

Invited Commentary

Personalized Breast Cancer Treatments Using Artificial Intelligence in Radiomics and Pathomics

William T. Tran, MRT(T), MSc, PhD^{abcd*},
Katarzyna Jerzak, MSc, MD, FRCPC^{adef}, Fang-I Lu, MD, FRCPC^{ag},
Jonathan Klein, MSc, MD, FRCPC, DABR^h, Sami Tabbarah, MSc^{ad},
Andrew Lagree, MSc^{ad}, Tina Wu, BSc (cand)^{ad},
Ivan Rosado-Mendez, MSc, PhDⁱ, Ethan Law, BSc (cand)^{ad},
Khadijeh Saednia, MSc^{dj}, and Ali Sadeghi-Naini, MSc, PhD^{adj}

^a Odette Cancer Program, Sunnybrook Health Sciences Centre, Toronto, Canada

^b Faculty of Medicine, Department Radiation Oncology, University of Toronto, Toronto, Canada

^c Faculty of Health and Wellbeing, Sheffield Hallam University, Sheffield, United Kingdom

^d Radiogenomics Laboratory, Sunnybrook Health Sciences Centre, Toronto, Canada

^e Division of Medical Oncology, Sunnybrook Health Sciences Centre, Toronto, Canada

^f Faculty of Medicine, Department of Medicine, University of Toronto, Toronto, Canada

^g Anatomic Pathology, Sunnybrook Health Sciences Centre, Toronto, Canada

^h Department of Radiation Oncology, Albert Einstein College of Medicine, Bronx, USA

ⁱ Instituto de Fisica, Universidad Nacional Autónoma de México, Mexico City, Mexico

^j Department of Electrical Engineering and Computer Science, Lassonde School of Engineering, York University, Toronto, Canada

ABSTRACT

Progress in computing power and advances in medical imaging over recent decades have culminated in new opportunities for artificial intelligence (AI), computer vision, and using radiomics to facilitate clinical decision-making. These opportunities are growing in medical specialties, such as radiology, pathology, and oncology. As medical imaging and pathology are becoming increasingly digitized, it is recently recognized that harnessing data from digital images can yield parameters that reflect the underlying biology and physiology of various malignancies. This greater understanding of the behaviour of cancer can potentially improve on therapeutic strategies. In addition, the use of AI is particularly appealing in oncology to facilitate the detection of malignancies, to predict the likelihood of tumor response to treatments, and to prognosticate the patients' risk of cancer-related mortality. AI will be critical for identifying candidate biomarkers from digital imaging and developing robust and reliable predictive models. These models will be used to personalize oncologic treatment strategies, and identify confounding variables that are related to the complex biology of tumors and diversity of patient-related factors (ie, mining "big data"). This commentary describes the growing body of work focussed on AI for precision oncology. Advances in AI-driven

computer vision and machine learning are opening new pathways that can potentially impact patient outcomes through response-guided adaptive treatments and targeted therapies based on radiomic and pathomic analysis.

RÉSUMÉ

Les progrès dans la puissance des ordinateurs et les avancées en imagerie médicale au fil des dernières décennies ont mené à de nouvelles avenues pour l'intelligence artificielle (IA), la vision artificielle et l'utilisation de la radiomique pour faciliter la prise de décision clinique. Ces avenues s'ouvrent de plus en plus largement dans des spécialités médicales comme la radiologie, la pathologie et l'oncologie. L'imagerie médicale et la pathologie étant de plus en plus numérisées, on a récemment reconnu que les données provenant des images numériques pouvaient fournir des paramètres reflétant la biologie et la physiologie sous-jacentes de différentes tumeurs malignes. Cette compréhension accrue du comportement du cancer pourrait potentiellement nous permettre d'améliorer nos stratégies thérapeutiques. De plus, le recours à l'intelligence artificielle est particulièrement attrayant en oncologie, pour faciliter la détection des tumeurs, prédire la probabilité de réponse tumorale aux traitements et établir un pronostic quant au risque de mortalité lié au cancer pour le patient. L'intelligence artificielle jouera un rôle essentiel dans l'identification des biomarqueurs candidats pour l'imagerie médicale, le développement de modèles prédictifs robustes et fiables pouvant être utilisés pour personnaliser les stratégies de traitement en oncologie et par la reconnaissance de tendances parmi le grand nombre de

* Corresponding author: William T. Tran, MRT(T), MSc, PhD, Department of Radiation Oncology, Sunnybrook Health Sciences Centre, 2075 Bayview Avenue, RM TB 097, Toronto, Ontario M4N3M5, Canada.

E-mail address: william.tran@sunnybrook.ca (W.T. Tran).

variables confondantes associées à la biologie complexe des tumeurs et à la variété des facteurs reliés aux patients (l'exploration des données massives). Ce commentaire décrit le nombre grandissant de travaux mettant l'accent sur l'intelligence artificielle et l'oncologie de précision. Les avancées dans la vision artificielle guidée par IA et

l'apprentissage machine ouvrent de nouvelles avenues qui pourraient potentiellement avoir une incidence sur les résultats pour les patients par des traitements adaptatifs guidés par la réaction et les thérapies ciblées basées sur l'analyse radiomique.

Keywords: Breast cancer; radiomics; pathomics; pathology; digital pathology; informatics

Introduction

In 1965, Gordon Moore, the cofounder of *Fairchild Semiconductors International Inc*, who later became the chief executive officer of *Intel Corporation*, released a white paper entitled, “*Cramming more components onto integrated circuits*” [1]. His paper described the rapid rate of development in computing hardware; specifically, he projected that the number of components per integrated circuit (ie, the computer chip) would double every 12 months, conferring increasingly greater computational power over time. Subsequently, Moore had revised his model with components doubling every 24 months, a phenomenon that is known today as *Moore’s Law*. Indeed, Moore’s Law has been well observed; there has been an unprecedented expansion in computer engineering, technology, and capability over recent decades.

Computers continue to be at the forefront of opportunities for industrial development, societal growth, and advancing medical science. In the modern medical era, computers (hardware and software components) are critical for aiding diagnosis and delivering medical treatment, with widespread uses in the medical specialties, such as ophthalmology, radiation oncology, radiology, and surgery. Computers are also virtual repositories for petabytes (10^{15} bytes) of data that contain medical images, clinical reports, medical progress notes, and patient demographic information. These large data sets are centralized in high-capacity computer servers that are, in principle, minable (ie, “*big data*”) using artificial intelligence (AI) algorithms to gain *actionable insight* for clinical decision-making.

