




Artificial intelligence in breast cancer imaging: risk stratification, lesion detection and classification, treatment planning and prognosis—a narrative review

Maurizio Cè^{1*}, Elena Caloro¹, Maria E. Pellegrino¹, Mariachiara Basile¹, Adriana Sorce¹, Deborah Fazzini², Giancarlo Oliva³, Michaela Cellina³

¹Postgraduate School in Diagnostic and Interventional Radiology, University of Milan, 20122 Milan, Italy

²Centro Diagnostico Italiano, 20147 Milan, Italy

³Department of Radiology, ASST Fatebenefratelli Sacco, 20121 Milan, Italy

***Correspondence:** Maurizio Cè, Postgraduate School in Diagnostic and Interventional Radiology, University of Milan, Via Festa del Perdono, 7, 20122 Milan, Italy. maurizioce.md1@gmail.com

Academic Editor: Valerio Nardone, University of Campania “L. Vanvitelli”, Italy

Received: August 25, 2022 **Accepted:** September 28, 2022 **Published:** December 27, 2022

Cite this article: Cè M, Caloro E, Pellegrino ME, Basile M, Sorce A, Fazzini D, et al. Artificial intelligence in breast cancer imaging: risk stratification, lesion detection and classification, treatment planning and prognosis—a narrative review. *Explor Target Antitumor Ther.* 2022;3:795–816. <https://doi.org/10.37349/etat.2022.00113>

Abstract

The advent of artificial intelligence (AI) represents a real game changer in today’s landscape of breast cancer imaging. Several innovative AI-based tools have been developed and validated in recent years that promise to accelerate the goal of real patient-tailored management. Numerous studies confirm that proper integration of AI into existing clinical workflows could bring significant benefits to women, radiologists, and healthcare systems. The AI-based approach has proved particularly useful for developing new risk prediction models that integrate multi-data streams for planning individualized screening protocols. Furthermore, AI models could help radiologists in the pre-screening and lesion detection phase, increasing diagnostic accuracy, while reducing workload and complications related to overdiagnosis. Radiomics and radiogenomics approaches could extrapolate the so-called imaging signature of the tumor to plan a targeted treatment. The main challenges to the development of AI tools are the huge amounts of high-quality data required to train and validate these models and the need for a multidisciplinary team with solid machine-learning skills. The purpose of this article is to present a summary of the most important AI applications in breast cancer imaging, analyzing possible challenges and new perspectives related to the widespread adoption of these new tools.

Keywords

Breast cancer imaging, artificial intelligence, machine learning, computer-aided detection, mammogram, digital breast tomosynthesis, magnetic resonance imaging



Introduction

Breast cancer is the most common type of cancer in women and the second leading cause of death after lung cancer [1]. With more than 2 million new cases in 2020, it represents a major public health concern for health systems and policymakers [2, 3]. In recent years, radiology has faced exponential growth in artificial intelligence (AI) applications in clinical practice with significant and encouraging results, especially in oncological imaging [4]. Various imaging modalities are currently used in breast cancer imaging: mammography and digital breast tomosynthesis (DBT), ultrasound (US), magnetic resonance (MR), and positron emission tomography (PET), and each could gain significant benefits from AI support [5]. Several studies show that incorporating an AI-based approach into the standard radiological workflow improves breast imaging diagnostic accuracy [6].

AI has multiple applications in breast cancer imaging: 1) risk stratification, in order to achieve individualized screening programs; 2) tumor-assisted detection, to increase diagnostic accuracy, reducing the rate of false negatives and false recalls, while improving radiologists' workload; 3) non-invasive tumor characterization (identification of tumor subtype, evaluation of tumor heterogeneity and microenvironment, etc.) to plan targeted therapy and follow-up; and finally, 4) prognostic/predictive applications regarding response to treatment, risk of relapse and overall survival [7–11].

However, despite the promising prospects, all that glitters is not gold. The integration of AI-powered tools into clinical practice presents new and fascinating challenges for radiologists, which must be considered to successfully address this new frontier [12–14]. An overview of AI applications in breast cancer imaging is presented in Figure 1.

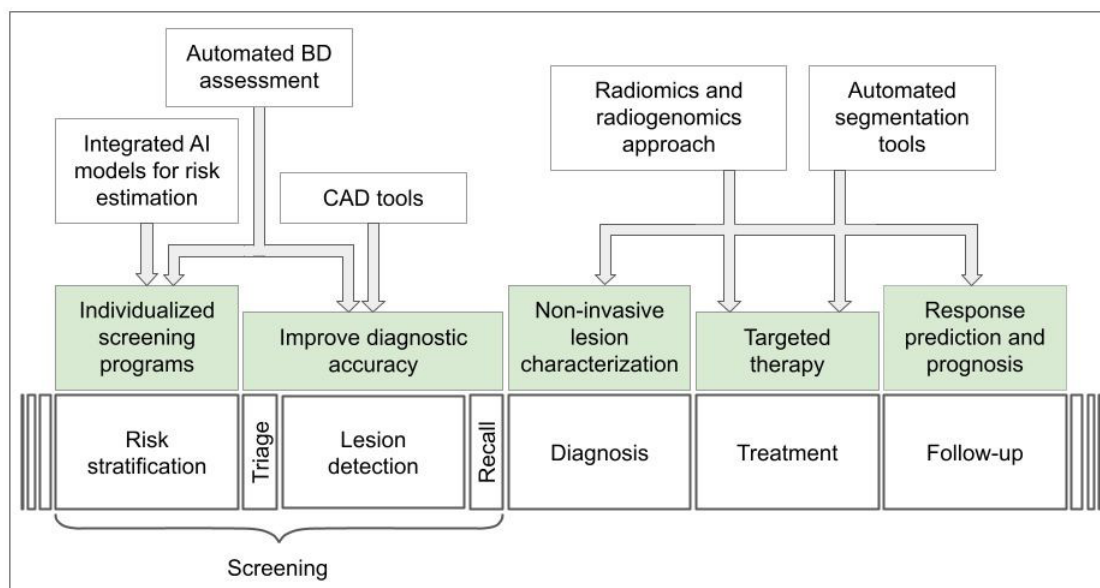


Figure 1. Overview of AI applications in breast cancer imaging. CAD: computer-aided detection; BD: breast density

A quick introduction to AI

AI is a vast field of knowledge. Technology that mimics human intelligence to solve problems is the core of what is collectively called AI. Its outermost shell is machine learning (ML), a term that refers to the automated detection of meaningful patterns in data [15, 16].

Normally a machine (i.e., a computer) performs operations on input (x) to obtain output (y). To perform such a task, the programmer is required to code the function (f) to be computed in the programming language (coding). Since breast cancer screening is challenging and time-consuming, radiologists would benefit from a machine aiding in the classification of breast lesions in mammograms as either benign or malignant. The problem is that, as radiologists are well aware, lesions detection is an extremely complex cognitive task that depends on intrinsic factors, such as the experience and skills of the radiologist, and extrinsic factors, such as the characteristics of the tumor (size, position, morphology, etc.) and the breast (BD, etc.). From a

computational point of view, it is not possible to translate into the code what the expert radiologist does in his work routine, mostly automatically (Figure 2).

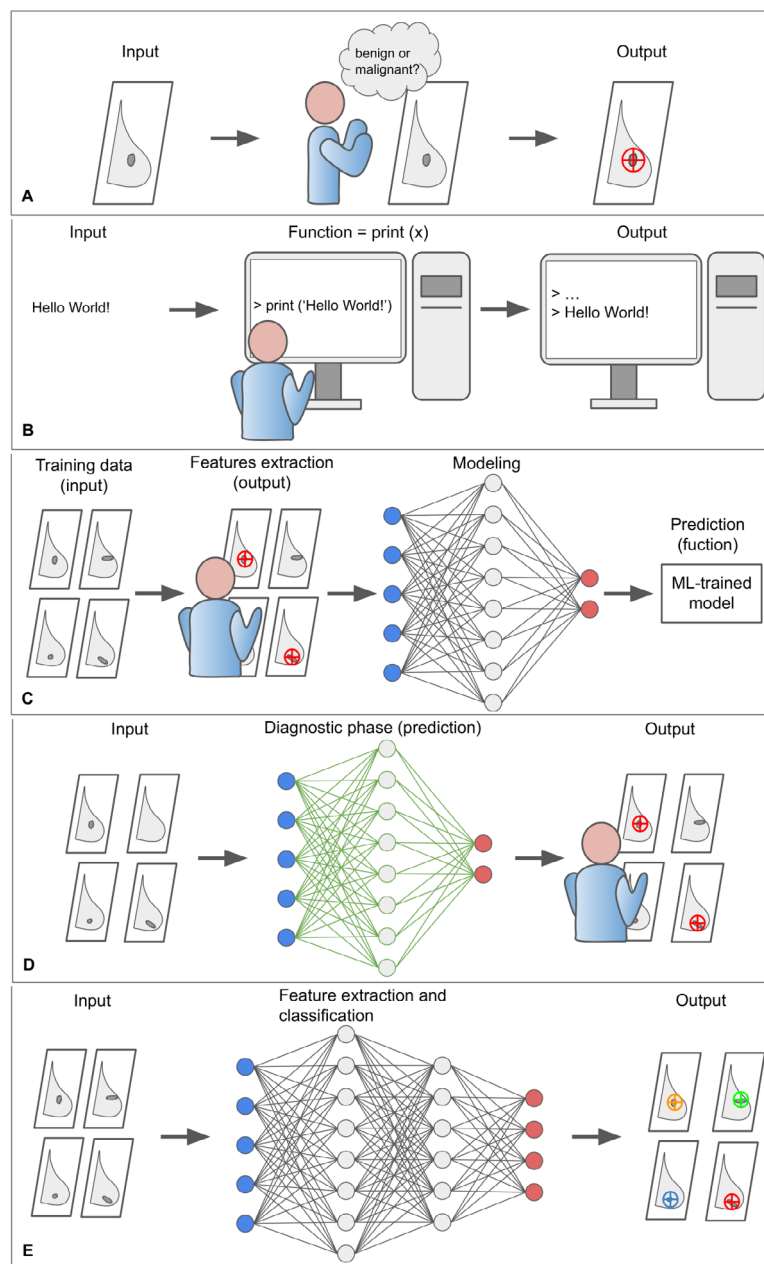


Figure 2. The core concepts of ML paradigm. A. In a normal workflow, the radiologist evaluates the mammogram (input) and determines whether the lesion is benign or malignant (output); B. to program a computer to perform a certain operation, it is necessary to specify the function that the machine is to perform. In the example, the function—print (x)—writes the argument of the function to the screen. However, the extremely complex cognitive process of the radiologist cannot be translated into the programming language; C. the classification problem can be addressed through a ML approach. The proposed example is a form of supervised learning that exploit a simple ANN to perform a binary classification task that consists in distinguishing between benign and malignant lesion on digital mammograms. The ML model requires a learning phase in which both raw data and exams already classified by the radiologist are provided. The learning phase of the model includes a training, validation, and testing phase, which are not represented for clarity. The model progressively reduces the uncertainty as it is exposed to more data samples and comes to approximate the “function” normally performed by the radiologists; D. after the ML algorithm has been trained and tested, it can be integrated into clinical practice and used to assist the radiologist in visual assessment; E. DL models exploit particularly complex neural networks to extract and analyse new and intricate image patterns (radiomic features) in large data sets that are usually not accessible to the human operator. ANN: artificial neural network; DL: deep learning

