

Artificial intelligence methods available for cancer research

Ankita Murmu^{1,2,3}, Balázs Györfly (✉)^{1,3,4}

¹Institute of Molecular Life Sciences, HUN-REN Research Centre for Natural Sciences, Budapest 1117, Hungary; ²National Laboratory for Drug Research and Development, Budapest 1117, Hungary; ³Department of Bioinformatics, Semmelweis University, Budapest 1094, Hungary; ⁴Department of Biophysics, University of Pecs, Pecs 7624, Hungary

© The Author(s) 2024. This article is published with open access at link.springer.com and journal.hep.com.cn

Abstract Cancer is a heterogeneous and multifaceted disease with a significant global footprint. Despite substantial technological advancements for battling cancer, early diagnosis and selection of effective treatment remains a challenge. With the convenience of large-scale datasets including multiple levels of data, new bioinformatic tools are needed to transform this wealth of information into clinically useful decision-support tools. In this field, artificial intelligence (AI) technologies with their highly diverse applications are rapidly gaining ground. Machine learning methods, such as Bayesian networks, support vector machines, decision trees, random forests, gradient boosting, and K-nearest neighbors, including neural network models like deep learning, have proven valuable in predictive, prognostic, and diagnostic studies. Researchers have recently employed large language models to tackle new dimensions of problems. However, leveraging the opportunity to utilize AI in clinical settings will require surpassing significant obstacles—a major issue is the lack of use of the available reporting guidelines obstructing the reproducibility of published studies. In this review, we discuss the applications of AI methods and explore their benefits and limitations. We summarize the available guidelines for AI in healthcare and highlight the potential role and impact of AI models on future directions in cancer research.

Keywords machine learning; artificial neural network; deep learning; natural language processing; prediction; guideline; diagnosis

Introduction

In recent years, significant progress has been made in comprehending the molecular and cellular mechanisms underlying tumor progression. However, several significant difficulties pause further progress. First, traditional imaging techniques such as magnetic resonance imaging (MRI) [1], computer tomography (CT) [2], and mammography [3], have been used for a long time as an approach to cancer screening. Obtaining clinical solutions from the data generated from these techniques is problematic as they require curation by trained professionals, which can be time-consuming. Second, cancer can be linked to changes in genes, and these can in turn be utilized as clinically relevant diagnostic [4], prognostic [5], and predictive biomarkers [6]. Unfortunately, most of the potential biomarker

candidates could not be translated into clinical practice due to variations in the presence of cancer metastasis, different treatment response rates of patients, and acquired resistance. Third, although new therapeutic strategies such as targeted and immune-based therapies have emerged as efficient options for combating cancer [7], the heterogeneity of cancer causes variation in the response rate of patients to anticancer drugs.

New artificial intelligence (AI) based methods promise solutions to some of these challenges. In recent years, AI models have been extensively utilized in drug development [8] and in cancer prediction and diagnosis [9]. Efforts of cancer researchers have resulted in several repositories containing cancer-related data that can be analyzed and integrated with AI approaches for different applications [10]. Among these, AI methods have become increasingly important in the evaluation of the diverse and complex data generated by next-generation sequencing. AI-based algorithms can identify genetic mutations [11] or gene signatures [12] that can aid in the early detection of cancer and the development of targeted cancer therapies. Handling and processing large volumes

Received January 3, 2024; accepted May 17, 2024

Correspondence: Balázs Györfly,

gyorffy.balazs@yahoo.com,

gyorffy.balazs@med.semmelweis-univ.hu

of these data requires augmentation of cloud computing and storage power, for which integrating AI systems can achieve state-of-the-art performance. AI may also directly assist oncologists at the bedside by providing estimations of clinical outcomes. Developing accurate AI models and implementing them in clinical settings remains challenging primarily because of the limitations of using heterogeneous data sets, biases in outcomes, and data privacy [13]. Furthermore, ethical, legal, and social considerations also play a role. Regardless, AI methods have demonstrated robustness, leading to improved clinical decision-making. Overall, in the past few decades, several AI methods were proposed which were utilized for different applications in cancer research (Fig. 1).

Notably, there is no such method, as “AI.” AI is a collection of methods and techniques which can be used to manipulate and interpret the original data. A major weakness of current oncology research is the sparse reporting of actual methods used, which prevents robust and reproducible research. In this review, we highlight in detail the application of different AI methods used in cancer research, including their advantages and limitations. The overall usage of the methods discussed in

our review in the last ten years is provided in Table 1. We also explore guidelines available on how AI models should be incorporated into clinical settings and how the emerging pre-trained language models can boost the personalization of cancer care strategies.

AI methods

The terms “Artificial Intelligence” and “Machine Learning” were coined in the 1950s [14]. Machine learning (ML) includes two major arms, unsupervised and supervised learning. In unsupervised learning, we look for the inherent structure of the data and it includes dimension reduction (principal component analysis) and clustering. In supervised learning, we assign samples in the training set to classes and teach the model to recognize these using the input data. Supervised learning includes regression and classification—the latter involves a broad set of methods. Traditional ML models such as Bayesian networks, support vector machines, and random forest models continuously incorporate data and produce an outcome. A major set of ML methods is based on neural network algorithms that allow machines to mimic the human brain’s ability. A neural network technique

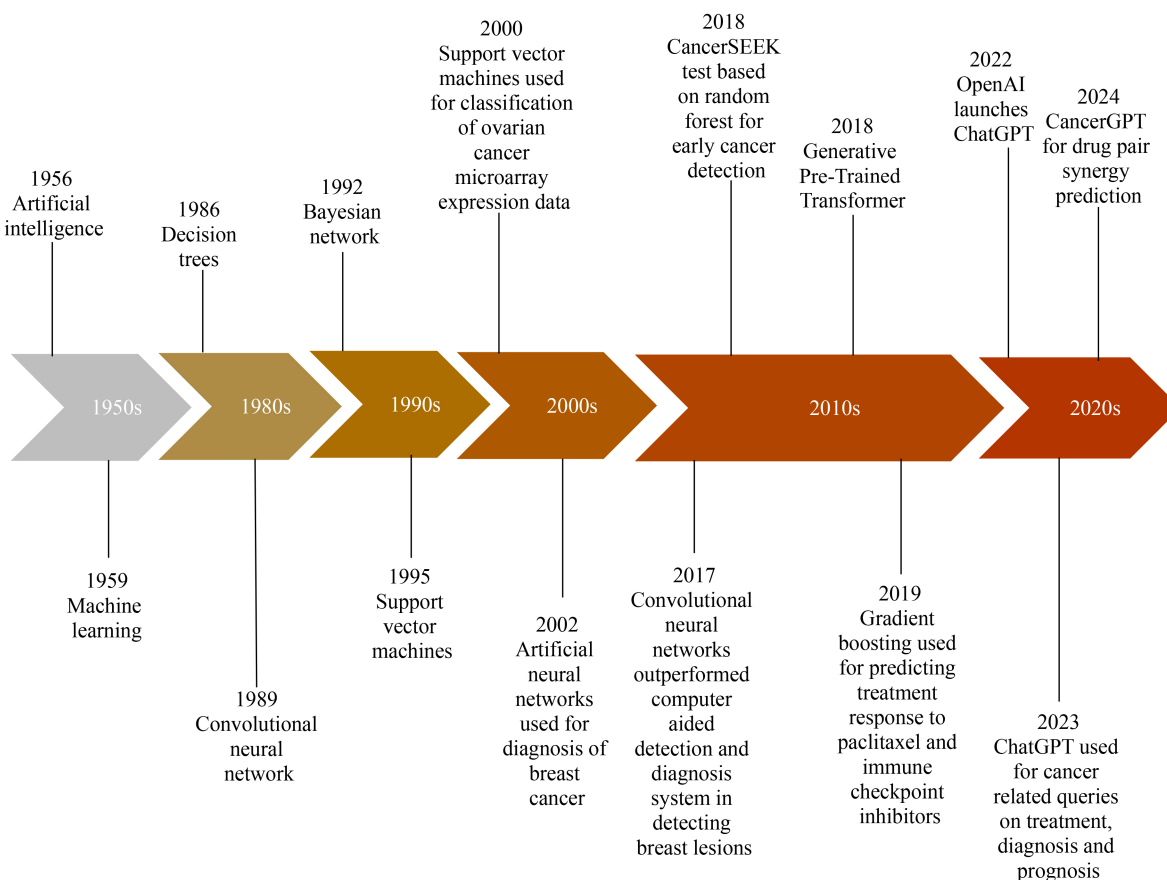


Fig. 1 A brief overview highlighting the timeline when the different AI methods were introduced for the first time. The timeline from 2000 also shows their applications in different areas of cancer research that are discussed in the review.

Table 1 Number of hits in PubMed in the last ten years for the combination of the listed keywords

+ AI method	cancer + diagnosis	cancer + prognosis	cancer + therapy	cancer + pathology	cancer + health records
Random forest	1042	726	618	927	36
Decision trees	539	270	406	430	20
Gradient boosting	211	142	121	148	21
Support vector machines	1335	525	426	939	27
K-nearest neighbors	70	21	25	44	2
Bayesian network	85	67	294	151	2
Artificial neural network	337	120	139	206	6
Deep learning	3446	994	1406	2630	94
Natural language processing	279	58	175	210	262

Search term example: “cancer” AND “diagnosis” AND “random forest.”

with multiple layers gaining popularity in cancer research is deep learning. Deep learning uses various hidden layers, which enhance processing power to explore more complex patterns in the data. Another AI algorithm gaining prominence is natural language processing which targets narrative texts and extracts useful information that can assist in decision-making.

The AI models in cancer research have been developed to utilize commonly used inputs including multi-omics and clinical information obtained from different sources like imaging, laboratory, clinical, and pathological data. The most common task of the ML models is classification, and the general approach used to validate and assess the performance of these models is the receiver operating characteristic analysis which assists in computing area under the curve (AUC), sensitivity, specificity, and precision [15]. Almost all ML algorithms use supervised learning for classification tasks based on conditional probabilities.

Below, we will first discuss traditional machine-learning tools including decision tree-based methods, support vector machines, Bayesian networks, and K-nearest neighbors and then we will endeavor to neural networks and large language models for natural language processing tasks. An overview of the described methods is provided in Fig. 2.

Decision tree based AI methods

Decision trees are supervised learning methods used in ML and data analysis. They are tree-like models used for decision-making or predicting the classification of data sets [16]. They are represented as structured graphs with nodes and branches indicating decisions and consequences, respectively. They learn by taking a subset of labeled training data and recursively splitting it until a decision is reached. Additionally, decision trees are recognized as one of the prominent ML algorithms as they are simple, easy to discern and quick to learn from the data [17].

Decision trees have shown potential in prognostic decision making. A novel decision tree molecular classifier identified molecular subgroups based on presence or absence of mutation or protein in patients with endometrial cancer [18]. They found that the patients who had polymerase- ϵ exonuclease domain mutations had the most favorable prognosis while the patients with p53 null/missense mutations had the worst prognosis.

Researchers in the past have tested and compared the accuracy of conventional methods like MRI, positron emission tomography-CT, and sentinel node biopsy in detecting lymph node status, an important characteristic for selecting appropriate treatment for patients [19]. But, with improvements in image analysis using radiomics [20], a model of clinical factors combined with a decision tree model achieved the best diagnostic performance with an AUC of 0.84 in the validation cohort for predicting lymph node metastasis in patients with cervical cancer [21]. Radiomics generates massive amounts of high-dimensional data and ML models can reduce this dimensionality and identify relevant features. So, researchers have linked radiomics with genomic features of tumors to predict mutations in lung cancer [22] and copy number variations in glioblastoma [23]. These show the potential to develop image-based biomarkers that can improve diagnostic accuracy and treatment selection. Besides the radiomics model, when a decision tree model was combined with multiparametric MRI features, the model demonstrated low performance in predicting pathological complete response, disease-specific survival, and recurrence-free survival in patients with breast cancer [24].