AI is a domain of computer science that employs mathematical and statistical algorithms to make machine-based inferences that would otherwise be performed by human cognition. The inferences generated through AI are grounded in underlying data including knowledge, symbols, perceptions, observations, patterns, reasoning, and constraints. Subdomains of AI also include machine learning. Machine learning uses algorithms to recognize patterns and relationships from available training data to cluster or classify new data samples [2]. Deep learning, a recently introduced branch of machine learning, applies a system of artificial neural networks (ANNs) with several hidden layers that compute a transformation of the underlying data that result in an output layer associated with a class [3].

Machine learning and deep learning models have also been adapted for studies in medicine; particularly, in oncology, there is ongoing research to address clinical challenges, which include (1) accurate computer-aided diagnoses, (2) monitoring drug efficacy, (3) predicting treatment response (ie, theranostics), and

(4) prognostication. Deep learning has been shown to be useful for detecting and segmenting malignancies captured in medical images and extracting relevant biomarkers from quantitative and functional imaging. The process of extracting information from images and studying high-dimensional imaging biomarkers for predictive and prognostic modelling is known as *radiomics* [4]. Radiomics and AI hold the promise of providing clinicians and patients with information that can possibly guide treatments, personalize therapeutic strategies, reduce the delays in diagnosis, and may also play a role in preventative oncology.

In this commentary, we present principles and applications of AI (including machine learning and deep learning), and radiomics for precision oncology within the context of breast cancer. This commentary comprises four sections. In Section [Radiomics](#), the concept of radiomics is presented. Here, we outline image processing techniques and the use of AI to attain radiomic features. In Section [Machine Learning Classification in Oncology](#), machine learning constructs in the context of modelling and classification are presented. Commonly implemented AI algorithms are presented for breast cancer studies. Section [Pathomics: Machine Learning Applications in Breast Oncology](#) describes emerging applications of AI in pathology as they are related to oncology (ie, pathomics). Finally, current challenges and opportunities for AI in breast cancer are discussed.

Radiomics

Digital imaging has succeeded analog radiology systems in modern clinics. This is attributed to the increased utilization of electronic signal detectors that constitute more recent devices such as computed tomography (CT), magnetic resonance imaging (MRI), and digital X-ray. Computers are an essential component to reconstruct digital images into formats, such as, DICOMs (digital imaging and communication in medicine) that can be retrieved, reviewed, and processed for systematic data mining through radiomics frameworks. The widespread implementation of digital radiology has afforded greater opportunities for radiomics analysis; for example, access to large medical imaging and informatics databases for the purpose of extracting biomarkers related to cancer detection, diagnosis, treatment, and surveillance. Medical images can also yield information about tissue phenotypes and the underlying physiology. Such radiomic descriptors include morphological features of lesions, as well as first-, second-, or higher-order intensity features (textures) that can be linked to tissue microstructure and heterogeneity. It may also be used to relay

functional characteristics of tumors, such as, blood flow, cell metabolism, and cell death. The synergy between radiomics and AI can confer insight into the tumor's behaviour (eg, identifying aggressive vs. indolent tumors), and this insight can potentially aid clinical decision-making in breast cancer management. However, major challenges involve robust data provenance; that is, machine learning and deep learning models are only as good as the input data. In essence, erroneous radiomic data sets will inevitably result in faulty, inaccurate, or overestimated models that are not reproducible, repeatable, or ultimately, clinically useful. Overcoming these challenges may be addressed, in part, by adhering to standard protocols within the radiomics pipeline (described in the following sections).

Radiomics Pipeline and Features

Radiomics analysis is a multistep process that includes (1) standard image acquisition protocols, (2) segmentation (region of interest delineation), (3) feature extraction, and (4) analysis and modelling (Figure 1) [5]. Standard imaging employs parameters that are optimal and reproducible during acquisition and data collection. Segmentation is the process of delineating the anatomical region that will be targeted for image analysis. Feature extraction is the process by which image processing yields candidate parameters, features, or biomarkers for classification. Finally, analysis and modelling organize the data and calculate statistical associations between features and the task or outcome. At each step, establishing robust data provenance is critical to obtain reproducible and repeatable results for clinical translation. Repeatability involves measuring the imaging biomarkers from the same subject (ie, patient), equipment, or software to ensure that the measurements are consistent from one test series to the next. Reproducibility focuses on obtaining the same results using different imaging devices (eg, two different MRI scanners with similar settings), different users, or different software [6]. Reproducibility and repeatability are major challenges within each step of the radiomics pipeline and may potentially be addressed by automated analysis that is driven by AI. Radiomics features can be derived using second-

or higher-order statistical approaches from image processing. Second-order analysis includes texture analysis, which quantitates image heterogeneity by quantifying pixel-to-pixel relationships that have been previously described elsewhere [7]. Briefly, grey-level textures, including Haralick features, have been widely studied for radiomics analysis [8,9]. Second-order texture features can be extracted from medical images using (1) grey-level co-occurrence matrix; (2) grey-level run length matrix; and (3) neighborhood grey tone difference matrix.

Machine Learning Classification in Oncology

Employing machine learning algorithms for categorization of disease traits and predicting clinical endpoints (eg, tumor recurrence risk) depend heavily on constraints to the classifier model itself. In this section, we present considerations for implementing machine learning classifiers within the oncology context.

Overall, machine learning algorithms are structured to explore training data, identify patterns and relationships, and devise models that relate the data to measured outcomes. In general, machine learning algorithms are trained to find the relationship between the independent variable(s) "X" and outcome/dependent variable(s) "Y." The independent variables (ie, "X" variables) are known as *descriptors*, *features*, or *attributes* and are extracted from measured observations or examples. In the radiomics context, these may include texture features and shape descriptors. Correct labels of data samples (ie, variable "Y") are referred to as the *ground truth* and are bound to empirical outcomes, classes, or events. These may include clinical endpoints, such as, cancer recurrence, death, or measures of drug resistance. Ground truths may also be tissue classifications, such as, benign vs. malignant types. Ground truth labels are also known as "gold-standard" classifiers and often require manual evaluation or input from human (expert) counterparts.

