

Review

Application of Artificial Intelligence Algorithms in the Comprehensive Care of Patients with Breast Cancer

Dorota Bartusik-Aebisher ¹, Sara Czech ², Jakub Szpara ², Avijit Paul ³, Marvin Xavierselvan ³ and David Aebisher ^{4,*}

¹ Department of Biochemistry and General Chemistry, Medical Faculty, Collegium Medicum, University of Rzeszów, 35-310 Rzeszów, Poland; dbartusikaebisher@ur.edu.pl

² English Division Science Club, Collegium Medicum, Faculty of Medicine, Rzeszów University, 35-310 Rzeszów, Poland; sc126240@stud.ur.edu.pl (S.C.); js126214@stud.ur.edu.pl (J.S.)

³ Department of Biomedical Engineering, Tufts University, Medford, MA 02155, USA; avijit.paul@tufts.edu (A.P.); marvin.xavierselvan@tufts.edu (M.X.)

⁴ Department of Photomedicine and Physical Chemistry, Medical Faculty, Collegium Medicum, University of Rzeszów, 35-310 Rzeszów, Poland

* Correspondence: daebisher@ur.edu.pl

Abstract

Breast cancer remains one of the most significant challenges in modern oncology, while advances in artificial intelligence (AI) are creating new opportunities to improve diagnosis, prognosis, and treatment personalization. The aim of this review was to summarize current and emerging applications of AI in the comprehensive care of patients with breast cancer. This study was conducted as a structured narrative review with elements of integrative evidence synthesis based on publications retrieved from PubMed/MEDLINE, Scopus, Web of Science, and Embase. The review included studies evaluating machine learning and deep learning approaches, such as support vector machines, random forests, convolutional neural networks, Vision Transformers, foundation models, self-supervised learning, federated learning, and multimodal AI systems. The strongest clinical evidence currently concerns AI-supported mammographic screening, where large prospective and real-world studies suggest improvements in cancer detection and workflow efficiency. Applications involving MRI, ultrasound, histopathology, molecular prediction, treatment-response assessment, and treatment selection have shown promising performance, but most remain investigational because of limited prospective multicenter validation. Emerging approaches integrating imaging, pathological, molecular, and clinical data show considerable potential for precision oncology. AI may also support treatment selection, patient monitoring, and survivorship care. Despite promising results, widespread clinical implementation remains limited by challenges related to data heterogeneity, model interpretability, external validation, and integration into clinical workflows. Further prospective multicenter studies are required to establish the safety, reliability, and clinical utility of AI-driven systems in breast cancer care.



Academic Editor: Francesc Pozo

Received: 26 May 2026

Revised: 20 June 2026

Accepted: 25 June 2026

Published: 29 June 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.

Keywords: breast cancer; artificial intelligence; machine learning; deep learning; mammography; magnetic resonance imaging; ultrasound; radiomics; Vision Transformers; foundation models; precision oncology

1. Introduction

1.1. Epidemiology

Breast cancer is an extremely heterogeneous disease, characterized by a variety of subtypes and distinct epidemiological patterns [1]. The incidence of breast cancer worldwide is shaped by the complex interplay of genetic, environmental, and lifestyle factors. In high-income countries, a higher number of new cases is generally observed compared to low- and middle-income countries. At the same time, mortality in these countries is often lower, mainly due to better access to early detection methods and more advanced treatment options [2]. Globally, breast cancer accounts for approximately one-third of all malignant neoplasms diagnosed in women. Moreover, deaths caused by this disease constitute nearly 15% of the total number of diagnosed cases [3].

Breast cancer carcinogenesis is a multistep process involving the accumulation of genetic alterations and the influence of environmental factors, in which normal cells progressively transition through stages of hyperplasia, premalignant lesions, carcinoma in situ, and invasive carcinoma. Key roles are played by mutations and changes in the expression of genes such as NF1, ESR1, ALDH2, GATA3, KMT2C, PTEN, FOXM1, YTHDF3, TP53, PIK3CA, and RB1, which affect signaling pathways associated with cell proliferation, metastasis, and immune evasion. The most important risk factors include genetic predispositions (e.g., mutations in BRCA1/BRCA2, polymorphisms in GSTM1 and NQO2); hormonal factors such as prolonged exposure to estrogens, early menarche, late menopause, nulliparity, and lack of breastfeeding; and lifestyle and environmental factors, including exposure to radiation, excessive alcohol consumption, smoking, a diet rich in fats and sugars, obesity, and low physical activity [4]. Below, Figure 1 presents a graphical illustration of factors that may contribute to the development of breast cancer.

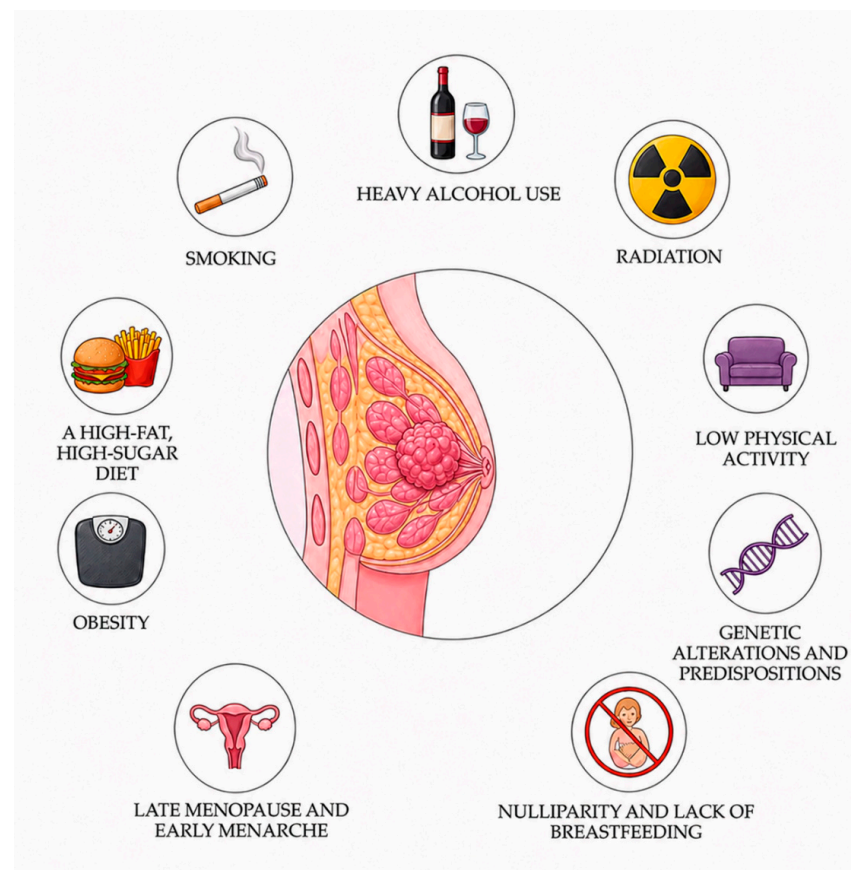


Figure 1. Graphical illustration of factors that may contribute to the development of breast cancer.

1.2. The Role of Early Diagnosis in Breast Cancer

Early diagnosis of breast cancer plays a key role in improving patient prognosis, as the disease often develops without clear symptoms in its early stages, which contributes to late detection and hampers effective treatment [5]. Studies indicate that screening mammography enables detection of cancer at an earlier stage and is associated with a reduction in breast cancer mortality. A cohort study demonstrated that earlier participation in screening was associated with a lower likelihood of diagnosis at an advanced stage and a reduced risk of death [6]. At the same time, the development and optimization of screening programs—including improvements in mammogram analysis, the use of computer-aided detection systems, and additional imaging modalities—may increase the number of cases detected at an early stage and enhance the effectiveness of population-level prevention [5], although research is still ongoing regarding the optimal age to begin screening, its frequency, and the most effective diagnostic strategies [7].

1.3. The Role of AI in Medicine

Artificial intelligence (AI) is playing an increasingly important role in the development of modern medicine, serving as a significant tool supporting the processes of diagnosis, treatment, and medical data management. The rapid advancement of machine learning and deep learning algorithms enables the analysis of large datasets, including clinical data, diagnostic images, and information derived from electronic health records, allowing for faster and more precise identification of patterns associated with disease development [8]. The application of AI contributes to improved diagnostic accuracy and supports physicians in clinical decision-making, while simultaneously reducing the risk of diagnostic errors and the workload of medical personnel [9].

AI is particularly widely applied in medical imaging, such as radiology, pathology, and oncological diagnostics, where AI algorithms can analyze medical images to detect abnormalities that may be invisible or difficult for humans to identify [10]. AI-based technologies are also used across various medical specialties, supporting, among others, disease diagnosis, treatment planning, and patient data analysis to improve the quality of healthcare [11]. Furthermore, the integration of artificial intelligence with healthcare systems may enhance the efficiency of healthcare delivery, streamline diagnostic processes, and foster the development of more personalized medicine, although its widespread implementation requires further research and careful consideration of ethical and data security issues [12].

1.4. Objective of the Review

The main objective of the review is to review the literature and demonstrate the application of AI in the context of breast cancer. The main aspects addressed include AI algorithms in breast cancer imaging, AI in treatment selection, AI in prognosis and prediction of response to breast cancer treatment.

2. Materials and Methods

2.1. Study Design

This study was designed as a structured narrative review with elements of integrative evidence synthesis concerning the application of artificial intelligence in the comprehensive care of patients with breast cancer. The review focused on AI applications in imaging diagnostics, histopathological analysis, prognosis prediction, treatment response assessment, treatment selection, and emerging precision oncology approaches. Particular emphasis was placed on the integration of clinical, imaging, pathological, and molecular data using modern AI methodologies. Although this was not a conventional systematic review or

meta-analysis, the PRISMA 2020 framework was used to transparently document study identification, screening, eligibility assessment and inclusion.

All databases were searched from their inception to 1 June 2026, without restrictions on publication date. Although no formal time limit was applied, greater emphasis in the narrative synthesis was placed on recent studies, particularly large prospective, multicenter, externally validated, and real-world investigations.

Unlike conventional systematic reviews focused on a single intervention or technology, this review aimed to provide a broad and clinically oriented overview of both established and emerging AI paradigms. The analysis included classical machine learning methods, deep learning architectures, transformer-based models, foundation models, self-supervised learning, multimodal learning frameworks, federated learning, explainable AI, causal AI, and other novel approaches with potential relevance to breast cancer management.

2.2. Review Question and Conceptual Scope

The primary research question was whether artificial intelligence can improve diagnosis, risk stratification, prediction of treatment response, treatment selection, and long-term management of patients with breast cancer.

The conceptual scope encompassed five major domains:

- AI applications in breast imaging, including mammography, MRI, ultrasound, and multimodal imaging;
- AI-assisted histopathology and computational pathology;
- Prognostic modeling and prediction of treatment response;
- AI-guided treatment selection and precision oncology;
- Emerging AI paradigms and technical challenges, including foundation models, self-supervised learning, multimodal AI, federated learning, explainability, uncertainty quantification, causal AI, and continual learning.

The specific objectives were:

- To summarize current applications of AI in breast cancer care;
- To compare the performance of AI-based approaches with conventional diagnostic and prognostic methods;
- To evaluate the clinical utility of emerging AI paradigms;
- To identify technical, ethical, and translational barriers to implementation;
- To discuss future directions for AI-driven precision oncology.

2.3. Information Sources

The literature review was conducted using PubMed/MEDLINE, Scopus, Web of Science, and Embase. These databases were selected because of their extensive coverage of oncology, radiology, pathology, biomedical informatics, and artificial intelligence research.

All databases were searched from their inception to 1 June 2026, without restrictions on publication date. Targeted supplementary searches and citation checking were conducted during manuscript revision until 19 June 2026 to identify relevant methodological references and recently published evidence. Although no formal time limit was applied, greater emphasis in the narrative synthesis was placed on recent studies, particularly large prospective, multicenter, externally validated, and real-world investigations.

To ensure comprehensive coverage of emerging technologies, additional manual screening of references from relevant reviews, landmark studies, and methodological publications was performed. Forward citation tracking was also used to identify recently published studies involving novel AI architectures and translational applications.

2.4. Search Strategy

The search strategy was designed to identify studies evaluating the use of artificial intelligence in breast cancer diagnosis, prognosis, treatment planning, and clinical decision support.

Keywords included combinations of:

“breast cancer”, “artificial intelligence”, “machine learning”, “deep learning”, “convolutional neural networks”, “CNN”, “Vision Transformer”, “foundation model”, “self-supervised learning”, “federated learning”, “multimodal learning”, “radiomics”, “radiogenomics”, “mammography”, “MRI”, “ultrasound”, “histopathology”, “treatment response”, “prognosis”, “precision medicine”, “precision oncology”, “explainable AI”, and “causal AI”.

Additional search terms were included to capture emerging AI paradigms and next-generation computational approaches. These terms included: “Vision Transformer”, “transformer”, “foundation model”, “self-supervised learning”, “federated learning”, “multimodal learning”, “synthetic data”, “diffusion model”, “explainable AI”, “uncertainty-aware AI”, “causal AI”, “continual learning”, “computational pathology”, and “precision oncology”.

An example PubMed search strategy was: ((“breast cancer”) AND (“artificial intelligence” OR “machine learning” OR “deep learning” OR “Vision Transformer” OR “foundation model”) AND (“imaging” OR “mammography” OR “MRI” OR “ultrasound” OR “histopathology” OR “radiomics”) AND (“prediction” OR “prognosis” OR “treatment response” OR “precision oncology”)). The search strategy was intentionally broad to capture both clinical and methodological studies, including investigations of emerging AI paradigms. The complete database-specific search strategies are provided in Supplementary Table S1. Search syntax was adapted to the indexing rules and controlled vocabulary of each database.

2.5. Eligibility Criteria

Studies included in the review met the following criteria:

Inclusion criteria:

- Original clinical studies, translational studies, and selected high-quality review articles;
- Studies concerning breast cancer;
- Publications analyzing the use of AI in diagnosis, prognosis, or treatment;
- Articles published in peer-reviewed scientific journals;
- Publications in English.

Exclusion criteria:

- Studies conducted exclusively on animal models or in vitro (unless mechanistically relevant);
- Conference abstracts without full text;
- Case reports, commentaries, and opinion articles without primary data;
- Publications not directly related to breast cancer or AI.

2.6. Study Selection

All retrieved records were compiled in a single reference library. Duplicate records were identified by matching DOI or PMID and, when these identifiers were unavailable, by comparing the title, first author, publication year and journal. Records with uncertain matches were checked manually before screening. Study selection was performed by three reviewers in two stages. First, titles and abstracts were screened for relevance to breast cancer and artificial intelligence. Second, the full texts of potentially eligible publications were assessed according to the predefined inclusion and exclusion criteria. Disagreements

regarding study eligibility were resolved through discussion among the three reviewers until consensus was reached. The review protocol was not prospectively registered.

Preference was given to studies:

- Involving well-characterized patient cohorts;
- Using advanced AI methods;
- Including model validation;
- Relating findings to clinical practice.

The study selection process is presented in Figure 2 according to the PRISMA 2020 framework.

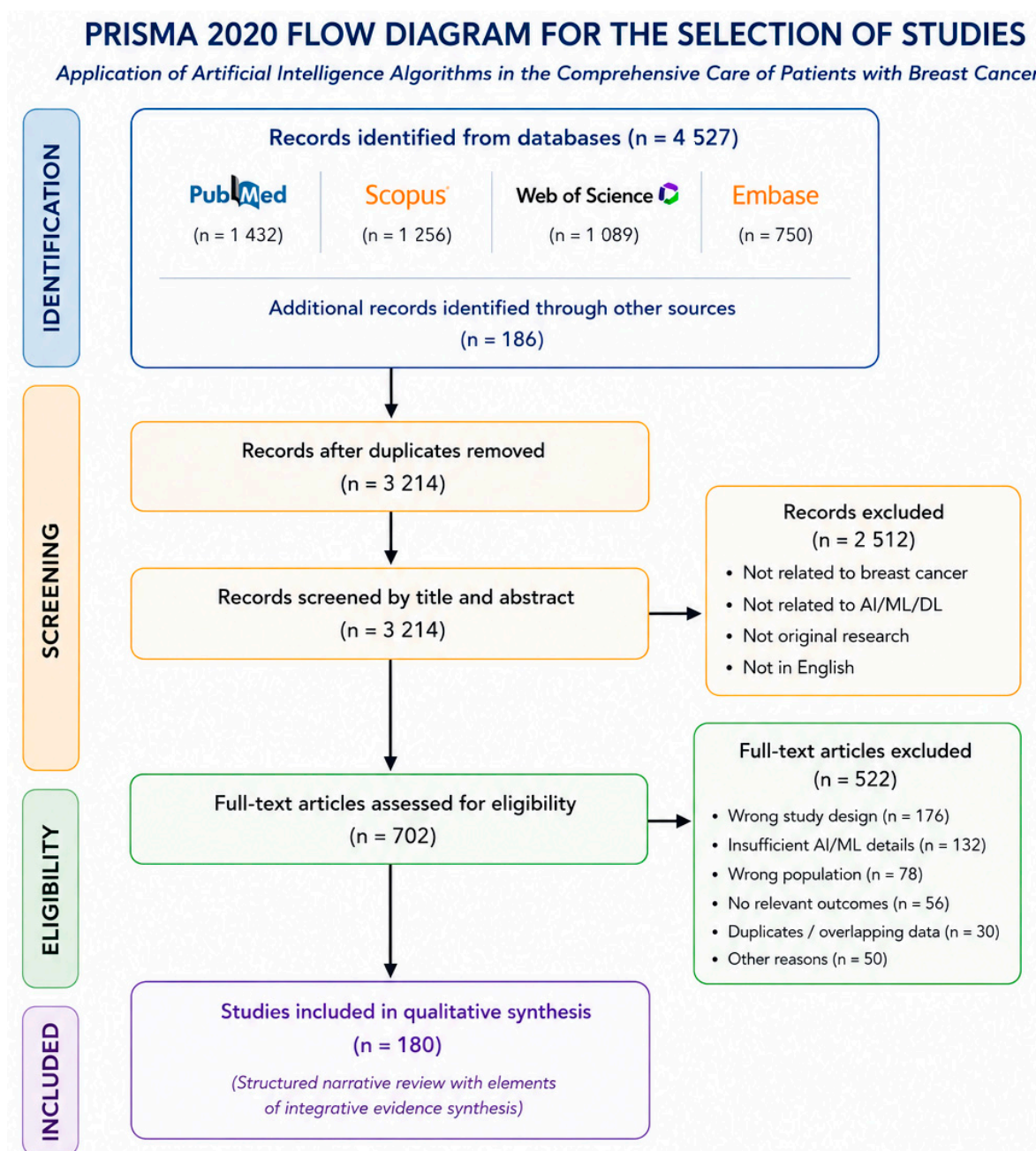


Figure 2. PRISMA 2020 flow diagram illustrating the study selection process.

2.7. Data Extraction

From each included study, the following information was extracted: publication year, study type, sample size, population characteristics, AI method used, type of data (e.g.,

imaging, clinical), main outcomes and clinical relevance. For studies involving advanced artificial intelligence methodologies, additional information was extracted regarding model architecture, validation strategy, interpretability methods, use of multimodal data integration, privacy-preserving approaches, and reported measures of model robustness and generalizability.

Particular attention was paid to:

- Type of AI model (ML vs. DL);
- Type of data used;
- Performance metrics (e.g., AUC, sensitivity, and specificity);
- Potential clinical application;
- Model interpretability and explainability techniques;
- Multimodal data integration strategies;
- Use of external or multicenter validation;
- Privacy-preserving approaches such as federated learning;
- Reported limitations regarding generalizability and domain shift.

Performance metrics were extracted and reported as presented in the original studies. AUC and AUROC values were standardized to three decimal places whenever sufficient numerical precision was available. Proportions were reported consistently as percentages, whereas absolute and relative changes were clearly distinguished. Confidence intervals and *p*-values were included when reported in the original publications; they were not independently estimated when the required data were unavailable.

2.8. Methodological Quality Appraisal

Because this review included highly heterogeneous evidence, including diagnostic-accuracy studies, prognostic models, randomized and observational studies, technical model-development studies, and translational investigations, no single formal risk-of-bias instrument could be applied consistently across all included publications. Therefore, no numerical quality score or formal study-level risk-of-bias classification was performed. Instead, a structured qualitative methodological appraisal was incorporated into data extraction and evidence synthesis. For each primary study, the reviewers considered study design, cohort size and class distribution, retrospective or prospective data collection, single-center or multicenter setting, internal or external validation, separation of training and test cohorts, reporting of uncertainty and calibration, clinical relevance of the evaluated endpoints, and potential generalizability. Studies based on small, highly selected, single-center cohorts or lacking external or prospective validation were interpreted more cautiously than large multicenter, prospective, externally validated, or real-world investigations.

Prediction-model development and validation studies were assessed using PROBAST+AI, randomized trials using the Cochrane RoB 2 tool, and non-randomized comparative or real-world implementation studies using ROBINS-I. The assessment was conducted independently by two reviewers, with disagreements resolved through discussion with a third reviewer. The resulting domain-level and overall assessments are presented in Supplementary Table S2. When reporting was insufficient to permit a definitive domain-level judgment, the corresponding risk of bias was classified as unclear rather than inferred from unreported methodological information.

2.9. Data Synthesis

Data synthesis was conducted in a narrative and thematic manner. The results were grouped according to the main areas of AI application, such as imaging diagnostics, prognosis prediction and treatment selection.

The analysis aimed to identify recurring patterns, evaluate the effectiveness of AI methods, and determine their potential clinical value. Particular emphasis was placed on integrating findings across different domains, evaluating emerging AI paradigms, and assessing their relevance to the development of personalized and precision oncology approaches in breast cancer care.

Because the included studies differed substantially in terms of clinical task, imaging modality, patient population, dataset size, outcome definition, performance metrics, and validation strategy, statistical pooling of the reported results was not considered methodologically appropriate. Therefore, no formal meta-analysis was performed, and the findings were synthesized descriptively. To facilitate comparison across studies, representative clinical studies were additionally summarized according to sample size, AI application, reported performance metrics and validation design.

For the purposes of descriptive comparison within this review, we adopted the following approximate interpretative framework: an AUC of 0.500 was considered to represent chance-level discrimination. AUC values greater than 0.500 but below 0.600 were interpreted as indicating very weak discrimination, values from 0.600 to below 0.700 as poor discrimination, values from 0.700 to below 0.800 as moderate discrimination, values from 0.800 to below 0.900 as good discrimination, and values of 0.900 or higher as excellent discrimination. These categories are descriptive rather than universal thresholds for clinical usefulness. The clinical value of a model also depends on the target population, disease prevalence, decision threshold, sensitivity, specificity, calibration, and consequences of false-positive and false-negative predictions.

3. Results

3.1. Review of Selected Algorithms

The following subsection provides only the methodological background necessary to interpret the breast cancer studies included in this review. Breast cancer-specific datasets, validation strategies, clinical endpoints, performance results, and study limitations are discussed in Sections 3.2–3.5.

3.1.1. Machine Learning

Support Vector Machines (SVM) are supervised learning models that identify a decision boundary maximizing the margin between classes. Kernel functions allow nonlinear relationships to be modeled in high-dimensional datasets. Their advantages include effectiveness in relatively small, high-dimensional datasets, whereas important limitations include sensitivity to hyperparameter selection and limited interpretability. Mathematically, the binary SVM classifier assigns a new observation x to a class using the decision function:

$$f(x) = \text{sign}\left(\sum_{i=1}^n \alpha_i y_i K(x_i, x) + b\right), \quad (1)$$

where x_i denotes a training observation, y_i is its class label, α_i represents the learned coefficient associated with the corresponding support vector, $K(x_i, x)$ is the kernel function, and b is the intercept. The kernel function enables the classifier to construct nonlinear decision boundaries without explicitly transforming the observations into a higher-dimensional feature space [13–15].

