

REVIEW

AI in spotting high-risk characteristics of medical imaging and molecular pathology

Chong Zhang^{1,2}, Jionghui Gu^{2,3}, Yangyang Zhu^{2,3}, Zheling Meng^{2,3}, Tong Tong^{2,3}, Dongyang Li^{2,3}, Zhenyu Liu^{2,3}, Yang Du^{2,3}, Kun Wang^{2,3,*} and Jie Tian^{2,3,4,*}

¹Department of Big Data Management and Application, School of International Economics and Management, Beijing Technology and Business University, Beijing 100048, China

²CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

³School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China

⁴Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, School of Medicine and Engineering, Beihang University, Beijing 100191, China

*Correspondence: Kun Wang, kun.wang@ia.ac.cn; Jie Tian, jie.tian@ia.ac.cn

Abstract

Medical imaging provides a comprehensive perspective and rich information for disease diagnosis. Combined with artificial intelligence technology, medical imaging can be further mined for detailed pathological information. Many studies have shown that the macroscopic imaging characteristics of tumors are closely related to microscopic gene, protein and molecular changes. In order to explore the function of artificial intelligence algorithms in in-depth analysis of medical imaging information, this paper reviews the articles published in recent years from three perspectives: medical imaging analysis method, clinical applications and the development of medical imaging in the direction of pathological molecular prediction. We believe that AI-aided medical imaging analysis will be extensively contributing to precise and efficient clinical decision.

Key words: machine learning; radiomics; medical imaging; pathology; artificial intelligence

Introduction

Medical imaging provides clinicians with comprehensive perspectives and rich information. It plays a vital role in disease screening, diagnosis, treatment selection and prognostic evaluation. The changes in the morphology or function of the lesion are determined by many factors such as the patient's individual genes, cells, physiological microenvironment, living habits and living

environment. Through data mining technology in conventional imaging diagnosis, the deep characteristics of diseases can be found, reflecting the changes of human tissues, cells and genes, which will have a significant promotion for clinical precision medicine. Since medical imaging can comprehensively, non-invasively and quantitatively observe the overall tumor morphology, and monitor the development process and treatment

Received: 24 October 2021; Revised: 26 November 2021; Accepted: 29 November 2021

© The Author(s) 2021. Published by Oxford University Press on behalf of the West China School of Medicine & West China Hospital of Sichuan University. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

response of the tumor at any time, it provides a reliable solution to the problem of tumor heterogeneity. Compared with traditional clinical medicine, which only interprets medical images from a visual level, radiomics can dig deeper into the biological features of images and provide clinical decision support.

Radiomics believes that the macroscopic imaging characteristics of tumors are closely related to microscopic gene, protein and molecular changes. Common medical imaging methods currently in clinical practice include computed tomography (CT), magnetic resonance imaging (MRI), ultrasound and positron emission computed tomography (PET). The above-mentioned imaging methods provide a wealth of disease information for clinical diagnosis. Among them, CT scans a thick layer with an X-ray beam and the detector can take a cross-sectional or three-dimensional image of the inspected part. MRI generates magnetic resonance phenomena by applying radio frequency pulses to the human body in a static magnetic field and realizes imaging through processes such as MR signal reception, spatial encoding, and image reconstruction. US scans the human body with ultrasonic sound beams and obtains images of internal organs by receiving and processing reflected signals. PET labels materials with short-lived radionuclides (^{18}F , ^{11}C , etc.), releasing positrons during the decay process and generating opposite photons. Captured by a highly sensitive photon camera and processed by a computer, a three-dimensional image of the aggregation in the organism can be obtained. Besides these non-invasive medical imaging methods, pathological examination is another important technique for clinical analysis and is the gold standard for judging the state of sampled tissues. Advances in equipment have made it possible to preserve and transmit the whole-slide pathological images, therefore promoting the application of artificial intelligent (AI) in pathological image analysis.

The large-scale application of AI technology provides an opportunity for in-depth analysis of medical images. The advantage of AI is that it can use more accurate and generalized models to capture subtle features, powerful classification and prediction capabilities, and provide the possibility to predict microscopic pathology for medical image analysis. AI combined with medical imaging further expands the analysis function of radiomics. Gradually, the macroscopic judgment of the lesion as simply benign or malignant is developing towards the microscopic prediction of molecular typing; the disease state analysis is developing in the direction of gene mutation prediction; the non-invasive imaging prediction is approaching the results of pathological analysis. AI and medical imaging together provide an effective way for precision medicine and is expected to provide a reliable auxiliary analysis method for clinical decision-making.

In the following, this paper will review the articles published in recent years from three perspectives: medical imaging analysis method, clinical applications and the development of medical imaging in the direction of pathological molecular prediction. The second part

focuses on the main methods of medical imaging analysis. In the third part, it introduces the application of CT, MRI, PET, Ultrasound and pathological examination, and analyzes the application scope and effect. The fourth part is a summary and prospect of the existing technology and application. It is hoped that this article can provide field overview and research ideas for researchers in related fields.

Methodology of medical imaging analysis

Methodological framework

The development of medical imaging analysis is based on the assumption that detailed information can be extracted, thereby providing useful diagnosis information for disease prediction.¹⁻³

Typical radiomics workflow includes four steps: 1) medical image acquisition; 2) segmentation, areas of interest (ROI) delineation; 3) data cleaning and data enhancement, feature extraction and selection; and 4) modeling and analysis (Fig. 1).⁴ The first important step is to obtain high-quality medical images with uniform standards. Ideally, image resolution, uniform collection parameters, imaging parameters, and others all need to be standardized.⁵⁻⁷ In the second step of feature division, target areas (ROI) need to be delineated. The method of segmentation can be manual, semi-automatic, or automated execution. The manual method needs to rely on experienced doctors to divide one by one, and the workload is large. Semi-automatic is a combination of manpower and machine. It requires experienced doctors to identify and modify the boundaries of automatic separation, saving a part of manpower. The automated method does not require human involvement and is more suitable for processing large data sets. The third step is feature extraction. Feature extraction can include shape and geometry features of the ROI area; texture features; intensity features, in which the density distribution reflects tumor heterogeneity. Medical imaging features can also be combined with clinical characteristics, such as clinical and pathological staging, etc. The fourth step of model construction and analysis is to establish a corresponding feasible model.

Traditional method

Traditional algorithms are often in three steps: medical image segmentation, feature extraction and model construction. Segmentation algorithm is mainly applied for automatic ROI delineation.

Feature extraction and selection directly affect the analysis results of the model. For clinical application, features are usually composed of two parts, including the clinical characteristics and the radiomics features (Table 1). The effect and repeatability of feature extraction from medical images affects the construction of radiomics model. The extracted features are mainly morphological characteristics, including volume, diameter,

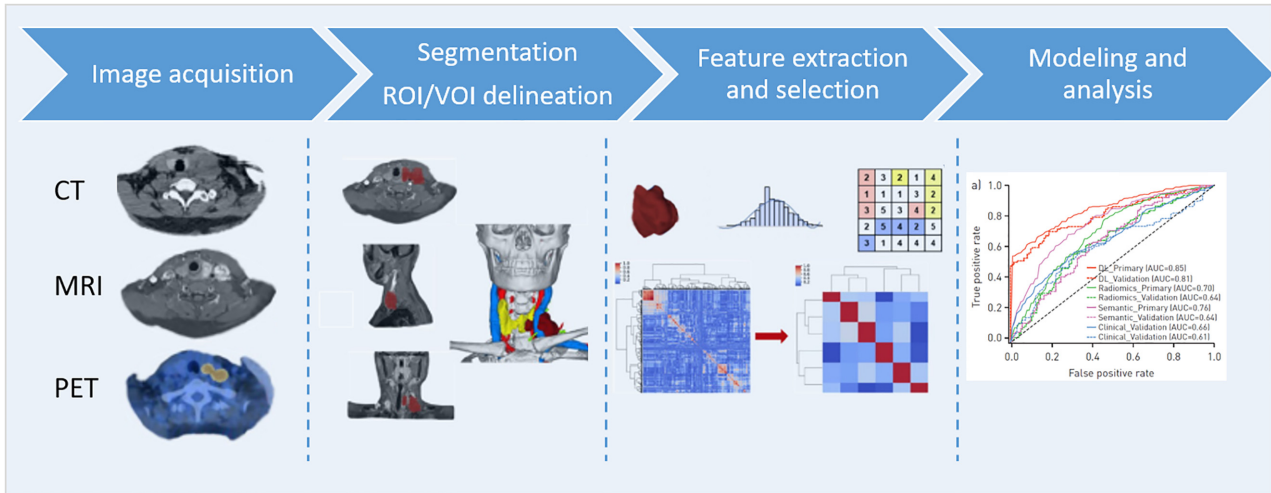


Figure 1. Typical radiomics workflow.

Table 1. Classification of clinical characteristics and radiomics features.

Classification of basic feature	Common features	Statistic feature
Clinical characteristics	Age, BMI, Sex, clinical dementia rating, histological type, clinical staging	Statistics and partition representation
Radiomics features ⁴	Volume, diameter, size, shape, location Texture	Histogram statistics Gray-level co-occurrence matrix (GLCM) Gray-level neighborhood difference matrix (GLNDM) Gray-level run length matrix (GLRLM) Gray-level size zone (GLSZM)
	Image feature	Fourier, Gabor, Wavelet and Laplacian transforms

texture and several transformation functions. The feature evaluation method can be statistics, which evaluates the distribution of gray values or the irregularity of the area; also based on transformation, and the spatial information is converted into frequency.⁸

The construction of medical imaging analysis model depends on the clinical problem to be solved and the function to be realized. For classification and prediction, the commonly used methods are statistical and machine learning methods. The commonly used machine learning methods include principal component analysis (PCA), decision tree, random forests (RF), Logistic regression, support vector machine (SVM), least absolute shrinkage and selection operator (LASSO) method, etc.⁹⁻¹³

Deep learning method

Convolutional neural network (CNN) has now been used extensively in medical image classification.¹⁴ The mostly commonly used models include LeNet,¹⁵ Alexnet,¹⁶ VGG19,¹⁷ GoogLeNet¹⁸ and ResNet.¹⁹

Nibali et al.²⁰ fine-tuned a pre-trained ResNet model to evaluate the classification performance of malignant tumors in lung nodules by using deep learning model. The results show that the combination of deep residual learning and transfer learning can achieve higher accuracy of nodule classification. Nishio et al.²¹ used VGG16 to extract features from CT images of lung nodules,

demonstrating that transfer learning methods outperformed hand-crafted features and traditional machine learning methods in lung cancer. Marentakis et al.²² found that in the classification of non-small cell lung cancer (NSCLC) into adenocarcinoma (AC) and squamous cell carcinoma (SCC), the use of CNN combined with long short-term memory (LSTM) network (CNN + LSTM accuracy = 0.74, AUC = 0.78) was more effective and would yield better results than the use of CNN only (the best CNN accuracy = 0.67, AUC = 0.74). An et al.²³ used CNN to extract high-risk features of images. Dual-energy computed tomography (DECT) is a new imaging technique that enables X-ray attenuation data to be obtained at two different energy levels, 100 and 150 keV. Their method uses CNN to extract image features at 100 keV, 150 keV and virtual monoenergetic images (VMI) at 40 keV, respectively; the extracted features were concatenated, and six models were built, including VMI 40 keV model, 100 keV model, 150 keV model, 100 + 150 keV model, 100 + 150 keV and clinical + 100 + 150 keV model. The dataset consisted of 148 patients and was divided into two groups. The first 113 patients were used to train the network and the remaining 35 patients were used to test the performance of the model. The results showed that the model combining 100 + 150 keV and clinical data predicted the most accurate results.