Efforts have been made by researchers to improve the accuracy and interpretation of diagnostic models. For example, a stacking-based decision tree ensemble learning method was proposed for detecting prostate cancer [25]. The learning process of this method involved base-level learning using regression trees, model selection, and stacking, as well as extracting decision rules from regression trees. However, the authors

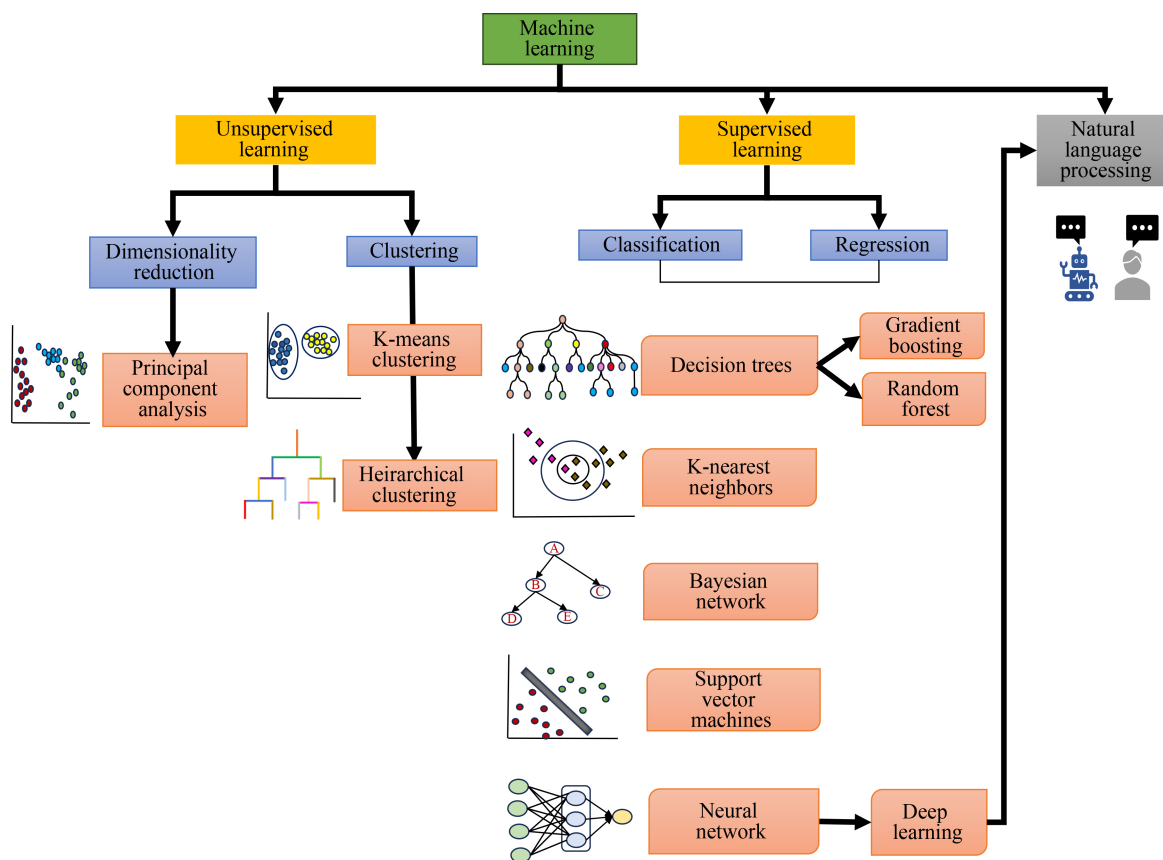


Fig. 2 Classification of AI methods discussed in the manuscript. Machine learning, the sub-domain of AI is divided into unsupervised and supervised learning methods. The unsupervised learning methods consist of dimensionality reduction and clustering algorithms. Supervised learning methods consist of classification and regression algorithms which can also be applied to natural language processing tasks.

observed longer training time compared to single classifiers and other ensemble methods. In another study the performance of an ensemble method called Extra Trees was evaluated [26]. The model achieved high diagnostic accuracy in classifying breast cancer types from the Wisconsin Breast Cancer Database. However, it was reported that Extra Trees are black box models and may be inefficient when using a large number of trees. Moreover, decision trees may also overfit the data, decreasing the performance of the algorithm [27].

Despite these limitations, decision tree based learning methods, such as random forest and gradient boosting, have been developed to further enhance predictive accuracy.

Random forest

Growing ensemble of trees has gained significant attention in cancer research due to its improvements in classification accuracy. The random forest algorithm uses several individual decision trees and ensemble learning techniques to produce a single output [28]. The random forest classifier repeatedly selects a random subset of features to train and generate many decision trees.

Finally, the class selected by most trees is considered as the output.

Random forest models have been employed for the analysis of cell line data to predict drug resistance [29] and for early cancer detection in patients using blood-based assays. For instance, a multi-analyte blood-based test called CancerSEEK uses assessments of genetic alterations and abundance of protein biomarkers to identify early malignant lesions [30]. Identifying stratification factors that define patient characteristics enabling early detection is crucial. In this regard, ML models could be valuable in recognizing complex patterns of different biomarkers to improve diagnostic accuracy. For instance, a random forest model utilizing DNA methylation biomarkers achieved an AUC of 0.95 for discriminating between patients with less or more aggressive prostate cancer and proved to be an independent predictor of recurrence free survival [31]. A different study focused on multiple peripheral blood biomarkers and patients' clinicopathological features developed a random forest model that achieved an AUC of 0.96 for distinguishing epithelial ovarian cancer from benign ovarian tumors, surpassing other ML models [32].

Extending to prognostic studies, random survival forest

models were developed to predict overall survival in colorectal cancer [33] and cancer specific survival in pancreatic cancer [34]. A random forest model based on The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus data sets of ovarian cancer patients, identified 17 metabolic pathways associated with prognosis [35]. Also, a random forest model integrating radiomics features demonstrated potential in predicting prognostic factors associated with breast cancer [36]. Further focus has also been given to treatment outcome prediction. A random forest classifier was trained in rectal cancer patients to discriminate their treatment outcomes and reached an AUC of 93% [37]. Another random forest classifier outperformed several machine algorithms in discriminating (chemo)radiotherapy outcome in a multi-cancer patient cohort [38].

Improved versions of random forest models have also been explored. A tumor feature selection strategy using a random forest based on a genetic algorithm showed good performance in predicting the clinical outcomes of patients with esophageal cancer [39]. The model demonstrated high classification accuracy for prediction and prognosis tasks, with AUC values of 0.82 and 0.80, respectively. For the diagnosis of breast cancer, an interpretable random forest model based on rule extraction methods was proposed for making better patient-centric decisions [40]. The performance metrics demonstrated that the proposed method outperforms other black box models and rule extraction methods in terms of accuracy.

In the past, a quantitative structure-activity relationship-based random forest prediction model has been developed to assess the structure-activity relationship of a large set of compounds for discriminating the epidermal growth factor (EGFR) inhibitors which is an important drug target in cancer [41]. In more recent studies, random forest models focused on synergistic drug combinations for accelerating drug discovery processes. In this regard, drug target, gene expression profiles [42], and large-scale phenotypic drug combinations data sets [43] were useful for building random forest models to predict synergistic drug combinations.

Based on these applications, it can be noted that random forest models have been widely employed from a prediction perspective, can show satisfactory performances, and achieve high accuracy and AUC for classification tasks. They are best suited for the analysis of many variables in relatively medium-sized data sets and could also address the challenge of limited sample size [44]. Random forest models are alternatives to conventional statistical methods that are not suitable to extract information with multiple input variables and to identify factors identifying cancer patients. However, their applicability in the early detection of cancer may still be limited because of the challenges associated with interpreting the model for making clinical decisions [45].

Gradient boosting

The boosting methods are based on a constructive strategy adding new features to the ensemble formation sequentially [46]. Among them, the gradient boosting algorithm is a novel and complex cooperative learning method that produces a strong prediction model or decision tree by using an ensemble of weak models [47]. One of the commonly used gradient boosting frameworks is Extreme Gradient Boosting (XGBoost), which uses decision trees as weak learners and outputs the sum of the predictions of all the individual trees [48]. The XGBoost algorithm also employs random seeds to improve its performance in repetitive tasks.

In oncology, the Cox proportional hazards model has been the standard prognostic model for survival outcomes but it does not capture nonlinearities of features and can be inadequate when using high dimensional data [49]. ML models like gradient boosting have demonstrated significant potential in predicting survival outcomes among cancer patients. For example, XGBoost successfully predicted five-year survival in non-metastatic breast cancer patient data obtained from the Netherlands Cancer Registry and showed comparable performance with the classical Cox proportional hazards regression model [50]. In addition, it was noted that unlike the Cox model, which did not assume any relationship between the features, XGBoost model effectively captured complex interaction between the features and modeled the nonlinearities. Additionally, studies on breast cancer [24] and clear cell renal cell carcinoma [51] patients found that the XGBoost models demonstrated better performance and higher accuracy compared to other ML methods when predicting survival outcomes.

The XGBoost algorithm can be a cost-effective solution for finding clinically relevant biomarkers for targeted therapies by detecting key mutations in clinical images. Clinically relevant biomarkers such as *EGFR* and *KRAS* are frequently mutated or altered in cancer, especially in patients with non-small-cell lung cancer [52]. In a recent study, an XGBoost model using radiomics features showed robust performance in detecting *EGFR* and *KRAS* mutations with 83% and 86% accuracy respectively from the Cancer Imaging Archive data of non-small-cell lung cancer patients [53]. Additionally, for boosting the survival chances in breast cancer, prediction of metastasis and recurrence were explored. For instance, one study showed that the XGBoost model identified 6 gene signatures among which *SQSTM1* was found to regulate metastasis in breast cancer [54] while another study showed that color features play an important role in detecting breast cancer metastasis and recurrence using XGBoost [55].

New developments in the treatment of cancer have led

to improved clinical outcomes where publicly available data from the TCGA cohort has been used to predict treatment response of breast cancer patients to paclitaxel [56] and immune checkpoint inhibitors [57] using an XGBoost model. Furthermore, XGBoost models could predict chemoradiation response for patients with esophageal cancer [58] as well as patient-reported outcomes at the 1-year follow-up of surgery for patients with breast cancer [59]. Another XGBoost model showed significantly better performance in distinguishing between cancer types such as early and late stages of renal clear cell carcinoma, renal papillary cell carcinoma, lung squamous cell carcinoma, and head and neck squamous cell carcinoma with DNA methylation data [60]. Hence, XGBoost models could be implemented for patient-centric decisions and promote targeted and effective treatment.

While gradient boosting algorithms such as XGBoost are significantly advancing cancer research, their main advantage lies when handling data with missing or incomplete values primarily without imputation [58]. Its ability to classify and utilize the data without imputation makes it a valuable tool for researchers. Moreover, how the gradient boosting model makes predictions is easy to decipher, as it has decision trees as the base learner [46].

Other supervised learning AI methods

Support vector machines

The concept of support vector machine (SVM) is different from decision tree-based methods. SVM is a supervised ML method that uses a hyperplane or decision boundaries for classification problems [61]. The algorithm uses kernel functions that make the computation faster; therefore, the choice of kernel functions influences the performance of the model [62]. In addition, they are suitable for nonlinear classifications as well.

SVM-based classifiers have been widely used in cancer research since the advent of high-throughput microarray gene expression. A very early application of SVMs in this area focused on the classification of cancerous ovarian tissue, normal ovarian tissue, and normal non-ovarian tissues [63]. In the later years, SVM applications expanded to include gene expression and copy number variation features for predicting breast cancer patients' response to chemotherapeutic agents like paclitaxel and gemcitabine using an online platform [64].

More recent applications of SVM include integrating radiomics features into SVM models. Predictive or prognostic modeling in radiomics can be useful to improve decision support in oncology. For example, a study used kernel SVM classifier with MRI radiomics features to predict local and distant failure in patients with

advanced nasopharyngeal carcinoma which could be useful for making decisions regarding treatment plans [65]. Another study constructed a radiomics signature consisting of 30 selected features using linear kernel SVM to distinguish whether patients with rectal cancer received a pathological complete response to neoadjuvant chemotherapy [66]. Additionally, an interesting study on colorectal cancer revealed that CT radiomics signature is highly correlated with *KRAS/NRAS/BRAF* mutation status [67]. These applications suggest that SVM holds high predictive or prognostic potential which could enhance the applications of non-invasive and cost-effective techniques like radiomics.

Identifying biomarkers is a crucial approach in cancer diagnosis. These biomarkers can serve as features for ML models to classify healthy and diseased samples [68]. For instance, an SVM classifier using integrated extracellular vesicle long RNA markers demonstrated high sensitivity and specificity in classifying hepatocellular carcinoma patients and healthy controls [69], while DNA methylation-based biomarkers were associated with recurrence in early-stage hepatocellular carcinoma [70]. In gastric cancer research, SVM models have been utilized to predict survival of patients. By incorporating immunomarkers and clinicopathologic features, a prognostic SVM classifier was developed to predict overall survival and disease-free survival in gastric cancer patients and identify the benefits of postoperative adjuvant chemotherapy in stage II and stage III patients [71]. Additionally, an SVM model based on 32-gene signature specific to gastric cancer generated risk scores that were prognostic of overall survival and response to treatments [72]. The stability of biomarker selection while developing these models is necessary for reproducibility of the classification so that the prediction model shows similar performance when classifying new samples.

Another important factor in diagnostic applications involves cancer staging and grading systems. For example, the Fuhrman nuclear grading system is used to assess the tumor aggressiveness in renal cell carcinoma, impacting clinical treatment selection [73]. Likewise, for cervical cancer, clinical staging is recommended by the International Federation of Gynecology and Obstetrics based on imaging and pathological findings [74]. In addition to the staging and grading systems, ML methods could play a vital role in rendering clinical decisions. An SVM classifier has demonstrated promising results in predicting high and low Fuhrman nuclear grades from CT texture features in clear cell renal cell carcinoma [75]. The model's performance was comparable to percutaneous biopsy—an invasive method available for Fuhrman nuclear grading. For prognosis prediction, high-risk surgical-pathological factors, including the International Federation of Gynecology and Obstetrics staging, were

used to investigate the accuracy of the SVM model in early-stage cervical cancer patients after surgery [76], showing that these factors could predict the recurrence with an accuracy of 69%. Another study utilized gene signatures to distinguish colon cancer patients at high risk of recurrence from those at low risk [77].

