



Predictive Radiomic Models for the Chemotherapy Response in Non-Small-Cell Lung Cancer based on Computerized-Tomography Images

Runsheng Chang¹, Shouliang Qi^{1,2*}, Yong Yue³, Xiaoye Zhang⁴, Jiangdian Song¹ and Wei Qian⁵

¹ College of Medicine and Biological Information Engineering, Northeastern University, Shenyang, China, ² Key Laboratory of Intelligent Computing in Medical Image, Ministry of Education, Northeastern University, Shenyang, China, ³ Department of Radiology, Shengjing Hospital of China Medical University, Shenyang, China, ⁴ Department of Oncology, Shengjing Hospital of China Medical University, Shenyang, China, ⁵ Department of Electrical and Computer Engineering, University of Texas at El Paso, El Paso, TX, United States

OPEN ACCESS

Edited by:

Bo Gao,
Affiliated Hospital of Guizhou Medical
University, China

Reviewed by:

Lin Lu,
Columbia University Irving Medical
Center, United States
Wei Mu,
Moffitt Cancer Center, United States

*Correspondence:

Shouliang Qi
qjisl@bmie.neu.edu.cn
orcid.org/0000-0003-0977-1939

Specialty section:

This article was submitted to
Cancer Imaging and Image-
directed Interventions,
a section of the journal
Frontiers in Oncology

Received: 25 December 2020

Accepted: 16 June 2021

Published: 07 July 2021

Citation:

Chang R, Qi S, Yue Y, Zhang X,
Song J and Qian W (2021)
Predictive Radiomic Models for the
Chemotherapy Response in Non-
Small-Cell Lung Cancer based on
Computerized-Tomography Images.
Front. Oncol. 11:646190.
doi: 10.3389/fonc.2021.646190

The heterogeneity and complexity of non-small cell lung cancer (NSCLC) tumors mean that NSCLC patients at the same stage can have different chemotherapy prognoses. Accurate predictive models could recognize NSCLC patients likely to respond to chemotherapy so that they can be given personalized and effective treatment. We propose to identify predictive imaging biomarkers from pre-treatment CT images and construct a radiomic model that can predict the chemotherapy response in NSCLC. This single-center cohort study included 280 NSCLC patients who received first-line chemotherapy treatment. Non-contrast CT images were taken before and after the chemotherapy, and clinical information were collected. Based on the Response Evaluation Criteria in Solid Tumors and clinical criteria, the responses were classified into two categories: response ($n = 145$) and progression ($n = 135$), then all data were divided into two cohorts: training cohort (224 patients) and independent test cohort (56 patients). In total, 1629 features characterizing the tumor phenotype were extracted from a cube containing the tumor lesion cropped from the pre-chemotherapy CT images. After dimensionality reduction, predictive models of the chemotherapy response of NSCLC with different feature selection methods and different machine-learning classifiers (support vector machine, random forest, and logistic regression) were constructed. For the independent test cohort, the predictive model based on a random-forest classifier with 20 radiomic features achieved the best performance, with an accuracy of 85.7% and an area under the receiver operating characteristic curve of 0.941 (95% confidence interval, 0.898–0.982). Of the 20 selected features, four were first-order statistics of image intensity and the others were texture features. For nine features, there were significant differences between the response and progression groups ($p < 0.001$). In the response group, three features, indicating heterogeneity, were overrepresented and one feature indicating homogeneity was underrepresented. The proposed radiomic model with pre-chemotherapy CT features can predict the chemotherapy response of patients with

non-small cell lung cancer. This radiomic model can help to stratify patients with NSCLC, thereby offering the prospect of better treatment.

Keywords: lung cancer, radiomics, CT images, chemotherapy response, machine learning

INTRODUCTION

According to the Global Cancer Incidence and Mortality Report in 2018, lung cancer was the most commonly diagnosed cancer (11.6% of all cases) and the leading cause of cancer deaths (18.4% of all cancer deaths) (1, 2), with non-small cell lung cancer (NSCLC) accounting for 80% to 85% of all lung cancers. However, despite considerable advances in diagnosis and treatment over the years, the 5-year survival rate of lung-cancer patients is currently less than 18% (54% for localized-stage disease, 26% for regional stage, and 4% for distant stage) (3–5). As reported, 74% of cases are diagnosed at the regional or distant stage (3), and any patient diagnosed as being at stage IIIA or IV is virtually unresectable and has no choice but to receive radiotherapy or chemotherapy with severe side effects.

The heterogeneity and complexity of NSCLC tumors mean that NSCLC patients at the same stage can have different chemotherapy prognoses (6). According to the Response Evaluation Criteria in Solid Tumors (RECIST), treatment responses can be divided into four types: (i) complete response (CR), (ii) partial response (PR), (iii) progressive disease (PD), and (iv) stable disease (SD) (7). However, there are currently few quantitative criteria or models that can predict the NSCLC chemotherapy response from pre-treatment information (8). Accurate predictive models could recognize NSCLC patients likely to respond to chemotherapy so that they can be given personalized and effective treatment.