Deep learning has been applied in the clinical diagnosis of many tumors and has shown beneficial effect.

The AI method benefits from the technological advancement of computer hardware (graphics processing unit, GPU), and its computational timeliness is also sharply improved.^{24,25} At the same time, more and more open-source and effective calculation framework codes and pre-training models also provide more evolutionary steps for the algorithm iteration to the AI of medical images. In the work of Zheng et al., 584 breast cancer patients were enrolled, and ultrasound images were collected. The prediction of axillary lymph node status was realized by radiomics model based on deep learning, with prediction AUC of 0.902.²⁶ Marentakis et al. Enrolled 102 lung cancer patients using LSTM and Inception model for histological classification analysis, with an AUC of 0.78.²² To predict EGFR mutation status in pulmonary adenocarcinoma, Zhao et al. used deep learning analysis method based on 3D DenseNets and the AUC reached 0.75.²⁷ Similarly, for the prediction of EGFR, Wang et al. enrolled 844 patients with lung adenocarcinoma and improved AUC to 0.81 by using the self-built deep learning model.²⁸ This noninvasive prediction of EGFR is beneficial for targeted therapy. Yan et al. applied bayesian regularization neural networks to the prediction of IDH and TERT status with 357 glioma patients.²⁹ Burlingame et al. proposed a deep learning system named 'speedy histopathological-to-immunofluorescent translation (SHIFT)³⁰ which used cycle GAN algorithm³¹ to learn the spatial pattern of paired hematoxylin and eosin (H&E) and immunohistochemistry IHC images. The model is capable of making predictions of DAPI, α -SMA and PanCK distribution using H&E stained images as inputs. This work further demonstrated the capacity of artificial intelligence in histological analysis tasks.

Applications of medical imaging analysis

We here summarize the use of CT, MRI, PET, ultrasound and pathological examinations, as well as their realization of clinical application and molecular pathological analysis, in diagnosing tumors including breast cancer, hepatocellular tumor, lung cancer, glioma, and others (Fig. 2).

CT clinical application

Medical images are now used for a wide range of medical applications such as early diagnosis, detection and evaluate patients for treatment in a non-invasive manner.³² Among the many imaging modalities, the wide range and high speed of imaging are the characteristics of computed tomography (CT).³³ CT produces a cross-sectional image of the measurement subject by rotating the X-ray tube with a detector at its relative position to collect the X-ray, using the difference in the attenuation coefficient of the X-ray as they pass through different tissues.³⁴

In CT images, the grey level corresponds to the attenuation of X-rays and reflects the proportion of X-rays that are scattered or absorbed.³⁵ Given the wealth of detail contained in CT images, their interpretation requires an

experienced clinician. Medical staffs can benefit from computer-aided decision-making.³⁶ As a result, machine learning and deep learning related to image processing have been widely used in the field of medical imaging analysis (Table 2).^{37,38}

MRI clinical application

Recent studies have shown significant heterogeneity in tumor lesions, including variability between tumors and within the same tumor. MRI is a more specific and sensitive method for lesion identification and diagnosis among all available tools because of its higher image resolution. However, for the diagnosis and prediction of tumor at the molecular level, there are limitations in MRI images observed by the naked eye alone. Fortunately, radiomics is increasingly used in tumor diagnosis, prognosis, and treatment selection. The potential of radiomics in precision medicine practice is further enhanced by the ability to consistently acquire microscopic features of medical images that are not visible to the naked eye of human experts (Fig. 3).

The application of MRI radiomics in molecular level of breast cancer (BC) mainly refers to the prediction of molecular subtype. Hormonal status of BC was analyzed by immunohistochemistry (IHC) and divided into three molecular subgroups: HR+ and HER2-, HER2+, and triple negative.⁵³⁻⁵⁵ Various molecular subtypes have been shown to correlate with treatment planning and prognosis. Among them, IHC only analyzes localized tissue samples of breast cancer, which may not accurately represent the microscopic state of the entire tumor due to the complexity and heterogeneity of the tumor. However, molecular subtypes are confirmed by IHC analysis of sample tissues, which may not reflect the complexity and heterogeneity of the whole tumor. Recent studies have shown that radiomics is expected to be a new imaging label for identifying molecular subtypes of BC patients because of its good performance.⁵⁶⁻⁶² Lee et al.⁵⁷ found that textural parameters correlated with hormone receptor (HR), HER2 and Ki67 status, and molecular subtypes ($p < 0.002$). The status of HR and Ki67, grading and molecular subtypes were also correlated with perfusion parameters ($P < 0.003$). Bitencourt et al.⁵⁸ used three MRI parameters (two clinical, one radiomic) to achieve the prediction of HER2 burden, then they predicted whether BC patients with HER2+ who received neoadjuvant chemotherapy (NAC) would achieve a pathologically complete response (PCR). One study confirmed that texture features extracted from quantitative ADC map and DCE Map (Flush and Rinse) were able to identify triple negative BC (TNBC) by histogram analysis. The AUC of these models were 0.710 (Luminal A vs. TNBC), 0.763 (HER2+ vs. TNBC) and 0.683 (non-TNBC vs. TNBC), respectively.⁶³

The molecular-level prediction tasks of MRI radiomics on hepatocellular carcinoma (HCC) include predicting molecular states related to tumor and immunotherapy. Chen et al. proposed a machine learning model based on

Table 2. Clinical application of radiomics in molecular and pathological analysis.

Study	Number of patients	Tumor characteristic	Imaging modality	Function and Prediction results	Segmentation and feature selection method/model	Machine learning algorithm
Yang et al. ³⁹	467	lung adenocarcinoma (LADC)	CT	Predicting EGFR mutation AUC 0.789	Nodule segmentation 3D U-net model and pyradiomics	RF
Feng et al. ⁴⁰	300	Breast cancer	CT	Predicting triple negative breast cancer AUC 0.851	Manual segmentation and LASSO logistic method	Statistics
Ma et al. ⁴¹	140	Solid Lung Adenocarcinoma	CT	Predicting Anaplastic Lymphoma Kinase Gene Rearrangement AUC 0.801	Pearson correlation coefficient and ANOVA or RFE	SVM
Marentakis et al. ²²	102	Lung cancer	CT	Histological classification AUC 0.78	Joint FDG-PET and MRI prediction of lung metastases	LSTM + Inception
Zhang et al. ⁴²	420	lung adenocarcinoma	CT	Predicting EGFR mutation status AUC 0.835	LASSO and Wilcoxon test, DT, logistic regression	SVM
Wu et al. ⁴³	74	hepatocellular carcinoma	CT	Predicting the Ki-67 marker index AUC 0.75	Statistics	Logistic regression
Zhao et al. ²⁷	579	pulmonary adenocarcinoma	CT	Predicting EGFR mutation status AUC 0.75	Manual delineate	3D DenseNets
Li et al. ⁴⁴	207	colon cancer	CT	Predicting perineural invasion and KRAS mutation AUC 0.793 and 0.862	Manual delineate	SVM
Wang et al. ²⁸	844	lung adenocarcinoma	CT	Predicting EGFR mutation status AUC 0.81	A cubic ROI containing the entire tumour manual select	Deep learning model
Sutton et al. ⁴⁵	273	breast cancer	MRI	Classifying pathologic response post-neoadjuvant chemotherapy AUC 0.83	GMM GLMNet-RF-RFE	Statistics
Fan et al. ⁴⁶	144	Breast Cancer	MRI	Predicting histological grade and Ki-67 expression level AUC 0.814 and 0.810	Spatial fuzzy C-means algorithm refined by a Markov random field	Multitask learning method
Shofly et al. ⁴⁷	47	low-grade gliomas	MRI	1p/19q codeletion status prediction AUC 0.87	AnalyzeDirect software segmentation	Ensemble Bagged Trees

Table 2. Continued

Study	Number of patients	Tumor characteristic	Imaging modality	Function and Prediction results	Segmentation and feature selection method/model	Machine learning algorithm
Park et al. ⁴⁸	121	low-grade gliomas	MRI	Predicting molecular features of glioblastoma in Isocitrate Dehydrogenase Wild-Type AUC 0.863	Clinical feature + VASARI + radiomics feature	RFE SVM
Yan et al. ²⁹	357	glioma	MRI	Predicting IDH and TERT status AUC 0.884 and 0.669	Wavelet LASSO	Bayesian-regularization neural networks
Wu et al. ⁴⁹	126	diffuse gliomas	MRI	Predicting isocitrate dehydrogenase genotype AUC 0.931	Automated segmentation	RF
Braman et al. ⁵⁰	117	Breast cancer	MRI	Predicting pathological complete response to neoadjuvant chemotherapy AUC 0.74	A combined intratumoral and peritumoral radiomics approach	Cluster
Niu et al. ⁵¹	182	High-Grade Gliomas	MRI	Estimating the IDH1 Genotype AUC 0.86	Statistics	LASSO
Umutlu et al. ⁵²	124	Breast cancer	PET/MRI	Breast Cancer Phenotyping and Tumor Decoding	Statistics	LASSO
Zheng et al. ²⁶	584	Breast cancer	US	Predicting axillary lymph node status AUC 0.902	Deep learning radiomics model	Deep learning radiomics model

ANOVA, analysis of variance; DT, decision tree; RF, random forest; LASSO, least absolute shrinkage and selection operator; RFE, recursive feature elimination; SVM, support vector machine.

