, and decision-curve analysis were performed.	First Affiliated Hospital of Soochow University and Jiangsu Province Hospital of Traditional Chinese Medicine, China.	163 TNBC patients (2016–2023); 54 ALN-metastatic and 109 nonmetastatic cases. Training cohort: 115 patients; validation cohort: 48 patients.	LR, RF, SVM.	Manual ROI segmentation using ITK-SNAP; Pyradiomics extraction; extraction of 1132 radiomics features; z-score normalisation; Pearson/Spearman correlation analysis; LASSO regression with 10-fold cross-validation; randomised-search hyperparameter optimisation; t2-weighted imaging–short tau inversion recovery (T2WI-STIR) MRI radiomics.	LASSO selected four optimal radiomics features. In the validation cohort, the LR-based combined clinical-radiomics model achieved the best performance with AUC 0.828 (95% CI 0.706–0.950), sensitivity 0.813, and accuracy 0.812. RF showed highest specificity (0.906) but poorer stability. Nottingham grade and tumor size were significant predictors of ALN metastasis.	(+) TNBC-specific predictive framework; multicenter dataset; integration of clinical and radiomics data; calibration and decision-curve analyses; interpretable nomogram. (–) Small cohort size; retrospective design; manual segmentation variability; no independent external validation; only T2WI sequence analysed; no deep-learning architecture evaluated.
[27]	To develop a machine-learning model integrating ultrasound radiomics and clinical features for predicting axillary lymph node metastasis (ALNM) in early breast cancer.	Retrospective radiomics ML study.	Intratumoral and peritumoral ultrasound radiomics extraction, clinical-feature analysis, ML model comparison, SHAP interpretability analysis, and cross-validation were performed.	First Affiliated Hospital of Shihezi University, China.	321 breast cancer patients (2019–2023); 121 ALNM-positive and 200 ALNM-negative cases. Training cohort: 224 patients; validation cohort: 97 patients.	XGBoost, RF, SVM, LR, DT, KNN.	Manual ROI segmentation using 3D-Slicer; intratumoral plus 1–5 mm peritumoral expansion; extraction of 107 radiomics features per model; ICC reproducibility analysis (>0.8); MRMR feature filtering; LASSO dimensionality reduction; 10-fold cross-validation; grid-search hyperparameter optimisation; SHAP feature attribution analysis.	The intratumoral + 3 mm peritumoral radiomics model achieved the best radiomics performance with AUCs of 0.847 (training) and 0.844 (validation). The combined clinical-radiomics XGBoost model achieved the highest diagnostic performance with AUCs of 0.917 (training) and 0.905 (validation), accuracy 0.841, sensitivity 0.856, specificity 0.819, PPV 0.877, and F1-score 0.867. SHAP analysis identified Rad score as the strongest predictor, followed by HER2 expression, microcalcifications, and Adler flow grade.	(+) Integrated intratumoral and peritumoral ultrasound radiomics; comparison of multiple ML algorithms; SHAP-based explainability; robust cross-validation and calibration analyses; clinically interpretable multimodal model. (–) Single-center retrospective study; limited geographic diversity; no external validation; manual segmentation variability; elastography and contrast-enhanced ultrasound were not included.
[55]	To develop AI models using clinical blood markers and ultrasound maximum lesion diameter to predict distant metastasis in breast cancer.	Retrospective multicenter ML study.	Clinical blood-marker analysis, ultrasound-feature integration, feature reduction, LightGBM model development, internal/external validation, feature-importance analysis, regression analysis, and decision-curve analysis were performed.	Guangxi Medical University centers, China.	416 female breast cancer patients from two centers. Training cohort: 274 patients; internal test cohort: 68 patients; external validation cohort: 74 patients. Included de novo bone, liver, and lung metastasis cases.	LightGBM	z-score normalisation; Spearman correlation filtering (>0.9 threshold); LASSO feature selection with mean squared error (MSE) optimisation; fivefold cross-validation; grid-search hyperparameter optimisation; feature-importance analysis; univariate and multivariate regression analyses.	The clinical LightGBM model achieved AUCs of 0.950 (training), 0.795 (internal test), and 0.883 (external validation). The combined model integrating blood markers and ultrasound diameter achieved higher AUCs of 0.955, 0.835, and 0.918, respectively. Key predictive features included CK-MB, CEA, cancer antigen 15-3 (CA15-3), albumin, creatine kinase, and maximum lesion diameter on ultrasound. CA153, CK-MB, lipoprotein(a), and lesion diameter positively correlated with distant metastasis.	(+) Included external validation cohort; used inexpensive and clinically accessible biomarkers; strong LightGBM performance; integrated laboratory and ultrasound data; feature-importance and regression analyses improved interpretability. (–) Retrospective regional dataset; no deep-learning architecture; limited ultrasound variables beyond lesion diameter; included only de novo bone/liver/lung metastases; multicenter diversity limited to one geograp-hic region.

Table A4. *Cont.*

Reference	Aim of the Study	Research Type	Study Methods	Countries/Setting/Location	Participant Details	AI System/Algorithm/Application Involved	Feature-Processing System or Technique Used	Results	Strengths (+) & Limitations (−)
[42]	To identify the optimal MRI-based machine-learning model for preoperative prediction of sentinel lymph node (SLN) metastasis in breast cancer.	Retrospective MRI radiomics ML study.	DWI-MRI feature extraction, feature-selection benchmarking, classifier comparison, repeated subset analysis, and 10-fold cross-validation were performed to identify the most stable and accurate predictive model.	Southern Medical University, Guangzhou, China.	172 histologically confirmed breast cancer patients: 74 SLN-metastatic and 98 nonmetastatic cases. Fifty random subsets of 138 patients each were generated for stability analysis.	RF, SVM, KNN, DT, generalised linear model (GLM), multivariate adaptive regression splines (MARS), bagging (BAG), boostingNB, partial least squares regression (PLSR), Linear Discriminant Analysis (L-DA).	Diffusion-weighted imaging magnetic resonance imaging (DWI-MRI) radiomics extraction; manual 3D ROI segmentation using ITK-SNAP; wavelet bandpass filtering; isotropic resampling; gray-level quantization; extraction of 10,320 texture and 4 non-texture features; min–max normalisation; feature-selection benchmarking using 24 methods including MRMR, l1-regularised support vector machine (L1-SVM), and sequential forward floating selection (SFFS); repeated 10-fold cross-validation; stability analysis using relative standard deviation (RSD).	The combination of MRMR + RF achieved the best predictive performance with AUC 0.97 ± 0.03 and accuracy 0.89 ± 0.05, while also demonstrating the best model stability (RSD 2.94). L1-SVM showed the best average feature-selection performance, and SFFS demonstrated the highest stability among feature-selection methods. Texture features outperformed non-texture MRI features for SLN prediction.	(+) Extensive benchmarking of 24 feature-selection and 11 classification methods; robust stability analysis using repeated subset validation; comprehensive MRI texture-feature engineering; strong predictive performance for non-invasive SLN assessment. (–) No independent external validation; retrospective single-center design; no deep-learning architecture; manual segmentation dependency; magnetic-field inhomogeneity effects not addressed; limited explainability of selected imaging biomarkers.
[36]	To develop a nomogram integrating radiomics and deep-learning MRI signatures for preoperative prediction of axillary lymph node (ALN) metastasis in breast cancer.	Retrospective multimodal AI imaging study.	DCE-MRI radiomics extraction, VGG-16 deep-learning feature extraction, feature selection, ML classifier comparison, logistic-regression nomogram construction, calibration analysis, and decision-curve analysis were performed.	Tongji Hospital, Wuhan, China.	151 invasive breast cancer patients. Training cohort: 106 patients; test cohort: 45 patients.	VGG-16 transfer-learning CNN; SVM, KNN, RF, DT, ExtraTrees, XGBoost, LightGBM; multivariable logistic-regression nomogram.	Manual 3D ROI segmentation using ITK-SNAP; Pyradiomics extraction of 120 radiomics features; transfer learning with ImageNet-pretrained VGG-16; image augmentation (flipping, rotation, displacement); two-stage feature selection using *t*-tests and LASSO regression; fivefold cross-validation; DCE-MRI-based feature extraction.	Five radiomics and two deep-learning features were retained after selection. In the test cohort, most radiomics models achieved AUCs > 0.80 except DecisionTree and ExtraTrees. KNN, XGBoost, and LightGBM using deep-learning features also achieved AUCs >0.80. The combined nomogram integrating the radiomics signature, deep-learning signature, and MRI-reported lymph node (LN) status achieved AUCs of 0.90 (95% CI, 0.85–0.96) in the training cohort and 0.90 (95% CI, 0.80–0.99) in the test cohort, outperforming either signature alone.	(+) Hybrid integration of radiomics and deep-learning MRI features; multimodal nomogram improved prediction performance; calibration and DCA analyses demonstrated clinical utility; comparison of multiple ML classifiers; non-invasive ALN prediction framework. (–) Retrospective single-center study; relatively small sample size; manual segmentation dependency; no external validation; possible overfitting risk due to limited training data.
[51]	To develop a pathomics-clinical ML nomogram for preoperative prediction of axillary lymph node metastasis (ALNM) in breast cancer.	Retrospective two-center pathomics ML study.	H&E whole-slide image digitisation, ROI tile selection, pathomics extraction, feature selection, ML classifier comparison, logistic-regression nomogram construction, calibration analysis, and decision-curve analysis were performed.	Chongqing University Cancer Hospital and First Affiliated Hospital of Chongqing Medical University, China.	407 women with stage I–III invasive breast cancer (2019–2023). Training cohort: 203 patients; internal validation cohort: 136 patients; external validation cohort: 68 patients. All underwent sentinel lymph node biopsy (SLNB) or ALND.	LR, DT, RF, KNN, LGBM, XGB; ALNET pathomics-clinical nomogram.	H&E WSI digitisation; seven non-overlapping 1000 × 1000 ROI tiles per case; CellProfiler extraction of 1121 pathomics features; normalisation; Mann–Whitney U-test filtering; Spearman correlation filtering (>0.9 removed); LASSO with 10-fold cross-validation; Path-score construction; ROC, calibration, DCA, and DeLong testing.	LASSO selected eight pathomics features to generate the Path-score. Independent predictors of ALNM were ER status, HER2 status, tumor size, and Path-score. Among six ML models, LR achieved the best validation performance. The final nomogram achieved AUCs of 0.783 in the internal validation cohort and 0.738 in the external validation cohort, outperforming the clinical-only model (external AUC 0.574). External validation accuracy was 0.779, sensitivity 0.742, and specificity 0.811. Calibration and DCA demonstrated good model reliability and clinical benefit.	(+) Two-center study with external validation; integration of microscopic pathomics and clinical variables; extensive ML comparison; good calibration and clinical-benefit analysis; practical ALNET mini-program for clinical prediction. (–) Retrospective design; moderate sample size; manual ROI selection; limited population diversity; no prospective multicenter validation; sensitivity remained moderate.