An important distinction exists between *supervised learning* and *unsupervised learning* in the context of the algorithm. Supervised machine learning involves presenting the

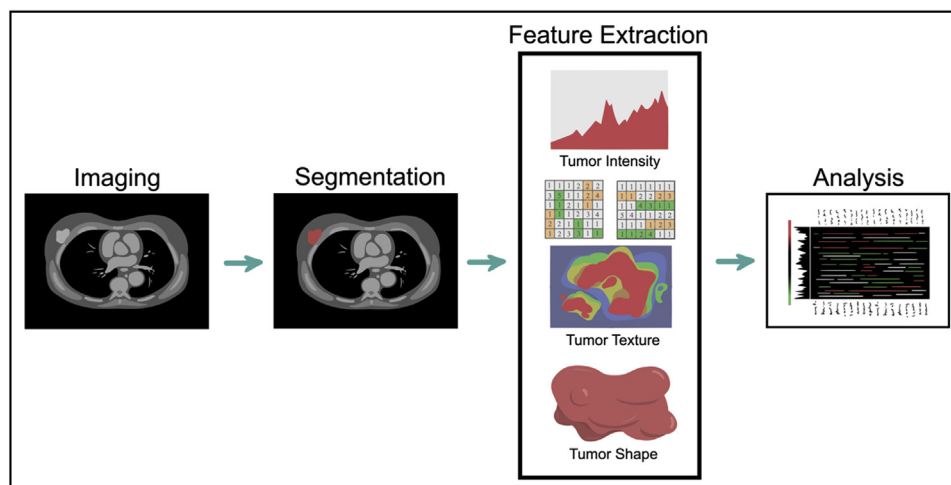


Figure 1. Radiomics pipeline. Images are acquired using standard and protocol-based techniques, followed by selecting the region of interest (segmentation). Imaging biomarkers are obtained using first-, second-, or higher-order features. The last step involves analysis, which uses artificial intelligence for modelling.

classification labels of the data samples to the learning algorithm upfront. This approach is useful for *post hoc* analysis, where the algorithm-learning from the data develops models related to the measured target or outcome. In contrast, unsupervised machine learning involves mining data samples represented by a set of features or attributes with no label presented to the algorithm, for example, a clustering method. By using clustering algorithms, such as *k-means* clustering, distinct clusters (groups) of the data samples can be identified based on the data distribution and structure and subsequently applied for classification. Finally, certain data sets may have a mixture of labels and hidden/unknown labels and these problems are approached with semi-supervised machine learning techniques. For training and validation design, obtaining the appropriate ground truth labels for computational oncology is still an ongoing challenge. First, many machine learning algorithms are structured for binary class labels, whereas clinical endpoints are often structured as ordinal, continuous, categorical, or descriptive (ie, qualitative) data. Thus, choosing the appropriate cutoff boundary within ordinal or continuous data sets remains a significant chal-

lenge. Second, there is still no consensus about choosing the appropriate clinical endpoint standard in many applications. For example, in locally advanced breast cancer, response to neoadjuvant chemotherapy (NAC) may be evaluated using a plethora of pathological assessment guidelines including Miller-Payne, Chevalier, synoptic pathology, Residual Cancer Burden Index, or using the American Joint Committee of Cancer Criteria (AJCC) [10]. Finally, it is important to note that some clinical variables are important for outcome prediction. For breast cancer, tumor size, nodal status, and residual cancer burden are predictors of survival outcomes. In addition, recurrence patterns and survival rates are highly dependent on breast cancer subtypes; for example, estrogen receptor (ER)-positive/HER2-negative breast cancer typically demonstrates a long metastasis latency period compared with triple-negative breast cancer [11].

Taken together, AI and machine learning algorithms should be chosen carefully, based on the data type and outcome measures. Here, we describe commonly utilized machine learning classifiers applicable to oncology.

Machine Learning Classifiers

k-Nearest Neighbor

The *k*-nearest neighbor (*k*-NN) classifier is a nonparametric algorithm used for classification. It is considered to be one of the least computationally demanding algorithms for supervised machine-learning. The *k*-NN algorithm makes no assumptions about the form of the data distribution (eg,

Gaussian distribution), and therefore, it is ideal for exploratory studies where there is *no prior knowledge* about the attributes and distribution of the data [12]. The *k*-NN classification uses a weighting function that varies in value based on the distance between a sample and its neighbor(s) (*k* = number of nearest neighbor(s) considered), seeking out patterns in the distribution of the data within a sample set [12]. The data samples (known as “instances”) are treated in groups or “bags” with a defined label [2]. The bags are then analysed in terms of their attributes or features (Figure 2).

The *k*-NN algorithm first organizes the bags of training samples into a feature space based on the values of the attributes. The test samples are assigned a label according to a majority vote that is dependent on the nearest neighbor as determined by the Euclidean distance calculation (Eq. 1, Figure 2).

A Euclidean distance between two samples, s_1 and s_k , is defined as follow: where $(a_n^{(s)})$ represents the attribute *n* of the sample *s*. Because different attributes may have varying scales or units of measure, data are often normalized between 0 and 1 for analysis within the *k*-NN feature space.

$$Distance(s_1, s_k) = \sqrt{\left(a_1^{(s_1)} - a_1^{(s_k)}\right)^2 + \left(a_2^{(s_1)} - a_2^{(s_k)}\right)^2 + \dots + \left(a_n^{(s_1)} - a_n^{(s_k)}\right)^2} \quad (1)$$

Naïve Bayes Classifier

Naïve Bayes classification can be used to predict the probability of a binary class membership. The algorithm uses the probabilities of the class label and its attributes to compute a probability prediction of a sample. An important assumption for naïve Bayes classification algorithms is that the individual attributes (a_1, a_2, \dots, a_n) of a class are independent to each other (conditional independence) [13]. The naïve Bayes classifier function is expressed as follows (Eq. 2):

$$P(X|C_i) = \prod_{k=1}^n [P(x_k|C_i)] \quad (2)$$

where $P(X|C_i)$ is the probability of sample (X) labelled into a class (C_i), and x_k represents the value of attribute a_k for the sample X [13].

Support Vector Machine Classifiers

Support vector machines (SVMs) are used to solve classification problems based on pattern recognition from sample features. There are four main elements in an SVM classifier: (1) a separating line or hyperplane within a feature space, (2) a margin, (3) support vectors, and sometimes (4) a kernel function (Figure 3). The hyperplane serves as a decision boundary between the two classes, with a maximum margin calculated for optimal discrimination. The hyperplane is represented as follows [14]:

