

Applications of Artificial Intelligence in Breast Pathology

Yueping Liu, MD; Dandan Han, MD; Anil V. Parwani, MD, PhD; Zaibo Li, MD, PhD

• **Context.**—Increasing implementation of whole slide imaging together with digital workflow and advances in computing capacity enable the use of artificial intelligence (AI) in pathology, including breast pathology. Breast pathologists often face a significant workload, with diagnosis complexity, tedious repetitive tasks, and semi-quantitative evaluation of biomarkers. Recent advances in developing AI algorithms have provided promising approaches to meet the demand in breast pathology.

Objective.—To provide an updated review of AI in breast pathology. We examined the success and challenges of current and potential AI applications in diagnosing and grading breast carcinomas and other pathologic changes,

Pathology is the gold standard for disease diagnosis. Pathologists are committed to assisting clinicians in the precise treatment of various diseases, especially malignant tumors. The advance in individualized treatment requires higher standards for pathologic diagnosis. The pathologic diagnosis of tumors involves many aspects of analysis and interpretation, such as morphologic identification of tumor cells, evaluation of mitotic counts, determination of lymph node metastasis, interpretation of various biomarkers, etc. However, the accuracy of pathologic diagnosis is limited by a shortage of pathologists, differences in pathologists' diagnostic skill, and the availability of ancillary studies.¹⁻⁴

The emergence of whole slide imaging together with a digital workflow has brought many changes to pathology practice, including digital primary sign-out, remote consultations, archiving slides for teaching, research, and conferences, etc. More importantly, a digital workflow with whole slide images (WSIs) paves the road to implement AI algorithms in routine pathology practice. AI can reduce tedious workload for pathologists, improve their efficiency and accuracy, provide new information of disease prognosis and therapy response, etc. Machine learning (ML), a subfield of AI, develops algorithms to learn repetitive data patterns from a large data set of cases and then to match new cases to the learned data patterns. In recent years, deep learning (DL) algorithms have gained increasing attention

detecting lymph node metastasis, quantifying breast cancer biomarkers, predicting prognosis and therapy response, and predicting potential molecular changes.

Data Sources.—We obtained data and information by searching and reviewing literature on AI in breast pathology from PubMed and based our own experience.

Conclusions.—With the increasing application in breast pathology, AI not only assists in pathology diagnosis to improve accuracy and reduce pathologists' workload, but also provides new information in predicting prognosis and therapy response.

(*Arch Pathol Lab Med.* doi: 10.5858/arpa.2022-0457-RA)

in big data and image processing. DL is a subset of ML using artificial neural networks composed of multiple layers (input layer, hidden layers, and output layer) to extract progressively higher-level features from data. The neural network learns data patterns by generating multiple hidden variables from data and also learns hierarchical representations of sophisticated data patterns that cannot be easily identified by humans. DL includes supervised learning, semisupervised learning, unsupervised learning, and transfer learning.⁵⁻⁸ Deep convolutional neural network, one of the DL algorithms, has shown superiority in image recognition and analysis and been used in image-based detection and segmentation to identify and quantify different cells. It can also provide accurate prognosis and identify potential drug target indicators.^{9,10} Currently, research using ML and pathologic images is a hot topic in medical fields.¹¹⁻¹³

Breast pathologists often face a significant workload, with diagnosis complexity and tedious repetitive tasks, such as the quantification of biomarkers and the evaluation of lymph node metastasis. These tasks are time-consuming, labor-intensive, and subjected to interobserver variability. Recent advances in digital pathology workflow integrated with AI have provided promising approaches to solve these challenges. In this review, we summarize current and potential AI applications in diagnosing and grading breast carcinomas, predicting prognosis and therapy response, quantifying breast cancer biomarkers, detecting lymph node metastasis, and predicting potential molecular changes.

AI-BASED DETECTION AND CLASSIFICATION OF BREAST CARCINOMA AND OTHER LESIONS

With the advent of the era of precision medicine, the accurate classification, grading, and determination of the tumor extent of breast cancers have become important for clinical management. Histopathologic diagnosis of breast

Accepted for publication November 15, 2022.

From the Department of Pathology, The Fourth Hospital of Hebei Medical University, Shijiazhuang, China (Liu, Han); and the Department of Pathology, The Ohio State University, Columbus (Parwani, Li).

The authors have no relevant financial interest in the products or companies described in this article.

Corresponding author: Zaibo Li, MD, PhD, Department of Pathology, The Ohio State University, 410 W 10th Ave, Columbus, OH 43210 (email: Zaibo.Li@osumc.edu).

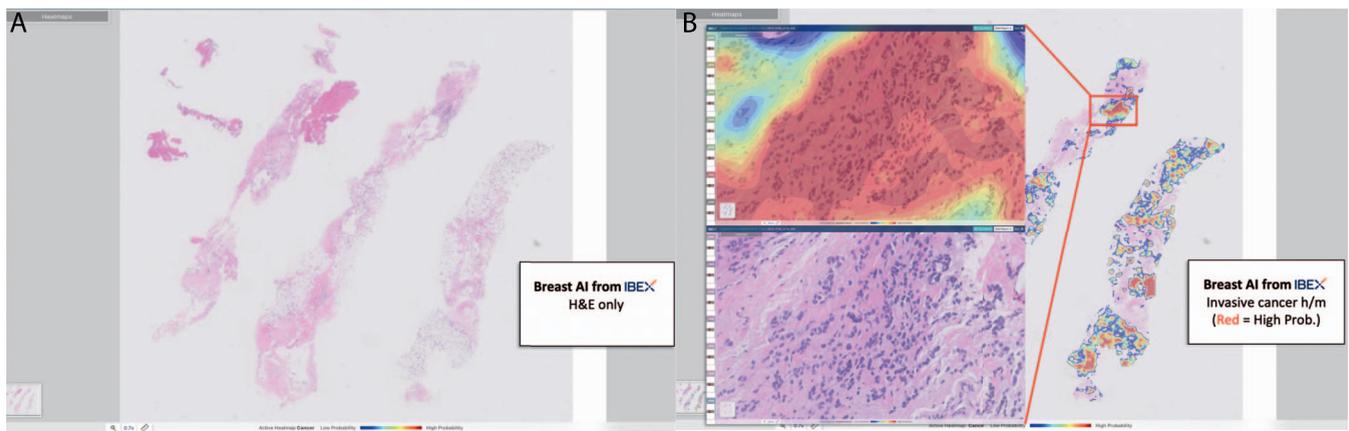


Figure 1. IBEX GALEN Breast detects invasive lobular carcinoma in breast core biopsy. A, Hematoxylin-eosin (H&E) slide. B, Annotated images with invasive lobular carcinoma area highlighted in red (H&E, original magnifications $\times 0.5$ [A] and $\times 10$ [B]).

cancer is the basis for clinical management, and accurate classification of breast cancer is crucial for treatment options. Recently, researchers have developed ML/DL algorithms to detect and classify breast cancers. Han et al¹⁴ proposed a novel DL model to automate multi-classification of breast cancer histopathologic types, such as ductal carcinoma, lobular carcinoma, mucinous carcinoma, papillary carcinoma, etc. The model was validated on a large-scale data set with a high level of performance (average accuracy of 93.2%). Cruz-Roa et al¹⁵ built a convolutional neural network (CNN) model to classify image blocks (“patches”) from breast cancer WSIs containing invasive ductal carcinoma, and then used a ConvNet classifier to estimate the extent of invasive foci and degree of infiltration on entire WSIs. The study used manually annotated region labels from 400 slides from multiple institutions to train the model and validated it on 200 annotated slides obtained from The Cancer Genome Atlas (TCGA) with a pixel-level F1 score of 75.86%. The model was able to detect invasive carcinoma regions on WSIs with high accuracy, even when tested on a validation set from a different cohort. They found that lesions of invasive carcinoma mixed with in situ carcinoma were most challenging, but the performance could be improved by training a more complex algorithm on a data set with more in situ carcinomas.

Several AI platforms (algorithms) are commercially available to detect/screen breast lesions from breast core biopsy specimens. One example is the GALEN Breast algorithm (IBEX; (Figure 1). The algorithm can screen entire breast core needle biopsy WSIs to produce heat maps for different breast lesions, including invasive carcinoma (ductal and lobular), in situ carcinoma (ductal and lobular), and atypical hyperplasia (ductal and lobular), and benign findings, such as sclerosing adenosis, fat necrosis, etc. The algorithm was based on an ensemble of CNNs trained on more than 2 million labeled image patches that were extracted from manual annotations on 2153 hematoxylin-eosin (H&E) slides. A separate different data set of 436 breast biopsies was used to test the algorithm’s performance and the results demonstrated an area under the curve (AUC) of 0.99 for the detection of invasive carcinoma and an AUC of 0.98 for the detection of ductal carcinoma in situ (DCIS). The algorithm differentiated well between subtypes/grades of invasive and in situ carcinoma, with an AUC of 0.97 for

invasive ductal carcinoma versus invasive lobular carcinoma and an AUC of 0.92 for high-grade DCIS versus low-grade DCIS/atypical ductal hyperplasia, respectively.¹⁶

AI-BASED GRADING OF BREAST CARCINOMA

The Nottingham histologic grading system is the most used grading system for invasive breast carcinoma and is well correlated with prognosis and outcome. The Nottingham grading system includes 3 components: tubule formation, nuclear pleomorphism, and mitotic activity. Pathologists usually evaluate these 3 components manually, and interobserver variability does exist.¹⁷ Several studies had shown that DL neural networks can improve the accuracy of histologic grading assessment.^{18–21}

Mitotic activity is an important component of Nottingham grading system and a key predictor for breast carcinoma aggressiveness. Mitotic activity is mainly determined by manual count, which is time-consuming and labor-intensive. Recently, studies have reported AI-assisted mitotic count in breast cancer.^{22–24} Nateghi et al²³ proposed a fully automated system to count mitosis. They first constructed a DL model to detect regions of interest (ROIs) by selecting hotspot regions from WSIs, then trained deep neural networks to identify nuclear mitosis in the selected ROIs. Their results showed the model significantly improved the accuracy of tumor proliferation. Li et al²⁴ proposed an AI model to identify mitosis by using a novel multistage DL framework with the following components: (1) a deep segmentation network for dividing mitotic regions when only weak markers were present, (2) a deep detection network for locating mitoses using information from the upper and lower regions, and (3) a validation network that improved the detection accuracy by eliminating false positive mitosis. They validated their method on the 2012 international conference on pattern recognition (ICPR) grand challenge data set and the 2014 ICPR MITOS-ATYPIA challenge data set to achieve the highest F-score of 0.832 on the ICPR 2012 grand challenge data set and F-score of 0.572 on the 2014 ICPR MITOS-ATYPIA challenge data set. Sebai et al²⁵ developed “MaskMitosis” DL framework to estimate mitosis on 2012 ICPR grand challenge and 2014 ICPR MITOS-ATYPIA challenge data sets and their method outperformed all state-of-the-art mitosis detection approaches on the 2014 ICPR data set by achieving an F-score of 0.475. Mahmood et al²⁶ developed a

mitotic cell detection method based on Faster region CNN (Faster R-CNN) and deep CNNs using ICPR 2012 and ICPR 2014 (MITOS-ATYPIA-14) data sets and demonstrated their method achieved the results of 0.876 precision, 0.841 recall, and 0.858 F1-measure for the ICPR 2012 data set, and 0.848 precision, 0.583 recall, and 0.691 F1-measure for the ICPR 2014 data set, which were higher than those obtained using previous methods.