Table A4. *Cont.*

Reference	Aim of the Study	Research Type	Study Methods	Countries/Setting/Location	Participant Details	AI System/Algorithm/Application Involved	Feature-Processing System or Technique Used	Results	Strengths (+) & Limitations (−)
[28]	To develop a machine learning-driven ultrasound radiomics model integrating intratumoral and peritumoral features with clinical variables for predicting high axillary lymph node burden (AHNB) in breast cancer.	Prospective and retrospective radiomics ML diagnostic study.	Ultrasound ROI segmentation, intratumoral/peritumoral radiomics extraction, feature selection, ML classifier comparison, nomogram construction, ROC analysis, calibration analysis, bootstrap validation, and decision-curve analysis were performed.	First Affiliated Hospital of Shihezi University, Xinjiang, China.	131 breast cancer patients with axillary lymph node metastasis (2019–2024): 79 low-burden ($\leq$2 metastatic nodes) and 52 high-burden (>2 metastatic nodes) cases. Training cohort: 92 patients; validation cohort: 39 patients.	LR, DT, SVM, XGBoost, RF, KNN	Manual ROI segmentation using 3D-Slicer; intratumoral plus 1–5 mm peritumoral expansion; extraction of 107 radiomics features per ROI; ICC reproducibility analysis (>0.8); MRMR feature filtering; LASSO regression with nested 10-fold cross-validation; grid-search hyperparameter tuning; bootstrap calibration; nomogram visualisation.	The intratumoral + 3 mm peritumoral radiomics model achieved the best radiomics performance (validation AUC = 0.771). The combined LR clinical-radiomics model showed the best overall performance, with AUCs of 0.845–0.857 in the training cohort and 0.806–0.820 in the validation cohort. Validation accuracy, sensitivity, specificity, PPV, NPV, and F1-score were 0.897, 0.879, 0.663, 0.860, 0.692, and 0.835, respectively. Independent predictors included vascular invasion, tumor diameter >2 cm, Adler blood-flow grade, and Rad-score. DCA demonstrated optimal clinical utility at a 52% probability threshold.	(+) Combined clinical and intratumoral/peritumoral radiomics integration; comparison of multiple ML classifiers; calibration, bootstrap validation, and DCA analyses; interpretable nomogram; biologically relevant peritumoral analysis. (–) Small single-center cohort; limited external validation; exclusion of tumors > 5 cm and multifocal lesions; potential sensitivity to radiomics parameter selection; manual segmentation and ultrasound operator dependence; no multimodal imaging integration.
[56]	To predict breast cancer metastasis and identify genomic biomarkers using machine learning and explainable AI (XAI).	Genomics + ML + XAI predictive modelling study.	Gene-expression analysis, elastic net feature selection, ML classifier comparison, SHAP explainability analysis, and biostatistical evaluation were performed.	Turkey-based research using NCBI GEO public genomic datasets.	98 sporadic primary breast cancer samples from lymph node-negative patients aged <55 years; 34 patients developed distant metastasis within 5 years and 44 remained disease-free for $\geq$5 years.	LightGBM, CatBoost, XGBoost, gradient boosting trees (GBT), AdaBoost; SHAP explainable AI framework.	Robust standardisation; elastic net feature selection; microarray gene-expression profiling (~24,481 genes); Mann–Whitney U-test; Spearman correlation analysis; SHAP feature-importance analysis.	Elastic net reduced 24,481 genes to 18 candidate biomarkers (99.93% reduction). LightGBM achieved the best performance with an accuracy 96%, an AUC 99.3%, and a Brier score of 0.024. SHAP analysis identified TSPYL5, CACTIN, ATP5E, carbonic anhydrase 9 (CA9), and NUP210 as the most important metastasis-associated genes. Increased expression of TSPYL5, ATP5E, CA9, NUP210, SLC37A1, ARIH1, PSMD7, UBQLN1, PRAME, and UBE2T was associated with increased metastasis risk, whereas reduced expression of CACTIN, TGFB3, SCUBE2, ARL4D, OR1F1, ALDH4A1, PHF1, and CROCC was also linked to metastasis.	(+) Strong explainable-AI framework using SHAP; high predictive performance; robust dimensionality reduction; interpretable genomic biomarker prioritisation; comparison of multiple boosting algorithms. (–) Small retrospective genomic dataset; no independent external validation; limited generalisability; no deep-learning comparison; requires larger and more diverse validation cohorts.
[71]	To develop MRI radiomics-based ML models for identifying the primary tumor site of brain metastases, including breast cancer metastases.	Retrospective radiomics + ML diagnostic modelling study.	Multiparametric MRI radiomics extraction, supervised ML classification, feature selection, SHAP explainability analysis, and ROC/AUC validation were performed to differentiate brain metastases from lung, breast, and gastrointestinal cancers.	West China Hospital, Sichuan University, China.	202 patients with 439 brain metastases: 244 lung cancer metastases, 135 breast cancer metastases, and 60 gastrointestinal tumor metastases.	Support Vector Machine (SVM), Logistic Regression (LR), LASSO Logistic Regression.	MRI radiomics extraction from T1WI, T2WI, fluid-attenuated inversion recovery (FLAIR), and Contrast-Enhanced T1-Weighted Imaging (CE-T1WI) sequences; semi-automatic 3D Slicer segmentation; ANOVA, Recursive Feature Elimination (RFE), and Kruskal–Wallis feature selection; SMOTE balancing; z-score normalisation; SHAP feature-importance analysis.	The combined radiomics + clinical model achieved AUCs of 0.965 for lung vs. breast metastases, 0.991 for lung vs. gastrointestinal metastases, and 0.935 for breast vs. gastrointestinal metastases. Radiomics-only models achieved AUCs of 0.915 (lung vs. GI), 0.954 (breast vs. GI), and 0.729 (lung vs. breast). SHAP analysis identified key MRI radiomics features that contribute to primary tumor classification.	(+) Multiparametric MRI radiomics framework; multiclass tumor-origin discrimination; integration of explainable AI using SHAP; robust feature-selection pipeline; clinically relevant differentiation of brain metastasis origin. (–) Retrospective single-center design; moderate sample size; semi-automatic ROI segmentation dependency; no deep-learning comparison; limited MRI sequence diversity; no independent external validation cohort.