Radiomics is a potential bridge between medical imaging and personalized medicine (9, 10). In this approach, artificial intelligence is used to convert image data from a lesion region into a high-dimensional feature space and to construct predictive models for various clinical outcomes (11, 12). Radiomics has been used successfully in biological oncology for detection, differential diagnosis, phenotype or subtype stratification, prognosis prediction, and even the prediction of invasiveness and gene mutation status (13–17).

Radiomics has achieved exceptional results in predicting the prognosis of NSCLC treatment with survival as the endpoint. For example, based on a dataset of 1194 NSCLC patients treated with either radiotherapy or surgery, Hosny et al. constructed convolutional neural network models that could predict 2-year overall survival from pre-treatment computerized-tomography (CT) images with an accuracy of 70% (18). For 179 stage-III NSCLC patients treated with definitive radiotherapy and chemotherapy, Xu et al. designed a deep-learning model using time-series CT scans and found that it was significantly predictive of survival and cancer-specific outcomes (4). Wang et al. collected CT images and clinical information for 173 NSCLC patients and trained a radiomic model that could predict the range of a patient's prognosis survival time (6).

Song et al. established a Cox regression model with a least absolute shrinkage and selection operator for CT images to predict the progression-free survival of stage-IV epidermal growth factor receptor (EGFR)-mutated NSCLC patients being treated with EGFR tyrosine kinase inhibitors (19). Paul et al. used a transfer-learning model to extract deep features to predict short- and long-term survivors with lung adenocarcinoma with an accuracy of 90% (20). Lou et al. used deep-learning methods of pre-treatment CT scans to analyze survival and found an individualized radiation dose that gave an estimated probability of treatment failure of below 5% (21).

Moreover, predicting the chemotherapy response in NSCLC earlier in the course treatment is very useful and promising. It can help clinicians make decisions on whether to adapt, intensify, or alter treatment plans early and improve patient outcomes (22). Compared with the long-term endpoint of survival, treatment response is a short-term prognosis endpoint that may help to identify precisely those NSCLC patients who are likely to benefit from chemotherapy.

However, to the best of our knowledge, few predictive models use chemotherapy response in NSCLC as the endpoint. Chen et al. proposed a radiomic model to predict NSCLC lesions shrinkage during treatment with either pembrolizumab or combinations of chemotherapy and pembrolizumab (23). The model used features extracted from lesions, margins, and blood vessels and reached an area under the curve (AUC) of 0.73 in a test cohort with 176 patients. Seki et al. had demonstrated the usefulness of CT and Positron Emission Tomography (PET)/CT in the early prediction of chemoradiotherapy in NSCLC (24). In the present study, we constructed a radiomic model that used pre-chemotherapy CT images to predict the NSCLC chemotherapy response.

MATERIALS AND METHODS

Data Acquisition

We enrolled 622 patients with lung cancer being treated at Shengjing Hospital of China Medical University between 2014 and 2019. The parameters for CT images acquisition are listed in **Table 1**. As shown in **Figure 1**, after two steps of exclusion criteria, 280 patients were included in our study. Their clinical characteristics are given in **Table 2**. This study was approved by the ethics committee of Shengjing Hospital of China Medical University and the waived informed consent forms were waived because it is a retrospective study.

Label of Treatment Response

NSCLC tumors were categorized according to RECIST jointly by an experienced radiologist and an experienced oncologist: (i) CR:

TABLE 1 | Parameters for CT image acquisition.

Parameter	Value
kVp (kV)	120
X-ray tube current (mean ±S.D.) (mA)	215.274 ± 70.816
Slice thickness (mm)	2.5 (n = 14); 3.0 (n = 244); 5.0 (n = 22)
Pixel size (mm)	0.783 ± 0.074
CT scanner manufacturer	GE Medical (n = 10), Siemens (n = 11), Toshiba (n = 13), Philips (n = 246)

disappearance of all target lesions, (ii) PR: at least 30% decrease in the sum of the diameters of the target lesions, (iii) PD: at least 20% increase in the sum of the diameters of the target lesions, and (iv) SD: neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD.

According to the requirement of clinical applications and radiologist’s advice, we had excluded the patients of SD in our study. The CR and PR patients were combined into a category named “response” and the PD patients were included into a category named “progression.” Finally, 145 NSCLC patients were labeled as response and 135 were labeled as progression.

Overview of Study Procedure

We split the total 280 patients into the training cohort (n = 224) and the independent test cohort (n=56). As shown in **Figure 2**, the study procedure had six steps. First, by comparing CT images taken before and after chemotherapy, responses were determined as being either response or progression. Second, in the preprocessing step, the tumor lesion in the pre-chemotherapy CT images was cropped to a cube. Third, radiomic features were extracted from the cropped cube. Fourth, discriminative features were selected with different methods and analyzed. Fifth, the selected radiomic features, labels, and clinical information were used to train the different models using the training cohort. Finally, the performance of the radiomic models was evaluated using the independent test cohort. The evaluation measures included the AUC of receiver operating characteristic (ROC) curve, confusion matrix, recall, precision, and F-score. The best cutoff value of ROC curve to calculate the confusion matrix and related measures was determined, whereas Youden index reach the maximum value.