Random Forest combines multiple decision trees trained using bootstrap samples and random feature selection. This approach can model nonlinear relationships and interactions while providing measures of variable importance. However, importance estimates may be biased, and the resulting ensemble is less directly interpretable than conventional statistical

models [16,17]. For regression tasks, the Random Forest prediction is obtained by averaging the outputs of B individual decision trees:

$$\hat{y}(x) = \frac{1}{B} \sum_{b=1}^B T_b(x) \quad (2)$$

where $T_b(x)$ denotes the prediction generated by the b -th decision tree. For classification tasks, the final class is typically determined by majority voting:

$$\hat{y}(x) = \text{mode}\{T_1(x), T_2(x), \dots, T_B(x)\} \quad (3)$$

where $T_b(x)$ represents the class predicted by the b -th tree [16,17].

The k-Nearest Neighbors method is a nonparametric algorithm that assigns an observation to a class according to the labels of its nearest neighbors. Its performance depends strongly on the selected distance metric, the value of k , feature scaling, and dataset dimensionality [18]. In a classification task, kNN assigns an observation x to the class most frequently represented among its k nearest neighbors:

$$\hat{y}(x) = \arg \max_c \sum_{i \in N_k(x)} I(y_i = c) \quad (4)$$

where $N_k(x)$ denotes the set of the k nearest observations, c represents a possible class, and $I(\cdot)$ is the indicator function. In regression tasks, the prediction is usually calculated as the mean or distance-weighted mean of the outcomes of the nearest neighbors [18].

Logistic regression is an interpretable statistical model for binary outcomes that estimates the relationship between predictor variables and the log-odds of an event. It enables calculation of adjusted odds ratios but assumes a correctly specified model and linear relationships between predictors and the logit of the outcome [19,20]. Logistic regression models the log-odds of a binary outcome as a linear combination of predictor variables:

$$\log\left(\frac{p(x)}{1-p(x)}\right) = \beta_0 + \beta^T x, \quad (5)$$

which can equivalently be expressed as:

$$p(x) = P(y = 1|x) = \frac{1}{1 + \exp[-(\beta_0 + \beta^T x)]}, \quad (6)$$

where $p(x)$ is the predicted probability of the outcome, β_0 is the intercept, and β represents the vector of regression coefficients. Exponentiating an individual coefficient yields the corresponding adjusted odds ratio, assuming that the remaining predictors are held constant [19,20].

Table 1 provides a conceptual comparison of the classical machine learning methods discussed in this section. It summarizes their principal characteristics, advantages, and limitations. Algorithm-specific performance ranges were not included because predictive performance depends strongly on the clinical task, dataset composition, preprocessing pipeline, and validation strategy.

Table 1. Conceptual comparison of classical machine learning methods.

Method	Characteristics	Advantages	Limitations	References
SVM	A model based on margin maximization; uses kernel functions (linear, polynomial, RBF)	High predictive performance; effectiveness in high-dimensional spaces; flexibility due to kernels	Low interpretability (“black box”); difficulty in parameter selection	[13–15]
RF	An ensemble of decision trees; bootstrap + random feature selection	High accuracy; resistance to overfitting; ability to assess variable importance	Bias in importance measures (e.g., for categorical variables); more difficult global interpretation	[16,17]
kNN	A nonparametric, lazy method; classification based on similarity (distance metrics)	Simplicity; no assumptions about data distribution; flexibility	Sensitivity to the choice of metric and k; computational cost for large datasets; no explicit model	[18]
Logistic regression	A statistical model based on the logit link function; estimation of odds ratios	High interpretability; ability to analyze multiple variables; solid statistical foundations	Sensitivity to multicollinearity; requires correct model specification; limited to linear relationships in the logit	[19,20]

3.1.2. Deep Learning

Convolutional Neural Networks (CNNs) are one of the fundamental deep learning techniques used for image analysis because they combine automatic feature extraction and classification within a single architecture [21]. Unlike classical machine learning methods, CNNs learn image features directly from input data, such as medical images [22]. Their core components are convolutional layers, in which filters detect patterns such as edges, textures, and anatomical structures. For an input feature map X and a convolutional kernel K , the two-dimensional convolution operation can be expressed as:

$$Y(i, j) = (X \times K)(i, j) = \sum_m \sum_n X(i - m, j - n) K(m, n) \quad (7)$$

where $Y(i, j)$ is the resulting feature-map value at spatial position (i, j) , while m and n index the elements of the convolutional kernel. Nonlinearity is commonly introduced using the rectified linear unit activation function:

$$\text{ReLU}(z) = \max(0, z) \quad (8)$$

This function preserves positive activation values while setting negative values to zero, thereby enabling the network to learn nonlinear relationships. CNNs also use pooling layers to reduce dimensionality and computational cost. For max pooling, the output at a given spatial position is defined as:

$$\text{MaxPool}(X)(i, j) = \max_{(m, n) \in \Omega} X(i + m, j + n), \quad (9)$$

where Ω denotes the local pooling window. This operation reduces the spatial dimensions of the feature map while retaining its strongest local activations [21–23].

Additional mechanisms including batch normalization, dropout, and residual connections improve training stability and reduce overfitting [21,24]. Thanks to these properties, CNNs are highly effective in medical image analysis, especially for detecting subtle pathological patterns [22,23]. Below, Figure 3 illustrates the architecture and workflow of a CNN.

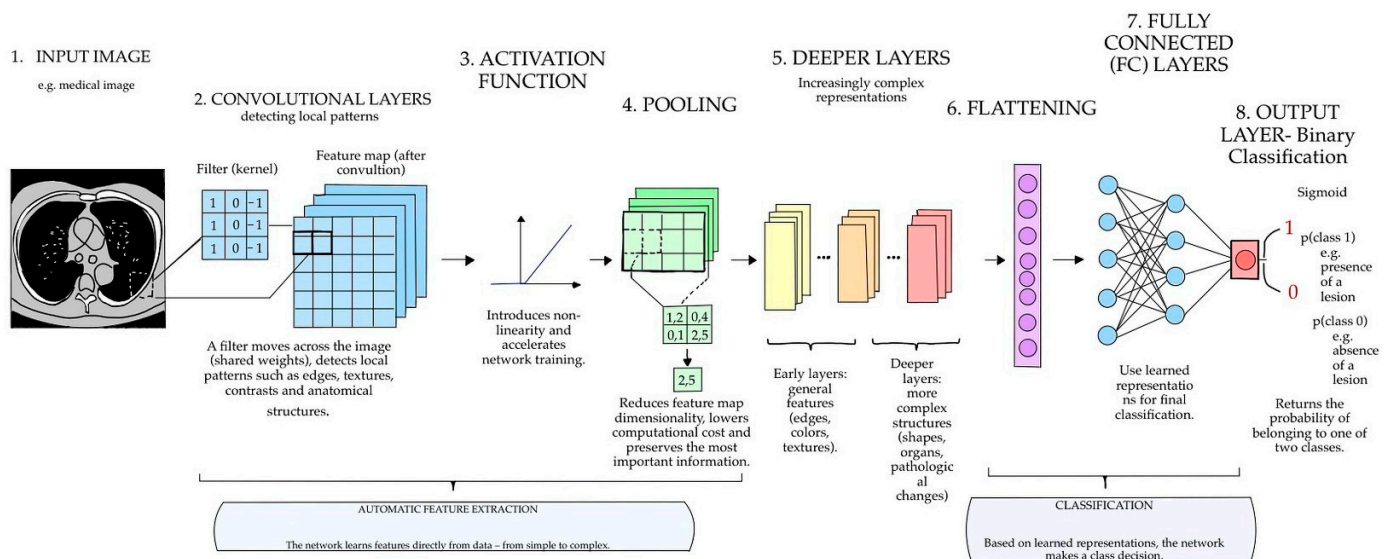


Figure 3. The figure illustrates the architecture and workflow of a CNN used for medical image analysis. The process includes input image acquisition, feature extraction through convolutional layers, application of an activation function, dimensionality reduction via pooling, processing in deeper layers, flattening, and final classification using fully connected layers. The model ultimately performs binary classification, estimating the probability of belonging to one of two classes.

The ResNet architecture extends classical CNNs by enabling effective training of much deeper networks through residual learning. Instead of learning full mappings directly, layers learn residual functions relative to the input data. This is achieved using shortcut or skip connections, which transfer information directly between layers [25]. These connections facilitate signal and gradient flow, reducing issues such as vanishing gradients and improving convergence in very deep models [25,26]. In practice, residual blocks sum convolutional outputs with original inputs, preserving earlier information while learning more complex representations [25,27]. The standard residual block can be represented as:

$$y = \mathcal{F}(x, \{W_i\}) + x, \quad (10)$$

where x is the input to the residual block, $\mathcal{F}$ denotes the residual mapping learned by the convolutional layers, $\{W_i\}$ represents the corresponding trainable weights and y is the block output [25–27].

This approach allows the construction of networks with hundreds of layers and inspired later modifications such as Res2Net, which improves multiscale feature representation [28]. In medical imaging, residual connections help maintain information flow and improve segmentation and classification accuracy [26,27].

The VGG architecture is a classical CNN model used in image recognition and classification, with VGG-16 and VGG-19 being the most common variants [29–32]. VGG-16 contains 13 convolutional layers and 3 fully connected layers, while VGG-19 is deeper. The model processes RGB images of size 224×224 , often after basic preprocessing such as mean RGB subtraction. Successive convolutional and pooling layers extract increasingly complex image features, with filter numbers growing from 64 to 512 in deeper blocks [29]. Fully connected layers then transform extracted features into classification outputs, typically ending with a softmax function for multiclass tasks. VGG can also serve as a feature extractor in hybrid models combined with classifiers such as SVM or Random Forest [30]. This transfer learning approach adapts pretrained models, often trained on ImageNet, to new tasks [29,30]. In breast imaging studies, VGG architectures are primarily used as pretrained feature extractors or components of transfer-learning and hybrid classification

models [29–32]. Mathematically, VGG applies repeated sequences of convolution, ReLU activation, and max-pooling operations, as defined in Equations (7)–(9), followed by fully connected layers that generate the final classification output.

EfficientNet is a modern deep learning architecture designed to optimize both accuracy and computational efficiency through compound scaling, which simultaneously scales network depth, width, and image resolution [33]. Compared with earlier methods that scaled only one parameter, EfficientNet achieves a better balance between performance and computational cost [33,34]. It uses MBConv blocks and depthwise separable convolutions for efficient feature extraction with relatively few parameters [33]. EfficientNet can also be combined with architectures such as ResNet in fusion models, further improving generalization and classification accuracy [33]. Due to these advantages, EfficientNet is considered one of the most efficient architectures for image classification tasks, especially when computational resources are limited [33–35].

Transfer learning involves applying knowledge acquired from one domain or task to improve performance on another related task, reducing the need for large target datasets [36]. In deep learning, this usually involves pretrained models whose layers may be frozen, fine-tuned, or extended with new layers [37,38]. This approach is important because deep learning models typically require large labeled datasets and high computational resources, both of which transfer learning helps reduce [38]. In CNNs, earlier layers learn general image features, while deeper layers capture more task-specific representations [37,38]. Common strategies include fine-tuning pretrained models, freezing feature extraction layers while training classification layers, and progressive learning with additional layers [38]. The effectiveness of transfer learning depends on the relationship between source and target domains, so results may vary across applications [37,38]. In breast cancer diagnostics, transfer learning is especially valuable because medical datasets are often small and difficult to annotate, while pretrained models enable more effective image classification [36,37].

CNNs have traditionally dominated breast imaging applications due to their ability to effectively extract local image features such as lesion margins, microcalcifications, and tissue texture. However, their limited receptive field may restrict the modeling of long-range spatial relationships and global contextual information. To address these limitations, Vision Transformers (ViTs) have emerged as a promising alternative architecture based on self-attention mechanisms, which enable the capture of global dependencies between distant image regions. Recent studies have demonstrated the potential of transformer-based approaches in breast cancer imaging. Abimouloud et al. reported that hybrid Vision Transformer-CNN architectures achieved improved mammographic classification performance by combining local feature extraction with global contextual awareness. Similarly, Kassis et al. demonstrated the effectiveness of Vision Transformers in digital breast tomosynthesis, where the ability to process complex volumetric information contributed to improved breast cancer detection. Furthermore, Jeny et al. developed a hybrid transformer-based model integrating both prior and current mammograms, showing that attention mechanisms can effectively exploit temporal information and improve classification performance. Despite these advantages, transformer-based models generally require larger datasets and greater computational resources than conventional CNNs, making them more susceptible to overfitting in smaller medical datasets. Consequently, current evidence suggests that hybrid CNN-Transformer architectures may represent the most promising direction for future breast imaging systems, combining the strengths of both local feature extraction and global representation learning [39–41].

Foundation models represent a rapidly emerging paradigm in medical artificial intelligence and are increasingly considered a potential next step beyond conventional task-

specific deep learning systems. Unlike traditional models trained for a single diagnostic task, foundation models are pretrained on extremely large and diverse datasets containing medical images, radiology reports, clinical text, and other healthcare data, enabling the learning of highly transferable representations that can be adapted to multiple downstream applications [42,43]. Recent developments in radiology have demonstrated that foundation models can function as generalist systems capable of supporting image classification, lesion detection, segmentation, report generation and clinical decision support within a unified framework [42]. Wu et al. developed a large-scale radiology foundation model trained on web-scale collections of both 2D and 3D medical images, demonstrating strong performance across diverse imaging tasks and highlighting the potential of generalized medical AI systems that can be transferred to multiple clinical applications [43]. Similarly, Paschali et al. emphasized that foundation models may substantially improve scalability, efficiency, and transferability in radiological workflows, while also discussing important challenges related to computational requirements, interpretability, data governance, and clinical validation [42]. Beyond radiology, foundation models are increasingly being explored in computational pathology. Vorontsov et al. developed a clinical-grade pathology foundation model trained on large collections of whole-slide images, demonstrating the ability to learn generalized histomorphological representations that can subsequently be applied to tumor classification, biomarker prediction, and analysis of the tumor microenvironment [44]. Despite their considerable promise, foundation models remain associated with important challenges, including substantial computational demands, limited interpretability, potential biases inherited from pretraining datasets, and the need for rigorous prospective validation before routine clinical implementation [42–44]. Nevertheless, they are increasingly viewed as a key component of future multimodal precision oncology systems capable of integrating imaging, pathology, genomics and clinical information within a unified AI framework [42,43].

Self-supervised learning (SSL) has emerged as a promising strategy for addressing one of the major limitations of breast imaging AI, namely the scarcity of large, accurately annotated datasets. Unlike conventional supervised learning, SSL enables models to learn informative feature representations from large collections of unlabeled medical images before fine-tuning on specific downstream clinical tasks, thereby reducing dependence on manual expert annotation [45,46]. This approach is particularly attractive in breast cancer imaging, where annotation of mammograms, MRI examinations, and histopathological images is time-consuming, expensive, and often affected by interobserver variability [46]. Miller et al. demonstrated that self-supervised pretraining can improve breast cancer detection performance in screening mammography and enhance label efficiency, allowing models to achieve strong diagnostic accuracy while requiring fewer annotated examinations for training [45]. More recently, SSL has been combined with transformer-based architectures, enabling more effective utilization of large-scale unlabeled imaging datasets and facilitating the development of transferable representations for breast cancer detection and classification tasks [47]. Figueiras et al. highlighted that modern SSL approaches, including contrastive learning, masked image modeling, and representation learning techniques, may improve model robustness, generalizability, and performance across different imaging modalities [46]. Furthermore, Wang emphasized the growing role of SSL as a foundation for transformer-based and foundation-model architectures in breast imaging, suggesting that large-scale pretraining on unlabeled datasets may contribute to the development of more generalizable and clinically applicable AI systems [47]. Despite these advantages, challenges remain regarding optimal pretraining strategies, computational requirements and the need for prospective validation in real-world clinical environments [46,47].

Multimodal large models represent an emerging direction in medical artificial intelligence, aiming to integrate diverse sources of information within a single predictive framework. Unlike conventional AI systems designed for specific tasks, these models can simultaneously analyze multiple data dimensions and generate predictions across several clinically relevant endpoints. Luo et al. developed a large-scale model for breast cancer management based on multiparametric MRI, demonstrating the potential of large AI systems to support non-invasive and personalized patient care. The model was designed to extract comprehensive imaging representations and perform multiple prediction tasks within a unified architecture, highlighting the feasibility of using large models for individualized risk assessment, tumor characterization, and treatment planning. Such approaches may contribute to the development of more generalized and scalable AI systems capable of supporting precision oncology; however, further validation across diverse clinical settings remains necessary before widespread implementation [48].

Federated learning (FL) has emerged as a promising approach for developing artificial intelligence models while preserving patient privacy and complying with increasingly stringent data protection regulations. Unlike conventional centralized training, which requires the aggregation of data from multiple institutions into a single repository, FL enables collaborative model development without direct sharing of sensitive patient information, as only model parameters or updates are exchanged between participating centers [49,50].

In the commonly used federated averaging approach, each participating client performs local model training, after which the central server constructs an updated global model by calculating a weighted average of the locally trained model parameters:

$$\theta^{(t+1)} = \sum_{k=1}^K \frac{n_k}{n} \theta_k^{(t+1)} \quad (11)$$

$$n = \sum_{k=1}^K n_k \quad (12)$$

where K denotes the number of participating clients, n_k is the number of training samples available at client k , n is the total number of samples across all participating clients, $\theta_k^{(t+1)}$ represents the model parameters obtained after local training at client k , and $\theta^{(t+1)}$ denotes the aggregated global model for the next communication round. Consequently, clients with larger local datasets contribute proportionally more to the global parameter update. Although federated averaging enables collaborative training without transferring raw patient data, differences in local sample sizes and data distributions may affect convergence and model performance [50].

This approach is particularly relevant in breast cancer research, where access to large and diverse datasets is essential for building robust AI systems, yet data sharing is often restricted by legal, ethical, and institutional constraints. Tzortzis et al. demonstrated that federated learning can improve the generalizability of breast imaging models across heterogeneous mammography datasets obtained from different healthcare institutions and imaging platforms, thereby reducing the risk of center-specific bias and improving model robustness [49]. Furthermore, Shukla et al. showed that integrating differential privacy mechanisms into FL frameworks can enhance data security and model integrity while maintaining strong diagnostic performance, highlighting the feasibility of privacy-preserving AI for breast cancer diagnosis [50]. Although federated learning offers significant advantages for multicenter collaboration and secure data utilization, challenges related to data heterogeneity, communication efficiency, and optimization across distributed environments remain important areas for future research [49,50].

Synthetic data generation has emerged as a promising strategy for addressing data scarcity, class imbalance, and privacy limitations in medical imaging. By using generative artificial intelligence techniques, researchers can create realistic artificial datasets that supplement real-world clinical data and support the development of more robust AI models. Synthetic data may facilitate data augmentation, algorithm validation, and multicenter research while reducing the need for direct sharing of sensitive patient information [51]. In breast imaging, diffusion-based generative models have demonstrated the ability to generate high-quality full-field digital mammograms and even synthesize lesion-containing images for training purposes. Montoya-del-Angel et al. developed the MAM-E framework, which uses diffusion models to generate realistic mammographic images and perform controlled lesion synthesis, highlighting the potential of synthetic data to enhance breast cancer detection systems and support AI model development in settings with limited annotated datasets. Mathematically, MAM-E uses a forward diffusion process in which Gaussian noise is progressively added to a latent representation of the mammogram according to a predefined noise schedule. A learned reverse process subsequently estimates and removes this noise to generate a mammographic image conditioned on a text prompt. For synthetic lesion generation, a binary mask constrains the reverse denoising and inpainting process to a predefined breast region while preserving the remaining image content [52]. Despite these advantages, further validation is required to ensure the clinical realism, diversity, and reliability of synthetic images before widespread implementation [51,52].

Diffusion models have recently gained increasing attention in breast cancer artificial intelligence due to their ability to learn complex data distributions and generate highly informative image representations. Beyond their original role in image synthesis, diffusion-based architectures are increasingly being applied to diagnostic and classification tasks. You et al. developed BreastDiff, a multi-condition guided diffusion model capable of analyzing breast cancer characteristics across different imaging modalities, demonstrating improved classification performance and enhanced feature extraction compared with conventional approaches [53]. Similarly, Akbari et al. employed a conditional denoising diffusion probabilistic model (DDPM) for breast cancer detection in histopathological images, combining diffusion-based tissue segmentation with transformer-based feature fusion to improve diagnostic accuracy across multiple datasets [54]. These findings suggest that diffusion models may offer a versatile framework for both image generation and diagnostic decision support in breast cancer care. However, further validation is required to determine their clinical utility, computational efficiency, and generalizability in real-world settings [53,54].

Uncertainty-aware artificial intelligence has emerged as an important approach for improving the reliability and trustworthiness of deep learning systems in breast cancer diagnosis. Unlike conventional AI models that provide only deterministic predictions, uncertainty-aware methods estimate the confidence associated with each prediction, allowing clinicians to identify cases that may require additional review or further diagnostic assessment [55,56].

Uncertainty should be interpreted according to the statistical framework used. In frequentist inference, uncertainty is commonly expressed using confidence intervals and *p*-values, whose interpretation is based on the behavior of repeated samples. In Bayesian inference, uncertainty is represented by a posterior distribution of model parameters conditional on the observed data, and interval estimates are expressed as credible intervals [57]. Practical model-based approaches, including Monte Carlo dropout and deep ensembles, estimate predictive variability by generating multiple predictions under different parameter realizations or model configurations. Monte Carlo dropout can be interpreted as an

approximate Bayesian approach, whereas ensemble variance reflects disagreement among independently trained models [58].

$$\text{Var}(y|x) = \underbrace{\text{Var}_{\theta}(\mathbb{E}[y|x, \theta])}_{\text{epistemic uncertainty}} + \underbrace{\mathbb{E}_{\theta}[\text{Var}(y|x, \theta)]}_{\text{aleatoric uncertainty}} \quad (13)$$

where x denotes the input data, y the predicted outcome and θ the model parameters. Epistemic uncertainty reflects uncertainty in the model parameters and may be reduced by increasing the amount or representativeness of training data. Aleatoric uncertainty arises from inherent noise or variability in the observations and generally cannot be eliminated by collecting more data of the same quality. Accordingly, confidence intervals, credible intervals, Monte Carlo dropout estimates, and ensemble variance should not be treated as interchangeable measures of uncertainty.

Chegini et al. developed an uncertainty-aware deep learning framework for breast cancer diagnosis using both mammographic and ultrasound images, demonstrating that uncertainty estimation can support more reliable classification across different imaging modalities [55]. In a subsequent study, the same group proposed a Bayesian deep learning model for molecular subtype prediction from full mammographic images, incorporating predictive uncertainty quantification through Monte Carlo Dropout techniques. The authors showed that uncertainty estimation not only provided information regarding prediction reliability but also supported clinically relevant classification of breast cancer molecular subtypes [56]. These findings suggest that uncertainty-aware AI may improve the safety, interpretability, and clinical acceptance of breast imaging algorithms, particularly in cases where prediction confidence is low and human oversight remains essential [55,56].

Explainable artificial intelligence (XAI) has become increasingly important in breast cancer diagnostics, as clinical adoption of AI systems requires not only high predictive performance but also transparent and interpretable decision-making. Techniques such as Grad-CAM, saliency maps and SHAP enable visualization of image regions or features that contribute most strongly to model predictions, thereby improving clinician trust and facilitating validation of AI outputs. Shen et al. developed an interpretable deep learning model for mammographic screening that generated localization maps highlighting suspicious regions associated with malignancy, demonstrating that explainability can be incorporated without substantially compromising diagnostic performance [59]. More recently, Sajid et al. integrated SHAP-based explanations, feature attribution methods, and uncertainty estimation within a breast cancer prediction framework, enabling clinicians to better understand the factors influencing model decisions and identify potentially unreliable predictions [60]. Such approaches may improve transparency, support clinical decision-making, and facilitate the safe implementation of AI systems in breast cancer care [59,60].