However, it is possible to address this challenge through an AI-based approach, in which a ML model is trained to infer the desired function by “learning for examples”, in some way simulating the learning modalities of the human brain. More formally, ML attempts to approximate a function f by analyzing the input data features that would produce the desired output, satisfying some predefined requirements [17].

It is also expected for the ML model to reduce the uncertainty of the approximation as it is exposed to more data samples. For example, in supervised learning (the simplest form of ML), the learning phase of the model would require a dataset (with at least two subparts, one for training and one for validation) comprising typical examples of inputs (e.g., digital mammograms) and corresponding outputs (e.g., tumor lesions already identified by an experienced radiologist). By feeding the model with enough data, it can learn to infer the relationship that binds them with a desired level of accuracy. After the learning phase, the model should be tested to verify its reliability in a different dataset [16].

From a practical point of view, a ML model consists of a set of rules that map relationships between data. Among different models, ANNs are a biologically-inspired programming paradigm that enables a computer to learn from observational data [18]. ANNs simulate the human brain and consist of several layers of interconnected 'nodes' or 'cells'. Through the training phase, the ANN progressively shapes the weights of its connections toward the implementation of the desired function. After the ANN is trained and tested, it can be applied to a new dataset to analyze and extract information from raw data (Figure 2) [18].

DL is a subdomain of AI that exploits complex ANNs with a large number of intermediate layers, each representing increasing levels of abstraction, to discover intricate patterns in large data sets that go beyond the features that could be extracted by the radiologist [19]. DL-based tools are particularly suitable for medical image analysis, including CAD, disease prediction, image segmentation, image generation, etc. [20]. DL models are built to capture the whole image context and learn complex correlations between local features, resulting in superior performance on image analysis tasks like classifying breast lesions in a screening mammogram as likely malignant or benign [21].

The application of AI tools in the field of diagnostic imaging is the basis of radiomics and radiogenomics approaches. In simple terms, radiomics could be considered synonymous with quantitative imaging [22]. The radiomics approach exploits sophisticated AI-based tools to extract and analyze large quantitative metrics (radiomics features) from medical images, that are typically inaccessible to the human reader, and try to correlate them to clinical features [23–25], in order to improve precision diagnosis and treatment [26, 27]. Radiogenomics could be considered a subset of radiomics applications, aiming to correlate lesion imaging phenotype ("radio") to the genotype ("genomics"), based on the assumption that phenotype is the expression of genotype [28]. If a specific imaging phenotype is related to a genotype, imaging features analysis could ideally predict cancer molecular patterns and behavior, allowing personalized treatment planning.

Risk stratification and screening

Early detection of breast cancer is essential for timely management [29]. However, breast cancer screening is a resource-consuming activity in modern radiology departments and radiologists take prime responsibility for image quality and diagnostic interpretation [30]. A position paper published in 2018 in the European Journal of Cancer by Autier and Boniol [29] concluded that "New, effective methods for breast screening are needed, as well as research on risk-based screening strategies". The importance of establishing more effective and individualized screening programs is well understood considering the limitations of current screening protocols [29]. Small lesions may be difficult to detect by the human eye during routine mammography and present as interval cancers or advanced cancers at the next examination, with a worse prognosis [31]. Although the recent development of DBT, which creates high-resolution three-dimensional (3D) images, increased breast cancer detection while simultaneously reducing false recalls, the impact of false positive results is still relevant, with increased patient anxiety, unnecessary invasive testing, and ultimately increased costs [32–34]. Furthermore, some authors pointed out the methodological limitations of some randomized trials that may have led to exaggerating the effectiveness of screening [29].

AI-based models could find application in the different stages of breast cancer screening: risk stratification, triage or pre-screening phase, exam interpretation, and patients' recall [35]. Breast cancer screening protocols in Europe generally rely on double-blind reading by two different radiologists, while in the United States a single reader plus CAD is more common [30, 36, 37]. To date, in the United States, the Food and Drug Administration (FDA) has approved several AI tools for application in breast cancer imaging:

10 for BD assessment, 3 for triage or pre-screening phase, 3 for lesion classification, 5 for lesion detection and classification [38], and further authorizations are being evaluated.

BD evaluation

A woman's risk of developing breast cancer depends on several factors such as age, personal and family history, and imaging features such as BD [39, 40], which is the radiographic appearance of the absolute amount or percentage of fibro glandular tissue in the breast [41]. High BD is an established major risk factor for breast cancer and a well-known factor limiting the sensitivity of mammographic screening as it may mask cancer [42, 43]. Mammographic evaluation of BD is traditionally held by the expert radiologists' visual assessment according to breast imaging-reporting and data system (BI-RADS) criteria. However, this approach presents several limitations such as high intra- and inter-observer variability, resulting in low reliability and reproducibility [44–47]. AI-based tools can help both BD quantitative assessment and lesion detection in high-BD mammograms.

Both BI-RADS-based qualitative BD assessment and computer-generated quantitative BD measures have been shown to be associated with breast cancer risk [48, 49]. In a recent study, a validated fully automated system for BD assessment [DenSeeMammo (DSM)] was found positively associated with breast cancer risk and non-inferior to radiologists' visual assessment [50].

Several automated or semi-automated software for reproducible assessment of BD have been developed as the Laboratory for Individualized Breast Radiodensity Assessment (LIBRA) [51], Quantra [52], and Volpara [53], and their performance has been evaluated in several studies. For example, BD estimates obtained by Quantra software tend to be lower than visual estimates, however, they correlate well with the BI-RADS BD categories visually assigned to the mammograms (Figure 3) [54]. According to Ekpo et al. [55], Quantra is a poor predictor of BI-RADS assessment on a four-grade scale, but well reproduces BI-RADS rating on a two-grade scale.

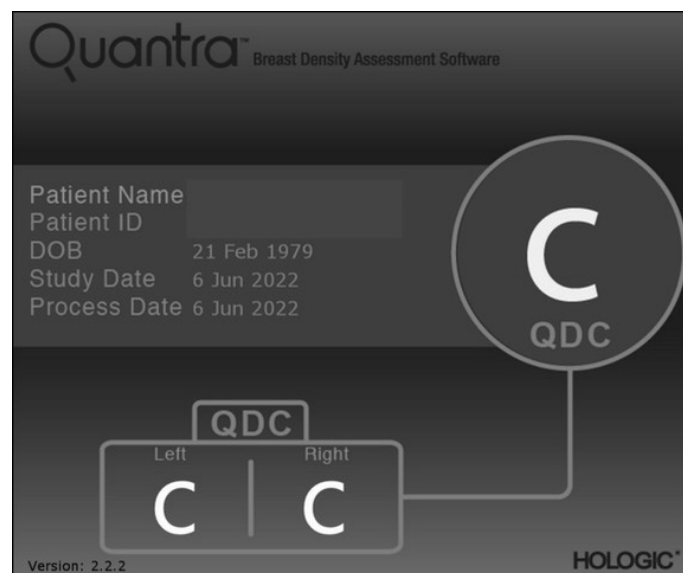


Figure 3. Quantra (HOLOGIC) automated detection of BD according to BI-RADS classification, classified as C in both breasts. The automatic assessment of BD refers to the mammography in Figure 4. QDC: Quantra density category

Deep-LIBRA, an AI system recently validated on a multi-racial, multi-institutional dataset of 15,661 images using convolutional neural network architectures, demonstrated a strong agreement of BD estimates between AI and gold-standard assessment by an expert reader [56].

Advanced DL algorithms enable the extraction of imaging texture features of breast tissue other than BD from standard screening mammographies, such as energy, contrast, correlation, and so on, which can be used to predict the risk of developing breast cancer [57]. A study by Arefan et al. [58] evaluated the feasibility and performance of two DL methods (denoted GoogLeNet and GoogLeNet-LDA) in a case-control setting: it

demonstrated that both exhibited superior performance in predicting breast cancer risk than the percentage of BD alone.

Risk stratification

Several individualized breast cancer risk prediction models have been proposed [59–61] and the effectiveness of individualized *versus* universal screening programs has been investigated in numerous randomized trials such as the Tailored Breast Screening Trial (TBST) [62], women informed to screen depending on measures of risk (WISDOM) [63], and My Personalized Breast Screening (MyPeBS) [64].

One of the most popular risk prediction models is the International Breast Intervention Study (IBIS) model, or Tyrer-Cuzick (TC) model, a scoring system guiding breast cancer screening and prevention by accounting for age, genotype, family history of breast cancer, age at menarche and at first birth, menopausal status, atypical hyperplasia, lobular carcinoma *in situ*, height, and body mass index (BMI) [60]. Despite its widespread use, however, IBIS/TC model demonstrated limited accuracy in some high-risk patient populations [65].