Feature Extraction

First, all the pre-treatment CT images for the patients were interpolated into voxels of 0.750 × 0.750 × 3.000 mm. To include the characteristic information [both the tumor lesion and its habitat (8, 12)], we cropped a 64 × 64 × 32 cube from the lesion center[determined by software 3D Slicer (25)]. In our study, this cube can include the largest lesion and no cube includes more than one lesion. Because all the patients are in the advanced stage

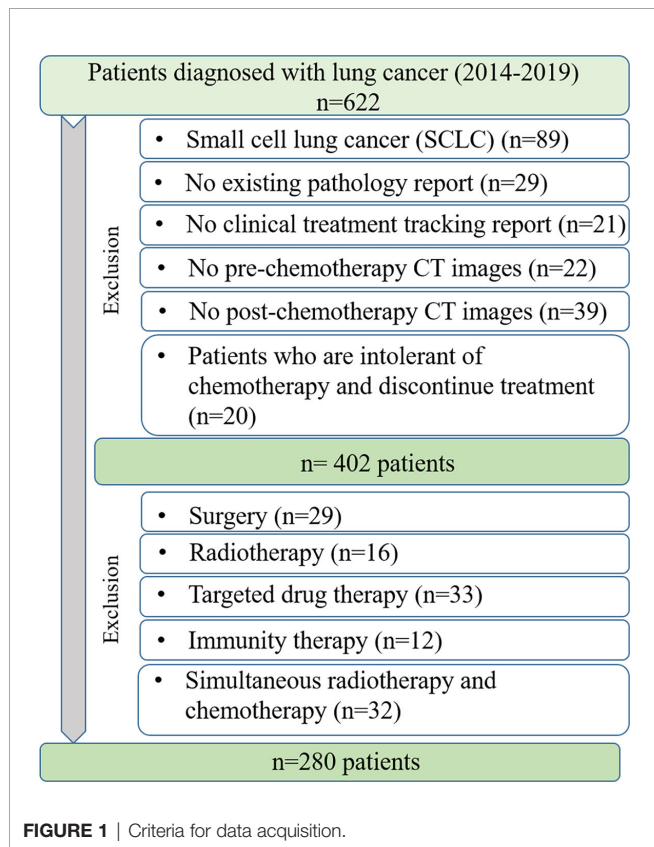


FIGURE 1 | Criteria for data acquisition.

TABLE 2 | Clinical characteristics of NSCLC patients.

Characteristics		Response group	Progression group	p value
Number of patients		145	135	–
Gender	Male	79	77	0.605
	Female	66	58	
Age(years)		63.864 ± 10.042	64.402 ± 9.713	0.437
Histological type	Adenocarcinoma	119	109	0.201
	Squamous cell carcinoma	26	26	
Smoking status	Ever	49	74	0.002
	Never	96	61	
Number of treatment courses		4.492 ± 1.603	3.681 ± 1.396	<0.001
Chemotherapy drug	AP	53	36	–
	GP	29	34	
	TP	31	28	
	DP	32	37	