Causal artificial intelligence (Causal AI) has recently emerged as a promising approach for overcoming some of the limitations of conventional correlation-based machine learning models. Unlike traditional AI systems that primarily identify statistical associations, causal AI aims to uncover cause-and-effect relationships that may better reflect underlying biological and clinical mechanisms. Ribeiro-Dantas et al. applied causal discovery methods to nearly 400,000 breast cancer patient records, demonstrating that large-scale causal network analysis can identify interpretable relationships between clinical variables, treatment patterns, and patient outcomes while distinguishing potential causal effects from spurious associations [61]. In the diagnostic setting, Chen et al. developed a causal explainable AI framework based on mammography reports, integrating causal graphs with interpretable machine learning techniques to support breast cancer diagnosis and improve model transparency [62]. By explicitly modeling causal relationships, such approaches may improve

interpretability, robustness, and clinical trustworthiness compared with purely correlation-driven algorithms. However, further research is required to validate causal models in prospective clinical settings and determine their practical impact on decision-making in breast cancer care [61,62].

Continual learning has emerged as a promising strategy for maintaining the performance of artificial intelligence systems in dynamic clinical environments, where imaging protocols, patient populations, and disease characteristics may change over time. Unlike conventional deep learning models that are trained once on a fixed dataset, continual learning enables AI systems to incrementally incorporate new information while retaining previously acquired knowledge [63]. This approach may be particularly important in breast imaging, where continuous updates are often required to accommodate evolving imaging technologies and clinical workflows. Li et al. demonstrated that implementing continuous learning in breast MRI artificial intelligence improved diagnostic performance over time when new clinical data were incorporated into the model development process [64]. More broadly, continual learning has been proposed as a potential solution to challenges such as dataset shift and catastrophic forgetting, thereby improving the long-term robustness and adaptability of medical imaging AI systems [61]. Despite its potential, further research is needed to establish effective continual learning strategies that can be safely deployed in routine clinical practice while maintaining model stability and reliability [63,64].

Below in Table 2 there is a summary of the discussed deep learning techniques.

Table 2. Summary of the discussed deep learning techniques.

Method	Characteristics	Advantages	Limitations	Typical Application	References
CNN (Convolutional Neural Networks)	Deep neural networks using convolutional layers for automatic feature extraction from images; include pooling, activation functions (e.g., ReLU), and fully connected layers	Automatic feature extraction; high effectiveness in image analysis; no need for manual feature engineering	Require large datasets; high computational cost; prone to overfitting	Classification and detection of changes in medical images (e.g., mammography, histopathology)	[21–24]
ResNet	An extension of CNN with residual learning mechanism and shortcut connections (skip connections) enabling training of very deep networks	Solves the vanishing gradient problem; enables building very deep models; better convergence	Greater architectural complexity; higher computational requirements	Advanced classification and segmentation of medical images	[25–28]
VGG (VGG-16, VGG-19)	A classical CNN architecture with a simple, sequential arrangement of convolutional and pooling layers; uses fixed-size images (224×224)	Simple and interpretable architecture; good as a feature extractor; widely used	Large number of parameters; high memory consumption; slower than newer models	Feature extraction, medical image classification, hybrid models	[29–32]
EfficientNet	A modern architecture using compound scaling (scaling of depth, width, and resolution); uses MBConv and separable convolutions	High accuracy with lower computational cost; parameter efficiency; good generalization	More complex design process; dependence on proper scaling	Medical image classification, diagnostic systems requiring high precision	[33–35]
Transfer Learning	Use of pretrained models on large datasets; includes fine-tuning, layer freezing, and progressive learning	Reduced data requirements; faster training; improved performance on small datasets	Effectiveness depends on domain similarity; risk of mismatch	Breast cancer diagnostics with limited data; adaptation of models to new tasks	[36–38]

Table 2. Cont.

Method	Characteristics	Advantages	Limitations	Typical Application	References
Vision Transformers (ViT) and Hybrid CNN-Transformer Models	Self-attention-based architectures capable of modeling global image dependencies; often combined with CNNs for local feature extraction	Capture long-range relationships; improved contextual understanding; enhanced classification performance	Require large datasets and substantial computational resources; susceptible to overfitting in small datasets	Mammography classification, digital breast tomosynthesis, longitudinal breast image analysis	[39–41]
Foundation Models	Large-scale pretrained models trained on multimodal medical data (images, reports, text) and adaptable to multiple downstream tasks	Strong transferability; support multiple clinical applications within a unified framework; scalability	High computational requirements; limited interpretability; risk of dataset bias; need for extensive validation	Classification, detection, segmentation, report generation, pathology and radiology AI systems	[42–44]
Self-Supervised Learning (SSL)	Representation learning from unlabeled data using contrastive learning, masked image modeling, and related techniques	Reduces annotation requirements; improves robustness and label efficiency; leverages large unlabeled datasets	Computationally intensive; optimal pretraining strategies remain uncertain	Mammography, MRI and pathology image analysis with limited annotations	[45–47]
Multimodal Large Models	Large AI systems integrating multiple data sources and performing multiple prediction tasks within a unified architecture	Comprehensive patient representation; supports personalized medicine and precision oncology	Complex implementation; requires extensive multimodal datasets and validation	Breast cancer risk assessment, tumor characterization, treatment planning	[48]
Federated Learning (FL)	Distributed training approach where institutions share model updates rather than patient data	Preserves privacy; enables multicenter collaboration; improves model generalizability	Data heterogeneity; communication overhead; optimization challenges	Collaborative breast imaging AI development across institutions	[49,50]
Synthetic Data Generation	Use of generative AI to create realistic artificial medical images for augmentation and research	Addresses data scarcity and class imbalance; supports privacy-preserving research	Clinical realism and diversity must be carefully validated	Mammographic image augmentation, algorithm development and validation	[51,52]
Diffusion Models	Generative probabilistic models capable of image synthesis, feature extraction and classification through iterative denoising processes	High-quality image generation; improved representation learning; versatile applications	Computationally expensive; limited clinical validation	Mammography synthesis, lesion generation, histopathology classification	[52–54]
Uncertainty-Aware AI	Models that estimate predictive uncertainty alongside diagnostic outputs using Bayesian approaches or Monte Carlo methods	Improved reliability; identifies low-confidence cases; supports clinical decision-making	Increased model complexity; calibration challenges	Breast cancer diagnosis, molecular subtype prediction, risk assessment	[55,56]
Explainable AI (XAI)	Methods such as Grad-CAM, saliency maps and SHAP providing interpretation of model predictions	Improves transparency, trust and clinical acceptance; facilitates validation of AI decisions	Explanations may not always reflect true model reasoning; additional computational burden	Mammography interpretation, breast cancer prediction systems	[59,60]
Causal AI	Approaches designed to identify cause-and-effect relationships rather than statistical associations	Improved interpretability; greater robustness; potentially better clinical reasoning	Difficult causal inference; requires high-quality data and validation	Outcome prediction, diagnostic support, treatment-effect analysis	[61,62]
Continual Learning	Incremental learning paradigm allowing models to incorporate new data while retaining previous knowledge	Adaptation to changing clinical environments; reduces dataset shift effects	Risk of catastrophic forgetting; implementation complexity	Long-term breast imaging AI systems and evolving diagnostic workflows	[63,64]

In the breast cancer studies reviewed here, deep learning architectures were applied to mammography and digital breast tomosynthesis [39–41,45–47,49,52,55,56], multiparametric MRI [48], histopathological imaging [44,54], ultrasound [55], multimodal classification [53], and molecular subtype prediction [56]. Their reported performance and clinical relevance depended not only on architecture selection but also on cohort composition, preprocessing methods, availability of multicenter data, and the use of internal or external validation [42–50]. Therefore, the clinical evidence presented in the following sections is organized primarily according to clinical application and validation quality rather than algorithm type.

3.1.3. Technical Challenges and Emerging AI Paradigms in Breast Cancer CNNs Versus Transformers in Breast Imaging

CNNs remain the dominant deep learning architecture in breast imaging due to their ability to effectively capture local spatial patterns such as lesion margins, microcalcifications, architectural distortions, and texture heterogeneity, which are critical for mammographic, ultrasound and MRI interpretation. However, the local receptive field inherent to convolutional operations may limit the ability of CNNs to model long-range spatial relationships and global contextual information distributed across the breast image. To address these limitations, transformer-based architectures have recently emerged as a promising alternative. Vision Transformers (ViTs) employ self-attention mechanisms that enable the modeling of global dependencies between distant image regions, potentially improving the analysis of complex imaging patterns and subtle contextual features that extend beyond localized lesions [36,65–68].

Several studies have demonstrated that transformer-based models achieve diagnostic performance comparable to or exceeding that of conventional CNNs in breast imaging applications. In mammography, hybrid CNN–transformer architectures have shown superior performance by combining the strong local feature extraction capabilities of convolutional layers with the global contextual awareness provided by self-attention mechanisms [39,65]. Similar findings have been reported in digital breast tomosynthesis, where Vision Transformers effectively process volumetric image information and improve lesion detection performance in highly complex imaging datasets [66]. In breast ultrasound, transformer-based models have demonstrated excellent classification accuracy and may outperform traditional CNN architectures when larger datasets are available, particularly in challenging lesion characterization tasks [67,68].

Nevertheless, transformer architectures present several practical limitations. Unlike CNNs, which incorporate strong inductive biases such as locality and translational invariance, transformers generally require substantially larger datasets to achieve optimal performance. This issue is particularly relevant in breast cancer imaging, where high-quality annotated datasets remain relatively limited and expensive to acquire. As a result, purely transformer-based models may be more susceptible to overfitting when trained on small cohorts [65,67]. Consequently, recent research increasingly favors hybrid CNN–transformer architectures that leverage the complementary strengths of both approaches. Such models have demonstrated improved robustness, generalizability, and classification accuracy across mammography, ultrasound and MRI applications [39,68]. Current evidence therefore suggests that transformers are unlikely to fully replace CNNs in the near future; instead, the most promising direction appears to involve integrated architectures combining convolutional feature extraction with attention-based global representation learning.

2D vs. 3D Deep Learning in Breast MRI

Deep learning approaches in breast MRI can generally be divided into 2D and 3D architectures, each associated with specific technical advantages and limitations. Traditional

2D convolutional neural networks analyze individual MRI slices independently and are computationally less demanding, allowing faster training and lower GPU memory requirements. However, because breast MRI is inherently volumetric, slice-based analysis may fail to capture important spatial relationships between adjacent slices and may incompletely represent tumor morphology, spatial heterogeneity, and enhancement patterns distributed throughout the lesion volume [69,70]. In contrast, 3D deep learning architectures process the entire volumetric dataset simultaneously, enabling improved modeling of tumor shape, contextual anatomical information, and inter-slice dependencies. Several studies demonstrated that 3D CNNs achieve superior performance in lesion classification, segmentation, and molecular subtype prediction compared with conventional 2D approaches [69–72].

The advantages of 3D models appear particularly important in dynamic contrast-enhanced breast MRI (DCE-MRI), where tumor characterization depends not only on local texture but also on volumetric enhancement kinetics and spatial tumor heterogeneity. Weakly supervised 3D CNN frameworks have shown the ability to simultaneously classify and localize breast lesions while reducing the need for labor-intensive voxel-level annotations [70]. Similarly, multi-scale and context-aware 3D architectures demonstrated improved malignancy classification by incorporating broader anatomical context extending beyond the lesion itself [73]. Hybrid models combining 2D and 3D convolutions have also emerged as a promising compromise between computational efficiency and volumetric feature extraction, allowing integration of local high-resolution slice information with global three-dimensional context [71].

Despite their advantages, 3D deep learning models remain associated with important practical limitations. Volumetric MRI analysis substantially increases computational complexity, memory consumption, and training time compared with 2D approaches. In addition, 3D architectures typically require larger annotated datasets to avoid overfitting, which represents a major challenge in breast MRI due to limited availability of standardized multicenter datasets [69,70]. Variability in MRI acquisition protocols, scanner vendors, and image resolution may further reduce the generalizability of 3D models across institutions. Consequently, current research increasingly focuses on hybrid and weakly supervised approaches that attempt to balance diagnostic performance, computational feasibility, and robustness in real-world clinical environments [70,71].

Domain Shift and Generalizability

One of the major challenges limiting the clinical translation of artificial intelligence in breast imaging is domain shift, which refers to differences between the data used for model development and the data encountered during real-world deployment. In breast cancer imaging, domain shift may arise from variations in scanner vendors, acquisition protocols, image reconstruction methods, patient demographics, disease prevalence and annotation practices across institutions [74–77]. As a result, models demonstrating excellent performance in internal validation frequently experience substantial performance degradation when applied to external datasets. This problem is particularly pronounced in breast MRI, where differences in imaging protocols and scanner characteristics can significantly alter image appearance and feature distributions, thereby reducing segmentation and classification accuracy [74,78].

Recent studies have demonstrated that improving model generalizability requires strategies that explicitly account for cross-domain variability. Domain adaptation techniques aim to align feature representations between source and target datasets, allowing models to maintain performance despite differences in image acquisition conditions [74]. Similarly, domain generalization approaches seek to learn robust representations that remain stable across previously unseen environments without requiring access to target-

domain data during training [76,77]. In mammography, scanner-specific effects and vendor-related image characteristics have been shown to substantially influence model behavior, highlighting the importance of multicenter validation and harmonization strategies before clinical implementation [75,79].

Beyond technical performance, limited generalizability may also affect model fairness and reliability. AI systems trained on narrowly selected populations may perform unevenly across demographic groups, healthcare settings, or imaging platforms, potentially introducing unintended biases into clinical decision-making [78]. Consequently, external validation across diverse patient cohorts has become increasingly recognized as a critical prerequisite for trustworthy deployment of breast imaging AI. Emerging approaches such as federated learning, multicenter collaborative training, and large-scale benchmark datasets may help mitigate domain shift by exposing models to greater variability during development while preserving data privacy [80]. Overall, improving robustness to domain shift remains essential for ensuring that high retrospective performance translates into reliable real-world clinical utility.

Weakly Supervised Learning for Mammography

One of the major barriers to the development of deep learning systems in mammography is the limited availability of high-quality annotated datasets. Precise lesion delineation requires expert radiologist input and is both time-consuming and costly, particularly for large-scale screening datasets. To address this challenge, increasing attention has been directed toward weakly supervised learning approaches, which enable model training using only image-level labels, such as the presence or absence of malignancy, without requiring detailed pixel- or lesion-level annotations [59,81–83]. Such approaches substantially reduce annotation burden while allowing the utilization of large clinical datasets that would otherwise be unsuitable for supervised learning.

Several weakly supervised frameworks have demonstrated the ability to simultaneously perform breast cancer classification and lesion localization in mammograms despite being trained exclusively on examination-level labels [81,82]. These methods commonly employ attention mechanisms, multiple instance learning strategies, or region-ranking architectures to identify image regions most strongly associated with malignancy. More recent studies have further improved localization performance through self-training and iterative pseudo-label generation, enabling models to progressively refine lesion identification without additional manual annotation [83]. Weak supervision has also been successfully applied to high-resolution mammography analysis, where interpretable models can generate localization maps highlighting suspicious regions while maintaining strong diagnostic performance [59,84].

Beyond lesion detection, annotation-efficient learning is increasingly being extended to more advanced clinical applications. Weakly supervised frameworks have been used to predict tumor biological characteristics, including hormone receptor status, directly from mammographic images without requiring extensive region-level annotations [85]. Furthermore, the integration of weak supervision with semi-supervised learning allows models to leverage large collections of weakly labeled examinations and radiology reports, thereby substantially increasing the amount of training data available for model development [86]. These approaches may be particularly important for future breast imaging AI systems, as they offer a practical pathway toward developing robust models while minimizing the substantial annotation effort traditionally required for supervised deep learning.

Self-Supervised Learning in Breast Imaging

Self-supervised learning (SSL) has emerged as a promising strategy for addressing one of the major limitations of breast imaging AI: the scarcity of large, accurately annotated datasets. Unlike conventional supervised learning, SSL enables models to learn informative feature representations from vast amounts of unlabeled imaging data through pretext tasks, contrastive learning, masked image modeling, or other representation-learning strategies before fine-tuning on downstream clinical tasks [43,45,46,87,88]. This approach is particularly attractive in breast imaging, where expert annotation of mammograms, MRI examinations, and other imaging modalities is time-consuming, expensive and often limited by interobserver variability.

Recent studies have demonstrated that self-supervised pretraining can significantly improve breast cancer detection performance while simultaneously reducing dependence on manually labeled data. In screening mammography, SSL-based models have shown improved diagnostic accuracy, enhanced label efficiency, and better generalization compared with conventional supervised approaches, particularly when annotated datasets are limited [45]. The combination of SSL with modern architectures such as Vision Transformers and hybrid CNN–transformer models has further improved lesion detection performance by enabling more effective utilization of large-scale unlabeled mammographic datasets [87]. In addition, anatomically aware self-supervised frameworks trained on millions of mammograms have demonstrated strong cross-dataset generalizability and robustness, suggesting that large-scale representation learning may provide a foundation for more transferable breast imaging AI systems [88].

The growing interest in SSL extends beyond its application to individual diagnostic models and increasingly encompasses the development of foundation models for breast imaging. These models are initially trained on large-scale and highly diverse datasets, after which they can be fine-tuned for a variety of downstream tasks, including lesion detection, classification, segmentation, risk prediction, and radiogenomic analysis [43,47,88,89]. Owing to their broad pretraining, foundation models have the potential to decrease reliance on task-specific model development and promote the creation of AI systems that generalize more effectively across institutions and imaging environments. While clinical implementation is still in its early stages, accumulating evidence suggests that SSL-based pretraining may evolve into a dominant framework for future breast imaging AI as the amount of available imaging data continues to grow [43,45,47,87–89].

Foundation Models for Radiology/Pathology

Foundation models represent a rapidly emerging paradigm in medical artificial intelligence and may significantly influence the future development of breast imaging and computational pathology. Unlike conventional task-specific deep learning systems, foundation models are pretrained on extremely large multimodal datasets containing medical images, radiological reports, histopathological slides, and clinical text, enabling the learning of highly transferable representations adaptable to multiple downstream tasks [90–94]. In radiology, these models are increasingly being developed as generalist systems capable of performing image classification, lesion detection, segmentation, report generation, radiogenomic prediction, and clinical decision support within a unified framework [90,91].

Recent advances in radiology foundation models have been driven largely by the integration of self-supervised learning, multimodal pretraining, and vision–language architectures. Large-scale systems such as RadFM have demonstrated that foundation models pretrained on millions of radiological images can generalize across both 2D and 3D imaging modalities while maintaining strong performance on diverse downstream applications [90]. Similarly, vision–language foundation models trained jointly on medical images and ra-

diology reports enable the integration of visual and textual information, facilitating more context-aware interpretation and improving transferability across clinical tasks [92,93]. Such multimodal architectures may be particularly relevant for breast cancer care, where imaging findings must often be interpreted together with histopathological, genomic, and clinical information.

Foundation models are also increasingly being explored in computational pathology. Pretraining on large collections of whole-slide images may allow models to learn generalized histomorphological representations that can subsequently be adapted for tumor classification, biomarker prediction, and microenvironment analysis [94]. However, current evidence suggests that pathology foundation models remain associated with substantial challenges, including biological heterogeneity of tissue architecture, variability in staining and slide preparation, limited interpretability, and reduced robustness across institutions [95]. Moreover, the extremely large computational requirements of foundation models, together with concerns regarding transparency, fairness, and clinical validation, continue to limit widespread implementation in routine clinical practice [91,96].

Despite these limitations, foundation models may ultimately provide the basis for highly generalizable and multimodal AI systems capable of integrating radiology, pathology, genomics, and clinical data within unified precision oncology frameworks. Still, they are increasingly viewed as one of the most important future directions in breast cancer artificial intelligence research [90–96].

Multimodal Fusion Models

Breast cancer is a biologically heterogeneous disease characterized by complex interactions between imaging phenotypes, histopathological features, molecular alterations, and clinical variables. Consequently, models relying on a single data modality may capture only a limited aspect of tumor biology. To overcome this limitation, increasing attention has been directed toward multimodal fusion approaches, which integrate information derived from imaging, pathology, genomics, and clinical data within a unified predictive framework [48,97–102]. Such models aim to exploit the complementary strengths of individual modalities, enabling a more comprehensive representation of disease characteristics and supporting precision oncology applications.

Recent studies have consistently demonstrated that multimodal models outperform unimodal approaches across a variety of clinically relevant tasks. The integration of MRI features with pathological findings, genomic profiles, and clinical variables has been associated with improved prediction of treatment response, recurrence risk, and molecular subtype classification compared with models based on imaging alone [98–102]. In the neoadjuvant setting, multimodal deep learning systems combining radiological, pathological and clinical information have shown enhanced ability to predict pathological complete response (pCR), facilitating more individualized treatment planning [99,100]. Similarly, the incorporation of gene expression data alongside pathological and clinical features has improved risk stratification and prediction of recurrence or metastatic progression [101].

Advances in deep learning have also enabled the development of increasingly sophisticated multimodal fusion architectures. Modern systems frequently employ attention mechanisms, transformer-based fusion strategies, cross-modal feature alignment, and mixture-of-experts frameworks to model complex relationships between heterogeneous data sources [48,97,98]. Such approaches are particularly attractive in breast cancer, where imaging findings often require contextual interpretation in light of molecular subtype, tumor microenvironment characteristics, and patient-specific clinical factors. Furthermore, multimodal fusion may provide an important foundation for future radiogenomic and precision medicine applications by linking non-invasive imaging biomarkers with under-

lying biological processes [98,102]. Although challenges related to data harmonization, missing modalities, and model interpretability remain substantial, multimodal learning is increasingly regarded as one of the most promising directions for the development of clinically meaningful breast cancer AI systems [48,97–102].

Federated Learning and Privacy-Preserving AI

The development of artificial intelligence in breast cancer increasingly depends on access to large, diverse, and multicenter datasets. However, sharing sensitive medical imaging and histopathological data between institutions remains restricted by privacy regulations, ethical concerns, and legal limitations. Federated learning (FL) has therefore emerged as a promising approach that enables collaborative model training across multiple institutions without direct exchange of patient-level data [49,50,103–106]. In federated learning, models are trained locally within participating centers, while only model parameters or gradients are shared and aggregated centrally, thereby reducing the need for direct data transfer and improving privacy protection.

Recent studies have demonstrated the feasibility of federated learning in breast imaging applications, including mammography, digital breast tomosynthesis, and breast cancer classification tasks [49,50,103,104]. FL-based systems have shown the ability to achieve diagnostic performance comparable to centralized training while simultaneously improving model robustness across heterogeneous datasets originating from different institutions and imaging devices [49,103]. In mammography, federated learning frameworks combined with modern deep learning architectures such as YOLOv6 have demonstrated effective breast cancer detection performance while incorporating privacy-preserving strategies including encrypted parameter sharing and differential privacy mechanisms [50,104]. Such approaches may be particularly important for real-world clinical deployment, where strict data protection requirements frequently limit the creation of large centralized datasets.

Federated learning is also increasingly being explored in computational pathology and histopathological image analysis. Multi-institutional FL frameworks may facilitate collaborative training on whole-slide images while preserving institutional control over highly sensitive pathology data [104,105]. In addition to privacy preservation, federated learning may improve model generalizability by exposing AI systems to greater demographic, biological and technical variability during training [49,105]. Nevertheless, important challenges remain, including communication efficiency, heterogeneity between participating institutions, differences in imaging protocols, data imbalance, and the risk of performance degradation caused by non-identically distributed data across centers [106]. Consequently, although federated learning represents one of the most promising strategies for privacy-preserving breast cancer AI, further large-scale prospective validation and technical standardization remain necessary before widespread clinical implementation can be achieved.