AI can help integrate imaging features into predictive risk models increasing accuracy. For example, a study by Yala et al. [7] evaluated the performance of a hybrid DL model considering both traditional risk factors and mammograms, in comparison with the IBIS/TC model alone: the hybrid model placed 31% of patients in the top risk category, compared with 18% identified by the IBIS/TC model, and was able to identify the features associated with long-term risk beyond early detection of the disease.

Pre-screening

AI tools are giving promising results in the pre-screening or triage phase to identify likely negative mammograms from those that need further evaluation by an experienced radiologist, with a significant reduction in the workload. A study by Rodriguez-Ruiz et al. [66] evaluated the performance of a new AI system to preliminary rule out mammograms with a low probability of cancer before the radiologist assessment. The balance between workload and diagnostic performance depended on the risk threshold chosen: with a low-risk threshold, they obtained an exclusion rate of only 1% of exams with cancer with a significant reduction in the workload (17%) as well [66]. Similar results have been obtained by Dembrower et al. [67] in a study proving that AI-assisted mammography analysis could reduce radiologist workload and increase cancer detection.

A recent study by Balta and colleagues [68] evaluated the performance of a single *versus* double reading screening program in pre-assessed exams by an AI system that assigns a score from 1 to 10 on each screening exam indicating the likelihood of cancer. It was found that the use of single reading instead of double reading in tests with a low probability of cancer (score 1 to 7), led to unaffected detection (no screen-detected cancers were missed and even the AI detection score was low the single-reader would recall the exam for double check), with a recall rate down 11.8% of the workload down 32.6% [68].

Lesion detection

Assisting the radiologist in improving lesion detection is one of the primary goals of AI-based tools (Figure 4). After the United States FDA approved CAD for mammography in 1998, several studies evaluated the accuracy of CAD software in mammography screening, with initially controversial results. Unfortunately, some authors observed that CAD systems adversely affect some radiologists' performance and increase recall rates [69]. Against an estimated cost of over \$400 million annually, a study by Lehman et al. [37] found no evidence that CAD applied to digital mammography significantly improves screening performance.

However, before 2020, AI algorithms were developed using small mammography data sets collected by one or two institutions, limiting verification of the robustness of those algorithms [70], instead, numerous evidence has emerged in the last two years that supports the use of AI in mammography screening as a complementary diagnostic tool, both in single and double reading settings, with significant performance gains.

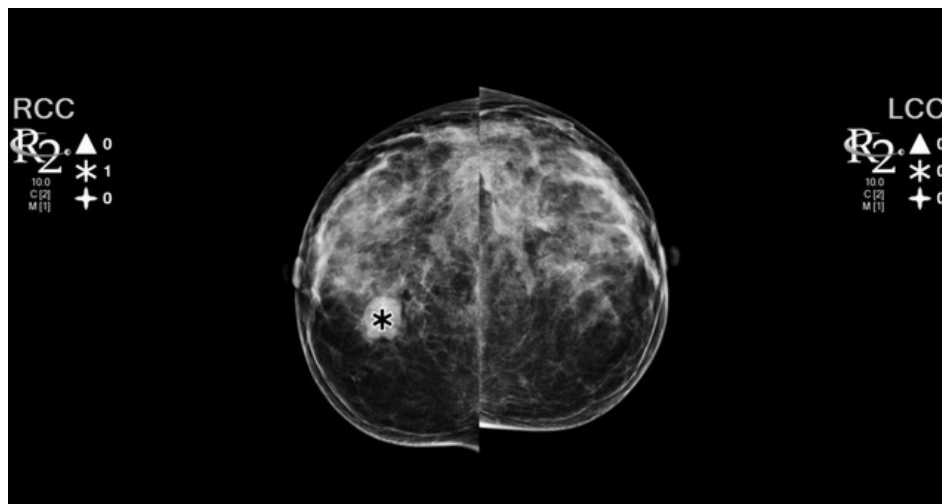


Figure 4. A nodule was detected by the CAD tool in the right breast (asterisk). LCC: left craniocaudal view; RCC: right craniocaudal view

In an international crowdsourced challenge, Schaffter et al. [71] underscored the potential of using ML methods to enhance mammography screening interpretation: even if no single AI algorithm outperformed radiologists, an ensemble of AI algorithms combined with radiologist assessment in a single-reader screening environment showed to improved overall accuracy.

A European study conducted on screening mammograms of 429 consecutive women diagnosed with interval cancer, positively addressed the question of whether a DL-based AI system could reduce interval cancer in mammography screening, without additional screening modalities [72].

Recent retrospective studies based on a large population have confirmed the usefulness of AI-based approaches. In a study conducted in Europe including 122,969 mammography examinations from 47,877 women, AI showed promising results according to the proportion of screen-detected cancers detected by AI at different preselected thresholds [73]. In another large retrospective study, a DL algorithm was developed and validated with 170,230 mammography examinations collected from five institutions in South Korea, the United States, and the UK, showing a better diagnostic performance in breast cancer detection compared with radiologists [70].

In a study by Kim et al. [74] the use of an AI-based system demonstrated significant added value in detecting mammographically occult breast cancers: 97.5% were found in heterogeneous or extremely dense breasts, 52.5% were asymptomatic, 86.5% were invasive, and 29.7% already had axillary lymph node metastases.

Many studies conducted to date present several limitations: most of them are small and retrospective in nature. Therefore, further prospective trials are required. The milestone MASAI study, a randomized controlled trial, still ongoing, with 100,000 participants, is aiming to assess whether AI can improve the efficacy of mammography screening with the application of AI in the main phases of mammographic screening, having interval cancer as the principal endpoint [75].

Unlike traditional mammography, DBT operates in a 3D domain and differs in acquisition time, acquisition angle, and the number of images. For these reasons, the application of AI models to DBT represents a considerable challenge. Several studies have evaluated the performance of DL-based algorithms in identifying suspicious masses, architectural distortions, and microcalcifications in DBT images [76, 77]. In a recent study, van Winkel et al. [78] evaluated the impact of AI on accuracy and reading time in DBT interpretation, demonstrating that radiologists improved their performances when using an AI-based support system. Another study by Conant et al. [79] compared the performance of 24 radiologists (13 of whom were breast subspecialists) in reading 260 DBT examinations (including 65 cancer cases), both with and without AI support. The AI-based tool was found to improve radiologist performance in the detection of malignant lesions, with a reduction in recall rate and reading time [79].

Still, the validation of these AI systems requires huge case series. Buda et al. [80] curated and annotated a data set of DBT studies including 22,032 reconstructed DBT volumes from 5,060 patients and made this data set publicly available at the Cancer Imaging Archive [81] a public data hosting service for medical images of various modalities, to develop and test a DL algorithm for breast cancer detection.

Concerning MR imaging (MRI), actually, the only FDA-approved CAD is QuantX™ from Qlarity Imaging. QuantX™ is indicated for the assessment and characterization of breast abnormalities from MRI data in patients presenting for high-risk screening, diagnostic imaging workup, or evaluation of the extent of known disease. In a retrospective, clinical reader study, the application of this tool improved radiologists' performance in the task of differentiating benign and malignant MRI breast lesions, showing an increase in the average area under the curve (AUC) of all readers from 0.71 to 0.76 ($P = 0.04$) when using the AI system [82].

Lesion characterization

Breast cancer is a clinically and biologically heterogeneous disease, with several recognized histotypes and molecular subtypes [83]. Subtype discrimination is essential for planning targeted therapy. Currently, four molecular subtypes have been identified: 1) luminal A; 2) luminal B; 3) human epidermal growth factor receptor 2 (HER2)-enriched; and 4) basal-like, which have critical differences in incidence, response to treatment, disease progression, survival, and imaging characteristics [84]. High-throughput discrimination of breast cancer subtypes and inter- and intra-tumoral heterogeneity is essential to set up targeted therapies [85]. However, this could only be partially accounted for by current diagnostic tools. In this landscape, valuable help could come from AI through its radiomics and radiogenomics approaches, representing a considerable boost toward personalized medicine [86–89].

Nowadays, tumor genotype and molecular characterization require invasive techniques like surgery or biopsy to collect tissue samples. In addition to common complications like pain, bleeding, hematoma, and breast implant damage, the most important bias in biopsy is the missed diagnosis due to insufficient material collected, especially in small lesions [90]. Moreover, since the tissue sample is a small portion of a heterogeneous lesion, molecular analysis results may not be accurate for the entire lesion [91] and large-scale genome-based cancer characterization is not habitually performed due to high costs, time consumption, and technical complexity [91]. Finally, the tumor genome may change over time making treatment less effective, however, the biopsy is not the ideal method for tracking tumor evolution.

By combining multimodality imaging data, AI tools could complement traditional histological assessment by overcoming some of the limitations of biopsy. First of all, AI image analysis software could evaluate the whole 3D tumor lesion and its surrounding microenvironment [92]. Moreover, AI makes it possible to assess multiple lesions and track them at different time points, allowing physicians to adapt targeted therapy over time [93].

To date, some early attempts to match genomic and imaging data have been made through the creation of two archives: The Cancer Genome Atlas (TCGA) which collects several genomic and clinical cancer biomarkers, and The Cancer Imaging Archive (TCIA) containing corresponding imaging data, with the limit of different protocols, sometimes used [24].

In the last decade, radiomics analyses have been applied to mammography and DBT, breast US, and MRI, with promising results, and AI tools have been successfully applied to extract radiomic features [88, 94]. In 2019, a radiomics approach was applied in a retrospective study on 331 cancer cases aiming to automatically extract radiomics features from digital mammograms. Both qualitative and agnostic features have been evaluated and four of them showed a statistically significant ($P < 0.05$) difference: concavity, correlation, roundness, and gray mean (calculated from the histogram of tumor voxel intensities). More specifically, triple-negative samples have shown a smaller concavity, a larger roundness, and a major gray mean than HER2-enhanced and luminal samples [95].