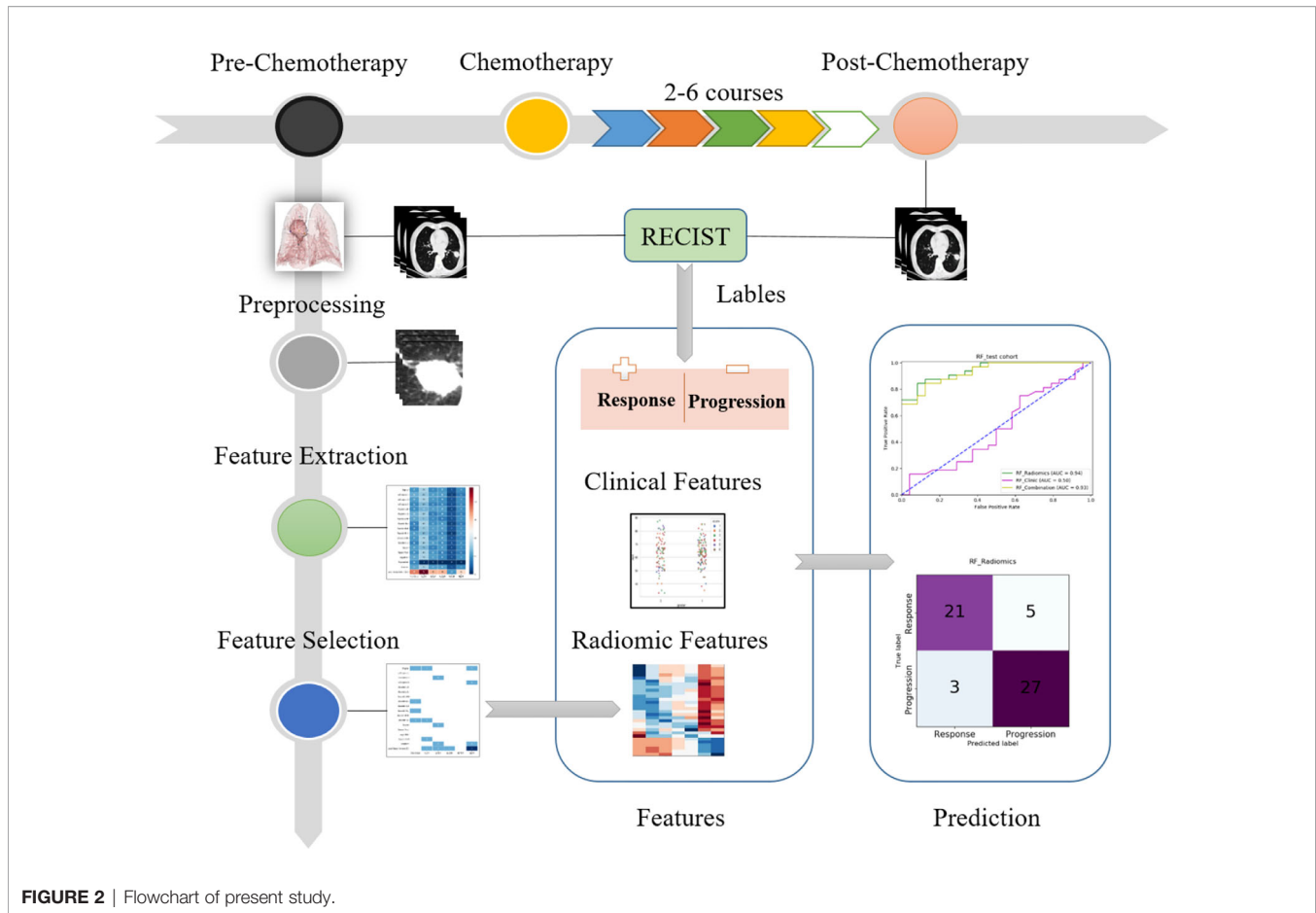


FIGURE 2 | Flowchart of present study.

and have taken chemotherapy, the lesion cannot be very small compared with the cube of $64 \times 64 \times 32$.

Then, we used the open-source Python package PyRadiomics to extract radiomic features from each cube (26). In total, 1,927 features were extracted from the original CT images, of which 1,629 meaningful features were used. Because the shapes and sizes were same for all cubes, the related features had no discriminative capability and were excluded ($n=298$).

It is should be noted that according to a previous study, besides the intra-tumor region, the extra-tumor marginal region may also provide predictive information for the treatment response and overall survival (8, 27, 28). Therefore, the features extracted from the cube in our study represent the characteristics not only of intra-tumor region but also of extra-tumor region.

Feature Selection

Next, we used three algorithms to select discriminative features and passed them into the model for training and testing: random forest (RF) (29), mRMR (max-relevance and min-redundancy) (30), and relief (31). RF can be easily applied to select the critical features by ranking the importance score of features. It belongs to the embedded feature selection using SelectFromModel. Actually, the package of sci-kit learn has provided two ways of feature selection by using RF: (1) mean decrease impurity; (2) mean decrease accuracy. In our study, we directly used the

way of “mean decrease accuracy.” Both mRMR and relief are the feature selection methods based on filter and publicly available (13).

Using the rule of thumb given by Gillies et al., with each feature corresponding to 10 samples in a binary classifier (12), we selected 20 features to represent each patient to do the next classification, and the performance of RF, mRMR, and relief for the feature selection was compared.

Construction of Predictive Models

We selected three representative machine-learning classifiers: support vector machine (SVM), RF, and logistic regression (LR). We constructed a model to clarify the role of clinical information (gender, smoking status, age, pathology, course of treatment, and medicine), and we constructed another model with both clinical and radiomic features to assess whether that combination increased the predictive performance. Moreover, we also constructed a model with two clinical features of smoking status and course of treatment because there was a significant difference between the response and progression groups for these two features (Table 2). Correspondingly, a model with the combination of two significant clinical features and the selected radiomic features was constructed.

The optimal parameters of the model were determined by grid search technique and 10-fold cross-validation. Specifically,

for each grid of parameters, the performance of the model was evaluated by the average of 10-fold cross-validation. The optimal parameters were determined after traversing all grids. During the 10-fold cross-validation, the training data were divided into 10 folds. For each of the 10 “folds,” a model was trained by using nine folds as the training data and validated by the remaining fold. With the determined optimal parameters, the model was retrained by all training data ($n=224$). Finally, the independent test cohort ($n=56$) was used to evaluate the retrained model and gave the performance measures. All these procedures were performed by strictly following the document given by the Scikit learn website (https://scikit-learn.org/stable/modules/cross_validation.html).

To find the optimal parameters in classification models, we used the grid search with cross-validation (GridSearchCV) to traverse the parameters within a certain range and with a specific interval. In SVM, the kernel parameter was set as “linear” or “rbf” (radial basis function); the parameter C ranged from 1 to 5 with the interval of 1; the gamma parameter was set as 0.125, 0.25, 0.5, 1, 2, or 4. Through the ten-fold cross-validation of the training cohort in each grid, the optimal value (or setting) of the kernel, C, and gamma were determined as “linear,” 3, and 1, respectively. In RF, $n_estimators$ parameter ranged from 20 to 2000 with the interval of 10; $max_features$ parameter was set as 2 or 3; min_sample_leaf ranged from 1 to 50 with the interval of 1. The optimal value of $n_estimators$, $max_features$, and min_sample_leaf was determined as 100, 3, and 2, respectively. In LR, C parameter ranged from 1 to 5 with the interval of 1; the penalty item was set as l1 or l2. By the same way, the optimal value of C and penalty item was determined as 3 and l1, respectively.