The Black-Box Challenge: Explainability and Human-AI Collaboration

One of the major barriers to the clinical implementation of artificial intelligence in oncology is the “black-box” nature of many deep learning models, where predictions are generated without providing transparent explanations of the underlying decision-making process. Consequently, increasing attention has been directed toward explainable artificial intelligence (XAI) techniques designed to improve model interpretability and clinician trust [107]. Commonly used approaches include Grad-CAM, saliency maps, attention maps, SHAP (Shapley Additive Explanations) and other feature attribution methods, which enable visualization and quantification of image regions or features that contribute most strongly to model predictions [107,108]. Shen et al. developed an interpretable deep learn-

ing model for mammographic screening that generated localization maps highlighting suspicious regions associated with malignancy, demonstrating that explainability can be incorporated into breast imaging algorithms while maintaining high diagnostic performance [108]. Beyond technical interpretability, explainable AI may also facilitate more effective collaboration between clinicians and AI systems. Calisto et al. demonstrated that personalized explanation strategies can improve clinician–AI interaction by enhancing understanding of model outputs, increasing user confidence, and supporting more informed diagnostic decision-making. Such human-in-the-loop approaches allow physicians to critically evaluate AI recommendations rather than passively accept algorithmic outputs, thereby combining computational efficiency with clinical expertise [109]. As a result, explainability is increasingly recognized as a key requirement for the safe, trustworthy and responsible implementation of artificial intelligence in breast cancer care [107–109].

3.2. AI in Breast Cancer Imaging

3.2.1. Mammography

The application of AI in mammography currently represents one of the most dynamically developing areas of medical imaging, encompassing both direct diagnostic support and optimization of screening program organization. In the classical approach, AI serves as CAD-type systems based on deep learning methods, which enable automatic detection of suspicious lesions, their localization, and classification of examinations as normal or abnormal [110–112]. The most advanced prospective and randomized studies indicate that the integration of AI into clinical practice can at least match the standard double reading of mammograms by radiologists. In the ScreenTrustCAD study involving 55,581 women, the strategy combining one radiologist with AI was non-inferior to standard double reading. This approach detected 261 cancers compared with 250 cancers using standard double reading, corresponding to 11 additional detected cancers, a 4.4% relative increase, and a relative proportion of 1.04 (95% CI: 1.00–1.09). The AI-supported strategy was also associated with a slightly lower recall rate, despite a higher number of examinations initially classified as abnormal [110]. In the MASAI study, which included more than 105,000 participants, AI-supported screening detected 338 cancers compared with 262 cancers in the control group. The corresponding cancer detection rates were 6.4 per 1000 examinations (95% CI: 5.7–7.1) and 5.0 per 1000 examinations (95% CI: 4.4–5.6), respectively. This corresponded to a significant 29% relative increase in cancer detection, with a proportion ratio of 1.29 (95% CI: 1.09–1.51; $P = 0.0021$). AI-supported screening also reduced the screen-reading workload by 44.2% [111]. A subsequent long-term analysis of the MASAI trial demonstrated significantly higher screening sensitivity in the AI-supported group than in the control group: 80.5% (95% CI: 76.4–84.2) versus 73.8% (95% CI: 68.9–78.3; $P = 0.031$). Interval cancer rates were 1.55 per 1000 participants (95% CI: 1.23–1.92) and 1.76 per 1000 participants (95% CI: 1.42–2.15), respectively, corresponding to a proportion ratio of 0.88 (95% CI: 0.65–1.18; $P = 0.41$). Specificity was identical in both groups at 98.5% (95% CI: 98.4–98.6; $P = 0.88$) [113].

In parallel with controlled studies, real-world clinical data confirm the effectiveness of AI in population settings. In a large multicenter implementation study in Germany involving over 463 thousand women, the use of AI increased breast cancer detection from 5.7 to 6.7 per 1000 women (+17.6%) without increasing the false positive rate, and the so-called “safety net” mechanism enabled the identification of at least 204 cancers potentially missed by radiologists [112]. Beyond the role of a “second reader,” AI is also applied in exam triage, enabling prioritization of patients requiring rapid diagnostics and significantly reducing the time to further imaging (25.6 vs. 19.1 days) and to biopsy diagnosis (55.9 vs. 39.2 days), with the high-risk group reaching an average of 3.5 days [114]. It is worth

emphasizing that the development of AI in mammography extends beyond diagnostics—it also includes educational applications, such as generating synthetic mammograms for training physicians, which may improve interpretative skills, leading to an increase in sensitivity by 7.43%, NPV by 5.05%, and overall accuracy by 6.49% [115].

Despite these results, the implementation of AI in mammography requires consideration not only of technical and clinical aspects, but also social factors. Studies on patient preferences indicate that the hybrid model, in which AI supports the radiologist, has the highest acceptance, while solutions that completely replace humans raise more concerns, with the key decision factor being the maximization of sensitivity while limiting false positive results [116].

3.2.2. Magnetic Resonance Imaging

The application of AI in breast magnetic resonance imaging (MRI) covers a wide spectrum of uses, from detection and classification of lesions, through risk stratification, to the analysis of tumor biological features and optimization of the imaging process itself. In particular, dynamic contrast-enhanced MRI (DCE-MRI) represents a key area of AI application, as it provides both morphological and kinetic information. Modern models based on deep learning, including Vision Transformer and UNETR architectures, enable simultaneous segmentation and classification of lesions, achieving high diagnostic performance. The BL4AS system was developed and evaluated using a multicenter dataset comprising 2803 BI-RADS 4 lesions from 2686 women. The model achieved an AUC of 0.896 (95% CI: 0.832–0.949) in external test set A and 0.930 (95% CI: 0.906–0.950) in external test set B. In the prospective test set, BL4AS achieved an AUC of 0.892 (95% CI: 0.813–0.951), with a sensitivity of 0.862 (95% CI: 0.773–0.944) and a specificity of 0.889 (95% CI: 0.750–1.000), compared with a pooled specificity of 0.491 for radiologists. BL4AS demonstrated significantly higher specificity than seven of the eight individual radiologists ($P < 0.05$). When used as a decision-support tool, BL4AS increased readers' specificity by 27.3 percentage points and reduced the mean false-positive rate from 50.9% to 23.6% ($p = 0.012$) [117]. The application of AI is particularly important in the BI-RADS 4 category, where models can further stratify lesions into subcategories 4A–4C, potentially reducing the number of unnecessary biopsies [117–119].

Multiparametric MRI (mpMRI), including post-contrast T1, DWI/ADC, and T2WI sequences, enables further improvement of AI model performance through the integration of information on perfusion, diffusion, and tissue structure. Models based on CNN (e.g., ResNet18) achieve very high accuracy in lesion differentiation, especially in difficult cases such as distinguishing triple-negative breast cancer from fibroadenoma—in one study, an AUC of 0.944, sensitivity of 0.926, specificity of 0.950, and accuracy of 0.940 were achieved in a test set of 319 patients [118]. Similarly, fully automatic CAD systems combining segmentation (Dice 0.831) with radiomic analysis achieve high diagnostic performance (AUC 0.946 internally and 0.842 in external validation), and integration with BI-RADS assessment allows increasing AUC up to 0.975 with sensitivity of 0.920 and specificity of 0.923 [119].

AI in breast MRI is also used for non-invasive assessment of tumor biology. Deep learning models enable prediction of molecular subtypes of breast cancer based on MRI images, achieving an AUC of 0.920 for the triple-negative subtype and 0.885 for HER2-enriched in an analysis of 136 patients [120]. Radiogenomic studies have also demonstrated the possibility of identifying more subtle relationships between imaging and molecular profiles, although performance remains moderate (AUC around 0.65) [121]. Classical diagnostic tasks, such as differentiation between benign and malignant lesions, can also be supported by transfer learning models, for example DenseNet201, which achieve very high

accuracy (98.01% in the test set) and good agreement with histopathology (kappa 0.749), although validation results are more moderate (AUROC 0.79) [122].

An important, often overlooked area of AI application is the technical optimization of MRI. Deep learning-based image reconstruction (DLR) allows significant reduction in DWI acquisition time from 3 min 36 s to 1 min 54 s (−47.2%) while simultaneously improving image quality (better CNR, fewer artifacts, better lesion visibility) and without significant changes in ADC values, which is crucial for diagnostic reliability [123]. In addition, AI can support more complex clinical decisions, such as predicting the presence of an invasive component in DCIS. In an analysis of 131 patients, a deep learning model achieved an AUC of 0.70 with sensitivity of 69% and specificity of 68%, indicating the potential of this method as a risk assessment support tool, although insufficient for independent clinical decision-making [124].

Nevertheless, it should be noted that despite the growing number of promising results, most studies are retrospective, include relatively small patient groups (often <300), and use data from single centers, which limits the generalizability of the results. Additionally, many models do not integrate clinical data or other imaging modalities.

3.2.3. Ultrasound, Other Techniques, and Combination Approaches

The application of AI in breast ultrasound as well as in approaches combining different imaging modalities extends classical AI applications beyond mammography and MRI alone. In ultrasound, systems supporting the classification of benign and malignant lesions, segmentation, and reduction in false positive results are particularly important. In a study on automated breast volume ultrasound (ABVS), a deep learning model analyzing images in three planes achieved very high diagnostic performance: AUC 0.984 in internal testing and 0.978 and 0.942 in two external validations, with sensitivity of 98.2%, specificity of 90.3%, and accuracy of 94.0%. Importantly, AI improved the performance of less experienced radiologists, sensitivity increased from approximately 60.2% to 79.7%, specificity from 78.2% to 88.7%, and accuracy from 69.9% to 84.4%, while also reducing the average interpretation time from 36.5 to 14.4 s [125].

An important direction of AI development in ultrasound is also the analysis of elastography. In the international multicenter INSPIRED 006 study, the AI-SWE model was developed using data from 924 patients and 4026 images. In two independent validation cohorts, the model achieved AUROC values of 0.940 and 0.930, respectively, with corresponding sensitivities of 97.9% and 97.8%. The model reduced false-positive findings by 62.1% and 38.1% in the respective validation cohorts [126]. Similar importance is observed in the analysis of screening ultrasound elastography based on CNN, where the model achieved AUC 0.895, sensitivity 0.800, specificity 0.966, and accuracy 0.898, outperforming classical elastographic indicators such as fat-to-lesion ratio and elasticity score [127]. AI can also support the assessment of axillary lymph node involvement; in a study using BPNN in 90 patients, segmentation accuracy reached 97.3%, and specificity in subsequent analyses ranged from 90.31 to 97.65% [128].

Beyond ultrasound itself, multimodal and personalized strategies are gaining increasing importance. AI can be used to select women after negative mammography who would benefit most from additional MRI. In the ScreenTrustMRI study, the AISmartDensity algorithm selected 4103 women at the highest risk out of 59,354 examined, i.e., the top 6.9%, and among 559 women who underwent additional MRI, 36 cancers were detected, corresponding to 64.4 cancers per 1000 MRI examinations. This strategy was approximately four times more effective than earlier approaches based solely on conventional breast density [129]. Multimodal models can also combine MRI with clinical and radiological features, for example in predicting lymphovascular invasion (LVI), where PCMM-Net achieved

AUC 0.843, outperforming the clinical-radiological model (0.743), radiomic model (0.795), and the model without additional clinical features (0.774) [130].

AI can also integrate data from mammography, ultrasound, and clinical features for non-invasive prediction of breast cancer biology. In a study involving 600 patients, interpretable machine learning models predicted molecular subtypes of breast cancer based on 56 variables, achieving for TNBC an AUC of 0.971, accuracy of 0.947, sensitivity of 0.905, and specificity of 0.941, and for luminal and HER2 subtypes AUC values of 0.900 and 0.855 respectively. AI support also improved radiologists' performance, especially less experienced ones, where in TNBC differentiation sensitivity, specificity, and accuracy increased by 0.090, 0.125, and 0.114 respectively [131].

Other promising, although still complementary, approaches include infrared thermography supported by AI. A mobile AI-IRT system based on MobileNetV3-Small, prospectively evaluated in 2202 individuals from 20 centers, achieved in binary classification an AUC of 0.9487, sensitivity of 0.9325, specificity of 0.8026, and accuracy of 0.8698 in internal validation, and an AUC of 0.9120 and accuracy of 0.8627 in external validation. Such a system does not replace mammography, ultrasound, or biopsy, but may serve as a low-cost, mobile preliminary triage tool, especially in populations with limited access to conventional diagnostics [132]. AI also has educational applications, interactive breast self-examination simulation increased nursing students' satisfaction (56.59 vs. 50.45 points), although it was less effective than simulation with a standardized patient in terms of practical skills (59.71 vs. 73.72 points) [133].

3.3. AI for Prognosis, Treatment-Response Prediction and Treatment Selection

3.3.1. Axillary Lymph-Node Prediction and Prognostic Stratification

AI in breast cancer is increasingly extending beyond classical detection tasks and is being applied to prognosis, selection of patients for specific therapeutic strategies, and prediction of treatment response. In the prognostic domain, models predicting axillary lymph node involvement are particularly important, as nodal status remains one of the key prognostic factors and influences the extent of surgical and systemic treatment. Yu et al., in a multicenter cohort of 1088 patients with invasive breast cancer, applied MRI radiomics and machine learning algorithms, including random forest for feature selection and SVM for predictive model construction. A radiomic signature including the primary tumor and the axillary lymph node region achieved an AUC of 0.88 in the training cohort and 0.87 in both validation cohorts, while a multi-omics model combining radiomics with clinicopathological data and molecular subtype achieved AUC values of 0.90, 0.91, and 0.93 [134]. A similar direction is represented by the study of Dihge et al., in which, based on the population-based SCAN-B cohort including 3023 patients, models based on clinicopathological data, gene expression, and their combination were compared. The best combined model, using among others gradient boosting machine, achieved an AUC of approximately 0.72 in validation, with the most important predictors being tumor size, lymphovascular invasion, age, and multifocality [133]. Such models may in the future support the identification of patients at low risk of nodal involvement, in whom limitation of surgical procedures such as sentinel lymph node biopsy or axillary lymphadenectomy may be considered [134,135].

AI is also used in more general prognostic stratification. Bhattarai et al. used a unique cohort of 92 patients in whom the tumor was retrospectively visible on earlier mammography, which allowed estimation of the actual in vivo tumor growth rate. Rapidly growing tumors were associated with larger histological size, higher grade, increased mitotic count, vascular invasion, worse Nottingham Prognostic Index, and shorter breast cancer-specific survival. The authors then developed the Surr-INVIGOR model based on routinely avail-

able parameters such as Ki-67, mitotic index, and tumor size, and validated its value in an independent cohort of 1241 patients [136]. Pan et al. developed a prognostic signature based on 10 lncRNAs associated with cuproptosis. The model stratified patients into high- and low-risk groups, and the risk score remained an independent prognostic factor. Prediction performance for overall survival reached AUC values of 0.755 for 1 year, 0.717 for 3 years, and 0.643 for 5 years [137]. In a similar context, Jin et al. used Random Forest to predict adverse events in patients with poor response to neoadjuvant chemotherapy. The model achieved an AUC of 0.810 for 1-year prediction and 0.829 for 5-year prediction, and in external validation on SEER data achieved an AUC of 0.779, outperforming logistic regression with an AUC of 0.619 [138]. These results suggest that AI-based models may integrate clinical, histopathological, imaging, and molecular data to support prediction of progression, recurrence, or death; however, their clinical utility requires further prospective validation [136–138].

3.3.2. Prediction of Treatment Response

A second important area is the prediction of treatment response. In breast cancer treated with neoadjuvant therapy, radiomic models can help assess the response of both the primary tumor and lymph nodes. Lee et al. analyzed 226 patients with initially node-positive disease treated with neoadjuvant chemotherapy. Pathological complete response in axillary lymph nodes was achieved in 120 patients (53.1%), while 106 patients (46.9%) had residual metastases. A qualitative CT assessment model after NAC achieved only an AUC of 0.642, whereas radiomic models achieved AUC values of 0.812 for intranodal ROI, 0.762 for perinodal ROI, and 0.832 for combined ROI. The best model was a clinical–radiomic model based on post-NAC CT, with an AUC of 0.866, sensitivity of 74.1%, specificity of 88.9%, and accuracy of 80.0% [139].

In advanced disease, AI-based models have been investigated for predicting clinical benefit from systemic therapy using imaging and molecular biomarkers. Dercle et al. evaluated 106 patients with advanced HR+/HER2– breast cancer treated with exemestane and everolimus with or without xentuzumab. Of eight tested radiomic biomarkers, seven showed significant predictive value, with key factors including liver tumor volume, total tumor volume, and their changes after 8 weeks of treatment. A multimodal model based on 40 clinical and imaging variables and a Random Forest algorithm achieved an AUC of approximately 0.75 [140]. Côté et al. showed that germline SNPs in ERBB3 and BARD1 may be associated with worse relapse-free survival in patients with HER2-positive breast cancer treated with the TCH regimen. In a study including 157 patients, SVM was used to predict RFS events based on 6 SNPs and follow-up time, with AI serving a supportive role while the main contribution was the identification of predictive biomarkers for treatment response [141].

Particularly promising within digital histopathology and molecular profiling are models predicting response to immunotherapy. Li et al. applied automated immunophenotyping of digital histopathology slides in a cohort including patients with triple-negative breast cancer from the IMpassion130 trial. Algorithms analyzed the spatial distribution of CD8+ lymphocytes relative to tumor cells and classified tumors as desert, excluded, or inflamed. The inflamed phenotype was associated with greater benefit from atezolizumab plus nab-paclitaxel and better overall survival and progression-free survival compared to desert or excluded phenotypes [142]. Greenwald et al. analyzed 103 patients with metastatic TNBC from the TONIC trial using highly multiplexed imaging of 37 proteins across 270 tumors. The SpaceCat pipeline extracted over 800 features per sample, including cell density, immune population diversity, spatial interactions, and functional marker expression. The best prediction of response to nivolumab was obtained from

on-treatment samples, where multivariable models achieved an AUC of approximately 0.90 [143]. Complementarily, Hernando-Calvo et al. showed that a 12-gene transcriptomic signature (VIGex) and ctDNA dynamics can predict outcomes of pembrolizumab therapy. Patients with VIGex-Hot tumors had higher response rates than intermediate-cold/cold groups (24% vs. 10%), and the combination of VIGex-Hot with decreasing ctDNA yielded an ORR of 53%, while no responses were observed in the VIGex I-Cold/Cold group with increasing ctDNA [144]. These data suggest that AI-based integration of digital pathology, spatial microenvironment analysis, transcriptomics, and ctDNA may support prediction of immunotherapy benefit. However, superiority over established biomarkers such as PD-L1 or TMB requires confirmation in prospective comparative studies [142–144].

3.3.3. AI-Guided Treatment Selection

A third area is active treatment selection, meaning the use of AI not only to assess risk or response, but also to recommend therapeutic strategies. Fan et al., in the LINUXtrial study, applied deep learning to classify patients with metastatic HR+/HER2– breast cancer after progression on CDK4/6 inhibitors into subtypes SNF1–SNF4 based on digital H&E slides. Precision therapy tailored to AI-defined subtypes improved the overall objective response rate from 23% in the control group to 51% in the precision treatment group. The greatest benefit was observed in subtypes SNF2 and SNF4: in SNF2, ORR was 65% vs. 30% and median PFS 8.1 vs. 4.3 months, while in SNF4 ORR was 70% vs. 20% and median PFS 7.0 vs. 3.4 months [145]. An even more direct example of AI as a treatment selection tool was presented by Ge et al., who developed the GDnet model integrating transcriptomic data with drug representations. The model was trained on 4371 patients from 31 datasets and aimed to predict pCR after neoadjuvant therapy and select the optimal treatment regimen. GDnet outperformed models based solely on transcriptomics, with AUROC around 0.71 vs. 0.68, and in simulated clinical trials increased treatment optimization intensity was associated with a linear increase in odds ratio for achieving pCR from approximately 1.6 to 2.5 [146]. These studies illustrate the emerging concept of AI-assisted “digital therapy testing”; however, its ability to guide treatment selection and improve patient outcomes remains to be established in prospective clinical studies [145,146].

3.3.4. Radiotherapy Planning and Procedure Selection

Treatment selection may also involve local procedures, radiotherapy planning, and patient-centered decisions. Mao et al. developed radiomic models to predict the optimal number of fields in postoperative IMRT radiotherapy in 242 breast cancer patients. The best radiomic model based on Random Forest achieved an accuracy of 0.76, while a combined model integrating radiomic score with clinical factors such as T stage, N stage, and type of surgery increased accuracy to 0.80 [147]. Kazemimoghadam et al. developed the SDL-Seg model for automatic segmentation of tumor bed volume after breast-conserving surgery. In a study including 29 patients and 145 CT images, a saliency-guided 3D U-Net achieved a mean Dice similarity coefficient of 76.4%, HD95 of 6.76 mm, and ASD of 1.9 mm, outperforming the classical U-Net (Dice 62.6%). Segmentation of a single image took approximately 10 s, which may be relevant for rapid radiotherapy planning workflows [148]. Another dimension of treatment selection is presented by the CINDERELLA study, a prospective, randomized, multicenter trial including approximately 1030 patients. The AI platform uses clinical images, biometric and clinical data, and case-based reasoning to predict aesthetic outcomes of locoregional treatment. This is not a prediction of oncological response, but a tool supporting shared decision-making by better aligning treatment choices with patient expectations [149].

It is also worth emphasizing that AI can support the selection of patients for further invasive diagnostics, which constitutes an indirect element of therapeutic decision-making. Liu et al. developed a deep learning model to predict malignancy of BI-RADS 4 microcalcifications in screening mammography. In a retrospective analysis of 384 patients and 414 histopathologically confirmed microcalcifications, including 221 malignant and 193 benign, the best combined model integrating mammography with clinical and radiological data achieved in the test set an AUC of 0.910, sensitivity of 85.3%, and specificity of 91.9%. The model outperformed standard BI-RADS assessment and improved the performance of junior radiologists: AUC increased from 0.816 to 0.854 and from 0.773 to 0.901 respectively, and interobserver agreement measured by kappa increased from 0.331 to 0.843. Although this study primarily concerns malignancy prediction, its clinical significance lies in better selection of patients requiring biopsy and reduction in unnecessary invasive procedures [150].

3.4. AI in Survivorship, Supportive Care, and Patient Communication

In recent years, an expansion of artificial intelligence applications in breast cancer has been observed, moving toward supporting patient care, quality of life, and optimization of the clinical pathway. A particularly important direction is the use of AI in survivorship care, i.e., long-term care after oncological treatment. The ASCAPE project represents an example of a comprehensive approach to quality of life prediction, in which machine learning models integrate data from medical records, QoL questionnaires, mobile applications, wearable devices, and environmental data to predict problems such as anxiety, depression, fatigue, joint pain, sleep disturbances, neurotoxicity, or lymphedema in subsequent 3-month follow-up periods. This project employs advanced technical solutions, including federated learning and homomorphic encryption, highlighting the growing importance of secure data analysis in multicenter and cross-border contexts. Although this approach does not directly concern oncological prognosis or tumor response to treatment, it demonstrates the potential of AI in predicting and preventing treatment-related complications and in personalizing supportive care [151].

Complementing this direction are digital interventions using AI to improve quality of life and reduce psychological symptoms. Jiang et al., in a randomized study involving 124 young patients (18–45 years), demonstrated that the AI-TA mobile application, using behavioral data analysis and content personalization, led to significant improvement in psychological symptoms (MSAS-SF), self-efficacy (CBI-B), social support, and quality of life (FACT-B) after just 1 month, with sustained effects after 3 months [152]. In turn, in a study by Schmitz et al., involving 42 patients with metastatic breast cancer, the virtual assistant Nurse AMIE monitored daily symptoms such as pain, fatigue, sleep, and distress, and proposed supportive interventions. The intervention was well evaluated (acceptability ~51%, feasibility 65%, satisfaction $\geq 70\%$), although no statistically significant improvement in clinical parameters was observed, which may be due to the limited sample size [153]. These results suggest that AI can serve as a tool supporting symptom monitoring and communication with the patient [152,153].