In some pioneering studies, AI-assisted tools have been tested and evaluated in US imaging, in order to double-check radiologist interpretation and eventually reduce unnecessary biopsies. In a study by Wang et al. [96] the performance of two AI-based downgrading stratification methods was evaluated using

histopathological results as the reference standards. Stratification method A was used to downgrade only if the assessments of both orthogonal sections were possibly benign and it showed promising results: forty-three lesions diagnosed as BI-RADS category 4A by conventional US received a hypothetical AI-based downgrade, reducing the biopsy rate from 100% to 67.4% ($P < 0.001$) without losing malignancies [96].

Several studies evaluated the performance of AI tools in the analysis of dynamic contrast-enhanced (DCE) MR (DCE-MR) and diffusion-weighted imaging (DWI) data. When using a radiogenomics approach, researchers are interested in investigating the correlation between DCE-MR or DWI characteristics such as tumor size, shape, morphology, and genomic features, such as protein expression and mutations [97, 98]. For example, a study by Zhu et al. [98] found that transcriptional activities of various genetic pathways were positively associated with tumor size, blurred margins, and irregular tumor shape.

Several studies have focused on evaluating the effectiveness of a radiomics-based approach in distinguishing between malignant and benign lesions, exploiting the possibility of identifying and quantifying features otherwise difficult to be recognized by the human reader [22]. Entropy is an important imaging feature in tumor lesions reflecting the tumoral heterogeneity and its vascular status [99]. In an MRI-based radiomics study, an advanced ML tool was evaluated and the entropy value was found to be a useful parameter to distinguish malignant lesions compared to benign lesions [100]. Another study based on DCE-MR data tried the additive benefit of a set of quantitative features such as irregularity and entropy over maximum linear size alone to differentiate luminal A breast cancers from benign breast lesions, with promising results [101].

The correlation between MR features and molecular breast cancer subtypes has been investigated using traditional and radiomics-based approaches [24, 102]. Leithner et al. [103] evaluated the performance of an AI-based open-source software (MaZda 4.6) in the extraction of radiomics features aiming to assess breast cancer molecular subtypes. The discrimination between luminal A and triple negative cancers had the best statistical results, with an overall median AUC of 0.8 and median accuracies of 74% in the training dataset and 68.2% in the validation dataset [103]. A study by Li et al. [104] evaluated the performance of a classifier model for molecular subtyping: statistically significant associations were found between tumor phenotypes and receptor status, with aggressive cancers that tend to be larger in size with more heterogeneity in their contrast enhancement.

In another study, Yeh et al. [105] performed a radiomic analysis of different MR features to study the underlying activity of multiple molecular pathways that regulate replication, proliferation, apoptosis, immune system, and extracellular signaling. The results showed that tumors with upregulation of immune signaling pathways such as T-cell receptor and chemokine signaling, as well as extracellular signaling pathways, are associated with typical imaging features. The results suggest the possibility to infer the most immunologically active tumors and predict the effectiveness of immunological therapies [105]. Couture et al. [8] trained a DL image analysis tool on a data set of 571 breast tumors to create an image-based classifier assessing tumor grade, estrogen receptor (ER) status, prediction analysis of microarray 50 (PAM50) profile, histologic subtype, and risk of recurrence score. DL-based image analysis was able to distinguish low-intermediate *versus* high tumor grade (82% accuracy), HER status (84% accuracy), basal-like *versus* non-basal-like (77% accuracy), ductal *versus* lobular (94% accuracy), high *versus* the low-medium risk of recurrence score (75% accuracy) [8].

Whole-breast and tumor segmentation

Tumor segmentation is an important task in oncological imaging [106]. It consists of image analysis and delimitation of the regions of interest (ROI) comprising the tumor from a 2D or 3D acquisition [107]. Manual segmentation is a time-consuming task, affected by a high degree of inter-reader variability due to the limitations of the human reader to solve the lesion-background relationship unambiguously, especially in lesions with blurred margins or in the high-density breast. AI-assisted tools based on DL algorithms can reduce segmentation time and significantly increase reproducibility and efficiency [106]. In breast cancer imaging, it could be potentially useful in different tasks such as treatment planning, and lesion follow-up, but

also prognostic and predictive evaluations (Figure 1). Furthermore, tumor segmentation is an essential step in the radiomic workflow, necessary for the extraction of radiomic features [17, 22].

Jiang et al. [108] developed a fully automated algorithm for accurate segmentation of the whole breast using 3D fat-suppressed DCE-MR images and demonstrated a good overlap with manual segmentation. Zhang et al. [109] tested two DL models (UNet and SegNet) as segmentation methods in diffusion-weighted MR images and compared them to manual segmentation used as the reference standard. The study demonstrates that the DL models could achieve promising segmentation results to help computer-aided quantitative analyses of breast DWI images [109].

The assessment of normal breast tissue is essential to improve lesion detection and involves some segmentation tasks. In the first instance, the total amount of fibroglandular tissue is a well-known breast cancer risk factor (see also [BD evaluation](#)) [110]. In a study by Huo et al. [111], a DL algorithm was tested on a dataset of 100 breast DCE-MR previously assessed by expert radiologists according to BI-RADS criteria, showing a high correlation coefficient (0.981) between the breast densities obtained with the DL-based segmentation and the manual assessment.

Furthermore, the level of background parenchymal enhancement (BPE) in DCE-MR images is considered an independent marker of breast cancer risk and breast cancer treatment outcomes [112]. In a case-control study, Saha et al. [113] evaluated an AI-based model to assess quantitative measures of BPE in 133 women at high risk for developing breast cancer. AI-extracted BPE offered a more precise and reproducible measuring than humans and may potentially be used to further stratify risk in patients undergoing high-risk screening MRI [113].

Fully automated methods for assessing both fibroglandular tissue and BPE were developed and validated. Fibroglandular tissue was segmented in T1-weighted, nonfat-saturated MRI, and then propagates this segmentation to DCE-MR to quantify BPE within the segmented fibroglandular areas [112]. High spatial correspondence was observed between the automatic and manual fibroglandular tissue segmentation and both fibroglandular tissue and BPE quantifications indicated a high correlation between automatic and manual segmentations. However, a poor correlation was found between segmentation and clinical rating [114]. Similar studies have been proposed by Ma et al. [110] and Ha et al. [115].

An AI-assisted lesion segmentation tool has been applied to automated whole breast US by Lee et al. [116], aiming to overcome the amount of speckle noise and the low contrast of the lesion boundaries typical of this ultrasonic image technique. The ground truth was assessed by two experienced radiologists in breast US, the proposed method demonstrated high accuracy in processing automated whole breast US images, segmenting lesions, calculating lesion volumes, and visualizing lesions to facilitate observation by physicians [116].

A better 3D reconstruction and a more precise margin evaluation can improve preoperative planning. In patients receiving radiotherapy, proper target delineation can reduce radiation doses to the nearby normal organs at risk. The feasibility of a DL-based auto-segmentation tool was demonstrated by Chung et al. [117], the correlation between the auto-segmented and manually segmented contours was acceptable and the differences in dosimetric parameters were minimal. Recently, Byun et al. [118] assessed the performance of another DL auto-contouring system with a group of experts. Manual contours, corrected auto-contours, and auto-contours were compared. Inter-physician variations among the experts were reduced in corrected auto contours, compared to variations in manual contours; furthermore, the DL tool revealed good user satisfaction [118].

While AI methods are unlikely to replace the work of radiation oncologists, they could be a useful tool with excellent potential for assisting radiation oncologists in the future, improving the quality of breast radiotherapy, and reducing inter-reader variability in clinical practice.

Prognosis

One of the most important AI applications in breast cancer imaging is the development of new predictive and prognostic models based on a radiomic approach [23, 88, 119]. Traditionally prognostic factors

include age, number of positive axillary lymph nodes, tumor size, tumor grade, lymphatic and vascular invasion, and immunohistochemical biomarkers such as ER/progesterone receptor (PR) status, HER2, and Ki-67 [120, 121]. Additional factors include grade, presence of lymphovascular invasion, age, and ethnicity. Certain biologic factors, including ER/PR and HER2/neu, are both prognostic and predictive [121].

AI enables the integration of quantitative radiological data from various imaging modalities with patient clinical data (e.g., family history, molecular and genomic data) with the aim of improving the predictions of relevant clinical endpoints such as disease-free survival (DFS), progression-free survival, complete response, and others [23, 122]. Furthermore, radiomics-based approach allows the potential identification of new imaging biomarkers [88].

MR-based AI models have been developed to predict response to neoadjuvant chemotherapy at an early stage or even prior to the beginning of the treatments. These AI tools could be used to avoid the administration of ineffective and potentially toxic therapies, as well as to expedite surgery in patients who would not benefit from neoadjuvant chemotherapy. Furthermore, surgery may be avoided in patients who have a pathologic complete response after neoadjuvant chemotherapy [123].

ML was tested in the early prediction of complete response to neoadjuvant chemotherapy and survival outcomes in breast cancer patients through the analysis of multiparametric MR examinations performed on a 3 Tesla equipment [124]. There were 23 features extracted for each lesion: qualitative T2-weighted and DCE-MRI features according to BI-RADS, quantitative pharmacokinetic DCE features (mean plasma flow, volume distribution, mean transit time), and DWI apparent diffusion coefficient values. To apply ML to multiparametric MR examinations, 8 classifiers including linear support vector machine, linear discriminant analysis, logistic regression, random forests, stochastic gradient descent, decision tree, adaptive boosting, and extreme gradient boosting (XGBoost) were applied to rank the features. Histopathologic residual cancer burden class, recurrence-free survival, and disease-specific survival were used as the standards of reference. The study demonstrated the high accuracy of the ML model in predicting both residual cancer burden (AUC, 0.86) and disease-specific survival (AUC, 0.83).