Nuclear pleomorphism is another component of the Nottingham grading system. Morphologic features of nuclear pleomorphism include nuclear size, nuclear contour, nucleolus, vesicular nuclei, and chromatin clump. AI algorithms have the potential to objectively assess these features with improved reproducibility compared with pathologists' visual assessment.²⁷

There has been a paucity of studies to evaluate tubule formation, another component of the Nottingham grading system. One study developed a DL classifier to automatically identify tubule nuclei from whole slide imaging and then obtain the ratio of tubule nuclei to overall number of nuclei (a tubule formation indicator) to correlate with the corresponding Oncotype DX (Genomic Health, Redwood City, California) risk categories. The study demonstrated a good correlation between a larger tubule formation indicator and low Oncotype DX score.²⁸ The Oncotype DX assay assesses the expression levels of 21 genes involved in the pathways of proliferation, invasion, estrogen and human epidermal growth factor receptor 2 (HER2) signaling to generate a recurrent score (RS) to predict possibility of recurrence and chemotherapy benefit.^{29,30} Oncotype DX RS has been widely accepted in clinical practice across the United States and Canada to guide decisions on adjuvant systemic chemotherapy if a patient has ER⁺, HER2⁻ breast cancer.³¹

Furthermore, Elsharawy et al²² proposed a supervised CNN model to render an AI grade of breast cancers based on nuclear features. Both AI grade and routine Nottingham grade were used to train models to evaluate their correlation with molecular changes and prognosis. Their results showed that AI grade was helpful to identify genetic changes with important prognostic significance. Wang et al¹⁷ developed and validated a novel histologic grading model using a DL method trained on WSIs and DL to analyze Nottingham grade 2 breast carcinomas. Their results demonstrated prognostic significance for stratifying patients with Nottingham grade 2 breast cancers.

AI-BASED DETECTION OF LYMPH NODE METASTASIS

Accurate assessment of axillary lymph node metastasis in breast cancer patients is crucial for their clinical management because lymph node metastasis status is significantly correlated with prognosis. Lymph node metastasis evaluation is time-consuming and labor-intensive. Although identifying macrometastasis is straightforward, it may be challenging to manually detect micrometastasis or isolated tumor cells. Recent studies have demonstrated AI algorithms can improve the accuracy and efficiency of lymph node assessment.³²⁻³⁴

During the period from 2015 to 2016, a researcher challenge competition (CAMELYON16) was launched to develop automated solutions for detecting lymph node metastases using a training data set of WSIs with (n = 110) and without (n = 160) nodal metastases verified by immunohistochemistry (IHC). Algorithm performance was eval-

uated in an independent test set of 129 WSIs (49 with and 80 without metastases). The AUC for the submitted algorithms ranged from 0.556 to 0.994 and the best algorithm (AUC = 0.994) performed significantly better than the pathologists.³⁵ Recently, Huang et al³⁶ developed an AI-assisted lymph node assessment workflow using CNNs trained on 5907 lymph node images. The results demonstrated the algorithm identified metastatic lymph nodes in gastric cancer with an AUC of 0.9936, and the algorithm was also highly robust (AUC = 0.9829) on cross-site evaluation. Steiner et al³⁴ reported the results of 6 pathologists who interpreted 70 lymph node WSIs according to 2 assessment methods (with and without AI assistance). The AI-assisted assessment demonstrated higher accuracy than AI alone or visual assessment by pathologists alone. The AI-assisted assessment significantly improved the sensitivity of micrometastasis detection (91% versus 83%, *P* = .02). Furthermore, the mean interpretation time per image was significantly shorter for AI-assisted assessment than for visual assessment alone. Liu et al³² developed and validated an AI algorithm using 270 lymph node images from 2 centers as a training set and 129 lymph node images as a validation set. The algorithm's optimal AUC reached 0.99, whereas the pathologist's optimal performance AUC was 0.88.

Several lymph node metastasis AI algorithms are commercially available. One example is the Visiopharm lymph node metastasis detection app. The app includes 4 steps: tissue detection, metastasis detection, measurement, and calculation. The process can be automated using directly streamed WSIs from an image management system and batch analysis. The results include annotated WSIs, calculated maximum length, and area of metastasis (Figure 2). Our unpublished data have demonstrated that the AI algorithm detected all 44 metastases (19 macrometastases, 25 micrometastases, 1 with isolated tumor cells) out of 233 lymph nodes with a sensitivity of 100%, specificity of 41.3%, positive predictive value of 28.4%, and negative predictive value of 100%, indicating its utility as a screening modality in routine clinical practice.

AI-BASED BREAST CANCER BIOMARKER QUANTIFICATION

Tissue biomarkers are surrogates for diagnosing disease, predicting prognosis, and selecting patients for targeted therapy. Protein biomarkers are commonly tested using IHC, and nucleic acid biomarkers are tested using *in situ* hybridization (ISH) or polymerase chain reaction. Clinically, tissue biomarker stains are manually examined and interpreted. Since the wide implementation of digital pathology, it is emerging that tissue biomarkers are assessed using AI tools, which can provide more objective and reproducible results.³⁷⁻⁴⁰ In breast pathology, these tissue biomarkers include estrogen receptor (ER), progesterone receptor (PR), HER2/neu, Ki-67, and programmed death ligand-1 (PD-L1) IHCs, as well as ISH analysis of HER2/neu.³⁹⁻⁴⁹

Estrogen Receptor and Progesterone Receptor Immunohistochemistry

American Society of Clinical Oncology (ASCO)/College of American Pathologists (CAP) guidelines recommend that all primary and recurrent/metastatic breast carcinomas should be tested for ER and PR expression, which has both prognostic and predictive values.⁵⁰⁻⁵³ ER/PR expression is tested by IHC and usually assessed by pathologists' manual

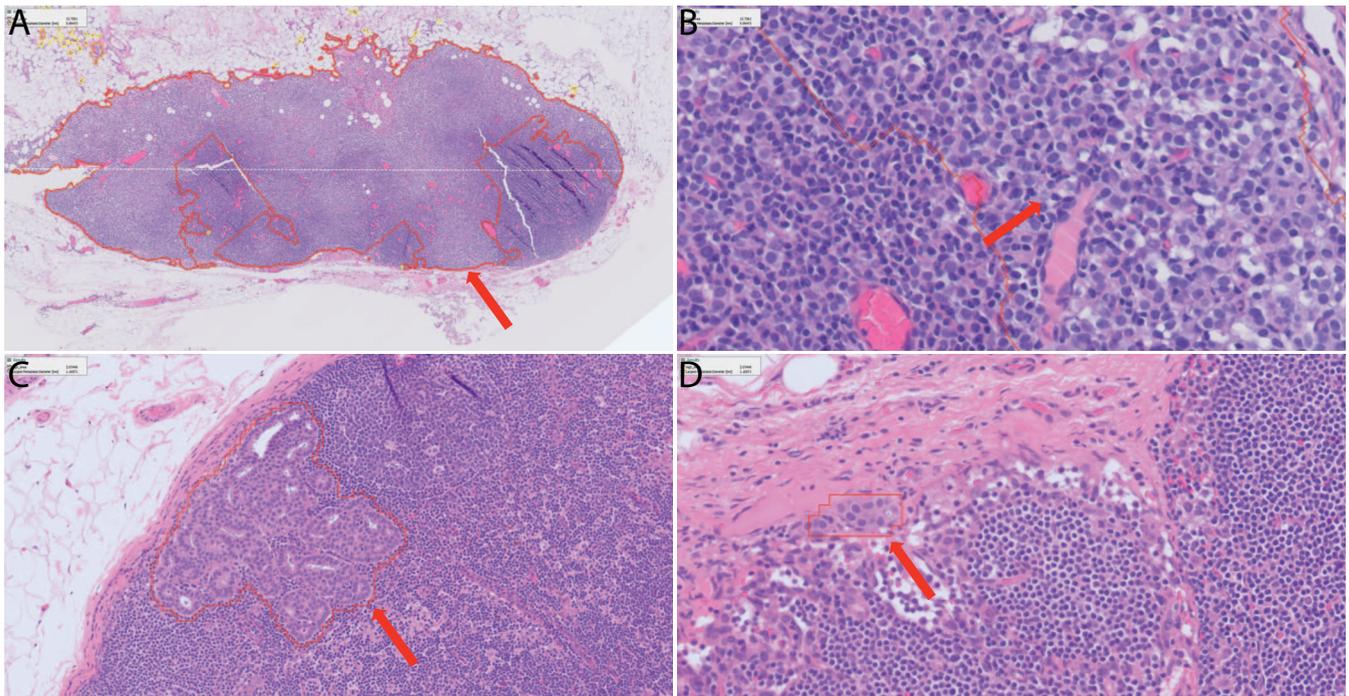


Figure 2. Representative images of lymph node metastasis detected by artificial intelligence (AI) algorithm. A, One annotated lymph node with macrometastasis from an invasive lobular carcinoma case. B, High magnification of the lymph node with macrometastasis from invasive lobular carcinoma. C, One annotated lymph node with micrometastasis. D, One annotated lymph node with isolated tumor cells (hematoxylin-eosin, original magnifications $\times 2$ [A], $\times 40$ [B], $\times 10$ [C], and $\times 20$ [D]). Arrows indicate metastatic carcinomas.