Another important area is the use of AI in patient education and improving communication during the treatment process. In a randomized study by Lee et al., involving 145 patients undergoing radiotherapy, a chatbot and video materials were compared with traditional forms of education in terms of their effect on treatment-related anxiety. In the overall population, no significant differences were observed in anxiety reduction measured by APAIS, STAI, and LASA scales; however, subgroup analysis showed a trend toward greater anxiety reduction in younger patients (≤ 50 years) using the chatbot [154]. Similarly, Al-Hilli et al. conducted a randomized trial involving 37 women with stage 0–III breast

cancer, comparing an AI chatbot with traditional genetic counseling. All patients opted for genetic testing, and pathogenic variants were detected in 13.5% (5/37). The chatbot achieved comparable results to traditional counseling both in terms of knowledge (median 11 vs. 12 points) and satisfaction (median 30 points), without affecting treatment delays or surgical decisions [155].

3.5. Comparative Clinical Readiness of AI Applications in Breast Cancer Care

The current landscape of artificial intelligence in breast cancer demonstrates substantial heterogeneity not only in diagnostic performance but also in validation quality, translational maturity, and real-world clinical applicability. While some AI applications are approaching integration into routine clinical workflows, others remain highly experimental despite reporting very high performance metrics.

Among all currently available applications, AI-assisted mammographic screening possesses the strongest level of clinical evidence and the highest degree of translational readiness. Large prospective and population-based studies, including the ScreenTrustCAD and MASAI trials, demonstrated that AI-supported mammographic screening can achieve performance comparable to or exceeding standard double-reading workflows while simultaneously reducing radiologists' workload [110,111,113]. Importantly, these studies included prospective evaluation, external validation and real-world implementation involving very large cohorts exceeding hundreds of thousands of examinations [110–114]. Therefore, mammography currently represents the AI application closest to routine clinical implementation in breast cancer care.

In contrast, AI applications in breast MRI remain less clinically mature despite frequently reporting very high diagnostic metrics. Deep learning and radiomic models demonstrated strong performance in lesion classification, BI-RADS stratification, and molecular subtype prediction, with several studies reporting AUC values above 0.90 [117–120]. However, most MRI studies remain retrospective, frequently originate from single centers, and rarely include prospective validation [117–124]. Moreover, substantial concerns remain regarding scanner dependence, protocol heterogeneity, and limited generalizability across patient populations. Although MRI-based AI demonstrates considerable promise, its routine clinical implementation still requires large multicenter prospective validation studies.

Similarly, AI systems in ultrasound and elastography have shown encouraging results, particularly in improving diagnostic accuracy among less experienced radiologists and reducing false-positive findings [125–127]. Some studies included external validation and demonstrated excellent diagnostic performance [125,126]. Nevertheless, operator dependence, acquisition variability, and the lack of prospective workflow-based validation continue to limit widespread implementation. Therefore, current ultrasound AI systems should primarily be considered assistive rather than autonomous clinical tools.

Beyond imaging diagnostics, AI is increasingly being applied to prognosis prediction, treatment response assessment, and precision oncology strategies. Radiomic and multi-modal models predicting axillary lymph node involvement, recurrence risk, or response to neoadjuvant therapy have demonstrated promising performance [134–139]. In addition, AI-driven integration of imaging, transcriptomics, digital pathology, and ctDNA analysis has shown potential in predicting immunotherapy benefit [140–142]. However, despite encouraging retrospective results, most of these models remain investigational due to limited external validation, relatively small cohorts, and the absence of prospective clinical utility studies. Similarly, AI-guided treatment selection systems, including the LINUX and GDnet studies, represent highly innovative approaches but remain early translational technologies rather than clinically established decision-support systems [145,146].

A particularly emerging area involves multimodal and patient-centered AI applications integrating imaging, genomics, pathology, electronic health records, symptom monitoring, and survivorship care [129–131,151–155]. These approaches may ultimately support comprehensive precision oncology ecosystems extending beyond diagnosis alone. Nevertheless, current evidence remains heterogeneous and many systems are still limited to proof-of-concept or pilot-stage evaluations.

Importantly, many published studies across imaging, prognostic, and therapeutic domains report extremely high performance metrics, frequently with AUC values exceeding 0.90. Such findings should be interpreted cautiously, as they may partly reflect methodological limitations including retrospective design, small or highly curated datasets, spectrum bias, hidden data leakage, overfitting, limited demographic diversity, and insufficient external or temporal validation. Therefore, excellent retrospective performance does not necessarily translate into real-world clinical robustness, fairness, or generalizability.

Direct comparison of performance metrics across studies remains challenging because of substantial methodological heterogeneity. The reviewed studies differed in patient selection, disease prevalence, imaging modality, reference standard, dataset size, class distribution, outcome definition, decision threshold, and validation strategy. Moreover, some studies reported results from internal test sets, whereas others used external, multicenter, prospective, or real-world validation. Consequently, apparently similar AUC, sensitivity, or specificity values should not be interpreted as directly equivalent across studies.

Future studies should consistently report cohort size and class distribution, patient selection criteria, model-development and validation procedures, point estimates with 95% confidence intervals, sensitivity and specificity at prespecified decision thresholds, and, where applicable, calibration measures. Clear distinction between internal, external, prospective, and real-world validation would further improve comparability and facilitate assessment of clinical readiness. For the purposes of this narrative review, clinical readiness was categorized using an author-defined qualitative framework based on study design, cohort size, availability of external or multicenter validation, prospective or randomized evidence, consistency of reported performance, evaluation in real-world clinical workflows, and evidence of clinical utility. “Moderate-High” readiness was assigned to applications supported by large prospective, randomized, multicenter, or real-world studies demonstrating consistent performance and potential workflow integration. “Low–Moderate” readiness indicated promising results with some external or multicenter validation but limited prospective or workflow-based evidence. “Low” readiness was assigned to applications supported predominantly by retrospective studies, with limited external validation and no established prospective clinical utility. “Very Low” readiness indicated early proof-of-concept or exploratory applications lacking robust external validation, prospective evidence, and demonstrated integration into clinical practice. These categories are intended for descriptive comparison within this review and do not represent a validated regulatory or clinical grading system. The comparative clinical readiness of major AI applications in breast cancer care is summarized in Table 3.

Table 3. Comparative clinical readiness of major AI applications in breast cancer care.

Application	Typical Performance	External Validation	Prospective Evidence	Current Clinical Readiness	References
Mammography screening AI	High diagnostic accuracy; improved cancer detection and workflow efficiency	Present in several large multicenter studies	Yes; randomized prospective trials available	Moderate-High	[110–114]
Breast MRI lesion classification	Moderate-to-high performance; several studies report AUC > 0.90	Limited; available in selected studies	Rare	Low–Moderate	[117–120]

Table 3. *Cont.*

Application	Typical Performance	External Validation	Prospective Evidence	Current Clinical Readiness	References
MRI molecular subtype prediction	Promising but variable radiogenomic performance	Rare	No	Low	[120,121]
Ultrasound and elastography AI	High performance in selected cohorts; reduction in false positives	Present in some multicenter studies	Rare	Low–Moderate	[125–127]
Multimodal imaging AI	Promising integration of imaging and clinical variables	Limited	No	Low	[129–131]
Prognostic and lymph node prediction models	Moderate-to-high performance in retrospective cohorts	Limited	Rare	Low	[134–138]
Prediction of neoadjuvant treatment response	Promising radiomic and clinical–radiomic performance	Rare	No	Low	[139,140]
Immunotherapy response prediction	Experimental but biologically promising multimodal approach	Rare	No	Very Low	[142–144]
AI-guided treatment selection systems	Early precision oncology applications with promising preliminary result	Very limited	No	Very Low	[145,146]
Radiotherapy planning and segmentation AI	Good technical performance in selected workflows	Limited	Rare	Low–Moderate	[147,148]
Patient support, survivorship and chatbot systems	Moderate improvements in symptom monitoring and patient engagement	Limited	Present in small prospective studies	Very Low	[151–155]

Note: Clinical readiness categories were assigned using the author-defined qualitative framework described in the preceding text.

In addition to this review-wide qualitative appraisal, a design-specific formal risk-of-bias assessment was performed for the representative primary clinical studies included in Table 4.

Table 4. Quantitative summary of representative studies evaluating AI applications in breast cancer care.

Study/Application	Sample Size	Study Design/Validation	Main Performance Results	Reference
ScreenTrustCAD—mammographic screening	55,581 women	Prospective, population-based, paired-reader non-inferiority study	261 vs. 250 detected cancers; 11 additional cancers, corresponding to a 4.4% relative increase	[110]
MASAI—AI-supported mammographic screening	>105,000 participants	Randomized controlled screening study	Cancer detection rate: 6.4 vs. 5.0 per 1000 examinations; 338 vs. 262 cancers; 29% relative increase; 44.2% reduction in radiological readings	[111]
Nationwide German implementation study	>463,000 women	Multicenter real-world implementation study	Cancer detection rate: 6.7 vs. 5.7 per 1000 women; 17.6% relative increase without an increase in the false-positive rate	[112]
BL4AS—breast MRI lesion classification	2686 patients; 2803 lesions	External validation and prospective evaluation	AUC 0.896–0.930 in external validation and 0.892 in prospective evaluation; specificity 0.889 vs. 0.491 for radiologists	[117]

Table 4. *Cont.*

Study/Application	Sample Size	Study Design/Validation	Main Performance Results	Reference
Multiparametric MRI model for differentiating TNBC from fibroadenoma	319 patients	Retrospective single-center study with internal split-sample testing; no external validation	AUC 0.944; sensitivity 0.926; specificity 0.950; accuracy 0.940	[118]
AI-SWE—shear-wave elastography	924 patients (4026 images) in the development set; 194 patients (562 images) and 176 patients (188 images) in two external validation sets	International multicenter model-development study with two external validation cohorts, including validation using updated SWE software	AUROC 0.94 and 0.93; sensitivity 97.9% and 97.8%	[126]
ScreenTrustMRI—AI-based selection for supplemental MRI	59,354 screened women; 559 underwent MRI	Interim analysis of prespecified secondary outcomes from the MRI intervention arm of a randomized population-based trial	36 cancers detected among 559 MRI examinations; 64.4 cancers per 1000 MRI examinations	[129]

4. Discussion

This review indicates that artificial intelligence is increasingly being investigated for applications in diagnosis, prognosis, prediction of treatment response, and the organization of care for patients with breast cancer. The analyzed studies suggest that AI should not be viewed solely as a tool for automatic detection of neoplastic lesions, but rather as a technology with the potential to support different stages of the clinical pathway, from screening and tumor biology assessment to treatment personalization and post-therapy care [110–155].

The evidence summarized in Tables 3 and 4 illustrates a clear difference in methodological maturity across clinical applications. AI-supported mammographic screening is supported by large prospective, randomized, and real-world studies, whereas evidence concerning MRI, ultrasound, molecular prediction, treatment-response assessment, and treatment selection is derived predominantly from retrospective cohorts, smaller datasets, and internal validation. This contrast directly informs the following discussion of generalizability, risk of overfitting, external validation, workflow integration, and clinical readiness.

From a methodological perspective, both classical machine learning algorithms such as SVM, Random Forest, kNN and logistic regression as well as modern deep learning architectures, including CNN, ResNet, VGG, EfficientNet, and transfer learning, are of particular importance [13–38]. Classical ML methods remain especially relevant in the analysis of tabular, clinical, radiomic and molecular data, where datasets are often smaller and model interpretability is of significant clinical importance [13–20]. In contrast, deep learning models have demonstrated strong performance in medical image analysis, as they enable automatic feature extraction without the need for manual design [21–24]. Transfer learning is especially relevant in breast cancer, as medical data are often limited, difficult to obtain, and require expert annotation [36–38].

The strongest clinical evidence currently concerns the application of AI in screening mammography. Results from the ScreenTrustCAD and MASAI studies suggest that AI can achieve performance comparable to standard double reading by radiologists while potentially increasing breast cancer detection rates and reducing radiologists' workload [110,111]. Notably, in the MASAI study, the use of AI was associated with an increase in cancer detection from 5.0 to 6.4 per 1000 screenings and a reduction in the number of radiological readings by 44.2% [111]. Real-world clinical data further support the potential utility of AI; a large German study involving over 463,000 women reported increased breast cancer detection without a corresponding increase in false-positive rates [112]. Although these

findings are encouraging, further prospective validation is required before broad clinical implementation.

AI has also demonstrated potential in breast MRI, particularly in differentiating benign and malignant lesions, stratifying BI-RADS 4 lesions, and assessing tumor biology [117–124]. Deep learning models have achieved high AUC values and, in some studies, improved specificity compared to radiologists, suggesting a possible role in reducing unnecessary biopsies [117–119]. Importantly, MRI represents an area where AI may extend beyond simple image classification. Deep learning models have shown the ability to predict molecular subtypes such as TNBC and HER2-enriched, while AI-based image reconstruction approaches have demonstrated the potential to shorten DWI acquisition times while maintaining diagnostic performance [120,123]. These findings indicate promising opportunities for AI to support both image interpretation and technical optimization of the diagnostic process, although prospective validation remains necessary.

In ultrasound and multimodal approaches, AI may provide support particularly for less experienced radiologists and may help reduce false-positive findings [125–128]. In a study on automated breast volume scanning (ABVS), AI was associated with improvements in sensitivity, specificity, and diagnostic accuracy among less experienced radiologists, while also reducing interpretation time [125]. Similarly, models analyzing elastography demonstrated promising diagnostic performance and may help reduce unnecessary biopsies in BI-RADS 3 and 4 lesions [126,127]. A particularly important direction is the integration of different imaging modalities and clinical data. An example is the ScreenTrustMRI strategy, where AI identified women with negative mammography who might benefit most from additional MRI [129]. This reflects the potential of AI to contribute to more personalized diagnostic strategies.

Applications of AI in prognosis and treatment selection suggest that this technology may support the development of personalized medicine. Models integrating radiomic, clinico-pathological, and molecular data have demonstrated promising performance in predicting axillary lymph node involvement, risk of progression, recurrence, or death [134–138]. In the context of neoadjuvant treatment, radiomic models outperformed conventional imaging in predicting response in axillary lymph nodes [129]. Particularly promising are models predicting response to immunotherapy, which integrate digital pathology, tumor microenvironment analysis, transcriptomics and ctDNA [142–144]. While these findings are encouraging, their clinical utility requires prospective validation in larger and more diverse patient populations.

An interesting direction is also the use of AI as a potential tool for treatment selection. The LINUXtrial study suggested that classification of patients into AI-defined subtypes may be associated with improved objective response rates in metastatic HR+/HER2– breast cancer [145]. Meanwhile, the GDnet model, integrating transcriptomic data with drug representations, demonstrated the potential for digitally comparing different therapeutic strategies prior to clinical application [146]. Such approaches may eventually expand the role of AI beyond risk assessment; however, their clinical value remains to be established in prospective studies.

It is also important to emphasize that AI applications in breast cancer extend beyond diagnosis and oncological treatment. Increasing attention is being given to tools supporting quality of life, symptom monitoring, patient education, and communication with the medical team [151–155]. The ASCAPE project demonstrated the potential of AI in long-term post-treatment care by predicting symptoms such as anxiety, depression, fatigue, pain, sleep disturbances, and lymphedema [151]. Mobile applications, virtual assistants, and chatbots may support patients in education, self-care and symptom monitoring, although current evidence suggests that their effectiveness depends on factors such as sample size, patient age, and the specific endpoints assessed [152–155].

4.1. Challenges, Limitations and Future Directions of AI in Breast Cancer

The implementation of AI in routine clinical practice in breast cancer still faces significant limitations. One of the main challenges is model generalizability, as many studies—particularly in MRI, tumor biology prediction and treatment planning—are retrospective, involve relatively small patient cohorts, and often originate from single centers [117–124,139–141,147,148]. Such models may achieve excellent performance on training or test datasets but may perform less effectively when applied to different populations, imaging systems, diagnostic protocols, or healthcare settings.

Another limitation is interpretability, especially in the case of nonlinear models such as SVM and complex deep learning architectures, which may function as “black boxes” [14,15,21–24]. In oncology, this is particularly important, as diagnostic and therapeutic decisions must not only be accurate but also clinically explainable. Data quality remains another key challenge: AI models require large, well-annotated, and representative datasets, whereas medical data are often limited, heterogeneous, difficult to standardize, and require time-consuming expert annotation [36–38].

Additionally, some studies report very high performance metrics but lack sufficient external or prospective validation, limiting their direct applicability in clinical practice [117–124,132,138]. Social and ethical aspects must also be considered. Patients tend to prefer hybrid models in which AI supports radiologists rather than fully autonomous systems replacing human decision-making [116].

Therefore, future development of AI in breast cancer should focus on large, prospective, multicenter validation studies, data and protocol standardization, development of explainable models, and integration of AI into existing clinical workflows. Particularly promising are multimodal systems combining data from mammography, MRI, ultrasound, genomics, transcriptomics, ctDNA and clinical information [129–131,134–146].

In the future, AI may support personalized screening, reduction in unnecessary biopsies, prediction of response to systemic and immunotherapy, radiotherapy planning, treatment selection, and long-term post-treatment care [60,70–86]. However, safe implementation will require maintaining physician oversight, ensuring model transparency, protecting data privacy and demonstrating real clinical benefit in prospective studies.

4.2. Methodological Weaknesses of Current AI Studies

Despite the rapidly growing number of studies investigating artificial intelligence in breast cancer, substantial methodological limitations continue to hinder the translation of many AI systems into routine clinical practice. Although numerous studies report excellent diagnostic and predictive performance, these findings should often be interpreted cautiously due to important issues related to dataset quality, validation strategies, model generalizability and reproducibility.

One of the most common limitations is the predominance of retrospective study designs. Many AI models in breast imaging, prognosis prediction, and treatment response assessment were developed and evaluated using retrospective datasets collected from single institutions [117–131,134–146]. While retrospective analyses are useful for initial model development, they are inherently associated with selection bias, spectrum bias and limited control over data heterogeneity. As a result, model performance observed in retrospective cohorts may not accurately reflect real-world clinical effectiveness.

The lack of robust external validation also remains a major issue. Although selected studies incorporated multicenter validation cohorts [110–114,125,126,134], many published AI models were evaluated only using internal validation or random train–test splits [117–124,129–146]. Internal validation alone is insufficient to ensure clinical robustness because patients from the same institution frequently share similar imaging protocols,

scanner characteristics, demographic features, and annotation practices. Consequently, model performance may deteriorate substantially when applied to data obtained from different healthcare systems or imaging platforms.

Closely related to this issue is the problem of scanner and protocol dependence. Imaging data used for AI training may vary considerably according to MRI vendor, ultrasound acquisition settings, mammography equipment, reconstruction algorithms, and institutional imaging protocols [117–131]. Therefore, models trained in one environment may fail to maintain comparable performance in external clinical settings. This challenge is particularly important in breast MRI and ultrasound, where image acquisition remains relatively heterogeneous and operator-dependent.

Another important but often underreported issue is hidden data leakage. In AI research, leakage may occur when information from the test set unintentionally influences model training, leading to artificially inflated performance metrics. Although many reviewed studies reported excellent diagnostic accuracy, sensitivity and AUC values, insufficient methodological transparency in some studies makes it difficult to fully exclude subtle forms of leakage, particularly in retrospective image-based datasets [117–131,134–146]. This issue may partially explain why some models demonstrate near-perfect performance despite relatively limited datasets.

Discrimination and calibration represent two distinct aspects of model performance. Discrimination describes the ability of a model to distinguish between patients with and without a particular outcome and is commonly assessed using AUC, sensitivity, and specificity. Calibration, in contrast, describes the agreement between predicted probabilities and the frequencies of outcomes observed [156]. Consequently, a model may demonstrate excellent discrimination while remaining poorly calibrated, meaning that its predicted probabilities systematically overestimate or underestimate the true risk.

One commonly used measure of overall probabilistic prediction error is the Brier score:

$$Brier = \frac{1}{N} \sum_{i=1}^N (\hat{p}_i - y_i)^2 \quad (14)$$

where N is the number of observations, $\hat{p}_i$ is the probability predicted for the i -th observation, and y_i is the corresponding observed binary outcome. Lower Brier scores indicate smaller differences between predicted probabilities and observed outcomes.

Calibration should also be evaluated graphically by comparing predicted probabilities with observed outcome frequencies across clinically relevant risk groups. In a well-calibrated model, the calibration curve should approximate the 45-degree identity line. However, among the studies summarized in this review, calibration measures and calibration plots were reported substantially less frequently than discrimination metrics. This limits assessment of whether the reported probabilities can be reliably used for individual clinical decision-making.

Class imbalance represents another important methodological concern. In breast cancer datasets, positive cases are often substantially less frequent than negative cases. Consequently, accuracy alone may provide a misleading assessment of model performance. For example, in a dataset with a disease prevalence of 5%, a classifier that assigns every case to the negative class would achieve an accuracy of 95% while having a sensitivity of zero. Therefore, studies using imbalanced datasets should report class distribution and complement accuracy with metrics such as sensitivity (recall), specificity, precision, F1-score, and the Matthews correlation coefficient (MCC). Where appropriate, precision–recall curves and the area under the precision–recall curve should also be considered. In the reviewed

literature, dataset balancing procedures and class-sensitive performance measures were not reported consistently, which limits direct comparison between studies.

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (15)$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (16)$$

where TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively. The F1-score summarizes the balance between precision and recall, whereas MCC incorporates all four components of the confusion matrix and is particularly informative when class distributions are imbalanced [157]. However, no single metric fully characterizes clinical performance; therefore, these measures should be interpreted together with sensitivity, specificity, disease prevalence, and the intended clinical application.

The reproducibility of AI studies also remains problematic. Many studies use proprietary datasets, institution-specific preprocessing pipelines, and non-standardized annotation methods [117–146]. In addition, differences in lesion segmentation, radiomic feature extraction, image preprocessing, and threshold selection may substantially influence final model performance. Such variability complicates independent validation and limits reproducibility across institutions.

Annotation variability constitutes another important challenge. In breast imaging and pathology, annotations often depend on expert interpretation, which may differ between radiologists and pathologists. Variability in BI-RADS classification, lesion segmentation, or histopathological labeling may introduce additional noise into training datasets and affect model stability [117–131].

Finally, many currently available AI systems remain insufficiently evaluated in prospective clinical workflows. Although mammography screening studies such as MASA and ScreenTrustCAD demonstrated encouraging prospective results [110,111,113], the majority of AI applications in MRI, ultrasound, prognosis prediction, immunotherapy response assessment, and treatment selection remain at the retrospective or proof-of-concept stage [117–146]. Therefore, despite highly promising preliminary findings, the real-world clinical utility of many AI systems in breast cancer remains incompletely established.

Although a design-specific formal risk-of-bias assessment was performed for the representative primary studies summarized in Table 4, the remaining heterogeneous literature was evaluated qualitatively. Therefore, the appraisal should not be interpreted as a comprehensive risk-of-bias assessment of all 180 publications included in the narrative synthesis.

Overall, current evidence suggests that artificial intelligence has considerable potential in breast cancer care; however, substantial methodological improvements, including larger multicenter datasets, rigorous external and prospective validation, improved reproducibility, standardized reporting, and robust evaluation of model calibration and generalizability, are necessary before widespread clinical implementation can be achieved.

Sample Size, Model Complexity and Risk of Overfitting

Another important methodological concern is the frequent use of relatively small and highly curated datasets. For example, some of the reviewed MRI studies included only 136 patients for molecular subtype prediction [120] or 131 patients for predicting occult invasive disease in ductal carcinoma in situ [124]. One of these studies reported an AUC exceeding 0.900 despite the limited cohort size [120]. Although high diagnostic performance in a small cohort does not necessarily indicate an invalid model, it increases uncertainty regarding the stability, reproducibility, and generalizability of the reported results.

Sample-size adequacy in artificial intelligence studies cannot be determined solely from the total number of model parameters or the number of available images. Multiple images or lesions obtained from the same patient are correlated and should not be interpreted as fully independent observations. Consequently, the number of patients and the number of outcome events within each class are generally more informative than the total number of images alone.