Table A4. *Cont.*

Reference	Aim of the Study	Research Type	Study Methods	Countries/Setting/Location	Participant Details	AI System/Algorithm/Application Involved	Feature-Processing System or Technique Used	Results	Strengths (+) & Limitations (−)
[41]	To develop and validate a multimodal ultrasound deep-learning radiomics model integrating B-mode ultrasound (BMUS) and color doppler ultrasound (CDUS) for noninvasive prediction of sentinel lymph node (SLN) metastasis in breast cancer.	Retrospective multicenter radiomics + deep-learning diagnostic modelling study.	Ultrasound radiomics extraction, dual-stream CNN deep-learning modelling, feature fusion, dimensionality reduction, feature selection, ROC/AUC evaluation, and class activation map (CAM) interpretability analysis was performed.	Shanghai Pudong New Area People's Hospital and Shuguang Hospital Affiliated to Shanghai University of Traditional Chinese Medicine, China.	450 women with unilateral unifocal invasive breast cancer (2021–2025). Training cohort: 276 patients; external validation cohort: 105 patients; testing cohort: 69 patients.	USCD-Net dual-stream MobileNetV2 architecture; SVM; transfer learning–based deep-learning radiomics framework.	BMUS + CDUS multimodal imaging integration; PyRadiomics extraction; PCA reduction of 2048 deep features to 256 dimensions; ICC reproducibility analysis; Pearson correlation filtering; LASSO regression with 10-fold cross-validation; z-score normalisation; feature-level fusion; CAM visualisation; data augmentation.	The COMB model, which integrates handcrafted and deep-learning features, achieved the best SLN metastasis prediction performance, with AUCs of 0.888 (training), 0.861 (validation), and 0.837 (testing), outperforming handcrafted-only and deep-feature-only models. The model integrated 10 final features (6 handcrafted and 4 deep learning features). Negative predictive values reached 88.89% (training), 76.60% (validation), and 71.23% (testing). MobileNetV2 outperformed ResNet18, ResNet50, EfficientNetV2-S, and Vision Transformer (ViT) backbones. CAM analysis identified lesion boundaries and lobulated regions as key predictive areas.	(+) Large multimodal ultrasound dataset; multicenter external validation; integration of handcrafted radiomics and deep-learning features; explainable AI using CAM visualisation; comparison of multiple CNN backbones; clinically relevant high NPV for reducing unnecessary SLNB. (–) Retrospective design; moderate cohort size; lower performance for three-class metastatic-burden prediction; exclusion of multifocal/bilateral tumors; semi-manual ROI delineation; possible overfitting; limited international generalisability.
[49]	To develop a multimodal deep-learning framework (Integrated Clinical, Sequencing and Digital Pathology Attention [ICSDA]) model integrating histopathological images, clinical variables, and gene-expression data for predicting breast cancer recurrence and metastasis risk.	Retrospective multimodal deep-learning prognostic prediction study.	Whole-slide image preprocessing, weakly supervised learning, transfer learning, attention-based MIL, multimodal feature fusion, ROC/AUC evaluation, Kaplan–Meier survival analysis, and external validation were performed.	China-based study using TCGA breast cancer dataset; external validation from Cancer Hospital of the Chinese Academy of Medical Sciences.	196 TCGA breast cancer patients with matched H&E images, clinical data, and RNA sequencing (RNA-seq) data: 124 non-recurrence/non-metastasis and 72 recurrence/metastasis cases. External validation cohort: 101 HER2-positive breast cancer patients, including 28 recurrence/metastasis cases.	Integrated clinical, sequencing and digital pathology attention model (ICSDA); ResNet18 encoder; attention-based weakly supervised deep-learning framework; RF, LR, and SVM comparison models.	Clustering-constrained attention multiple instance learning (CLAM)-based tissue segmentation; 256 × 256 patch extraction at 20× magnification; ResNet18 feature encoding; transfer learning; attention pooling; random forest gene selection using Gini index; multimodal fusion of pathology, clinical, and sequencing data; SGD optimisation; dropout regularisation; fivefold cross-validation; attention heatmap visualisation.	ICSDA achieved a sensitivity of 0.67, a specificity of 0.87, and a test AUC of 0.75, outperforming image-only and traditional ML models. Incorporating H&E image information improved prediction performance compared with clinical or sequencing data alone. Kaplan–Meier analysis showed significant survival separation between predicted high- and low-risk groups ($p = 0.0181$). Fivefold cross-validation achieved AUC 0.73. External validation achieved AUC 0.604.	(+) Multimodal integration of pathology, clinical, and genomic data; weakly supervised whole-slide analysis without manual ROI annotation; interpretable attention heatmaps; comparison against RF/LR/SVM baselines; inclusion of external validation cohort. (–) Small sample size; retrospective design; limited external validation performance; arbitrary 7:3 train-test split; heterogeneous breast cancer subtypes; limited generalisability; moderate predictive performance.
[67]	To characterise plasma proteomic and metabolomic signatures associated with bone, liver, and lung metastases in breast cancer and identify predictive biomarkers using machine learning.	Retrospective multi-omics biomarker discovery and validation study.	Plasma proteomics using ultra-performance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS), untargeted metabolomics, orthogonal partial least squares discriminant analysis (OPLS-DA), kyoto encyclopedia of genes and genomes (KEGG) and Metascape enrichment analyses, Mfuzz clustering, LASSO regression, Random Forest ranking, ROC/AUC analysis, and enzyme-linked immunosorbent assay (ELISA) validation in an independent cohort were performed.	Zhejiang Cancer Hospital, China.	Discovery cohort: 51 female breast cancer patients (10 non-metastatic, 11 bone metastasis, 19 liver metastasis, 11 lung metastasis). Validation cohort: 19 additional patients (2 non-metastatic, 4 bone, 4 liver, 9 lung metastasis).	LASSO regression model; Random Forest machine-learning classifier; OPLS-DA statistical modelling.	Multi-omics integration of plasma proteomics and metabolomics; UPLC-MS/MS; Pareto scaling; K-nearest-neighbor imputation; OPLS-DA; KEGG and Metascape enrichment; Mfuzz clustering; Random Forest mean decreases accuracy ranking; ELISA validation for fibronectin 1 (FN1) and heparan sulfate proteoglycan 2 (HSPG2).	Distinct molecular signatures were identified for each metastatic site. Combined biomarker panels achieved AUCs of 0.955 for bone metastasis, 0.941 for liver metastasis, and 0.989 for lung metastasis. External validation cohort AUCs were 0.900 (bone), 0.783 (liver), and 0.800 (lung). Bone metastasis was associated with extracellular matrix remodeling and platelet activation; liver metastasis with retinol metabolism; and lung metastasis with leukocyte transendothelial migration and inflammation. Validated biomarkers included FN1, HSPG2, leucyl-tryptophan, and LysoPC(P-16:0/0:0) for bone metastasis; dUDP, LPE (18:1/0:0), and aspartylphenylalanine for liver metastasis; and testosterone sulfate, dUDP, and PE (14:0/20:5) for lung metastasis.	(+) Integrated proteomics and metabolomics with machine-learning-driven biomarker discovery; included external validation cohort; biologically interpretable pathway analysis; site-specific biomarker panels with high diagnostic performance. (–) Small sample size with possible overfitting risk; retrospective single-center design; limited external validation; some metabolites unavailable for targeted validation; uncertain generalisability across broader populations and breast cancer subtypes.

Table A4. *Cont.*

Reference	Aim of the Study	Research Type	Study Methods	Countries/Setting/Location	Participant Details	AI System/Algorithm/Application Involved	Feature-Processing System or Technique Used	Results	Strengths (+) & Limitations (−)
[29]	To develop and validate MRI radiomics-based machine-learning models for preoperative prediction of axillary lymph node metastasis (ALNM) and explore associations between radiomics and the tumor microenvironment in early-stage invasive breast cancer.	Retrospective multicent machine-learning radiomics study with external and prospective-retrospective validation cohorts.	Multi-sequence breast MRI radiomics extraction, tumor and ALN ROI segmentation, Random Forest feature selection, SVM model construction, ROC/AUC evaluation, decision-curve analysis, radiologist comparison, and tumor microenvironment correlation analyses were performed using TCGA/TCIA datasets.	Four institutions in China: Sun Yat-sen Memorial Hospital, Sun Yat-sen University Cancer Center, Shunde Hospital of Southern Medical University, and Tungwah Hospital of Sun Yat-sen University.	1088 women with stage I–III invasive breast cancer after quality control. Training cohort: 803; external validation cohort: 179; prospective-retrospective validation cohort: 106 patients receiving neoadjuvant chemotherapy.	RF feature-selection algorithm; SVM; cell-type identification by estimating relative subsets of RNA transcripts (CIBERSORT) immune-cell deconvolution algorithm.	Multi-sequence MRI radiomics from tumor and ALN ROIs using contrast-enhanced t1-weighted imaging (T1 + C), T2WI, and DWI-ADC sequences; n4 insight toolkit bias-field correction algorithm (N4ITK) bias-field correction; semi-automatic 3D Slicer segmentation; PyRadiomics extraction; z-score normalisation; GLCM, GLRLM, GLSZM, GLDM, and NGTDM texture features; wavelet features; multiomic integration with clinicopathologic and molecular-subtype data.	The ALN-tumor radiomic signature achieved AUCs of 0.88 (training), 0.87 (external validation), and 0.87 (prospective-retrospective validation). The multiomic signature improved performance to AUCs of 0.90, 0.91, and 0.93, respectively, outperforming radiologists in ALNM prediction. Significant radiomic changes after neoadjuvant chemotherapy were observed, particularly in GLDM features. Radiomic features correlate immune cells, long non-coding RNAs, and methylation profiles within the tumor microenvironment.	(+) Very large multicenter cohort; external and prospective validation; integration of tumor and ALN MRI radiomics; multi-sequence MRI analysis; comparison against radiologists; biologically interpretable tumor microenvironment analysis; strong clinical decision-support potential. (–) Retrospective design; MRI heterogeneity across centers; semi-manual ROI segmentation; incomplete TCIA MRI sequence availability; indirect biological validation only; uncertainty in sentinel lymph-node diagnosis variability.
[52]	To develop a weakly supervised AI framework for predicting breast cancer lymph node metastasis (BCLNM) from whole-slide pathology images (WSIs), with improved micro-metastasis detection and interpretability.	Weakly supervised deep-learning computational pathology study using multiple instance learning (MIL).	Whole-slide image tiling, CNN feature extraction, prototype discovery using unsupervised clustering, prototype-based slide embedding, metric learning, top-K patch selection, ROC/AUC evaluation, cross-center validation, and visualisation analyses were performed.	Zhujiang Hospital, Southern Medical University, China, and the public Camelyon16 multicenter dataset.	Camelyon16 dataset: 399 WSIs (270 training, 129 testing). ZBraln dataset: 635 H&E WSIs from breast cancer patients (247 lymph node metastasis (LNM)-positive, 388 LNM-negative) collected 2015–2020.	Prototypical multiple instance learning (PMIL); ResNet34 backbone; affinity propagation clustering (APC); metric-learning framework; top-K selector.	WSI binarization and 256 × 256 patch extraction; vocabulary-based MIL; prototype discovery module using two-level APC clustering; prototype-based slide embedding; cosine-similarity metric learning; selective top-K pooling; t-SNE visualisation; weak supervision using slide-level labels only; ImageNet-pretrained ResNet34 feature extraction.	PMIL achieved state-of-the-art performance with AUCs of 88.2% on Camelyon16 and 98.4% on ZBraln at 40× magnification, consistently outperforming STREAM, ABMIL, dual-stream multiple instance learning (DSMIL), recurrent neural network multiple instance learning (RNN-MIL), maximum-pooling multiple instance learning (MAX-MIL), and DeepAttnMISL. PMIL substantially improved micro-metastasis identification and demonstrated better cross-centre generalisation than most comparison models. Metric learning improved AUC from 88.8% to 95.8% on validation experiments. Visualisation analyses showed interpretable pathological prototypes and biologically meaningful heatmaps.	(+) Strong weakly supervised performance without pixel-level annotation; interpretable prototype-based pathology representations; improved micro-metastasis detection; robust cross-center generalisation; comprehensive ablation and visualisation analyses; biologically meaningful heatmaps. (–) Computationally intensive WSI processing; dependence on clustering quality and hyperparameter tuning (λ and K); reduced performance on highly heterogeneous datasets such as Camelyon16; prototype discovery requires large-scale clustering operations; micro-metastasis detection still remained challenging.
[30]	To develop and validate a radiomics-based ML nomogram using contrast-enhanced computed tomography (CECT) images of axillary lymph nodes for preoperative prediction of axillary lymph node metastasis (ALNM) in breast cancer.	Retrospective radiomics machine-learning study.	Axillary lymph-node ROI segmentation on venous-phase CECT images, radiomics extraction, Pearson correlation filtering, LASSO logistic regression, multivariate logistic regression, ROC/AUC evaluation, calibration analysis, and decision-curve analysis were performed.	Affiliated Hospital of Southwest Medical University, Luzhou, China.	193 patients with invasive breast cancer (February 2021–February 2022); training cohort: 135 patients; validation cohort: 58 patients; 97 pN0 and 96 pN+ cases.	Radiomics nomogram integrating radiomics score and clinical N stage; LASSO-based ML radiomics model.	2D venous-phase CECT lymph-node segmentation using ITK-SNAP; Pyradiomics extraction of 1218 features including first-order, shape, GLCM, GLDM, GLRLM, GLSZM, wavelet, and LoG features; Pearson correlation filtering (>0.9); normalisation; LASSO feature selection reduced features to 21 predictors; multivariate logistic regression nomogram.	Clinical N stage and radiomics score were independent predictors of ALNM. The radiomics signature achieved an AUC of 0.88 (95% CI, 0.80–0.97) in the validation cohort. The combined nomogram achieved AUCs of 0.95 (95% CI 0.93–0.98) in the training cohort and 0.93 (95% CI 0.87–0.99) in the validation cohort, outperforming radiomics or clinical staging alone. Calibration and decision-curve analyses demonstrated good model fit and superior clinical utility.	(+) non-invasive ALNM prediction using lymph-node-focused radiomics; strong predictive performance; integration of imaging and clinical variables improved discrimination; clinically interpretable nomogram; decision-curve and calibration validation performed. (–) Retrospective single-center design; relatively small sample size; no external validation; potential underfitting risk; manual segmentation required; interobserver reproducibility not assessed.