When compared to other classifiers, the XGBoost achieved the most stable performance with high accuracy. Changes in lesion size, complete pattern of shrinkage, mean transit time on DCE-MRI, minimum apparent diffusion coefficient on DWI, and peritumoral edema on T2-weighted imaging, were the most important features for predicting residual cancer burden. On the other hand, volume distribution, mean plasma flow and mean transit time, DCE-MRI lesion size, and minimum, maximum, and mean ADC with DWI were the most important features for predicting recurrence-free survival. On DCE-MRI, the most important features for predicting disease-specific survival were lesion size, volume distribution, and mean plasma flow, as well as maximum ADC with DWI. In a multicenter study, Liu et al. [125] obtained similar results using a radiomics multiparametric model with four radiomic signatures.

Bitencourt et al. [126] used AI in conjunction with clinical variables to assess the complete pathologic response after neoadjuvant chemotherapy in overexpressing HER2 breast cancer and found that it was 83.9% accurate. Another study on HER2-positive cancer responses was conducted by Braman et al. [127], who analyzed intra and peritumoral features. In both validation cohorts, their model was able to identify the HER2 breast cancer subtype with an AUC of 0.89 and predict neoadjuvant chemotherapy response to HER2-targeted therapy (AUC of 0.80 and 0.69, respectively). Cain et al. [128] used pretreatment MR of 288 patients to predict response to neoadjuvant chemotherapy developing a multivariate ML-based model. Twelve features were chosen, six from the tumor alone, five from the fibroglandular tissue alone, and one from both. The “change in variance of uptake”, a tumor-based feature that quantifies the change in variance of tumor uptake in two consecutive time points, was found to be the most relevant feature [128].

Sutton et al. [129] in a study on 273 women with 278 invasive breast cancers, developed a model that combined MRI-extracted radiomics features with molecular subtypes to identify pathologic complete response post neoadjuvant treatment, which showed an AUC of 0.78 on the test set.

The radiomics approach has the potential to extract quantitative imaging biomarkers that can be used to predict DFS. In 620 patients with invasive breast cancer, Xiong et al. [130] evaluated the added value of the US

radiomics signature. Independent of clinicopathological predictors, the radiomics signature was significantly associated with DFS and outperformed the clinicopathological nomogram. Other authors have confirmed the usefulness of the radiogenomic approach in the subgroup of patients with triple-negative breast cancer [131].

A recent study evaluated preoperative MRI of 294 patients with invasive breast cancer and concluded that radiomics nomograms could significantly improve DFS prediction individualization [132]. A study that included patients with HER2-positive invasive breast cancer came to similar conclusions [133].

Cancer recurrence prediction is another relevant clinical issue in patient management and AI-based MRI models have demonstrated their potential in recurrence prediction [134, 135]. For example, Ha et al. [136] used a CNN to predict Oncotype Dx recurrence score, using MRI datasets for the distinction of the low, medium, and high-risk patients, with an overall accuracy of 81% in a three-class prediction with a specificity of 90%, a sensitivity of 60%, and AUC of 0.92.

Kim et al. [137] collected data on 679 breast cancer surgery patients, including histological grade, tumor size, number of metastatic lymph nodes, ER, lymphovascular invasion, local invasion of the tumor, and number of tumors, to develop AI-based models to estimate the risk of recurrence. One of the models obtained demonstrated high sensitivity (0.89), specificity (0.73), positive predictive values (0.75), and negative predictive values (0.89). The authors were adamant that grouping patients into high-risk and low-risk categories help with treatment and follow-up planning.

Finally, studies demonstrated that a radiomics-based AI model was able to predict the presence of sentinel or axillary lymph node metastases [138, 139]. For example, Dietzel et al. [140] demonstrated that a breast MRI-based ANN can predict axillary lymph node metastasis with an AUC of 0.74.

Challenges and perspectives

Publications on AI in medicine have increased exponentially in recent years, demonstrating the great interest in these applications. However, despite the rapidly growing hype, some challenges initially frustrated its clinical application.

In the first instance, clinicians are not always adequately equipped to deal with this paradigm shift. Radiologists in particular feel a pressing threat to their professional expertise coming from these new tools [12]. However, the applications of AI in medical imaging could open up new scenarios allowing radiologists to perform more value-added tasks while avoiding repetitive and time-consuming ones playing a pivotal role in multidisciplinary clinical teams [12, 14].

It is unlikely that radiologists will be replaced because their work includes not only image interpretation [141]. In breast imaging, in particular, radiologists have a close relationship with the patient that includes diagnosis communication, the outline of the diagnostic path based on patient values and preferences, the overall medical judgment integrating different types of clinical information that go beyond imaging, interventional procedures interventional procedure and quality insurance [12, 14]. To take full advantage of the opportunities presented by a quantitative imaging approach radiologists must develop new cross-disciplinary and multidisciplinary skills. This knowledge can hardly be managed by a single radiologist, especially in the research field, which is why multidisciplinary teamwork is becoming more and more important. This fact can be a barrier for small research groups and institutions that lack the financial resources to recruit non-medical professionals to dedicate exclusively to research and support activity in AI.

Other crucial topics concern technical and methodological issues. The need for large amounts of data is a limitation to the development of these models which should be trained, validated, and tested on big data sets. In recent years, the growing popularity of open-source image repositories has partially addressed these issues and encouraged the development of AI-based tools through data sharing, open-access education, and collaborative research [142].

The lack of reproducibility and validation of radiomic studies is considered to be a major challenge in this field. In order to address this problem, the image biomarker standardization initiative (IBSI) was built

as an independent international collaboration that works towards standardizing the extraction of image biomarkers from acquired imaging for the purpose of high-throughput quantitative image analysis [143, 144].

Finally, even validated AI tools with high performance would likely fail to significantly improve patient management if they were not well integrated into existing clinical workflows. Further studies are needed to verify the actual impact that the application of these tools has on patient management and outcomes. Such a type of translational research requires close collaboration between breast radiologists, physicists, statisticians, and suppliers of AI tools. For this reason, creating a shared language and methods is an even more pressing need to face these challenges.

Conclusions

The application of AI in breast imaging has rapidly evolved from the research phase to clinical application. Softwares for automatic lesion detection, macroscopic classification (e.g., malignant/benign), and BD assessment are progressively integrated into the workflows of radiology departments, albeit with some latency in Europe compared to the United States, where these tools are already part of the clinical routine. Furthermore, AI-based algorithms for advanced tasks such as risk stratification, non-invasive molecular characterization of lesions, treatment response prediction, and prognosis are showing promising results, but need to be refined and validated in large case series. It is also essential to understand their real-world impact and how to properly integrate them into the clinical workflow. However, the widespread adoption of these techniques will have a great impact on patient management, moving toward personalized medicine.

Abbreviations

3D: three-dimensional

AI: artificial intelligence

ANN: artificial neural network

AUC: area under the curve

BD: breast density

BI-RADS: breast imaging-reporting and data system

BPE: background parenchymal enhancement

CAD: computer-aided detection

DBT: digital breast tomosynthesis

DCE-MR: dynamic contrast-enhanced magnetic resonance

DFS: disease-free survival

DL: deep learning

DWI: diffusion-weighted imaging

ER: estrogen receptor

FDA: Food and Drug Administration

HER2: human epidermal growth factor receptor 2

IBIS: International Breast Intervention Study

ML: machine learning

MR: magnetic resonance

MRI: magnetic resonance imaging

TC: Tyrer-Cuzick

US: ultrasound

Declarations

Author contributions

M Cè: Conceptualization, Writing—original draft, Writing—review and editing. EC: Investigation, Writing—original draft. MEP: Writing—original draft. MB: Writing—original draft. AS: Writing—original draft. DF: Writing—original draft, Writing—review and editing. GO: Data curation, Supervision. M Cellina: Conceptualization, Supervision, Validation, Writing—review and editing.

Conflicts of interest

The authors declare no conflicts of interest in this manuscript.

Ethical approval

Not applicable.

Consent to participate

Not applicable.

Consent to publication

Not applicable.

Availability of data and materials

Not applicable.

Funding

Not applicable.

Copyright

© The Author(s) 2022.

References

1. Tao Z, Shi A, Lu C, Song T, Zhang Z, Zhao J. Breast cancer: epidemiology and etiology. *Cell Biochem Biophys*. 2015;72:333–8.
2. Veronesi U, Boyle P, Goldhirsch A, Orecchia R, Viale G. Breast cancer. *Lancet*. 2005;365:1727–41.
3. Łukasiewicz S, Czezelewski M, Forma A, Baj J, Sitarz R, Stanisławek A. Breast cancer-epidemiology, risk factors, classification, prognostic markers, and current treatment strategies-an updated review. *Cancers (Basel)*. 2021;13:4287.
4. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. *Nat Rev Cancer*. 2018;18:500–10.
5. Iranmakani S, Mortezaadeh T, Sajadian F, Ghaziani MF, Ghafari A, Khezerloo D, et al. A review of various modalities in breast imaging: technical aspects and clinical outcomes. *Egypt J Radiol Nucl Med*. 2020;51:57.
6. Crivelli P, Ledda RE, Parascandolo N, Fara A, Soro D, Conti M. A new challenge for radiologists: radiomics in breast cancer. *Biomed Res Int*. 2018;2018:6120703.
7. Yala A, Lehman C, Schuster T, Portnoi T, Barzilay R. A deep learning mammography-based model for improved breast cancer risk prediction. *Radiology*. 2019;292:60–6.
8. Couture HD, Williams LA, Geradts J, Nyante SJ, Butler EN, Marron JS, et al. Image analysis with deep learning to predict breast cancer grade, ER status, histologic subtype, and intrinsic subtype. *NPJ Breast Cancer*. 2018;4:30.