It is worth noting from these applications that SVM models show versatile performance when using a wide variety of biological data, such as multi-omics, imaging, and clinical data, for cancer diagnosis and prediction. SVM models have also paved their way toward drug discovery by outperforming other ML methods in predicting the inhibition of breast cancer resistance protein [78]. Nonetheless, challenges still lie in making a good SVM classification model. Therefore, to improve the classification accuracy, several researchers proposed different approaches. For example, the feature clustering recursive feature elimination [79] could be a suitable approach to reduce computational complexity, redundancy among genes, and increase classification accuracy. On the other hand, a weighted AUC ensemble learning based on SVM [80] could significantly increase accuracy of breast cancer diagnosis. Finally, the performance of SVM relies heavily on the choice of kernel functions, and the expertise of the user.

Bayesian network

Another useful ML model for classification is the Bayesian network classifiers which produce probabilistic estimation of the variables [81]. They are represented as a directed acyclic graph that explains the relationships or dependencies amid the random variables based on inferences. Among the Bayesian networks, the Naïve Bayes classifier is the most effective [82].

Naïve Bayes has been successful in lung cancer [83] and rectal cancer [84] prognosis which involved prediction of patient survival. Additionally, it was also used for the diagnosis of diffuse large B cell lymphoma genetic subtypes based on mutation, copy number variation, and BCL2 or BCL6 rearrangement data, providing the likelihood that a patient's lymphoma belongs to one of the six defined genetic subtypes [85].

Hospitals generate a vast amount of data, and conducting research on these data can be challenging due to ethical, legal, and administrative issues. To address this issue, Jochems *et al.* used a distributed learning approach to train a Bayesian Network model on clinical data of patients with lung cancer treated with chemoradiation or radiotherapy at five hospitals [86]. The model underwent external validation with hospital data not included in the training set and achieved an AUC ranging from 0.59 to 0.71. Targeted therapy, especially for breast cancer treatment, is a crucial focus in oncology. The ability to predict pathological complete response to neoadjuvant

chemotherapy is a significant advancement for improving patient outcomes. The development of Naïve Bayes model based on radiomic features represented a sophisticated approach for predicting pathological complete response to neoadjuvant therapy [87]. The model demonstrated high performance and achieved an AUC of 0.93 for triple-negative and in human epidermal growth factor (HER2) positive patients. Similarly, another Naïve Bayes prediction model exhibited significant positive correlation with pathological complete response to neoadjuvant chemotherapy [88]. Hence, Bayesian Network models could enable targeted administration of neoadjuvant therapy and prevent delay of the clinically effective treatment for breast cancer patients.

Furthermore, to evaluate the performance, studies have compared Bayesian network models with other ML models. In one study, the accuracy of a Naïve Bayes classifier was assessed against other ML classifiers for classifying benign and malignant breast tumors [89]. In another study, a Bernoulli Naïve Bayes algorithm was compared with traditional ML methods for predicting the binding of estrogen receptors [90]. However, in these studies, Naïve Bayes models did not show satisfactory performance in terms of accuracy and AUC compared to other ML methods.

A small body of literature suggests that modifying Naïve Bayes classifiers could improve classification accuracy for breast cancer detection. For instance, a weighted Naïve Bayes classifier, proposed for the detection of breast cancer, achieved an accuracy of 98.5% when trained and tested on attributes from the Wisconsin Breast Cancer Database [91]. In another study, authors proposed a two-layer ensemble hybrid classifier for detecting malignant and benign tumors, indicating potential to enhance classification accuracy of traditional Naïve Bayes classifier in breast cancer detection [92]. These studies are limited to breast cancer data but could be expanded to other cancer types.

Naïve Bayes classifiers assume conditional independence of features without considering their relationship which can lead to poor performance with subjective observations [82]. In this regard, developing a Naïve Bayes classifier by selecting a subset of attributes may improve accuracy [93,94].

K-nearest neighbors

One of the simplest yet popular ML algorithms, the K-nearest neighbors (kNN), relies on the distance and assumes that a similar 'k' number of data points are close to each other [95]. It selects the class with the highest probability based on the likelihood of the test data belonging to the 'k' training data. This classification method does not require any prior knowledge about the distribution of the data [96].

Several notable studies of kNN have used imaging data from mammograms [97] and breast ultrasound image segments [98] for breast cancer classification. Based on the texture features of the lesions obtained from MRI, a kNN model classified breast cancer subtype images with a ROC AUC value of 0.81 [99]. In a similar study [100], radiomics features were extracted from contrast-enhanced MRI for classifying breast cancer receptor status and molecular subtypes. In a study by García-Laencina *et al.*, four ML methods including kNN were used to predict five-year survival of breast cancer patients with incomplete clinical data and accuracy of more than 81% and an AUC of more than 0.78 without any imputation [101].

Due to the utilization of a random seed at the kNN commencement, repeated runs can produce different outcomes. This is probably the major limitation and could be the reason why kNN models demonstrated low accuracy compared to other ML techniques in applications on different cancer types, such as lung cancer [102] and brain tumors [103]. To address this issue and improve performance, researchers combined a wrapper-based feature selection method with a kNN classifier which could be suitable for microarray or RNA-Seq data that has thousands of features [104]. However, common drawbacks of wrapper-based methods for feature selection are that they are prone to overfitting and can be computationally intensive [105]. To address this, other researchers [106] followed a different approach by proposing a combination of particle swarm optimization methods along with adaptive kNN for gene selection from microarray data. A study by Zhang *et al.* also proposed three methods for finding optimal 'k' values for efficient classification of test or new data [107]. Furthermore, as building classical kNN models could be time-consuming when dealing with large data sets, some fast versions of the kNN algorithm [108] have been developed by researchers for disease prediction.

Neural network algorithms

A subset of ML, called the neural networks, have been a great deal of excitement among the scientific audience. The concept of artificial neural network (ANN) emerged from the way neurons in the brain work [109]. They use hidden layers connecting each node or artificial neurons to generate output from the input variables. These hidden layers are connected in a hierarchical manner like the organization of neurons in the brain. The strength of the neural connections of an ANN depends on an optimization technique called back-propagation. ANN models may overfit the data and show poor generalization capability when several neurons are allowed in the hidden layers.

In the early 2000s, ANNs found widespread use in

diagnostics [110,111] and prognostic outcome prediction [112,113] applications. In the following years, evolved neural network approaches like particle swarm-optimized wavelet neural networks [114] and genetically optimized neural networks [115] were proposed for the detection and diagnosis of breast cancer. ANNs have also been employed with one or two hidden layers for diagnosis and prediction of breast cancer [116] and pancreatic cancer [117]. Over the years, many important cancer-related applications which attracted attention have been based on the concept of deep neural networks which is discussed in the next section.

Deep learning

Neural networks consisting of more than one hidden layer are termed "deep." The fundamental architecture of deep learning (DL) is based on deep neural networks consisting of multilayered interconnected nodes or artificial neurons for categorization [118]. This neural network architecture uses nonlinear functions like the rectified linear unit (ReLU), to pass the result of the weighted sum of inputs from the previous layer to the next layer.

Convolutional neural network

Convolutional neural network (CNN) [119], a type of deep feedforward neural network that has become dominant in various computer vision tasks, is attracting interest across a variety of domains, including radiology. CNN is designed to learn spatial hierarchies of features automatically and adaptively through backpropagation by using multiple building blocks, such as convolution layers, pooling layers, and fully connected layers. A review article by Yamashita *et al.* offers a perspective on the basic concepts of CNN and its application to various radiological tasks [120]. The review also highlights that the two challenges in applying CNN to radiological tasks are small data sets and overfitting which can be overcome by training more data. Being familiar with the concepts and advantages, as well as limitations of CNN, is essential to leverage its potential in diagnostic radiology, to augment the performance of radiologists and improve patient care.

Deep CNN is capable of showing excellent performance in supervised learning for image classifications [121]. Recent studies have shown that DL models have revolutionized the analysis of medical images in oncology by learning representative features from raw input like tumor tissue images. A study reported that a DL-based classifier could extract more prognostic information, like survival, from the tumor tissue of colorectal cancer than experienced pathologists [122]. CNN models based on hematoxylin and eosin-stained

tumor sections of colorectal cancer patients were used for survival prediction outcome [123,124] and identification of molecular subtypes associated with prognosis [125]. Routine histological images, when analyzed with DL models, can provide useful information for directly identifying genetic mutations, tumor molecular subtypes, gene expression signatures, and pathological biomarkers of potential clinical relevance [126]. Furthermore, DL model based on histological images could reduce the time-consuming diagnostic process of predicting recurrence and metastasis in HER2-positive breast cancer patients [127]. Besides histological images, CNN based on time series CT images were used to address the challenge of capturing the evolving phenotype of tumors and predicted several clinical endpoints from patients with locally advanced non-small cell lung cancer to improve clinical outcomes [128].

Computer aided detection and diagnosis systems have been used for a long time to help radiologists to analyze mammogram screenings [129]. However, since the inception of AI technologies, DL models could be used to reduce human bias by learning directly from training data. CNN outperformed the conventional computer aided detection and diagnosis system in detecting solid and malignant lesions and showed best AUC of 0.90 on the validation set [130]. A computer aided diagnosis system based on Faster R-CNN [131] was used to detect and classify malignant or benign lesions on mammogram. The system showed the highest AUC of 0.95 on the INbreast [132] data sets. Using similar INbreast data sets, another CNN model achieved AUC of 0.95 per image [133].

Furthermore, CNN can be a powerful algorithm that may surpass human experts when applied to classification tasks. For instance, a CNN trained on histopathological images of melanoma and nevi showed good concordance with the pathologists [134]. In lung cancer application, a CNN model outperformed six radiologists in predicting risk of lung cancer from CT imaging data [135]. A DL model even surpassed radiologists in identifying breast cancer from cancer screening program mammograms in the USA and the UK and demonstrated improvement in absolute specificity (1.2%–5.7%) and absolute sensitivity (2.7%–9.4%) [136]. However, these studies aimed to illustrate the potential of DL models rather than replace human experts. Instead, using a blend of DL methods and radiologists could facilitate better interpretation in mammogram screening [137,138].

In cancer subtype classification, renal cell carcinoma subtypes were distinguished using a CNN from histopathological images [139]. The authors showed that training CNN in whole-slide images of renal cell carcinoma achieved very high accuracy in distinguishing tumors from normal tissue. Another CNN model discriminated four molecular subtypes: canonical luminal, immunogenic, proliferative, and receptor tyrosine

kinase-driven from hormone receptor positive/HER2 negative breast cancer patients, based on which precise treatments were proposed [140]. A more recent CNN based study followed an integrative approach by using gene expression and methylation data of glioma patients to classify subtypes into low grade gliomas and glioblastoma multiforme [141]. These developments have the potential to improve treatment selection by enabling more tailored and effective approaches based on the specific molecular subtypes identified.

Focusing on predicting gene mutations, a CNN model identified several significant mutated genes related to prognosis from histopathology images of hepatocellular carcinoma with AUC values between 0.71 and 0.89 [142]. The study also reported that the performance of DL classifiers was nearly comparable to that of a 5-year experienced pathologist for tumor classification and differentiation. In another analysis, mutations of six key genes, including *STK11*, *EGFR*, *FAT1*, *SETBP1*, *KRAS*, and *TP53*, were predicted from pathology images of lung adenocarcinoma, with AUCs ranging from 0.73 to 0.86 [143]. These findings support the idea that DL models will effectively assist pathologists in the detection of cancer mutations.

One of the major challenges in cancer treatment is investigating the effect of potential therapeutic agents. Several studies show that DL is an important approach to consider for predicting drug responses in cancer patients. For example, a multi-omics late integration method based on deep neural networks was developed to predict drug response using somatic mutation, copy number aberration, and gene expression data as input [144]. In another study, a deep neural network was trained on gene expression and drug response data of cancer cell lines from the Genomics of Drug Sensitivity in Cancer database to predict drug responses [145]. The model was tested on multiple unseen clinical cohorts where it outperformed other ML algorithms. Recognizing the challenge of interpreting these models, Kuenzi *et al.* developed a better DL model that is interpretable and could be used in clinical settings for predicting drug response and identifying synergistic drug combinations [146]. In the process of drug discovery, identifying a large number of drug combinations from pharmacogenomics databases can be time-consuming and investigating medium or large scales of these data in real-life settings can be challenging. Focusing on these challenges, an improved DL model, Deep-Resp-Forest, based on deep forest architecture, was developed, which could adapt to different scales of data by automatically learning the depth of the forest cascade [147]. Also, the development of publicly available online platforms like DeepSynergy, based on a feed-forward neural network model, has shown advantages in prioritizing and screening anti-cancer drug combination data sets [148].

Recent developments in DL have unlocked valuable insights from electronic health records, addressing the challenge of analyzing these messy data. The incorporation of free-text clinical records from electronic health records has had an outstanding impact on the performance of state-of-the-art DL models in predicting clinical problems and outcomes [149]. In cancer research, DL models have incorporated electronic health records in tasks like predicting the risk of breast cancer [150] and the onset of pancreatic cancer [151]. DeepPatient, a novel unsupervised deep learning method showed great potential in creating a general-purpose set of patient features from raw electronic health records of various cancer patients that may be used for building predictive clinical models [152]. However, despite these retrospective studies, extensive experiments and prospective trials may be required to demonstrate the accurate predictive ability of these models.