Table A4. *Cont.*

Reference	Aim of the Study	Research Type	Study Methods	Countries/Setting/Location	Participant Details	AI System/Algorithm/Application Involved	Feature-Processing System or Technique Used	Results	Strengths (+) & Limitations (−)
[37]	To develop a DCE-MRI-based machine-learning radiomics model for predicting axillary lymph node metastasis (ALNM) in breast cancer.	Retrospective MRI radiomics machine-learning diagnostic study.	Pre-treatment DCE-MRI acquisition, manual tumor segmentation, 2D/3D radiomics extraction, feature selection, LASSO-based ML modelling, comparative model analysis, ROC/AUC evaluation, 10-fold cross-validation, and nomogram construction were performed.	The First Central Hospital of Baoding, China.	166 patients with pathologically confirmed breast cancer (2017–2020). Training cohort: 133 patients; test cohort: 33 patients. Patients with multifocal, bilateral lesions, or poor-quality imaging were excluded.	LASSO-based machine-learning radiomics model; integrated radiomics-clinical nomogram.	Manual 3D volume of interest (VOI) segmentation using 3D Slicer; PyRadiomics extraction of 9863 3D features and 7224 2D features from multi-phase DCE-MRI; ICC reproducibility testing; Mann–Whitney U-test filtering; correlation filtering (R >0.8 removed); feature normalisation; LASSO feature selection; multi-period MRI feature fusion; heatmap visualisation.	The 3D multi-period MRI radiomics model achieved the best imaging performance, with training/test AUCs of 0.82 and 0.80, respectively. Mean 10-fold cross-validation AUC, accuracy, sensitivity, and specificity were 0.82, 0.82, 0.83, and 0.81. The integrated radiomics-clinical model combining MRI semantic features and radiomics achieved a C-index and AUC of 0.90, specificity 0.91, sensitivity 0.77, accuracy 0.84, PPV 0.91, and NPV 0.78. Multi-period 3D MRI features outperformed 2D, pre-enhancement, and first- and post-enhancement models.	(+) Multi-period 3D DCE-MRI radiomics integration; comprehensive comparison of 2D vs. 3D and temporal MRI models; inclusion of MRI semantic clinical features; reproducibility assessment using ICC; strong diagnostic performance for non-invasive ALNM prediction. (–) Retrospective single-center study; small test cohort; no external or prospective validation; exclusion of multifocal/bilateral tumors reduced generalisability; manual segmentation dependency; ALN images themselves were not directly analysed.
[31]	To develop an interpretable dual-sequence MRI radiomics ML model using DWI and DCE MRI for predicting axillary lymph node metastasis (ALNM) in invasive breast cancer (IBC).	Retrospective MRI radiomics machine-learning study.	MRI ROI delineation, radiomics extraction, ICC reproducibility testing, variance-threshold filtering, SelectKBest and LASSO feature selection, Random Forest modelling, ROC/AUC analysis, DeLong testing, IDI analysis, DCA, and SHAP interpretability analysis were performed.	Affiliated Hospital of North Sichuan Medical College, Nanchong, China.	183 pathologically confirmed IBC patients (107 ALNM, 76 non-ALNM) evaluated between June 2021 and December 2023. Training cohort: 128 patients; testing cohort: 55 patients.	RF radiomics models using DWI, DCE, and combined DWI + DCE MRI sequences; SHAP explainable AI framework.	3D Slicer ROI segmentation; Pyradiomics extraction of 1223 features; dual-sequence DWI and DCE MRI radiomics; ICC filtering ($\geq$0.75 retained); z-score normalisation; variance-threshold filtering; SelectKBest; LASSO feature selection; SHAP (Shapley Additive Explanations); texture features including GLCM, GLDM, GLRLM, GLSZM, shape, and first-order features.	The combined DWI + DCE model achieved the best predictive performance with AUCs of 0.864 (training) and 0.859 (testing), outperforming single-sequence models. DWI and DCE models achieved testing AUCs of 0.728 and 0.760, respectively. IDI improved by 13.5% versus DWI and 10.2% versus DCE ($p < 0.05$). SHAP analysis identified the DCE "original shape Flatness" feature as the most influential predictor. Decision-curve analysis demonstrated superior clinical benefit for the combined model.	(+) Dual-sequence MRI integration improved predictive accuracy; interpretable AI using SHAP; robust radiomics feature-selection pipeline; clinically relevant non-invasive ALNM prediction; DCA and IDI analyses strengthened translational value. (–) Retrospective single-center study; limited sample size; no external validation; manual ROI segmentation dependency; single-scanner dataset may reduce generalisability; clinical variables were not effectively integrated into the final model.
[57]	To develop and validate an AI-based clinical blood biomarker model for predicting de novo distant bone metastasis in newly diagnosed breast cancer patients.	Retrospective dual-center machine-learning diagnostic study.	Clinical blood biomarker collection, z-score normalisation, Spearman correlation filtering, LASSO feature selection, LightGBM model construction, 5-fold cross-validation, ROC analysis, DCA, and univariate/multivariate logistic regression analyses were performed.	Guangxi Medical University Cancer Hospital and affiliated centers, China.	239 female breast cancer patients from two centers. Primary center: 176 cases split into training (n = 123) and testing (n = 53) cohorts. External validation cohort: 63 patients from a second center.	LightGBM machine-learning model for predicting de novo breast cancer bone metastasis.	Clinical blood biomarker preprocessing using z-score normalisation; Spearman correlation filtering ($\rho > 0.9$); LASSO regression with 10-fold cross-validation; feature-importance analysis; logistic regression analysis; DCA. Biomarkers included tumor markers, liver/kidney function, lipid profile, and cardiovascular markers.	The LightGBM model achieved AUCs of 0.945 (training), 0.892 (internal testing), and 0.908 (external validation). The combined model achieved AUCs of 0.955, 0.835, and 0.918, respectively. The model demonstrated high sensitivity, specificity, and accuracy, with external validation accuracy reaching 0.841 and specificity 0.906. Key predictive biomarkers included carcinoembryonic antigen (CEA), creatine kinase, albumin-globulin ratio, apolipoprotein B (ApoB), and CA153. Lipoprotein(a), CA153, γ-glutamyl transferase, alpha-hydroxybutyrate dehydrogenase (α-HBDH), alkaline phosphatase, and creatine kinase were positively associated with bone metastasis, while total cholesterol and white blood cell count were negatively associated.	(+) Dual-center validation improved robustness; high predictive accuracy using low-cost routinely available biomarkers; external validation cohort included clinically accessible and non-invasive variables; practical potential for resource-limited settings. (–) Retrospective design; limited geographic diversity; focused only on bone metastasis; no imaging/genomic integration; biomarker specificity limitations; possible overfitting despite validation; external cohort originated from a geographically similar region.