9. Shah SM, Khan RA, Arif S, Sajid U. Artificial intelligence for breast cancer analysis: trends & directions. *Comput Biol Med.* 2022;142:105221.
10. Skarping I, Larsson M, Förnvik D. Analysis of mammograms using artificial intelligence to predict response to neoadjuvant chemotherapy in breast cancer patients: proof of concept. *Eur Radiol.* 2022;32:3131–41.
11. Hayashi M, Yamamoto Y, Iwase H. Clinical imaging for the prediction of neoadjuvant chemotherapy response in breast cancer. *Chin Clin Oncol.* 2020;9:31.
12. Pesapane F, Codari M, Sardanelli F. Artificial intelligence in medical imaging: threat or opportunity? Radiologists again at the forefront of innovation in medicine. *Eur Radiol Exp.* 2018;2:35.
13. Houssami N, Lee CI, Buist DSM, Tao D. Artificial intelligence for breast cancer screening: opportunity or hype? *Breast.* 2017;36:31–3.
14. Sechopoulos I, Mann RM. Stand-alone artificial intelligence - the future of breast cancer screening? *Breast.* 2020;49:254–60.
15. Shalev-Shwartz S, Ben-David S. *Understanding machine learning: from theory to algorithms.* 1st ed. Cambridge University Press; 2014.
16. Tan PN, Steinbach M, Karpatne A. *Introduction to data mining, second edition.* Pearson; 2018.
17. Santosh KC, Das N, Ghosh S. *Deep learning models for medical imaging.* 1st ed. Elsevier; 2021.
18. Nielsen MA. *Neural networks and deep learning.* Determination Press; 2015.
19. Lecun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015;521:436–44.
20. Kim M, Yun J, Cho Y, Shin K, Jang R, Bae HJ, et al. Deep learning in medical imaging. *neurospine.* 2019;16:657–68. Erratum in: *Neurospine.* 2020;17:471–2.
21. Vobugari N, Raja V, Sethi U, Gandhi K, Raja K, Surani SR. Advancements in oncology with artificial intelligence-a review article. *Cancers (Basel).* 2022;14:1349.
22. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer.* 2012;48:441–6.
23. Reginelli A, Nardone V, Giacobbe G, Belfiore MP, Grassi R, Schettino F, et al. Radiomics as a new frontier of imaging for cancer prognosis: a narrative review. *Diagnostics (Basel).* 2021;11:1796.
24. Lo Gullo R, Daimiel I, Morris EA, Pinker K. Combining molecular and imaging metrics in cancer: radiogenomics. *Insights Imaging.* 2020;11:1.
25. Cellina M, Pirovano M, Ciocca M, Gibelli D, Floridi C, Oliva G. Radiomic analysis of the optic nerve at the first episode of acute optic neuritis: an indicator of optic nerve pathology and a predictor of visual recovery? *Radiol Med.* 2021;126:698–706.
26. Koçak B, Durmaz EŞ, Ateş E, Kılıçkesmez Ö. Radiomics with artificial intelligence: a practical guide for beginners. *Diagn Interv Radiol.* 2019;25:485–95.
27. Mayerhoefer ME, Materka A, Langs G, Häggström I, Szczypiński P, Gibbs P, et al. Introduction to radiomics. *J Nucl Med.* 2020;61:488–95.
28. Mazurowski MA. Radiogenomics: what it is and why it is important. *J Am Coll Radiol.* 2015;12:862–6.
29. Autier P, Boniol M. Mammography screening: a major issue in medicine. *Eur J Cancer.* 2018;90:34–62.
30. European Commission, Directorate-General for Health and Consumers. *European guidelines for quality assurance in breast cancer screening and diagnosis: fourth edition, supplements.* Publications Office; 2013.
31. Ghoncheh M, Pournamdar Z, Salehiniya H. Incidence and mortality and epidemiology of breast cancer in the world. *Asian Pac J Cancer Prev.* 2016;17:43–6.
32. Pace LE. False-positive results of mammography screening in the era of digital breast tomosynthesis. *JAMA Netw Open.* 2022;5:e222445.

33. Morris E, Feig SA, Drexler M, Lehman C. Implications of overdiagnosis: impact on screening mammography practices. *Popul Health Manag.* 2015;18:S3–11.
34. Friedewald SM, Rafferty EA, Rose SL, Durand MA, Plecha DM, Greenberg JS, et al. Breast cancer screening using tomosynthesis in combination with digital mammography. *JAMA.* 2014;311:2499–507.
35. Schünemann HJ, Lerda D, Quinn C, Follmann M, Alonso-Coello P, Rossi PG, et al.; European Commission Initiative on Breast Cancer (ECIBC) Contributor Group. Breast cancer screening and diagnosis: a synopsis of the European breast guidelines. *Ann Intern Med.* 2020;172:46–56.
36. Taylor-Phillips S, Stinton C. Double reading in breast cancer screening: considerations for policy-making. *Br J Radiol.* 2020;93:20190610.
37. Lehman CD, Wellman RD, Buist DS, Kerlikowske K, Tosteson AN, Miglioretti DL; Breast Cancer Surveillance Consortium. Diagnostic accuracy of digital screening mammography with and without computer-aided detection. *JAMA Intern Med.* 2015;175:1828–37.
38. Bahl M. Updates in artificial intelligence for breast imaging. *Semin Roentgenol.* 2022;57:160–7.
39. Rojas K, Stuckey A. Breast cancer epidemiology and risk factors. *Clin Obstet Gynecol.* 2016;59:651–72.
40. Sun YS, Zhao Z, Yang ZN, Xu F, Lu HJ, Zhu ZY, et al. Risk factors and preventions of breast cancer. *Int J Biol Sci.* 2017;13:1387–97.
41. Vinnicombe SJ. Breast density: why all the fuss? *Clin Radiol.* 2018;73:334–57.
42. Boyd NF. Mammographic density and risk of breast cancer. *Am Soc Clin Oncol Educ Book.* 2013;33:e57–62.
43. Pinsky RW, Helvie MA. Mammographic breast density: effect on imaging and breast cancer risk. *J Natl Compr Canc Netw.* 2010;8:1157–64; quiz 1165.
44. ACR statement on reporting breast density in mammography reports and patient summaries [Internet]. American College of Radiology. [cited 2022 Oct 10]. Available from: <https://www.acr.org/Advocacy-and-Economics/ACR-Position-Statements/Reporting-Breast-Density>
45. Nicholson BT, LoRusso AP, Smolkin M, Bovbjerg VE, Petroni GR, Harvey JA. Accuracy of assigned BI-RADS breast density category definitions. *Acad Radiol.* 2006;13:1143–9.
46. Sprague BL, Conant EF, Onega T, Garcia MP, Beaber EF, Herschorn SD, et al. Variation in mammographic breast density assessments among radiologists in clinical practice: a multicenter observational study. *Ann Intern Med.* 2016;165:457–64.
47. Holland K, van Zelst J, den Heeten GJ, Imhof-Tas M, Mann RM, van Gils CH, et al. Consistency of breast density categories in serial screening mammograms: a comparison between automated and human assessment. *Breast.* 2016;29:49–54.
48. Destounis S, Arieno A, Morgan R, Roberts C, Chan A. Qualitative *versus* quantitative mammographic breast density assessment: applications for the US and abroad. *Diagnostics (Basel).* 2017;7:30.
49. Engmann NJ, Golmakani MK, Miglioretti DL, Sprague BL, Kerlikowske K; Breast Cancer Surveillance Consortium. Population-attributable risk proportion of clinical risk factors for breast cancer. *JAMA Oncol.* 2017;3:1228–36. Erratum in: *JAMA Oncol.* 2019;5:1643.
50. Giorgi Rossi P, Djuric O, Hélin V, Astley S, Mantellini P, Nitrosi A, et al. Validation of a new fully automated software for 2D digital mammographic breast density evaluation in predicting breast cancer risk. *Sci Rep.* 2021;11:19884.
51. Keller BM, Nathan DL, Wang Y, Zheng Y, Gee JC, Conant EF, et al. Estimation of breast percent density in raw and processed full field digital mammography images via adaptive fuzzy c-means clustering and support vector machine segmentation. *Med Phys.* 2012;39:4903–17.
52. Ciatto S, Bernardi D, Calabrese M, Durando M, Gentilini MA, Mariscotti G, et al. A first evaluation of breast radiological density assessment by QUANTRA software as compared to visual classification. *Breast.* 2012;21:503–6.