scoring, which has interobserver/intraobserver variability.^{54–62} Computational assessment of ER/PR IHC using AI algorithms provides opportunities to improve precision performance (Figure 3). AI algorithms not only demonstrate excellent correlation with pathologists' manual scoring but also yield higher reproducibility than pathologists' scoring.^{39,41–43,63–66} Furthermore, AI algorithms can be coupled with a digital pathology laboratory information system to provide an automated workflow.^{67,68} Like manual scoring, AI algorithms provide the ratio of positively staining tumor cells. Additionally, AI algorithms can further divide tumor cells into different staining intensity and then calculate an H-score by multiplying the percentages of nuclei with their corresponding staining intensity. Challenges still exist in evaluating ER/PR IHC using automated AI algorithms. False-positive results can be caused by intermixed benign glands in tumor area and DCIS components in invasive carcinoma. False-negative results can be caused by faint IHC staining that is not detected by AI algorithms.⁶⁹ Therefore, a pathologist's final review to confirm AI algorithms' analysis is necessary.

HER2 Immunohistochemistry

HER2 is another important prognostic and therapeutic biomarker in breast pathology. Up to 20% of breast carcinomas harbor *HER2* protein overexpression/gene amplification.^{70–73} ASCO/CAP guideline recommends that HER2 status should be determined for all invasive breast carcinomas by HER2 IHC and/or by ISH.⁷⁴ Like ER/PR IHCs, HER2 IHC is usually evaluated by pathologists' manual scoring, which often shows interobserver variability.^{74–76} According to the guidelines, HER2 IHC is classified into negative (0 and 1+), equivocal (2+), and positive (3+) based on HER2 membranous staining intensity together

with the percentage of tumor cells with membranous staining. Studies have demonstrated AI algorithms provide equally accurate, but more objective and reproducible assessments than manual scoring.^{44–48,77} Different approaches have been applied in HER2 IHC AI algorithms. Some AI algorithms segment tumor cells, classify each cell into a different staining category, and then calculate to obtain a final score. Other AI algorithms evaluate the connectivity of HER2-stained membrane to determine HER2 IHC score.⁴⁴ (Figure 4) Studies have used such AI algorithms to accurately determine HER2 IHC score, and to discriminate HER2⁺ and HER2⁻ cases.^{44,78–80} Most AI algorithms couple with WSIs or ROI images to evaluate HER2 IHC. An AI-assisted microscope has been developed by equipping a conventional microscope with an HER2 scoring AI algorithm and an augmented reality module. This AI-assisted microscope enables pathologists to obtain real-time HER2 IHC results for each view field and improves the consistency and accuracy of pathologist scoring.⁸¹ As the scientific understanding of HER2 in breast cancer is evolving, it may become necessary to further stratify HER2 negative breast cancers into HER2 IHC 0 and HER2-low (HER2 IHC 1+ and HER2 IHC 2+/ISH negative) categories.^{82–84} It is suggested that AI may play an important role in classifying these categories.

Finally, the ASCO/CAP guidelines have acknowledged AI algorithms as diagnostic modalities to evaluate HER2 IHC, and the CAP has published guidelines to facilitate the adoption of AI algorithms into clinical practice.^{74,85}

HER2 ISH

HER2 gene amplification is examined by HER2 ISH, including fluorescence ISH (FISH), silver ISH, chromogenic ISH, and dual ISH.^{86,87} Studies have demonstrated excellent

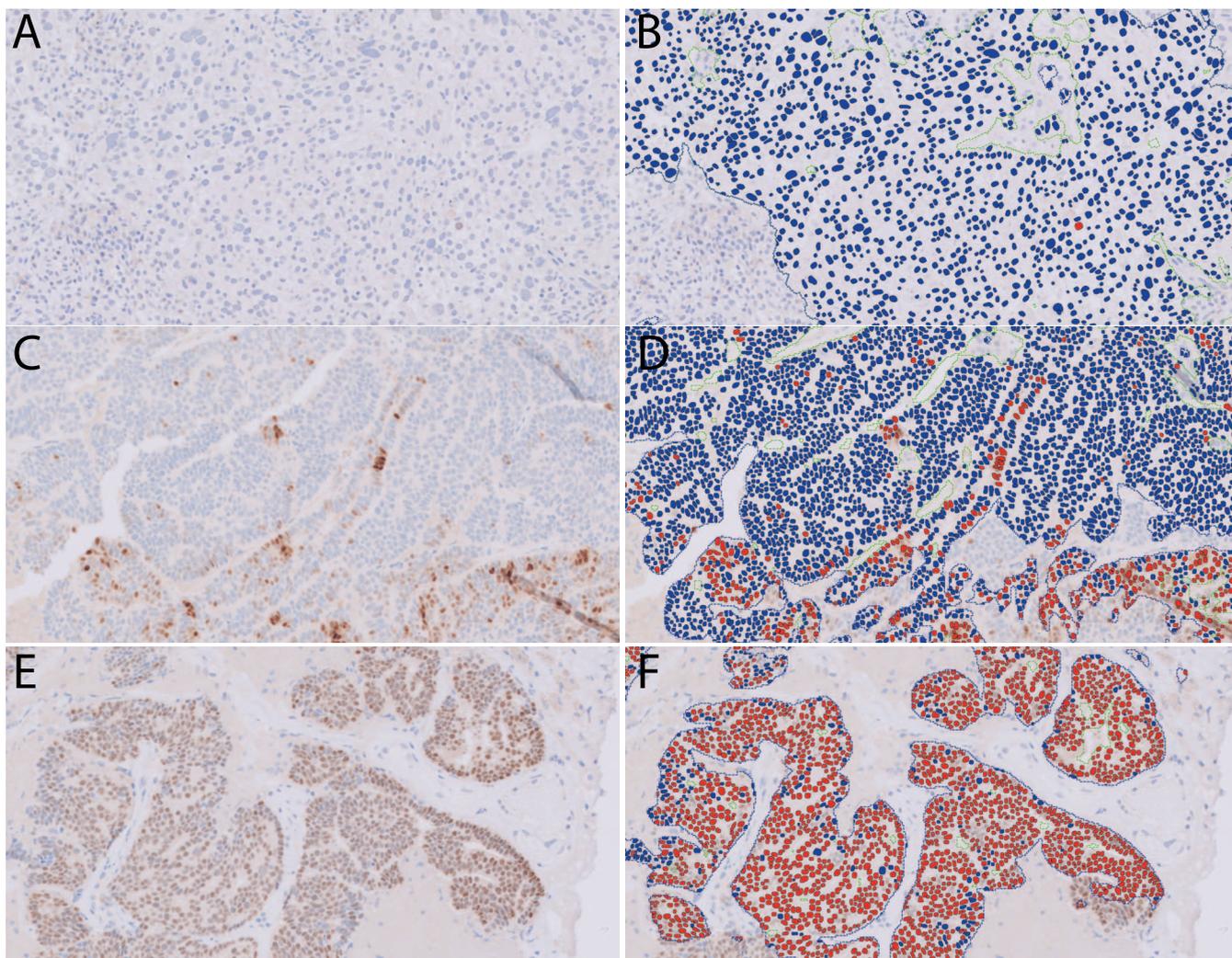


Figure 3. Example of estrogen receptor (ER) quantification by artificial intelligence algorithm. The left panels (A, C, and E) show ER immunohistochemistry staining, and the right panels (B, D, and F) show cell segmentation and ER quantification with pseudocolors (blue, negative staining; red, positive staining). Invasive carcinoma is automatically detected and outlined by the algorithm. A and B, ER with $<1\%$ positive staining. C and D, ER with 25% positive staining. E and F, ER with 90% positive staining (immunostains of ER, original magnification $\times 10$).

concordance between HER2 ISH and IHC in determining HER2 status.⁸⁸ AI algorithms have been developed to assist pathologists in detecting, classifying, and counting HER2 signals in cells of interest, and excellent concordance between manual scoring and AI scoring of HER2 FISH has been detected.⁸⁹ The benefits of using AI to analyze HER2 FISH include improved efficiency and productivity because manual counting of HER2 FISH signals is time consuming.^{89–93}

Ki-67 Immunohistochemistry

Ki-67 is a surrogate marker for cell proliferation and is expressed in all cell cycle phases except in G₀.^{94–97} Ki-67 expression is related to breast tumor grade and biologic behavior, with high expression associated with worse prognosis.^{97–100} Cyclin-dependent kinases 4 and 6 (CDK4/6) inhibitor was recently approved for early-stage ER⁺, HER2⁻ breast cancer with high risk of recurrence and a Ki67 score of 20% or higher.^{101,102} The Ki-67 IHC MIB-1 pharmDx assay was also approved as a companion diagnostic test for this indication.