Based on the studies described above, DL models can be highly efficient in analyzing cancer imaging data and have shown robust performance in prediction, detection, and classification tasks initiating a surge of interest among pathologists and radiologists. The success behind DL models showing excellent performance on imaging data is their sensitivity to minuscule details and intricate structures using the backpropagation algorithm [118]. Hence, these models can be useful in extracting a large number of features missed by humans to discover underlying disease characteristics and patterns in cancer patients. While the success of DL models holds promise to overcome the challenge of translating drug response research to actual patients, it is important to note that because of their complexity they function as black box models, making it difficult to interpret what features are used for learning and how decisions are made [153]. Moreover, a lack of retraining on large patient cohorts may hinder the performance of DL diagnostic models [126]. Therefore, an interpretable and retrained DL model may show better performance in clinical settings.

Natural language processing

In recent years, extensive research has been conducted on algorithms, such as natural language processing (NLP), which can process human text, speech, or language using computer algorithms [154]. The performance of the NLP system depends on the task and shows better results on the data sets on which they are built [155]. AI techniques like neural networks, SVM, and decision tree-based methods can also be applied to NLP tasks.

In the field of cancer research, manual analysis of clinical data are time-consuming and inaccurate. The need to decrease manual abstraction of information from clinical charts and reports has shifted the focus of researchers to automatic extraction. In this context, early

NLP models were proposed to serve as an alternative and efficient strategy for extracting structured and unstructured clinical texts related to breast cancer patients. For instance, an automated NLP system was developed for the extraction of breast imaging reporting and data system categories from breast radiology reports [156] and to automatically extract unstructured text from mammography and pathology reports [157]. Furthermore, an NLP system identified 92% of breast cancer recurrences from electronic health records with high sensitivity and specificity [158]. Another study focusing on drug repurposing used NLP to extract drug exposure information from unstructured clinical data of cancer patients [159]. DeepPhe software, based on an NLP model, was developed to perform automatic and detailed extraction of phenotypes from the electronic health records of cancer patients [160]. This software produces a summary of the characteristics of cancer-related phenotypes, which is useful for further clinical investigations.

From these applications, it is evident that NLP works best with data that contain structured or unstructured texts, such as electronic health records or clinical notes. They facilitate rapid analysis of unstructured data and reduce human error in clinical settings. In addition, NLP models can support the development of oncology databases which still require manual annotation of free-text clinical data [161]. However, in real-world clinical reports, having uncertainty in parameters like lymph node status, NLP models might display low performance [162]. Therefore, to create a robust model, careful planning and multiple iterations are required, which could be invaluable for extracting important information from medical records.

Large language models for NLP task

In recent years, large language models that are based on transformers [163], have been explored. These models are AI systems that use neural network architectures to generate content by training on massive data sets consisting of words [164]. Large pre-trained language models have shown benefits for NLP tasks [165]. Hence, to improve NLP tasks, models such as the Generative Pre-trained Transformer (GPT) [166] have been explored. There has been a surge in GPT models, which has attracted widespread interest, especially since OpenAI launched the ChatGPT chatbot in 2022.

In 2023, several researchers have explored the performance of ChatGPT in cancer-specific queries. A ChatGPT model was evaluated based on its responses to questions on hepatocellular carcinoma [167]. The model showed comprehensive and correct responses about basic knowledge, treatment lifestyle, and diagnosis. However, the model incorrectly answered questions related to

hepatocellular carcinoma screening. In a different query, ChatGPT responded to questions related to the diagnosis, prognosis, and treatment of lung cancer and pancreatic cancer with a similar quality as Google's feature snippet [168]. For breast tumor management, ChatGPT was evaluated as a clinical decision support tool for 10 patients [169]. Surprisingly, ChatGPT's recommendations for surgery were similar to the tumor board's decision in 7 out of 10 cases.

Although ChatGPT has shown remarkable capabilities, its limitations in the field of cancer have been proven in several studies. The accuracy of ChatGPT on cancer myths and misconceptions was compared with that of the National Cancer Institute information [170]. The overall accuracy of 13 questions was 100% for National Cancer Institute and 96.9% for GPT. Another interesting study shows that ChatGPT provided incorrect treatment recommendations along with correct ones for breast, prostate, and lung cancer [171]. Hence, the accuracy of cancer-related information is still not reliable. Similar study on prostate cancer showed that the accuracy and precision of ChatGPT content was low and the information was not always consistent when compared to a reference source [172]. Hence, ChatGPT does not provide any references, generates multiple answers, and shows incorrect references, which remain major limitations. Therefore, caution must be exercised by clinicians when dealing with ChatGPT's responses or while introducing them into the clinical settings.

The more recent GPT-4 [173], which can process both texts and images as input, has shown some potential. While comparing ChatGPT with the GPT-4 for lung cancer applications, GPT-4 performed better in translating chest CT reports into plain language [174] and extracting phenotypes from free-text CT reports [175]. Hence, GPT-4 can be a promising tool in radiology and may also provide another possible avenue for decision-making and treatment recommendations based on cancer imaging data as well.

AI guidelines

In recent years, there has been growing concern about the risks associated with the use of AI. To ensure unbiased and better formulation and delivery of AI-based studies in medical research, experts have developed or suggested several reporting guidelines (Table 2). These guidelines are based on recommendations from the EQUATOR (Enhancing the Quality and Transparency of Health Research) network, which promotes and develops guidelines to improve the quality of healthcare. The first reporting guidelines, such as the SPIRIT-AI (Standard Protocol Items: Recommendations for Interventional Trials-Artificial Intelligence) and CONSORT-AI (Consolidated Standards of Reporting Trials-Artificial

Intelligence) were developed for AI interventions in clinical trials [176]. SPIRIT-AI guidelines are centered on the use of AI in clinical trial protocols [177], whereas CONSORT-AI guidelines are centered on the use of AI in clinical trial reports [178]. Both guidelines were developed simultaneously and show similarities in terms of reports and protocols. In addition to these guidelines, establishing how the models should be developed and tested is also important for transparency. The MI-CLAIM (Minimum Information about Clinical Artificial Intelligence Modeling) checklist was suggested for reporting AI algorithms in the field of medicine [179]. MI-CLAIM also overlaps with MINIMAR (Minimum Information for Medical AI Reporting) [180], which focuses on guidelines for developing AI algorithms.

Furthermore, to report research that uses AI for diagnostic test accuracy, the STARD-AI (Standards for Reporting of Diagnostic Accuracy Studies-AI) was published [181], addressing the limitations of the previous STARD 2015 [185] for utilizing AI models. Similarly, TRIPOD-AI (Transparent Reporting of a multivariable prediction model of Individual Prognosis Or Diagnosis-AI) and PROBAST-AI (Prediction model Risk Of Bias Assessment Tool-AI) were published for improving the diagnostic, prognostic, and prediction of ML models [182]. A risk bias tool called QUADAS-AI (Quality Assessment of Diagnostic Accuracy Studies-AI) was developed to evaluate the risk of bias and applicability in AI-centered diagnostic accuracy studies [183]. The guidelines are not limited to building models but have also been extended to decision-making systems. AI technologies have been developed to assist decision-making in healthcare, but only a few have been successful and benefited patient care. The DECIDE-AI (Developmental and Exploratory Clinical Investigations of Decision Support Systems driven by Artificial Intelligence) reporting guideline based on a checklist that includes 17 AI-specific and ten generic reporting items was developed to improve clinical decision outcomes during the early stages [184].

The major focus of these reporting guidelines is on building an AI model with complete transparency on its algorithms, architecture, accuracy, performance, and study design. They were established to serve as standard documenting for developers, investigators, data scientists, and clinicians. Apart from the reporting guidelines, several researchers proposed how AI prediction models should be implemented in healthcare settings. Smith *et al.* proposed a "plan, do, study, adjust" approach while deploying AI for patient care [186]. Wiens *et al.* proposed testing the AI system in real-time and checking it with country-specific government regulations before deploying the model to the market [187]. Larson *et al.* proposed strategies that can address the shortcomings of developing and evaluating diagnostic AI algorithms before

Table 2 Overview of described AI guidelines for medical studies

Guideline	Purpose	Utility					Citations (as of Nov. 2023)	Reference
		Diagnosis	Prognosis	Pathology	Decision- making	Clinical studies		
SPIRIT-AI	AI interventions in clinical trial protocols	Yes	No	No	Yes	Yes	82	[177]
CONSORT-AI	AI interventions in clinical trial reports	Yes	No	No	Yes	Yes	160	[178]
MI-CLAIM	Transparent reporting of AI algorithms in medicine	No	No	No	Yes	No	157	[179]
MINIMAR	Minimum information necessary for reporting AI-based studies in medicine	No	No	No	Yes	No	106	[180]
STARD-AI	Diagnostic test accuracy studies using AI	Yes	No	Yes	No	No	58	[181]
TRIPOD-AI	Diagnostic and prognostic prediction studies based on AI	Yes	Yes	No	No	No	182	[182]
PROBAST-AI	Risk bias tool for diagnostic and prognostic prediction studies based on AI	Yes	Yes	No	No	No	182	[182]
QUADAS-AI	Risk bias tool for AI-centered diagnostic test accuracy studies	Yes	No	No	No	No	34	[183]
DECIDE-AI	Early-stage clinical evaluation of decision support systems driven by AI	No	No	No	Yes	Yes	35	[184]

Abbreviations: SPIRIT-AI, Standard Protocol Items: Recommendations for Interventional Trials - Artificial Intelligence; CONSORT-AI, Consolidated Standards of Reporting Trials - Artificial Intelligence; MI-CLAIM, Minimum Information about Clinical Artificial Intelligence Modeling; MINIMAR, Minimum Information for Medical AI Reporting; STARD-AI, Standards for Reporting of Diagnostic Accuracy Studies-AI; TRIPOD-AI, Transparent Reporting of a multivariable prediction model of Individual Prognosis Or Diagnosis-AI; PROBAST-AI, Prediction model Risk Of Bias Assessment Tool-AI; QUADAS-AI, Quality Assessment of Diagnostic Accuracy Studies-AI; DECIDE-AI, Developmental and Exploratory Clinical Investigations of Decision Support Systems driven by Artificial Intelligence.

implementing them in healthcare [188]. They also suggested continuous monitoring and evaluation of the AI system throughout its life cycle. Despite these recommendations, more transparency and guidance are required in terms of software scrutiny, cost-effectiveness, retraining of the data sets, and a checklist of conditions required to make use of an AI system for a particular task. To date, there is no specific guideline available for the use of AI for the study design in the cancer domain.

Future directions

In the rapidly evolving world of technologies, AI holds great potential in cancer research. There is a growing need for AI technologies that not only provides good accuracy but that can be trustworthy and understandable. In this regard, explainable AI [189] has been an emerging trend. Explainable AI helps to understand how a model functions and interpret the predictions generated by the model, accounting for the limitations of the black box AI systems. This ensures transparency and leaves space of improvements. In addition, the experts would have the means to understand the whole decision making process of AI which would ultimately prevent barriers to implement AI into clinical settings.

In the long-term, ML algorithms will have a major role in personalized and targeted therapies. With data explosion, soon every facet of multi-omics data such as transcriptomics, genomics, epigenomics, metabolomics, and proteomics for individual patients would be stored in databases which could be used for therapy selection.

Early adoption of matching drugs to patients based on the multi-omics features will improve the personalized medicine paradigm. Open source AI platforms like MatchMiner [190] have the capability to assist clinicians to match candidates to precision medicine trials based on their genomic profiles. Further, data-driven AI tools will be useful for accelerating clinical trials through linking individual patients to trials which would overcome the current challenge of labor intensive work.

While there has been tremendous amount of research in common cancer types in tissues, like breast and lung, focusing on obtaining data from rare cancer types or tissues will be a major necessity for future research. Recently, a group of researchers developed large language based prediction model, CancerGPT [191] which successfully predicted drug pair synergy in rare cancer tissues with limited data. Their model could help researchers to promptly identify potential targets and biomarkers.

Non-invasive AI tools with high accuracy represent the future for early cancer detection and diagnosis. For instance, DermaSensor [192], an FDA approved device, uses an AI algorithm to analyze spectral data of skin lesions for the detection of skin cancer. However, the development and commercialization of these software and devices will take a long time due to regulatory limitations and clinical lengthy clinical trials. Regardless, with continued research efforts and innovation, AI technologies will revolutionize cancer detection and improve patient outcomes.

Conclusions

Building AI prediction models has been a crucial area of focus in cancer research. In this review, we discussed and summarized how different AI methods have shown remarkable progress in cancer-related applications. The traditional ML methods, particularly, the supervised learning algorithms, have outperformed conventional statistical tests in classification tasks. These traditional ML algorithms have been used widely with multi-omics and clinical data for cancer classification and for diagnosis of cancer, predicting survival of patients, and treatment response.

Moreover, DL models have opened new possibilities for better accuracy than the traditional ML models in prediction tasks. They constitute a more recent approach and have been widely used in several cancer-related applications, specifically with imaging data. The ability of CNN models to provide clinician level interpretation has shown the potential of DL in oncology. Furthermore, recent studies comparing validation metrics of AI methods for feature selection and classification have shown promising results.