Table A4. *Cont.*

Reference	Aim of the Study	Research Type	Study Methods	Countries/Setting/Location	Participant Details	AI System/Algorithm/Application Involved	Feature-Processing System or Technique Used	Results	Strengths (+) & Limitations (−)
[58]	To develop and validate AI models that integrate clinical blood biomarkers, ultrasound findings, and breast biopsy pathology features to predict distant metastasis in breast cancer.	Retrospective dual-center machine-learning diagnostic study.	Clinical blood biomarker, ultrasound, and biopsy pathology feature extraction; z-score normalisation; Spearman correlation filtering; LASSO feature selection; ML model construction; internal/external validation; ROC/AUC analysis; DCA; and feature-importance analysis were performed.	Guangxi Medical University Cancer Hospital and the First Affiliated Hospital of Guangxi Medical University, China.	417 female breast cancer patients from two centers. Training cohort: 274 patients; internal testing cohort: 68 patients; external validation cohort: 75 patients. Distant metastases mainly involved bone, liver, and lung sites.	LR, LightGBM, gradient boosting, and AdaBoost machine-learning models.	Z-score normalisation; Spearman correlation filtering (>0.9); LASSO regression with 10-fold cross-validation (optimal λ = 0.0193); 27 final selected features; multimodal integration of blood biomarkers, ultrasound lesion diameter, and estrogen receptor/progesterone receptor/human epidermal growth factor receptor 2 (ER/PR/HER2)/Ki-67 proliferation index (Ki67) pathology features; feature-importance analysis; DCA.	The LightGBM model achieved the best predictive performance with AUCs of 0.971 (training), 0.861 (internal testing), and 0.890 (external validation), outperforming the LR model (AUCs 0.892, 0.816, and 0.817, respectively). DCA demonstrated greater net clinical benefit for LightGBM. Key predictive features included creatine kinase-mb isoenzyme (CK-MB), CA153, α-HBDH, ApoB, and CEA. In the LR model, α-HBDH, Ki67, alkaline phosphatase (ALP), ultrasound lesion diameter, and CEA contributed most strongly.	(+) Dual-center design with external validation; integration of multimodal clinical, ultrasound, and pathology data; strong LightGBM predictive performance; clinically accessible low-cost features; potential to reduce dependence on expensive imaging; comparison of multiple ML algorithms. (–) Retrospective design; geographically limited Chinese cohorts; brain and post-treatment metastases excluded; possible institutional/equipment variability; limited international generalisability; further multicenter validation required.
[32]	To develop a CT-based deep-learning radiomics system for distinguishing bone tumors originating from breast cancer metastasis from those of non-breast-cancer origin.	Retrospective CT radiomics and deep-learning diagnostic study.	CT preprocessing, automated tumor segmentation, radiomics extraction, ICC reproducibility analysis, PCA dimensionality reduction, LASSO feature selection, ML classifier comparison, 10-fold cross-validation, and ROC/AUC evaluation were performed.	Zhejiang Institute of Economics and Trade and Wuxi Hospital of Traditional Chinese Medicine, China.	178 patients with bone tumors on CT imaging, including 78 breast cancer bone metastases and 100 non-breast-cancer bone metastases.	Medical Image Segmentation via Self-distilling TransUNet (MISSU); SVM; LR; KNN; RF; Extra Trees; XGBoost; LightGBM; Multi-Layer Perceptron (MLP).	Automated CT tumor segmentation using a CNN-transformer hybrid MISSU framework; Pyradiomics extraction; ICC filtering (ICC > 0.9); PCA dimensionality reduction; LASSO regression; texture-feature extraction including GLCM, GLRLM, and GLSZM; 10-fold cross-validation.	The SVM model achieved the best diagnostic performance, with AUCs of 0.936 (training) and 0.953 (testing) and accuracies of 0.864 and 0.853, respectively. Automated segmentation achieved a dice similarity coefficient (DSC) of 0.915. The integrated deep-learning radiomics framework effectively differentiated breast cancer bone metastases from non-breast-cancer bone tumors.	(+) Combined deep-learning segmentation and radiomics workflow; automated tumor segmentation; high diagnostic accuracy; comparison of multiple ML classifiers; strong segmentation performance; clinically relevant CT-based diagnostic framework. (–) Relatively small retrospective dataset; single-country cohort; no external multicenter validation; CT-only imaging modality; possible overfitting risk for some algorithms; limited generalisability across tumor subtypes and imaging platforms.
[68]	To develop ML models for predicting diagnosis and 5-year survival of breast cancer bone metastasis (BCBM) patients.	Retrospective population-based machine-learning prediction study.	Surveillance, epidemiology, and end results (SEER) database extraction, univariable/multivariable logistic regression, Cox proportional hazards regression, five-fold cross-validation, grid-search hyperparameter tuning, SHAP feature-importance analysis, ROC/AUC, PR curves, calibration curves, and DCA were performed.	United States—SEER database (2010–2016).	283,373 breast cancer patients overall; 3492 BCBM patients included for prognostic analysis. Training cohort: 198,364 breast cancer patients and 2448 BCBM patients; validation cohort: 85,009 breast cancer patients and 1044 BCBM patients.	XGBoost; RF DT; Extra Trees; Gaussian NB; LR.	Feature selection using logistic and Cox regression analyses; five-fold cross-validation; grid-search optimisation; SHAP explainability analysis; ROC and precision-recall analysis; calibration curves; DCA; ensemble ML modelling.	XGBoost achieved the best diagnostic performance with AUCs of 0.987 (training) and 0.940 (validation), and the best prognostic performance with AUCs of 0.880 (training) and 0.800 (validation). Surgery was the most influential predictor for both diagnosis and prognosis. Liver and lung metastases were strongly associated with poorer survival. The XGBoost model outperformed traditional American Joint Committee On Cancer (AJCC) staging in DCA analyses.	(+) Very large national population-based dataset; comparison of multiple ML algorithms; explainable AI using SHAP; separate diagnostic and prognostic modelling; clinically accessible variables; web-based prediction tools developed. (–) Retrospective design; U.S.-only cohort; potential selection bias; lack of laboratory/biochemical variables and Charlson comorbidity index (CCI) data; no prospective validation; limited external international validation.

Table A4. *Cont.*

Reference	Aim of the Study	Research Type	Study Methods	Countries/Setting/Location	Participant Details	AI System/Algorithm/Application Involved	Feature-Processing System or Technique Used	Results	Strengths (+) & Limitations (–)
[43]	To determine whether CT-based radiomics can detect thoracic vertebral bone marrow microenvironment changes before breast cancer bone metastasis and predict future metastasis risk.	Retrospective radiomics and machine-learning prediction study.	Retrospective cohort analysis, thoracic vertebral bone marrow ROI segmentation on non-contrast chest CT, radiomics extraction, Pearson correlation filtering, mRMR feature selection, ML model construction, ROC/AUC analysis, DCA, multivariable logistic regression, and nomogram development were performed.	The First Affiliated Hospital of Zhejiang Chinese Medical University and The Second Affiliated Hospital of Fujian Medical University, China.	106 breast cancer patients followed for 5 years (2013–2023); 37 developed bone metastases. Training cohort: 29 metastatic cases; testing cohort: 8 metastatic cases.	LR, SVM, KNN, RF, ExtraTrees, LightGBM, MLP.	Manual ITK-SNAP segmentation of thoracic vertebral bone marrow (primarily T6 vertebra); extraction of 1197 radiomics features; Pearson correlation filtering (>0.90 removed); minimum redundancy maximum relevance (mRMR) selection of top 20 features; texture analysis including GLCM, GLDM, GLRLM, GLSZM, and NGTDM features; multivariable logistic-regression nomogram integration.	The logistic regression radiomics model achieved the best predictive performance, with AUCs of 0.959 (training) and 0.929 (testing). Radiomics significantly outperformed both the clinical model and the combined nomogram in DCA and ROC analyses. Significant differences in Rad-scores were observed between metastasis and non-metastasis groups ($p < 0.05$). CT radiomics detected subtle vertebral bone marrow microenvironment changes before visible bone metastasis, supporting prediction of 5-year metastasis risk.	(+) novel analysis of pre-metastatic thoracic bone marrow microenvironment; use of routine non-contrast chest CT; comparison of multiple ML algorithms; clinically practical non-invasive workflow; strong predictive performance; DCA demonstrated clinical utility. (–) Small retrospective cohort; exact timing of metastasis unavailable; no external validation; manual ROI segmentation; limited population diversity; possible selection bias; limited generalisability.
[38]	To evaluate the value of contrast-enhanced magnetic resonance imaging (CE-MRI) radiomics combined with machine learning for preoperative prediction of sentinel lymph node metastasis (SLNM) in breast cancer.	Retrospective radiomics and machine-learning prediction study.	Retrospective analysis, CE-MRI ROI segmentation, radiomics extraction, normalisation, feature filtering, LASSO selection, ML classification, ROC/AUC evaluation, DeLong testing, and combined-model construction were performed.	The First Affiliated Hospital of Soochow University, China.	177 breast cancer patients (81 SLN-positive, 96 SLN-negative) examined between January 2015 and May 2021. Training cohort: 123 patients; validation cohort: 54 patients.	SVM, RF, LR, GBDT, DT.	DCE-MRI phase-2 imaging; manual 3D ROI segmentation using the Darwin Scientific Research Platform; extraction of 1316 radiomics features including first-order, shape, GLDM, GLRLM, GLCM, GLSZM, and NGTDM features; min-max normalisation; variance F-value feature filtering; LASSO regression; 10-fold cross-validation.	Thirteen optimal radiomics features were selected. In the validation cohort, SVM achieved the best performance among single ML models, with an AUC of 0.86. The combined model integrating CE-MRI radiomics, tumor size, and Breast Imaging Reporting and Data System (BI-RADS) classification improved predictive performance, achieving a validation AUC of 0.88. DT demonstrated lower validation performance due to overfitting.	(+) Comparison of multiple ML algorithms; clinically relevant non-invasive SLNM prediction; integration of radiomics and clinical variables improved performance; standardised MRI protocol; independent validation cohort; DeLong comparison analysis performed. (–) Retrospective single-center study; relatively small sample size; lack of external multicenter validation; manual segmentation variability; potential overfitting in DT/RF/GBDT models; limited generalisability.
[60]	To develop multimodal radiomics and deep-learning models using mammography (MG) and MRI for preoperative prediction of axillary lymph node metastasis (ALNM) in breast cancer.	Retrospective multimodal AI imaging study.	Clinical and multimodal imaging data were retrospectively analysed using radiomics extraction, 3D deep-learning feature extraction, feature fusion, and ML classification. Multiple radiomics, deep learning, and combined models were evaluated using ROC, calibration, decision curve, DeLong, NRI, and IDI analyses.	Third Hospital of Shanxi Medical University, China.	270 breast cancer patients (2022–2024); training set: 189 patients; test set: 81 patients.	MLP, LR, SVM, KNN, XGBoost, LightGBM, AdaBoost, NB; pretrained 3D ResNet18 deep-learning model.	N4 bias-field correction; grayscale normalisation; voxel resampling ($1 \times 1 \times 1$ mm); manual ROI segmentation using ITK-SNAP; ICC reproducibility testing (>0.9); SMOTE oversampling; Z-score normalisation; univariate filtering; Spearman correlation filtering; mRMR feature selection; LASSO dimensionality reduction; multimodal MG + MRI feature fusion.	The combined-MLP model achieved the best performance with test AUC = 0.846, accuracy = 0.765, F1-score = 0.782, and sensitivity = 0.944, outperforming radiomics-only (AUC = 0.756) and deep-learning-only (AUC = 0.712) models.	(+) Integrated MG and multiparametric MRI with radiomics and deep learning; extensive ML comparison; robust preprocessing and feature-selection pipeline; calibration and clinical-utility analyses. (–) Single-center retrospective study; relatively small sample size; no external validation; only second-phase DCE-MRI features analysed.
[63]	To evaluate CT-based radiomics and ML models for predicting efficacy of anti-HER2 therapy in breast cancer liver metastases (BCLM).	Retrospective CT radiomics ML study.	Contrast-enhanced CT radiomics features from arterial, portal venous, and delayed phases were extracted and analysed using multiple ML classifiers. Models were trained and validated using ROC analysis to predict prognosis following anti-HER2 therapy.	Three hospitals in China.	83 HER2-positive BCLM patients (2011–2021); training cohort: 58 patients; validation cohort: 25 patients.	SVM, KNN, XGBoost, RF, LR, DT.	Manual ROI delineation of largest liver metastasis; grayscale normalisation; extraction of 1409 radiomics features; ICC reproducibility filtering (>0.85); variance threshold filtering; SelectKBest filtering; LASSO dimensionality reduction; multiphase CT texture analysis.	The portal venous phase SVM model showed the best predictive performance, with a validation AUC of 0.865, sensitivity of 0.77, and specificity of 0.83 for poor prognosis prediction. Portal venous phase models consistently outperformed arterial and delayed-phase models.	(+) Evaluated multiphase contrast-enhanced CT; compared multiple ML classifiers; robust multistep feature-selection pipeline; clinically relevant prediction of anti-HER2 treatment response. (–) Small retrospective cohort; manual segmentation subjectivity; no external validation; heterogeneous anti-HER2 regimens and imaging intervals.