The sample size required for reliable model development depends on several factors, including the clinical task, outcome prevalence, class balance, number of candidate predictors, dimensionality of the input data, model complexity, use of transfer learning or regularization, and the planned validation strategy. Therefore, a single universal sample-size threshold cannot be applied equally to conventional regression models, classical machine learning algorithms, and deep neural networks.

Future studies should report the number of patients, lesions or examinations, outcome events in each class, allocation of observations to training, validation, and test sets, and whether data partitioning was performed at the patient level. Authors should also justify the adequacy of the sample size and use independent external or temporal validation whenever possible. Very high performance estimates obtained from small internal test sets should be interpreted cautiously, particularly when confidence intervals, calibration measures, or external validation are unavailable.

4.3. Bias, Fairness, and Equity in AI-Based Breast Cancer Care

Despite the rapidly growing adoption of artificial intelligence in breast cancer diagnostics and precision oncology, important concerns remain regarding bias, fairness, and health equity. AI systems are highly dependent on the quality and representativeness of the datasets used during model development and biased datasets may lead to unequal performance across different patient populations. In medical imaging AI, biases may originate from demographic imbalance, differences in disease prevalence, imaging acquisition protocols, scanner variability and annotation practices, potentially resulting in systematic disparities in model performance [158]. These concerns are particularly important in breast cancer imaging, where breast density, imaging accessibility, socioeconomic factors, and disease presentation may differ substantially across populations.

Algorithmic bias may also emerge when underrepresented demographic groups are insufficiently represented in training datasets. Healthcare AI systems trained predominantly on data from specific racial, geographic, or socioeconomic populations may unintentionally perpetuate existing healthcare inequalities and demonstrate reduced generalizability in underserved patient groups [159]. In oncology, unequal access to high-quality imaging, molecular diagnostics, and digital healthcare infrastructure may further amplify disparities in AI-supported cancer care [160,161]. Consequently, AI systems developed using highly curated datasets from tertiary academic centers may not maintain equivalent performance in lower-resource clinical environments.

Another important concern is the possibility that AI systems unintentionally learn hidden demographic information unrelated to disease biology. Recent studies demonstrated that deep learning algorithms can identify patient race directly from medical imaging data with unexpectedly high accuracy, even when such information is not visually recognizable to clinicians [162]. This finding raises concerns that AI models may rely on hidden surrogate features associated with race, institutional imaging characteristics, or technical acquisition parameters rather than clinically meaningful tumor features.

Bias associated with imaging hardware and acquisition heterogeneity also remains a major challenge in breast imaging AI. As discussed in previous sections, MRI and ultrasound studies frequently involve data acquired using specific scanners, vendors, and

institutional imaging protocols [117–131]. Such dependence may reduce model robustness and disproportionately affect healthcare systems with different technological resources or imaging standards. Therefore, external multicenter validation across diverse patient populations and imaging environments remains essential before broad clinical implementation can be considered.

4.4. Explainable AI and Clinician-AI Collaboration

One of the major barriers limiting the clinical adoption of artificial intelligence in breast cancer care is the limited interpretability of many machine learning and deep learning models. Complex architectures such as convolutional neural networks and ensemble models often function as “black boxes,” meaning that the reasoning behind individual predictions may not be directly understandable to clinicians [13–15,21–24]. In oncology, where diagnostic and therapeutic decisions may have critical consequences for patient outcomes, transparency and explainability are particularly important for establishing clinical trust, accountability, and safe implementation [107,162–164].

To address these limitations, increasing attention has been directed toward XAI methods. In breast cancer imaging and prediction models, explainability techniques such as Grad-CAM, SHAP, saliency maps, attention maps and feature attribution methods are increasingly being used to visualize regions of interest and identify the features contributing most strongly to model predictions [107,162,165]. In mammography, MRI, ultrasound and histopathological imaging, Grad-CAM-based heatmaps may help clinicians understand which image regions influenced AI classification decisions, potentially improving transparency and facilitating human-AI collaboration [165]. Similarly, SHAP-based approaches may provide feature-level interpretability for radiomic, clinical, and multimodal prediction models by estimating the contribution of individual variables to a specific prediction [166].

Explainability may also improve clinician trust and facilitate integration of AI into clinical workflows. Studies evaluating clinician-oriented explanations demonstrated that interpretable outputs may improve physicians’ understanding of AI-generated recommendations and support more informed decision-making processes [165]. In breast imaging, explainability tools may additionally help radiologists identify potential model failures, recognize spurious correlations, and verify whether AI systems focus on clinically meaningful imaging features rather than irrelevant artifacts [107,162,167–169].

However, explainability methods themselves also possess important limitations. Saliency maps and heatmap-based techniques may appear visually convincing despite not always accurately reflecting the true internal reasoning of a model. Similarly, SHAP and LIME explanations may be affected by feature collinearity, instability and sensitivity to input perturbations, potentially leading to misleading interpretations [167]. Therefore, explainability should not be interpreted as proof of model correctness or reliability, but rather as an additional tool supporting transparency and critical evaluation.

Another emerging concept is clinician-AI collaboration through “human-in-the-loop” systems, where AI serves as a decision-support tool rather than a fully autonomous system [107,162,164]. Current evidence suggests that hybrid approaches combining physician expertise with AI assistance may achieve better clinical acceptance and safer implementation than fully automated workflows. This is particularly relevant in breast cancer care, where diagnostic uncertainty, treatment complexity, and individualized patient management require continuous clinical oversight.

Overall, explainable AI represents an important step toward improving the transparency, trustworthiness and clinical usability of AI systems in breast cancer care. Nevertheless, further research is still required to standardize explainability methods, evaluate

their clinical utility, and determine whether interpretable AI systems truly improve patient outcomes and decision-making reliability in real-world oncology practice.

4.5. Emerging AI Approaches in Breast Cancer

Recent advances in artificial intelligence are rapidly expanding beyond conventional convolutional neural network architectures and task-specific machine learning models. Emerging approaches such as ViTs, foundation models, self-supervised learning, federated learning, uncertainty-aware AI, continual learning, and multimodal large models may significantly influence the future development of breast cancer diagnostics and precision oncology [170–177].

One of the most important recent developments is the growing use of transformer-based architectures in medical imaging. Unlike classical CNNs, Vision Transformers are capable of capturing long-range spatial relationships and global contextual information within images, potentially improving performance in complex imaging tasks such as lesion detection, segmentation, and radiogenomic analysis [177]. Transformer-based architectures may be particularly useful in breast MRI and multimodal imaging, where interpretation often requires integration of distributed imaging features across multiple sequences and scales.

Another rapidly developing field involves foundation models and generalist medical AI systems. These models are trained using extremely large multimodal datasets and self-supervised learning strategies, enabling adaptation to multiple downstream clinical tasks with limited task-specific training. Foundation models may eventually integrate imaging, pathology, genomics, electronic health records and clinical narratives into unified predictive systems capable of supporting diagnosis, prognosis, and treatment planning [170].

Self-supervised learning has also emerged as a promising strategy for overcoming one of the major limitations of medical AI, namely the scarcity of large, well-annotated datasets. Instead of relying exclusively on manually labeled data, self-supervised models learn generalized image representations from large unlabeled datasets and can subsequently be fine-tuned for specific clinical applications [171]. This approach may be particularly important in breast imaging, where expert annotation is time-consuming, expensive and often limited by interobserver variability.

Increasing attention is also being directed toward federated learning, which enables collaborative model training across multiple institutions without direct sharing of patient data [172]. Such approaches may improve generalizability while simultaneously addressing privacy concerns and regulatory requirements associated with sensitive medical information. Federated learning may be particularly relevant in breast cancer AI, where multicenter collaboration is essential for obtaining sufficiently diverse datasets while maintaining compliance with GDPR, HIPAA, and other privacy regulations.

Another emerging direction involves synthetic data generation and diffusion models. Diffusion-based generative models may produce realistic synthetic medical images that can potentially augment limited training datasets, improve robustness and reduce overfitting [176]. Synthetic imaging data may additionally help address demographic imbalance and underrepresentation in AI training datasets, although concerns remain regarding realism, hidden artifacts and potential propagation of existing biases.

Model robustness and reliability are also increasingly recognized as critical requirements for safe clinical implementation. Uncertainty-aware AI systems aim to estimate prediction confidence and identify cases where model outputs may be unreliable or ambiguous [174]. Such approaches may improve safety in high-risk clinical environments by enabling clinicians to recognize uncertain predictions and maintain appropriate oversight. Similarly, continual learning approaches aim to allow AI systems to adapt dynam-

ically to evolving clinical data, imaging protocols and population characteristics after deployment [175]. This may be particularly important in breast imaging, where scanner technologies, acquisition protocols and disease patterns continuously evolve over time.

Finally, domain adaptation techniques are increasingly being investigated to address one of the major limitations of current breast imaging AI systems: poor generalizability across institutions and imaging devices. Differences in scanner vendors, acquisition protocols, image quality, and patient populations may substantially reduce external model performance [173]. Domain adaptation methods seek to improve robustness across heterogeneous datasets and may represent an important step toward reliable multicenter clinical deployment.

Despite their considerable promise, most of these emerging AI approaches remain at relatively early stages of clinical translation. Large prospective validation studies, standardized evaluation frameworks, transparent reporting and careful regulatory oversight will remain essential before these technologies can be safely integrated into routine breast cancer care.

4.6. Regulatory and Ethical Landscape

Despite the rapidly growing development of artificial intelligence in breast cancer diagnostics and precision oncology, regulatory approval, ethical oversight and governance frameworks remain major challenges for safe clinical implementation. AI systems intended for medical use are increasingly being recognized as high-risk technologies due to their direct influence on diagnostic and therapeutic decision-making, emphasizing the need for robust regulatory supervision, transparency, accountability, and continuous monitoring [178–184].

The number of FDA-approved AI-based medical devices has increased substantially in recent years, particularly in radiology and medical imaging applications [178]. However, regulatory approval alone does not necessarily guarantee long-term clinical reliability or generalizability. Many AI systems are developed using retrospective datasets and may experience performance degradation after deployment because of domain shift, evolving imaging protocols, demographic variability, and changes in clinical practice [181]. Consequently, regulatory agencies such as the FDA increasingly emphasize the importance of post-market surveillance, continuous performance evaluation, and lifecycle monitoring for AI-based Software as a Medical Device (SaMD) systems [182].

An additional challenge involves the emergence of adaptive and continually learning AI systems. Unlike traditional static medical software, some AI models may evolve after deployment through ongoing exposure to new clinical data. This creates significant regulatory complexity because continuously updating algorithms may change their behavior over time, potentially affecting safety, reproducibility, and clinical reliability [182,184]. Therefore, current regulatory discussions increasingly focus on predefined update protocols, algorithmic transparency, auditability and continuous validation frameworks.

In Europe, the recently introduced European Union Artificial Intelligence Act (EU AI Act) classifies many medical AI applications as high-risk systems requiring strict compliance with transparency, safety, risk management, and human oversight obligations [179]. Similarly, ethical frameworks proposed by the World Health Organization emphasize that AI systems in healthcare should ensure transparency, explainability, equity, privacy protection, and maintenance of physician oversight during clinical decision-making [184]. These principles are particularly relevant in oncology, where AI-assisted decisions may directly influence diagnosis, treatment selection and patient outcomes.

Data governance and patient privacy also remain central concerns in medical AI development. The use of large-scale imaging, genomic and clinical datasets requires compliance

with privacy regulations such as the General Data Protection Regulation (GDPR) and other healthcare data protection frameworks [183]. Federated learning and privacy-preserving AI approaches may partially address these concerns by enabling collaborative multicenter model training without direct sharing of sensitive patient data [172].

Another important issue involves explainability and trustworthiness. Ethical analyses emphasize that clinicians must be able to understand, critically evaluate, and appropriately supervise AI-generated recommendations before these systems can be safely integrated into routine care [180,184]. The lack of transparency in complex deep learning systems may complicate accountability and legal responsibility when diagnostic errors occur. Consequently, questions regarding liability, auditability and allocation of responsibility between clinicians, healthcare institutions and AI developers remain incompletely resolved [184].

5. Conclusions

Artificial intelligence is being investigated across multiple aspects of breast cancer care, extending beyond traditional image classification tasks. Current evidence is strongest for AI-supported mammographic screening, whereas applications in MRI, ultrasound, histopathology, prognostic modeling, treatment-response prediction, and treatment selection remain at varying stages of clinical validation. Machine learning and deep learning algorithms have shown substantial potential in supporting personalized treatment planning and precision oncology approaches.

Recent advances indicate that the field is evolving toward more sophisticated paradigms, including Vision Transformers, foundation models, self-supervised learning, multimodal large models, federated learning, and explainable AI. These approaches may enable the integration of imaging, pathology, genomic, and clinical data within unified predictive frameworks, providing a more comprehensive representation of tumor biology and supporting individualized patient management. Furthermore, privacy-preserving and uncertainty-aware AI systems may contribute to safer and more trustworthy clinical implementation.

Despite these developments, important challenges remain. Many published studies are retrospective, based on relatively small or single-center datasets, and lack rigorous prospective validation. Additional concerns include limited interpretability, domain shift, data heterogeneity, algorithmic bias, regulatory requirements, and integration into existing clinical workflows. Addressing these limitations will be essential before AI can be routinely adopted in everyday oncological practice.

Overall, artificial intelligence should be viewed not as a replacement for clinicians but as a powerful decision-support tool capable of augmenting human expertise. Future research should focus on prospective multicenter studies, robust external validation, transparent and explainable models, and the development of clinically deployable multimodal systems. Such advances may ultimately contribute to earlier diagnosis, more effective treatment selection, improved patient outcomes, and the realization of precision oncology in breast cancer care.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/a19070524/s1>, Table S1: Complete database-specific search strategies used in PubMed/MEDLINE, Scopus, Web of Science and Embase; Table S2. Design-specific risk-of-bias assessment of representative primary clinical studies included in Table 4.

Author Contributions: Conceptualization, D.B.-A., S.C., J.S., A.P., M.X. and D.A.; methodology, D.B.-A., S.C., J.S., A.P., M.X. and D.A.; software, D.B.-A., S.C., J.S., A.P., M.X. and D.A.; validation, D.B.-A., S.C., J.S., A.P., M.X. and D.A.; formal analysis, D.B.-A., S.C., J.S., A.P., M.X. and D.A.; investigation, D.B.-A., S.C., J.S., A.P., M.X. and D.A.; resources, D.B.-A., S.C., J.S., A.P., M.X. and D.A.; data curation,

D.B.-A., S.C., J.S., A.P., M.X. and D.A.; writing—original draft preparation, D.B.-A., S.C., J.S., A.P., M.X. and D.A.; writing—review and editing, D.B.-A., S.C., J.S., A.P., M.X. and D.A.; visualization, D.B.-A., S.C., J.S., A.P., M.X. and D.A.; supervision, D.A.; project administration, D.A.; funding acquisition, D.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

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

References

1. Wu, J.; Fan, D.; Shao, Z.; Xu, B.; Ren, G.; Jiang, Z.; Wang, Y.; Jin, F.; Zhang, J.; Zhang, Q.; et al. CACA Guidelines for Holistic Integrative Management of Breast Cancer. *Holist. Integr. Oncol.* **2022**, *1*, 7. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Bray, F.; Laversanne, M.; Sung, H.; Ferlay, J.; Siegel, R.L.; Soerjomataram, I.; Jemal, A. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2024**, *74*, 229–263. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Siegel, R.L.; Giaquinto, A.N.; Jemal, A. Cancer statistics, 2024. *CA Cancer J. Clin.* **2024**, *74*, 12–49. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Xiong, X.; Zheng, L.W.; Ding, Y.; Chen, Y.F.; Cai, Y.W.; Wang, L.P.; Huang, L.; Liu, C.C.; Shao, Z.M.; Yu, K.D. Breast cancer: Pathogenesis and treatments. *Signal Transduct. Target. Ther.* **2025**, *10*, 49. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Milosevic, M.; Jankovic, D.; Milenkovic, A.; Stojanov, D. Early diagnosis and detection of breast cancer. *Technol. Health Care* **2018**, *26*, 729–759. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Huang, S.; Westvold, S.J.; Soulos, P.R.; Fan, J.; Winer, E.P.; Zhan, H.; Lustberg, M.B.; Lewin, J.; Robinson, T.J.; Dinan, M.A. Screening history, stage at diagnosis, and mortality in screen-detected breast cancer. *JAMA Netw. Open* **2025**, *8*, e255322. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Henderson, J.T.; Webber, E.M.; Weyrich, M.S.; Miller, M.; Melnikow, J. Screening for breast cancer: Evidence report and systematic review for the US Preventive Services Task Force. *JAMA* **2024**, *331*, 1931–1946. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Rajpurkar, P.; Chen, E.; Banerjee, O.; Topol, E.J. AI in health and medicine. *Nat. Med.* **2022**, *28*, 31–38. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Al Kuwaiti, A.; Nazer, K.; Al-Reedy, A.; Al-Shehri, S.; Al-Muhanna, A.; Subbarayalu, A.V.; Al Muhanna, D.; Al-Muhanna, F.A. A review of the role of artificial intelligence in healthcare. *J. Pers. Med.* **2023**, *13*, 951. [\[CrossRef\]](#)
10. Umaphathy, V.R.; Rajinikanth, B.S.; Samuel Raj, R.D.; Yadav, S.; Munavarah, S.A.; Anandapandian, P.A.; Mary, A.V.; Padmavathy, K.; R, A. Perspective of artificial intelligence in disease diagnosis: A review of current and future endeavours in the medical field. *Cureus* **2023**, *15*, e45684. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Popover, J.L.; Wallace, S.P.; Feldman, J.; Chastain, G.; Kalathia, C.; Imam, A.; Almasri, M.; Toomey, P.G. Artificial intelligence in medicine: A specialty-level overview of emerging AI trends. *JSLs* **2025**, *29*, e2025.00041. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Baklola, M.; Reda Elmahdi, R.; Ali, S.; Elshenawy, M.; Mohamed Mossad, A.; Al-Bawah, N.; Mohamed Mansour, R. Artificial intelligence in disease diagnostics: A comprehensive narrative review of current advances, applications, and future challenges in healthcare. *Ann. Med. Surg.* **2025**, *87*, 4237–4245. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Van Belle, V.; Van Calster, B.; Van Huffel, S.; Suykens, J.A.K.; Lisboa, P. Explaining support vector machines: A color based nomogram. *PLoS ONE* **2016**, *11*, e0164568. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Gaonkar, B.; Shinohara, R.T.; Davatzikos, C.; Alzheimer’s Disease Neuroimaging Initiative. Interpreting support vector machine models for multivariate group-wise analysis in neuroimaging. *Med. Image Anal.* **2015**, *24*, 190–204. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Rodríguez-Pérez, R.; Bajorath, J. Evolution of support vector machine and regression modeling in chemoinformatics and drug discovery. *J. Comput. Aided Mol. Des.* **2022**, *36*, 355–362. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Couronné, R.; Probst, P.; Boulesteix, A.L. Random forest versus logistic regression: A large-scale benchmark experiment. *BMC Bioinform.* **2018**, *19*, 270. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Strobl, C.; Boulesteix, A.L.; Zeileis, A.; Hothorn, T. Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinform.* **2007**, *8*, 25. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Ehsani, R.; Drabløs, F. Robust distance measures for kNN classification of cancer data. *Cancer Inform.* **2020**, *19*, 1176935120965542. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Ranganathan, P.; Pramesh, C.S.; Aggarwal, R. Common pitfalls in statistical analysis: Logistic regression. *Perspect. Clin. Res.* **2017**, *8*, 148–151. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Sperandei, S. Understanding logistic regression analysis. *Biochem. Med.* **2014**, *24*, 12–18. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Roy, D.; Panda, P.; Roy, K. Tree-CNN: A hierarchical deep convolutional neural network for incremental learning. *Neural Netw.* **2020**, *121*, 148–160. [\[CrossRef\]](#) [\[PubMed\]](#)

22. Rani, S.; Memoria, M.; Almogren, A.; Bharany, S.; Joshi, K.; Altameem, A.; Rehman, A.U.; Hamam, H. Deep learning to combat knee osteoarthritis and severity assessment by using CNN-based classification. *BMC Musculoskelet. Disord.* **2024**, *25*, 817. [[CrossRef](#)] [[PubMed](#)]
23. Shah, A.A.; Malik, H.A.M.; Muhammad, A.; Alourani, A.; Butt, Z.A. Deep learning ensemble 2D CNN approach towards the detection of lung cancer. *Sci. Rep.* **2023**, *13*, 2987. [[CrossRef](#)] [[PubMed](#)]
24. Yang, Z.; Tong, K.; Jin, S.; Wang, S.; Yang, C.; Jiang, F. CNN-Siam: Multimodal siamese CNN-based deep learning approach for drug–drug interaction prediction. *BMC Bioinform.* **2023**, *24*, 110. [[CrossRef](#)] [[PubMed](#)]
25. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
26. Wang, Z.; Peng, Y.; Li, D.; Guo, Y.; Zhang, B. MMNet: A multi-scale deep learning network for the left ventricular segmentation of cardiac MRI images. *Appl. Intell.* **2022**, *52*, 5225–5240.
27. Wang, C.; Wang, Z.; Xi, W.; Yang, Z.; Bai, G.; Wang, R.; Duan, M. MufiNet: Multiscale fusion residual networks for medical image segmentation. In *Proceedings of the International Joint Conference on Neural Networks (IJCNN)*, Glasgow, UK, 19–24 July 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–8.
28. Gao, S.H.; Cheng, M.M.; Zhao, K.; Zhang, X.Y.; Yang, M.H.; Torr, P. Res2Net: A new multi-scale backbone architecture. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *43*, 652–662.
29. Jameela, T.; Athotha, K.; Singh, N.; Gunjan, V.K.; Kahali, S. Deep learning and transfer learning for malaria detection. *Comput. Intell. Neurosci.* **2022**, *2022*, 2221728. [[CrossRef](#)] [[PubMed](#)]
30. Maity, R.; Raja Sankari, V.M.; Snehalatha, U.; Velu, S.; Alahmadi, T.J.; Alhababi, Z.A.; Alkahtani, H.K. Early detection of Alzheimer’s disease in structural and functional MRI. *Front. Med.* **2024**, *11*, 1520878. [[CrossRef](#)] [[PubMed](#)]
31. Feng, S.; Zhang, R.; Zhang, W.; Yang, Y.; Song, A.; Chen, J.; Wang, F.; Xu, J.; Liang, C.; Liang, X.; et al. Predicting acute exacerbation phenotype in chronic obstructive pulmonary disease patients using VGG-16 deep learning. *Respiration* **2025**, *104*, 1–14. [[CrossRef](#)] [[PubMed](#)]
32. Ali, R.; Lei, R.; Shi, H.; Xu, J. Cranio-maxillofacial post-operative face prediction by deep spatial multiband VGG-NET CNN. *Am. J. Transl. Res.* **2022**, *14*, 2527–2539. [[PubMed](#)]
33. Alruwaili, M.; Mohamed, M. An integrated deep learning model with EfficientNet and ResNet for accurate multi-class skin disease classification. *Diagnostics* **2025**, *15*, 551. [[CrossRef](#)] [[PubMed](#)]
34. Atila, Ü.; Uçar, M.; Akyol, K.; Uçar, E. Plant leaf disease classification using EfficientNet deep learning model. *Ecol. Inform.* **2021**, *61*, 101182. [[CrossRef](#)]
35. Talukder, M.A.; Layek, M.A.; Kazi, M.; Uddin, M.A.; Aryal, S. Empowering COVID-19 detection: Optimizing performance through fine-tuned EfficientNet deep learning architecture. *Comput. Biol. Med.* **2024**, *168*, 107789. [[CrossRef](#)] [[PubMed](#)]
36. Zhuang, F.; Qi, Z.; Duan, K.; Xi, D.; Zhu, Y.; Zhu, H.; Xiong, H.; He, Q. A comprehensive survey on transfer learning. *Proc. IEEE* **2020**, *109*, 43–76. [[CrossRef](#)]
37. Gupta, J.; Pathak, S.; Kumar, G. Deep learning (CNN) and transfer learning: A review. *J. Phys. Conf. Ser.* **2022**, *2273*, 012012. [[CrossRef](#)]
38. Iman, M.; Arabnia, H.R.; Rasheed, K. A review of deep transfer learning and recent advancements. *Technologies* **2023**, *11*, 40. [[CrossRef](#)]
39. Jeny, A.A.; Hamzehei, S.; Jin, A.; Baker, S.A.; Van Rathe, T.; Bai, J.; Yang, C.; Nabavi, S. Hybrid transformer-based model for mammogram classification by integrating prior and current images. *Med. Phys.* **2025**, *52*, 2999–3014. [[CrossRef](#)] [[PubMed](#)]
40. Kassis, I.; Lederman, D.; Ben-Arie, G.; Giladi Rosenthal, M.; Shelef, I.; Zigel, Y. Detection of breast cancer in digital breast tomosynthesis with vision transformers. *Sci. Rep.* **2024**, *14*, 22149. [[CrossRef](#)] [[PubMed](#)]
41. Firouzbakht, M.; Amirmazlaghani, M. Breast cancer detection in mammography images using Neighborhood Attention transformer and Shearlet Transform. *Comput. Biol. Med.* **2025**, *198*, 111239. [[CrossRef](#)] [[PubMed](#)]
42. Paschali, M.; Chen, Z.; Blankemeier, L.; Varma, M.; Youssef, A.; Bluethgen, C.; Langlotz, C.; Gatidis, S.; Chaudhari, A. Foundation models in radiology: What, how, why, and why not. *Radiology* **2025**, *314*, e240597. [[CrossRef](#)] [[PubMed](#)]
43. Wu, C.; Zhang, X.; Zhang, Y.; Hui, H.; Wang, Y.; Xie, W. Towards generalist foundation model for radiology by leveraging web-scale 2D and 3D medical data. *Nat. Commun.* **2025**, *16*, 7866. [[CrossRef](#)] [[PubMed](#)]
44. Vorontsov, E.; Bozkurt, A.; Casson, A.; Shaikovski, G.; Zelechowski, M.; Severson, K.; Zimmermann, E.; Hall, J.; Tenenholtz, N.; Fusi, N.; et al. A foundation model for clinical-grade computational pathology and rare cancers detection. *Nat. Med.* **2024**, *30*, 2924–2935. [[CrossRef](#)] [[PubMed](#)]
45. Miller, J.D.; Arasu, V.A.; Pu, A.X.; Margolies, L.R.; Sieh, W.; Shen, L. Self-supervised deep learning to enhance breast cancer detection on screening mammography. *arXiv* **2022**, arXiv:2203.08812.
46. Figueiras, H.; Domingues, J.; Matela, N.; Garcia, N. Self-supervised learning for breast cancer detection: A review. *Comput. Biol. Med.* **2025**, *198*, 111245. [[CrossRef](#)] [[PubMed](#)]