$$\vec{w} \cdot \vec{x} + b = 0 \quad (3)$$

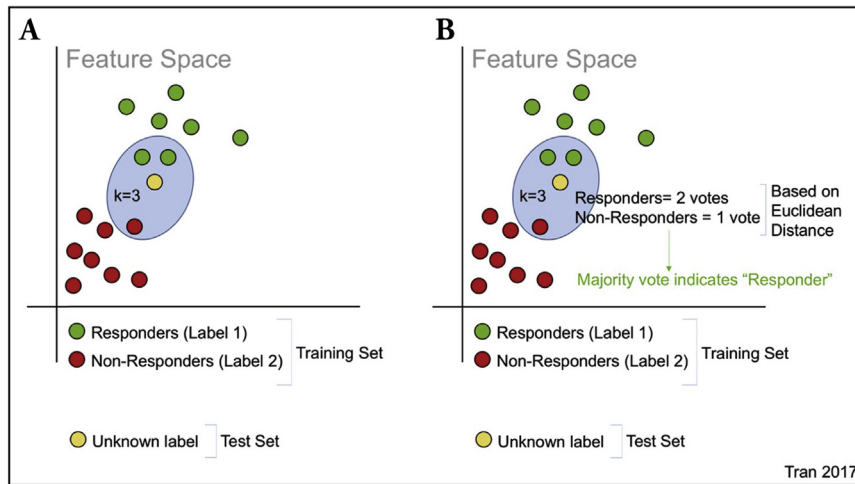


Figure 2. The k-nearest neighbor (k-NN) classifiers: An example of predicting response vs. nonresponse to a drug agent. The k is three in this representation. The bags are labelled as responders (green) or nonresponders (red) in the training set (A). A test sample is used to find the three nearest neighbors using a Euclidean distance calculation. Depending on the labels of the nearest neighbors, a vote is tabulated, and the majority vote determines how the test sample is labelled (B).

where \vec{x} is the feature vector, \vec{w} is the normal vector to the hyperplane, and b is the offset of the hyperplane from the origin. In a normalized feature space, the margin is defined as the region bounded by the two boundaries expressed by the following equations [14]:

$$\vec{w} \cdot \vec{x} + b = -1 \quad (4)$$

$$\vec{w} \cdot \vec{x} + b = +1 \quad (5)$$

The goal is to maximize the margin $\left(\frac{2}{\|\vec{w}\|}\right)$ subject to the following constraints [14]:

For linearly separable data, maximize $\frac{2}{\|\vec{w}\|}$, such that, $\vec{w} \cdot \vec{x} + b \geq +1, \forall x$ of label "class 1" and $\vec{w} \cdot \vec{x} + b \leq -1, \forall x$, of label "class 2."

Limitations of the support vector machine learning are challenges with scalability within the feature space, as well as the computational burden when there is highly dimensional data (ie, many features to model).

Artificial Neural Networks

The basic framework of neural networks consists of many neurons (units) that are structured to activate from a threshold command of another neuron's weighted input signal. Thus, the network is composed of several interconnected neurons that work through structured input/output functions. The earliest neural network was introduced as a single-layer perceptron, which consisted of only an input layer and an output layer (the input layer is not considered a network layer) [15]. Subsequently, multilayer neural networks were developed, defined by the addition of hidden layers to overcome the challenges of single-layer perceptrons such as inefficient handling of large, complex, and multidimensional data. The architecture of a neural network is organized as multiple structured layers with a fixed number of units within each layer (Figure 4). The hidden layers are tasked with finding attributes associated with the output layer. Concisely, the

architecture of neural networks is often predicated by (1) an input layer, (2) a number of hidden layers, (3) an output layer, (4) fixed units within each layer, (5) fully connected units between neighboring layers, (6) neurons within the same layer that are not connected, and finally in feed-forward ANNs, (7) the neural connections are unidirectional [15]. Using a two-layer neural network as an example, we can express the input vector as follows [15]:

$$v = [v_i] \in \mathbb{R}^D \quad (6)$$

With respect to the output neuron y_k , the estimation function can be expressed as follows [15]:

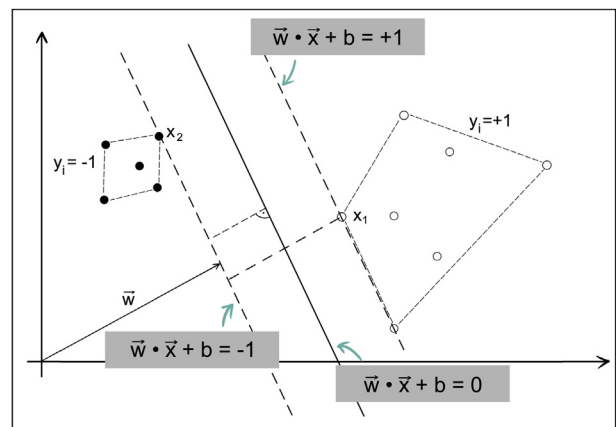


Figure 3. Support vector machine learning. The support vector machine comprises a separating line or hyperplane in the feature space, as well as a maximal margin that is determined by the support vectors. Figure adapted from Hearst et al [14].

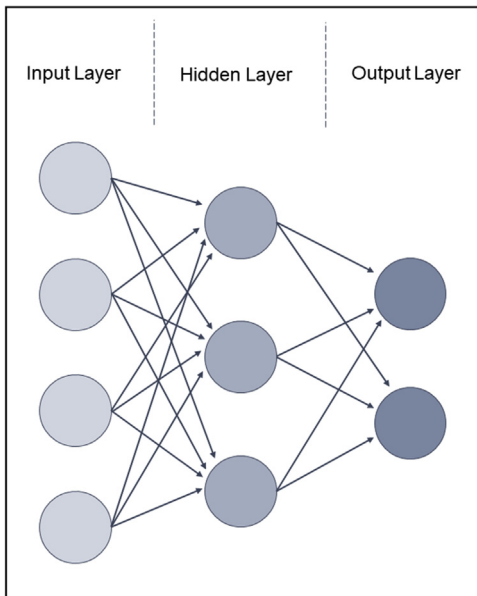


Figure 4. Artificial neural networks. The artificial neural networks contain multiple layers, each with several units (neurons) that work on input/output functions.

$$y_k(v; \theta) = f^{(2)} \left(\sum_{j=1}^M W_{kj}^{(2)} f^{(1)} \left(\sum_{i=1}^D W_{ji}^{(1)} v_i + b_j^{(1)} \right) + b_k^{(2)} \right) \quad (7)$$

where $f^{(1)}(\cdot)$ and $f^{(2)}(\cdot)$ are considered nonlinear activation functions of units pertaining to the respective layer, and M is the number of hidden units within the layer [15]. Also, θ denotes parameter sets (including connections), where $\theta = \{W^1, W^2, b^1, b^2\}$ [15].

In the cases where the hidden layers are used to extract features (such as in radiomics analysis), then the output layer is expressed as a linear model [15]:

$$y_k(v; \theta) = f^{(2)} \left(\sum_{j=1}^M W_{kj}^{(2)} \phi_j(v) + b_k^{(2)} \right) \quad (8)$$

As part of the training process, errors in the learning algorithm can be detected by finding inefficiencies in the parameter set, using a back-propagation method.

