New research progress on 18F-FDG PET/CT radiomics for EGFR mutation prediction in lung adenocarcinoma: a review

Xinyu Ge[1,2,3], Jianxiong Gao[1,2,3], Rong Niu[1,2,3], Yunmei Shi[1,2,3], Xiaoliang Shao[1,2,3], Yuetao Wang[1,2,3] and Xiaonan Shao[1,2,3]*

[1]Department of Nuclear Medicine, the Third Affiliated Hospital of Soochow University, Changzhou, China, [2]Institute of Clinical Translation of Nuclear Medicine and Molecular Imaging, Soochow University, Changzhou, China, [3]Department of Nuclear Medicine, Changzhou Clinical Medical Center, Changzhou, China

Lung cancer, the most frequently diagnosed cancer worldwide, is the leading cause of cancer-associated deaths. In recent years, significant progress has been achieved in basic and clinical research concerning the epidermal growth factor receptor (EGFR), and the treatment of lung adenocarcinoma has also entered a new era of individualized, targeted therapies. However, the detection of lung adenocarcinoma is usually invasive. 18F-FDG PET/CT can be used as a noninvasive molecular imaging approach, and radiomics can acquire high-throughput data from standard images. These methods play an increasingly prominent role in diagnosing and treating cancers. Herein, we reviewed the progress in applying 18F-FDG PET/CT and radiomics in lung adenocarcinoma clinical research and how these data are analyzed via traditional statistics, machine learning, and deep learning to predict EGFR mutation status, all of which achieved satisfactory results. Traditional statistics extract features effectively, machine learning achieves higher accuracy with complex algorithms, and deep learning obtains significant results through end-to-end methods. Future research should combine these methods to achieve more accurate predictions, providing reliable evidence for the precision treatment of lung adenocarcinoma. At the same time, facing challenges such as data insufficiency and high algorithm complexity, future researchers must continuously explore and optimize to better apply to clinical practice.

KEYWORDS

radiomics, machine learning, deep learning, EGFR, positron-emission tomography/computed tomography

Introduction

Lung cancer is responsible for the highest cancer-associated mortality worldwide, among which lung adenocarcinoma, the most common pathological type, accounts for 40% of all lung cancers (1–3). Because of the insidious onset of symptoms at an early stage, advanced tumors have developed in nearly 80% of patients by the time of diagnosis, which

results in a poor prognosis (4). Approximately 50% of Asian patients with adenocarcinoma harbor epidermal growth factor receptor (EGFR) mutations (5, 6). EGFR-mutant tumors exhibit an increased response rate to tyrosine kinase inhibitors (TKIs) versus EGFR wild-type tumors (7), while cisplatin-based chemotherapy yields more optimal outcomes in patients with EGFR wild-type lung cancer (8). Exon 19 deletion (19 del) and exon 21 L858R missense (21 L858R) are the two most frequent mutant isoforms of EGFR, which represent approximately 90% of all mutations (9). It has been shown that the 19 del mutation is more sensitive to TKIs than the 21 L858R mutation, leading to a longer median survival time (10). Afatinib or osimertinib is currently the first choice for treating patients with the 19 del mutation, whereas erlotinib plus bevacizumab therapy is recommended for patients with the 21 L858R mutation (11). Therefore, identifying EGFR mutation status and subtypes before treatment is crucial for offering accurate guidance regarding individualized patient treatment strategies.

Molecular detection of tumor tissues obtained by biopsy or surgical resection is the gold standard for identifying EGFR mutations (12). Because of the high risk of invasive procedures and the high heterogeneity of tumor tissues, pathological detection is not always feasible (13). Detection of circulating tumor DNA (ctDNA) has recently emerged as an alternative to EGFR mutation testing, but plasma ctDNA mutation detection can be easily affected by tumor burden and has a relatively high false-negative rate (14). Therefore, a noninvasive, rapid, and accurate method is urgently needed to identify EGFR mutations and subtypes.

Positron emission tomography (PET) imaging can visualize tissues at the biochemical level and detect potential neoplastic lesions earlier than computed tomography (CT) and magnetic resonance imaging (MRI) (15), while PET/CT can simultaneously provide metabolic and anatomical information regarding lesions. Recently, the relationship between the conventional metabolic parameters (such as maximum standardized uptake value (SUVmax)) of 18F-fluorodeoxyglucose (FDG) PET/CT and EGFR mutations were analyzed by many researchers, yet the results were controversial, probably because of the small sample size and complicated tumor microenvironment (16, 17).

First proposed by Lambin in 2012 (18), radiomics is continuously being perfected. Currently, radiomics refers to the interpretation of imaging data of regions of interest (ROIs), and thus, data with high-resolution spatial features are created using automated high-throughput feature extraction algorithms, which are further analyzed and interpreted by statistics and machine learning techniques, thereby supporting clinical decision-making (19). Over the past decade, 18F-FDG PET/CT radiomics has been applied to multiple aspects of cancer, such as the identification of benign and malignant solitary pulmonary nodules (20), lung cancer pathological typing and staging (21), prediction of gene mutation and molecular phenotypes (22), and evaluation of efficacy and prognosis (23). Amassed evidence substantiates the practicality and promising benefits of 18F-FDG PET/CT radiomics for prognosticating EGFR mutation classifications and subtypes (24). Radiomics methodology significantly mitigates patient discomfort and potential complications while delivering expedited diagnostic feedback and a comprehensive evaluation of tumor heterogeneity (25). Moreover, it serves as an invaluable tool for the dynamic monitoring of disease progression. Such advancements considerably improve patient experiences and promote efficient allocation of medical resources, thereby paving the way for developing tailored treatment regimens. This review aimed to summarize the research progress on 18F-FDG PET/CT radiomics for predicting EGFR mutation status and subtypes in lung adenocarcinoma.

Literature search strategy

To gather information relevant to our topic, we conducted a comprehensive literature search in PubMed, using the following search terms: ((("Lung Neoplasms"[MeSH]) OR (Lung adenocarcinoma[Title/Abstract]) OR (Squamous cell carcinoma [Title/Abstract]) OR (Pulmonary Neoplasm[Title/Abstract]) OR (Lung Cancer[Title/Abstract]) OR (Pulmonary Cancer[Title/Abstract]) OR (NSCLC[Title/Abstract])) AND ((EGFR [Title/Abstract]) OR (epidermal growth factor receptor[Title/Abstract]) OR (targeted therapy[Title/Abstract]))) AND ((radiomics [Title/Abstract]) OR (deep learning[Title/Abstract]) OR (machine learning [Title/Abstract]) OR (transfer learning [Title/Abstract]) OR (CNN [Title/Abstract]) OR (Convolutional Neural Networks [Title/Abstract]) OR (features [Title/Abstract]) OR (radiogenomics [Title/Abstract]) OR (Artificial intelligence [Title/Abstract]))AND (("Positron Emission Tomography Computed Tomography"[MeSH]) OR (emission-computed tomography [Title/Abstract]) OR (PET [Title/Abstract])). Initially, we identified 83 articles. After a careful review of titles and abstracts, we selected 26 articles that met our topic.

Procedures and classification of radiomics

Traditional radiomics (TR) procedures mainly consist of image acquisition, reconstruction, segmentation, feature extraction and filtering, model establishment, and performance evaluation (26). Multiple classifiers are used during the model establishment. Based on the different classifiers, the radiomics were classified in this study into traditional statistics-based radiomics (TSR), machine learning-based radiomics (MLR, specifically shallow learning), and deep learning (DL) to reflect iterations in omics technique. Figure 1 illustrates the differences in the procedures for the three radiomics techniques.

TSR can identify the most valued features with non-zero coefficients by extracting TR features (e.g., shape-based features, histogram features, and texture features) and applying mathematical-statistical methods integrated with algorithms such as the least absolute shrinkage and selection operator (LASSO) (27). A radiomics score (Rad-score) is then calculated and established for each lesion with a linear combination of the selected features correspondingly weighted by their coefficients. A classical logistic regression (LR) approach is then adopted as the classifier to build radiomics models. This method is the earliest, simplest, and most interpretable omics technique.