47. Wang, L. Self-supervised learning and transformer-based technologies in breast cancer imaging. *Front. Radiol.* **2025**, *5*, 1684436. [[CrossRef](#)] [[PubMed](#)]
48. Luo, L.; Wu, M.; Li, M.; Xin, Y.; Wang, Q.; Vardhanabhuti, V.; Chu, W.C.; Li, Z.; Zhou, J.; Rajpurkar, P.; et al. A large model for non-invasive and personalized management of breast cancer from multiparametric MRI. *Nat. Commun.* **2025**, *16*, 3647. [[CrossRef](#)] [[PubMed](#)]
49. Tzortzis, I.N.; Gutierrez-Torre, A.; Sykiotis, S.; Agulló, F.; Bakalos, N.; Doulamis, A.; Doulamis, N.; Berral, J.L. Towards generalizable federated learning in medical imaging: A real-world case study on mammography data. *Comput. Struct. Biotechnol. J.* **2025**, *28*, 106–117. [[CrossRef](#)] [[PubMed](#)]
50. Shukla, S.; Rajkumar, S.; Sinha, A.; Esha, M.; Elango, K.; Sampath, V. Federated learning with differential privacy for breast cancer diagnosis enabling secure data sharing and model integrity. *Sci. Rep.* **2025**, *15*, 13061. [[CrossRef](#)] [[PubMed](#)] [[PubMed Central](#)]
51. Koetzier, L.R.; Wu, J.; Mastrodicasa, D.; Lutz, A.; Chung, M.; Koszek, W.A.; Pratap, J.; Chaudhari, A.S.; Rajpurkar, P.; Lungren, M.P.; et al. Generating synthetic data for medical imaging. *Radiology* **2024**, *312*, e232471. [[CrossRef](#)] [[PubMed](#)]
52. Montoya-Del-Angel, R.; Sam-Millan, K.; Vilanova, J.C.; Martí, R. MAM-E: Mammographic synthetic image generation with diffusion models. *Sensors* **2024**, *24*, 2076. [[CrossRef](#)] [[PubMed](#)]
53. You, Y.; Zhuang, C.; Gan, H.S.; Rulaningtyas, R.; Ramlee, M.H.; Wahab, A.A. BreastDiff: A multi-condition guided diffusion model for breast cancer classification in diverse modalities. *Biomed. Signal Process. Control* **2026**, *112*, 108708. [[CrossRef](#)]
54. Akbari, Y.; Abdullakutty, F.; Al Maadeed, S.; Bouridane, A.; Hamoudi, R. Breast cancer detection based on histological images using fusion of diffusion model outputs. *Sci. Rep.* **2025**, *15*, 21463. [[CrossRef](#)] [[PubMed](#)]
55. Chegini, M.; Mahloojifar, A. Uncertainty-aware deep learning-based CAD system for breast cancer classification using ultrasound and mammography images. *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.* **2024**, *12*, 2297983. [[CrossRef](#)]
56. Chegini, M.; Mahloojifar, A. Reliable breast cancer molecular subtype prediction based on uncertainty-aware Bayesian deep learning by mammography. *arXiv* **2024**, arXiv:2412.11953.
57. Hespanhol, L.; Vallio, C.S.; Costa, L.M.; Saragiotto, B.T. Understanding and Interpreting Confidence and Credible Intervals around Effect Estimates. *Braz. J. Phys. Ther.* **2019**, *23*, 290–301. [[CrossRef](#)] [[PubMed](#)]
58. Kurz, A.; Hauser, K.; Mehrtens, H.A.; Krieghoff-Henning, E.; Hekler, A.; Kather, J.N.; Fröhling, S.; von Kalle, C.; Brinker, T.J. Uncertainty Estimation in Medical Image Classification: Systematic Review. *JMIR Med. Inform.* **2022**, *10*, e36427. [[CrossRef](#)] [[PubMed](#)]
59. Shen, Y.; Wu, N.; Phang, J.; Park, J.; Liu, K.; Tyagi, S.; Geras, K.J. An interpretable classifier for high-resolution breast cancer screening images utilizing weakly supervised localization. *Med. Image Anal.* **2021**, *68*, 101908. [[CrossRef](#)] [[PubMed](#)]
60. Sajid, M.Z.; Hamid, M.F.; Qureshi, I. Explainable and uncertainty-aware ensemble framework with causal analysis for breast cancer detection. *Front. Oncol.* **2026**, *15*, 1751090. [[CrossRef](#)] [[PubMed](#)]
61. da Câmara Ribeiro-Dantas, M.; Li, H.; Cabeli, V.; Dupuis, L.; Simon, F.; Hettal, L.; Isambert, H. Learning interpretable causal networks from very large datasets: Application to 400,000 medical records of breast cancer patients. *iScience* **2024**, *27*, 109736. [[CrossRef](#)]
62. Chen, D.; Zhao, H.; He, J.; Pan, Q.; Zhao, W. A causal XAI diagnostic model for breast cancer based on mammography reports. In Proceedings of the 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Houston, TX, USA, 9–12 December 2021; pp. 3341–3349. [[CrossRef](#)]
63. Qazi, M.A.; Hashmi, A.U.R.; Sanjeev, S.; Almakky, I.; Saeed, N.; Gonzalez, C.; Yaqub, M. Continual learning in medical imaging: A survey and practical analysis. *ACM Comput. Surv.* **2026**, *58*, 1–25. [[CrossRef](#)]
64. Li, H.; Whitney, H.M.; Ji, Y.; Edwards, A.; Papaioannou, J.; Liu, P.; Giger, M.L. Impact of continuous learning on diagnostic breast MRI AI: Evaluation on an independent clinical dataset. *J. Med. Imaging* **2022**, *9*, 034502. [[CrossRef](#)] [[PubMed](#)]
65. Abimouloud, M.L.; Bensid, K.; Elleuch, M.; Aiadi, O.; Kherallah, M. Vision transformer-convolution for breast cancer classification using mammography images: A comparative study. *Int. J. Hybrid Intell. Syst.* **2024**, *20*, 67–83. [[CrossRef](#)]
66. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An image is worth 16×16 words: Transformers for image recognition at scale. In Proceedings of the International Conference on Learning Representations (ICLR), Vienna, Austria, 3–7 May 2021.
67. Zhou, G.; Mosadegh, B. Distilling knowledge from an ensemble of vision transformers for improved classification of breast ultrasound. *Acad. Radiol.* **2024**, *31*, 104–120. [[CrossRef](#)] [[PubMed](#)]
68. Ayana, G.; Choe, S.W. BUViTNet: Breast ultrasound detection via Vision Transformers. *Diagnostics* **2022**, *12*, 2654. [[CrossRef](#)] [[PubMed](#)]
69. Fontes, J.P.P.; Raimundo, J.N.C.; Magalhães, L.G.M.; Lopez, M.A.G. Accurate phenotyping of luminal A breast cancer in magnetic resonance imaging: A new 3D CNN approach. *Comput. Biol. Med.* **2025**, *189*, 109903. [[CrossRef](#)] [[PubMed](#)]
70. Zhou, J.; Luo, L.Y.; Dou, Q.; Chen, H.; Chen, C.; Li, G.J.; Jiang, Z.F.; Heng, P.A. Weakly supervised 3D deep learning for breast cancer classification and localization of the lesions in MR images. *J. Magn. Reson. Imaging* **2019**, *50*, 1144–1151. [[CrossRef](#)] [[PubMed](#)]

71. Wang, H.; Cao, J.; Feng, J.; Xie, Y.; Yang, D.; Chen, B. Mixed 2D and 3D convolutional network with multi-scale context for lesion segmentation in breast DCE-MRI. *Biomed. Signal Process. Control* **2021**, *68*, 102607. [\[CrossRef\]](#)
72. Liu, G.; Mitra, D.; Jones, E.F.; Franc, B.L.; Behr, S.C.; Nguyen, A.; Bolouri, M.S.; Wisner, D.J.; Joe, B.N.; Esserman, L.J.; et al. Mask-guided convolutional neural network for breast tumor prognostic outcome prediction on 3D DCE-MR images. *J. Digit. Imaging* **2021**, *34*, 630–636. [\[CrossRef\]](#) [\[PubMed\]](#)
73. Haarbuerger, C.; Baumgartner, M.; Truhn, D.; Broeckmann, M.; Schneider, H.; Schrading, S.; Merhof, D. Multi scale curriculum CNN for context-aware breast MRI malignancy classification. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2019), Shenzhen, China, 13–17 October 2019; pp. 495–503. [\[CrossRef\]](#)
74. Kuang, S.; Woodruff, H.C.; Granzier, R.; van Nijmegen, T.J.; Lobbes, M.B.; Smidt, M.L.; Mehrkanoon, S. MSCDA: Multi-level semantic-guided contrast improves unsupervised domain adaptation for breast MRI segmentation in small datasets. *Neural Netw.* **2023**, *165*, 119–134. [\[CrossRef\]](#) [\[PubMed\]](#)
75. Akyüz, U.; Katircioglu-Öztürk, D.; Süslü, E.K.; Keleş, B.; Kaya, M.C.; Durhan, G.; Akar, G.B. DoSRMC: Domain shift resilient mammography classification using batch normalization adaptation. *arXiv* **2025**, arXiv:2508.15452.
76. Yoon, J.S.; Oh, K.; Shin, Y.; Mazurowski, M.A.; Suk, H.I. Domain generalization for medical image analysis: A review. *Proc. IEEE* **2024**, *112*, 1583–1609. [\[CrossRef\]](#)
77. Matta, S.; Lamard, M.; Zhang, P.; Le Guilcher, A.; Borderie, L.; Cochener, B.; Quéllec, G. A systematic review of generalization research in medical image classification. *Comput. Biol. Med.* **2024**, *183*, 109256. [\[CrossRef\]](#) [\[PubMed\]](#)
78. Garrucho, L.; Joshi, S.; Kushibar, K.; Osuala, R.; Bobowicz, M.; Bargalló, X.; Lekadir, K. The MAMA-MIA Challenge: Advancing generalizability and fairness in breast MRI tumor segmentation and treatment response prediction. *arXiv* **2026**, arXiv:2603.01250.
79. Pan, H.; Durak, G.; Aktas, H.E.; Bejar, A.M.; Tutun, B.; Uysal, E.; Bagci, U. LUMINA: A multi-vendor mammography benchmark with energy harmonization protocol. *arXiv* **2026**, arXiv:2603.14644.
80. Tayebi Arasteh, S.; Kuhl, C.; Saehn, M.J.; Isfort, P.; Truhn, D.; Nebelung, S. Enhancing domain generalization in the AI-based analysis of chest radiographs with federated learning. *Sci. Rep.* **2023**, *13*, 22576. [\[CrossRef\]](#) [\[PubMed\]](#)
81. Bakalo, R.; Ben-Ari, R.; Goldberger, J. Classification and detection in mammograms with weak supervision via dual branch deep neural net. In Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), Venice, Italy, 8–11 April 2019; pp. 1905–1909. [\[CrossRef\]](#)
82. Bakalo, R.; Goldberger, J.; Ben-Ari, R. Weakly and semi supervised detection in medical imaging via deep dual branch net. *Neurocomputing* **2021**, *421*, 15–25. [\[CrossRef\]](#)
83. Liang, G.; Wang, X.; Zhang, Y.; Jacobs, N. Weakly-supervised self-training for breast cancer localization. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2020), Montreal, QC, Canada, 20–24 July 2020; pp. 1124–1127. [\[CrossRef\]](#) [\[PubMed\]](#)
84. Liu, K.; Shen, Y.; Wu, N.; Chłędowski, J.; Fernandez-Granda, C.; Geras, K.J. Weakly-supervised high-resolution segmentation of mammography images for breast cancer diagnosis. *Proc. Mach. Learn. Res.* **2021**, *143*, 268–278. [\[PubMed\]](#)
85. Zhang, M.; Wang, C.; Cai, L.; Zhao, J.; Xu, Y.; Xing, J.; Zhang, Y. Developing a weakly supervised deep learning framework for breast cancer diagnosis with HR status based on mammography images. *Comput. Struct. Biotechnol. J.* **2023**, *22*, 17–26. [\[CrossRef\]](#) [\[PubMed\]](#)
86. Tang, Y.; Cao, Z.; Zhang, Y.; Yang, Z.; Ji, Z.; Wang, Y.; Chang, P. Leveraging large-scale weakly labeled data for semi-supervised mass detection in mammograms. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2021), Virtual Conference, 19–25 June 2021; pp. 3855–3864. [\[CrossRef\]](#)
87. Chen, H.; Martel, A.L. Enhancing breast cancer detection on screening mammogram using self-supervised learning and a hybrid deep model of Swin Transformer and convolutional neural networks. *J. Med. Imaging* **2025**, *12*, S22007. [\[CrossRef\]](#) [\[PubMed\]](#)
88. Zhou, S.; Wu, L.; Xiao, C.; Bhatia, P.; Kass-Hout, T. MammoDINO: Anatomically aware self-supervision for mammographic images. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2026), Barcelona, Spain, 4–8 May 2026; pp. 8182–8186. [\[CrossRef\]](#)
89. Jiao, L.; Wang, S.; Zhang, Y.; Liu, H.; Li, X.; Zhao, W.; Chen, H. Foundation Models Meet Medical Image Interpretation. *Research* **2026**, *9*, 1024. [\[CrossRef\]](#) [\[PubMed\]](#)
90. D’Antonoli, T.A.; Bluethgen, C.; Cuocolo, R.; Klontzas, M.E.; Ponsiglione, A.; Kocak, B. Foundation models for radiology: Fundamentals, applications, opportunities, challenges, risks, and prospects. *Diagn. Interv. Radiol.* **2026**, *32*, 259–272.
91. Tavakoli, N.; Shakeri, Z.; Gowda, V.; Samsel, K.; Bedayat, A.; Ghasemiesfe, A.; Rahsepar, A.A. Generative AI and foundation models in radiology: Applications, opportunities, and potential challenges. *Radiology* **2025**, *317*, e242961. [\[CrossRef\]](#) [\[PubMed\]](#)
92. Ryu, J.S.; Kang, H.; Chu, Y.; Yang, S. Vision-language foundation models for medical imaging: A review of current practices and innovations. *Biomed. Eng. Lett.* **2025**, *15*, 809–830. [\[CrossRef\]](#) [\[PubMed\]](#)
93. van Veldhuizen, V.; Botha, V.; Lu, C.; Cesur, M.E.; Lipman, K.G.; de Jong, E.D.; Teuwen, J. Foundation models in medical imaging: A review and outlook. *arXiv* **2025**, arXiv:2506.09095.
94. Tizhoosh, H.R. Beyond the failures: Rethinking foundation models in pathology. *arXiv* **2025**, arXiv:2510.23807.

95. Niu, C.; Wu, P.; De Man, B.; Wang, G. Foundation models for medical imaging: Status, challenges, and directions. *arXiv* **2026**, arXiv:2602.15913.
96. Nakach, F.Z.; Idri, A.; Goceri, E. A comprehensive investigation of multimodal deep learning fusion strategies for breast cancer classification. *Artif. Intell. Rev.* **2024**, *57*, 327. [[CrossRef](#)]
97. Li, T.; Song, S.; Pan, Y.; Song, W.; Fong, S.; Gao, J.; Wang, Q.; Zhang, X.; Mohammed, S. Deep learning in multi-modal breast cancer data fusion: A literature review. *Quant. Imaging Med. Surg.* **2025**, *15*, 11578–11610. [[CrossRef](#)] [[PubMed](#)]
98. Krasniqi, E.; Filomeno, L.; Arcuri, T.; Ferretti, G.; Gasparro, S.; Fulvi, A.; Vici, P. Multimodal deep learning for predicting neoadjuvant treatment outcomes in breast cancer: A systematic review. *Biol. Direct* **2025**, *20*, 72. [[CrossRef](#)] [[PubMed](#)]
99. Eskreis-Winkler, S.; Vega, F.S.; Kohli, A.; Moiso, E.; Alto, M.; Fong, C.; Razavi, P. Multimodal analyses of clinical, radiology, pathology and genomic information for enhanced prediction of response to neoadjuvant therapy in breast cancer. *Clin. Cancer Res.* **2025**, *31*, P1-03-16. [[CrossRef](#)]
100. 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](#)]
101. Zhang, T.; Tan, T.; Han, L.; Appelman, L.; Veltman, J.; Wessels, R.; Duvivier, K.M.; Loo, C.; Gao, Y.; Wang, X.; et al. Predicting breast cancer types on and beyond molecular level in a multi-modal fashion. *NPJ Breast Cancer* **2023**, *9*, 16. [[CrossRef](#)] [[PubMed](#)]
102. Alhussan, A.A.; Nhidi, W.; Filali, I.; Benhmida, F.; Ejbal, R. Federated learning architecture for 3D breast cancer image classification. *Cancers* **2025**, *17*, 3450. [[CrossRef](#)] [[PubMed](#)]
103. Gupta, C.; Gill, N.S.; Gulia, P.; Alduaiji, N.; Shreyas, J.; Shukla, P.K. Applying YOLOv6 as an ensemble federated learning framework to classify breast cancer pathology images. *Sci. Rep.* **2025**, *15*, 3769. [[CrossRef](#)] [[PubMed](#)]
104. Babu, G.M.; Wong, K.W.; Parry, J. Federated learning for digital pathology: A pilot study. *Procedia Comput. Sci.* **2022**, *207*, 736–743. [[CrossRef](#)]
105. Schoenpflug, L.A.; Nie, Y.; Sheikhzadeh, F.; Koelzer, V.H. A review on federated learning in computational pathology. *Comput. Struct. Biotechnol. J.* **2024**, *23*, 3938–3945. [[CrossRef](#)] [[PubMed](#)]
106. Ghosh, D.; Mehjabin, M.; Rayed, M.E.; Mridha, M.F.; Kabir, M.M. Advancements and challenges of federated learning in medical imaging: A systematic literature review. *Artif. Intell. Rev.* **2026**, in press.
107. Ghasemi, A.; Hashtarkhani, S.; Schwartz, D.L.; Shaban-Nejad, A. Explainable artificial intelligence in breast cancer detection and risk prediction: A systematic scoping review. *Cancer Innov.* **2024**, *3*, e136. [[CrossRef](#)] [[PubMed](#)]
108. Ameen, M.; Alshamrani, A.; Alghamdi, M.; Alzahrani, N.; Alsubaie, M. Explainable mammogram analysis with EfficientNetV2 and Grad-CAM++ for breast cancer diagnosis. *Diagnostics* **2025**, *16*, 105. [[CrossRef](#)] [[PubMed](#)]
109. Calisto, F.M.; Abrantes, J.M.; Santiago, C.; Nunes, N.J.; Nascimento, J.C. Personalized explanations for clinician-AI interaction in breast imaging diagnosis by adapting communication to expertise levels. *Int. J. Hum.-Comput. Stud.* **2025**, *197*, 103444. [[CrossRef](#)]
110. Dembrower, K.; Crippa, A.; Colón, E.; Eklund, M.; Strand, F.; ScreenTrustCAD Trial Consortium. Artificial intelligence for breast cancer detection in screening mammography in Sweden: A prospective, population-based, paired-reader, non-inferiority study. *Lancet Digit. Health* **2023**, *5*, e703–e711. [[CrossRef](#)] [[PubMed](#)]
111. Hernström, V.; Josefsson, V.; Sartor, H.; Schmidt, D.; Larsson, A.M.; Hofvind, S.; Andersson, I.; Rosso, A.; Hagberg, O.; Lång, K. Screening performance and characteristics of breast cancer detected in the Mammography Screening with Artificial Intelligence trial (MASAI): A randomised, controlled, parallel-group, non-inferiority, single-blinded, screening accuracy study. *Lancet Digit. Health* **2025**, *7*, e175–e183. [[CrossRef](#)] [[PubMed](#)]
112. Eiseemann, N.; Bunk, S.; Mukama, T.; Baltus, H.; Elsner, S.A.; Gomille, T.; Hecht, G.; Heywang-Köbrunner, S.; Rathmann, R.; Siegmann-Luz, K.; et al. Nationwide real-world implementation of AI for cancer detection in population-based mammography screening. *Nat. Med.* **2025**, *31*, 917–924. [[CrossRef](#)] [[PubMed](#)]
113. Gommers, J.; Hernström, V.; Josefsson, V.; Sartor, H.; Schmidt, D.; Hjelmgren, A.; Larsson, A.M.; Hofvind, S.; Andersson, I.; Rosso, A.; et al. Interval cancer, sensitivity, and specificity comparing AI-supported mammography screening with standard double reading without AI in the MASAI study: A randomised, controlled, non-inferiority, single-blinded, population-based, screening-accuracy trial. *Lancet* **2026**, *407*, 505–514. [[CrossRef](#)] [[PubMed](#)]
114. Friedewald, S.M.; Sieniek, M.; Jansen, S.; Mahvar, F.; Kohlberger, T.; Schacht, D.; Bhole, S.; Gupta, D.; Prabhakara, S.; McKinney, S.M.; et al. Triaging mammography with artificial intelligence: An implementation study. *Breast Cancer Res. Treat.* **2025**, *211*, 1–10. [[CrossRef](#)] [[PubMed](#)]
115. Rangarajan, K.; Manivannan, V.V.; Singh, H.; Gupta, A.; Maheshwari, H.; Gogoi, R.; Gogoi, D.; Das, R.J.; Hari, S.; Vyas, S.; et al. Simulation training in mammography with AI-generated images: A multireader study. *Eur. Radiol.* **2025**, *35*, 562–571. [[CrossRef](#)] [[PubMed](#)]
116. Woode, M.E.; De Silva Perera, U.; Degeling, C.; Aquino, Y.S.J.; Houssami, N.; Carter, S.M.; Chen, G. Preferences for the use of artificial intelligence for breast cancer screening in Australia: A discrete choice experiment. *Patient* **2025**, *18*, 495–510. [[CrossRef](#)] [[PubMed](#)]