Ki-67 expression is assessed by IHC and scored as a percentage of tumor cells stained by Ki-67 antibody. Several methodologies are used to score Ki-67 IHC, including visual estimation, manual counting of ROIs, and manual comprehensive counting of the whole slide. Visual estimation has notable interobserver variability.^{96,97,103} Similarly, interobserver variability also exists in selecting ROIs for counting.^{104,105} Manual comprehensive counting of the whole slide improves reproducibility, but it is very time-consuming and impractical during clinical routine practice.¹⁰⁶

Automated scoring methods using AI algorithms have shown promise for assessing Ki-67 IHC.^{37,107–113} AI algorithms automatically detect tumor nuclei based on their shape and size, identify positively stained tumor nuclei based on their color, and then calculate the positivity rate. However, automated AI algorithms have pitfalls. First, AI algorithms can confuse tumor cells with surrounding stromal cells, inflammatory cells, or artefacts. Second, AI algorithm can misclassify lesions, such as by including in situ components as invasive carcinoma for analysis. Therefore, pathologist intervention or extensive supervised learning for sophisticated cell segmentation and classifica-

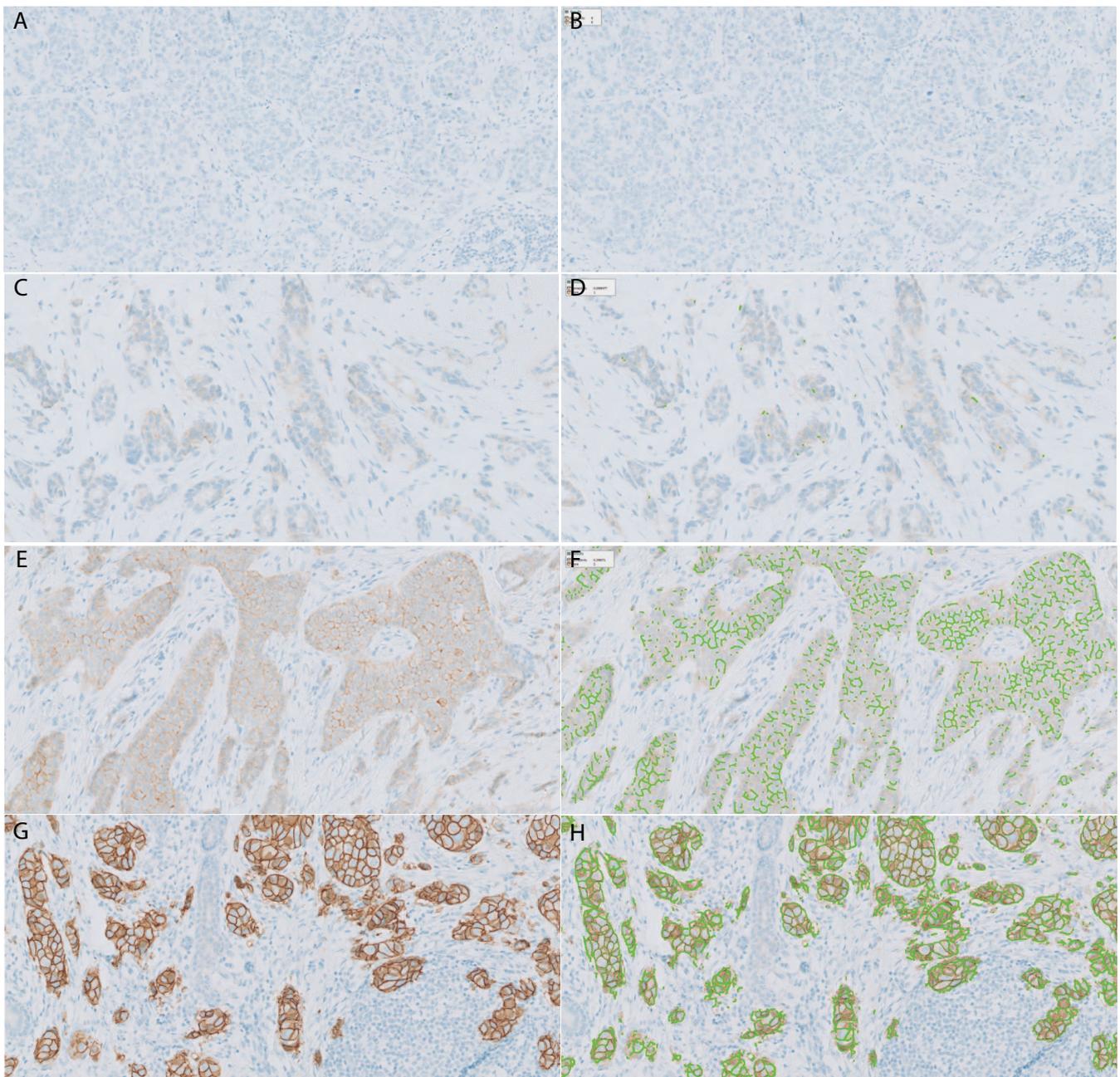


Figure 4. Human epidermal growth factor receptor 2 (HER2) immunohistochemistry (IHC) and the connectivity analyzed by artificial intelligence (AI) algorithm. A and B, One case with HER2 IHC 0. C and D, One case with HER2 IHC 1+. E and F, One case with HER2 IHC 2+. G and H, One case with HER2 IHC 3+. HER2 IHC (A, C, E, and G); HER2 connectivity (green color line) detected by AI algorithm (B, D, F, and H) (immunostain of HER2, original magnifications $\times 5$ [A and B] and $\times 10$ [C through G]).

tion may be necessary. Recently, new AI algorithms have been developed to more accurately score Ki-67 using dual IHC (Ki-67 plus cytokeratin stain for labeling tumor cells) or sequential IHC stains followed by virtual image reconstruction.¹¹⁴ Nevertheless, AI algorithms are capable of analyzing the whole slide efficiently, which may result in higher accuracy than counting on ROIs.¹¹⁵

PD-L1 IHC

PD-L1 IHC is an emerging test to select triple-negative breast cancer (TNBC) patients for immunotherapy. Two PD-L1 assays (SP142 and 22C3) had been approved as

companion diagnostic tests in breast cancer; however, SP142 was recently withdrawn because of failed clinical trial.^{116–121} Currently, only the 22C3 test is used for breast cancer. PD-L1 22C3 (Agilent Technologies, Santa Clara, California) uses a combined positive score (CPS) to assess PD-L1 expression. The CPS score is calculated as the number of PD-L1–stained cells (tumor cells, lymphocytes, and macrophages) divided by the total number of tumor cells multiplied by 100.¹²¹

Although studies have demonstrated AI algorithms' assessment of PD-L1 IHC showed excellent correlation with pathologists' manual scoring, only a few PD-L1 AI

algorithms have been validated for clinical practice.^{122,123} This may be due to the relatively short time of PD-L1 IHC in clinical practice and the complexity of PD-L1 testing (multiple antibodies and different scoring systems).

AI-BASED EVALUATION OF TUMOR-INFILTRATING LYMPHOCYTES IN BREAST CANCER

In breast cancer, especially triple-negative breast cancer and HER2⁺ breast carcinoma, the presence of tumor-infiltrating lymphocytes (TILs) in the tumor microenvironment is associated with better response to therapy and overall survival.^{124–128} TILs have been emerging as a biomarker in breast carcinoma.^{128,129} However, manual assessment of TILs is subjective, with significant interobserver variability and poor reproducibility.^{124,128}

With the widespread implementation of digital pathology, AI algorithms have been developed to evaluate TILs aiming for more accurate and reproducible results.^{130–135} In 1 study, a computer-aided diagnosis scheme was developed to automatically detect and grade the extent of lymphocytic infiltrate in HER2⁺ breast carcinoma using WSIs, and the results showed the architectural feature set successfully distinguished samples of high and low lymphocytic infiltrate levels with a classification accuracy greater than 90%.¹³¹ In another study, an automated method based on grid subsampling of microscopy image analysis data was developed to extract the tumor-stroma interface and compute immunogradient indicators for TIL density profiles, which provided a strong and independent prognostic stratification in ER⁺ breast cancer patients.¹³²

AI-BASED PREDICTION OF PROGNOSIS, SURVIVAL, AND THERAPY RESPONSE IN BREAST CANCER

The combined antiestrogen therapy and chemotherapy have significantly reduced recurrence frequency and improved survival rate in certain populations of ER⁺ breast cancer patients.¹³⁶ However, these therapies, especially chemotherapy, have toxic side effects; therefore, identifying patients who are more likely to benefit from chemotherapy is important. ASCO Clinical Practice Guideline recommends the clinician may use Oncotype DX RS to guide decisions on adjuvant systemic chemotherapy if a patient has ER⁺, HER2⁻ breast cancer and the Oncotype DX assay has been widely accepted in clinical practice across the United States and Canada.³¹ However, Oncotype DX is a prohibitively expensive test. Multiple studies have suggested that standard histopathologic variables (including tumor grade, tubule formation, nuclear pleomorphism, and mitotic activity) together with breast cancer biomarkers (ER, PR, HER2), can provide information similar to that provided by the Oncotype DX RS.^{137–140} An AI-based algorithm using histopathologic WSIs, Image-based Risk Score (IbRiS), was developed to serve as a proxy for Oncotype DX, and demonstrated a mean accuracy of 84% in distinguishing low-risk from high-risk specimens.¹⁴¹

Breast cancer patients' prognosis and survival are closely related to histopathologic grade (Nottingham grade), as mentioned above. Studies have shown a considerable amount of interobserver variability in breast cancer histopathologic grade determined by pathologists.¹⁴² AI-based approaches have been developed to measure the morphologic components in the Nottingham grading system and correlated to the patient's prognosis and recurrent risk. A DL algorithm was developed to automatically measure

tubule formation in breast carcinoma WSIs to generate a tubule formation indicator, which correlated with histologic grade and recurrence risk determined by molecular test (Oncotype DX).²⁸ One study evaluated the ability of computer-extracted nuclear morphology features from histologic images to train 4 ML classifiers (Random Forest, Neural Network, Support Vector Machine, and Linear Discriminant Analysis) for breast cancer risk categorization and demonstrated per-patient accuracies ranging from 75% to 86% by correlating with Oncotype DX risk categories.¹⁴³ Additional study quantitated computer-extracted image features of nuclear shape and orientation on digitized breast carcinoma images and demonstrated that quantitative histomorphometric features of nuclear shape and orientation were strongly and independently predictive of patient survival.¹⁴⁴ In another study, a deep neural network classifier was also developed to quantify mitosis in breast cancer WSIs and demonstrated high accuracy (83%) and good correlation with Oncotype DX risk categories.¹⁴⁵