FIGURE 1
The differences in the procedures for the three radiomics.
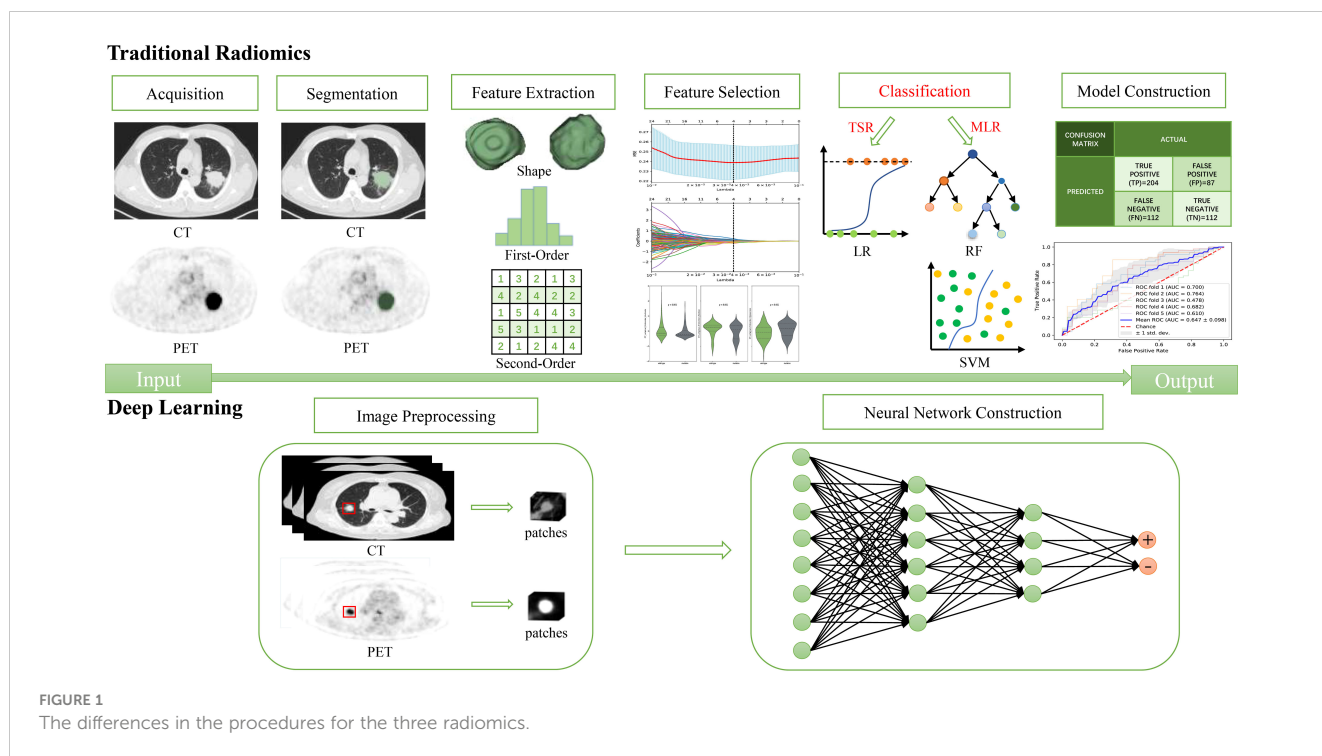
MLR is the most mature and mainstream approach that establishes classification or prediction models using machine learning algorithms after feature extraction and optimal radiomics feature screening. Common machine learning classifiers encompass random forest (RF), support vector machine (SVM), decision tree (DT), Bayesian network (BN), and k-nearest neighbor (KNN) algorithms (28). After MLR, there is a unique method named deep learning-based radiomics (DLR) that can extract DL features using artificial neural networks (ANNs), such as convolutional neural networks (CNNs), and then construct models using machine learning algorithms (29, 30).

DL, different from the TR represented by TSR and MLR, relies on multilayer nonlinear neural networks and bypasses the cumbersome feature extraction and selection procedures to automatically learn features from images, by which prediction models are established according to the "end-to-end" workflow without any human intervention (31).

Prediction of EGFR mutations by the TSR model

Before the emergence of radiomics, the clinical characteristics of patients with lung adenocarcinoma were frequently applied to predict EGFR mutation status, and multiple clinical characteristics (e.g., female, non-smoker, and adenocarcinoma histology) were related to EGFR mutations (32). It has also been shown that several CT features, such as maximum tumor diameter, tumor location, density, ground-glass opacity, pleural traction, and air bronchogram, denote EGFR mutation status in lung adenocarcinoma (33). A recent study has illustrated that tumors with reduced long diameters have a slightly higher risk of EGFR mutations though tumors with ground-glass opacity have markedly risk of EGFR mutations (34). Yip et al. (35, 36) demonstrated that PET radiomic features have great potential to predict EGFR mutation status by quantifying the tumor metabolic phenotype. However, the radiomics features usually can only reflect information regarding the image and cannot comprehensively reflect the patient's condition. To increase the model's accuracy in identifying EGFR mutation status and subtypes, several researchers incorporated clinical information, including CT features, into 18F-FDG PET/CT radiomics models (34).

In recent years, it has been revealed that the diagnostic performance and goodness-of-fit of models utilizing clinical features integrated with TR features are more optimal, with greater clinical benefit in predicting EGFR mutations. Eight studies on TSR are listed in Table 1, six of which used combined models of radiomics and clinical features (22, 34, 37, 40–42) and achieved superior performance as compared to single radiomics models and single clinical feature models, with the area under the curve (AUC) values of the test set results ranging from 0.81 to 0.87. These data demonstrate the effectiveness of this integrated imaging tool in predicting EGFR mutations.

Regarding discriminating the main subtypes of EGFR mutations, studies have shown a relatively low predictive accuracy of 18F-FDG PET/CT radiomics to distinguish the 19 del mutation from 21 L858R. Li et al. (22) reported AUC values of 0.708 and 0.652 in predicting the 19 del mutation and 21 L858R mutation in the training and testing sets, respectively. In the study of Zhang et al. (40), after clinical information was integrated into the predictive model, only one PET feature could discriminate between the 19 del mutation and 21 L858R mutation, showing low predictive ability (AUC = 0.661). Additionally, Nair et al. (38) suggested that PET/CT features were superior (AUC = 0.86) for discriminating mutations in exons 19 and 21 compared to CT features alone. However, this

TABLE 1 Studies on applying 18F-FDG PET/CT TSR in predicting EGFR mutations in patients with lung adenocarcinoma.

Author	Year	Samples	Radiomics features	Training set	Testing set
Li et al. (22)	2022	179 (EGFR+: 105, EGFR-: 74; 19 del: 46, 21 L858R: 53)	2 PET radiomics features; 4 CT radiomics features	mutant/wild model PET/CT + clinical: AUC= 0.882; 19 del/21 L858R model: PET/CT: AUC = 0.708	mutant/wild model PET/CT + clinical: AUC= 0.837; 19 del/21 L858R model: PET/CT: AUC = 0.652
Zhao et al. (37)	2022	88	6 PET radiomics features; 6 CT radiomics features	PET/CT + clinical: AUC = 0.864	–
Nair et al. (38)	2021	50 (EGFR+: 21, EGFR-: 29; 19 del: 11, 21 mut: 10)	14 PET/CT radiomics features	mutant/wild model: PET/CT: AUC = 0.8713 ± 0.05 exon19/21 model: PET/CT: AUC = 0.860 ± 0.07	–
Yang et al. (39)	2021	114 (EGFR+:51, EGFR-:63)	7 PET/CT radiomics features	PET/CT: AUC = 0.866	–
Chang et al. (34)	2021	583 (EGFR+: 295, EGFR-: 288)	30 PET/CT radiomics features; 12 PET features; 15 CT features	PET/CT: AUC= 0.76; PET/CT + clinical: AUC= 0.84	PET/CT: AUC= 0.75; PET/CT + clinical: AUC= 0.81
Zhang et al. (40)	2020	173 (EGFR+: 71, EGFR-: 102; 19 del: 29, 21 L858R: 38)	2 PET radiomics features; 4 CT radiomics features	PET/CT: AUC= 0.868, PET/CT + clinical: AUC= 0.866	mutant/wild model PET/CT: AUC= 0.769, PET/CT + clinical: AUC= 0.827, 19 del/21 L858R model: AUC = 0.661;
Zhang et al. (41)	2020	248 (EGFR+: 133, EGFR-: 115)	5 PET radiomics features; 5 CT radiomics features	PET/CT: AUC= 0.79; PET/CT + clinical: AUC= 0.86	PET/CT: AUC= 0.85; PET/CT + clinical: AUC = 0.87
Li et al. (42)	2019	115 (EGFR+: 64, EGFR-: 51)	2 PET radiomic features; 4 CT radiomic features	PET/CT: AUC= 0.805; PET/CT + clinical: AUC= 0.822	–