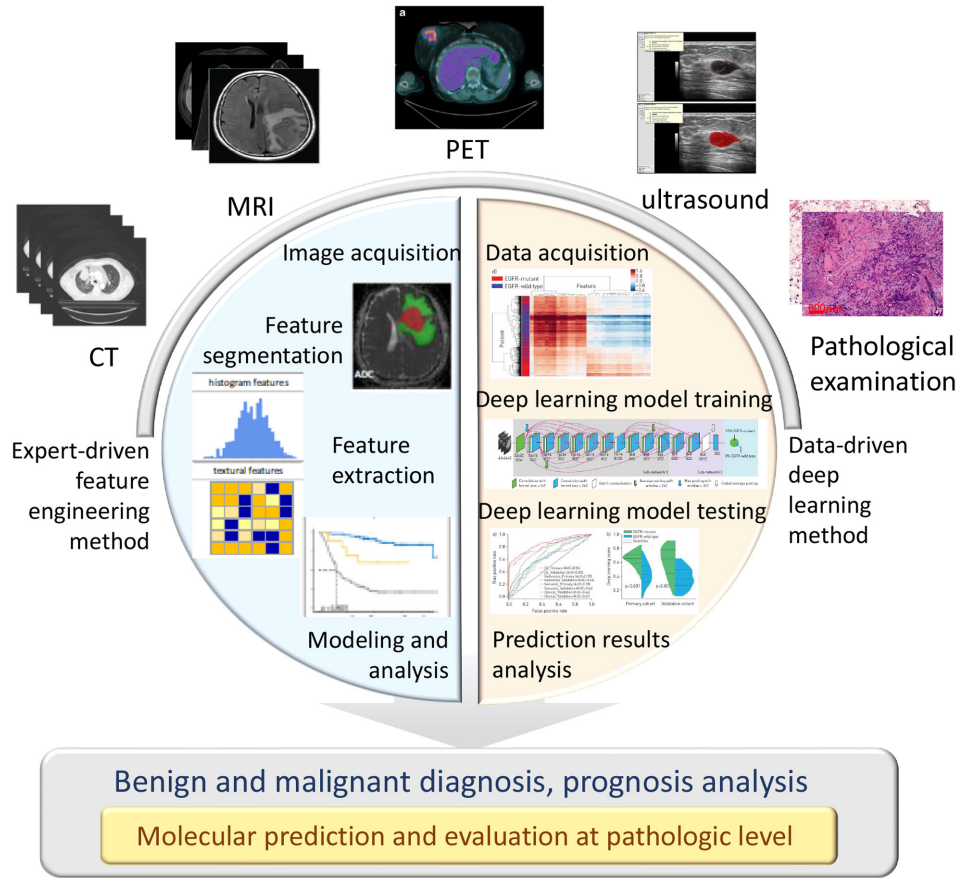


Figure 2. Medical imaging analysis model for pathological and molecular prediction.

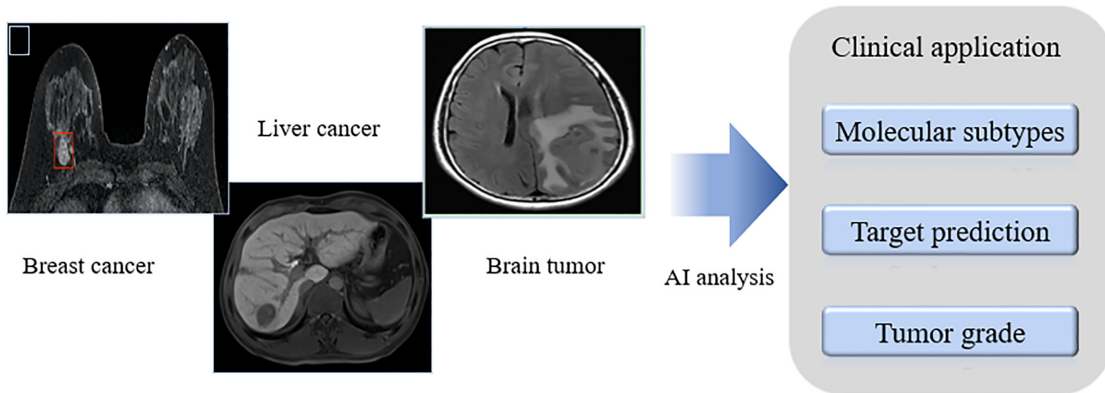


Figure 3. MRI radiomics applied at the molecular level of various tumors.

gadolinium-ethoxybenzyl-diethylenetriamine (Gd-EOB-DTPA)-enhanced MRI to predict the immunoscore related to the density of CD3+ and CD8+.⁶⁴ The AUC of this model based on the selected intratumoral and peritumoral radiomics features and clinical data is up to 0.926. They also illustrated that using combined radiomics features can obtain better predicting performance than only using intratumoral radiomics features. Gu *et al.* established a nomogram model based on ten radiomics features and a clinical characteristics (α -fetoprotein (AFP)) by multivariable logistic regression to predict the

Glypican 3 (GPC3) expression which is a biomarker associated with the prognosis of HCC patients.⁶⁵ The imaging sequence used in their research is contrast-enhanced T₁-weight MRI. This model achieved higher AUCs of 0.926 in training and 0.914 in validation cohorts compared with only using radiomics features. Wang *et al.* incorporated the level of AFP, enhancement patterns of tumor in the arterial phase, irregular margin of tumor, and 11 radiographic features extracted from gadoteric acid-enhanced MRI images into the final nomogram model to predict the status of cytokeratin (CK) 19 in

HCC patients.⁶⁶ The AUCs of this model for predicting CK 19 status can achieve 0.959 and 0.846 in training and validation cohorts. Hectors et al. explored the correlation between qualitative or quantitative radiomics features based on contrast-enhanced T1-weighted and diffusion-weighted images.⁶⁷ This study proved that radiomics features are correlated with immunohistochemical cells such as CD3, CD68 and CD31. At protein and mRNA expression level, the radiomics features also have correlation with PD-L1, PD1 and CTLA4.

Noninvasive preoperative grading and detection of key tumor markers in glioma are useful for surgeons. Recent studies have shown that MRI radiomics can extract micropathological features of gliomas from medical images, which may further understand some physiological behavior of gliomas. Studies have found that image fusion models combining radiomic features based on contrast-enhanced T1-weighted imaging (cT1WI) and apparent diffusion coefficient (ADC) achieved AUCs of 0.884 for status of IDH prediction and 0.669 for predicting TERT status. The cT1WI radiomic features performed well with AUC of 0.815 for 1p/19q status prediction.⁶⁸ In addition, this study confirmed that MRI-based radiomics can be used to noninvasively detect molecular populations and predict glioma survival regardless of glioma grade. These findings have been confirmed by other investigators.^{69–74} Yogananda et al.⁷⁵ proposed a new deep learning network named MGMT-net that can efficiently identify the status of MGMT promoter methylation based on T2-weighted images (T2WI), with good accuracy of 0.947 in the validation cohort confirmed by tri-fold cross-validation. Akbari et al.⁷⁶ constructed a study cohort of 129 patients with new glioblastoma. Then they obtained imaging labels of EGFRvIII from MRI images using radiomics techniques and achieved accuracy of 0.853 and 0.87 for identifying EGFRvIII mutated status in validation and test cohort, confirming that imaging signature of EGFRvIII can reveal a complex and unique macro glioblastoma phenotype.

One common childhood brain tumor is medulloblastoma (MB), which has a very high degree of malignancy. A study by Yan et al.⁷⁷ found that clinical and MRI imaging information from routine preoperative examinations could predict molecular subgroups of MB with high accuracy by using a machine learning algorithm. The model showed excellent predictive performance for wingless with an AUC of 0.9097 and accuracy of 0.8, and for sonic hedgehog with an AUC of 0.8654 and accuracy of 0.867. Recent studies have shown that some patients with astrocytoma, the most common type of glioma with a poor prognosis, have improved survival by responding well to temozolomide (TMZ) chemotherapy.^{78,79} The reason is that these patients were grade II-IV astrocytomas with methylation of the oxy-6-methylguanine-DNA methyltransferase.^{80,81} Wei et al.⁸² found that in the training and validation cohort, fusion radiomic features exhibited the highest ability in predicting methylation of MGMT promoter, with AUCs of 0.925 and 0.902 respectively. Additionally, MRI radiomics performed well

in predicting overall survival for patients who completed TMZ chemotherapy ($P = 0.003$ for high risk vs. low risk). Zormpas-Petridis et al.⁸³ revealed that MRI-based functional imaging can detect apoptotic responses to MYCN-targeted small-molecule inhibitors in a genetically engineered murine model of MYCN-driven neuroblastoma.

In summary, fusion radiomic features based on MRI images are important for predicting molecular subtype and prognostic analysis of lesions in various malignancies. Current studies have aimed to assess potential associations between tumor microscopic features and MRI radiomic features, but results have not yet been uniform. From the development of various radiomic models to their actual application in clinical practice (such as prediction of histopathological features), more and more extensive studies are necessary.

PET clinical application

PET is a functional imaging modality that noninvasively shows the metabolic processes of disease *in vivo*. It is widely applied in clinical practice for diagnosis, staging, assessment of therapeutic response, and prediction of gene mutations and prognosis. The introduction of radiomics has stimulated a new platform for non-invasive tissue characterization based on functional imaging. Some studies have predicted clinical outcomes based only on the radiomics of PET images. In the study of Li et al.,⁸⁴ radiological characteristics were extracted and evaluated from static images, early summary images, and dynamic ¹⁸F-FET-PET images to predict TERTp-mutation status in isocitrate dehydrogenase (IDH) gene-wild-type high-grade glioma patients, where recursive feature elimination and logistic regression are used; the study showed that the model based on dynamic ¹⁸F-FET-PET features could anticipate TERTp-mutation status (AUC 0.82, sensitivity 92.1%). In two other studies also targeting gliomas, they used different radioactive tracers for functional imaging. Qian et al.⁸⁵ demonstrated that radiological features of ¹⁸F-DOPA-PET imaging can be used to predict pathological O⁶-methylguanine methyltransferase status (accuracy 80%±10%). Li et al.⁸⁶ used support vector machines to generate ¹⁸F-FDG-PET imaging features combined with clinical features to predict the potential of IDH genotype status in patients with glioma, and verified the effectiveness of the model.