Feed-forward ANNs are just one approach within a multitude of neural network algorithms. Multilayer neural networks encompass a broader domain that include deep learning methods, which are algorithms that have greater than two layers. This multilayer architecture can approach more complex problems [15]. Although describing all the deep learning models goes well beyond the scope of this review, it is important to mention that other models such as the deep Boltzmann machine and convolutional neural networks (CNNs) are increasingly utilized for radiomic analysis and machine learning to extract imaging-based features for modelling.

Pathomics: Machine Learning Applications in Breast Oncology

In contrast to radiomics, which largely employs *in vivo* imaging, pathomics concerns analysis of digital microscopy images of tissue, cells, and subcellular structures *ex vivo*. Pathomics involves preparing whole-slide specimens, applying conventional microscopy techniques, transferring the whole-slide specimens into digital formats (ie, digitizing the images), and performing *in silico* analysis using AI to relate image features (biomarkers) to diagnosis, treatment, and clinical endpoints. There are immense opportunities for pathomics in oncology, primarily, because conventional (manual) pathology reporting already plays a critical role in the clinic. For example, the mitotic count, nuclear pleomorphism, and tubule formation are used to calculate the Nottingham tumor grade in breast cancer, which relays information about the tumor's aggressiveness. The tumor grade is also associated with clinical endpoints (eg, tumor recurrence), and this plays a role in clinical decision-making in breast radiation and medical oncology [16,17]. Another pathology-based analysis is tumor cellularity, which is correlated to survival endpoints after NAC for breast cancer [11]. Highly clustered nuclei are associated with dense cellular areas, and aberrant nuclear morphology. These characteristics are associated with malignancies that are distinct from normal or benign tissue. The driving factors that cause morphological and spatial differences in tumor cells are attributed to abnormal cell function and increased cell proliferation [18]. Within this framework, pathomic studies aim to develop computational methods to characterize and automate the detection of conventional attributes such as histomorphology (ie, nuclear shape, size, and color) and tissue-spatial characteristics such as cellular distribution. Pathomics also aims to extract handcrafted or data-driven attributes (ie, biomarkers) to be used for cancer diagnosis and clinical decision-making. The benefits include higher efficiency, greater standardization for pathologists and clinicians, and discovery of previously unknown predictive and prognostic biomarkers.

Substantial research has focused on AI for automated analysis of digital whole-slide specimens. One promising research avenue is deep learning algorithms that will match the diagnostic performance of manual inspection by the pathologist. Wang et al (2014) [19] developed a computerized system to quantitate mitotic activity from high-power field (400 \times) digital pathology images. Their group used a CNN architecture consisting of three layers: two sequential convolutional and sequential layers and then an output layer consisting of two neurons reflecting binary classes (mitosis vs. nonmitosis) [19]. The classification accuracy of the CNNs was evaluated using the following performance measures: (1) recall (sensitivity), (2) precision, and (3) the *F-measure*. The results demonstrated that modelling handcrafted and CNN-based features demonstrated high classification accuracy (*F-measure* = 0.734) [19]. Other research groups have explored using machine learning to determine tumor grade. Petushi et al

(2004) [20] developed an automated analysis pipeline consisting of (1) grayscale conversion, (2) threshold segmentation, (3) object (nuclei, extracellular matrix, adipose tissue) labeling, (4) feature extraction, and finally, (5) object classification using supervised machine learning. The machine learning classifiers that were tested consisted of probability-density functions (eg, linear discriminant functions, gaussian, or gaussian mixture), neural network posterior-probability classifiers (eg, multilayer perceptrons), and boundary-forming classifiers (eg, k-nearest neighbor, decision tree). Data comprised handcrafted (input) features using high-order texture parameters and feature vectors based on the distance between neighboring nuclei. The results of the study showed that the quadratic classifiers had the least classification errors (error = 0.304), when compared with ground truths, suggesting the potential to achieve concordance between machine-classified and pathologist-reported tumor grades [20]. More recent research is addressing common challenges of whole-slide preparation, such as overlapping tumor cells that can erroneously train the learning algorithms; for example, detection of two distinct cells with overlapping nuclei, but classification as one cell undergoing mitosis [21]. Researchers from Chongqing University (China) approached this problem by using a curvature scale space corner detection method to “split” the overlapped cells to improve detection and classification. The feature-set that was used for classification included shape-based features ($n = 4$) and grey-level components of multiple color spaces to attain texture features ($n = 138$) [21]. Their study demonstrated excellent classification accuracies using a SVM classifier, which was able to distinguish normal from malignant areas (accuracy = $96.19\% \pm 0.31\%$) [21].

Computer vision for pathological analysis plays an important role in detecting the relevant areas (ie, tumor location) from whole-slide specimens. Following this, it is possible to generate predictive or prognostic assays by applying traditional statistical frameworks (eg, Cox regression models) or machine learning classifiers of the digital pathology features to find associations to ground truth labels (eg, clinical and survival endpoints, or tissue-type classification). The Cambridge group has recently reported results using digitized pre-treatment tumor biopsies and post-treatment surgical specimens obtained from the ARTemis trial [22] and from the Neo-tAnGo trial [23]. The primary aim of those studies was to investigate pathomic markers of NAC response in breast cancer, which included 1,223 baseline samples. Machine learning algorithms using a k-NN and support vector machine learning was used to classify tumor, stroma, and lymphocytes of the digitized pathology samples [22,23]. The results showed that lymphocyte density, as measured in the pre-treatment biopsy, was an independent predictor of pathological complete response (pCR); defined as a complete disappearance of invasive cancer cells after treatment (odds ratio = 2.92–4.46, $P < .001$) [22]. Other research from Romo-Bucheli et al (2016) [24] have used deep neural networks to identify tubule-nuclei structures from digitized whole-slide

images ($n = 174$) of ER-positive breast tumors. For this study, the deep neural network architecture was structured using a convolutional neural network, a rectifier linear unit (ReLU) and a maximum pooling (max pool) operator [24], and the output layer that consisted of a binary class label for (\pm) tubule nuclei. The ratio between tubule nuclei and the total number of nuclei was indexed as the tubule formation indicator (TFI). Their algorithm demonstrated modest classification results, showing an F-score of 0.59 ± 0.14 and a precision of 0.72 ± 0.12 [24]. The research group’s secondary aim was to evaluate the association between the TFI score to the Oncotype DX prognostic risk score. The results showed that low Oncotype DX scores (when modelled with the low tumor grade) were associated with a larger TFI ($P < .01$). In contrast, a high Oncotype DX score and high tumor grade had a smaller TFI value ($P < .01$) [24]. The results of this study suggest that pathomics-based analysis using deep learning architecture may provide information about breast cancer prognosis. Such prognostic pathomic markers have been studied for oropharyngeal cancer [25] and lung cancer [26–28].