AUC, area under the curve; EGFR, epidermal growth factor receptor; 19 del, 19 deletion; 21 L858R, 21 L858R missense; PET, positron emission tomography; CT, computed tomography.

study had a smaller sample size of only 21 patients and lacked an independent test set. It was considered that the poor predictive performance for EGFR mutation subtypes was mainly attributed to overfitting of the trained model caused by inherent technical limitations of TSR, such as but not limited to the possible loss of essential information with the use of single LASSO regression due to high dimensionality of the radiomics features (43).

Radiomics has rapidly developed as an emerging field, but TSR continues to be used mainly due to its advantages of good interpretability and easy application in clinical practice. TSR can effectively transform the data into scores, usually in the form of a nomogram (23, 34, 39), which allows intuitive calculation and easier understanding and analysis of the associations and differences between the data. However, its limitation is the low predictive efficacy of the traditional logistic regression model relative to most machine learning classifiers, which is also why the use of MLR is becoming increasingly common.

Prediction of EGFR mutations by the MLR model

Most machine learning algorithms aim to develop an optimal model to solve one problem. The selection of machine learning

classifiers requires comprehensive consideration of multiple factors such as data characteristics, model performance, computational resources, and usage, and experiments are necessary to determine the optimal classifier (44, 45). Studies on EGFR mutation identification in lung cancer based on different machine learning classifiers are presented in Table 2.

LR is a simple and easily implementable algorithm for binary classification problems. Zhang et al. (47) found that the fusion of PET and CT image features did not significantly enhance the classifier's performance compared to using only PET image features. Furthermore, simple and efficient classifiers tend to be more advantageous in situations with limited data. Therefore, they utilized a small amount of single-mode PET images, employed various methods to extract features and perform dimensionality reduction, and then input the most significant features into the LR model for classification. This approach ultimately yielded the highest AUC value of 0.843. This method used less data than previous research but achieved better results.

The RF algorithm, one of the most classical machine learning algorithms, is a decision tree-based ensemble learning algorithm that integrates the results obtained by several weak classifiers to improve the model performance. Gao et al. (46) found that compared with LR and SVM, the RF performed best among the

TABLE 2 Studies on applying 18F-FDG PET/CT MLR in predicting EGFR mutations in patients with lung adenocarcinoma.

Author	Year	Samples	Model	Training set	Testing set
Gao et al. (46)	2023	515 (EGFR+:313, EGFR-:202)	**RF,** LR, SVM	AUC=0.760	AUC=0.730
Zhang et al. (47)	2023	115 (EGFR+:64, EGFR-: 51)	**LR**, RF, SVM, Adaboost	AUC=0.843	–
Zuo et al. (48)	2023	767 (EGFR+:450, EGFR-: 317)	**LGBM**, XGB, RF,LR	–	mutant/wild model: AUC=0.80; 19 del/21 L858R model: AUC=0.76
Yang et al. (49)	2022	313 (EGFR+:181, EGFR-:132)	**SVM**, DT, RF	mutant/wild model: AUC=0.881; 19 del/21 L858R model: $AUC_{19\ del}$=0.849; $AUC_{21\ L858R}$=0.851	mutant/wild model: AUC=0.926; 19 del/21 L858R model: $AUC_{19\ del}$=0.859; $AUC_{21\ L858R}$=0.805
Agüloğlu et al. (50)	2022	159 (EGFR+:59, EGFR-:100)	RF, **NB**, KNN, LR, SVM, DT	AUC=0.751	AUC=0.797
Shiri et al. (51)	2022	136	RF	AUC=0.92–0.94	–
Ruan et al. (52)	2022	100 (EGFR+:46, EGFR-: 54)	LR, **SVM**	AUC=0.746	AUC=0.741
Huang et al. (30)	2022	138 (EGFR+:64, EGFR-: 74)	CNN	AUC=0.91	AUC=0.85
Wang et al. (43)	2021	238 (EGFR+: 126, EGFR-: 112)	KNN, SVM, **Adaboost**	19 del/21 L858R model: AUC=0.87	–
Shiri et al. (53)	2020	300	SVM, KNN, DT, QDA, MLP, SGD, LR, NB, GNB, **RF,** AB, BAG	AUC=0.82	–
Yang et al. (54)	2020	174 (EGFR+:109, EGFR-:65)	RF	mutant/wild model: AUC=0.77; 19 del/21 L858R model: AUC=0.82	mutant/wild model: AUC=0.71; 19 del/21 L858R model: AUC=0.73
Liu et al. (55)	2020	148 (EGFR+: 75, EGFR-: 73)	XGB	mutant/wild model: AUC=0.93; 19 del/21 L858R model: $AUC_{19\ del}$=0.91; $AUC_{21\ L858R}$=1.0	mutant/wild model: AUC=0.87; 19 del/21 L858R model: $AUC_{19\ del}$=0.77; $AUC_{21\ L858R}$=0.92
Koyasu et al. (56)	2020	138 (EGFR+:38, EGFR-:100)	RF, **XGB**	AUC=0.659	–
Jiang et al. (57)	2019	80 (EGFR+:30, EGFR-:50)	SVM	AUC=0.953	–

KNN, K-nearest neighbor; DT, decision tree; QDA, quadratic discriminant analysis; MLP, multilayer perceptron; SGD, stochastic gradient descent; LR, logistic regression; NB, naive Bayes; GNB, Gaussian naive Bayes; RF, random forest; AB, adaptive boosting; BAG, bagging; SVM, support vector machine; XGBoost, eXtreme gradient boosting; Adaboost, adaptive boosting; LGBM, light gradient boosting machine classifier; CNN, convolutional neural network. Bold indicates the best performance model.

three radiomics models of CT, PET, and PET/CT (testing sets AUC: 0.726, 0.678, and 0.704). Yang et al. (54) adopted the RF algorithm to develop a predictive model for EGFR mutations and determined its AUC values of 0.77 and 0.71 in the training set and testing set, respectively.

XGBoost algorithm, derived from the RF algorithm, is a gradient-boosting decision tree-based model algorithm with a strong modeling effect and ultrafast computational speed. It increases computational efficiency with its parallel computation and cache optimization. Also, there is greater stability associated with the XGBoost algorithm due to its application of a regularization technique in the training process. Liu et al. (55) employed XGBoost to predict EGFR mutations and found that the 18F-FDG PET/CT radiomics model achieved satisfactory capability for identifying EGFR mutation status (AUC = 0.87). It was reported by Koyasu et al. (56) that a stronger performance was observed with a radiomics model employing

XGBoost and multiple types of imaging features as compared to RF, and this model showed the most optimal performance in the classification of EGFR mutations (AUC = 0.659).

SVM, another frequently used supervised learning algorithm, accomplishes satisfactory nonlinear data-fitting because SVM is a very commonly used strong learner with excellent performance in finding the optimal solution in high-dimensional space using a kernel function. After a comparison, Ruan et al. (52) revealed that the SVM model outperformed the LR model in EGFR mutation identification, and particularly, the SVM model based on the radial basis kernel function exhibited the most optimal performance in the testing set (AUC = 0.741). Yang et al. (49) retained the radiomics features and clinical factors to establish integrated models. The SVM model exhibited a stronger performance than the RF and DT models, yielding an AUC value of 0.926 in the testing set. Jiang et al. (57) found that combining qualitative and quantitative features outperformed qualitative or quantitative features alone, and the SVM model achieved excellent performance with a high AUC of 0.953.