Table A4. *Cont.*

Reference	Aim of the Study	Research Type	Study Methods	Countries/Setting/Location	Participant Details	AI System/Algorithm/Application Involved	Feature-Processing System or Technique Used	Results	Strengths (+) & Limitations (−)
[73]	To identify genomic biomarker genes and predict distant metastasis risk in breast cancer using machine-learning models.	Genomic ML classification study.	Gene-expression data from breast cancer patients with and without distant metastasis were analysed using feature selection and ML classification. Models were validated using 10-fold cross-validation and evaluated with accuracy, sensitivity, specificity, PPV, NPV, and F1-score.	Public NCBI GEO genomic dataset.	97 breast cancer patients with 24,481 gene-expression features; 46 patients developed distant metastases and 51 were lymph node-negative without distant metastasis.	XGBoost, NB, MLP.	Logistic Regression-based recursive feature elimination (LR-RFE); high-dimensional genomic feature reduction; candidate biomarker-gene selection; 10-fold cross-validation. Twenty metastasis-associated genes were retained after feature selection.	XGBoost achieved the best predictive performance with accuracy = 0.990, sensitivity = 0.978, specificity = 1.000, PPV = 1.000, NPV = 0.981, and F1-score = 0.989. Top biomarker genes included AL080059, Ubiquilin-1, CA9, PEX12, and CCN4.	(+) High predictive performance; effective dimensionality reduction for high-dimensional genomic data; identified metastasis-associated genes; compared multiple ML classifiers. (−) Small retrospective public dataset; no independent external validation; limited clinical/demographic integration; overfitting risk due to high feature-to-sample ratio.

References

1. World Health Organization. Breast Cancer. 2026. Available online: https://www.who.int/news-room/fact-sheets/detail/breast-cancer (accessed on 22 August 2026).
2. The International Agency for Research on Cancer. Breast Cancer Cases and Deaths Are Projected to Rise Globally. 2025. Available online: https://www.iarc.who.int/wp-content/uploads/2025/02/pr361_E.pdf (accessed on 20 May 2026).
3. Gillies, R.J.; Kinahan, P.E.; Hricak, H. Radiomics: Images are more than pictures, they are data. *Radiology* **2016**, *278*, 563–577. [CrossRef] [PubMed]
4. Miao, S.; An, Y.; Liu, P.; Mu, S.; Zhou, W.; Jia, H.; Huang, W.; Li, J.; Wang, R. Pectoralis muscle predicts distant metastases in breast cancer by deep learning radiomics. *Acta Radiol.* **2023**, *64*, 2561–2569. [CrossRef] [PubMed]
5. Corredor, G.; Bharadwaj, S.; Pathak, T.; Viswanathan, V.S.; Toro, P.; Madabhushi, A. A Review of AI-Based Radiomics and Computational Pathology Approaches in Triple-Negative Breast Cancer: Current Applications and Perspectives. *Clin. Breast Cancer* **2023**, *23*, 800–812. [CrossRef] [PubMed]
6. Topol, E.J. As artificial intelligence goes multimodal, medical applications multiply. *Science* **2023**, *381*, eadk6139. [CrossRef] [PubMed]
7. Hosny, A.; Parmar, C.; Quackenbush, J.; Schwartz, L.H.; Aerts, H.J.W.L. Artificial intelligence in radiology. *Nat. Rev. Cancer* **2018**, *18*, 500–510. [CrossRef] [PubMed]
8. Koçak, B.; Durmaz, E.Ş.; Ateş, E.; Kılıçkesmez, Ö. Radiomics with artificial intelligence: A practical guide for beginners. *Diagn. Interv. Radiol.* **2019**, *25*, 485–495. [CrossRef] [PubMed]
9. Liu, H.; Zou, L.; Xu, N.; Shen, H.; Zhang, Y.; Wan, P.; Wen, B.; Zhang, X.; He, Y.; Gui, L.; et al. Deep learning radiomics based prediction of axillary lymph node metastasis in breast cancer. *npj Breast Cancer* **2024**, *10*, 22. [CrossRef] [PubMed]
10. Vrdoljak, J.; Krešo, A.; Kumrić, M.; Martinović, D.; Cvitković, I.; Grahovac, M.; Vickov, J.; Bukić, J.; Božic, J. The Role of AI in Breast Cancer Lymph Node Classification: A Comprehensive Review. *Cancers* **2023**, *15*, 2400. [CrossRef] [PubMed]
11. Chen, M.; Kong, C.; Lin, G.; Chen, W.; Guo, X.; Chen, Y.; Cheng, X.; Chen, M.; Shi, C.; Xu, M.; et al. Development and validation of convolutional neural network-based model to predict the risk of sentinel or non-sentinel lymph node metastasis in patients with breast cancer: A machine learning study. *eClinicalMedicine* **2023**, *63*, 102176. [CrossRef] [PubMed]
12. Lambin, P.; Leijenaar, R.T.H.; Deist, T.M.; Peerlings, J.; de Jong, E.E.C.; van Timmeren, J.; Sanduleanu, S.; Larue, R.T.H.M.; Even, A.J.G.; Jochems, A.; et al. Radiomics: The bridge between medical imaging and personalized medicine. *Nat. Rev. Clin. Oncol.* **2017**, *14*, 749–762. [CrossRef] [PubMed]
13. World Health Organization. Ethics and Governance of Artificial Intelligence for Health: WHO Guidance. Available online: https://www.who.int/publications/i/item/9789240029200 (accessed on 12 May 2026).
14. Bera, K.; Schalper, K.A.; Rimm, D.L.; Velcheti, V.; Madabhushi, A. Artificial intelligence in digital pathology—New tools for diagnosis and precision oncology. *Nat. Rev. Clin. Oncol.* **2019**, *16*, 703–715. [CrossRef] [PubMed]
15. Pesapane, F.; De Marco, P.; Rapino, A.; Lombardo, E.; Nicosia, L.; Tantrige, P.; Rotili, A.; Bozzini, A.C.; Penco, S.; Dominelli, V.; et al. How Radiomics Can Improve Breast Cancer Diagnosis and Treatment. *J. Clin. Med.* **2023**, *12*, 1372. [CrossRef] [PubMed]
16. Bhalla, K.; Xiao, Q.; Luna, J.M.; Podany, E.; Ahmad, T.; Ademuyiwa, F.O.; Davis, A.; Bennett, D.L.; Gastounioti, A. Radiologic imaging biomarkers in triple-negative breast cancer: A literature review about the role of artificial intelligence and the way forward. *BJR | Artif. Intell.* **2024**, *1*, ubae016. [CrossRef] [PubMed]
17. Champendal, M.; Marmy, L.; Malamateniou, C.; Reis, C.S.D. Artificial intelligence to support person-centred care in breast imaging—A scoping review. *J. Med. Imaging Radiat. Sci.* **2023**, *54*, 511–544. [CrossRef] [PubMed]
18. Bi, W.L.; Hosny, A.; Schabath, M.B.; Giger, M.L.; Birkbak, N.J.; Mehrtash, A.; Allison, T.; Arnaout, O.; Abbosh, C.; Dunn, I.F.; et al. Artificial intelligence in cancer imaging: Clinical challenges and applications. *CA Cancer J. Clin.* **2019**, *69*, 127–157. [CrossRef] [PubMed]
19. Marias, K. The Constantly Evolving Role of Medical Image Processing in Oncology: From Traditional Medical Image Processing to Imaging Biomarkers and Radiomics. *J. Imaging* **2021**, *7*, 124. [CrossRef] [PubMed]
20. Eldaly, A.S.; Avila, F.R.; Torres-Guzman, R.A.; Maita, K.; Garcia, J.P.; Serrano, L.P.; Forte, A.J. Radiomics and artificial intelligence in predicting axillary lymph node metastasis in breast cancer: A systematic review. *Curr. Med. Imaging Rev.* **2023**, *19*, 564–578. [CrossRef] [PubMed]
21. Liu, C.-J.; Zhang, L.; Sun, Y.; Geng, L.; Wang, R.; Shi, K.-M.; Wan, J.-X. Application of CT and MRI images based on an artificial intelligence algorithm for predicting lymph node metastasis in breast cancer patients: A meta-analysis. *BMC Cancer* **2023**, *23*, 1134. [CrossRef] [PubMed]
22. Mota, A.M. AI-Based Characterization of Breast Cancer in Mammography and Tomosynthesis: A Review of Radiomics and Deep Learning for Subtyping, Staging, and Prognosis. *Cancers* **2025**, *17*, 3387. [CrossRef] [PubMed]
23. Cox, J.; Bhatti, A.; Atapour-Abarghouei, A. Using artificial intelligence in the analysis of CT scans of the axillary nodes in breast cancer: A systematic review. *Eur. J. Radiol. Artif. Intell.* **2025**, *4*, 100040. [CrossRef]
24. Attride-Stirling, J. Thematic networks: An analytic tool for qualitative research. *Qual. Res.* **2001**, *1*, 385–405. [CrossRef]