RESULTS

Clinical Characteristics

As shown in **Figure 3A** and **Table 2**, there was no significant difference in gender between the response and progression groups. Similarly, there was no significant difference for age or histological type. For both groups, there were more patients with

adenocarcinoma than with squamous cell carcinoma (119 vs. 26; 109 vs. 26). The progression group had a higher percentage of smokers than the response group [54.8% (74/135) vs. 33.8% (49/145)]. The response group had more treatment courses than the progression group (4.492 ± 1.603 vs. 3.681 ± 1.396).

A platinum-based dual-drug regimen is the gold standard for the first-line treatment of advanced NSCLC. In our study, we included four common chemotherapy regimens: (i) AP: cisplatin or carboplatin combined with pemetrexed ($n = 53$ for response and $n = 36$ for progression), (ii) GP: cisplatin or carboplatin combined with gemcitabine ($n = 29$ for response and $n = 34$ for progression), (iii) TP: cisplatin or carboplatin combined with paclitaxel ($n = 31$ for response and $n = 28$ for progression), and (iv) DP: cisplatin or carboplatin combined with docetaxel ($n = 32$ for response and $n = 37$ for progression).

As shown in **Figure 3B**, for adenocarcinoma treated by AP, the response group had more patients than the progression group (47 vs. 27), but the opposite was the case for adenocarcinoma treated by GP (16 vs. 28). The situation for squamous cell carcinoma was the opposite of that for adenocarcinoma. For squamous cell carcinoma treated by AP, the response group had fewer patients than the progression group (6 vs. 9); for adenocarcinoma treated by GP, the response group had more patients than the progression group (13 vs. 6).

Radiomic Characteristics

Figure 4A shows the distribution of the 1,629 selected radiomic features. Of the six feature classes (columns), GLCM (gray-level co-occurrence matrix) had the most features (430/1629, 26.4%). Of the 18 filter classes (rows), local binary pattern (LBP) (3D) had the most features (279/1629, 17.1%).

Through three dimensionality reduction algorithms, the 20 most-informative features were selected from the 1629 radiomic features and input into the machine-learning classifiers. The distribution of these 20 features is shown in **Figure 4B**: gray-level dependence matrix (GLDM) had seven features, and first order, GLCM, and GLRLM (gray-level run-length matrix) each had four features. For the filter classes (rows), LBP (3D) had the most features (7/20, 35.0%). **Figure 4C** shows the mean values of these 20 highly informative features. In summary, 11 radiomic features

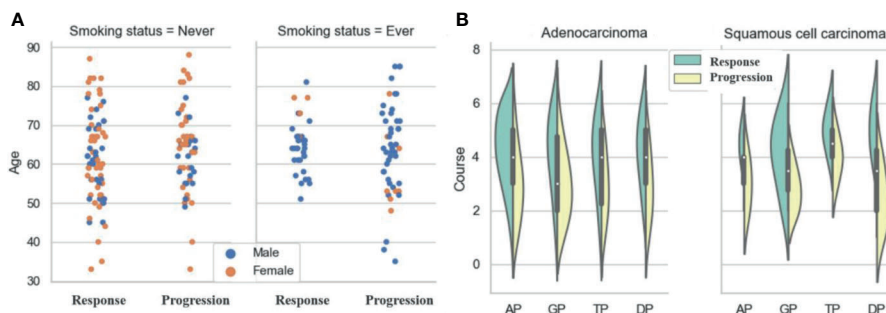


FIGURE 3 | Analysis of clinical characteristics: **(A)** Statistics of ages, genders and smoking status; **(B)** Statistics of treatment courses and chemotherapy drugs.