The naive Bayes (NB) algorithm is a supervised learning algorithm based on Bayes' theorem that is computationally efficient, rapid, and simple to operate. Agüloğlu et al. (50) found that the NB machine learning algorithm established using the GLZLM_GLNU feature and clinical data contributed to the most successful prediction of EGFR mutations (AUC = 0.751). Shiri et al. (53) developed models using multimodal PET/CT image features and trained them using multiple machine learning methods for EGFR mutation prediction, achieving an AUC value of 0.82, which was increased to 0.94 by ComBat harmonization (51). Further optimization, such as data correction, may improve the model's performance in case of unsatisfactory results.

Studies have demonstrated the potential of 18F-FDG PET/CT radiomics to discriminate EGFR subtypes. Liu et al. (55) obtained two sets of prognostic radiomics features for specific EGFR mutation subtypes and found that their XGBoost classifier for the 19 del and 21 L858R mutations yielded a prediction accuracy of 0.77 and 0.92 with respect to AUC values, respectively. The study of Yang et al. (54) utilized an RF classifier, and the resultant AUC values of the models for discriminating 19/21 site mutations were 0.82 and 0.73 in the training test and testing set, respectively.

Wang et al. (43) developed a PET/CT radiomics model using an Adaboost classifier and reported its predictive value for EGFR mutation subtypes (AUC = 0.86). Adaboost is also an ensemble learning algorithm that can train a weak classifier through multiple iterations, improving overall accuracy. Zuo et al. (48) discovered that the XGBoost classifier combined with SVM feature selection method achieved the best performance in predicting EGFR subtypes (AUC reached 0.76, 0.63, and 0.61 in the internal test cohort and two external test cohorts, respectively). Also, Yang et al. (49) revealed that the SVM model was superior to the RF model and the DT model, achieving AUC values of 0.805 and 0.859 for predicting the 19 del and 21 L858R mutations in the testing set, respectively. Based on the results of subtype prediction, the predictive performance for the exon 21 mutation was more optimal than that for the exon 19 mutation, and the overall results for MLR were also superior to those of TSR.

Particularly, with the continuous development of AI technologies, DL, represented by CNNs, has shown satisfactory performance. Radiomics gradually relies on DL techniques to address challenges and limitations in routine radiomics workflows, including automated detection and segmentation procedures and harmonizing images by synthetic generation (58). To reduce or balance the high cost of data acquisition, several researchers have proposed DLR to integrate the feature outputs from the DL networks with classical machine learning classifiers. Huang et al. (30) constructed a DLR model to predict EGFR mutations using three-dimensional (3D) CNN, and they found that this DLR model was superior to TR (AUC: 0.79 *vs*. 0.68). DLR, as a transition between TR and DL, extracts and classifies advanced features and image data using multilayer neural networks, improving the accuracy of data analysis. There are few studies on DLR in this field, and additional studies may be required in the future to continually explore and validate the feasibility and performance of this proposal.

EGFR prediction models established using machine learning algorithms can identify EGFR mutation status and subtypes, displaying satisfactory discrimination, prediction accuracy, correction performance, and gain values. There have been more reports on MLR in the last few years, and the machine learning algorithms applied in this method have also been more mature than those of TSR. In the field of machine learning, there is no single algorithm that performs best in all situations. The performance of an algorithm depends on various factors, such as the data's characteristics, the problem's complexity, and the algorithm's parameter tuning. For a specific problem, it is necessary to experimentally compare different algorithms' performance to determine which is the most suitable. Generally, MLR selects the optimal combination to achieve stronger results using multiple feature screening and multiple classifiers. Based on our summary, RF and SVM are the classifiers that usually perform best in prediction efficacy. In radiomics, fusion using the stacking algorithm is an effective approach (59) because it incorporates the obtained multiple results of classifiers, where the prediction results of various models are selected as a new input and then trained into one model to improve the classification performance. No studies concerning this approach in MLR have been performed, but this approach may become more common with advanced technologies and data accumulation.

Prediction of EGFR mutations by the DL model

Radiomics features can be classified into two types: the first refers to predefined or manually extracted features established by an image processing expert, also known as traditional features; the second refers to deep features, and several DL algorithms assign a task in their extraction layer to self-design and select features without any human intervention (31, 58). TSR and MLR extract similar and traditional features, which mainly differ concerning different classifiers, while DLR and DL extract deep features. Several studies have suggested that deep features are superior to traditional features (60, 61).

Recently, DL has shown great potential in improving feature engineering in medical imaging and classification and prediction accuracy (62). Several investigators are now focusing on applying DL models in predicting EGFR mutation status, which has shown promising performance. Xiao et al. (63) proposed a deep learning framework based on the EfficientNet-V2 model. First, 32 2D views are extracted from each 3D cube of lung nodules. Then, deep features are extracted from these 32 views to predict whether EGFR is mutated. The results show that this deep learning model outperforms the radiomics model, with AUCs of 83.64% and 82.41%, respectively, effectively predicting EGFR mutations.

Using the robust deep convolutional neural network structure, the squeeze-and-excitation residual network (SE-ResNet) module, Yin et al. (64) developed two DL models (SE_{CT} and SE_{PET}) for EGFR mutation identification in CT and PET images, respectively. They also integrated the results of SE_{CT} and SE_{PET} using stacked generalization and increased the AUC value to 0.84, which was significantly higher than that of either SE_{CT} or SE_{PET}.

In another study, Chen et al. (65) applied the Stack DL model to effectively integrate anatomical bioimaging data from PET/CT images with clinical data, which displayed a strong predictive ability for EGFR mutations (AUC = 0.85 ± 0.09) and was superior to the ResNet PET/CT model (AUC = 0.81 ± 0.07) and the radiomics model (AUC = 0.60 ± 0.06). The most optimal features were constructed from the image data rather than selected from a predefined and limited set of feature candidates. With sufficient training data, CNN will outperform the feature selection scheme.

Mu et al. (66) developed an 18F-FDG PET/CT-based DL model using the 2D small-residual-convolutional-network (SResCNN), which accurately classified EGFR mutation status. Also, after constructing an integrated model that included EGFR-DLS (DL score) and histological and clinical features (smoking), the integrated model showed superior performance to the DL or clinical model (AUC = 0.88, 0.83, and 0.78, respectively). Overall, the multitask artificial intelligence system incorporating DL clinical features achieved relatively satisfactory results in predicting EGFR

status. This finding illustrates that incorporating multiple data can partly enhance the model's predictive accuracy.

At present, the use of 18F-FDG PET/CT DL models for EGFR mutation status prediction has been scarcely studied, while its use in EGFR mutation subtype identification has not been discussed. Although three retrieved articles (Table 3) demonstrated clear discrimination of EGFR mutations, all of them acquired section samples from 3D input images and thus obtained sufficient data to train 2D-CNN *de novo*, which would result in the loss of rich 3D anatomical information (67). In addition, DL-dependent methods in radiomics face some problems and need to address new challenges, one of which is the limited size of the data set. Generally, DL training requires less human input than traditional ML algorithms, but effective DL often involves a mass of training data.

In medical imaging, obtaining such many training samples is generally difficult. This limitation can theoretically be addressed by the similarity of visual features in the problem domain using the transfer learning (TL) strategy (44). Indeed, although TL is effective in DL, its application is limited because of the lack of new pre-trained models for medical imaging (68, 69). In recent years, scholars have also proposed leveraging unlabeled data (i.e., semi-supervised learning) to address the issue of small sample sizes, including techniques like pseudo-labeling and Generative Adversarial Networks (GANs) (70, 71). The potential value of these innovative approaches in predicting EGFR mutations remains an area ripe for further exploration and research in the future (72). Second, considerable computational resources are usually required to train DL models, including central processing units (CPUs), graphics processing units (GPUs), and memory. If computational resources are limited, TR may be a more optimal choice. The lack of interpretability of DL network-based models is an additional issue that remains incompletely addressed, and this algorithm is usually called the "black box" (73). The high performance of deep neural networks is achieved at the cost of high complexity and many parameters. Although one can, in principle, follow every processing step, one cannot understand the internal decision-making process because many parameters make it

TABLE 3   Studies on applying 18F-FDG PET/CT DL in predicting EGFR mutations in patients with lung adenocarcinoma.