25. Liu, Y.; Li, X.; Zhu, L.; Zhao, Z.; Wang, T.; Zhang, X.; Cai, B.; Li, L.; Ma, M.; Ma, X.; et al. Preoperative Prediction of Axillary Lymph Node Metastasis in Breast Cancer Based on Intratumoral and Peritumoral DCE-MRI Radiomics Nomogram. *Contrast Media Mol. Imaging* **2022**, *2022*, 6729473. [CrossRef] [PubMed]
26. Romeo, V.; Kapetas, P.; Clauser, P.; Rasul, S.; Cuocolo, R.; Caruso, M.; Helbich, T.H.; Baltzer, P.A.T.; Pinker, K. Simultaneous 18F-FDG PET/MRI Radiomics and Machine Learning Analysis of the Primary Breast Tumor for the Preoperative Prediction of Axillary Lymph Node Status in Breast Cancer. *Cancers* **2023**, *15*, 5088. [CrossRef] [PubMed]
27. Wang, S.R.; Cao, C.L.; Du, T.T.; Wang, J.L.; Li, J.; Li, W.X.; Chen, M. Machine Learning Model for Predicting Axillary Lymph Node Metastasis in Clinically Node Positive Breast Cancer Based on Peritumoral Ultrasound Radiomics and SHAP Feature Analysis. *J. Ultrasound Med.* **2024**, *43*, 1611–1625. [CrossRef] [PubMed]
28. Wang, S.R.; Tian, F.; Zhu, T.; Cao, C.L.; Wang, J.L.; Li, W.X.; Li, J.; Hou, J.X. Machine learning-driven ultrasound radiomics for assessing axillary lymph node burden in breast cancer. *Front. Endocrinol.* **2025**, *16*, 1548888. [CrossRef] [PubMed]
29. Yu, Y.; He, Z.; Ouyang, J.; Tan, Y.; Chen, Y.; Gu, Y.; Mao, L.; Ren, W.; Wang, J.; Lin, L.; et al. Magnetic resonance imaging radiomics predicts preoperative axillary lymph node metastasis to support surgical decisions and is associated with tumor microenvironment in invasive breast cancer: A machine learning, multicenter study. *EBioMedicine* **2021**, *69*, 103460. [CrossRef] [PubMed]
30. Zhang, J.; Cao, G.; Pang, H.; Li, J.; Yao, X. Development and validation of radiomics machine learning model based on contrast-enhanced computed tomography to predict axillary lymph node metastasis in breast cancer. *Biomol. Biomed.* **2023**, *23*, 317–326. [CrossRef] [PubMed]
31. Zhang, D.; Shen, M.; Zhang, L.; He, X.; Huang, X. Establishment of an interpretable MRI radiomics-based machine learning model capable of predicting axillary lymph node metastasis in invasive breast cancer. *Sci. Rep.* **2025**, *15*, 26030. [CrossRef] [PubMed]
32. Zhao, X.; Dong, Y.-H.; Xu, L.-Y.; Shen, Y.-Y.; Qin, G.; Zhang, Z.-B. Deep bone oncology Diagnostics: Computed tomography based Machine learning for detection of bone tumors from breast cancer metastasis. *J. Bone Oncol.* **2024**, *48*, 100638. [CrossRef] [PubMed]
33. Chen, D.; Liu, X.; Hu, C.; Hao, R.; Wang, O.; Xiao, Y. Radiomics-based signature of breast cancer on preoperative contrast-enhanced MRI to predict axillary metastasis. *Futur. Oncol.* **2022**, *22*, 553–566. [CrossRef] [PubMed]
34. Li, L.; Tian, H.; Zhang, B.; Wang, W.; Li, B. Prediction for Distant Metastasis of Breast Cancer Using Dynamic Contrast-Enhanced Magnetic Resonance Imaging Images under Deep Learning. *Comput. Intell. Neurosci.* **2022**, *2022*, 6126061. [CrossRef] [PubMed]
35. Luo, H.; Chen, Z.; Xu, H.; Ren, J.; Zhou, P. Peritumoral edema enhances MRI-based deep learning radiomic model for axillary lymph node metastasis burden prediction in breast cancer. *Sci. Rep.* **2024**, *14*, 18900. [CrossRef] [PubMed]
36. Wang, D.; Hu, Y.; Zhan, C.; Zhang, Q.; Wu, Y.; Ai, T. A nomogram based on radiomics signature and deep-learning signature for preoperative prediction of axillary lymph node metastasis in breast cancer. *Front. Oncol.* **2022**, *12*, 940655. [CrossRef] [PubMed]
37. Zhang, Q.; Lou, Y.; Liu, X.; Liu, C.; Ma, W. DCE-MRI-based machine learning model for predicting axillary lymph node metastasis in breast cancer. *Gland Surg.* **2025**, *14*, 228–237. [CrossRef] [PubMed]
38. Zhu, Y.; Yang, L.; Shen, H. Value of the Application of CE-MRI Radiomics and Machine Learning in Preoperative Prediction of Sentinel Lymph Node Metastasis in Breast Cancer. *Front. Oncol.* **2021**, *11*, 757111. [CrossRef] [PubMed]
39. Chen, J.; Qiu, X.; Zheng, Y.; Liao, Y.; Yang, W.; Gu, J.; Yan, C.; Meng, L.; Cheng, J.; Jiang, T.; et al. Non-Invasive Prediction of Axillary Lymph Node Metastasis in Breast Cancer Using Combined Clinical, Radiomics, and Deep Learning Features. *Int. J. Imaging Syst. Technol.* **2025**, *35*, e70248. [CrossRef]
40. Sha, Y.; Ge, S.; Wang, Y.; Cai, S.; Wang, C.; Zhuang, H.; Shi, J.; He, S.; Sun, X.; Ma, L.; et al. Ultrasound-based radiomics combined with B3GALT4 level to predict sentinel lymph node metastasis in primary breast cancer. *Front. Oncol.* **2025**, *15*, 1570493. [CrossRef] [PubMed]
41. Yang, Z.Q.; Zhang, Y.; Lu, F.; Yang, T.; Shan, J.; Jiang, Q.; Lim, G.H.; Bertucci, F.; Du, H.; Zhu, Y.C. Integrating multimodal ultrasound imaging for improved radiomics sentinel lymph node assessment in breast cancer. *Gland Surg.* **2025**, *14*, 1348–1365. [CrossRef] [PubMed]
42. Wang, J.; Gao, X.; Zhang, S.; Zhang, Y. Machine-learning methods based on the texture and non-texture features of MRI for the preoperative prediction of sentinel lymph node metastasis in breast cancer. *Transl. Cancer Res.* **2023**, *12*, 3471–3485. [CrossRef] [PubMed]
43. Zhu, H.N.; Guo, Y.F.; Lin, Y.M.; Sun, Z.C.; Zhu, X.; Li, Y.Z. Radiomics analysis of thoracic vertebral bone marrow microenvironment changes before bone metastasis of breast cancer based on chest CT. *J. Bone Oncol.* **2025**, *50*, 100653. [CrossRef] [PubMed]
44. Chen, Y.; Wang, L.; Dong, X.; Luo, R.; Ge, Y.; Liu, H.; Zhang, Y.; Wang, D. Deep Learning Radiomics of Preoperative Breast MRI for Prediction of Axillary Lymph Node Metastasis in Breast Cancer. *J. Digit. Imaging* **2023**, *36*, 1323–1331. [CrossRef] [PubMed]
45. Fan, J.; Zhang, L.; Lv, T.; Liu, Y.; Sun, H.; Miao, K.; Jiang, C.; Li, L.; Pan, X. MEAI: An artificial intelligence platform for predicting distant and lymph node metastases directly from primary breast cancer. *J. Cancer Res. Clin. Oncol.* **2023**, *149*, 9229–9241. [CrossRef] [PubMed]

46. Gu, C.; He, Y.; Lin, J.; Wang, Z.; Guo, S.; Yang, H.; Wang, W.; Sun, J.; Gan, H.; Li, H. Preoperative prediction of axillary lymph node metastasis in breast invasive ductal carcinoma patients using a deep learning model based on dynamic contrast-enhanced magnetic resonance imaging: A multicenter study. *Gland Surg.* **2025**, *14*, 2288–2301. [CrossRef] [PubMed]
47. Oshino, T.; Enda, K.; Shimizu, H.; Sato, M.; Nishida, M.; Kato, F.; Oda, Y.; Hosoda, M.; Kudo, K.; Iwasaki, N.; et al. Artificial intelligence can extract important features for diagnosing axillary lymph node metastasis in early breast cancer using contrast-enhanced ultrasonography. *Sci. Rep.* **2025**, *15*, 5648. [CrossRef] [PubMed]
48. Park, D.; Lee, Y.M.; Eo, T.; An, H.J.; Kang, H.; Park, E.; Cha, Y.J.; Park, H.; Kwon, D.; Kwon, S.Y.; et al. Multimodal AI model for preoperative prediction of axillary lymph node metastasis in breast cancer using whole slide images. *npj Precis. Oncol.* **2025**, *9*, 131. [CrossRef] [PubMed]
49. Yao, Y.; Lv, Y.; Tong, L.; Liang, Y.; Xi, S.; Ji, B.; Zhang, G.; Li, L.; Tian, G.; Tang, M.; et al. ICSDA: A multi-modal deep learning model to predict breast cancer recurrence and metastasis risk by integrating pathological, clinical and gene expression data. *Brief. Bioinform.* **2022**, *23*, bbac448. [CrossRef] [PubMed]
50. Jarkman, S.; Karlberg, M.; Pocevičiūtė, M.; Bodén, A.; Bándi, P.; Litjens, G.; Lundström, C.; Treanor, D.; van der Laak, J. Generalization of Deep Learning in Digital Pathology: Experience in Breast Cancer Metastasis Detection. *Cancers* **2022**, *14*, 5424. [CrossRef] [PubMed]
51. Wang, L.; Qu, F.; Wen, P.; Luo, Y.; Zhang, H.; Li, S.; Yin, X.; Zhao, Y.; Zeng, X. Development of a Machine Learning Model Integrating Pathomics and Clinical Data to Predict Axillary Lymph Node Metastasis in Breast Cancer: A Two-Center Study. *Cancer Rep.* **2025**, *8*, e70302. [CrossRef] [PubMed]
52. Yu, J.G.; Wu, Z.; Ming, Y.; Deng, S.; Li, Y.; Ou, C.; He, C.; Wang, B.; Zhang, P.; Wang, Y. Prototypical multiple instance learning for predicting lymph node metastasis of breast cancer from whole-slide pathological images. *Med. Image Anal.* **2023**, *85*, 102748. [CrossRef] [PubMed]
53. Chen, H.; Lan, X.; Yu, T.; Li, L.; Tang, S.; Liu, S.; Jiang, F.; Wang, L.; Huang, Y.; Cao, Y.; et al. Development and validation of a radiogenomics model to predict axillary lymph node metastasis in breast cancer integrating MRI with transcriptome data: A multicohort study. *Front. Oncol.* **2022**, *12*, 1076267. [CrossRef] [PubMed]
54. Qin, J.N.; Dai, W.B.; Zhang, W.H.; Chen, B.J.; Liang, L.; Liang, C.F.; Lu, C.G.; Tan, Q.X.; Wei, C.Y.; Tan, Y.; et al. Identification of optimal biomarkers associated with distant metastasis in breast cancer using Boruta and Lasso machine learning algorithms. *BMC Cancer* **2025**, *25*, 1311. [CrossRef] [PubMed]
55. Tan, Y.; Zhang, W.H.; Huang, Z.; Tan, Q.X.; Zhang, Y.M.; Wei, C.Y.; Feng, Z.B. AI models predicting breast cancer distant metastasis using LightGBM with clinical blood markers and ultrasound maximum diameter. *Sci. Rep.* **2024**, *14*, 15561. [CrossRef] [PubMed]
56. Yagin, B.; Yagin, F.H.; Colak, C.; Inceoglu, F.; Kadry, S.; Kim, J. Cancer Metastasis Prediction and Genomic Biomarker Identification through Machine Learning and eXplainable Artificial Intelligence in Breast Cancer Research. *Diagnostics* **2023**, *13*, 3314. [CrossRef] [PubMed]
57. Zhang, W.H.; Tan, Y.; Huang, Z.; Tan, Q.X.; Zhang, Y.M.; Wei, C.Y. Development and validation of an artificial intelligence model for predicting de novo distant bone metastasis in breast cancer: A dual-center study. *BMC Women's Health* **2024**, *24*, 442. [CrossRef] [PubMed]
58. Zhang, W.H.; Tan, Y.; Huang, Z.; Tan, Q.X.; Zhang, Y.M.; Chen, B.J.; Wei, C.Y. Development and validation of AI models using LR and LightGBM for predicting distant metastasis in breast cancer: A dual-center study. *Front. Oncol.* **2024**, *14*, 1409273. [CrossRef] [PubMed]
59. Van Gómez, O.; Herraiz, J.L.; Udías, J.M.; Haug, A.; Papp, L.; Cioni, D.; Neri, E. Analysis of Cross-Combinations of Feature Selection and Machine-Learning Classification Methods Based on [18F]F-FDG PET/CT Radiomic Features for Metabolic Response Prediction of Metastatic Breast Cancer Lesions. *Cancers* **2022**, *14*, 2922. [CrossRef] [PubMed]
60. Guo, F.; Sun, S.; Deng, X.; Wang, Y.; Yao, W.; Yue, P.; Wu, S.; Yan, J.; Zhang, X.; Zhang, Y. Predicting axillary lymph node metastasis in breast cancer using a multimodal radiomics and deep learning model. *Front. Immunol.* **2024**, *15*, 1482020. [CrossRef] [PubMed]
61. Agraz, M. Comparison of Feature Selection Methods in Breast Cancer Microarray Data. *Med. Rec.* **2023**, *5*, 284–289. [CrossRef]
62. Cheng, Y.; Wu, M.; Tang, W.; Zhang, X.; Zhou, L.; Huang, Z.; Tang, X.; Pei, L.; Li, Z.; Lam, S.; et al. Prediction of axillary lymph node metastasis in breast cancer based on MRI: A novel domain adaptative radiomics pipeline for multicenter studies. *Med. Phys.* **2025**, *52*, e70122. [CrossRef] [PubMed]
63. He, M.; Hu, Y.; Wang, D.; Sun, M.; Li, H.; Yan, P.; Meng, Y.; Zhang, R.; Li, L.; Yu, D.; et al. Value of CT-Based Radiomics in Predicating the Efficacy of Anti-HER2 Therapy for Patients with Liver Metastases From Breast Cancer. *Front. Oncol.* **2022**, *12*, 852809. [CrossRef] [PubMed]
64. Lai, J.; Chen, Z.; Liu, J.; Zhu, C.; Huang, H.; Yi, Y.; Cai, G.; Liao, N. A radiogenomic multimodal and whole-transcriptome sequencing for preoperative prediction of axillary lymph node metastasis and drug therapeutic response in breast cancer: A retrospective, machine learning and international multicohort study. *Int. J. Surg.* **2024**, *110*, 2162–2177. [CrossRef] [PubMed]
65. Shen, Y.; Huang, R.; Zhang, Y.; Zhu, J.; Li, Y. Prediction of axillary lymph node metastasis in triple negative breast cancer using MRI radiomics and clinical features. *Sci. Rep.* **2025**, *15*, 21923. [CrossRef] [PubMed]