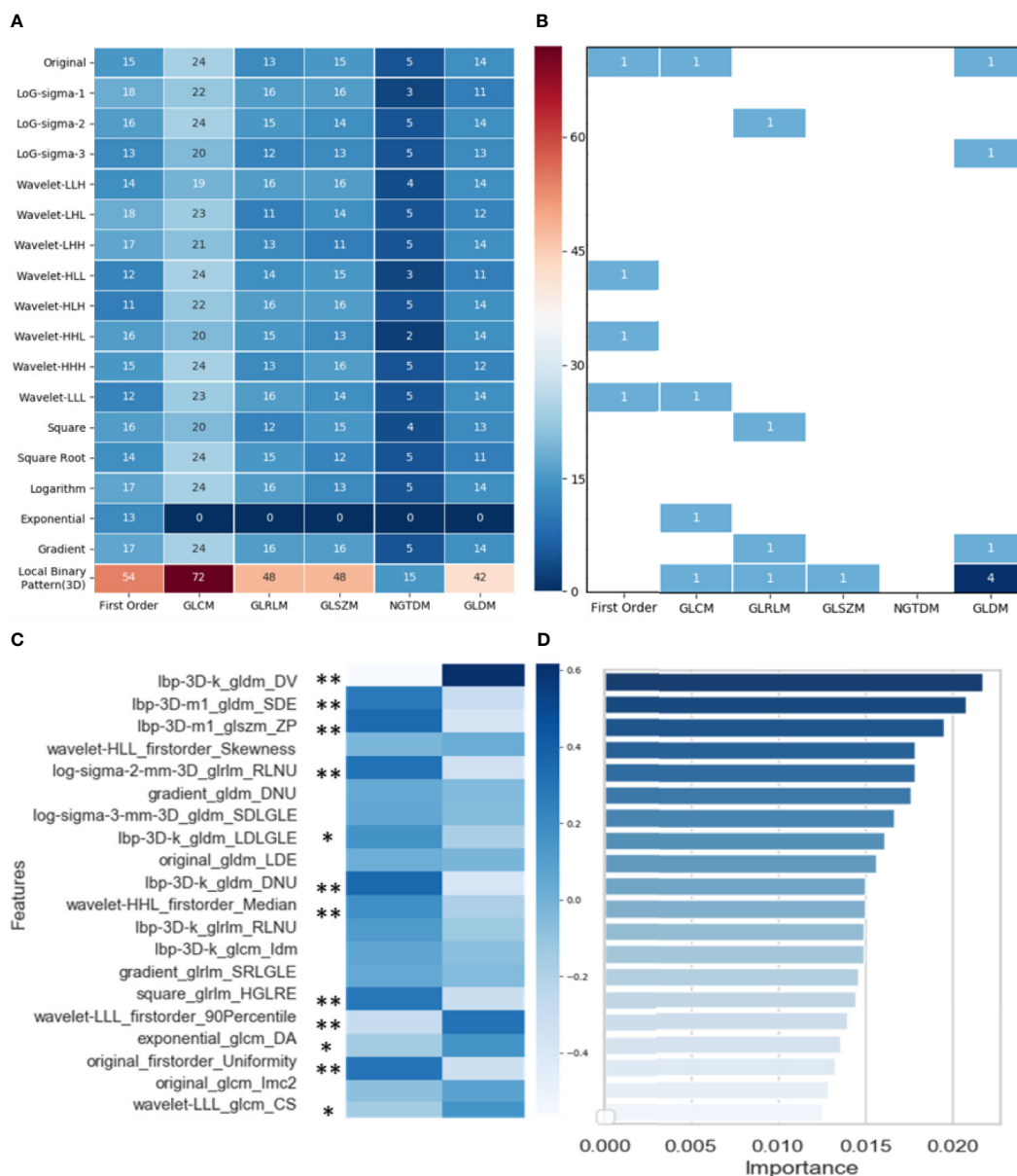


FIGURE 4 | Analysis and selection of radiomic features: **(A)** Distribution of 1629 extracted features; **(B)** Distribution of 20 selected features; **(C)** Mean values of 20 highly informative features and significance analysis between two groups (* $p < 0.05$, ** $p < 0.001$); **(D)** Importance of 20 selected features.

differed significantly between the response and progression groups [nine features with $p < 0.001$ (**) and two features with $p < 0.05$ (*)]. **Figure 4D** shows the importance of the 20 features selected *via* dimensionality reduction.

In the response group, the features small dependence emphasis (SDE), run length non-uniformity (RLNU), dependence non-uniformity (DNU), high gray level run emphasis (HGLRE), and uniformity are overrepresented, whereas dependence variance (DV) is underrepresented. SDE measures the distribution of small dependencies, with a larger value indicating less dependence and less-homogeneous textures. Similarly, larger values for RLNU and DNU indicate that there is

less homogeneity among run lengths and dependencies in the image, respectively. DV measures the variance independence size in the image. Overall, the representation of these features indicates that NSCLC tumors in the response group are more likely to be heterogeneous in CT images than are those in the progression group.

Dependence of Performance on the Feature Selection Method

We tried three different feature selection methods, RF, relief, and mRMR, to clarify their impact on the classification results. In **Figure 5**, in the training cohort, for the feature selection method