53. Alain G, Bengio Y. Understanding intermediate layers using linear classifier probes. arXiv:1610.01644 [Preprint]. [posted 2016 Oct 5; revised 2016 Oct 10; revised 2016 Oct 14; revised 2018 Nov 22; cited 2022 Oct 10]. Available from: <https://doi.org/10.48550/arXiv.1610.01644>
54. Pahwa S, Hari S, Thulkar S, Angraal S. Evaluation of breast parenchymal density with QUANTRA software. *Indian J Radiol Imaging*. 2015;25:391–6.
55. Ekpo EU, McEntee MF, Rickard M, Brennan PC, Kunduri J, Demchig D, et al. Quantra™ should be considered a tool for two-grade scale mammographic breast density classification. *Br J Radiol*. 2016;89:20151057.
56. Haji Maghsoudi O, Gastouniotti A, Scott C, Pantalone L, Wu FF, Cohen EA, et al. Deep-LIBRA: an artificial-intelligence method for robust quantification of breast density with independent validation in breast cancer risk assessment. *Med Image Anal*. 2021;73:102138.
57. Gastouniotti A, Conant EF, Kontos D. Beyond breast density: a review on the advancing role of parenchymal texture analysis in breast cancer risk assessment. *Breast Cancer Res*. 2016;18:91.
58. Arefan D, Mohamed AA, Berg WA, Zuley ML, Sumkin JH, Wu S. Deep learning modeling using normal mammograms for predicting breast cancer risk. *Med Phys*. 2020;47:110–8.
59. Gail MH, Costantino JP, Pee D, Bondy M, Newman L, Selvan M, et al. Projecting individualized absolute invasive breast cancer risk in African American women. *J Natl Cancer Inst*. 2007;99:1782–92. Erratum in: *J Natl Cancer Inst*. 2008;100:1118. Erratum in: *J Natl Cancer Inst*. 2008;100:373.
60. Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. *Stat Med*. 2004;23:1111–30. Erratum in: *Stat Med*. 2005;24:156.
61. Tice JA, Cummings SR, Smith-Bindman R, Ichikawa L, Barlow WE, Kerlikowske K. Using clinical factors and mammographic breast density to estimate breast cancer risk: development and validation of a new predictive model. *Ann Intern Med*. 2008;148:337–47.
62. Paci E, Mantellini P, Giorgi Rossi P, Falini P, Puliti D; TBST Working Group. Tailored breast screening trial (TBST). *Epidemiol Prev*. 2013;37:317–27. Italian.
63. Esserman LJ; WISDOM Study and Athena Investigators. The WISDOM study: breaking the deadlock in the breast cancer screening debate. *NPJ Breast Cancer*. 2017;3:34.
64. My personalized breast screening (MyPeBS) [Internet]. Source: National Library of Medicine; [cited 2022 Oct 10]. Available from: <https://clinicaltrials.gov/ct2/show/study/NCT03672331>
65. Valero MG, Zabor EC, Park A, Gilbert E, Newman A, King TA, et al. The Tyrer-Cuzick model inaccurately predicts invasive breast cancer risk in women with LCIS. *Ann Surg Oncol*. 2020;27:736–40.
66. Rodriguez-Ruiz A, Lång K, Gubern-Merida A, Teuwen J, Broeders M, Gennaro G, et al. Can we reduce the workload of mammographic screening by automatic identification of normal exams with artificial intelligence? A feasibility study. *Eur Radiol*. 2019;29:4825–32.
67. Dembrower K, Wählin E, Liu Y, Salim M, Smith K, Lindholm P, et al. Effect of artificial intelligence-based triaging of breast cancer screening mammograms on cancer detection and radiologist workload: a retrospective simulation study. *Lancet Digit Health*. 2020;2:e468–74.
68. Balta C, Rodriguez-Ruiz A, Mieskes C, Karssemeijer N, Heywang-Köbrunner SH. Going from double to single reading for screening exams labeled as likely normal by AI: what is the impact? In: Bosmans H, Marshall N, Van Ongeval C, editors. 15th International workshop on breast imaging (IWBI2020). SPIE Proceedings; 2020.
69. Le EPV, Wang Y, Huang Y, Hickman S, Gilbert FJ. Artificial intelligence in breast imaging. *Clin Radiol*. 2019;74:357–66.
70. Kim HE, Kim HH, Han BK, Kim KH, Han K, Nam H, et al. Changes in cancer detection and false-positive recall in mammography using artificial intelligence: a retrospective, multireader study. *Lancet Digit Health*. 2020;2:e138–48.
71. Schaffter T, Buist DSM, Lee CI, Nikulin Y, Ribli D, Guan Y, et al.; DM DREAM Consortium; Mackey L, Cahoon J, Shen L, Sohn JH, Trivedi H, Shen Y, et al. Evaluation of combined artificial intelligence and

- radiologist assessment to interpret screening mammograms. *JAMA Netw Open*. 2020;3:e200265. Erratum in: *JAMA Netw Open*. 2020;3:e204429.
72. Lång K, Hofvind S, Rodríguez-Ruiz A, Andersson I. Can artificial intelligence reduce the interval cancer rate in mammography screening? *Eur Radiol*. 2021;31:5940–7.
 73. Larsen M, Aglen CF, Lee CI, Hoff SR, Lund-Hanssen H, Lång K, et al. Artificial intelligence evaluation of 122 969 mammography examinations from a population-based screening program. *Radiology*. 2022;303:502–11.
 74. Kim HJ, Kim HH, Kim KH, Choi WJ, Chae EY, Shin HJ, et al. Mammographically occult breast cancers detected with AI-based diagnosis supporting software: clinical and histopathologic characteristics. *Insights Imaging*. 2022;13:57.
 75. Mammography screening with artificial intelligence (MASAI) (MASAI) [Internet]. Source: National Library of Medicine; [cited 2022 Oct 10]. Available from: <https://www.clinicaltrials.gov/ct2/show/NCT04838756>
 76. Bai J, Posner R, Wang T, Yang C, Nabavi S. Applying deep learning in digital breast tomosynthesis for automatic breast cancer detection: a review. *Med Image Anal*. 2021;71:102049.
 77. Geras KJ, Mann RM, Moy L. Artificial intelligence for mammography and digital breast tomosynthesis: current concepts and future perspectives. *Radiology*. 2019;293:246–59.
 78. van Winkel SL, Rodríguez-Ruiz A, Appelman L, Gubern-Mérida A, Karssemeijer N, Teuwen J, et al. Impact of artificial intelligence support on accuracy and reading time in breast tomosynthesis image interpretation: a multi-reader multi-case study. *Eur Radiol*. 2021;31:8682–91.
 79. Conant EF, Toledano AY, Periaswamy S, Fotin SV, Go J, Boatsman JE, et al. Improving accuracy and efficiency with concurrent use of artificial intelligence for digital breast tomosynthesis. *Radiol Artif Intell*. 2019;1:e180096.
 80. Buda M, Saha A, Walsh R, Ghate S, Li N, Swiecicki A, et al. A data set and deep learning algorithm for the detection of masses and architectural distortions in digital breast tomosynthesis images. *JAMA Netw Open*. 2021;4:e2119100.
 81. Breast cancer screening – digital breast tomosynthesis (BCS-DBT) [Internet]. TCIA; c2014–2020 [cited 2022 Oct 10]. Available from: <https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=64685580>
 82. Jiang Y, Edwards AV, Newstead GM. Artificial intelligence applied to breast MRI for improved diagnosis. *Radiology*. 2021;298:38–46.
 83. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature*. 2012;490:61–70.
 84. Johnson KS, Conant EF, Soo MS. Molecular subtypes of breast cancer: a review for breast radiologists. *J Breast Imaging*. 2021;3:12–24.
 85. Roulot A, Héquet D, Guinebretière JM, Vincent-Salomon A, Lerebours F, Dubot C, et al. Tumoral heterogeneity of breast cancer. *Ann Biol Clin (Paris)*. 2016;74:653–60.
 86. Scapicchio C, Gabelloni M, Barucci A, Cioni D, Saba L, Neri E. A deep look into radiomics. *Radiol Med*. 2021;126:1296–311.
 87. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol*. 2017;14:749–62.
 88. Satake H, Ishigaki S, Ito R, Naganawa S. Radiomics in breast MRI: current progress toward clinical application in the era of artificial intelligence. *Radiol Med*. 2022;127:39–56.
 89. Scimeca M, Urbano N, Toschi N, Bonanno E, Schillaci O. Precision medicine in breast cancer: from biological imaging to artificial intelligence. *Semin Cancer Biol*. 2021;72:1–3.

90. Consensus guideline on image-guided percutaneous biopsy of palpable and nonpalpable breast lesions [Internet]. The American Society of Breast Surgeons; c2018 [cited 2022 Oct 10]. Available from: <https://www.breastsurgeons.org/docs/statements/Consensus-Guideline-on-Concordance-Assessment-of-Image-Guided-Breast-Biopsies.pdf>
91. Bai HX, Lee AM, Yang L, Zhang P, Davatzikos C, Maris JM, et al. Imaging genomics in cancer research: limitations and promises. *Br J Radiol*. 2016;89:20151030.
92. Sala E, Mema E, Himoto Y, Veeraraghavan H, Brenton JD, Snyder A, et al. Unravelling tumour heterogeneity using next-generation imaging: radiomics, radiogenomics, and habitat imaging. *Clin Radiol*. 2017;72:3–10.
93. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun*. 2014;5:4006. Erratum in: *Nat Commun*. 2014;5:4644.
94. Vicini S, Bortolotto C, Rengo M, Ballerini D, Bellini D, Carbone I, et al. A narrative review on current imaging applications of artificial intelligence and radiomics in oncology: focus on the three most common cancers. *Radiol Med*. 2022;127:819–36.
95. Ma Y, Shan D, Wei J, Chen A. Application of intravoxel incoherent motion diffusion-weighted imaging in differential diagnosis and molecular subtype analysis of breast cancer. *Am J Transl Res*. 2021;13:3034–43.
96. Wang XY, Cui LG, Feng J, Chen W. Artificial intelligence for breast ultrasound: an adjunct tool to reduce excessive lesion biopsy. *Eur J Radiol*. 2021;138:109624.
97. Yamamoto S, Han W, Kim Y, Du L, Jamshidi N, Huang D, et al. Breast cancer: radiogenomic biomarker reveals associations among dynamic contrast-enhanced MR imaging, long noncoding RNA, and metastasis. *Radiology*. 2015;275:384–92.
98. Zhu Y, Li H, Guo W, Drukker K, Lan L, Giger ML, et al. Deciphering genomic underpinnings of quantitative MRI-based radiomic phenotypes of invasive breast carcinoma. *Sci Rep*. 2015;5:17787.
99. Fiz F, Viganò L, Gennaro N, Costa G, La Bella L, Boichuk A, et al. Radiomics of liver metastases: a systematic review. *Cancers (Basel)*. 2020;12:2881.
100. Parekh VS, Jacobs MA. Integrated radiomic framework for breast cancer and tumor biology using advanced machine learning and multiparametric MRI. *NPJ Breast Cancer*. 2017;3:43.
101. Whitney HM, Taylor NS, Drukker K, Edwards AV, Papaioannou J, Schacht D, et al. Additive benefit of radiomics over size alone in the distinction between benign lesions and luminal A cancers on a large clinical breast MRI dataset. *Acad Radiol*. 2019;26:202–9.
102. Elias SG, Adams A, Wisner DJ, Esserman LJ, van't Veer LJ, Mali WP, et al. Imaging features of HER2 overexpression in breast cancer: a systematic review and meta-analysis. *Cancer Epidemiol Biomarkers Prev*. 2014;23:1464–83.
103. Leithner D, Mayerhoefer ME, Martinez DF, Jochelson MS, Morris EA, Thakur SB, et al. Non-invasive assessment of breast cancer molecular subtypes with multiparametric magnetic resonance imaging radiomics. *J Clin Med*. 2020;9:1853.
104. Li H, Zhu Y, Burnside ES, Huang E, Drukker K, Hoadley KA, et al. Quantitative MRI radiomics in the prediction of molecular classifications of breast cancer subtypes in the TCGA/TCIA data set. *NPJ Breast Cancer*. 2016;2:16012.
105. Yeh AC, Li H, Zhu Y, Zhang J, Khramtsova G, Drukker K, et al. Radiogenomics of breast cancer using dynamic contrast enhanced MRI and gene expression profiling. *Cancer Imaging*. 2019;19:48.
106. Homayoun H, Ebrahimpour-Komleh H. Automated segmentation of abnormal tissues in medical images. *J Biomed Phys Eng*. 2021;11:415–24.