66. Koh, J.; Jeong, D.; Park, S.Y.; Han, D.; Kim, D.S.; Kim, H.Y.; Kim, H.; Yang, S.; Kim, S.; Ryu, H.S. Identification of VWA5A as a novel biomarker for inhibiting metastasis in breast cancer by machine-learning based protein prioritization. *Sci. Rep.* **2024**, *14*, 2459. [CrossRef] [PubMed]
67. Ye, H.; Shen, X.; Li, Y.; Zou, W.; Hassan, S.S.U.; Feng, Y.; Wang, X.; Tian, J.; Shao, X.; Tao, Y.; et al. Proteomic and metabolomic characterization of bone, liver, and lung metastases in plasma of breast cancer patients. *Proteom. Clin. Appl.* **2024**, *18*, 2300136. [CrossRef] [PubMed]
68. Zhong, X.; Lin, Y.; Zhang, W.; Bi, Q. Predicting diagnosis and survival of bone metastasis in breast cancer using machine learning. *Sci. Rep.* **2023**, *13*, 18301. [CrossRef] [PubMed]
69. Maouche, I.; Terrissa, L.S.; Benmohammed, K.; Zerhouni, N. An Explainable AI Approach for Breast Cancer Metastasis Prediction Based on Clinicopathological Data. *IEEE Trans. Biomed. Eng.* **2023**, *70*, 3321–3329. [CrossRef] [PubMed]
70. Cai, R.; Deng, L.; Zhang, H.; Zhang, H.; Wu, Q. A non-invasive preoperative prediction model for predicting axillary lymph node metastasis in breast cancer based on a machine learning approach: Combining ultrasonographic parameters and breast gamma specific imaging features. *Radiat. Oncol.* **2024**, *19*, 63. [CrossRef] [PubMed]
71. Yang, W.L.; Su, X.R.; Li, S.; Zhao, K.Y.; Yue, Q. Utilizing machine-learning techniques on MRI radiomics to identify primary tumors in brain metastases. *Front. Neurol.* **2025**, *15*, 1474461. [CrossRef] [PubMed]
72. Jiang, X.; Xu, C. Deep Learning and Machine Learning with Grid Search to Predict Later Occurrence of Breast Cancer Metastasis Using Clinical Data. *J. Clin. Med.* **2022**, *11*, 5772. [CrossRef] [PubMed]
73. Inceoğlu, F.; Yağin, F.H. Genomic Biomarkers of Metastasis in Breast Cancer Patients: A Machine Learning Approach. *J. Cogn. Syst.* **2022**, *7*, 29–32. [CrossRef]
74. Botlagunta, M.; Botlagunta, M.D.; Myneni, M.B.; Lakshmi, D.; Nayyar, A.; Gullapalli, J.S.; Shah, M.A. Classification and diagnostic prediction of breast cancer metastasis on clinical data using machine learning algorithms. *Sci. Rep.* **2023**, *13*, 485. [CrossRef] [PubMed]
75. Gong, X.; Guo, Y.; Zhu, T.; Peng, X. Diagnostic performance of radiomics in predicting axillary lymph node metastasis in breast cancer: A systematic review and meta-analysis. *Front. Oncol.* **2022**, *12*, 1046005. [CrossRef] [PubMed]
76. Huang, S.; Franc, B.L.; Harnish, R.J.; Liu, G.; Mitra, D.; Copeland, T.P.; Arasu, V.A.; Kornak, J.; Jones, E.F.; Behr, S.C.; et al. Exploration of PET and MRI radiomic features for decoding breast cancer phenotypes and prognosis. *npj Breast Cancer* **2018**, *4*, 24. [CrossRef] [PubMed]
77. Pfaehler, E.; Zhovannik, I.; Wei, L.; Boellaard, R.; Dekker, A.; Monshouwer, R.; El Naqa, I.; Bussink, J.; Gillies, R.; Wee, L.; et al. A systematic review and quality of reporting checklist for repeatability and reproducibility of radiomic features. *Phys. Imaging Radiat. Oncol.* **2021**, *20*, 69–75. [CrossRef] [PubMed]
78. Trojani, V.; Bassi, M.C.; Verzellesi, L.; Bertolini, M. Impact of Preprocessing Parameters in Medical Imaging-Based Radiomic Studies: A Systematic Review. *Cancers* **2024**, *16*, 2668. [CrossRef] [PubMed]
79. Alberich-Bayarri, A.; Sourbron, S.; Golay, X.; deSouza, N.; Smits, M.; van der Lugt, A.; Boellard, R. ESR Statement on the Validation of Imaging Biomarkers. *Insights Imaging* **2020**, *11*, 76. [CrossRef] [PubMed]
80. Esteva, A.; Chou, K.; Yeung, S.; Naik, N.; Madani, A.; Mottaghi, A.; Liu, Y.; Topol, E.; Dean, J.; Socher, R. Deep learning-enabled medical computer vision. *npj Digit. Med.* **2021**, *4*, 5. [CrossRef] [PubMed]
81. Pascuzzo, R.; Garattini, S.K.; Doniselli, F.M. Clinical Application of Radiomics in Oncology: Where Do We Stand? *J. Magn. Reson. Imaging* **2024**, *60*, 2745–2746. [CrossRef] [PubMed]
82. Schiano, C.; Franzese, M.; Pane, K.; Garbino, N.; Soricelli, A.; Salvatore, M.; de Nigris, F.; Napoli, C. Hybrid 18F-FDG-PET/MRI Measurement of Standardized Uptake Value Coupled with Yin Yang 1 Signature in Metastatic Breast Cancer. A Preliminary Study. *Cancers* **2019**, *11*, 1444. [CrossRef] [PubMed]
83. Ming, Y.; Wu, N.; Qian, T.; Li, X.; Wan, D.Q.; Li, C.; Li, Y.; Wu, Z.; Wang, X.; Liu, J.; et al. Progress and Future Trends in PET/CT and PET/MRI Molecular Imaging Approaches for Breast Cancer. *Front. Oncol.* **2020**, *10*, 507927. [CrossRef] [PubMed]
84. Tjoa, E.; Guan, C. A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *32*, 4793–4813. [CrossRef] [PubMed]
85. Arun, N.; Gaw, N.; Singh, P.; Chang, K.; Aggarwal, M.; Chen, B.; Hoebel, K.; Gupta, S.; Patel, J.; Gidwani, M.; et al. Assessing the trustworthiness of saliency maps for localizing abnormalities in medical imaging. *Radiol. Artif. Intell.* **2021**, *3*, e200267. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.