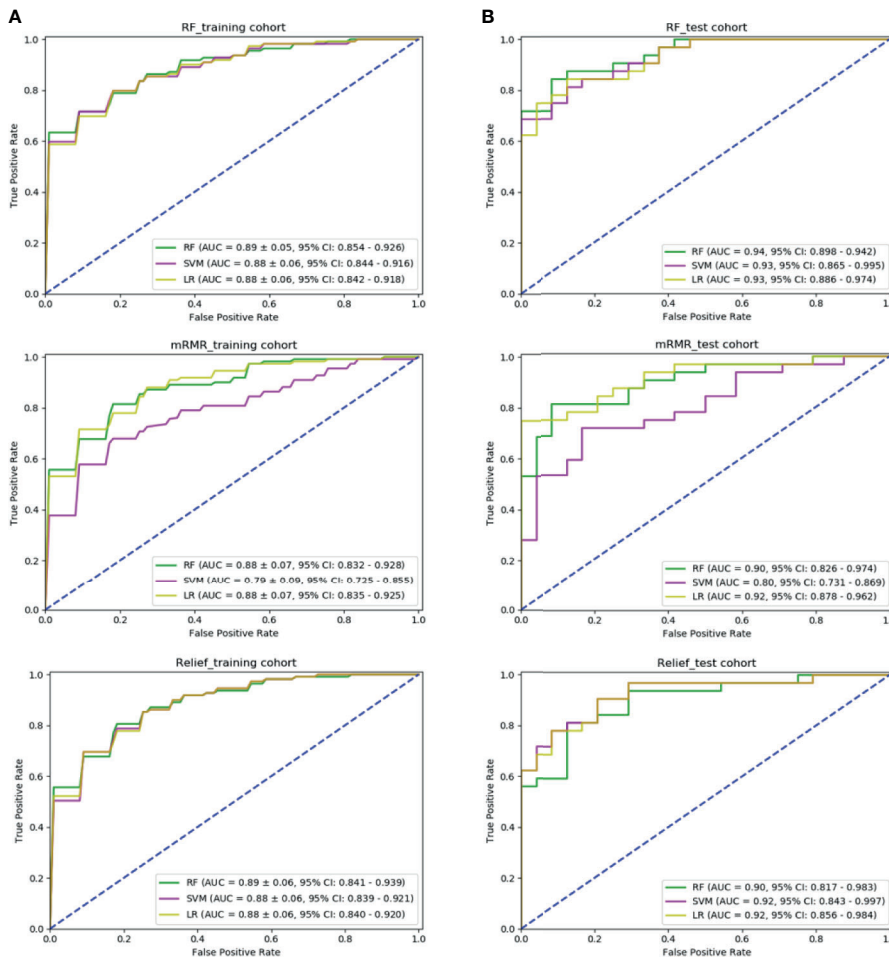


FIGURE 5 | Comparison of predictive models with different classifiers and different methods of feature selection: **(A)** ROC curve of three models using features selected by RF, mRMR, and relief in the training cohort; **(B)** ROC curve of three models using features selected by RF, mRMR, and relief in the independent test cohort.

of RF, the AUC of RF, SVM, and LR classification models was 0.891 ± 0.05 (95% confidence interval (CI), 0.854–0.926), 0.882 ± 0.06 (95% CI, 0.844–0.916), and 0.883 ± 0.06 (95% CI, 0.842–0.918), respectively. For mRMR, the AUC was 0.886 ± 0.07 (95% CI, 0.832–0.928), 0.798 ± 0.09 (95% CI, 0.725–0.855), and 0.889 ± 0.07 (95% CI, 0.835–0.925), respectively. For relief method, the AUC was 0.890 ± 0.06 (95% CI, 0.841–0.939), 0.886 ± 0.06 (95% CI, 0.839–0.921), and 0.888 ± 0.06 (95% CI, 0.840–0.920), respectively. In the independent test cohort, for the feature selection method of RF, the AUC of RF, SVM, and LR classification models were 0.941 (95% CI, 0.898–0.942), 0.932 (95% CI, 0.865–0.995), and 0.935 (95% CI, 0.886–0.974), respectively. For mRMR, the AUC was 0.901 (95% CI, 0.826–0.974), 0.804 (95% CI, 0.731–0.869), and 0.923 (95% CI, 0.878–0.962), respectively. For relief method, the AUC was 0.902 (95% CI, 0.817–0.983), 0.921 (95% CI, 0.843–0.997), and 0.926 (95% CI, 0.856–0.984), respectively. The combination of the feature selection by RF and the classification model of RF generated the