Author	Year	Samples	Model	Training set	Testing set
Xiao et al. (63)	2023	150 (EGFR+:57, EGFR-: 93)	EfficientNet-V2	–	AUC%=83.64 ± 2.41
Chen et al. (65)	2022	147 (EGFR+:37, EGFR-: 110)	ResNet	StackPET-CT+ clinical: AUC=0.85 ± 0.09	–
Yin et al. (64)	2021	301 (EGFR+:153, EGFR-: 148)	SE-ResNet	StackPET-CT: AUC=0.86	StackPET-CT: AUC=0.84
Mu et al. (66)	2020	681 (EGFR+:312, EGFR-:369)	SResCNN	DLS+ clinical: AUC=0.88	DLS+ clinical: AUC=0.88

ResNet, residual network; SE-ResNet, squeeze-and-excitation residual network; SResCNN, small-residual-convolutional-network; DLS, deep learning score.

difficult to obtain meaningful explanations of model behaviors in this way.

There is insufficient ability for traditional machine learning techniques to process raw data, while DL models are well-suited for model training with large-scale data in the current big data era, and their high accuracy is due to their multilayer structure that can capture complex relations in the data and process high-dimensional and nonlinear data (31). DL models can learn more representative features than TR models, which is critical for stronger analytical performance. TR methods require rigorous procedures, such as detection, segmentation, feature extraction, and selection, which are cumbersome and time-consuming (74). DL can greatly improve efficiency through its "end-to-end" learning mode. Furthermore, radiomics features can be easily affected by artificial segmentation and scanning parameters, while DL models can adaptively learn features based on the data, thus effectively improving the model's robustness and generalization ability (29). Currently, several visualization methods have been developed to interpret the decision-making process of DL models. The most common method is using gradient-weighted class activation mapping (Grad-CAM) to generate "heatmaps" for input images, showing the influence weights of different parts of the image on the classification results, thereby explaining the learning process (75). In summary, deep learning and machine learning are advantageous in different applications. The most suitable approach in practice often depends on the specific requirements and constraints of the task. In the era of precision medicine, MLR will combine with DL to achieve automated imaging analysis and diagnosis, thereby improving diagnostic accuracy and efficiency. The advantages and disadvantages of the three types of radiomics are listed in Table 4.

Limitations and prospects

Identifying EGFR mutation status in lung adenocarcinoma using 18F-FDG PET/CT radiomics is promising, but limitations remain. First, the scanner and scanning parameters have not yet been standardized, and there has been no consensus on delineating lesional ROIs, which will partly affect the objectivity and reproducibility of experiments. As a standardization guideline, the Image Biomarker Standardization Initiative (IBSI) has proposed uniform standards for image quality, feature extraction, and preprocessing modalities to ensure the robustness and reproducibility of experiments (76, 77). Also, data sharing is a huge obstacle for researchers. Sharing training parameters instead of data is a simple solution (78, 79).

Second, similar to multidisciplinary teams in the clinical treatment of lung cancer, integrating different features is also conducive to constructing prediction models (80), yielding the concept of multi-omics (81). Combining knowledge from different fields and multiple disciplines, such as imaging, pathology, statistics, and clinical data, and the convergence of old and new techniques is a worthwhile objective for the future.

TABLE 4 Advantages and disadvantages of different types of radiomics.

Radiomics type	Advantages	Disadvantages
TSR	1. Methods are based on explicit assumptions and statistical models, and the results can be mathematically demonstrated. 2. The data are easy to interpret and understand and can provide meaningful biological results. 3. The requirement for data is low, and a mass of data is not required.	1. Provides a poor fit to complex data relations. 2. The analytical performance for high-dimensional data is limited, and complex nonlinear relations are poorly processed. 3. Extensive feature engineering may be required, which requires researchers to have extensive knowledge of biology.
MLR	1. Accuracy is high with satisfactory applicability. 2. The model is more interpretable and comprehensible. 3. High-dimensional data and nonlinear relationships can be processed.	1. Overfitting is easy in insufficient training data or noise. 2. Only labeled data can be processed. 3. Adjusting parameters is difficult and requires the strong technical skills of researchers.
DL	1. High performance for large-scale and high-dimensional data processing. 2. Data features can be automatically learned with a satisfactory fit to complex data relationships. 3. Unstructured data processing performance is satisfactory.	1. Poor interpretability with difficulty in understanding the results. 2. Requires substantial computational resources and data. 3. There are strict requirements for training data, and poor data quality may affect the results.

TSR, traditional statistics-based radiomics; MLR, machine learning-based radiomics; DL, deep learning.

Additionally, the current studies on radiomics are mostly single-center, small-scale retrospective studies, and prospective studies and external independent validation are scarce, which may lead to the overfitting of data and insufficient generalization ability of models. Thus, multi-center collaboration will be necessary to build large databases to increase models' accuracy. Furthermore, limited specificity and sensitivity are the inherent bottlenecks in the application of the tracer 18F-FDG in predicting lung cancer driver genes, and new imaging methods and new molecular probes such as 18F-MPG are required to improve the model performance (66, 82).

Furthermore, there are currently few studies on the application of PET/CT in identifying EGFR mutation subtypes, and two binary classification models are adopted in most studies on EGFR mutation subtypes. The ternary classification model can be directly applied in the future to obtain richer information that better describes the true situation of the samples, thus improving the prediction accuracy of models. However, the training and evaluation of ternary classification models are more complicated, requiring additional data and a longer time.

Moreover, habitat imaging is an important component of radiomics and a hot topic in cancer quantitative imaging. This generally refers to the analysis of different sub-regions within a

tumor. These sub-regions are referred to as "habitats", each of which may represent different biological characteristics and pathophysiological processes within the tumor (83). By analyzing these habitats, we can better understand the heterogeneity of the cancer, thereby facilitating the development of personalized treatment plans.

Lastly, although these models have clear advantages, their widespread clinical application still faces multiple challenges, such as data heterogeneity, interpretability of the models, difficulties in clinical integration, the need for external validation, regulatory barriers, economic considerations, and ethical and privacy concerns (84). In the future, we can further explore these models' practical application potential and clinical translation ability through prospective clinical research.

Conclusions

Radiomics is a well-established research method, forming a complete theoretical system and research process, but AI-based radiomics in medicine remains in the initial stage and is not widely utilized. With the advancement of machine learning methods and the continuous development of DL technologies, identifying EGFR mutation status by radiomics in lung adenocarcinoma will continue to be an area of intense clinical medicine research, providing more precise imaging diagnosis and individualized therapeutic guidance for patients.

Author contributions

XG is responsible for literature search and manuscript writing; JG, RN, XLS, YS, and YW are accountable for content planning and editing. XNS is responsible for proposing the topic and revising the

content. All authors have read the final manuscript and approved the version to be published. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Lortet-Tieulent J, Soerjomataram I, Ferlay J, Rutherford M, Weiderpass E, Bray F. International trends in lung cancer incidence by histological subtype: adenocarcinoma stabilizing in men but still increasing in women. *Lung Cancer* (2014) 84(1):13–22. doi: 10.1016/j.lungcan.2014.01.009

2. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. *CA Cancer J Clin* (2023) 73(1):17–48. doi: 10.3322/caac.21763

3. Lemjabbar-Alaoui H, Hassan OU, Yang YW, Buchanan P. Lung cancer: biology and treatment options. *Biochim Biophys Acta* (2015) 1856(2):189–210. doi: 10.1016/j.bbcan.2015.08.002

4. Torre LA, Siegel RL, Jemal A. Lung cancer statistics. *Adv Exp Med Biol* (2016) 893:1–19. doi: 10.1007/978-3-319-24223-1_1

5. Yu S, Liu D, Shen B, Shi M, Feng J. Immunotherapy strategy of egfr mutant lung cancer. *Am J Cancer Res* (2018) 8(10):2106–15.