107. Cappella A, Gibelli D, Cellina M, Mazzarelli D, Oliva AG, De Angelis D, et al. Three-dimensional analysis of sphenoid sinus uniqueness for assessing personal identification: a novel method based on 3D-3D superimposition. *Int J Legal Med.* 2019;133:1895–901.
108. Jiang L, Hu X, Xiao Q, Gu Y, Li Q. Fully automated segmentation of whole breast using dynamic programming in dynamic contrast enhanced MR images. *Med Phys.* 2017;44:2400–14.
109. Zhang L, Mohamed AA, Chai R, Guo Y, Zheng B, Wu S. Automated deep learning method for whole-breast segmentation in diffusion-weighted breast MRI. *J Magn Reson Imaging.* 2020;51:635–43.
110. Ma X, Wang J, Zheng X, Liu Z, Long W, Zhang Y, et al. Automated fibroglandular tissue segmentation in breast MRI using generative adversarial networks. *Phys Med Biol.* 2020;65:105006.
111. Huo L, Hu X, Xiao Q, Gu Y, Chu X, Jiang L. Segmentation of whole breast and fibroglandular tissue using nnU-Net in dynamic contrast enhanced MR images. *Magn Reson Imaging.* 2021;82:31–41.
112. Liao GJ, Henze Bancroft LC, Strigel RM, Chitalia RD, Kontos D, Moy L, et al. Background parenchymal enhancement on breast MRI: a comprehensive review. *J Magn Reson Imaging.* 2020;51:43–61.
113. Saha A, Grimm LJ, Ghate SV, Kim CE, Soo MS, Yoon SC, et al. Machine learning-based prediction of future breast cancer using algorithmically measured background parenchymal enhancement on high-risk screening MRI. *J Magn Reson Imaging.* 2019;50:456–64.
114. Wei D, Jahani N, Cohen E, Weinstein S, Hsieh MK, Pantalone L, et al. Fully automatic quantification of fibroglandular tissue and background parenchymal enhancement with accurate implementation for axial and sagittal breast MRI protocols. *Med Phys.* 2021;48:238–52.
115. Ha R, Chang P, Mema E, Mutasa S, Karcich J, Wynn RT, et al. Fully automated convolutional neural network method for quantification of breast MRI fibroglandular tissue and background parenchymal enhancement. *J Digit Imaging.* 2019;32:141–7.
116. Lee CY, Chang TF, Chou YH, Yang KC. Fully automated lesion segmentation and visualization in automated whole breast ultrasound (ABUS) images. *Quant Imaging Med Surg.* 2020;10:568–84.
117. Chung SY, Chang JS, Choi MS, Chang Y, Choi BS, Chun J, et al. Clinical feasibility of deep learning-based auto-segmentation of target volumes and organs-at-risk in breast cancer patients after breast-conserving surgery. *Radiat Oncol.* 2021;16:44.
118. Byun HK, Chang JS, Choi MS, Chun J, Jung J, Jeong C, et al. Evaluation of deep learning-based autosegmentation in breast cancer radiotherapy. *Radiat Oncol.* 2021;16:203.
119. Nardone V, Reginelli A, Grassi R, Boldrini L, Vacca G, D’Ippolito E, et al. Delta radiomics: a systematic review. *Radiol Med.* 2021;126:1571–83.
120. Martín M, González Palacios F, Cortés J, de la Haba J, Schneider J. Prognostic and predictive factors and genetic analysis of early breast cancer. *Clin Transl Oncol.* 2009;11:634–42.
121. Cianfrocca M, Goldstein LJ. Prognostic and predictive factors in early-stage breast cancer. *Oncologist.* 2004;9:606–16.
122. Reig B, Heacock L, Geras KJ, Moy L. Machine learning in breast MRI. *J Magn Reson Imaging.* 2020;52:998–1018.
123. Bitencourt A, Daimiel Naranjo I, Lo Gullo R, Rossi Saccarelli C, Pinker K. AI-enhanced breast imaging: where are we and where are we heading? *Eur J Radiol.* 2021;142:109882.
124. Tahmassebi A, Wengert GJ, Helbich TH, Bago-Horvath Z, Alaei S, Bartsch R, et al. Impact of machine learning with multiparametric magnetic resonance imaging of the breast for early prediction of response to neoadjuvant chemotherapy and survival outcomes in breast cancer patients. *Invest Radiol.* 2019;54:110–7.
125. Liu Z, Li Z, Qu J, Zhang R, Zhou X, Li L, et al. Radiomics of multiparametric MRI for pretreatment prediction of pathologic complete response to neoadjuvant chemotherapy in breast cancer: a multicenter study. *Clin Cancer Res.* 2019;25:3538–47.

126. Bitencourt AGV, Gibbs P, Rossi Saccarelli C, Daimiel I, Lo Gullo R, Fox MJ, et al. MRI-based machine learning radiomics can predict HER2 expression level and pathologic response after neoadjuvant therapy in HER2 overexpressing breast cancer. *EBioMedicine*. 2020;61:103042.
127. Braman N, Prasanna P, Whitney J, Singh S, Beig N, Etesami M, et al. Association of peritumoral radiomics with tumor biology and pathologic response to preoperative targeted therapy for HER2 (ERBB2)-positive breast cancer. *JAMA Netw Open*. 2019;2:e192561.
128. Cain EH, Saha A, Harowicz MR, Marks JR, Marcom PK, Mazurowski MA. Multivariate machine learning models for prediction of pathologic response to neoadjuvant therapy in breast cancer using MRI features: a study using an independent validation set. *Breast Cancer Res Treat*. 2019;173:455–63.
129. Sutton EJ, Onishi N, Fehr DA, Dashevsky BZ, Sadinski M, Pinker K, et al. A machine learning model that classifies breast cancer pathologic complete response on MRI post-neoadjuvant chemotherapy. *Breast Cancer Res*. 2020;22:57.
130. Xiong L, Chen H, Tang X, Chen B, Jiang X, Liu L, et al. Ultrasound-based radiomics analysis for predicting disease-free survival of invasive breast cancer. *Front Oncol*. 2021;11:621993.
131. Yu F, Hang J, Deng J, Yang B, Wang J, Ye X, et al. Radiomics features on ultrasound imaging for the prediction of disease-free survival in triple negative breast cancer: a multi-institutional study. *Br J Radiol*. 2021;94:20210188.
132. Park H, Lim Y, Ko ES, Cho HH, Lee JE, Han BK, et al. Radiomics signature on magnetic resonance imaging: association with disease-free survival in patients with invasive breast cancer. *Clin Cancer Res*. 2018;24:4705–14.
133. Li Q, Xiao Q, Li J, Duan S, Wang H, Gu Y. MRI-based radiomic signature as a prognostic biomarker for HER2-positive invasive breast cancer treated with NAC. *Cancer Manag Res*. 2020;12:10603–13.
134. Ashraf AB, Daye D, Gavenonis S, Mies C, Feldman M, Rosen M, et al. Identification of intrinsic imaging phenotypes for breast cancer tumors: preliminary associations with gene expression profiles. *Radiology*. 2014;272:374–84.
135. Wan T, Bloch BN, Plecha D, Thompson CL, Gilmore H, Jaffe C, et al. A radio-genomics approach for identifying high risk estrogen receptor-positive breast cancers on DCE-MRI: preliminary results in predicting OncotypeDX risk scores. *Sci Rep*. 2016;6:21394.
136. Ha R, Chang P, Mutasa S, Karcich J, Goodman S, Blum E, et al. Convolutional neural network using a breast MRI tumor dataset can predict Oncotype Dx recurrence score. *J Magn Reson Imaging*. 2019;49:518–24.
137. Kim W, Kim KS, Lee JE, Noh DY, Kim SW, Jung YS, et al. Development of novel breast cancer recurrence prediction model using support vector machine. *J Breast Cancer*. 2012;15:230–8.
138. Dong Y, Feng Q, Yang W, Lu Z, Deng C, Zhang L, et al. Preoperative prediction of sentinel lymph node metastasis in breast cancer based on radiomics of T2-weighted fat-suppression and diffusion-weighted MRI. *Eur Radiol*. 2018;28:582–91.
139. Han L, Zhu Y, Liu Z, Yu T, He C, Jiang W, et al. Radiomic nomogram for prediction of axillary lymph node metastasis in breast cancer. *Eur Radiol*. 2019;29:3820–9.
140. Dietzel M, Baltzer PA, Dietzel A, Vag T, Gröschel T, Gajda M, et al. Application of artificial neural networks for the prediction of lymph node metastases to the ipsilateral axilla - initial experience in 194 patients using magnetic resonance mammography. *Acta Radiol*. 2010;51:851–8.
141. Pesapane F, Tantrige P, Patella F, Biondetti P, Nicosia L, Ianniello A, et al. Myths and facts about artificial intelligence: why machine- and deep-learning will not replace interventional radiologists. *Med Oncol*. 2020;37:40.
142. Prior F, Almeida J, Kathiravelu P, Kurc T, Smith K, Fitzgerald TJ, et al. Open access image repositories: high-quality data to enable machine learning research. *Clin Radiol*. 2020;75:7–12.

143. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology*. 2020;295:328–38.
144. Nan Y, Ser JD, Walsh S, Schönlieb C, Roberts M, Selby I, et al. Data harmonisation for information fusion in digital healthcare: a state-of-the-art systematic review, meta-analysis and future research directions. *Inf Fusion*. 2022;82:99–122.