ML algorithms were also developed to predict neoadjuvant chemotherapy response in breast cancer using clinical and pathologic features. In 1 study, a standard multivariable logistic regression (MLR) was developed using patient demographics, histologic characteristics (ER status, HER2 status, Nottingham grade, tumor size, and nodal status), molecular status, and staging information. MLR was compared with 5 ML models (k-nearest neighbor classifier, random forest classifier, naive Bayes algorithm, support vector machine, and multilayer perceptron model) for their performance in predicting neoadjuvant chemotherapy response in breast cancer. The AUC for the MLR was 0.64. Among the 5 ML models mentioned above, the random forest classifier performed best, with an AUC of 0.88.¹⁴⁶

AI-BASED PREDICTION OF GENETIC ABNORMALITY IN BREAST CANCER

Multiple studies have sought to predict HER2 status (amplified or nonamplified) using histologic images in breast cancer. Farahmand et al¹³ developed a novel CNN classifier trained on 188 WSIs manually annotated for tumor ROIs to be able to predict HER2 status with an AUC of 0.90 in cross-validation of slide-level HER2 status and 0.81 on an independent TCGA test set. Additionally, the classifier was tested on pretreatment samples from 187 HER2⁺ patients that subsequently received trastuzumab therapy to predict trastuzumab response with an AUC of 0.80. Another study trained a classification pipeline to determine HER2 overexpression status of H&E-stained WSIs which achieved an AUC of 0.82 (CI, 0.65–0.98) on held-out cases and an AUC of 0.76 (CI, 0.61–0.89) on the independent data set from TCGA.¹⁴⁷ Recently, a group of researchers presented the HER2 on hematoxylin and eosin (HEROHE) challenge, which aimed to predict the HER2 status in breast cancer using a large, publicly available, annotated H&E whole slide imaging data set (n = 509).¹⁴⁸ This was a parallel event of the 16th European Congress on Digital Pathology, and 21 teams worldwide submitted their models, and the best-performing models were presented with detailing their models' architectures and key parameters.

Besides predicting HER2 genetic status, AI algorithm was also developed based on H&E images to predict BRCA gene mutation in breast cancer. Wang et al¹⁴⁹ trained a deep CNN of ResNet WSIs and validated their model through an external data set that contained 17 BRCA mutated and 47

wild-type cases with AUCs (95% CI) as 0.766 (0.763–0.769), 0.763 (0.758–0.769), 0.750 (0.738–0.761), and 0.551 (0.526–0.575), using $\times 40$, $\times 20$, $\times 10$, and $\times 5$ magnification tiles, respectively.

Recently, an explainable ML approach was developed to investigate integrated profiling of morphologic, molecular, and clinical features from breast cancer histology. This ML approach was able to predict molecular features in breast cancer, including DNA methylation, gene expression, copy number variations, and somatic mutations with balanced accuracy up to 78% and very high accuracy ($>95\%$) in subgroup of patients for specific genes, such as p53, PTEN, etc.¹⁵⁰ Another study used histologic images through a DL-based model to predict chromosomal instability status in breast cancer and achieved an AOC of 0.822 in correctly classifying chromosomal instability status.¹⁵¹

LIMITATIONS TO AI APPLICATIONS IN BREAST PATHOLOGY

Several key factors must be considered when adopting AI algorithms in breast pathology. The first one is the quantity and quality of training data from which the AI algorithm was developed. Significant variations exist in file formats of WSIs, scanner quality, and glass slide quality, including staining intensity, coverslip, tissue size, folded tissue, air bubble etc. Manual selection of WSIs with artifact-free and proper quality must be used in the training of an algorithm to develop comprehensive ML/DL models. WSIs with low resolution and low quality have indistinguishable features, making it difficult for an ML/DL model to accurately assess detailed features. Therefore, standardization and normalization of the training data set are necessary for successful AI algorithm development and the application of AI algorithms across different laboratories. Recently, an open-source quality control tool for digital pathology slides, HistoQC, was described to not only identify and delineate artefacts but also discover cohort-level outliers.¹⁵² Subsequent study demonstrated HistoQC substantially improved overall concordance in identifying unsuitable WSIs for computational analysis among pathologists (moderate Gwet AC1 range 0.43–0.59 to excellent range 0.79–0.93).¹⁵³

Second, validating AI algorithms is required before adoption for clinical use. Many above-mentioned AI algorithms are experimental, except for some breast biomarker quantification AI algorithms. Recent studies have demonstrated the successful validation of AI algorithms, and their institutions have started to implement such AI algorithms for clinical practice.^{69,80}

Thirdly, a digital pathology workflow (digital sign-out) is always preferred or necessary to implement AI algorithms for routine pathology practice to achieve an automated workflow. Whole slide imaging technology has been implemented in many institutions worldwide; however, a fully digital pathology workflow has been limited to very few pathology laboratories.^{67,154}

Lastly, pathologists' trust is also a key limitation in adopting AI algorithms in routine pathology practice. Many AI algorithms are frequently referred to as "black boxes" because they cannot explain clearly what features trigger ML/DL and how neural networks piece together all information. If pathologists do not understand what features AI algorithms use and how the algorithms make decisions,

they will be reluctant to believe AI results and adopt AI algorithms.

CONCLUSIONS

With the widespread implementation of WSIs in breast pathology practice, the application of AI algorithms becomes increasingly popular. Integration of AI algorithms into digital workflow is becoming feasible to significantly improve pathologists' efficiency and diagnostic accuracy by reducing workload (repetitive tasks, such as lymph node metastasis detection etc) and diagnostic errors. Some AI algorithms may provide pathologists with new tools to tackle emerging pathologic assessment, such as TIL quantification, prediction of therapy response, and recurrence risk stratification. Furthermore, AI algorithms may complement or replace some expensive molecular tests in breast pathology, such as Oncotype DX and genetic testing.

References

1. Fernandez AI, Liu M, Bellizzi A, et al. Examination of low ERBB2 protein expression in breast cancer tissue. *JAMA Oncol.* 2022;8(4):1–4.
2. Sirinukunwattana K, Pluim JPW, Chen H, et al. Gland segmentation in colon histology images: the glas challenge contest. *Med Image Anal.* 2017;35:489–502.
3. Denkert C, Wienert S, Poterie A, et al. Standardized evaluation of tumor-infiltrating lymphocytes in breast cancer: results of the ring studies of the international immuno-oncology biomarker working group. *Mod Pathol.* 2016;29(1):1155–1164.
4. Yamamoto S, Chishima T, Mastubara Y, et al. Variability in measuring the Ki-67 labeling index in patients with breast cancer. *Clin Breast Cancer.* 2015;15(1):e35–e39.
5. Bejnordi BE, Zuidhof G, Balkenhol M, et al. Context-aware stacked convolutional neural networks for classification of breast carcinomas in whole-slide histopathology images. *J Med Imaging.* 2017;4(4):044504.
6. Cireşan DC, Giusti A, Gambardella LM, Schmidhuber J. Mitosis detection in breast cancer histology images with deep neural networks. *Med Image Comput Assist Interv.* 2013;16(pt 2):411–418.
7. Vandenbergh ME, Scott ML, Scorer PW, Söderberg M, Balcerzak D, Barker C. Relevance of deep learning to facilitate the diagnosis of HER2 status in breast cancer. *Sci Rep.* 2017;7:45938.
8. Tellez D, Balkenhol M, Otte-Holler I, et al. Whole-slide mitosis detection in H&E breast histology using PHH3 as a reference to train distilled stain-invariant convolutional networks [published online ahead of print April 28, 2018]. *IEEE Trans Med Imaging.* 2018. doi:10.1109/TMI.2018.2820199
9. Houssami N, Lee CI, Buist DSM, Tao D. Artificial intelligence for breast cancer screening: opportunity or hype? *Breast.* 2017;36:31–33.
10. Ferroni P, Zanzotto FM, Riondino S, Scarpato N, Guadagni F, Roselli M. Breast cancer prognosis using a machine learning approach. *Cancers (Basel).* 2019;11(3):328.
11. Xu Y, Hosny A, Zeleznik R, et al. Deep learning predicts lung cancer treatment response from serial medical imaging. *Clin Cancer Res.* 2019;25(11):3266–3275.
12. Yang J, Ju J, Guo L, et al. Prediction of HER2-positive breast cancer recurrence and metastasis risk from histopathological images and clinical information via multimodal deep learning. *Comput Struct Biotechnol J.* 2022;20:333–342.
13. Farahmand S, Fernandez AI, Ahmed FS, et al. Deep learning trained on hematoxylin and eosin tumor region of interest predicts HER2 status and trastuzumab treatment response in HER2+ breast cancer. *Mod Pathol.* 2022;35(1):44–51.
14. Han Z, Wei B, Zheng Y, Yin Y, Li K, Li S. Breast cancer multi-classification from histopathological images with structured deep learning model. *Sci Rep.* 2017;7(1):4172.
15. Cruz-Roa A, Gilmore H, Basavanthally A, et al. Accurate and reproducible invasive breast cancer detection in whole-slide images: a deep learning approach for quantifying tumor extent. *Sci Rep.* 2017;7:46450.
16. USCAP 2022 abstracts: breast pathology (74-204). *Mod Pathol.* 2022;35(2):153–305.
17. Wang Y, Acs B, Robertson S, et al. Improved breast cancer histological grading using deep learning. *Ann Oncol.* 2022;33(1):89–98.
18. Veta M, van Diest PJ, Willems SM, et al. Assessment of algorithms for mitosis detection in breast cancer histopathology images. *Med Image Anal.* 2015;20(1):237–248.
19. Das A, Nair MS, Peter DS. Batch mode active learning on the Riemannian manifold for automated scoring of nuclear pleomorphism in breast cancer. *Artif Intell Med.* 2020;103:101805.
20. Mantrala S, Ginter PS, Mitkar A, et al. Concordance in breast cancer grading by artificial intelligence on whole slide images compares with a multi-

institutional cohort of breast pathologists. *Arch Pathol Lab Med.* 2022;146(11):1369–1377.

21. Balkenhol MCA, Tellez D, Vreuls W, et al. Deep learning assisted mitotic counting for breast cancer. *Lab Invest.* 2019;99(11):1596–1606.