Focusing on an emerging AI technology, we also highlighted that pre-trained language models (GPT) could provide useful solutions when prompted with cancer-related queries. These large language models have the power to extract and analyze crucial insights from massive data sets and may have extensive utility in cancer research by extracting data to look for correlations between patients, identifying drug candidates, and assisting in personalized treatment options. Hence, their rapid advancements show that virtual assistants and specialized AI chatbots for oncology will soon become important in clinical settings.

While AI models are making significant advances in cancer research, human judgement remains a crucial part in areas such as patient-centric decision making, validation of predicted drug targets, better interpretation of imaging data, ethical challenges, and using tools like ChatGPT. Therefore, at the end, a blend of AI and human experts may lead to improved diagnosis, prognosis, and treatment outcomes in clinical settings.

Finally, we must note a few major limitations of AI applications. First, choosing the appropriate algorithm can be intricate and depends on various factors like the type and complexity of the data. Second, to integrate AI in clinical settings, detailed application, and explanation and transparency of the algorithms must be achieved. Third, monitoring the quality of the AI tools for robust performance will be important. A detailed discussion will be necessary to establish which AI models and algorithms are acceptable and can provide valuable outcomes for cancer patients. Overall, AI has already significantly impacted cancer research, and addressing the challenges

and validating the AI-generated results can lead the future of oncology research.

Acknowledgements

This work was supported by the National Research, Development, and Innovation Office (PharmaLab, RRF-2.3.1-21-2022-00015 and TKP2021-NVA-15). The manuscript has been edited using a GPT platform to improve grammar. Ankita Murmu and Balázs Györfly acknowledge the support of ELIXIR Hungary. Ankita Murmu is grateful to Tempus Public Foundation (Hungary) for the Stipendium Hungaricum Ph.D. Scholarship.

Compliance with ethics guidelines

Conflict of interest Ankita Murmu and Balázs Györfly declares no potential conflict of interest.

This manuscript is a review article and does not involve a research protocol requiring approval by the relevant institutional review board or ethics committee.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>.

Funding note Open access funding provided by HUN-REN Research Centre for Natural Sciences.

References

1. Morrow M, Waters J, Morris E. MRI for breast cancer screening, diagnosis, and treatment. *Lancet* 2011; 378(9805): 1804–1811
2. Boieselle PM. Computed tomography screening for lung cancer. *JAMA* 2013; 309(11): 1163–1170
3. Gøtzsche PC, Jørgensen KJ. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2013; 2013(6): CD001877
4. Nair M, Sandhu SS, Sharma AK. Cancer molecular markers: a guide to cancer detection and management. *Semin Cancer Biol* 2018; 52(Pt 1): 39–55
5. Györfly B. Discovery and ranking of the most robust prognostic biomarkers in serous ovarian cancer. *Geroscience* 2023; 45(3):

- 1889–1898
6. Kovács SA, Fekete JT, Gyórfy B. Predictive biomarkers of immunotherapy response with pharmacological applications in solid tumors. *Acta Pharmacol Sin* 2023; 44(9): 1879–1889
 7. Seebacher NA, Stacy AE, Porter GM, Merlot AM. Clinical development of targeted and immune based anti-cancer therapies. *J Exp Clin Cancer Res* 2019; 38(1): 156
 8. Liang G, Fan W, Luo H, Zhu X. The emerging roles of artificial intelligence in cancer drug development and precision therapy. *Biomed Pharmacother* 2020; 128: 110255
 9. Kumar Y, Gupta S, Singla R, Hu YC. A systematic review of artificial intelligence techniques in cancer prediction and diagnosis. *Arch Comput Methods Eng* 2022; 29(4): 2043–2070
 10. Pavlopoulou A, Spandidos DA, Michalopoulos I. Human cancer databases (review). *Oncol Rep* 2015; 33(1): 3–18
 11. Rios Velazquez E, Parmar C, Liu Y, Coroller TP, Cruz G, Stringfield O, Ye Z, Makrigiorgos M, Fennessy F, Mak RH, Gillies R, Quackenbush J, Aerts HJWL. Somatic mutations drive distinct imaging phenotypes in lung cancer. *Cancer Res* 2017; 77(14): 3922–3930
 12. Hou Q, Bing ZT, Hu C, Li MY, Yang KH, Mo Z, Xie XW, Liao JL, Lu Y, Horie S, Lou MW. RankProd combined with genetic algorithm optimized artificial neural network establishes a diagnostic and prognostic prediction model that revealed C1QTNF3 as a biomarker for prostate cancer. *EBioMedicine* 2018; 32: 234–244
 13. Carter SM, Rogers W, Win KT, Frazer H, Richards B, Houssami N. The ethical, legal and social implications of using artificial intelligence systems in breast cancer care. *Breast* 2020; 49: 25–32
 14. Kaul V, Enslin S, Gross SA. History of artificial intelligence in medicine. *Gastrointest Endosc* 2020; 92(4): 807–812
 15. Fekete JT, Gyórfy B. ROCplot. org: validating predictive biomarkers of chemotherapy/hormonal therapy/anti-HER2 therapy using transcriptomic data of 3, 104 breast cancer patients. *Int J Cancer* 2019; 145(11): 3140–3151
 16. Quinlan JR. Induction of decision trees. *Mach Learn* 1986; 1(1): 81–106
 17. Kingsford C, Salzberg SL. What are decision trees? *Nat Biotechnol* 2008; 26(9): 1011–1013
 18. Talhouk A, McConechy MK, Leung S, Yang W, Lum A, Senz J, Boyd N, Pike J, Anglesio M, Kwon JS, Karnezis AN, Huntsman DG, Gilks CB, McAlpine JN. Confirmation of ProMisE: a simple, genomics-based clinical classifier for endometrial cancer. *Cancer* 2017; 123(5): 802–813
 19. Selman TJ, Mann C, Zamora J, Appleyard TL, Khan K. Diagnostic accuracy of tests for lymph node status in primary cervical cancer: a systematic review and meta-analysis. *CMAJ* 2008; 178(7): 855–862
 20. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, Zegers CM, Gillies R, Boellard R, Dekker A, Aerts HJ. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer Oxf Engl* 2012; 48(4): 441–446
 21. Wu Q, Wang S, Chen X, Wang Y, Dong L, Liu Z, Tian J, Wang M. Radiomics analysis of magnetic resonance imaging improves diagnostic performance of lymph node metastasis in patients with cervical cancer. *Radiother Oncol* 2019; 138: 141–148
 22. Gevaert O, Echegaray S, Khuong A, Hoang CD, Shrager JB, Jensen KC, Berry GJ, Guo HH, Lau C, Plevritis SK, Rubin DL, Napel S, Leung AN. Predictive radiogenomics modeling of EGFR mutation status in lung cancer. *Sci Rep* 2017; 7(1): 41674
 23. Hu LS, Ning S, Eschbacher JM, Baxter LC, Gaw N, Ranjbar S, Plasencia J, Dueck AC, Peng S, Smith KA, Nakaji P, Karis JP, Quarles CC, Wu T, Loftus JC, Jenkins RB, Sicotte H, Kollmeyer TM, O'Neill BP, Elmquist W, Hoxworth JM, Frakes D, Sarkaria J, Swanson KR, Tran NL, Li J, Mitchell JR. Radiogenomics to characterize regional genetic heterogeneity in glioblastoma. *Neuro-oncol* 2017; 19(1): 128–137
 24. Tahmassebi A, Wengert GJ, Helbich TH, Bago-Horvath Z, Alaei S, Bartsch R, Dubsy P, Baltzer P, Clauser P, Kapetas P, Morris EA, Meyer-Baese A, Pinker K. Impact of machine learning with multiparametric magnetic resonance imaging of the breast for early prediction of response to neoadjuvant chemotherapy and survival outcomes in breast cancer patients. *Invest Radiol* 2019; 54(2): 110–117
 25. Wang Y, Wang D, Geng N, Wang Y, Yin Y, Jin Y. Stacking-based ensemble learning of decision trees for interpretable prostate cancer detection. *Appl Soft Comput* 2019; 77: 188–204
 26. Ghiasi MM, Zendejboudi S. Application of decision tree-based ensemble learning in the classification of breast cancer. *Comput Biol Med* 2021; 128: 104089
 27. De Felice F, Crocetti D, Parisi M, Maiuri V, Moscarelli E, Caiazzo R, Bulzonetti N, Musio D, Tombolini V. Decision tree algorithm in locally advanced rectal cancer: an example of over-interpretation and misuse of a machine learning approach. *J Cancer Res Clin Oncol* 2020; 146(3): 761–765
 28. Breiman L. Random forests. *Mach Learn* 2001; 45(1): 5–32
 29. Tibor Fekete J, Gyórfy B. A unified platform enabling biomarker ranking and validation for 1562 drugs using transcriptomic data of 1250 cancer cell lines. *Comput Struct Biotechnol J* 2022; 20: 2885–2894
 30. Cohen JD, Li L, Wang Y, Thoburn C, Afsari B, Danilova L, Douville C, Javed AA, Wong F, Mattox A, Hruban RH, Wolfgang CL, Goggins MG, Dal Molin M, Wang TL, Roden R, Klein AP, Ptak J, Dobbyn L, Schaefer J, Silliman N, Popoli M, Vogelstein JT, Browne JD, Schoen RE, Brand RE, Tie J, Gibbs P, Wong HL, Mansfield AS, Jen J, Hanash SM, Falconi M, Allen PJ, Zhou S, Bettegowda C, Diaz LA Jr, Tomasetti C, Kinzler KW, Vogelstein B, Lennon AM, Papadopoulos N. Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science* 2018; 359(6378): 926–930
 31. Toth R, Schiffmann H, Hube-Magg C, Büschek F, Höflmayer D, Weidemann S, Lebok P, Fraune C, Minner S, Schlomm T, Sauter G, Plass C, Assenov Y, Simon R, Meiners J, Gerhäuser C. Random forest-based modelling to detect biomarkers for prostate cancer progression. *Clin Epigenetics* 2019; 11(1): 148
 32. Kawakami E, Tabata J, Yanaihara N, Ishikawa T, Koseki K, Iida Y, Saito M, Komazaki H, Shapiro JS, Goto C, Akiyama Y, Saito R, Saito M, Takano H, Yamada K, Okamoto A. Application of artificial intelligence for preoperative diagnostic and prognostic prediction in epithelial ovarian cancer based on blood biomarkers. *Clin Cancer Res* 2019; 25(10): 3006–3015
 33. Zhang Z, Huang L, Li J, Wang P. Bioinformatics analysis reveals immune prognostic markers for overall survival of colorectal cancer patients: a novel machine learning survival predictive system. *BMC Bioinformatics* 2022; 23(1): 124