Since PET presents limited details on the morphology of lesions, it is often used in combination with CT (PET/CT, Fig. 4) or MRI (PET/MRI) to obtain both metabolic and structural information. Such multimodal imaging analysis shows good application results.

PET/CT-based radiomics applications

The prediction of specific biomarkers in lung cancer is a typical example of PET/CT-based radiomics applications. Mu et al. used small residual convolutional networks to analyze ¹⁸F-FDG-PET/CT images of 697 non-small cell lung cancer (NSCLC) patients from three

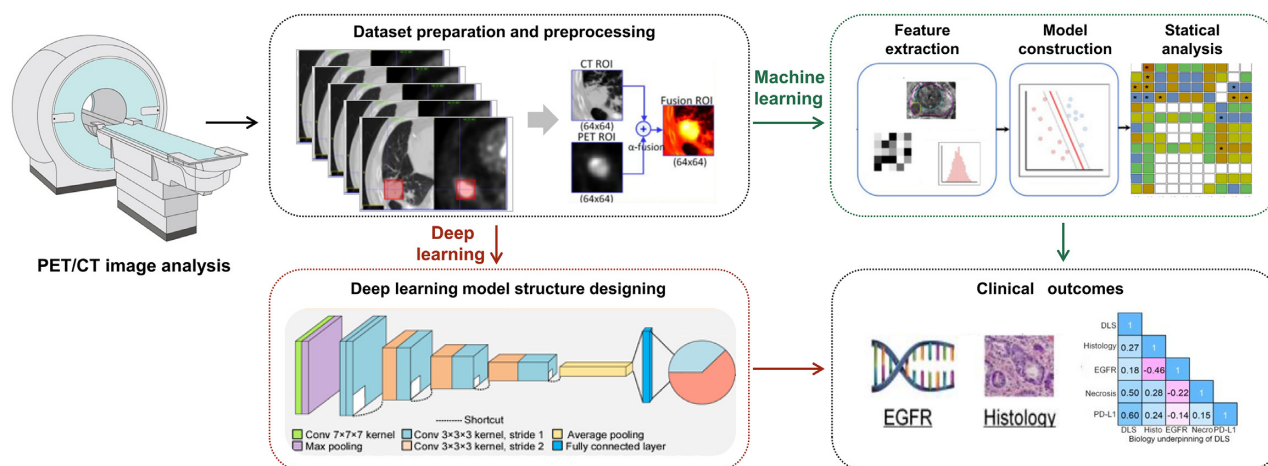


Figure 4. PET combined with CT for medical imaging analysis.

institutions and to develop deep learning scores to anticipate PD-L1 expression status. Their work was well validated in training, and testing cohorts ($AUC \geq 0.82$), which could be used to identify patients with immune checkpoint inhibitor sensitive NSCLC.⁸⁷ Similarly, Zhang *et al.*⁸⁸ used LASSO algorithm to select radiologic features that predict EGFR mutations based on ^{18}F -FDG PET/CT images from 248 NSCLC patients. They found that the best prediction for EGFR mutations was with the nomogram developed using a combination of radiological characteristics score and clinical variables ($AUC = 0.87$). While another study used AI software to extract radiological features from PET/CT images of patients with lung adenocarcinoma, 22 radiological features and 3 clinical features were selected to establish a model which predicted ALK rearrangement status ($AUC = 0.88$).⁸⁹ ^{18}F -PET/CT-based radiomics are also suitable for patients with NSCLC who are preparing for stereotactic body radiation therapy (SBRT), which can predict the potential of circulating tumor cells before and after SBRT to co-guide patients' subsequent treatment.⁹⁰ In addition, artificial intelligence radiomics has made new advances in other malignancies based on PET/CT images. Feature extraction (SUVmax, histogram parameters and texture features) was performed on ^{18}F -FDG PET images of 38 non-metastatic luminal breast cancer patients in one study and correlation of these extracted features with progesterone receptor and estrogen receptor expression was observed.⁹¹ While another study population was early-stage cervical cancer patients, Li *et al.* demonstrated that PET-CT-based radiomics integrating the primary tumor and perineural regions could predict E-cadherin.⁹² Different from the above studies with better predictive value, Saadani *et al.* explored the use of radiomic ^{18}F -FDG PET/CT features to predict B-rapidly accelerated fibrosarcoma valine 600 (BRAFV600) mutation status in melanoma patients and explored six different methods of feature selection. The results showed that the AUC for radiomics prediction of BRAFV600 mutation in melanoma patients ranged from 0.54 to 0.62 and are influenced by the feature selection method.⁹³

PET/MRI-based radiomics applications

A new trend in PET imaging technology introduced in the last few years is the PET-Magnetic Resonance Imaging (PET-MRI) system. This system uses MRI as an alternative to CT to eliminate additional radiation dose and can greatly improve the clarity of soft tissue imaging that CT tissue cannot provide. In Umutlu *et al.*'s study,⁵² the performance of synchronous ^{18}F -FDG PET/MRI as a comprehensive radiological platform for breast cancer subtype analysis, hormone receptor status, and proliferation rate was investigated. They used LASSO regression to select the most significant radiological characteristics in ^{18}F -FDG PET/MRI images from 124 breast cancer patients. Then, support vector machines were used for five-fold cross-validation to form a prediction model for the combination of various imaging data series. Finally, ^{18}F -FDG PET/MRI can be used to obtain morphological, functional and metabolic data simultaneously for comprehensive and high-quality radiomic analysis of breast cancer phenotype and tumor decoding. Zaragori *et al.*⁹⁴ used ^{18}F -FET PET-MRI radiomics to non-invasively predict IDH genotype, O⁶-methylguanine methyltransferase promoter methylation status and alpha thalassemia/mental retardation syndrome X-linked (ATRX) genotype with 1p/19q coding deletion, both of which showed good prediction.

Four-modalities model

Different from above, Matsui⁹⁵ diagnose the molecular subtype of lower-grade gliomas based deep learning (DL) and multi-modalities analysis. They designed a four-modalities (MRI, PET, CT and clinical information) deep learning model to learn and extract relative features automatically.

Ultrasound clinical application

Ultrasound imaging is a simple, flexible, low-cost and radiation-free imaging modality, especially suitable for imaging thyroid, breast and liver tissues. However, because ultrasound imaging has a lower resolution than

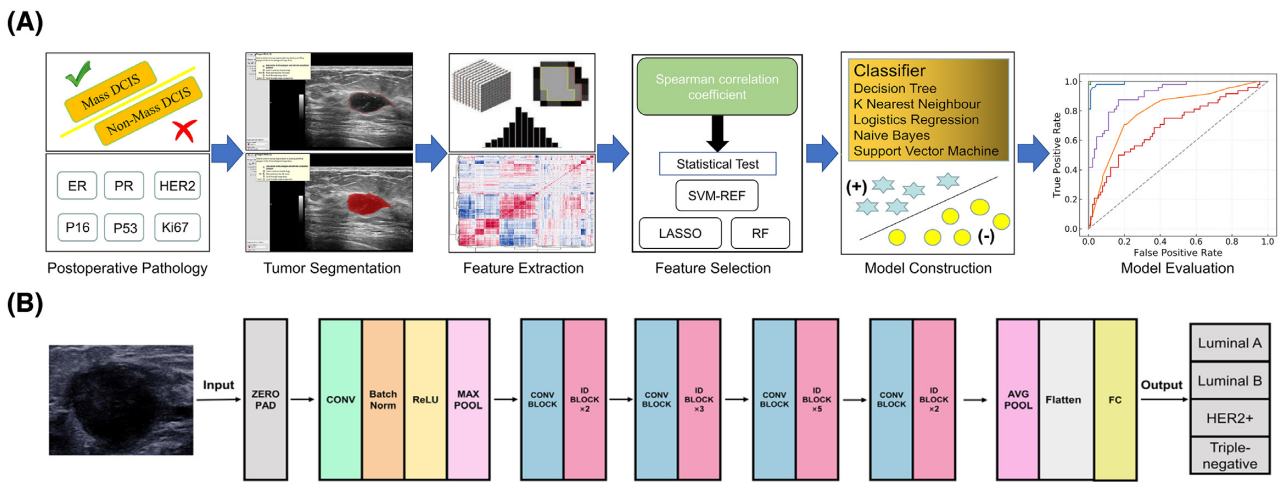


Figure 5. The workflow of (A) machine learning and (B) deep learning ultrasound radiomics for molecular subtype prediction.