Taken together, the opportunities for AI in oncology are immense. Machine learning and deep learning applications are demonstrating promising new assays for predicting treatment response and survival outcomes for high-risk breast cancer.

Discussion

It is anticipated that clinical support tools generated from AI will gain widespread adoption at the patient’s bedside and in the clinicians’ workspace. However, clinicians’ confidence in computer-assisted medicine and the incorporation of computational oncology will demand robust validation in large prospective randomized trials. With the rapid development of computer software and hardware, AI will increase its presence in the diagnostic workup and may enhance treatment strategies, as well as surveillance practices based on the predicted individualized recurrence risk. Ultimately, there are immense opportunities for AI to increase efficiency in the clinical workflow to improve patient care. Here, we discuss the pertinent opportunities for AI in the clinical management of breast cancer.

Neoadjuvant Chemotherapy for Breast Cancer

Breast cancer heterogeneity confers a therapeutic challenge. This is, in part, due to genetic and molecular alterations that result in variable tumor phenotypes, which ultimately affects the tumors’ responses to cytotoxic agents. Hundreds of genetic and molecular drivers of breast oncogenesis and progression have been identified, including genes responsible for proliferation, cell cycling, invasion, and metastasis [29,30]. Breast cancer intrinsic subtypes (luminal A, luminal B, basal like, normal like, and HER2-overexpressed) demonstrate variable clinical presentations and therapeutic-response patterns. The CTNeoBC pooled analysis (2014) showed low pCR rates

after NAC in women with luminal A breast cancer (7.5%; CI, 6.3%–8.7%), whereas 33.6% (CI, 30.9%–36.4%) of triple-negative breast cancer patients achieved a pCR [31]. Indeed, the pCR rates for all subtypes analyzed ranged from 7.5% to 50.5% [31]. These results underscore the variance in clinical endpoints. Further to this, the German Breast Group has also shown differences in long-term outcomes to NAC according to intrinsic subtype [32]. In their study, women with luminal B/HER2-negative ($P < .013$), HER2-positive ($P < .001$), and triple-negative ($P < .001$) breast cancer [32] who achieved a pCR were shown to survive longer than women who did not achieve a pCR [32]. On the other hand, pCR after NAC did not confer a survival advantage among patients with luminal A breast cancers [32]. Although these important studies have stratified oncologic risk at a population level, there is still an opportunity to better understand the individualized risk factors, such that chemotherapy may be personalized for each patient to (1) improve tumor response rates and (2) to prolong survival. If an AI-assisted algorithm could reliably predict pCR upfront, before initiation of NAC, this could allow medical oncologists to personalize patients' treatment regimens. AI-assisted strategies could also include dose-tailoring, allowing clinicians to administer dose-escalated systemic therapies based on AI-predictive models.

Radiotherapy Dose De-escalation in Early-Stage, Low-Risk Breast Cancer

The role of radiation therapy (XRT) following breast conserving surgery is to eradicate any potential microscopic tumor deposits remaining in the surgical bed. Prognosis for early-stage ER-positive/HER2-negative breast cancer is generally excellent with the 5-year disease-free survival of up to 87% [33]. The *National Surgical Adjuvant Breast and Bowel Project B-06* (NSABP B-06) trial demonstrated that women with tumor-free margins following breast conserving surgery, who received adjuvant whole-breast XRT, had a 20-year cumulative incidence of recurrence of 14.3% compared with 39.2% for patients who received surgery alone ($P < .001$) [34]. Following this, there was interest in ascertaining the role of endocrine therapies alone for locoregional control. Endocrine therapies such as tamoxifen (TAM) were studied intensely; for example, the NSABP B-21 trial showed that TAM + XRT resulted in a reduction in ipsilateral breast tumor recurrence (IBTR) compared with TAM alone, for small (<1 cm) tumors (Hazard ratio = 0.19, $P < .0001$) [35]. Similarly, Fyles et al (2000) demonstrated that TAM + XRT significantly reduced the five-year IBTR rates in women older than 50 years and with clinical stage T1-T2 tumors (Hazard ratio = 8.3, $P < .001$). This study reported a disease-free survival benefit with XRT added to TAM ($P = .004$), but there was no overall survival advantage between groups ($P = .83$). The CALGB 9343 trial showed omitting XRT in women >70 years of age, with T1, ER+, lymph-node-negative breast cancer increased the rate of locoregional recurrence by 8% ($P < .001$) [36]. Likewise, the PRIME II

trial reported on women >65 years of age, ER-positive, lymph-node-negative breast cancer that XRT added to hormonal therapy reduced the 5-year IBTR rates to 1.3% from 4.1% ($P = .0002$) [37]. These studies show that better locoregional control can be achieved with XRT in a large cohort of patients with low-risk characteristics. However, it is still unclear if, based on the tumor biology, these patients were inherently low risk for IBTR. In essence, these studies have not yet shown the individualized risk, at the tumor and patient level, associated with IBTR and invasive disease-free survival. Thus, there is an opportunity to study if AI-assisted decision support tools can help determine if dose de-escalated radiotherapy might be feasible in low-risk breast patients. Dose de-escalation studies for adjuvant chemotherapy have also shown promising results using genetic markers, such as the NSABP-B14 trial that has now progressed into the Oncotype DX-validated gene assay (Genomic Health Inc, Redwood) [38]. The Oncotype DX assay uses a 21-gene panel to stratify patients with ER-positive, HER2-negative breast cancer into low-, intermediate-, and high-risk categories based on distant recurrence risk. This information has now enhanced the dialogue between medical oncologists and patients when discussing the relative benefit and risks of adjuvant chemotherapy to make more personalized, patient-centered treatment decisions. Similar to the development of Oncotype DX, there are still immense opportunities within the context of radiotherapy, to use machine learning to model relative risks associated with tumor behaviour, progression, radiotherapy-response, and locoregional recurrence.