34. Lin J, Yin M, Liu L, Gao J, Yu C, Liu X, Xu C, Zhu J. The development of a prediction model based on random survival forest for the postoperative prognosis of pancreatic cancer: a SEER-based study. *Cancers (Basel)* 2022; 14(19): 4667
35. Zhang H, Chi M, Su D, Xiong Y, Wei H, Yu Y, Zuo Y, Yang L. A random forest-based metabolic risk model to assess the prognosis and metabolism-related drug targets in ovarian cancer. *Comput Biol Med* 2023; 153: 106432
36. Lee JY, Lee KS, Seo BK, Cho KR, Woo OH, Song SE, Kim EK, Lee HY, Kim JS, Cha J. Radiomic machine learning for predicting prognostic biomarkers and molecular subtypes of breast cancer using tumor heterogeneity and angiogenesis properties on MRI. *Eur Radiol* 2022; 32(1): 650–660
37. Horvat N, Veeraraghavan H, Khan M, Blazic I, Zheng J, Capanu M, Sala E, Garcia-Aguilar J, Gollub MJ, Petkovska I. MR imaging of rectal cancer: radiomics analysis to assess treatment response after neoadjuvant therapy. *Radiology* 2018; 287(3): 833–843
38. Deist TM, Dankers FJWM, Valdes G, Wijsman R, Hsu IC, Oberije C, Lustberg T, van Soest J, Hoebers F, Jochems A, El Naqa I, Wee L, Morin O, Raleigh DR, Bots W, Kaanders JH, Belderbos J, Kwint M, Solberg T, Monshouwer R, Bussink J, Dekker A, Lambin P. Machine learning algorithms for outcome prediction in (chemo)radiotherapy: an empirical comparison of classifiers. *Med Phys* 2018; 45(7): 3449–3459
39. Paul D, Su R, Romain M, Sébastien V, Pierre V, Isabelle G. Feature selection for outcome prediction in oesophageal cancer using genetic algorithm and random forest classifier. *Comput Med Imaging Graph* 2017; 60: 42–49
40. Wang S, Wang Y, Wang D, Yin Y, Wang Y, Jin Y. An improved random forest-based rule extraction method for breast cancer diagnosis. *Appl Soft Comput* 2020; 86: 105941
41. Singh H, Singh S, Singla D, Agarwal SM, Raghava GPS. QSAR based model for discriminating EGFR inhibitors and non-inhibitors using Random forest. *Biol Direct* 2015; 10: 10
42. Li X, Xu Y, Cui H, Huang T, Wang D, Lian B, Li W, Qin G, Chen L, Xie L. Prediction of synergistic anti-cancer drug combinations based on drug target network and drug induced gene expression profiles. *Artif Intell Med* 2017; 83: 35–43
43. Sidorov P, Naulaerts S, Arieu-Bonnet J, Pasquier E, Ballester PJ. Predicting synergism of cancer drug combinations using NCI-ALMANAC data. *Front Chem* 2019; 7: 509
44. Yang C, Huang X, Li Y, Chen J, Lv Y, Dai S. Prognosis and personalized treatment prediction in TP53-mutant hepatocellular carcinoma: an *in silico* strategy towards precision oncology. *Brief Bioinform* 2021; 22(3): bbaa164
45. Hu C, Steingrimsson JA. Personalized risk prediction in clinical oncology research: applications and practical issues using survival trees and random forests. *J Biopharm Stat* 2018; 28(2): 333–349
46. Natekin A, Knoll A. Gradient boosting machines, a tutorial. *Front Neurorobot* 2013; 7: 21
47. Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Stat* 2001; 29(5): 1189–1232
48. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (Pages 785–794)*. 2016; arXiv: 1603.02754
49. Mallett S, Royston P, Waters R, Dutton S, Altman DG. Reporting performance of prognostic models in cancer: a review. *BMC Med* 2010; 8(1): 21
50. Moncada-Torres A, van Maaren MC, Hendriks MP, Siesling S, Geleijnse G. Explainable machine learning can outperform Cox regression predictions and provide insights in breast cancer survival. *Sci Rep* 2021; 11(1): 6968
51. Nazari M, Shiri I, Zaidi H. Radiomics-based machine learning model to predict risk of death within 5-years in clear cell renal cell carcinoma patients. *Comput Biol Med* 2021; 129: 104135
52. Arrieta O, Cardona AF, Martín C, Más-López L, Corrales-Rodríguez L, Bramuglia G, Castillo-Fernandez O, Meyerson M, Amieva-Rivera E, Campos-Parra AD, Carranza H, Gómez de la Torre JC, Powazniak Y, Aldaco-Sarvide F, Vargas C, Trigo M, Magallanes-Maciel M, Otero J, Sánchez-Reyes R, Cuello M. Updated frequency of EGFR and KRAS mutations in non-small-cell lung cancer in Latin America: the Latin-American Consortium for the Investigation of Lung Cancer (CLICaP). *J Thorac Oncol* 2015; 10(5): 838–843
53. Le NQK, Kha QH, Nguyen VH, Chen YC, Cheng SJ, Chen CY. Machine learning-based radiomics signatures for EGFR and KRAS mutations prediction in non-small-cell lung cancer. *Int J Mol Sci* 2021; 22(17): 9254
54. Li Q, Yang H, Wang P, Liu X, Lv K, Ye M. XGBoost-based and tumor-immune characterized gene signature for the prediction of metastatic status in breast cancer. *J Transl Med* 2022; 20(1): 177
55. Liu X, Yuan P, Li R, Zhang D, An J, Ju J, Liu C, Ren F, Hou R, Li Y, Yang J. Predicting breast cancer recurrence and metastasis risk by integrating color and texture features of histopathological images and machine learning technologies. *Comput Biol Med* 2022; 146: 105569
56. Bomane A, Gonçalves A, Ballester PJ. Paclitaxel response can be predicted with interpretable multi-variate classifiers exploiting DNA-methylation and miRNA data. *Front Genet* 2019; 10: 1041
57. Polano M, Chierici M, Dal Bo M, Gentilini D, Di Cintio F, Baboci L, Gibbs DL, Furlanello C, Toffoli G. A pan-cancer approach to predict responsiveness to immune checkpoint inhibitors by machine learning. *Cancers (Basel)* 2019; 11(10): 1562
58. Ji GW, Jiao CY, Xu ZG, Li XC, Wang K, Wang XH. Development and validation of a gradient boosting machine to predict prognosis after liver resection for intrahepatic cholangiocarcinoma. *BMC Cancer* 2022; 22(1): 258
59. Pfof A, Mehrara BJ, Nelson JA, Wilkins EG, Pusic AL, Sidey-Gibbons C. Towards patient-centered decision-making in breast cancer surgery: machine learning to predict individual patient-reported outcomes at 1-year follow-up. *Ann Surg* 2023; 277(1): e144–e152
60. Ma B, Meng F, Yan G, Yan H, Chai B, Song F. Diagnostic classification of cancers using extreme gradient boosting algorithm and multi-omics data. *Comput Biol Med* 2020; 121: 103761
61. Cortes C, Vapnik V. Support-vector networks. *Mach Learn* 1995; 20(3): 273–297
62. Noble WS. What is a support vector machine? *Nat Biotechnol* 2006; 24(12): 1565–1567
63. Furey TS, Cristianini N, Duffy N, Bednarski DW, Schummer M, Haussler D. Support vector machine classification and validation

- of cancer tissue samples using microarray expression data. *Bioinformatics* 2000; 16(10): 906–914
64. Dorman SN, Baranova K, Knoll JHM, Urquhart BL, Mariani G, Carcangiu ML, Rogan PK. Genomic signatures for paclitaxel and gemcitabine resistance in breast cancer derived by machine learning. *Mol Oncol* 2016; 10(1): 85–100
 65. Zhang B, He X, Ouyang F, Gu D, Dong Y, Zhang L, Mo X, Huang W, Tian J, Zhang S. Radiomic machine-learning classifiers for prognostic biomarkers of advanced nasopharyngeal carcinoma. *Cancer Lett* 2017; 403: 21–27
 66. Liu Z, Zhang XY, Shi YJ, Wang L, Zhu HT, Tang Z, Wang S, Li XT, Tian J, Sun YS. Radiomics analysis for evaluation of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. *Clin Cancer Res* 2017; 23(23): 7253–7262
 67. Yang L, Dong D, Fang M, Zhu Y, Zang Y, Liu Z, Zhang H, Ying J, Zhao X, Tian J. Can CT-based radiomics signature predict KRAS/NRAS/BRAF mutations in colorectal cancer? *Eur Radiol* 2018; 28(5): 2058–2067
 68. Abeel T, Helleputte T, Van de Peer Y, Dupont P, Saeys Y. Robust biomarker identification for cancer diagnosis with ensemble feature selection methods. *Bioinformatics* 2010; 26(3): 392–398
 69. Li Y, Zhao J, Yu S, Wang Z, He X, Su Y, Guo T, Sheng H, Chen J, Zheng Q, Li Y, Guo W, Cai X, Shi G, Wu J, Wang L, Wang P, He X, Huang S. Extracellular vesicles long RNA sequencing reveals abundant mRNA, circRNA, and lncRNA in human blood as potential biomarkers for cancer diagnosis. *Clin Chem* 2019; 65(6): 798–808
 70. Qiu J, Peng B, Tang Y, Qian Y, Guo P, Li M, Luo J, Chen B, Tang H, Lu C, Cai M, Ke Z, He W, Zheng Y, Xie D, Li B, Yuan Y. CpG methylation signature predicts recurrence in early-stage hepatocellular carcinoma: results from a multicenter study. *J Clin Oncol* 2017; 35(7): 734–742
 71. Jiang Y, Xie J, Han Z, Liu W, Xi S, Huang L, Huang W, Lin T, Zhao L, Hu Y, Yu J, Zhang Q, Li T, Cai S, Li G. Immunomarker support vector machine classifier for prediction of gastric cancer survival and adjuvant chemotherapeutic benefit. *Clin Cancer Res* 2018; 24(22): 5574–5584
 72. Cheong JH, Wang SC, Park S, Porembka MR, Christie AL, Kim H, Kim HS, Zhu H, Hyung WJ, Noh SH, Hu B, Hong C, Karalis JD, Kim IH, Lee SH, Hwang TH. Development and validation of a prognostic and predictive 32-gene signature for gastric cancer. *Nat Commun* 2022; 13(1): 774
 73. Delahunt B, Eble JN, Egevad L, Samarantunga H. Grading of renal cell carcinoma. *Histopathology* 2019; 74(1): 4–17
 74. FIGO Committee on Gynecologic Oncology. FIGO staging for carcinoma of the vulva, cervix, and corpus uteri. *Int J Gynaecol Obstet* 2014; 125(2): 97–98
 75. Bektas CT, Kocak B, Yardimci AH, Turkcanoglu MH, Yucetas U, Koca SB, Erdim C, Kilickesmez O. Clear cell renal cell carcinoma: machine learning-based quantitative computed tomography texture analysis for prediction of Fuhrman nuclear grade. *Eur Radiol* 2019; 29(3): 1153–1163
 76. Xie L, Chu R, Wang K, Zhang X, Li J, Zhao Z, Yao S, Wang Z, Dong T, Yang X, Su X, Qiao X, Song K, Kong B. Prognostic assessment of cervical cancer patients by clinical staging and surgical-pathological factor: a support vector machine-based approach. *Front Oncol* 2020; 10: 1353
 77. Xu G, Zhang M, Zhu H, Xu J. A 15-gene signature for prediction of colon cancer recurrence and prognosis based on SVM. *Gene* 2017; 604: 33–40
 78. Jiang D, Lei T, Wang Z, Shen C, Cao D, Hou T. ADMET evaluation in drug discovery. 20. Prediction of breast cancer resistance protein inhibition through machine learning. *J Cheminform* 2020; 12(1): 16
 79. Huang X, Zhang L, Wang B, Li F, Zhang Z. Feature clustering based support vector machine recursive feature elimination for gene selection. *Appl Intell* 2018; 48(3): 594–607
 80. Wang H, Zheng B, Yoon SW, Ko HS. A support vector machine-based ensemble algorithm for breast cancer diagnosis. *Eur J Oper Res* 2018; 267(2): 687–699
 81. Cooper GF, Herskovits E. A Bayesian method for the induction of probabilistic networks from data. *Mach Learn* 1992; 9(4): 309–347
 82. Friedman N, Geiger D, Goldszmidt M. Bayesian network classifiers. *Mach Learn* 1997; 29(2/3): 131–163
 83. Johnson M, Albizri A, Simsek S. Artificial intelligence in healthcare operations to enhance treatment outcomes: a framework to predict lung cancer prognosis. *Ann Oper Res* 2022; 308(1-2): 275–305
 84. Li R, Zhang C, Du K, Dan H, Ding R, Cai Z, Duan L, Xie Z, Zheng G, Wu H, Ren G, Dou X, Feng F, Zheng J. Analysis of prognostic factors of rectal cancer and construction of a prognostic prediction model based on Bayesian network. *Front Public Health* 2022; 10: 842970
 85. Wright GW, Huang DW, Phelan JD, Coulibaly ZA, Roulland S, Young RM, Wang JQ, Schmitz R, Morin RD, Tang J, Jiang A, Bagaev A, Plotnikova O, Kotlov N, Johnson CA, Wilson WH, Scott DW, Staudt LM. A probabilistic classification tool for genetic subtypes of diffuse large B cell lymphoma with therapeutic implications. *Cancer Cell* 2020; 37(4): 551–568.e14
 86. Jochems A, Deist TM, van Soest J, Eble M, Bulens P, Coucke P, Dries W, Lambin P, Dekker A. Distributed learning: developing a predictive model based on data from multiple hospitals without data leaving the hospital—a real life proof of concept. *Radiother Oncol* 2016; 121(3): 459–467
 87. Braman NM, Etesami M, Prasanna P, Dubchuk C, Gilmore H, Tiwari P, Plecha D, Madabhushi A. Intratumoral and peritumoral radiomics for the pretreatment prediction of pathological complete response to neoadjuvant chemotherapy based on breast DCE-MRI. *Breast Cancer Res* 2017; 19(1): 57
 88. Yang L, Fu B, Li Y, Liu Y, Huang W, Feng S, Xiao L, Sun L, Deng L, Zheng X, Ye F, Bu H. Prediction model of the response to neoadjuvant chemotherapy in breast cancers by a Naive Bayes algorithm. *Comput Methods Programs Biomed* 2020; 192: 105458
 89. Rouhi R, Jafari M, Kasaei S, Keshavarzian P. Benign and malignant breast tumors classification based on region growing and CNN segmentation. *Expert Syst Appl* 2015; 42(3): 990–1002
 90. Russo DP, Zorn KM, Clark AM, Zhu H, Ekins S. Comparing multiple machine learning algorithms and metrics for estrogen receptor binding prediction. *Mol Pharm* 2018; 15(10): 4361–4370
 91. Karabatak M. A new classifier for breast cancer detection based on Naïve Bayesian. *Measurement* 2015; 72: 32–36
 92. Abdar M, Zomorodi-Moghadam M, Zhou X, Gururajan R, Tao X,