other modalities and is greatly affected by the operator, the image quality is relatively poor and unstable, and the tumor boundary is often not particularly clear. These disadvantages often limit the accuracy of direct analysis of ultrasound images by radiologists. With the development of artificial intelligence, ultrasound radiomics has been widely used in diagnosis and prediction tasks with a good performance. In recent years, ultrasound radiomics has gradually involved cancer-related molecular prediction tasks. The application of ultrasound radiomics at the molecular level is mainly focused on the classification of breast cancer molecular subtypes. Machine learning is also an important technique for feature selection and model development (Fig. 5A). Cui *et al.* used the ultrasound features, according to the fifth edition of Breast Imaging Reporting and Data System (BI-RADS) and the elasticity ultrasound features specified by the WFUMB guidelines to construct Ki67 and P53 expression prediction models.⁹⁶ It was found that the high expression of Ki67 was associated with the loss of echo halo, posterior acoustic enhancement and high BI-RADS category. The high expression of P53 was correlated with the loss of echo halo and high classification of BI-RADS. The AUC of prediction models reached 0.78 for Ki67 and 0.71 for P53. The above research only used the clinical appearance features of ultrasound images, ignoring more valuable image features such as grayscale and texture features, thus the prediction accuracy is not high. Wu *et al.* proposed six machine learning models with a series of handcraft image features that are able to predict the expression of multiple breast ductal carcinoma-related molecules including ER, PR, HER2, Ki67, P16 and P53.⁹⁷ The features used in this study contained 5234 image features with mathematical meaning, and different machine learning models for feature selection were used for different molecular prediction tasks. In the end, the AUC of the test cohort was up to 0.84 for ER, 0.78 for PR, 0.74 for HER2, 0.86 for Ki67, 0.78 for P16 and 0.74 for P53 expression prediction. Recently, more and more studies have applied

deep learning to the prediction of breast cancer molecular subtypes. Compared with machine learning methods based on manual selected features, deep learning model can adaptively extract high-level advanced features of images that are highly correlated with these molecules through data-driven learning (Fig. 5B). In the case of big data, the prediction performance of molecular expression can be further improved by using deep learning models. Zhang *et al.* proposed one deep learning model for molecular subtype diagnosis.⁹⁸ This study contains a multicenter and big dataset including more than 3 000 ultrasound images and 2 000 patients to train and test the model. The model achieved high performance in identifying different molecular subtypes of breast cancer. The AUCs of the test cohort was 0.864 for identifying triple-negative breast cancer, 0.811 for HER2+ and 0.837 for HR+ breast cancer. Similarly, Jiang *et al.* also developed the deep learning model for breast cancer molecular subtype assessment based on a multicenter dataset with more than 4 828 ultrasound images and 1 275 patients.⁹⁹ This model achieved higher prediction performance, identifying the four breast cancer molecular subtypes including Lumina A, Lumina B, HER2+ and triple-negative, with accuracy ranging from 80.07% to 97.02% and 87.94% to 98.83% in two test cohorts, respectively. Furthermore, this study also proposed an additional deep learning model for distinguishing luminal disease from non-luminal disease. The positive predictive value on the two test cohorts reached 93.29% and 88.21%, respectively. Zheng *et al.* proposed a model by using deep learning method for predicting axillary lymph node status in early-stage breast cancer, with AUC value of 0.902 for lymph node metastasis.²⁶

Pathological examination

Over the past few decades, faster computation capacity and cheaper storage have enabled digital pathology to gain rapid development.^{100–105} Pathologists can study digital pathological images more easily and flexibly than

Table 3. Application of AI in pathological examination.

Application category	Application examples
Diagnosis	Benign and malignant diagnosis ¹¹³ Tumor aggressiveness classification ^{114,115} Tumor differentiation prediction ¹¹⁶
Efficacy prediction	Tumor regression grade of neoadjuvant therapy prediction ¹¹⁷
Gene prediction	Classification and mutation prediction from non-small cell lung cancer ¹¹⁸ Microsatellite instability prediction ²⁵
Biomarker prediction	Prediction of fluorescence label distribution in unlabeled images ¹⁰⁹ Predictions of DAPI, α -SMA and PanCK distribution ³¹
Prognosis prediction	Predictions of the distribution of PanCK, DAPI, CD3 and CD20 biomarkers ^{111,112} Survival prediction ¹¹⁹⁻¹²³

in the past. The past 20 years have seen the beginning and development of digital imaging in pathology, which allows whole slides to be imaged and stored permanently at high resolution. These progresses have facilitated the application of deep learning-based methods in pathological examination analysis. AI is an emerging approach and has been widely applied in medical image analysis, such as nuclei segmentation,¹⁰⁶ cancer diagnoses^{103,107} and cancer subtype classification.¹⁰⁸

Christiansen *et al.* proposed a method based on deep learning to predict the distribution of fluorescent markers in unlabeled images by using a deep neural network that trains unlabeled and labeled images in pairs.¹⁰⁹ The results showed that the fluorescence labeling algorithm based on deep learning can be used to predict the fluorescence labeling of transmitted light images. This demonstrates the potential of the DL-based approach in histological image synthesis tasks. Diao *et al.* proposed a computational pathology process that integrated high-resolution cell-level information from whole-slide images to predict molecular derived phenotypes for five different cancer types.¹¹⁰ The proposed approach combines the deep learning method with the interpretability of human interpretable features. It can integrate prior knowledge and achieves the performance of an end-to-end model. Bian *et al.* utilized convolutional neural networks to learn the spatial patterns incorporated within H&E images and make predictions of the distribution of PanCK, DAPI, CD3 and CD20 biomarkers.^{111,112} The proposed works showed that artificial intelligence can be used in biomarker distribution prediction tasks, which can help aid pathologists in image analysis. The applications of AI in pathological examination are shown in Table 3.

Artificial intelligence demonstrates exciting potential in the challenge of digital pathological image quantitative analysis and prediction model construction. The method can be summarized into expert-driven feature engineering methods and data-driven deep learning methods.^{124,125} The expert-driven feature engineering method usually includes three steps: artificially defined pathology image feature extraction, feature selection, and modeling.^{119,120} The extracted features are used to quantitatively describe cell-level events such as cell mitosis, multi-nucleation process, and cell subtypes, which was further for the description of spatial

architecture and arrangement of cells. After these features are screened, the features are integrated through machine learning, and models are constructed for specific prediction tasks. The data-driven deep learning method employs convolutional neural networks (CNN)¹²⁶ to predict the patch of whole slide image (WSI) end-to-end, and then integrates the prediction results of multiple patches through a voting machine to obtain patient-level prediction results based on several patch-level results.^{118,121} Compared with expert-driven feature engineering method, data-driven deep learning method reduces the dependence on prior pathological knowledge and automatically learns meaningful features from data. In this way, feature design iterations and repeated expert discussions are reduced greatly.^{114,124,127} Currently, AI-based quantitative analysis methods of pathological images have been successfully tried in clinical needs such as diagnosis,^{113-116,127} efficacy prediction,¹¹⁷ gene prediction,^{25,118} and prognosis prediction^{119,121,122} of multiple cancers, and have shown good patient benefits. The cancer types involved include breast cancer, prostate cancer, rectal cancer, and many others. The clinical challenges that have been solved involve benign and malignant diagnosis,¹¹³ tumor aggressiveness classification,^{114,115} tumor differentiation prediction,¹¹⁶ microsatellite instability prediction,²⁵ tumor regression grade of neoadjuvant therapy prediction¹¹⁷ and survival prediction.¹¹⁹⁻¹²³ The AI showed significant improvement in the evaluation of classification accuracy, sensitivity, specificity, and other indicators.²⁴ The combination of pathologists' diagnosis results and AI methods can effectively improve the consistency of diagnosis among doctors of different experiences and achieve better performance than doctors in some specific tumor problems.^{114,115} Hopefully, AI is able to provide an effective micro-information reference for precise and personalized treatment, thereby improving patient benefits.

Discussion and conclusion

Medical imaging provides an important source of clinical information for medical diagnosis, efficacy evaluation, and patient prognosis and survival prediction. With the development of machine learning, artificial intelligent technology can extract key information, which is

vulnerable to be missed by human observation, from a large amount of data. The combination of artificial intelligence and medical imaging provides a possible technical means for precision medical analysis. Especially correlation studies have found that pathological and molecular level information is highly correlated with medical imaging features.^{128–130} This allows medical imaging to provide the possibility of molecular-level biological feature diagnosis.

This article mainly analyzes the current research status of medical imaging technology combined with artificial intelligence, especially the systematic research on this method to explore molecular pathology information. The survey results have also proved the feasibility of radiomics to analyze molecular pathology.

However, there are also some limitations and challenges, which require continuous research in the future to further overcome the deficiencies of medical imaging. Specifically, for data acquisition, the protocol and instrument changes between different institutions will reduce the robustness of the radiomics model. Data sharing between institutions involves challenges in patient privacy, which may be a limiting factor in the development of a unified model. This requires standard imaging protocols, repeatable and consistent image processing, and collaboration between agencies to create large amounts of annotated datasets. In addition, it is necessary to design robust imaging features,¹³¹ and specify unified algorithm design standards and evaluation standards, which will help establish a reliable radiomic model. Furthermore, studies have shown that radiomics, when combined with histopathology, genomics or molecular classification, and immunophenotype, can achieve more precise results in predicting patient prognostic characteristics.¹³²

In the future development, medical imaging analysis will focus on data sharing, carefully designed, forward-looking and comprehensive research. It has great potential in helping understanding of the *in vivo* physiology of tumors and provides opportunities for optimizing patient clinical care.¹³³ By incorporating more data and making the model more robust, the increasing prediction accuracy of medical imaging analysis will continue to contribute to personalized medicine. Real-time predictive analysis can be obtained from large semi-automatic patient data sets and electronic medical records to provide insights into various disease processes. In conclusion, in the new era of precision medicine, AI-aided medical imaging analysis will be extensively contributing to clinical decision.

Acknowledgements

This paper is supported by the Ministry of Science and Technology of China (Grant No. 2017YFA0205200), the National Natural Science Foundation of China (Grants No. 62027901, 81227901, and 81930053), Chinese Academy of Sciences (Grant No. QYZDJ-SSW-JSC005), the

Excellent Member Project of Youth Innovation Promotion Association CAS, and the Project of High-Level Talents Team Introduction in Zhuhai City (Zhuhai HLH-PTP201703). The authors would like to acknowledge the technical support of Multimodal Biomedical Imaging Experimental Platform, Institute of Automation, Chinese Academy of Sciences.

Conflict of interest

The authors declare no conflicts of interest. Besides, as an Editorial Board Member of *Precision Clinical Medicine*, the corresponding author Jie Tian was blinded from reviewing and making decision on this manuscript.