22. Elsharawy KA, Gerds TA, Rakha EA, Dalton LW. Artificial intelligence grading of breast cancer: a promising method to refine prognostic classification for management precision. *Histopathology.* 2021;79(2):187–199.

23. Nateghi R, Danyali H, Helfroush MS. A deep learning approach for mitosis detection: application in tumor proliferation prediction from whole slide images. *Artif Intell Med.* 2021;114:102048.

24. Li C, Wang X, Liu W, Latecki LJ. DeepMitosis: mitosis detection via deep detection, verification and segmentation networks. *Med Image Anal.* 2018;45:121–133.

25. Sebai M, Wang X, Wang T. MaskMitosis: a deep learning framework for fully supervised, weakly supervised, and unsupervised mitosis detection in histopathology images. *Med Biol Eng Comput.* 2020;58(7):1603–1623.

26. Mahmood T, Arsalan M, Owais M, Lee MB, Park KR. Artificial intelligence-based mitosis detection in breast cancer histopathology images using faster R-CNN and deep CNNs. *J Clin Med.* 2020;9(3):749.

27. Katayama A, Toss MS, Parkin M, Ellis IO, Quinn C, Rakha EA. Atypia in breast pathology: what pathologists need to know. *Pathology.* 2022;54(1):20–31.

28. Romo-Bucheli D, Janowczyk A, Gilmore H, Romero E, Madabhushi A. Automated tubule nuclei quantification and correlation with Oncotype DX risk categories in ER+ breast cancer whole slide images. *Sci Rep.* 2016;6:32706.

29. Paik S, Tang G, Shak S, et al. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol.* 2006;24(23):3726–3734.

30. Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med.* 2004;351(27):2817–2826.

31. Andre F, Ismaila N, Allison KH, et al. Biomarkers for adjuvant endocrine and chemotherapy in early-stage breast cancer: ASCO guideline update. *J Clin Oncol.* 2022;40(16):1816–1837.

32. Liu Y, Kohlberger T, Norouzi M, et al. Artificial intelligence-based breast cancer nodal metastasis detection: insights into the black box for pathologists. *Arch Pathol Lab Med.* 2019;143(7):859–868.

33. Zheng X, Yao Z, Huang Y, et al. Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. *Nat Commun.* 2020;11(1):1236.

34. Steiner DF, MacDonald R, Liu Y, et al. Impact of deep learning assistance on the histopathologic review of lymph nodes for metastatic breast cancer. *Am J Surg Pathol.* 2018;42(12):1636–1646.

35. Ehteshami Bejnordi B, Johannes Veta M, van Diest P, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *JAMA.* 2017;318(22):2199–2210.

36. Huang SC, Chen CC, Lan J, et al. Deep neural network trained on gigapixel images improves lymph node metastasis detection in clinical settings. *Nat Commun.* 2022;13(1):3 347.

37. Stalhammar G, Fuentes Martinez N, Lippert M, et al. Digital image analysis outperforms manual biomarker assessment in breast cancer. *Mod Pathol.* 2016;29(4):318–329.

38. Aeffner F, Zarella M, Buchbinder N, et al. Introduction to digital image analysis in whole-slide imaging: a white paper from the digital pathology association. *J Pathol Inform.* 2019;10(1):9.

39. Bolton KL, Garcia-Closas M, Pfeiffer RM, et al. Assessment of automated image analysis of breast cancer tissue microarrays for epidemiologic studies. *Cancer Epidemiol Biomarkers Prev.* 2010;19(4):992–999.

40. Webster JD, Dunstan RW. Whole-slide imaging and automated image analysis: considerations and opportunities in the practice of pathology. *Vet Pathol.* 2014;51(1):211–223.

41. Diaz LK, Sahin A, Sneige N. Interobserver agreement for estrogen receptor immunohistochemical analysis in breast cancer: a comparison of manual and computer-assisted scoring methods. *Ann Diagn Pathol.* 2004;8(1):23–27.

42. Faratian D, Kay C, Robson T, et al. Automated image analysis for high-throughput quantitative detection of ER and PR expression levels in large-scale clinical studies: the TEAM Trial Experience. *Histopathology.* 2009;55(5):587–593.

43. Rizzardi AE, Johnson AT, Vogel RI, et al. Quantitative comparison of immunohistochemical staining measured by digital image analysis versus pathologist visual scoring. *Diagn Pathol.* 2012;7:42.

44. Brugmann A, Eld M, Lelkaitis G, et al. Digital image analysis of membrane connectivity is a robust measure of HER2 immunostains. *Breast Cancer Res Treat.* 2012;132(1):41–49.

45. Helin HO, Tuominen VJ, Ylino O, Helin HJ, Isola J. Free digital image analysis software helps to resolve equivocal scores in HER2 immunohistochemistry. *Virchows Archiv.* 2016;468(2):191–198.

46. Laurinaviciene A, Dasevicus D, Ostapenko V, Jarmalaite S, Lazutka J, Laurinavicius A. Membrane connectivity estimated by digital image analysis of HER2 immunohistochemistry is concordant with visual scoring and fluorescence in situ hybridization results: algorithm evaluation on breast cancer tissue microarrays. *Diagn Pathol.* 2011;6:87.

47. Skaland I, Ovestad I, Janssen EA, et al. Comparing subjective and digital image analysis HER2/neu expression scores with conventional and modified FISH scores in breast cancer. *J Clin Pathol.* 2008;61(1):68–71.

48. Dobson L, Conway C, Hanley A, et al. Image analysis as an adjunct to manual HER-2 immunohistochemical review: a diagnostic tool to standardize interpretation. *Histopathology.* 2010;57(1):27–38.

49. Hou Y, Peng Y, Li Z. Update on prognostic and predictive biomarkers of breast cancer. *Semin Diagn Pathol.* 2022;39(5):322–332.

50. Allred DC, Harvey JM, Berardo M, Clark GM. Prognostic and predictive factors in breast cancer by immunohistochemical analysis. *Mod Pathol.* 1998;11(2):155–168.

51. Hahnel R, Woodings T, Vivian AB. Prognostic value of estrogen receptors in primary breast cancer. *Cancer.* 1979;44(2):671–675.

52. Fitzgibbons PL, Page DL, Weaver D, et al. Prognostic factors in breast cancer: College of American Pathologists Consensus Statement 1999. *Arch Pathol Lab Med.* 2000;124(7):966–978.

53. Hammond ME, Hayes DF, Dowsett M, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. *J Clin Oncol.* 2010;28(16):2784–2795.

54. Nadjji M, Gomez-Fernandez C, Ganjei-Azar P, Morales AR. Immunohistochemistry of estrogen and progesterone receptors reconsidered: experience with 5,993 breast cancers. *Am J Clin Pathol.* 2005;123(1):21–27.

55. Mann GB, Fahey VD, Feleppa F, Buchanan MR. Reliance on hormone receptor assays of surgical specimens may compromise outcome in patients with breast cancer. *J Clin Oncol.* 2005;23(22):5148–5154.

56. Hede K. Breast cancer testing scandal shines spotlight on black box of clinical laboratory testing. *J Natl Cancer Inst.* 2008;100(12):836–837, 844.

57. Collins LC, Botero ML, Schnitt SJ. Bimodal frequency distribution of estrogen receptor immunohistochemical staining results in breast cancer: an analysis of 825 cases. *Am J Clin Pathol.* 2005;123(1):16–20.

58. Badve SS, Baehner FL, Gray RP, et al. Estrogen- and progesterone-receptor status in ECOG 2197: comparison of immunohistochemistry by local and central laboratories and quantitative reverse transcription polymerase chain reaction by central laboratory. *J Clin Oncol.* 2008;26(15):2473–2481.

59. Ciocca DR, Elledge R. Molecular markers for predicting response to tamoxifen in breast cancer patients. *Endocrine.* 2000;13(1):1–10.

60. Gelber RD, Gelber S. Facilitating consensus by examining patterns of treatment effects. *Breast.* 2009;18:S2–S8.

61. Reisenbichler ES, Lester SC, Richardson AL, Dillon DA, Ly A, Brock JE. Interobserver concordance in implementing the 2010 ASCO/CAP recommendations for reporting ER in breast carcinomas: a demonstration of the difficulties of consistently reporting low levels of ER expression by manual quantification. *Am J Clin Pathol.* 2013;140(4):487–494.

62. Viale G, Regan MM, Maiorano E, et al. Prognostic and predictive value of centrally reviewed expression of estrogen and progesterone receptors in a randomized trial comparing letrozole and tamoxifen adjuvant therapy for postmenopausal early breast cancer: BIG 1-98. *J Clin Oncol.* 2007;25(25):3846–3852.

63. Thomsen C, Nielsen S, Nielsen BS, Pedersen SH, Vyberg M. Estrogen receptor-alpha quantification in breast cancer: concordance between immunohistochemical assays and mRNA in situ hybridization for ESR1 gene. *Appl Immunohistochem Mol Morphol.* 2020;28(5):347–353.

64. Turbin DA, Leung S, Cheang MC, et al. Automated quantitative analysis of estrogen receptor expression in breast carcinoma does not differ from expert pathologist scoring: a tissue microarray study of 3,484 cases. *Breast Cancer Res Treat.* 2008;110(3):417–426.

65. Gokhale S, Rosen D, Sneige N, et al. Assessment of two automated imaging systems in evaluating estrogen receptor status in breast carcinoma. *Appl Immunohistochem Mol Morphol.* 2007;15(4):451–455.

66. Rexhepaj E, Brennan DJ, Holloway P, et al. Novel image analysis approach for quantifying expression of nuclear proteins assessed by immunohistochemistry: application to measurement of oestrogen and progesterone receptor levels in breast cancer. *Breast Cancer Res.* 2008;10(5):R89.

67. Lujan GM, Savage J, Shana'ah A, et al. Digital pathology initiatives and experience of a large academic institution during the coronavirus disease 2019 (COVID-19) pandemic. *Arch Pathol Lab Med.* 2021;145(9):1051–1061.