- Barua PD, Gururajan R. A new nested ensemble technique for automated diagnosis of breast cancer. *Pattern Recognit Lett* 2020; 132: 123–131
93. Farid DM, Zhang L, Rahman CM, Hossain MA, Strachan R. Hybrid decision tree and naïve Bayes classifiers for multi-class classification tasks. *Expert Syst Appl* 2014; 41(4): 1937–1946
94. Chen S, Webb GI, Liu L, Ma X. A novel selective naïve Bayes algorithm. *Knowl Base Syst* 2020; 192: 105361
95. Cover T, Hart P. Nearest neighbor pattern classification. *IEEE Trans Inf Theory* 1967; 13(1): 21–27
96. Peterson LE. K-nearest neighbor. *Scholarpedia J* 2009; 4(2): 1883
97. Dhahbi S, Barhoumi W, Zagrouba E. Breast cancer diagnosis in digitized mammograms using curvelet moments. *Comput Biol Med* 2015; 64: 79–90
98. Huang Q, Huang Y, Luo Y, Yuan F, Li X. Segmentation of breast ultrasound image with semantic classification of superpixels. *Med Image Anal* 2020; 61: 101657
99. Waugh SA, Purdie CA, Jordan LB, Vinnicombe S, Lerski RA, Martin P, Thompson AM. Magnetic resonance imaging texture analysis classification of primary breast cancer. *Eur Radiol* 2016; 26(2): 322–330
100. Leithner D, Horvat JV, Marino MA, Bernard-Davila B, Jochelson MS, Ochoa-Albiztegui RE, Martinez DF, Morris EA, Thakur S, Pinker K. Radiomic signatures with contrast-enhanced magnetic resonance imaging for the assessment of breast cancer receptor status and molecular subtypes: initial results. *Breast Cancer Res* 2019; 21(1): 106
101. Garcia-Laencina PJ, Abreu PH, Abreu MH, Afonosos N. Missing data imputation on the 5-year survival prediction of breast cancer patients with unknown discrete values. *Comput Biol Med* 2015; 59: 125–133
102. Wu W, Parmar C, Grossmann P, Quackenbush J, Lambin P, Bussink J, Mak R, Aerts HJ. Exploratory study to identify radiomics classifiers for lung cancer histology. *Front Oncol* 2016; 6: 71
103. Tandel GS, Balestrieri A, Jujaray T, Khanna NN, Saba L, Suri JS. Multiclass magnetic resonance imaging brain tumor classification using artificial intelligence paradigm. *Comput Biol Med* 2020; 122: 103804
104. Wang A, An N, Chen G, Li L, Alterovitz G. Accelerating wrapper-based feature selection with K-nearest-neighbor. *Knowl Base Syst* 2015; 83: 81–91
105. Saeyns Y, Inza I, Larrañaga P. A review of feature selection techniques in bioinformatics. *Bioinformatics* 2007; 23(19): 2507–2517
106. Kar S, Das Sharma K, Maitra M. Gene selection from microarray gene expression data for classification of cancer subgroups employing PSO and adaptive K-nearest neighborhood technique. *Expert Syst Appl* 2015; 42(1): 612–627
107. Zhang S, Li X, Zong M, Zhu X, Wang R. Efficient kNN classification with different numbers of nearest neighbors. *IEEE Trans Neural Netw Learn Syst* 2018; 29(5): 1774–1785
108. Uddin S, Haque I, Lu H, Moni MA, Gide E. Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction. *Sci Rep* 2022; 12(1): 6256
109. Krogh A. What are artificial neural networks? *Nat Biotechnol* 2008; 26(2): 195–197
110. Khan J, Wei JS, Ringnér M, Saal LH, Ladanyi M, Westermann F, Berthold F, Schwab M, Antonescu CR, Peterson C, Meltzer PS. Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. *Nat Med* 2001; 7(6): 673–679
111. Abbas HA. An evolutionary artificial neural networks approach for breast cancer diagnosis. *Artif Intell Med* 2002; 25(3): 265–281
112. Wei JS, Greer BT, Westermann F, Steinberg SM, Son CG, Chen QR, Whiteford CC, Bilke S, Krasnoselsky AL, Cenacchi N, Catchpole D, Berthold F, Schwab M, Khan J. Prediction of clinical outcome using gene expression profiling and artificial neural networks for patients with neuroblastoma. *Cancer Res* 2004; 64(19): 6883–6891
113. Delen D, Walker G, Kadam A. Predicting breast cancer survivability: a comparison of three data mining methods. *Artif Intell Med* 2005; 34(2): 113–127
114. Dheeba J, Albert Singh N, Tamil Selvi S. Computer-aided detection of breast cancer on mammograms: a swarm intelligence optimized wavelet neural network approach. *J Biomed Inform* 2014; 49: 45–52
115. Bhardwaj A, Tiwari A. Breast cancer diagnosis using Genetically Optimized Neural Network model. *Expert Syst Appl* 2015; 42(10): 4611–4620
116. Alshayegi MH, Ellethy H, Abed S, Gupta R. Computer-aided detection of breast cancer on the Wisconsin dataset: an artificial neural networks approach. *Biomed Signal Process Control* 2022; 71: 103141
117. Almeida PP, Cardoso CP, de Freitas LM. PDAC-ANN: an artificial neural network to predict pancreatic ductal adenocarcinoma based on gene expression. *BMC Cancer* 2020; 20(1): 82
118. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015; 521(7553): 436–444
119. LeCun Y, Boser B, Denker J, Henderson D, Howard R, Hubbard W, Jackel L. Handwritten digit recognition with a back-propagation network. *NIPS'89: Proceedings of the 2nd International Conference on Neural Information Processing Systems*. 1989; 396–404
120. Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. *Insights Imaging* 2018; 9(4): 611–629
121. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017; 542(7639): 115–118
122. Bychkov D, Linder N, Turkki R, Nordling S, Kovanen PE, Verrill C, Walliander M, Lundin M, Haglund C, Lundin J. Deep learning based tissue analysis predicts outcome in colorectal cancer. *Sci Rep* 2018; 8(1): 3395
123. Kather JN, Krisam J, Charoentong P, Luedde T, Herpel E, Weis CA, Gaiser T, Marx A, Valous NA, Ferber D, Jansen L, Reyes-Aldasoro CC, Zörnig I, Jäger D, Brenner H, Chang-Claude J, Hoffmeister M, Halama N. Predicting survival from colorectal cancer histology slides using deep learning: a retrospective multicenter study. *PLoS Med* 2019; 16(1): e1002730
124. Skrede OJ, De Raedt S, Kleppe A, Hveem TS, Liestøl K, Maddison J, Askautrud HA, Pradhan M, Nesheim JA, Albrechtsen F, Farstad IN, Domingo E, Church DN, Nesbakken A, Shepherd

- NA, Tomlinson I, Kerr R, Novelli M, Kerr DJ, Danielsen HE. Deep learning for prediction of colorectal cancer outcome: a discovery and validation study. *Lancet* 2020; 395(10221): 350–360
125. Sirinukunwattana K, Domingo E, Richman SD, Redmond KL, Blake A, Verrill C, Leedham SJ, Chatzipli A, Hardy C, Whalley CM, Wu CH, Beggs AD, McDermott U, Dunne PD, Meade A, Walker SM, Murray GI, Samuel L, Seymour M, Tomlinson I, Quirke P, Maughan T, Rittscher J, Koelzer VH; S:CORT consortium. Image-based consensus molecular subtype (imCMS) classification of colorectal cancer using deep learning. *Gut* 2021; 70(3): 544–554
 126. Kather JN, Heij LR, Grabsch HI, Loeffler C, Echle A, Muti HS, Krause J, Niehues JM, Sommer KAJ, Bankhead P, Kooreman LFS, Schulte JJ, Cipriani NA, Buelow RD, Boor P, Ortiz-Brüchle NN, Hanby AM, Speirs V, Kochanny S, Patnaik A, Srisuwananukorn A, Brenner H, Hoffmeister M, van den Brandt PA, Jäger D, Trautwein C, Pearson AT, Luedde T. Pan-cancer image-based detection of clinically actionable genetic alterations. *Nat Cancer* 2020; 1(8): 789–799
 127. Yang J, Ju J, Guo L, Ji B, Shi S, Yang Z, Gao S, Yuan X, Tian G, Liang Y, Yuan P. 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
 128. Xu Y, Hosny A, Zeleznik R, Parmar C, Coroller T, Franco I, Mak RH, Aerts HJWL. Deep learning predicts lung cancer treatment response from serial medical imaging. *Clin Cancer Res* 2019; 25(11): 3266–3275
 129. Freer TW, Ullissey MJ. Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center. *Radiology* 2001; 220(3): 781–786
 130. Kooi T, Litjens G, van Ginneken B, Gubern-Mérida A, Sánchez CI, Mann R, den Heeten A, Karssemeijer N. Large scale deep learning for computer aided detection of mammographic lesions. *Med Image Anal* 2017; 35: 303–312
 131. Ren S, He K, Girshick R, Sun J. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Trans Pattern Anal Mach Intell* 2017; 39(6): 1137–1149
 132. Moreira IC, Amaral I, Domingues I, Cardoso A, Cardoso MJ, Cardoso JS. INbreast: toward a full-field digital mammographic database. *Acad Radiol* 2012; 19(2): 236–248
 133. Shen L, Margolies LR, Rothstein JH, Fluder E, McBride R, Sieh W. Deep learning to improve breast cancer detection on screening mammography. *Sci Rep* 2019; 9(1): 12495
 134. Hekler A, Utikal JS, Enk AH, Berking C, Klode J, Schadendorf D, Jansen P, Franklin C, Holland-Letz T, Krahl D, von Kalle C, Fröhling S, Brinker TJ. Pathologist-level classification of histopathological melanoma images with deep neural networks. *Eur J Cancer* 2019; 115: 79–83
 135. Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, Tse D, Etemadi M, Ye W, Corrado G, Naidich DP, Shetty S. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat Med* 2019; 25(6): 954–961
 136. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, Back T, Chesus M, Corrado GS, Darzi A, Etemadi M, Garcia-Vicente F, Gilbert FJ, Halling-Brown M, Hassabis D, Jansen S, Karthikesalingam A, Kelly CJ, King D, Ledam JR, Melnick D, Mostofi H, Peng L, Reicher JJ, Romera-Paredes B, Sidebottom R, Suleyman M, Tse D, Young KC, De Fauw J, Shetty S. International evaluation of an AI system for breast cancer screening. *Nature* 2020; 577(7788): 89–94
 137. Schaffter T, Buist DSM, Lee CI, Nikulin Y, Ribli D, Guan Y, Lotter W, Jie Z, Du H, Wang S, Feng J, Feng M, Kim HE, Albiol F, Albiol A, Morrell S, Wojna Z, Ahsen ME, Asif U, Jimeno Yepes A, Yohanandan S, Rabinovici-Cohen S, Yi D, Hoff B, Yu T, Chaibub Neto E, Rubin DL, Lindholm P, Margolies LR, McBride RB, Rothstein JH, Sieh W, Ben-Ari R, Harrer S, Trister A, Friend S, Norman T, Sahiner B, Strand F, Guinney J, Stolovitzky G; and the DM DREAM Consortium; Mackey L, Cahoon J, Shen L, Sohn JH, Trivedi H, Shen Y, Buturovic L, Pereira JC, Cardoso JS, Castro E, Kalleberg KT, Pelka O, Nedjar I, Geras KJ, Nensa F, Goan E, Koitka S, Caballero L, Cox DD, Krishnaswamy P, Pandey G, Friedrich CM, Perrin D, Fookes C, Shi B, Cardoso Negrie G, Kawczynski M, Cho K, Khoo CS, Lo JY, Sorensen AG, Jung H. Evaluation of combined artificial intelligence and radiologist assessment to interpret screening mammograms. *JAMA Netw Open* 2020; 3(3): e200265
 138. Wu N, Phang J, Park J, Shen Y, Huang Z, Zorin M, Jastrzębski S, Févry T, Katsnelson J, Kim E, Wolfson S, Parikh U, Gaddam S, Lin LLY, Ho K, Weinstein JD, Reig B, Gao Y, Toth H, Pysarenko K, Lewin A, Lee J, Airola K, Mema E, Chung S, Hwang E, Samreen N, Kim SG, Heacock L, Moy L, Cho K, Geras KJ. Deep neural networks improve radiologists' performance in breast cancer screening. *IEEE Trans Med Imaging* 2020; 39(4): 1184–1194
 139. Tabibu S, Vinod PK, Jawahar CV. Pan-renal cell carcinoma classification and survival prediction from histopathology images using deep learning. *Sci Rep* 2019; 9(1): 10509
 140. Jin X, Zhou YF, Ma D, Zhao S, Lin CJ, Xiao Y, Fu T, Liu CL, Chen YY, Xiao WX, Liu YQ, Chen QW, Yu Y, Shi LM, Shi JX, Huang W, Robertson JFR, Jiang YZ, Shao ZM. Molecular classification of hormone receptor-positive HER2-negative breast cancer. *Nat Genet* 2023; 55(10): 1696–1708
 141. Munquad S, Das AB. DeepAutoGlioma: a deep learning autoencoder-based multi-omics data integration and classification tools for glioma subtyping. *BioData Min* 2023; 16(1): 32
 142. Chen M, Zhang B, Topatana W, Cao J, Zhu H, Juengpanich S, Mao Q, Yu H, Cai X. Classification and mutation prediction based on histopathology H&E images in liver cancer using deep learning. *NPJ Precis Oncol* 2020; 4(1): 14
 143. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyö D, Moreira AL, Razavian N, Tsirigos A. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med* 2018; 24(10): 1559–1567
 144. Sharifi-Noghabi H, Zolotareva O, Collins CC, Ester M. MOLI: multi-omics late integration with deep neural networks for drug response prediction. *Bioinformatics* 2019; 35(14): i501–i509
 145. Sakellaropoulos T, Vougas K, Narang S, Koinis F, Kotsinas A, Polyzos A, Moss TJ, Piha-Paul S, Zhou H, Kardala E, Damianidou E, Alexopoulos LG, Aifantis I, Townsend PA, Panayiotidis MI, Sfrikakis P, Bartek J, Fitzgerald RC, Thanos D, Mills Shaw KR, Petty R, Tsirigos A, Gorgoulis VG. A deep learning framework for predicting response to therapy in cancer.