References

1. Zwanenburg A, Vallières M, Abdalah MA, et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology* 2020;**295**:328–38. doi:10.1148/radiol.2020191145.
2. Sun R, Limkin EJ, Vakalopoulou M, et al. A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study. *Lancet Oncol* 2018;**19**:1180–91. doi:10.1016/S1470-2045(18)30413-3.
3. Pinker K, Chin J, Melsaether AN, et al. Precision medicine and radiogenomics in breast cancer: new approaches toward diagnosis and treatment. *Radiology* 2018;**287**:732–47. doi:10.1148/radiol.2018172171.
4. Horvat N, Bates DD, Petkovska I. Novel imaging techniques of rectal cancer: what do radiomics and radiogenomics have to offer? A literature review. *Abdom Radiol* 2019;**44**:3764–74. doi:10.1007/s00261-019-02042-y.
5. Papp L, Rausch I, Grahovac M, et al. Optimized feature extraction for radiomics analysis of 18F-FDG PET imaging. *J Nucl Med* 2019;**60**:864–72. doi:10.2967/jnumed.118.217612.
6. Liu Z, Zhang X-Y, Shi Y-J, et al. Radiomics analysis for evaluation of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. *Clin Cancer Res* 2017;**23**:7253–62. doi:10.1158/1078-0432.CCR-17-1038.
7. Dalal V, Carmicheal J, Dhaliwal A, et al. Radiomics in stratification of pancreatic cystic lesions: Machine learning in action. *Cancer Lett* 2020;**469**:228–37. doi:10.1016/j.canlet.2019.10.023.
8. Lubner MG, Smith AD, Sandrasegaran K, et al. CT texture analysis: definitions, applications, biologic correlates, and challenges. *Radiographics* 2017;**37**:1483–503. doi:10.1148/rg.2017170056.
9. Lu C-F, Hsu F-T, Hsieh KLC, et al. Machine learning-based radiomics for molecular subtyping of gliomas. *Clin Cancer Res* 2018;**24**:4429–36. doi:10.1158/1078-0432.CCR-17-3445.
10. Park YW, Oh J, You SC, et al. Radiomics and machine learning may accurately predict the grade and histological subtype in meningiomas using conventional and diffusion tensor imaging. *Eur Radiol* 2019;**29**:4068–76. doi:10.1007/s00330-018-5830-3.
11. Gu Q, Feng Z, Liang Q, et al. Machine learning-based radiomics strategy for prediction of cell proliferation in non-small cell lung cancer. *Eur J Radiol* 2019;**118**:32–7. doi:10.1016/j.ejrad.2019.06.025.

12. Ji G-W, Zhu F-P, Xu Q, et al. Machine-learning analysis of contrast-enhanced CT radiomics predicts recurrence of hepatocellular carcinoma after resection: A multi-institutional study. *EBioMedicine* 2019;**50**:156–65. doi:10.1016/j.ebiom.2019.10.057.
13. Mukherjee P, Cintra M, Huang C, et al. CT-based radiomic signatures for predicting histopathologic features in head and neck squamous cell carcinoma. *Radiol Imaging Cancer* 2020;**2**:e190039. doi:10.1148/rycan.2020190039.
14. Anwar SM, Majid M, Qayyum A, et al. Medical image analysis using convolutional neural networks: a review. *J Med Syst* 2018;**42**:226. doi:10.1007/s10916-018-1088-1.
15. LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition. *Proc IEEEProc IEEE* 1998;**86**:2278–324. doi:10.1109/5.726791.
16. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Commun ACM* 2017;**60**:84–90. doi:10.1145/3065386.
17. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.
18. Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1–9.
19. He K, Zhang X, Ren S, et al. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–8.
20. Nibali A, He Z, Wollersheim D. Pulmonary nodule classification with deep residual networks. *Int J Comput Assist Radiol Surg* 2017;**12**:1799–808. doi:10.1007/s11548-017-1605-6.
21. Nishio M, Sugiyama O, Yakami M, et al. Computer-aided diagnosis of lung nodule classification between benign nodule, primary lung cancer, and metastatic lung cancer at different image size using deep convolutional neural network with transfer learning. *PLoS One* 2018;**13**:e0200721. doi:10.1371/journal.pone.0200721.
22. Marentakis P, Karaiskos P, Kouloulis V, et al. Lung cancer histology classification from CT images based on radiomics and deep learning models. *Med Biol Eng Comput* 2021;**59**:215–26. doi:10.1007/s11517-020-02302-w.
23. An C, Li D, Li S, et al. Deep learning radiomics of dual-energy computed tomography for predicting lymph node metastases of pancreatic ductal adenocarcinoma. *Eur J Nucl Med Mol Imaging* 2021. doi:10.1007/s00259-021-05573-z.
24. Bera K, Schalper KA, Rimm DL, et al. Artificial intelligence in digital pathology—new tools for diagnosis and precision oncology. *Nat Rev Clin Oncol* 2019;**16**:703–15. doi:10.1038/s41571-019-0252-y.
25. Kather JN, Pearson AT, Halama N, et al. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat Med* 2019;**25**:1054–6. doi:10.1038/s41591-019-0462-y
26. Zheng X, Yao Z, Huang Y, et al. Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. *Nat Commun* 2020;**11**:1236. doi:10.1038/s41467-020-15027-z.
27. Zhao W, Yang J, Ni B, et al. Toward automatic prediction of EGFR mutation status in pulmonary adenocarcinoma with 3D deep learning. *Cancer Med* 2019;**8**:3532–43. doi:10.1002/cam4.2233.
28. Wang S, Shi J, Ye Z, et al. Predicting EGFR mutation status in lung adenocarcinoma on computed tomography image using deep learning. *Eur Respir J* 2019;**53**:1800986. doi:10.1183/13993003.00986-2018.
29. Yan J, Zhang B, Zhang S, et al. Quantitative MRI-based radiomics for noninvasively predicting molecular subtypes and survival in glioma patients. *NPJ Precis Oncol* 2021;**5**:72. doi:10.1038/s41698-021-00205-z.
30. Burlingame EA, Margolin AA, Gray JW, et al. SHIFT: speedy histopathological-to-immunofluorescent translation of whole slide images using conditional generative adversarial networks. In *Medical Imaging 2018: Digital Pathology*, International Society for Optics and Photonics, 2018, 1058105.
31. Zhu J-Y, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 2223–32.
32. Kasban H, El-Bendary M, Salama D. A comparative study of medical imaging techniques. *Int J Inform Sci Intell Syst* 2015;**4**:37–58.
33. Domingues I, Pereira G, Martins P, et al. Using deep learning techniques in medical imaging: a systematic review of applications on CT and PET. *Artif Intell Rev* 2020;**53**:4093–160. doi:10.1007/s10462-019-09788-3.
34. Bushberg JT, Boone JM. *The Essential Physics of MEDICAL IMAGING*. Lippincott Williams & Wilkins, 2011.
35. McCollough CH, Leng S, Yu L, et al. Dual-and multi-energy CT: principles, technical approaches, and clinical applications. *Radiology* 2015;**276**:637–53. doi:10.1148/radiol.2015142631.
36. Chen G-M, Chou Y-H, Tagawa N, et al. Computer-aided detection and diagnosis in medical imaging. *Comput Math Methods Med* 2013;**2013**:790608. doi:10.1155/2013/790608.
37. Shen D, Wu G, Suk H-I. Deep learning in medical image analysis. *Annu Rev Biomed Eng* 2017;**19**:221–48. doi:10.1146/annurev-bioeng-071516-044442.
38. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal* 2017;**42**:60–88. doi:10.1016/j.media.2017.07.005.
39. Yang X, Dong X, Wang J, et al. Computed tomography-based radiomics signature: a potential indicator of epidermal growth factor receptor mutation in pulmonary adenocarcinoma appearing as a subsolid nodule. *Oncologist* 2019;**24**:e1156–64. doi:10.1634/theoncologist.2018-0706.
40. Feng Q, Hu Q, Liu Y, et al. Diagnosis of triple negative breast cancer based on radiomics signatures extracted from preoperative contrast-enhanced chest computed tomography. *BMC Cancer* 2020;**20**:579. doi:10.1186/s12885-020-07053-3.
41. Ma D-N, Gao X-Y, Dan Y-B, et al. Evaluating solid lung adenocarcinoma anaplastic lymphoma kinase gene rearrangement using noninvasive radiomics biomarkers. *OncoTargets Ther* 2020;**13**:6927–35. doi:10.2147/OTT.S257798.
42. Zhang G, Cao Y, Zhang J, et al. Predicting EGFR mutation status in lung adenocarcinoma: development and validation of a computed tomography-based radiomics signature. *Am J Cancer Res* 2021;**11**:546–60.
43. Wu H, Han X, Wang Z, et al. Prediction of the Ki-67 marker index in hepatocellular carcinoma based on CT radiomics features. *Phys Med Biol* 2020;**65**:235048. doi:10.1088/1361-6560/abac9c.
44. Li Y, Eresen A, Shangguan J, et al. Preoperative prediction of perineural invasion and KRAS mutation in colon cancer using machine learning. *J Cancer Res Clin Oncol* 2020;**146**:3165–74. doi:10.1007/s00432-020-03354-z.
45. Sutton EJ, Onishi N, Fehr DA, et al. A machine learning model that classifies breast cancer pathologic complete