6. Shi Y, Au JS, Thongprasert S, Srinivasan S, Tsai CM, Khoa MT, et al. A prospective, molecular epidemiology study of egfr mutations in Asian patients with advanced non-small-cell lung cancer of adenocarcinoma histology (Pioneer). *J Thorac Oncol* (2014) 9(2):154–62. doi: 10.1097/jto.0000000000000033

7. Okamoto I, Morita S, Tashiro N, Imamura F, Inoue A, Seto T, et al. Real world treatment and outcomes in egfr mutation-positive non-small cell lung cancer: long-

term follow-up of a large patient cohort. *Lung Cancer* (2018) 117:14–9. doi: 10.1016/j.lungcan.2018.01.005

8. Wu YL, Zhou C, Hu CP, Feng J, Lu S, Huang Y, et al. Afatinib versus cisplatin plus gemcitabine for first-line treatment of Asian patients with advanced non-small-cell lung cancer harbouring egfr mutations (Lux-lung 6): an open-label, randomised phase 3 trial. *Lancet Oncol* (2014) 15(2):213–22. doi: 10.1016/S1470-2045(13)70604-1

9. Suda K, Mitsudomi T, Shintani Y, Okami J, Ito H, Ohtsuka T, et al. Clinical impacts of egfr mutation status: analysis of 5780 surgically resected lung cancer cases. *Ann Thorac Surg* (2021) 111(1):269–76. doi: 10.1016/j.athoracsur.2020.05.041

10. Sutiman N, Tan SW, Tan EH, Lim WT, Kanesvaran R, Ng QS, et al. Egfr mutation subtypes influence survival outcomes following first-line gefitinib therapy in advanced Asian nsclc patients. *J Thorac Oncol* (2017) 12(3):529–38. doi: 10.1016/j.jtho.2016.11.2225

11. Stewart EL, Tan SZ, Liu G, Tsao MS. Known and putative mechanisms of resistance to egfr targeted therapies in nsclc patients with egfr mutations-a review. *Trans Lung Cancer Res* (2015) 4(1):67–81. doi: 10.3978/j.issn.2218-6751.2014.11.06

12. Ettinger DS, Wood DE, Aisner DL, Akerley W, Bauman J, Chirieac LR, et al. Non-small cell lung cancer, version 5.2017, nccn clinical practice guidelines in oncology. *J Natl Compr Canc Netw* (2017) 15(4):504–35. doi: 10.6004/jnccn.2017.0050

13. Taniguchi K, Okami J, Kodama K, Higashiyama M, Kato K. Intratumor heterogeneity of epidermal growth factor receptor mutations in lung cancer and its correlation to the response to gefitinib. *Cancer Sci* (2008) 99(5):929–35. doi: 10.1111/j.1349-7006.2008.00782.x

14. Goldman JW, Noor ZS, Remon J, Besse B, Rosenfeld N. Are liquid biopsies a surrogate for tissue egfr testing? *Ann Oncol* (2018) 29(suppl_1):i38–46. doi: 10.1093/annonc/mdx706

15. Cherry SR, Jones T, Karp JS, Qi J, Moses WW, Badawi RD. Total-body pet: maximizing sensitivity to create new opportunities for clinical research and patient care. *J Nucl Med* (2018) 59(1):3–12. doi: 10.2967/jnumed.116.184028

16. Gao XC, Wei CH, Zhang RG, Cai Q, He Y, Tong F, et al. (18)F-fdg pet/ct suv (Max) and serum cea levels as predictors for egfr mutation state in Chinese patients with non-small cell lung cancer. *Oncol Lett* (2020) 20(4):61. doi: 10.3892/ol.2020.11922

17. Minamimoto R, Jamali M, Gevaert O, Echegaray S, Khuong A, Hoang CD, et al. Prediction of egfr and kras mutation in non-small cell lung cancer using quantitative (18)F fdg-pet/ct metrics. *Oncotarget* (2017) 8(32):52792–801. doi: 10.18632/oncotarget.17782

18. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* (2012) 48(4):441–6. doi: 10.1016/j.ejca.2011.11.036

19. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology* (2016) 278(2):563–77. doi: 10.1148/radiol.2015151169

20. Guo HY, Lin JT, Huang HH, Gao Y, Yan MR, Sun M, et al. Development and validation of a (18)F-fdg pet/ct-based clinical prediction model for estimating Malignancy in solid pulmonary nodules based on a population with high prevalence of Malignancy. *Clin Lung Cancer* (2020) 21(1):47–55. doi: 10.1016/j.cllc.2019.07.014

21. Ren C, Zhang J, Qi M, Zhang J, Zhang Y, Song S, et al. Machine learning based on clinico-biological features integrated (18)F-fdg pet/ct radiomics for distinguishing squamous cell carcinoma from adenocarcinoma of lung. *Eur J Nucl Med Mol Imaging* (2021) 48(5):1538–49. doi: 10.1007/s00259-020-05065-6

22. Li S, Li Y, Zhao M, Wang P, Xin J. Combination of (18)F-fluorodeoxyglucose pet/ct radiomics and clinical features for predicting epidermal growth factor receptor mutations in lung adenocarcinoma. *Korean J Radiol* (2022) 23(9):921–30. doi: 10.3348/kjr.2022.0295

23. Yang B, Zhong J, Zhong J, Ma L, Li A, Ji H, et al. Development and validation of a radiomics nomogram based on (18)F-fluorodeoxyglucose positron emission tomography/computed tomography and clinicopathological factors to predict the survival outcomes of patients with non-small cell lung cancer. *Front Oncol* (2020) 10:1042. doi: 10.3389/fonc.2020.01042

24. Hu N, Yan G, Wu Y, Wang L, Wang Y, Xiang Y, et al. Recent and current advances in pet/ct imaging in the field of predicting epidermal growth factor receptor mutations in non-small cell lung cancer. *Front Oncol* (2022) 12:879341. doi: 10.3389/fonc.2022.879341

25. Brody H. Lung cancer. *Nature* (2020) 587(7834):S7. doi: 10.1038/d41586-020-03152-0

26. Mayerhoefer ME, Materka A, Langs G, Haggstrom I, Szczypinski P, Gibbs P, et al. Introduction to radiomics. *J Nucl Med* (2020) 61(4):488–95. doi: 10.2967/jnumed.118.222893

27. Ranstam J, Cook JA. Lasso regression. *Br J Surg* (2018) 105(10):1348. doi: 10.1002/bjs.10895

28. Yin X, Liao H, Yun H, Lin N, Li S, Xiang Y, et al. Artificial intelligence-based prediction of clinical outcome in immunotherapy and targeted therapy of lung cancer. *Semin Cancer Biol* (2022) 86(Pt 2):146–59. doi: 10.1016/j.semcancer.2022.08.002

29. Zhao W, Yang J, Ni B, Bi D, Sun Y, Xu M, et al. Toward automatic prediction of egfr mutation status in pulmonary adenocarcinoma with 3d deep learning. *Cancer Med* (2019) 8(7):3532–43. doi: 10.1002/cam4.2233

30. Huang W, Wang J, Wang H, Zhang Y, Zhao F, Li K, et al. Pet/ct based egfr mutation status classification of nsclc using deep learning features and radiomics features. *Front Pharmacol* (2022) 13:898529. doi: 10.3389/fphar.2022.898529

31. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* (2015) 521(7553):436–44. doi: 10.1038/nature14539

32. Zhou JY, Zheng J, Yu ZF, Xiao WB, Zhao J, Sun K, et al. Comparative analysis of clinicoradiologic characteristics of lung adenocarcinomas with alk rearrangements or egfr mutations. *Eur Radiol* (2015) 25(5):1257–66. doi: 10.1007/s00330-014-3516-z