68. Ahern TP, Beck AH, Rosner BA, et al. Continuous measurement of breast tumour hormone receptor expression: a comparison of two computational pathology platforms. *J Clin Pathol.* 2017;70(5):428–434.

69. Shafi S, Kellough DA, Lujan G, Satturwar S, Parwani AV, Li Z. Integrating and validating automated digital imaging analysis of estrogen receptor immunohistochemistry in a fully digital workflow for clinical use. *J Pathol Inform.* 2022;13:100122.

70. Slamon DJ, Clark GM, Wong WJ, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science.* 1987;235(4785):177–182.

71. Tandon AK, Clark GM, Chamness GC, Ullrich A, McGuire WL. HER-2/neu oncogene protein and prognosis in breast cancer. *J Clin Oncol.* 1989;7(8):1120–1128.

72. Press MF, Pike MC, Chazin VR, et al. Her-2/neu expression in node-negative breast cancer: direct tissue quantitation by computerized image analysis and association of overexpression with increased risk of recurrent disease. *Cancer Res.* 1993;53(20):4960–4970.

73. Park JW, Neve RM, Szollosi J, Benz CC. Unraveling the biologic and clinical complexities of HER2. *Clin Breast Cancer.* 2008;8(5):392–401.

74. Wolff AC, Hammond MEH, Allison KH, et al. Human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/

College of American Pathologists clinical practice guideline focused update. *J Clin Oncol*. 2018;36(20):2105–2122.

75. Gancberg D, Jarvinen T, di Leo A, et al. Evaluation of HER-2/NEU protein expression in breast cancer by immunohistochemistry: an interlaboratory study assessing the reproducibility of HER-2/NEU testing. *Breast Cancer Res Treat*. 2002;74(2):113–120.

76. Jacobs TW, Gown AM, Yaziji H, Barnes MJ, Schnitt SJ. Comparison of fluorescence in situ hybridization and immunohistochemistry for the evaluation of HER-2/neu in breast cancer. *J Clin Oncol*. 1999;17(7):1974–1982.

77. Holten-Rossing H, Moller Talman ML, Kristensson M, Vainer B. Optimizing HER2 assessment in breast cancer: application of automated image analysis. *Breast Cancer Res Treat*. 2015;152(2):367–375.

78. Koopman T, Buikema HJ, Hollema H, de Bock GH, van der Vegt B. What is the added value of digital image analysis of HER2 immunohistochemistry in breast cancer in clinical practice?: a study with multiple platforms. *Histopathology*. 2019;74(6):917–924.

79. Koopman T, de Bock GH, Buikema HJ, et al. Digital image analysis of HER2 immunohistochemistry in gastric- and oesophageal adenocarcinoma: a validation study on biopsies and surgical specimens. *Histopathology*. 2018;72(2):191–200.

80. Hartage R, Li AC, Hammond S, Parwani AV. A validation study of human epidermal growth factor receptor 2 immunohistochemistry digital imaging analysis and its correlation with human epidermal growth factor receptor 2 fluorescence in situ hybridization results in breast carcinoma. *J Pathol Inform*. 2020;11:2.

81. Yue M, Zhang J, Wang X, et al. Can AI-assisted microscope facilitate breast HER2 interpretation?: a multi-institutional ring study. *Virchows Archiv*. 2021;479(3):443–449.

82. Zhang H, Moisini I, Ajabnoor RM, Turner BM, Hicks DG. Applying the new guidelines of HER2 testing in breast cancer. *Curr Oncol Rep*. 2020;22(5):51.

83. Modi S, Park H, Murthy RK, et al. Antitumor activity and safety of trastuzumab deruxtecan in patients with HER2-low-expressing advanced breast cancer: results from a phase Ib study. *J Clin Oncol*. 2020;38(17):1887–1896.

84. Modi S, Jacot W, Yamashita T, et al. Trastuzumab deruxtecan in previously treated HER2-low advanced breast cancer. *New Engl J Med*. 2022;387(1):9–20.

85. Bui MM, Riben MW, Allison KH, et al. Quantitative image analysis of human epidermal growth factor receptor 2 immunohistochemistry for breast cancer: guideline from the College of American Pathologists. *Arch Pathol Lab Med*. 2019;143(10):1180–1195.

86. Kiyose S, Igarashi H, Nagura K, et al. Chromogenic in situ hybridization (CISH) to detect HER2 gene amplification in breast and gastric cancer: comparison with immunohistochemistry (IHC) and fluorescence in situ hybridization (FISH). *Pathol Int*. 2012;62(11):728–734.

87. Lichter P, Cremer T, Borden J, Manuelidis L, Ward DC. Delineation of individual human chromosomes in metaphase and interphase cells by in situ suppression hybridization using recombinant DNA libraries. *Hum Genet*. 1988;80(3):224–234.

88. Nitta H, Kelly BD, Allred C, et al. The assessment of HER2 status in breast cancer: the past, the present, and the future. *Pathol Int*. 2016;66(6):313–324.

89. Furrer D, Jacob S, Caron C, Sanschagrin F, Provencher L, Diorio C. Validation of a new classifier for the automated analysis of the human epidermal growth factor receptor 2 (HER2) gene amplification in breast cancer specimens. *Diagn Pathol*. 2013;8:17.

90. Theodosiou Z, Kasampalidis IN, Karayannopoulou G, et al. Evaluation of FISH image analysis system on assessing HER2 amplification in breast carcinoma cases. *Breast*. 2008;17(1):80–84.

91. Prins MJD, Ruurda JP, Van Diest PJ, Van Hillegersberg R, Ten Kate FJW. Evaluation of the HER2 amplification status in oesophageal adenocarcinoma by conventional and automated FISH: a tissue microarray study. *J Clin Pathol*. 2014;67(1):26–32.

92. Reljin B, Paskas M, Reljin I, Konstanty K. Breast cancer evaluation by fluorescent dot detection using combined mathematical morphology and multifractal techniques. *Diagn Pathol*. 2011;6(suppl 1):S21.

93. Tafe LJ, Allen SF, Steinmetz HB, et al. Automated processing of fluorescence in-situ hybridization slides for HER2 testing in breast and gastro-oesophageal carcinomas. *Exp Mol Pathol*. 2014;97(1):116–119.

94. Luo G, Hu Y, Zhang Z, et al. Clinicopathologic significance and prognostic value of Ki-67 expression in patients with gastric cancer: a meta-analysis. *Oncotarget*. 2017;8(30):50273–50283.

95. de Azambuja E, Cardoso F, de Castro G, et al. Ki-67 as prognostic marker in early breast cancer: a meta-analysis of published studies involving 12 155 patients. *Br J Cancer*. 2007;96(10):1504–1513.

96. Pollack A, Desilvio M, Khor LY, et al. Ki-67 staining is a strong predictor of distant metastasis and mortality for men with prostate cancer treated with radiotherapy plus androgen deprivation: Radiation Therapy Oncology Group trial 92-02. *J Clin Oncol*. 2004;22(11):2133–2140.

97. Yerushalmi R, Woods R, Ravdin PM, Hayes MM, Gelmon KA. Ki67 in breast cancer: prognostic and predictive potential. *Lancet Oncol*. 2010;11(2):174–183.

98. Pathmanathan N, Balleine RL, Jayasinghe UW, et al. The prognostic value of Ki67 in systemically untreated patients with node-negative breast cancer. *J Clin Pathol*. 2014;67(3):222–228.

99. Cheang MC, Chia SK, Voduc D, et al. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. *J Natl Cancer Inst*. 2009;101(10):736–750.

100. Petrelli F, Viale G, Cabiddu M, Barni S. Prognostic value of different cut-off levels of Ki-67 in breast cancer: a systematic review and meta-analysis of 64,196 patients. *Breast Cancer Res Treat*. 2015;153(3):477–491.

101. Harbeck N, Rastogi P, Martin M, et al. Adjuvant abemaciclib combined with endocrine therapy for high-risk early breast cancer: updated efficacy and Ki-67 analysis from the monarchE study. *Ann Oncol*. 2021;32(12):1571–1581.

102. Johnston SRD, Harbeck N, Hegg R, et al. Abemaciclib combined with endocrine therapy for the adjuvant treatment of HR+, HER2-, node-positive, high-risk, early breast cancer (monarchE). *J Clin Oncol*. 2020;38(34):3987–3998.

103. Dowsett M, Nielsen TO, A'Hern R, et al. Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in Breast Cancer working group. *J Natl Cancer Inst*. 2011;103(22):1656–1664.

104. Christgen M, von Ahnen S, Christgen H, Länger F, Kreipe H. The region-of-interest size impacts on Ki67 quantification by computer-assisted image analysis in breast cancer. *Hum Pathol*. 2015;46(9):1341–1349.

105. Leung SCY, Nielsen TO, Zabaglo LA, et al. Analytical validation of a standardised scoring protocol for Ki67 immunohistochemistry on breast cancer excision whole sections: an international multicentre collaboration. *Histopathology*. 2019;75(2):225–235.

106. Shui R, Yu B, Bi R, Yang F, Yang W. An interobserver reproducibility analysis of Ki67 visual assessment in breast cancer. *PLoS One*. 2015;10(5):e0125131.

107. Grala B, Markiewicz T, Kozłowski W, Osowski S, Słodkowska J, Papierz W. New automated image analysis method for the assessment of Ki-67 labeling index in meningiomas. *Folia Histochem Cytobiol*. 2009;47(4):587–592.

108. Remes SM, Tuominen VJ, Helin H, Isola J, Arola J. Grading of neuroendocrine tumors with Ki-67 requires high-quality assessment practices. *Am J Pathol*. 2012;36(9):1359–1363.

109. Ács B, Madaras L, Kovács KA, et al. Reproducibility and prognostic potential of Ki-67 proliferation index when comparing digital-image analysis with standard semi-quantitative evaluation in breast cancer. *Pathol Oncol Res*. 2018;24(1):115–127.

110. Ståhlhammar G, Robertson S, Wedlund L, et al. Digital image analysis of Ki67 in hot spots is superior to both manual Ki67 and mitotic counts in breast cancer. *Histopathology*. 2018;72(6):974–989.