- response on MRI post-neoadjuvant chemotherapy. *Breast Cancer Res* 2020;**22**:57. doi:10.1186/s13058-020-01291-w.
46. Fan M, Yuan W, Zhao W, et al. Joint prediction of breast cancer histological grade and Ki-67 expression level based on DCE-MRI and DWI radiomics. *IEEE J Biomed Health Inform* 2020;**24**:1632–42. doi:10.1109/JBHI.2019.2956351.
 47. Shofty B, Artzi M, Bashat DB, et al. MRI radiomics analysis of molecular alterations in low-grade gliomas. *Int J Comput Assist Radiol Surg* 2018;**13**:563–71. doi:10.1007/s11548-017-1691-5.
 48. Park C, Han K, Kim H, et al. MRI features may predict molecular features of glioblastoma in isocitrate dehydrogenase wild-type lower-grade gliomas. *Am J Neuroradiol* 2021;**42**:448–56. doi:10.3174/ajnr.A6983.
 49. Wu S, Meng J, Yu Q, et al. Radiomics-based machine learning methods for isocitrate dehydrogenase genotype prediction of diffuse gliomas. *J Cancer Res Clin Oncol* 2019;**145**:543–50. doi:10.1007/s00432-018-2787-1.
 50. Braman NM, Etesami M, Prasanna P, et al. 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**:57. doi:10.1186/s13058-017-0846-1.
 51. Niu L, Feng W-H, Duan C-F, et al. The value of enhanced MR radiomics in estimating the IDH1 genotype in high-grade gliomas. *Biomed Res Int* 2020;**2020**:4630218. doi:10.1155/2020/4630218.
 52. Umutlu L, Kirchner J, Bruckmann NM, et al. Multiparametric integrated 18F-FDG PET/MRI-based radiomics for breast cancer phenotyping and tumor decoding. *Cancers* 2021;**13**:2928. doi:10.3390/cancers13122928.
 53. Perou CM, Sørlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature* 2000;**406**:747–52. doi:10.1038/35021093.
 54. Blows FM, Driver KE, Schmidt MK, et al. Subtyping of breast cancer by immunohistochemistry to investigate a relationship between subtype and short and long term survival: a collaborative analysis of data for 10,159 cases from 12 studies. *PLoS Med* 2010;**7**:e1000279. doi:10.1371/journal.pmed.1000279.
 55. Perou CM, Børresen-Dale AL. Systems biology and genomics of breast cancer. *Cold Spring Harb Perspect Biol* 2011;**3**:a003293. doi:10.1101/cshperspect.a003293.
 56. Choudhery S, Gomez-Cardona D, Favazza CP, et al. MRI radiomics for assessment of molecular subtype, pathological complete response, and residual cancer burden in breast cancer patients treated with neoadjuvant chemotherapy. *Acad Radiol* 2020;**S1076-6332(20)**30607-3. doi:10.1016/j.acra.2020.10.020.
 57. Lee JY, Lee KS, Seo BK, et al. Radiomic machine learning for predicting prognostic biomarkers and molecular subtypes of breast cancer using tumor heterogeneity and angiogenesis properties on MRI. *Eur Radiol* 2021. doi:10.1007/s00330-021-08146-8.
 58. Bitencourt AGV, Gibbs P, Rossi Saccarelli C, et al. MRI-based machine learning radiomics can predict HER2 expression level and pathologic response after neoadjuvant therapy in HER2 overexpressing breast cancer. *EBioMedicine* 2020;**61**:103042. doi:10.1016/j.ebiom.2020.103042.
 59. Zhang Y, Chen JH, Lin Y, et al. Prediction of breast cancer molecular subtypes on DCE-MRI using convolutional neural network with transfer learning between two centers. *Eur Radiol* 2021;**31**:2559–67. doi:10.1007/s00330-020-07274-x.
 60. Saha A, Harowicz MR, Grimm LJ, et al. A machine learning approach to radiogenomics of breast cancer: a study of 922 subjects and 529 DCE-MRI features. *Br J Cancer* 2018;**119**:508–16. doi:10.1038/s41416-018-0185-8.
 61. Li H, Zhu Y, Burnside ES, et al. Quantitative MRI radiomics in the prediction of molecular classifications of breast cancer subtypes in the TCGA/TCIA data set. *NPJ Breast Cancer* 2016;**2**:16012. doi:10.1038/npjbcancer.2016.12.
 62. Sutton EJ, Dashevsky BZ, Oh JH, et al. Breast cancer molecular subtype classifier that incorporates MRI features. *J Magn Reson Imaging* 2016;**44**:122–9. doi:10.1002/jmri.25119.
 63. Ma W, Ji Y, Qi L, et al. Breast cancer Ki67 expression prediction by DCE-MRI radiomics features. *Clin Radiol* 2018;**73**:909.e901–5. doi:10.1016/j.crad.2018.05.027.
 64. Chen S, Feng S, Wei J, et al. Pretreatment prediction of immunoscore in hepatocellular cancer: a radiomics-based clinical model based on Gd-EOB-DTPA-enhanced MRI imaging. *Eur Radiol* 2019;**29**:4177–87. doi:10.1007/s00330-018-5986-x.
 65. Gu D, Xie Y, Wei J, et al. MRI-based radiomics signature: a potential biomarker for identifying glypican 3-positive hepatocellular carcinoma. *J Magn Reson Imaging* 2020;**52**:1679–87. doi:10.1002/jmri.27199.
 66. Wang W, Gu D, Wei J, et al. A radiomics-based biomarker for cytokeratin 19 status of hepatocellular carcinoma with gadoteric acid-enhanced MRI. *Eur Radiol* 2020;**30**:3004–14. doi:10.1007/s00330-019-06585-y.
 67. Hectors SJ, Lewis S, Besa C, et al. MRI radiomics features predict immuno-oncological characteristics of hepatocellular carcinoma. *Eur Radiol* 2020;**30**:3759–69. doi:10.1007/s00330-020-06675-2.
 68. Yan J, Zhang B, Zhang S, et al. Quantitative MRI-based radiomics for noninvasively predicting molecular subtypes and survival in glioma patients. *NPJ Precis Oncol* 2021;**5**:72. doi:10.1038/s41698-021-00205-z.
 69. Ali MB, Gu IY, Berger MS, et al. Domain mapping and deep learning from multiple MRI clinical datasets for prediction of molecular subtypes in low grade gliomas. *Brain Sci* 2020;**10**:463. doi:10.3390/brainsci10070463.
 70. Zhang X, Tian Q, Wang L, et al. Radiomics strategy for molecular subtype stratification of lower-grade glioma: detecting IDH and TP53 mutations based on multimodal MRI. *J Magn Reson Imaging* 2018;**48**:916–26. doi:10.1002/jmri.25960.
 71. Li Y, Qian Z, Xu K, et al. MRI features predict p53 status in lower-grade gliomas via a machine-learning approach. *Neuroimage Clin* 2017;**17**:306–11. doi:10.1016/j.nicl.2017.10.030.
 72. Han Y, Xie Z, Zang Y, et al. Non-invasive genotype prediction of chromosome 1p/19q co-deletion by development and validation of an MRI-based radiomics signature in lower-grade gliomas. *J Neurooncol* 2018;**140**:297–306. doi:10.1007/s11060-018-2953-y.
 73. Decuyper M, Bonte S, Deblaere K, et al. Automated MRI based pipeline for segmentation and prediction of grade, IDH mutation and 1p19q co-deletion in glioma. *Comput Med Imaging Graph* 2021;**88**:101831. doi:10.1016/j.compmedimag.2020.101831.
 74. Casale R, Lavrova E, Sanduleanu S, et al. Development and external validation of a non-invasive molecular status predictor of chromosome 1p/19q co-deletion based on MRI radiomics analysis of Low Grade Glioma patients. *Eur J Radiol* 2021;**139**:109678. doi:10.1016/j.ejrad.2021.109678.
 75. Yogananda CGB, Shah BR, Nalawade SS, et al. MRI-based deep-learning method for determining glioma MGMT

- promoter methylation status. *AJNR Am J Neuroradiol* 2021;**42**:845–52. doi:10.3174/ajnr.A7029.
76. Akbari H, Bakas S, Pisapia JM, et al. In vivo evaluation of EGFRvIII mutation in primary glioblastoma patients via complex multiparametric MRI signature. *Neuro Oncol* 2018;**20**:1068–79. doi:10.1093/neuonc/noy033.
 77. Yan J, Liu L, Wang W, et al. Radiomic features from multiparameter MRI combined with clinical parameters predict molecular subgroups in patients with medulloblastoma. *Front Oncol* 2020;**10**:558162. doi:10.3389/fonc.2020.558162.
 78. Wen PY, Kesari S. Malignant gliomas in adults. *N Engl J Med* 2008;**359**:492–507. doi:10.1056/NEJMra0708126.
 79. Walker DG, Kaye AH. Diagnosis and management of astrocytomas, oligodendrogliomas and mixed gliomas: a review. *Australas Radiol* 2001;**45**:472–82. doi:10.1046/j.1440-1673.2001.00959.x.
 80. Hegi ME, Diserens AC, Gorlia T, et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. *N Engl J Med* 2005;**352**:997–1003. doi:10.1056/NEJMoa043331.
 81. Cairncross G, Wang M, Shaw E, et al. Phase III trial of chemoradiotherapy for anaplastic oligodendroglioma: long-term results of RTOG 9402. *J Clin Oncol* 2013;**31**:337–43. doi:10.1200/JCO.2012.43.2674.
 82. Wei J, Yang G, Hao X, et al. A multi-sequence and habitat-based MRI radiomics signature for preoperative prediction of MGMT promoter methylation in astrocytomas with prognostic implication. *Eur Radiol* 2019;**29**:877–88. doi:10.1007/s00330-018-5575-z.
 83. Zormpas-Petridis K, Poon E, Clarke M, et al. Noninvasive MRI native T(1) mapping detects response to MYCN-targeted therapies in the Th-MYCN model of neuroblastoma. *Cancer Res* 2020;**80**:3424–35. doi:10.1158/0008-5472.CAN-20-0133.
 84. Li Z, Kaiser L, Holzgreve A, et al. Prediction of TERTp-mutation status in IDH-wildtype high-grade gliomas using pre-treatment dynamic [18F] FET PET radiomics. *Eur J Nucl Med Mol Imaging* 2021;**48**:4415–25. doi:10.1007/s00259-021-05526-6.
 85. Qian J, Herman MG, Brinkmann DH, et al. Prediction of MGMT status for glioblastoma patients using radiomics feature extraction from 18F-DOPA-PET imaging. *Int J Radiat Oncol Biol Phys* 2020;**108**:1339–46. doi:10.1016/j.ijrobp.2020.06.073.
 86. Li L, Mu W, Wang Y, et al. A non-invasive radiomic method using 18F-FDG PET predicts isocitrate dehydrogenase genotype and prognosis in patients with glioma. *Front Oncol* 2019;**9**:1183. doi:10.3389/fonc.2019.01183.
 87. Mu W, Jiang L, Shi Y, et al. Non-invasive measurement of PD-L1 status and prediction of immunotherapy response using deep learning of PET/CT images. *J Immunother Cancer* 2021;**9**:e002118. doi:10.1136/jitc-2020-002118.
 88. Zhang J, Zhao X, Zhao Y, et al. Value of pre-therapy 18F-FDG PET/CT radiomics in predicting EGFR mutation status in patients with non-small cell lung cancer. *Eur J Nucl Med Mol Imaging* 2020;**47**:1137–46. doi:10.1007/s00259-019-04592-1.
 89. Chang C, Sun X, Wang G, et al. A machine learning model based on PET/CT radiomics and clinical characteristics predicts ALK rearrangement status in lung adenocarcinoma. *Front Oncol* 2021;**11**:603882. doi:10.3389/fonc.2021.603882.
 90. Lee SH, Kao GD, Feigenberg SJ, et al. Multiblock discriminant analysis of integrative 18F-FDG-PET/CT radiomics for predicting circulating tumor cells in early-stage non-small cell lung cancer treated with stereotactic body radiation therapy. *Int J Radiat Oncol Biol Phys* 2021;**110**:1451–65. doi:10.1016/j.ijrobp.2021.02.030.
 91. Aide N, Elie N, Blanc-Fournier C, et al. Hormonal receptor immunohistochemistry heterogeneity and 18F-FDG metabolic heterogeneity: preliminary results of their relationship and prognostic value in luminal non-metastatic breast cancers. *Front Oncol* 2021;**10**:599050. doi:10.3389/fonc.2020.599050.
 92. Li X-R, Jin J-J, Yu Y, et al. PET-CT radiomics by integrating primary tumor and peritumoral areas predicts E-cadherin expression and correlates with pelvic lymph node metastasis in early-stage cervical cancer. *Eur Radiol* 2021;**31**:5967–79. doi:10.1007/s00330-021-07690-7.
 93. Saadani H, Van der Hiel B, Aalbersberg EA, et al. Metabolic biomarker-based BRAFV600 mutation association and prediction in melanoma. *J Nucl Med* 2019;**60**:1545–52. doi:10.2967/jnumed.119.228312.
 94. Haubold J, Demircioglu A, Gratz M, et al. Non-invasive tumor decoding and phenotyping of cerebral gliomas utilizing multiparametric 18 F-FET PET-MRI and MR fingerprinting. *Eur J Nucl Med Mol Imaging* 2020;**47**:1435–45. doi:10.1007/s00259-019-04602-2.
 95. Matsui Y, Maruyama T, Nitta M, et al. Prediction of lower-grade glioma molecular subtypes using deep learning. *J Neurooncol* 2020;**146**:321–7. doi:10.1007/s11060-019-03376-9.
 96. Cui H, Zhang D, Peng F, et al. Identifying ultrasound features of positive expression of Ki67 and P53 in breast cancer using radiomics. *Asia Pac J Clin Oncol* 2021;**17**:e176–84. doi:10.1111/ajco.13397.
 97. Wu L, Zhao Y, Lin P, et al. Preoperative ultrasound radiomics analysis for expression of multiple molecular biomarkers in mass type of breast ductal carcinoma in situ. *BMC Med Imaging* 2021;**21**:84. doi:10.1186/s12880-021-00610-7.
 98. Zhang X, Li H, Wang C, et al. Evaluating the accuracy of breast cancer and molecular subtype diagnosis by ultrasound image deep learning model. *Front Oncol* 2021;**11**:623506. doi:10.3389/fonc.2021.623506.
 99. Jiang M, Zhang D, Tang S-C, et al. Deep learning with convolutional neural network in the assessment of breast cancer molecular subtypes based on US images: a multicenter retrospective study. *Eur Radiol* 2021;**31**:3673–82. doi:10.1007/s00330-020-07544-8.
 100. Salvi M, Acharya UR, Molinari F, et al. The impact of pre- and post-image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis. *Comput Biol Med* 2021;**128**:104129. doi:10.1016/j.compbiomed.2020.104129.
 101. Shao Y, Gao Y, Guo Y, et al. Hierarchical lung field segmentation with joint shape and appearance sparse learning. *IEEE Trans Med Imaging* 2014;**33**:1761–80. doi:10.1109/TMI.2014.2305691.
 102. Shin SJ, You SC, Jeon H, et al. Style transfer strategy for developing a generalizable deep learning application in digital pathology. *Comput Methods Programs Biomed Comput Methods Programs Biomed* 2021;**198**:105815. doi:10.1016/j.cmpb.2020.105815.
 103. Sun T, Zhang R, Wang J, et al. Computer-aided diagnosis for early-stage lung cancer based on longitudinal and balanced data. *PLoS One* 2013;**8**:e63559. doi:10.1371/journal.pone.0063559.
 104. Truong AH, Sharmanska V, Limbck-Stanic C, et al. Optimization of deep learning methods for visualization of tumor heterogeneity and brain tumor grading through digital pathology. *Neurooncol Adv* 2020;**2**:vdaa110. doi:10.1093/noajnl/vdaa110.