33. Rizzo S, Petrella F, Buscarino V, De Maria F, Raimondi S, Barberis M, et al. Ct radiogenomic characterization of egfr, K-ras, and alk mutations in non-small cell lung cancer. *Eur Radiol* (2016) 26(1):32–42. doi: 10.1007/s00330-015-3814-0

34. Chang C, Zhou S, Yu H, Zhao W, Ge Y, Duan S, et al. A clinically practical radiomics-clinical combined model based on pet/ct data and nomogram predicts egfr mutation in lung adenocarcinoma. *Eur Radiol* (2021) 31(8):6259–68. doi: 10.1007/s00330-020-07676-x

35. Yip SS, Kim J, Coroller TP, Parmar C, Velazquez ER, Huynh E, et al. Associations between somatic mutations and metabolic imaging phenotypes in non-small cell lung cancer. *J Nucl Med* (2017) 58(4):569–76. doi: 10.2967/jnumed.116.181826

36. Yip SSF, Parmar C, Kim J, Huynh E, Mak RH, Aerts H. Impact of experimental design on pet radiomics in predicting somatic mutation status. *Eur J Radiol* (2017) 97:8–15. doi: 10.1016/j.ejrad.2017.10.009

37. Zhao HY, Su YX, Zhang LH, Fu P. Prediction model based on 18f-fdg pet/ct radiomic features and clinical factors of egfr mutations in lung adenocarcinoma. *Neoplasma* (2022) 69(1):233–41. doi: 10.4149/neo_2021_201222N1388

38. Nair JKR, Saeed UA, McDougall CC, Sabri A, Kovacina B, Raidu BVS, et al. Radiogenomic models using machine learning techniques to predict egfr mutations in non-small cell lung cancer. *Can Assoc Radiol J* (2021) 72(1):109–19. doi: 10.1177/0846537119899526

39. Yang B, Ji H, Zhong J, Ma L, Zhong J, Dong H, et al. Value of (18)F-fdg pet/ct-based radiomics nomogram to predict survival outcomes and guide personalized targeted therapy in lung adenocarcinoma with egfr mutations. *Front Oncol* (2020) 10:567160. doi: 10.3389/fonc.2020.567160

40. Zhang M, Bao Y, Rui W, Shangguan C, Liu J, Xu J, et al. Performance of (18)F-fdg pet/ct radiomics for predicting egfr mutation status in patients with non-small cell lung cancer. *Front Oncol* (2020) 10:568857. doi: 10.3389/fonc.2020.568857

41. Zhang J, Zhao X, Zhao Y, Zhang J, Zhang Z, Wang J, et al. Value of pre-therapy (18)F-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(5):1137–46. doi: 10.1007/s00259-019-04592-1

42. Li X, Yin G, Zhang Y, Dai D, Liu J, Chen P, et al. Predictive power of a radiomic signature based on (18)F-fdg pet/ct images for egfr mutational status in nsclc. *Front Oncol* (2019) 9:1062. doi: 10.3389/fonc.2019.01062

43. Wang Z, Yin G, Li X, Xu W. Value of machine learning and 18F-FDG PET/CT radiomics features in lung adenocarcinoma EGFR mutation subtypes prediction. *Chin J Nucl Med Mol Imaging* (2021) 41(08):479–85. doi: 10.3760/cma.j.cn321828-20201105-00401

44. Uribe CF, Mathotaarachchi S, Gaudet V, Smith KC, Rosa-Neto P, Bénard F, et al. Machine learning in nuclear medicine: part 1-introduction. *J Nucl Med* (2019) 60 (4):451–8. doi: 10.2967/jnumed.118.223495

45. Chiu HY, Chao HS, Chen YM. Application of artificial intelligence in lung cancer. *Cancers (Basel)* (2022) 14(6):1370. doi: 10.3390/cancers14061370

46. Gao J, Niu R, Shi Y, Shao X, Jiang Z, Ge X, et al. The predictive value of [(18)F] Fdg pet/ct radiomics combined with clinical features for egfr mutation status in different clinical staging of lung adenocarcinoma. *EJNMMI Res* (2023) 13(1):26. doi: 10.1186/s13550-023-00977-4

47. Zhang T, Liu Z, Lin L, Han T, Long F, Guo H, et al. Detection of the gene mutation of epidermal growth factor receptor in lung adenocarcinoma by radiomic features from a small amount of pet data. *Nucl Med Commun* (2023) 44(9):795–802. doi: 10.1097/MNM.0000000000001718

48. Zuo Y, Liu Q, Li N, Li P, Zhang J, Song S. Optimal (18)F-fdg pet/ct radiomics model development for predicting egfr mutation status and prognosis in lung adenocarcinoma: A multicentric study. *Front Oncol* (2023) 13:1173355. doi: 10.3389/fonc.2023.1173355

49. Yang L, Xu P, Li M, Wang M, Peng M, Zhang Y, et al. Pet/ct radiomic features: A potential biomarker for egfr mutation status and survival outcome prediction in nsclc patients treated with tkis. *Front Oncol* (2022) 12:894323. doi: 10.3389/fonc.2022.894323

50. Agüloğlu N, Aksu A, Akyol M, Katgı N, Doksöz T. Importance of pretreatment 18f-fdg pet/ct texture analysis in predicting egfr and alk mutation in patients with non-small cell lung cancer. *Nuklearmedizin Nucl Med* (2022) 61(6):433–9. doi: 10.1055/a-1868-4918

51. Shiri I, Amini M, Nazari M, Hajianfar G, Haddadi Avval A, Abdollahi H, et al. Impact of feature harmonization on radiogenomics analysis: prediction of egfr and kras mutations from non-small cell lung cancer pet/ct images. *Comput Biol Med* (2022) 142:105230. doi: 10.1016/j.compbiomed.2022.105230

52. Ruan D, Fang J, Teng X. Efficient 18f-fluorodeoxyglucose positron emission tomography/computed tomography-based machine learning model for predicting epidermal growth factor receptor mutations in non-small cell lung cancer. *Q J Nucl Med Mol Imaging* (2022). doi: 10.23736/s1824-4785.22.03441-0

53. Shiri I, Maleki H, Hajianfar G, Abdollahi H, Ashrafinia S, Hatt M, et al. Next-generation radiogenomics sequencing for prediction of egfr and kras mutation status in nsclc patients using multimodal imaging and machine learning algorithms. *Mol Imaging Biol* (2020) 22(4):1132–48. doi: 10.1007/s11307-020-01487-8

54. Yang B, Ji HS, Zhou CS, Dong H, Ma L, Ge YQ, et al. (18)F-fluorodeoxyglucose positron emission tomography/computed tomography-based radiomic features for prediction of epidermal growth factor receptor mutation status and prognosis in patients with lung adenocarcinoma. *Trans Lung Cancer Res* (2020) 9(3):563–74. doi: 10.21037/tlcr-19-592

55. Liu Q, Sun D, Li N, Kim J, Feng D, Huang G, et al. Predicting egfr mutation subtypes in lung adenocarcinoma using (18)F-fdg pet/ct radiomic features. *Trans Lung Cancer Res* (2020) 9(3):549–62. doi: 10.21037/tlcr.2020.04.17

56. Koyasu S, Nishio M, Isoda H, Nakamoto Y, Togashi K. Usefulness of gradient tree boosting for predicting histological subtype and egfr mutation status of non-small cell lung cancer on (18)F fdg-pet/ct. *Ann Nucl Med* (2020) 34(1):49–57. doi: 10.1007/s12149-019-01414-0

57. Jiang M, Zhang Y, Xu J, Ji M, Guo Y, Guo Y, et al. Assessing egfr gene mutation status in non-small cell lung cancer with imaging features from pet/ct. *Nucl Med Commun* (2019) 40(8):842–9. doi: 10.1097/mnm.0000000000001043

58. Papadimitroulas P, Brocki L, Christopher Chung N, MarChadour W, Vermet F, Gaubert L, et al. Artificial intelligence: deep learning in oncological radiomics and

challenges of interpretability and data harmonization. *Physica Med PM* (2021) 83:108–21. doi: 10.1016/j.ejmp.2021.03.009