117. Liang, Y.; Wei, Z.; Dai, Y.; Chen, X.; Du, S.; Wong, C.; Xu, Z.; Gao, W.; Han, C.; Chen, K.; et al. An interpretable AI system reduces false-positive MRI diagnoses by stratifying high-risk breast lesions. *Nat. Commun.* **2026**, *17*, 2263. [[CrossRef](#)] [[PubMed](#)]
118. Yin, H.L.; Jiang, Y.; Xu, Z.; Jia, H.H.; Lin, G.W. Combined diagnosis of multiparametric MRI-based deep learning models facilitates differentiating triple-negative breast cancer from fibroadenoma magnetic resonance BI-RADS 4 lesions. *J. Cancer Res. Clin. Oncol.* **2023**, *149*, 2575–2584. [[CrossRef](#)] [[PubMed](#)]
119. Zhang, J.; Zhan, C.; Zhang, C.; Song, Y.; Yan, X.; Guo, Y.; Ai, T.; Yang, G. Fully automatic classification of breast lesions on multi-parameter MRI using a radiomics model with minimal number of stable, interpretable features. *Radiol. Med.* **2023**, *128*, 160–170. [[CrossRef](#)] [[PubMed](#)]
120. Yin, H.; Bai, L.; Jia, H.; Lin, G. Noninvasive assessment of breast cancer molecular subtypes on multiparametric MRI using convolutional neural network with transfer learning. *Thorac. Cancer* **2022**, *13*, 3183–3191. [[CrossRef](#)] [[PubMed](#)]
121. Zhu, Z.; Albadawy, E.; Saha, A.; Zhang, J.; Harowicz, M.R.; Mazurowski, M.A. Deep learning for identifying radiogenomic associations in breast cancer. *Comput. Biol. Med.* **2019**, *109*, 85–90. [[CrossRef](#)] [[PubMed](#)]
122. Meng, M.; Zhang, M.; Shen, D.; He, G. Differentiation of breast lesions on dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) using deep transfer learning based on DenseNet201. *Medicine* **2022**, *101*, e31214. [[CrossRef](#)] [[PubMed](#)]
123. Lee, E.J.; Chang, Y.W.; Sung, J.K.; Thomas, B. Feasibility of deep learning k-space-to-image reconstruction for diffusion weighted imaging in patients with breast cancers: Focus on image quality and reduced scan time. *Eur. J. Radiol.* **2022**, *157*, 110608. [[CrossRef](#)] [[PubMed](#)]
124. Zhu, Z.; Harowicz, M.; Zhang, J.; Saha, A.; Grimm, L.J.; Hwang, E.S.; Mazurowski, M.A. Deep learning analysis of breast MRIs for prediction of occult invasive disease in ductal carcinoma in situ. *Comput. Biol. Med.* **2019**, *115*, 103498. [[CrossRef](#)] [[PubMed](#)]
125. Liu, H.; Zhang, Y.; Tan, B.; Yin, Y.F.; Yan, L.X.; Xiang, L.H.; Shan, D.D.; Zhang, Y.Y.; Ding, S.S.; Xu, G.; et al. Deep learning based on automated breast volume scanner images for the diagnosis of breast lesions: A multicenter diagnostic study. *Int. J. Med. Sci.* **2025**, *22*, 3924–3937. [[CrossRef](#)] [[PubMed](#)]
126. Cai, L.; Pfof, A.; Barr, R.G.; Duda, V.; Alwafai, Z.; Balleyguier, C.; Clevert, D.A.; Fastner, S.; Gomez, C.; Goncalo, M.; et al. Deep learning model for breast shear wave elastography to improve breast cancer diagnosis (INSPiRED 006): An international, multicenter analysis. *J. Clin. Oncol.* **2025**, *43*, 3482–3493. [[CrossRef](#)] [[PubMed](#)]
127. Fukuda, T.; Tsunoda, H.; Yagishita, K.; Naganawa, S.; Hayashi, K.; Kurihara, Y. Deep learning for differentiation of breast masses detected by screening ultrasound elastography. *Ultrasound Med. Biol.* **2023**, *49*, 989–995. [[CrossRef](#)] [[PubMed](#)]
128. Zhang, L.; Jia, Z.; Leng, X.; Ma, F. Artificial intelligence algorithm-based ultrasound image segmentation technology in the diagnosis of breast cancer axillary lymph node metastasis. *J. Healthc. Eng.* **2021**, *2021*, 8830260. [[CrossRef](#)] [[PubMed](#)]
129. Salim, M.; Liu, Y.; Sorkhei, M.; Ntola, D.; Foukakis, T.; Fredriksson, I.; Wang, Y.; Eklund, M.; Azizpour, H.; Smith, K.; et al. AI-based selection of individuals for supplemental MRI in population-based breast cancer screening: The randomized ScreenTrustMRI trial. *Nat. Med.* **2024**, *30*, 2623–2630. [[CrossRef](#)] [[PubMed](#)]
130. Zheng, H.; Jian, L.; Li, L.; Liu, W.; Chen, W. Prior clinico-radiological features informed multimodal MR images convolution neural network: A novel deep learning framework for prediction of lymphovascular invasion in breast cancer. *Cancer Med.* **2024**, *13*, e6932. [[CrossRef](#)] [[PubMed](#)]
131. Ma, M.; Liu, R.; Wen, C.; Xu, W.; Xu, Z.; Wang, S.; Wu, J.; Pan, D.; Zheng, B.; Qin, G.; et al. Predicting the molecular subtype of breast cancer and identifying interpretable imaging features using machine learning algorithms. *Eur. Radiol.* **2022**, *32*, 1652–1662. [[CrossRef](#)] [[PubMed](#)]
132. Wang, X.; Chou, K.; Zhang, G.; Zuo, Z.; Zhang, T.; Zhou, Y.; Mao, F.; Lin, Y.; Shen, S.; Zhang, X.; et al. Breast cancer pre-clinical screening using infrared thermography and artificial intelligence: A prospective, multicentre, diagnostic accuracy cohort study. *Int. J. Surg.* **2023**, *109*, 3021–3031. [[CrossRef](#)] [[PubMed](#)]
133. Simsek-Cetinkaya, S.; Cakir, S.K. Evaluation of the effectiveness of artificial intelligence assisted interactive screen-based simulation in breast self-examination: An innovative approach in nursing students. *Nurse Educ. Today* **2023**, *127*, 105857. [[CrossRef](#)] [[PubMed](#)]
134. 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](#)]
135. Dihge, L.; Vallon-Christersson, J.; Hegardt, C.; Saal, L.H.; Häkkinen, J.; Larsson, C.; Ehinger, A.; Loman, N.; Malmberg, M.; Bendahl, P.O.; et al. Prediction of lymph node metastasis in breast cancer by gene expression and clinicopathological models: Development and validation within a population-based cohort. *Clin. Cancer Res.* **2019**, *25*, 6368–6381. [[CrossRef](#)] [[PubMed](#)]
136. Bhattarai, S.; Klimov, S.; Aleskandarany, M.A.; Burrell, H.; Wormall, A.; Green, A.R.; Rida, P.; Ellis, I.O.; Osan, R.M.; Rakha, E.A.; et al. Machine learning-based prediction of breast cancer growth rate in vivo. *Br. J. Cancer* **2019**, *121*, 497–504. [[CrossRef](#)] [[PubMed](#)]

137. Pan, Y.; Zhang, Q.; Zhang, H.; Kong, F. Prognostic and immune microenvironment analysis of cuproptosis-related lncRNAs in breast cancer. *Funct. Integr. Genom.* **2023**, *23*, 38. [[CrossRef](#)] [[PubMed](#)]
138. Jin, Y.; Lan, A.; Dai, Y.; Jiang, L.; Liu, S. Development and testing of a random forest-based machine learning model for predicting events among breast cancer patients with a poor response to neoadjuvant chemotherapy. *Eur. J. Med. Res.* **2023**, *28*, 394. [[CrossRef](#)] [[PubMed](#)]
139. Lee, H.J.; Nguyen, A.T.; Song, M.W.; Lee, J.E.; Park, S.B.; Jeong, W.G.; Park, M.H.; Lee, J.S.; Park, I.; Lim, H.S. Prediction of residual axillary nodal metastasis following neoadjuvant chemotherapy for breast cancer: Radiomics analysis based on chest computed tomography. *Korean J. Radiol.* **2023**, *24*, 498–511. [[CrossRef](#)] [[PubMed](#)]
140. Dercle, L.; McGale, J.; Zhao, B.; Schmitt, J.; Peltzer, A.; Schwartz, L.H.; Amend, M. Artificial intelligence and radiomics biomarkers for treatment response prediction in advanced HER2-negative breast cancer. *Breast* **2025**, *84*, 104571. [[CrossRef](#)] [[PubMed](#)]
141. Coté, D.; Eustace, A.; Toomey, S.; Cremona, M.; Milewska, M.; Furney, S.; Carr, A.; Fay, J.; Kay, E.; Kennedy, S.; et al. Germline single nucleotide polymorphisms in ERBB3 and BARD1 genes result in a worse relapse-free survival response for HER2-positive breast cancer patients treated with adjuvant-based docetaxel, carboplatin and trastuzumab (TCH). *PLoS ONE* **2018**, *13*, e0200996. [[CrossRef](#)] [[PubMed](#)]
142. Li, X.; Eastham, J.; Giltnane, J.M.; Zou, W.; Zijlstra, A.; Tabatsky, E.; Banchereau, R.; Chang, C.W.; Nabat, B.Y.; Patil, N.S.; et al. Automated tumor immunophenotyping predicts clinical benefit from anti-PD-L1 immunotherapy. *J. Pathol.* **2024**, *263*, 190–202. [[CrossRef](#)] [[PubMed](#)]
143. Greenwald, N.F.; Nederlof, I.; Sowers, C.; Ding, D.Y.; Park, S.; Kong, A.; Houlahan, K.E.; Varra, S.R.; de Graaf, M.; Geurts, V.; et al. Temporal and spatial composition of the tumor microenvironment predicts response to immune checkpoint inhibition in metastatic TNBC. *Nat. Cancer* **2026**, *7*, 435–450. [[CrossRef](#)] [[PubMed](#)]
144. Hernando-Calvo, A.; Yang, S.Y.C.; Vila-Casadesús, M.; Han, M.; Liu, Z.A.; Berman, A.H.K.; Spreafico, A.; Razak, A.A.; Lheureux, S.; Hansen, A.R.; et al. Combined transcriptome and circulating tumor DNA longitudinal biomarker analysis associates with clinical outcomes in advanced solid tumors treated with pembrolizumab. *JCO Precis. Oncol.* **2024**, *8*, e2400100. [[CrossRef](#)] [[PubMed](#)]
145. Fan, L.; Zhang, W.J.; Li, H.P.; Zeng, X.H.; Teng, Y.E.; Gong, Y.; Jin, X.; Zhao, S.; Sun, T.; Chen, W.Y.; et al. Precision treatment with artificial intelligence-assisted subtyping enhances therapeutic efficacy in HR+/HER2– breast cancer: The LINUX trial. *Cancer Cell* **2026**, *44*, 355–365.e3. [[CrossRef](#)] [[PubMed](#)]
146. Ge, H.; Mo, H.; Wei, Y.; Wang, J.; Qi, Y.; Li, L.; Ma, F. Biologically informed integration of drug representations for breast cancer treatment using deep learning. *Nat. Commun.* **2025**, *17*, 10. [[CrossRef](#)] [[PubMed](#)]
147. Mao, Y.; Di, W.; Zong, D.; Mu, Z.; He, X. Machine learning-based radiomics nomograms to predict number of fields in postoperative IMRT for breast cancer. *J. Appl. Clin. Med. Phys.* **2024**, *25*, e14194. [[CrossRef](#)] [[PubMed](#)]
148. Kazemimoghadam, M.; Chi, W.; Rahimi, A.; Kim, N.; Alluri, P.; Nwachukwu, C.; Lu, W.; Gu, X. Saliency-guided deep learning network for automatic tumor bed volume delineation in post-operative breast irradiation. *Phys. Med. Biol.* **2021**, *66*, ac176d. [[CrossRef](#)] [[PubMed](#)]
149. Kaidar-Person, O.; Antunes, M.; Cardoso, J.S.; Ciani, O.; Cruz, H.; Di Micco, R.; Gentilini, O.D.; Gonçalves, T.; Gouveia, P.; Heil, J.; et al. Evaluating the ability of an artificial intelligence cloud-based platform designed to provide information prior to locoregional therapy for breast cancer in improving patient satisfaction with therapy: The CINDERELLA trial. *PLoS ONE* **2023**, *18*, e0289365. [[CrossRef](#)] [[PubMed](#)]
150. Liu, H.; Chen, Y.; Zhang, Y.; Wang, L.; Luo, R.; Wu, H.; Wu, C.; Zhang, H.; Tan, W.; Yin, H.; et al. A deep learning model integrating mammography and clinical factors facilitates the malignancy prediction of BI-RADS 4 microcalcifications in breast cancer screening. *Eur. Radiol.* **2021**, *31*, 5902–5912. [[CrossRef](#)] [[PubMed](#)]
151. Tzelves, L.; Manolitsis, I.; Varkarakis, I.; Ivanovic, M.; Kokkonidis, M.; Useros, C.S.; Kosmidis, T.; Muñoz, M.; Grau, I.; Athanatos, M.; et al. Artificial intelligence supporting cancer patients across Europe—The ASCAPE project. *PLoS ONE* **2022**, *17*, e0265127. [[CrossRef](#)] [[PubMed](#)]
152. Jiang, L.; Xu, J.; Wu, Y.; Liu, Y.; Wang, X.; Hu, Y. Effects of the “AI-TA” mobile app with intelligent design on psychological and related symptoms of young survivors of breast cancer: A randomized controlled trial. *JMIR Mhealth Uhealth* **2024**, *12*, e50783. [[CrossRef](#)] [[PubMed](#)]
153. Schmitz, K.H.; Kanski, B.; Gordon, B.; Caru, M.; Vasakar, M.; Truica, C.I.; Wang, M.; Doerksen, S.; Lorenzo, A.; Winkels, R.; et al. Technology-based supportive care for metastatic breast cancer patients. *Support. Care Cancer* **2023**, *31*, 401. [[CrossRef](#)] [[PubMed](#)]
154. Lee, J.; Byun, H.K.; Kim, Y.T.; Shin, J.; Kim, Y.B. A study on breast cancer patient care using chatbot and video education for radiation therapy: A randomized controlled trial. *Int. J. Radiat. Oncol. Biol. Phys.* **2025**, *122*, 84–92. [[CrossRef](#)] [[PubMed](#)]
155. Al-Hilli, Z.; Noss, R.; Dickard, J.; Wei, W.; Chichura, A.; Wu, V.; Renicker, K.; Pederson, H.J.; Eng, C. A randomized trial comparing the effectiveness of pre-test genetic counseling using an artificial intelligence automated chatbot and traditional in-person genetic counseling in women newly diagnosed with breast cancer. *Ann. Surg. Oncol.* **2023**, *30*, 5990–5996. [[CrossRef](#)] [[PubMed](#)]

156. Steyerberg, E.W.; Vickers, A.J.; Cook, N.R.; Gerdts, T.; Gonen, M.; Obuchowski, N.; Pencina, M.J.; Kattan, M.W. Assessing the Performance of Prediction Models: A Framework for Traditional and Novel Measures. *Epidemiology* **2010**, *21*, 128–138. [[CrossRef](#)] [[PubMed](#)]
157. Chicco, D.; Jurman, G. The Advantages of the Matthews Correlation Coefficient (MCC) over F1 Score and Accuracy in Binary Classification Evaluation. *BMC Genom.* **2020**, *21*, 6. [[CrossRef](#)] [[PubMed](#)]
158. Drukker, K.; Chen, W.; Gichoya, J.; Grusauskas, N.; Kalpathy-Cramer, J.; Koyejo, S.; Myers, K.; Sá, R.C.; Sahiner, B.; Whitney, H.; et al. Toward fairness in artificial intelligence for medical image analysis: Identification and mitigation of potential biases in the roadmap from data collection to model deployment. *J. Med. Imaging* **2023**, *10*, 061104. [[CrossRef](#)] [[PubMed](#)]
159. Chen, R.J.; Wang, J.J.; Williamson, D.F.K.; Chen, T.Y.; Lipkova, J.; Lu, M.Y.; Sahai, S.; Mahmood, F. Algorithmic fairness in artificial intelligence for medicine and healthcare. *Nat. Biomed. Eng.* **2023**, *7*, 719–742. [[CrossRef](#)] [[PubMed](#)]
160. Viswanathan, V.S.; Parmar, V.; Madabhushi, A. Towards equitable AI in oncology. *Nat. Rev. Clin. Oncol.* **2024**, *21*, 628–637. [[CrossRef](#)] [[PubMed](#)]
161. Istasy, P.; Lee, W.S.; Iansavichene, A.; Upshur, R.; Gyawali, B.; Burkell, J.; Sadikovic, B.; Lazo-Langner, A.; Chin-Yee, B. The impact of artificial intelligence on health equity in oncology: Scoping review. *J. Med. Internet Res.* **2022**, *24*, e39748. [[CrossRef](#)] [[PubMed](#)]
162. Gichoya, J.W.; Banerjee, I.; Bhimireddy, A.R.; Burns, J.L.; Celi, L.A.; Chen, L.C.; Correa, R.; Dullerud, N.; Ghassemi, M.; Huang, S.C.; et al. AI recognition of patient race in medical imaging: A modelling study. *Lancet Digit. Health* **2022**, *4*, e406–e414. [[CrossRef](#)] [[PubMed](#)]
163. 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](#)]
164. Sadeghi, Z.; Alizadehsani, R.; Cifci, M.A.; Kausar, S.; Rehman, R.; Mahanta, P.; Pardalos, P.M. A review of explainable artificial intelligence in healthcare. *Comput. Electr. Eng.* **2024**, *118*, 109370. [[CrossRef](#)]
165. Muhammad, D.; Bendeche, M. Unveiling the black box: A systematic review of explainable artificial intelligence in medical image analysis. *Comput. Struct. Biotechnol. J.* **2024**, *24*, 542–560. [[CrossRef](#)] [[PubMed](#)]
166. Hur, S.; Lee, Y.; Park, J.; Jeon, Y.J.; Cho, J.H.; Cho, D.; Lim, D.; Hwang, W.; Cha, W.C.; Yoo, J. Comparison of SHAP and clinician friendly explanations reveals effects on clinical decision behaviour. *NPJ Digit. Med.* **2025**, *8*, 578. [[CrossRef](#)] [[PubMed](#)]
167. Salih, A.M.; Raisi-Estabragh, Z.; Galazzo, I.B.; Radeva, P.; Petersen, S.E.; Lekadir, K.; Menegaz, G. A perspective on explainable artificial intelligence methods: SHAP and LIME. *Adv. Intell. Syst.* **2025**, *7*, 2400304. [[CrossRef](#)]
168. Ciardiello, A.; D’Angelo, A.; De Angelis, L.; Giagu, S.; Sala, E.; Gigante, G. Beyond the black box: Lessons in explainability from AI in mammography. *Artif. Intell. Rev.* **2026**, *59*, 130. [[CrossRef](#)]
169. Talaat, F.M.; Gamel, S.A.; El-Balka, R.M.; Shehata, M.; ZainEldin, H. Grad-CAM enabled breast cancer classification with a 3D Inception-ResNet V2: Empowering radiologists with explainable insights. *Cancers* **2024**, *16*, 3668. [[CrossRef](#)] [[PubMed](#)]
170. Moor, M.; Banerjee, O.; Abad, Z.S.H.; Krumholz, H.M.; Leskovec, J.; Topol, E.J.; Rajpurkar, P. Foundation models for generalist medical artificial intelligence. *Nature* **2023**, *616*, 259–265. [[CrossRef](#)] [[PubMed](#)]
171. Azizi, S.; Mustafa, B.; Ryan, F.; Beaver, Z.; Freyberg, J.; Deaton, J.; Norouzi, M. Big self-supervised models advance medical image classification. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV 2021), Montreal, QC, Canada, 11–17 October 2021; pp. 3478–3488. [[CrossRef](#)]
172. Sheller, M.J.; Edwards, B.; Reina, G.A.; Martin, J.; Pati, S.; Kotrotsou, A.; Milchenko, M.; Xu, W.; Marcus, D.; Colen, R.R.; et al. Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data. *Sci. Rep.* **2020**, *10*, 12598. [[CrossRef](#)] [[PubMed](#)]
173. Guan, H.; Liu, M. Domain adaptation for medical image analysis: A survey. *IEEE Trans. Biomed. Eng.* **2022**, *69*, 1173–1185. [[CrossRef](#)] [[PubMed](#)]
174. Begoli, E.; Bhattacharya, T.; Kusnezov, D. The need for uncertainty quantification in machine-assisted medical decision making. *Nat. Mach. Intell.* **2019**, *1*, 20–23. [[CrossRef](#)]
175. Pianykh, O.S.; Lings, G.; Dewey, M.; Enzmann, D.R.; Herold, C.J.; Schoenberg, S.O.; Brink, J.A. Continuous learning AI in radiology: Implementation principles and early applications. *Radiology* **2020**, *297*, 6–14. [[CrossRef](#)] [[PubMed](#)]
176. Kazerouni, A.; Aghdam, E.K.; Heidari, M.; Azad, R.; Fayyaz, M.; Hacihaliloglu, I.; Merhof, D. Diffusion models in medical imaging: A comprehensive survey. *Med. Image Anal.* **2023**, *88*, 102846. [[CrossRef](#)] [[PubMed](#)]
177. He, K.; Gan, C.; Li, Z.; Rekik, I.; Yin, Z.; Ji, W.; Shen, D. Transformers in medical image analysis. *Intell. Med.* **2023**, *3*, 59–78. [[CrossRef](#)]
178. Benjamins, S.; Dhunoo, P.; Meskó, B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: An online database. *NPJ Digit. Med.* **2020**, *3*, 118. [[CrossRef](#)] [[PubMed](#)]
179. Artificial Intelligence Act. Available online: <https://artificialintelligenceact.eu/> (accessed on 24 May 2026).
180. Char, D.S.; Shah, N.H.; Magnus, D. Implementing machine learning in health care—Addressing ethical challenges. *N. Engl. J. Med.* **2018**, *378*, 981–983. [[CrossRef](#)] [[PubMed](#)]

181. Kelly, C.J.; Karthikesalingam, A.; Suleyman, M.; Corrado, G.; King, D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med.* **2019**, *17*, 195. [[CrossRef](#)] [[PubMed](#)]
182. U.S. Food and Drug Administration. Artificial Intelligence Software as a Medical Device (SaMD). Available online: <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-software-medical-device> (accessed on 24 May 2026).
183. Voigt, P.; von dem Bussche, A. *The EU General Data Protection Regulation (GDPR): A Practical Guide*, 1st ed.; Springer International Publishing: Cham, Switzerland, 2017.
184. World Health Organization. *Ethics and Governance of Artificial Intelligence for Health: WHO Guidance*; World Health Organization: Geneva, Switzerland, 2021. Available online: <https://www.who.int/publications/i/item/9789240029200> (accessed on 24 May 2026).

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.