- Cell Rep 2019; 29(11): 3367–3373.e4
146. Kuenzi BM, Park J, Fong SH, Sanchez KS, Lee J, Kreisberg JF, Ma J, Ideker T. Predicting drug response and synergy using a deep learning model of human cancer cells. *Cancer Cell* 2020; 38(5): 672–684.e6
 147. Su R, Liu X, Wei L, Zou Q. Deep-Resp-Forest: a deep forest model to predict anti-cancer drug response. *Methods* 2019; 166: 91–102
 148. Preuer K, Lewis RPI, Hochreiter S, Bender A, Bulusu KC, Klambauer G. DeepSynergy: predicting anti-cancer drug synergy with deep learning. *Bioinformatics* 2018; 34(9): 1538–1546
 149. Rajkumar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, Liu PJ, Liu X, Marcus J, Sun M, Sundberg P, Yee H, Zhang K, Zhang Y, Flores G, Duggan GE, Irvine J, Le Q, Litsch K, Mossin A, Tansuwan J, Wang D, Wexler J, Wilson J, Ludwig D, Volchenboum SL, Chou K, Pearson M, Madabushi S, Shah NH, Butte AJ, Howell MD, Cui C, Corrado GS, Dean J. Scalable and accurate deep learning with electronic health records. *NPJ Digit Med* 2018; 1(1): 18
 150. Yala A, Lehman C, Schuster T, Portnoi T, Barzilay R. A deep learning mammography-based model for improved breast cancer risk prediction. *Radiology* 2019; 292(1): 60–66
 151. Rasmy L, Xiang Y, Xie Z, Tao C, Zhi D. Med-BERT: pretrained contextualized embeddings on large-scale structured electronic health records for disease prediction. *NPJ Digit Med* 2021; 4(1): 86
 152. Miotto R, Li L, Kidd BA, Dudley JT. Deep Patient: an unsupervised representation to predict the future of patients from the electronic health records. *Sci Rep* 2016; 6(1): 26094
 153. Hosny A, Parmar C, Coroller TP, Grossmann P, Zeleznik R, Kumar A, Bussink J, Gillies RJ, Mak RH, Aerts HJWL. Deep learning for lung cancer prognostication: a retrospective multi-cohort radiomics study. *PLoS Med* 2018; 15(11): e1002711
 154. Hirschberg J, Manning CD. Advances in natural language processing. *Science* 2015; 349(6245): 261–266
 155. Yim WW, Yetisgen M, Harris WP, Kwan SW. Natural language processing in oncology: a review. *JAMA Oncol* 2016; 2(6): 797–804
 156. Castro SM, Tseytlin E, Medvedeva O, Mitchell K, Visweswaran S, Bekhuis T, Jacobson RS. Automated annotation and classification of BI-RADS assessment from radiology reports. *J Biomed Inform* 2017; 69: 177–187
 157. Patel TA, Puppala M, Ogunti RO, Ensor JE, He T, Shewale JB, Ankerst DP, Kaklamani VG, Rodriguez AA, Wong ST, Chang JC. Correlating mammographic and pathologic findings in clinical decision support using natural language processing and data mining methods. *Cancer* 2017; 123(1): 114–121
 158. Carrell DS, Halgrim S, Tran DT, Buist DSM, Chubak J, Chapman WW, Savova G. Using natural language processing to improve efficiency of manual chart abstraction in research: the case of breast cancer recurrence. *Am J Epidemiol* 2014; 179(6): 749–758
 159. Xu H, Aldrich MC, Chen Q, Liu H, Peterson NB, Dai Q, Levy M, Shah A, Han X, Ruan X, Jiang M, Li Y, Julien JS, Warner J, Friedman C, Roden DM, Denny JC. Validating drug repurposing signals using electronic health records: a case study of metformin associated with reduced cancer mortality. *J Am Med Assoc* 2015; 22(1): 179–191
 160. Savova GK, Tseytlin E, Finan S, Castine M, Miller T, Medvedeva O, Harris D, Hochheiser H, Lin C, Chavan G, Jacobson RS. DeepPhe: a natural language processing system for extracting cancer phenotypes from clinical records. *Cancer Res* 2017; 77(21): e115–e118
 161. Savova GK, Danciu I, Alamudun F, Miller T, Lin C, Bitterman DS, Tourassi G, Warner JL. Use of natural language processing to extract clinical cancer phenotypes from electronic medical records. *Cancer Res* 2019; 79(21): 5463–5470
 162. Kehl KL, Elmarakeby H, Nishino M, Van Allen EM, Lepisto EM, Hassett MJ, Johnson BE, Schrag D. Assessment of deep natural language processing in ascertaining oncologic outcomes from radiology reports. *JAMA Oncol* 2019; 5(10): 1421–1429
 163. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I. Attention is all you need. In: *Advances in Neural Information Processing Systems*. 2017; 30
 164. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. *Nat Med* 2023; 29(8): 1930–1940
 165. Qiu X, Sun T, Xu Y, Shao Y, Dai N, Huang X. Pre-trained models for natural language processing: a survey. *Sci China Technol Sci* 2020; 63(10): 1872–1897
 166. Radford A, Narasimhan K, Salimans T, Sutskever I. Improving Language Understanding by Generative Pre-Training. 2018. Available at the website of [cdn.openai.com](https://openai.com)
 167. Yeo YH, Samaan JS, Ng WH, Ting PS, Trivedi H, Vipani A, Ayoub W, Yang JD, Liran O, Spiegel B, Kuo A. Assessing the performance of ChatGPT in answering questions regarding cirrhosis and hepatocellular carcinoma. *Clin Mol Hepatol* 2023; 29(3): 721–732
 168. Hopkins AM, Logan JM, Kichenadasse G, Sorich MJ. Artificial intelligence chatbots will revolutionize how cancer patients access information: ChatGPT represents a paradigm-shift. *JNCI Cancer Spectr* 2023; 7(2): pkad010
 169. Sorin V, Klang E, Sklair-Levy M, Cohen I, Zippel DB, Balint Lahat N, Konen E, Barash Y. Large language model (ChatGPT) as a support tool for breast tumor board. *NPJ Breast Cancer* 2023; 9(1): 44
 170. Johnson SB, King AJ, Warner EL, Aneja S, Kann BH, Bylund CL. Using ChatGPT to evaluate cancer myths and misconceptions: artificial intelligence and cancer information. *JNCI Cancer Spectr* 2023; 7(2): pkad015
 171. Chen S, Kann BH, Foote MB, Aerts HJWL, Savova GK, Mak RH, Bitterman DS. Use of artificial intelligence chatbots for cancer treatment information. *JAMA Oncol* 2023; 9(10): 1459–1462
 172. Coskun B, Ocakoglu G, Yetemen M, Kaygisiz O. Can ChatGPT, an artificial intelligence language model, provide accurate and high-quality patient information on prostate cancer? *Urology* 2023; 180: 35–58
 173. Sanderson K. GPT-4 is here: what scientists think. *Nature* 2023; 615(7954): 773
 174. Lyu Q, Tan J, Zapadka ME, Ponnatapura J, Niu C, Myers KJ, Wang G, Whitlow CT. Translating radiology reports into plain language using ChatGPT and GPT-4 with prompt learning: results, limitations, and potential. *Vis Comput Ind Biomed Art* 2023; 6(1): 9
 175. Fink MA, Bischoff A, Fink CA, Moll M, Kroschke J, Dulz L, Heußel CP, Kauczor HU, Weber TF. Potential of ChatGPT and

- GPT-4 for data mining of free-text CT reports on lung cancer. *Radiology* 2023; 308(3): e231362
176. Ibrahim H, Liu X, Denniston AK. Reporting guidelines for artificial intelligence in healthcare research. *Clin Exp Ophthalmol* 2021; 49(5): 470–476
 177. Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ; SPIRIT-AI and CONSORT-AI Working Group; SPIRIT-AI and CONSORT-AI Steering Group; SPIRIT-AI and CONSORT-AI Consensus Group. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Nat Med* 2020; 26(9): 1351–1363
 178. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK; SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nat Med* 2020; 26(9): 1364–1374
 179. Norgeot B, Quer G, Beaulieu-Jones BK, Torkamani A, Dias R, Gianfrancesco M, Arnaout R, Kohane IS, Saria S, Topol E, Obermeyer Z, Yu B, Butte AJ. Minimum information about clinical artificial intelligence modeling: the MI-CLAIM checklist. *Nat Med* 2020; 26(9): 1320–1324
 180. Hernandez-Boussard T, Bozkurt S, Ioannidis JPA, Shah NH. MINIMAR (MINimum Information for Medical AI Reporting): developing reporting standards for artificial intelligence in health care. *J Am Med Inform Assoc* 2020; 27(12): 2011–2015
 181. Sounderajah V, Ashrafian H, Golub RM, Shetty S, De Fauw J, Hooft L, Moons K, Collins G, Moher D, Bossuyt PM, Darzi A, Karthikesalingam A, Denniston AK, Mateen BA, Ting D, Treanor D, King D, Greaves F, Godwin J, Pearson-Stuttard J, Harling L, McInnes M, Rifai N, Tomasev N, Normahani P, Whiting P, Aggarwal R, Vollmer S, Markar SR, Panch T, Liu X; STARD-AI Steering Committee. Developing a reporting guideline for artificial intelligence-centred diagnostic test accuracy studies: the STARD-AI protocol. *BMJ Open* 2021; 11(6): e047709
 182. Collins GS, Dhiman P, Andaur Navarro CL, Ma J, Hooft L, Reitsma JB, Logullo P, Beam AL, Peng L, Van Calster B, van Smeden M, Riley RD, Moons KG. Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence. *BMJ Open* 2021; 11(7): e048008
 183. Sounderajah V, Ashrafian H, Rose S, Shah NH, Ghassemi M, Golub R, Kahn CE Jr, Esteva A, Karthikesalingam A, Mateen B, Webster D, Milea D, Ting D, Treanor D, Cushman D, King D, McPherson D, Glocker B, Greaves F, Harling L, Ordish J, Cohen JF, Deeks J, Leeftang M, Diamond M, McInnes MDF, McCradden M, Abramoff MD, Normahani P, Markar SR, Chang S, Liu X, Mallett S, Shetty S, Denniston A, Collins GS, Moher D, Whiting P, Bossuyt PM, Darzi A. A quality assessment tool for artificial intelligence-centered diagnostic test accuracy studies: QUADAS-AI. *Nat Med* 2021; 27(10): 1663–1665
 184. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, Denniston AK, Faes L, Geerts B, Ibrahim M, Liu X, Mateen BA, Mathur P, McCradden MD, Morgan L, Ordish J, Rogers C, Saria S, Ting DSW, Watkinson P, Weber W, Wheatstone P, McCulloch P; DECIDE-AI expert group. Reporting guideline for the early stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *BMJ* 2022; 377: e070904
 185. Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, Irwig L, Levine D, Reitsma JB, de Vet HC, Bossuyt PM. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open* 2016; 6(11): e012799
 186. Smith M, Sattler A, Hong G, Lin S. From code to bedside: implementing artificial intelligence using quality improvement methods. *J Gen Intern Med* 2021; 36(4): 1061–1066
 187. Wiens J, Saria S, Sendak M, Ghassemi M, Liu VX, Doshi-Velez F, Jung K, Heller K, Kale D, Saeed M, Ossorio PN, Thadaneysrani S, Goldenberg A. Do no harm: a roadmap for responsible machine learning for health care. *Nat Med* 2019; 25(9): 1337–1340
 188. Larson DB, Harvey H, Rubin DL, Irani N, Tse JR, Langlotz CP. Regulatory frameworks for development and evaluation of artificial intelligence-based diagnostic imaging algorithms: summary and recommendations. *J Am Coll Radiol* 2021; 18(3 Pt A): 413–424
 189. Gunning D, Stefik M, Choi J, Miller T, Stumpf S, Yang GZ. XAI-explainable artificial intelligence. *Sci Robot* 2019; 4(37): eaay7120
 190. Klein H, Mazor T, Siegel E, Trukhanov P, Ovalle A, Vecchio Fitz CD, Zwiesler Z, Kumari P, Van Der Veen B, Marriott E, Hansel J, Yu J, Albayrak A, Barry S, Keller RB, MacConaill LE, Lindeman N, Johnson BE, Rollins BJ, Do KT, Beardslee B, Shapiro G, Hector-Barry S, Methot J, Sholl L, Lindsay J, Hassett MJ, Cerami E. MatchMiner: an open-source platform for cancer precision medicine. *NPJ Precis Oncol* 2022; 6(1): 69
 191. Li T, Shetty S, Kamath A, Jaiswal A, Jiang X, Ding Y, Kim Y. CancerGPT for few shot drug pair synergy prediction using large pretrained language models. *NPJ Digit Med* 2024; 7(1): 40
 192. Hartman RI, Trepanowski N, Chang MS, Tepedino K, Gianacas C, McNiff JM, Fung M, Braghiroli NF, Grant-Kels JM. Multicenter prospective blinded melanoma detection study with a handheld elastic scattering spectroscopy device. *JAAD Int* 2024; 15: 24–31