Multi-Omics and Comprehensive AI Modelling

Big data can be retrieved from multi-omics, by combining information from such disciplines as genomics, transcriptomics, proteomics, metabolomics, radiomics, and pathomics. Each of these branches hold great promise for understanding cancer's aberrant lifecycle; however, together may provide a more comprehensive, multi-level portrait of the tumor's biology. Multi-omics data may also provide better models to predicting treatment response. Stetson et al (2017) [39] used multi-omic data to investigate drug response signatures in several cell lines, including breast cancer. Machine learning models included using a random forest classifier and SVM. Their results showed that multi-omic transcriptomic and genomic markers were predictive of breast cancer response to taxane chemotherapy [39]. Other studies have explored cross-modality multi-omics, such as *radiogenomics* (ie, combining radiomics and genomics), underscoring the relationship between genetic alterations and the resulting phenotype, ultimately, detected by quantitative imaging [40]. Radiogenomic studies have been intensely focused on quantitative MRI and genomic relationships in recent years, as quantitative MRI techniques, such as dynamic contrast-enhanced MRI techniques, and genotyping platforms have become increasingly sophisticated. A recent study by Saha et al (2018) [41] studied a large patient cohort ($n = 922$) and extracted 529 radiomic features. Genomic data consisted

of markers for ER, progesterone, HER2, and the proliferative marker ki67 [41]. The aim of the study was to find associations between radiomic and genomic markers to classify breast cancer subtypes. This study showed that radiogenomic markers can successfully distinguish luminal A breast cancer, corresponding to area under the curve of 0.697 (CI, 0.647–0.746; $P < .0001$) [41].

Taken together, the major challenge for both clinicians and the health system is to find efficient strategies to tackle problems that might improve patient outcomes. AI may offer such informatics-based solutions to mine the data, seek patterns, and ultimately build models that confer a deeper understanding of breast cancer.

Opportunities for Radiomics in Contexts of Limited Resources

Access to medical imaging services involves large capital costs, highly skilled personnel, and infrastructural support. An important challenge regarding the emergence of radiomics is the cost of accessing imaging resources (MRI, CT, and PET scanners), data banks, and the significant time required for allied health professionals, radiologists, and pathologists to curate images and interpret findings. This challenge is more relevant in health systems from low- and middle-income countries. In response, there has been significant effort to use less-expensive imaging modalities to make access to radiomics more affordable and accessible.

Within this framework, various research groups are exploring the use of quantitative ultrasound (QUS). QUS is relatively inexpensive compared with MRI, CT, or PET, and other advantages include system portability and mitigating operator and system dependence during image acquisition and processing. There is current work to validate QUS-radiomics in the long term, particularly in breast cancer [42–44]. Novel QUS imaging techniques to attain imaging biomarkers aim for objective, reproducible, and have the high potential for AI-driven automation and interpretation. These attributes can be useful for low- and middle-income countries that have limited capital and human resources. The translation and standardization of QUS techniques into the clinic will benefit from clinical and scientific organizations such as the Quantitative Imaging Network of the National Cancer Institute or the Quantitative Imaging Biomarker Alliance of the Radiological Society of North America. The prospective use of well-developed imaging protocols will homogenize databases and make ultrasound-based radiomics more reproducible and accessible.

Conclusion

AI and machine learning, including deep learning, are poised to play an important role in managing cancer diagnosis, treatment, and follow-up. Developing such systems will lead to increasing presence of digital imaging and electronic medical information systems in oncology clinics. The shift to these digital platforms will allow immense data sets

to be created and mined to extract new information that can be used for clinical decision-making. Challenges include robust data provenance, optimization, and tuning parameters for machine learning and deep learning algorithms, standardizing imaging acquisition and its parameters, and aligning ground truth labels with well-defined clinical endpoints. Taken together, it is expected that exploiting AI in oncology will be commonplace as technology evolves, with significant potential for data-driven diagnostic and treatment strategies for cancer in the future.

Acknowledgment

The authors thank Dr. Calvin Law, Dr. Eileen Rakovitch, Ms. Jan Stewart, and Mr. Steve Russell for research support. This work is funded, in part, by the Terry Fox Research Institute, Canada, Natural Sciences and Engineering Research Council of Canada, Canada, the Kavelman-Fonn Foundation, and the Women's Health Golf Classic Fund.

References

- [1] Moore, G. E. (2006). Cramping more components onto integrated circuits, Reprinted Electronics, volume 38, number 8, April 19, 1965, pp.114 ff. *IEEE Solid State Circuits Soc Newsl* 11(3), 33–35.
- [2] Wang, S., & Summers, R. M. (2012). Machine learning and radiology. *Med Image Anal* 16(5), 933–951.
- [3] McBee, M. P., Awan, O. A., & Colucci, A. T., et al. (2018). Deep learning in radiology. *Acad Radiol* 25(11), 1472–1480.
- [4] Gillies, R. J., Kinahan, P. E., & Hricak, H. (2016). Radiomics: images are more than pictures, they are data. *Radiology* 278(2), 563–577.
- [5] Lambin, P., Rios-Velazquez, E., & Leijenaar, R., et al. (2012). Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 48(4), 441–446.
- [6] O'Connor, J. P., Aboagye, E. O., & Adams, J. E., et al. (2017). Imaging biomarker roadmap for cancer studies. *Nat Rev Clin Oncol* 14, 169–186.
- [7] Tran, W. T., Childs, C., Probst, H., Farhat, G., & Czarnota, G. J. (2018). Imaging biomarkers for precision medicine in locally advanced breast cancer. *J Med Imaging Radiat Sci* 49(4), 342–351.
- [8] Haralick, R. M. (1979). Statistical and structural approaches to texture. *Proc IEEE* 67(5), 786–804.
- [9] Yip, S. S. F., & Aerts, H. J. W. L. (2016). Applications and limitations of radiomics. *Phys Med Biol* 61(13), R150–R166.
- [10] Marchio, C., & Sapino, A. (2011). The pathologic complete response open question in primary therapy. *J Natl Cancer Inst Monogr* 2011(43), 86–90.
- [11] Symmans, W. F., Peintinger, F., & Hatzis, C., et al. (2007). Measurement of residual breast cancer burden to predict survival after neoadjuvant chemotherapy. *J Clin Oncol* 25(28), 4414–4422.
- [12] Dudani, S. A. (1976). The distance-weighted k-nearest-neighbor rule. *IEEE Trans Syst Man Cybern* 6(4), 325–327.
- [13] Han, J., & Kamber, M. (2006). Data mining: concepts and techniques. 2nd ed. The Morgan Kaufmann series in data management systems (pp. 770). Amsterdam; Boston, San Francisco, CA: Elsevier; Morgan Kaufmann, xxviii.
- [14] Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., & Scholkopf, B. (1998). Support vector machines. *IEEE Intell Syst Appl* 13(4), 18–28.
- [15] Shen, D., Wu, G., & Suk, H.-I. (2017). Deep learning in medical image analysis. *Annu Rev Biomed Eng* 19, 221–248.
- [16] Green, A. R., Soria, D., & Stephen, J., et al. (2016). Nottingham prognostic index plus: validation of a clinical decision making tool in breast cancer in an independent series. *J Pathol Clin Res* 2(1), 32–40.