105. Zheng L, Liu Z, Yang Y, et al. Accurate inference of gene regulatory interactions from spatial gene expression with deep contrastive learning. *Bioinformatics* 2021. doi:10.1093/bioinformatics/btab718.
106. Jung H, Lodhi B, Kang J. An automatic nuclei segmentation method based on deep convolutional neural networks for histopathology images. *BMC Biomed Eng* 2019;1:24. doi:10.1186/s42490-019-0026-8.
107. Wang D, Khosla A, Gargeya R, et al. Deep learning for identifying metastatic breast cancer. arXiv preprint arXiv:1606.05718, 2016.
108. Dai X, Li T, Bai Z, et al. Breast cancer intrinsic subtype classification, clinical use and future trends. *Am J Cancer Res* 2015;5:2929–43.
109. Christiansen EM, Yang SJ, Ando DM, et al. In silico labeling: predicting fluorescent labels in unlabeled images. *Cell* 2018;173:792–803. doi:10.1016/j.cell.2018.03.040.
110. Diao JA, Wang JK, Chui WF, et al. Human-interpretable image features derived from densely mapped cancer pathology slides predict diverse molecular phenotypes. *Nat Commun* 2021;12:1613. doi:10.1038/s41467-021-21896-9.
111. Bian C, Wang Y, Lu Z, et al. ImmunoAIzer: A deep learning-based computational framework to characterize cell distribution and gene mutation in tumor microenvironment. *Cancers* 2021;13:1659. doi:10.3390/cancers13071659.
112. Bian C, Wang Y, An Y, et al. A computational prediction method based on modified U-Net for cell distribution in tumor microenvironment. In *Medical Imaging 2021: Digital Pathology, International Society for Optics and Photonics*, 2021, 116030C.
113. Song Z, Zou S, Zhou W, et al. Clinically applicable histopathological diagnosis system for gastric cancer detection using deep learning. *Nat Commun* 2020;11:4294. doi:10.1038/s41467-020-18147-8.
114. Bulten W, Pinckaers H, van Boven H, et al. Automated deep-learning system for Gleason grading of prostate cancer using biopsies: a diagnostic study. *Lancet Oncol* 2020;21:233–41. doi:10.1016/S1470-2045(19)30739-9.
115. Nagpal K, Foote D, Tan F, et al. Development and validation of a deep learning algorithm for Gleason grading of prostate cancer from biopsy specimens. *JAMA Oncol* 2020;6:1372–80. doi:10.1001/jamaoncol.2020.2485.
116. Campanella G, Hanna MG, Geneslaw L, et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat Med* 2019;25:1301–9. doi:10.1038/s41591-019-0508-1.
117. Shao L, Liu Z, Feng L, et al. Multiparametric MRI and whole slide image-based pretreatment prediction of pathological response to neoadjuvant chemoradiotherapy in rectal cancer: a multicenter radiopathomic study. *Ann Surg Oncol* 2020;27:4296–306. doi:10.1245/s10434-020-08659-4.
118. Coudray N, Ocampo PS, Sakellaropoulos T, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med* 2018;24:1559–67. doi:10.1038/s41591-018-0177-5.
119. Yu K-H, Zhang C, Berry GJ, et al. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. *Nat Commun* 2016;7:12474. doi:10.1038/ncomms12474.
120. Corredor G, Wang X, Zhou Y, et al. Spatial architecture and arrangement of tumor-infiltrating lymphocytes for predicting likelihood of recurrence in early-stage non-small cell lung cancer. *Clin Cancer Res* 2019;25:1526–34. doi:10.1158/1078-0432.CCR-18-2013.
121. Mobadersany P, Yousefi S, Amgad M, et al. Predicting cancer outcomes from histology and genomics using convolutional networks. *Proc Natl Acad Sci U S A* 2018;115:E2970–9. doi:10.1073/pnas.1717139115.
122. Kather JN, Krisam J, Charoentong P, et al. Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study. *PLoS Med* 2019;16:e1002730. doi:10.1371/journal.pmed.1002730.
123. Courtiol P, Maussion C, Moarii M, et al. Deep learning-based classification of mesothelioma improves prediction of patient outcome. *Nat Med* 2019;25:1519–25. doi:10.1038/s41591-019-0583-3.
124. van der Laak J, Litjens G, Ciompi F. Deep learning in histopathology: the path to the clinic. *Nat Med* 2021;27:775–84. doi:10.1038/s41591-021-01343-4.
125. Moen E, Bannan D, Kudo T, et al. Deep learning for cellular image analysis. *Nat Methods* 2019;16:1233–46. doi:10.1038/s41592-019-0403-1.
126. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;521:436–44. doi:10.1038/nature14539.
127. Zhang Z, Chen P, McGough M, et al. Pathologist-level interpretable whole-slide cancer diagnosis with deep learning. *Nat Mach Intell* 2019;1:236–45. doi:10.1038/s42256-019-0052-1.
128. Huang C, Cintra M, Brennan K, et al. Development and validation of radiomic signatures of head and neck squamous cell carcinoma molecular features and subtypes. *EBioMedicine* 2019;45:70–80. doi:10.1016/j.ebiom.2019.06.034.
129. Zwirner K, Hilke FJ, Demidov G, et al. Radiogenomics in head and neck cancer: correlation of radiomic heterogeneity and somatic mutations in TP53, FAT1 and KMT2D. *Strahlenther Onkol* 2019;195:771–9. doi:10.1007/s00066-019-01478-x.
130. Shi Z, Jiang W, Chen X, et al. Prognostic and clinicopathological value of Ki-67 expression in patients with nasopharyngeal carcinoma: a meta-analysis. *Ther Adv Med Oncol* 2020;12:1758835920951346. doi:10.1177/1758835920951346.
131. Traverso A, Wee L, Dekker A, et al. Repeatability and reproducibility of radiomic features: a systematic review. *Int J Radiat Oncol Biol Phys* 2018;102:1143–58. doi:10.1016/j.ijrobp.2018.05.053.
132. Paul R, Cherezov D, Schabath MB, et al. Towards deep radiomics: nodule malignancy prediction using CNNs on feature images. In *Medical Imaging 2019: Computer-Aided Diagnosis, International Society for Optics and Photonics*, 2019, 109503Z.
133. Taha B, Boley D, Sun J, et al. State of radiomics in glioblastoma. *Neurosurgery* 2021;89:177–84. doi:10.1093/neuros/nyab124.