111. Klauschen F, Wienert S, Schmitt WD, et al. Standardized Ki67 diagnostics using automated scoring—clinical validation in the GeparTrio breast cancer study. *Clin Cancer Res*. 2015;21(16):3651–3657.

112. Koopman T, Buikema HJ, Hollema H, de Bock GH, van der Vegt B. Digital image analysis of Ki67 proliferation index in breast cancer using virtual dual staining on whole tissue sections: clinical validation and inter-platform agreement. *Breast Cancer Res Treat*. 2018;169(1):33–42.

113. Roge R, Riber-Hansen R, Nielsen S, Vyberg M. Proliferation assessment in breast carcinomas using digital image analysis based on virtual Ki67/cytokeratin double staining. *Breast Cancer Res Treat*. 2016;158(1):11–19.

114. Serna G, Simonetti S, Fasani R, et al. Sequential immunohistochemistry and virtual image reconstruction using a single slide for quantitative Ki67 measurement in breast cancer. *Breast*. 2020;53:102–110.

115. Goodell PP, Krasinskas AM, Davison JM, Hartman DJ. Comparison of methods for proliferative index analysis for grading pancreatic well-differentiated neuroendocrine tumors. *Am J Clin Pathol*. 2012;137(4):576–582.

116. Nanda R, Chow LQM, Dees EC, et al. Pembrolizumab in patients with advanced triple-negative breast cancer: phase Ib keynote-012 study. *J Clin Oncol*. 2016;34(21):2460–2467.

117. Adams S, Loi S, Toppmeyer D, et al. Pembrolizumab monotherapy for previously untreated, PD-L1-positive, metastatic triple-negative breast cancer: cohort B of the phase II KEYNOTE-086 study. *Ann Oncol*. 2019;30(3):405–411.

118. Schmid P, Cortes J, Pusztai L, et al. Pembrolizumab for early triple-negative breast cancer. *N Engl J Med*. 2020;382(9):810–821.

119. Schmid P, Rugo HS, Adams S, et al. Atezolizumab plus nab-paclitaxel as first-line treatment for unresectable, locally advanced or metastatic triple-negative breast cancer (IMpassion130): updated efficacy results from a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Oncol*. 2020;21(1):44–59.

120. Miles D, Gligorov J, André F, et al. Primary results from IMpassion131, a double-blind, placebo-controlled, randomised phase III trial of first-line paclitaxel with or without atezolizumab for unresectable locally advanced/metastatic triple-negative breast cancer. *Ann Oncol*. 2021;32(8):994–1004.

121. Schmid P, Adams S, Rugo HS, et al. Atezolizumab and nab-paclitaxel in advanced triple-negative breast cancer. *N Engl J Med*. 2018;379(22):2108–2121.

122. Bankhead P, Loughrey MB, Fernández JA, et al. QuPath: open source software for digital pathology image analysis. *Sci Rep*. 2017;7(1):16878.

123. Humphries MP, Hynes S, Bingham V, et al. Automated tumour recognition and digital pathology scoring unravels new role for PD-L1 in predicting good outcome in ER-/HER2+ breast cancer. *J Oncol*. 2018;2018:2937012.

124. Denkert C, Loibl S, Noske A, et al. Tumor-associated lymphocytes as an independent predictor of response to neoadjuvant chemotherapy in breast cancer. *J Clin Oncol*. 2010;28(1):105–113.

125. Issa-Nummer Y, Darb-Esfahani S, Loibl S, et al. Prospective validation of immunological infiltrate for prediction of response to neoadjuvant chemotherapy in HER2-negative breast cancer—a substudy of the neoadjuvant GeparQuinto trial. *PLoS one*. 2013;8(12):e79775.

126. Ono M, Tsuda H, Shimizu C, et al. Tumor-infiltrating lymphocytes are correlated with response to neoadjuvant chemotherapy in triple-negative breast cancer. *Breast Cancer Res Treat*. 2012;132(3):793–805.

127. Yamaguchi R, Tanaka M, Yano A, et al. Tumor-infiltrating lymphocytes are important pathologic predictors for neoadjuvant chemotherapy in patients with breast cancer. *Hum Pathol*. 2012;43(10):1688–1694.
128. Savas P, Teo ZL, Lefevre C, et al. The subclonal architecture of metastatic breast cancer: results from a prospective community-based rapid autopsy program “CASCADE”. *PLoS Med*. 2016;13(12):e1002204.
129. Adams S, Diamond JR, Hamilton EP, et al. Phase Ib trial of atezolizumab in combination with nab-paclitaxel in patients with metastatic triple-negative breast cancer (mTNBC). Poster presented at the American Society of Clinical Oncology meeting; June 3–7, 2016; Chicago, Illinois. Abstract 1009.
130. Amgad M, Stovgaard ES, Balslev E, et al. Report on computational assessment of tumor-infiltrating lymphocytes from the International Immunology Biomarker Working Group. *NPJ Breast Cancer*. 2020;6:16.
131. Basavanthally AN, Ganesan S, Agner S, et al. Computerized image-based detection and grading of lymphocytic infiltration in HER2+ breast cancer histopathology. *IEEE Trans Biomed Eng*. 2010;57(3):642–653.
132. Rasmuson A, Zilenaite D, Nestarenkaitė A, et al. Immunogradient indicators for antitumor response assessment by automated tumor-stroma interface zone detection. *Am J Pathol*. 2020;190(6):1309–1322.
133. Saltz J, Gupta R, Hou L, et al. Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images. *Cell Rep*. 2018;23(1):181–193.e7.
134. Heindl A, Sestak I, Naidoo K, Cuzick J, Dowsett M, Yuan Y. Relevance of spatial heterogeneity of immune infiltration for predicting risk of recurrence after endocrine therapy of ER+ breast cancer. *J Natl Cancer Inst*. 2018;110(2):166–175.
135. Maley CC, Koelble K, Natrajan R, Aktipis A, Yuan Y. An ecological measure of immune-cancer colocalization as a prognostic factor for breast cancer. *Breast Cancer Res*. 2015;17(1):131.
136. Early Breast Cancer Trialists’ Collaborative Group. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet*. 2005;365(9472):1687–1717.
137. Farrugia DJ, Landmann A, Zhu L, et al. Magee equation 3 predicts pathologic response to neoadjuvant systemic chemotherapy in estrogen receptor positive, HER2 negative/equivocal breast tumors. *Mod Pathol*. 2017;30(8):1078–1085.
138. Hou Y, Tozbikian G, Zynger DL, Li Z. Using the modified Magee equation to identify patients unlikely to benefit from the 21-gene recurrence score assay (Oncotype DX assay). *Am J Clin Pathol*. 2017;147(6):541–548.
139. Hou Y, Zynger DL, Li X, Li Z. Comparison of Oncotype DX with modified Magee equation recurrence scores in low-grade invasive carcinoma of breast. *Am J Clin Pathol*. 2017;148(2):167–172.
140. Turner BM, Skinner KA, Tang P, et al. Use of modified Magee equations and histologic criteria to predict the Oncotype DX recurrence score. *Mod Pathol*. 2015;28(7):921–931.
141. Basavanthally A, Jun X, Madabhushi A, Ganesan S. Computer-aided prognosis of ER+ breast cancer histopathology and correlating survival outcome with Oncotype DX assay. Paper presented at the 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro; June 28–July 1, 2009; Boston, MA.
142. Li H, Wang J, Li Z, et al. Deep learning-based pathology image analysis enhances Magee feature correlation with Oncotype DX breast recurrence score. *Front Med*. 2022;9:886763.
143. Whitney J, Corredor G, Janowczyk A, et al. Quantitative nuclear histomorphometry predicts oncotype DX risk categories for early stage ER+ breast cancer. *BMC Cancer*. 2018;18(1):610.
144. Lu C, Romo-Bucheli D, Janowczyk A, Gilmore H, Romero E, Madabhushi A. Nuclear shape and orientation features from H&E images predict survival in early-stage estrogen receptor-positive breast cancers. *Lab Invest*. 2018;98(11):1438–1448.
145. Romo-Bucheli D, Janowczyk A, Gilmore H, Romero E, Madabhushi A. A deep learning based strategy for identifying and associating mitotic activity with gene expression derived risk categories in estrogen receptor positive breast cancers. *Cytometry A*. 2017;91(6):566–573.
146. Meti N, Saednia K, Lagree A, et al. Machine learning frameworks to predict neoadjuvant chemotherapy response in breast cancer using clinical and pathological features. *JCO Clinical Cancer Inform*. 2021(5):66–80.
147. Anand D, Kurian NC, Dhage S, et al. Deep learning to estimate human epidermal growth factor receptor 2 status from hematoxylin and eosin-stained breast tissue images. *J Pathol Inform*. 2020;11:19.
148. Conde-Sousa E, Vale J, Feng M, et al. HEROHE Challenge: predicting HER2 status in breast cancer from hematoxylin-eosin whole-slide imaging. *J Imaging*. 2022;8(8):213.
149. Wang X, Zou C, Zhang Y, et al. Prediction of BRCA gene mutation in breast cancer based on deep learning and histopathology images. *Front Genet*. 2021;12:661109.
150. Binder A, Bockmayr M, Hägele M, et al. Morphological and molecular breast cancer profiling through explainable machine learning. *Nat Machine Intell*. 2021;3(4):355–366.
151. Xu Z, Verma A, Naveed U, Bakhom SF, Khosravi P, Elemento O. Deep learning predicts chromosomal instability from histopathology images. *iScience*. 2021;24(5):102394.
152. Janowczyk A, Zuo R, Gilmore H, Feldman M, Madabhushi A. HistoQC: an open-source quality control tool for digital pathology slides. *JCO Clin Cancer Inform*. 2019;3:1–7.
153. Chen Y, Zee J, Smith A, et al. Assessment of a computerized quantitative quality control tool for whole slide images of kidney biopsies. *J Pathol*. 2021;253(3):268–278.
154. Hanna MG, Reuter VE, Ardon O, et al. Validation of a digital pathology system including remote review during the COVID-19 pandemic. *Mod Pathol*. 2020;33(11):2115–2127.