- [17] Tyler, S., Truong, P. T., & Lesperance, M., et al. (2018). Close margins less than 2 mm are not associated with higher risks of 10-year local recurrence and breast cancer mortality compared with negative margins in women treated with breast-conserving therapy. *Int J Radiat Oncol Biol Phys* 101(3), 661–670.
- [18] Esposito, V. (2004). Analysis of cell cycle regulator proteins in non-small cell lung cancer. *J Clin Pathol* 57(1), 58–63.
- [19] Wang, H., Cruz-Roa, A., & Basavanthally, A., et al. (2014). Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features. *J Med Imaging (Bellingham)* 1(3), 034003.
- [20] Petushi, S., Garcia, F. U., Haber, M. M., Katsinis, C., & Tozeren, A. (2006). Large-scale computations on histology images reveal grade-differentiating parameters for breast cancer. *BMC Med Imaging* 6(1), 14.
- [21] Wang, P., Hu, X., Li, Y., Liu, Q., & Zhu, X. (2016). Automatic cell nuclei segmentation and classification of breast cancer histopathology images. *Signal Process* 122, 1–13.
- [22] Ali, H. R., Dariush, A., & Thomas, J., et al. (2017). Lymphocyte density determined by computational pathology validated as a predictor of response to neoadjuvant chemotherapy in breast cancer: secondary analysis of the ARTemis trial. *Ann Oncol* 28(8), 1832–1835.
- [23] Ali, H. R., Dariush, A., & Provenzano, E., et al. (2016). Computational pathology of pre-treatment biopsies identifies lymphocyte density as a predictor of response to neoadjuvant chemotherapy in breast cancer. *Breast Cancer Res* 18(1), 21.
- [24] Romo-Bucheli, D., Janowczyk, A., Gilmore, H., Romero, E., & Madabhushi, A. (2016). Automated tubule nuclei quantification and correlation with oncotype DX risk categories in ER+ breast cancer whole slide images. *Sci Rep* 6(1), 32706.
- [25] Lewis, J. S., Ali, S., Luo, J., Thorstad, W. L., & Madabhushi, A. (2014). A quantitative histomorphometric classifier (QuHbIC) identifies aggressive versus indolent p16-positive oropharyngeal squamous cell carcinoma. *Am J Surg Pathol* 38(1), 128–137.
- [26] Luo, X., Zang, X., & Yang, L., et al. (2017). Comprehensive computational pathological image analysis predicts lung cancer prognosis. *J Thorac Oncol* 12(3), 501–509.
- [27] Yu, K.-H., Zhang, C., & Berry, G. J., et al. (2016). Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. *Nat Commun* 7(1), 12474.
- [28] Wang, X., Janowczyk, A., & Zhou, Y., et al. (2017). Prediction of recurrence in early stage non-small cell lung cancer using computer extracted nuclear features from digital H&E images. *Sci Rep* 7(1), 13543.
- [29] Sotiriou, C., Neo, S.-Y., & McShane, L. M., et al. (2003). Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proc Natl Acad Sci U S A* 100(18), 10393–10398.
- [30] van 't Veer, L. J., Dai, H., & van de Vijver, M. J., et al. (2002). Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415(6871), 530–536.
- [31] Cortazar, P., Zhang, L., & Untch, M., et al. (2014). Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. *Lancet* 384(9938), 164–172.
- [32] Von Minckwitz, G., Untch, M., & Blohmer, J. U., et al. (2012). Definition and impact of pathologic complete response on prognosis after neoadjuvant chemotherapy in various intrinsic breast cancer subtypes. *J Clin Oncol* 30(15), 1796–1804.
- [33] Onitilo, A. A., Engel, J. M., Greenlee, R. T., & Mukesh, B. N. (2009). Breast cancer subtypes based on ER/PR and Her2 expression: comparison of clinicopathologic features and survival. *Clin Med Res* 7(1–2), 4–13.
- [34] Fisher, B., Anderson, S., & Bryant, J., et al. (2002). Twenty-year follow-up of a randomized trial comparing total mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer. *N Engl J Med* 347(16), 1233–1241.
- [35] Fisher, B., Bryant, J., & Dignam, J. J., et al. (2002). Tamoxifen, radiation therapy, or both for prevention of ipsilateral breast tumor recurrence after lumpectomy in women with invasive breast cancers of one centimeter or less. *J Clin Oncol* 20(20), 4141–4149.
- [36] Hughes, K. S., Schnaper, L. A., & Bellon, J. R., et al. (2013). Lumpectomy plus tamoxifen with or without irradiation in women age 70 years or older with early breast cancer: long-term follow-up of CALGB 9343. *J Clin Oncol* 31(19), 2382–2387.
- [37] Kunkler, I. H., Williams, L. J., Jack, W. J. L., Cameron, D. A., & Dixon, J. M. (2015). Breast-conserving surgery with or without irradiation in women aged 65 years or older with early breast cancer (PRIME II): a randomised controlled trial. *Lancet Oncol* 16(3), 266–273.
- [38] Paik, S., Shak, S., & Tang, G., et al. (2004). A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 351(27), 2817–2826.
- [39] Stetson, L. C., Pearl, T., Chen, Y., & Barnholtz-Sloan, J. S. (2014). Computational identification of multi-omic correlates of anticancer therapeutic response. *BMC Genomics* 15(Suppl 7), S2.
- [40] Pinker, K., Chin, J., Melsaether, A. N., Morris, E. A., & Moy, L. (2018). Precision medicine and radiogenomics in breast cancer: new approaches toward diagnosis and treatment. *Radiology* 287(3), 732–747.
- [41] Saha, A., Harowicz, M. R., & Grimm, L. J., et al. (2018). A machine learning approach to radiogenomics of breast cancer: a study of 922 subjects and 529 DCE-MRI features. *Br J Cancer* 119(4), 508–516.
- [42] Tadayyon, H., Sannachi, L., & Gangeh, M., et al. (2016). Quantitative ultrasound assessment of breast tumor response to chemotherapy using a multi-parameter approach. *Oncotarget* 7(29), 45094–45111.
- [43] Nasief, H. G., Rosado-Mendez, I. M., Zagzebski, J. A., & Hall, T. J. (2019). A quantitative ultrasound-based multi-parameter classifier for breast masses. *Ultrasound Med Biol* 45(7), 1603–1616.
- [44] Oelze, M. L., & Mamou, J. (2016). Review of quantitative ultrasound: envelope statistics and backscatter coefficient imaging and contributions to diagnostic ultrasound. *IEEE Trans Ultrason Ferroelectr Freq Control* 63(2), 336–351.