59. Naimi AI, Balzer LB. Stacked generalization: an introduction to super learning. *Eur J Epidemiol* (2018) 33(5):459–64. doi: 10.1007/s10654-018-0390-z

60. Li Z, Wang Y, Yu J, Guo Y, Cao W. Deep learning based radiomics (Dlr) and its usage in noninvasive idh1 prediction for low grade glioma. *Sci Rep* (2017) 7(1):5467. doi: 10.1038/s41598-017-05848-2

61. Ypsilantis PP, Siddique M, Sohn HM, Davies A, Cook G, Goh V, et al. Predicting response to neoadjuvant chemotherapy with pet imaging using convolutional neural networks. *PloS One* (2015) 10(9):e0137036. doi: 10.1371/journal.pone.0137036

62. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, 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

63. Xiao Z, Cai H, Wang Y, Cui R, Huo L, Lee EY, et al. Deep learning for predicting epidermal growth factor receptor mutations of non-small cell lung cancer on pet/ct images. *Quant Imaging Med Surg* (2023) 13(3):1286–99. doi: 10.21037/qims-22-760

64. Yin G, Wang Z, Song Y, Li X, Chen Y, Zhu L, et al. Prediction of egfr mutation status based on 18f-fdg pet/ct imaging using deep learning-based model in lung adenocarcinoma. *Front Oncol* (2021) 11:709137. doi: 10.3389/fonc.2021.709137

65. Chen S, Han X, Tian G, Cao Y, Zheng X, Li X, et al. Using stacked deep learning models based on pet/ct images and clinical data to predict egfr mutations in lung cancer. *Front Med* (2022) 9:1041034. doi: 10.3389/fmed.2022.1041034

66. Mu W, Jiang L, Zhang J, Shi Y, Gray JE, Tunali I, et al. Non-invasive decision support for nsclc treatment using pet/ct radiomics. *Nat Commun* (2020) 11(1):5228. doi: 10.1038/s41467-020-19116-x

67. Boopathiraja S, Punitha V, Kalavathi P, Surya Prasath VB. Computational 2d and 3d medical image data compression models. *Arch Comput Methods Eng* (2022) 29 (2):975–1007. doi: 10.1007/s11831-021-09602-w

68. Hosseinzadeh Taher MR, Haghighi F, Feng R, Gotway MB, Liang J. A systematic benchmarking analysis of transfer learning for medical image analysis. *Domain Adapt Represent Transf Afford Healthc AI Resour Divers Global Health* (2021) 12968:3–13. doi: 10.1007/978-3-030-87722-4_1

69. Yang J, Huang X, He Y, Xu J, Yang C, Xu G, et al. Reinventing 2d convolutions for 3d images. *IEEE J Biomed Health Inf* (2021) 25(8):3009–18. doi: 10.1109/jbhi.2021.3049452

70. You A, Kim JK, Ryu IH, Yoo TK. Application of generative adversarial networks (Gan) for ophthalmology image domains: A survey. *Eye Vis (Lond)* (2022) 9(1):6. doi: 10.1186/s40662-022-00277-3

71. Apostolopoulos ID, Papathanasiou ND, Apostolopoulos DJ, Panayiotakis GS. Applications of generative adversarial networks (Gans) in positron emission tomography (Pet) imaging: A review. *Eur J Nucl Med Mol Imaging* (2022) 49 (11):3717–39. doi: 10.1007/s00259-022-05805-w

72. Pan Z, Song L, Song W, Jin Z. Progress of deep learning in imaging prediction of lung adenocarcinoma aggressiveness. *Chin J Radiol* (2023) 57(09):1018–22. doi: 10.3760/cma.j.cn112149-20230316-00192

73. Hutson M. Has artificial intelligence become alchemy? *Sci (New York NY)* (2018) 360(6388):478. doi: 10.1126/science.360.6388.478

74. Aerts HJ. The potential of radiomic-based phenotyping in precision medicine: A review. *JAMA Oncol* (2016) 2(12):1636–42. doi: 10.1001/jamaoncol.2016.2631

75. Zhang H, Ogasawara K. Grad-cam-based explainable artificial intelligence related to medical text processing. *Bioeng (Basel Switzerland)* (2023) 10(9):1070. doi: 10.3390/bioengineering10091070

76. Ibrahim A, Primakov S, Beuque M, Woodruff HC, Halilaj I, Wu G, et al. Radiomics for precision medicine: current challenges, future prospects, and the proposal of a new framework. *Methods* (2021) 188:20–9. doi: 10.1016/j.ymeth.2020.05.022

77. Zwanenburg A, Vallières M, Abdalah MA, Aerts H, Andrearczyk V, Apte A, et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology* (2020) 295 (2):328–38. doi: 10.1148/radiol.2020191145

78. Jochems A, Deist TM, El Naqa I, Kessler M, Mayo C, Reeves J, et al. Developing and Validating a Survival Prediction Model for Nsclc Patients through Distributed Learning across 3 Countries. *Int J Radiat Oncol Biol Phys* (2017) 99(2):344–52. doi: 10.1016/j.ijrobp.2017.04.021

79. Jochems A, Deist TM, van Soest J, Eble M, Bulens P, Coucke P, et al. 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):459–67. doi: 10.1016/j.radonc.2016.10.002

80. Giang TT, Nguyen TP, Tran DH. Stratifying patients using fast multiple kernel learning framework: case studies of Alzheimer's disease and cancers. *BMC Med Inf Decision Making* (2020) 20(1):108. doi: 10.1186/s12911-020-01140-y

81. Gao Y, Zhou R, Lyu Q. Multiomics and machine learning in lung cancer prognosis. *J Thorac Dis* (2020) 12(8):4531–5. doi: 10.21037/jtd-2019-itm-013

82. Theodoropoulos AS, Gkiozos I, Kontopyrgias G, Charpidou A, Kotteas E, Kyrgias G, et al. Modern radiopharmaceuticals for lung cancer imaging with positron emission tomography/computed tomography scan: A systematic review. *SAGE Open Med* (2020) 8:2050312120961594. doi: 10.1177/2050312120961594

83. Napel S, Mu W, Jardim-Perassi BV, Aerts H, Gillies RJ. Quantitative imaging of cancer in the postgenomic era: radio(Geno)Mics, deep learning, and habitats. *Cancer* (2018) 124(24):4633–49. doi: 10.1002/cncr.31630

84. Huang S, Yang J, Shen N, Xu Q, Zhao Q. Artificial intelligence in lung cancer diagnosis and prognosis: current application and future perspective. *Semin Cancer Biol* (2023) 89:30–7. doi: 10.1016/j.semcancer.2023.01.006

Glossary

19 del	exon 19 deletion
21 L858R	exon 21 L858R missense
3D	three-dimensional
AB/Adaboost	adaptive boosting
ANN	artificial neural networks
AUC	area under the curve
BAG	bagging
BN	Bayesian network
ctDNA	circulating tumor DNA
CNN	convolutional neural network
CPU	central processing unit
DL	deep learning
DLR	deep learning-based radiomics
DLS	DL score
DT	decision tree
EGFR	epidermal growth factor receptor
FDG	fluorodeoxyglucose
GANs	Generative Adversarial Networks
GNB	Gaussian naive Bayes
GPU	graphics processing unit
Grad-CAM	gradient-weighted class activation mapping
IBSI	image biomarker standardization initiative
KNN	k-nearest neighbor
LASSO	least absolute shrinkage and selection operator
LR	logistic regression
MLP	multilayer perceptron
MLR	machine learning-based radiomics
MRI	magnetic resonance imaging
NB	naive Bayes
PET	positron emission tomography
QDA	quadratic discriminant analysis
Rad-score	radiomics score
ResNet	Residual network
RF	random forest
ROI	region of interest
SE-ResNet	Squeeze-and-excitation residual network
SGD	stochastic gradient descent
SResCNN	small-residual-convolutional-network

(Continued)

Continued

SUVmax	maximum standardized uptake value
SVM	support vector machine
TKI	tyrosine kinase inhibitors
TR	traditional radiomics
TSR	traditional statistics-based radiomics
XGBoost	eXtreme gradient boosting