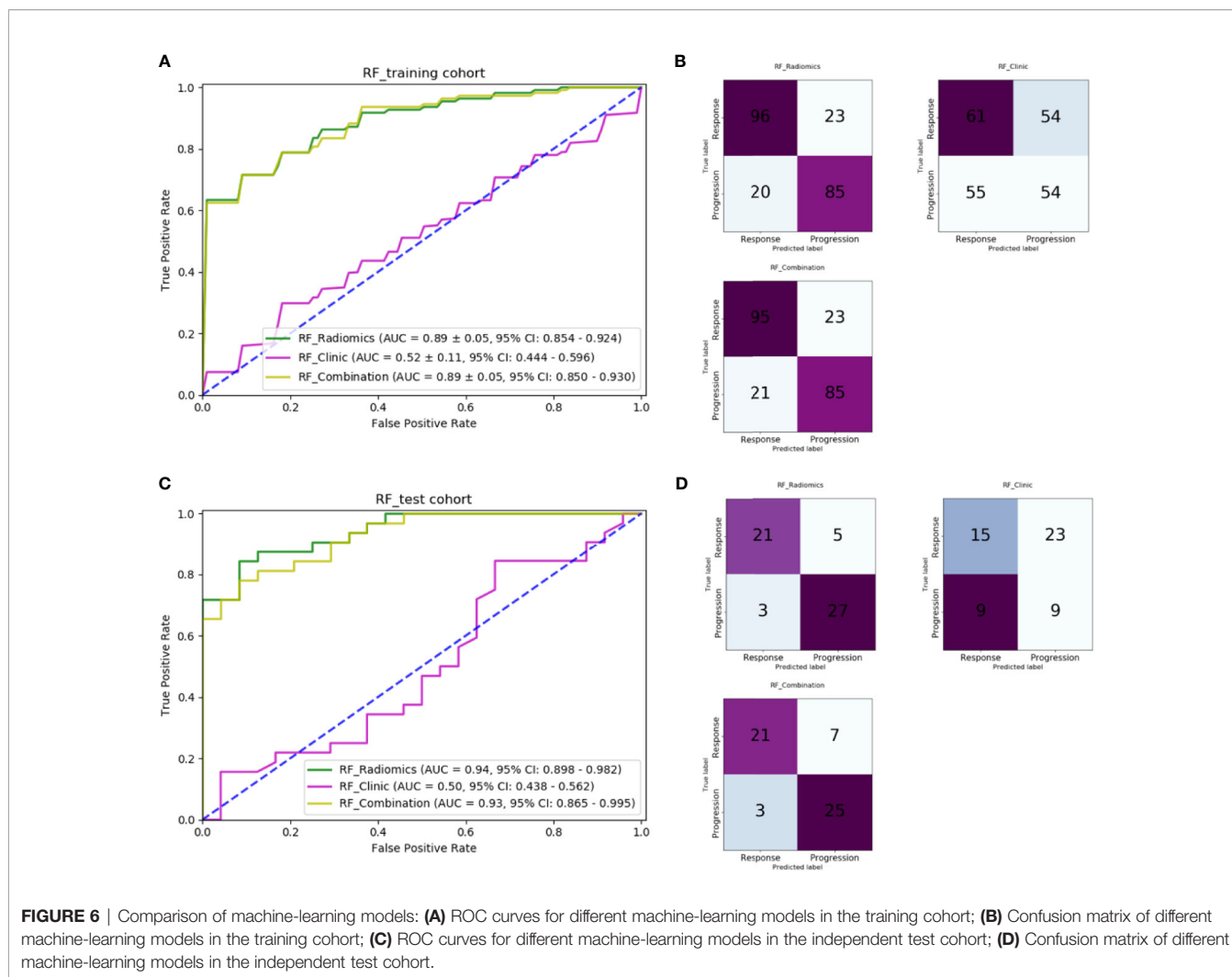
best predictive performance in both the training cohort and the independent test cohort.

Performance of Machine-Learning Models

Table 3 lists the performance of the three machine-learning models, and **Figure 6** shows the receiver operating characteristic (ROC) curves and the areas under the curve (AUC). In the training cohort, the RF model with radiomic features had the best performance, its AUC was 0.891 ± 0.05 (95% CI, 0.854–0.924). In the independent test cohort, the RF model with

TABLE 3 | Predictive performance of machine-learning models with radiomic features, clinical features, and combined features in the independent test cohort.

Classifier	Accuracy	AUC	Recall	Precision	F-score
RF_Radiomics	85.7%	0.941	0.875	0.808	0.840
RF_Clinic	42.9%	0.503	0.625	0.395	0.484
RF_Combination	82.1%	0.936	0.875	0.750	0.808



radiomic features had the best performance. Its AUC was 0.941 (95% CI, 0.898–0.982), and its accuracy, recall, precision, and F-score were 85.7%, 0.875, 0.808, and 0.840, respectively. The cutoff of ROC curve was 0.438.

The RF model with five clinical features had an AUC of only 0.523 ± 0.11 (95% CI: 0.444–0.596) in the training and 0.503 (95% CI: 0.438–0.562) in the independent test cohort, which indicates that clinical features played hardly any role in predicting chemotherapy response in our study. The cutoff of ROC curve in the independent test cohort was 0.459.

The RF model with combined clinical and radiomic features did not perform better than the RF model with only radiomic features. The AUC of training cohort was 0.890 ± 0.05 (95% CI: 0.850–0.930). In the independent test cohort, the accuracy, recall, precision, and F-score of the former were 82.1%, 0.875, 0.750, and 0.808, respectively, which are lower than those of the RF model with only radiomic features. The AUC was 0.930 (95% CI: 0.865–0.995) with a cutoff of 0.543.

The RF models with two significant clinical features are compared with those with five clinical features in **Figure 7**. The RF model with two significant clinical features had an AUC of 0.498 ± 0.15 (95% CI: 0.378–0.602) and 0.456 (95% CI: 0.398–0.502) in the training and independent test cohort, respectively. The RF model with the combination of two significant clinical features and the selected radiomic features had an AUC of 0.882 ± 0.06 (95% CI: 0.846–0.914) and 0.936 (95% CI: 0.868–0.992) in the training and independent test cohort, respectively. The performance of models with two significant clinical features was not as good as that of models with five clinical features.

Performance for Different Chemotherapy Drugs and Histological Subtypes

Table 4 presents the prediction accuracy of the RF model with 20 radiomic features for different chemotherapy drugs and histological subtypes. The prediction accuracy was higher for adenocarcinoma

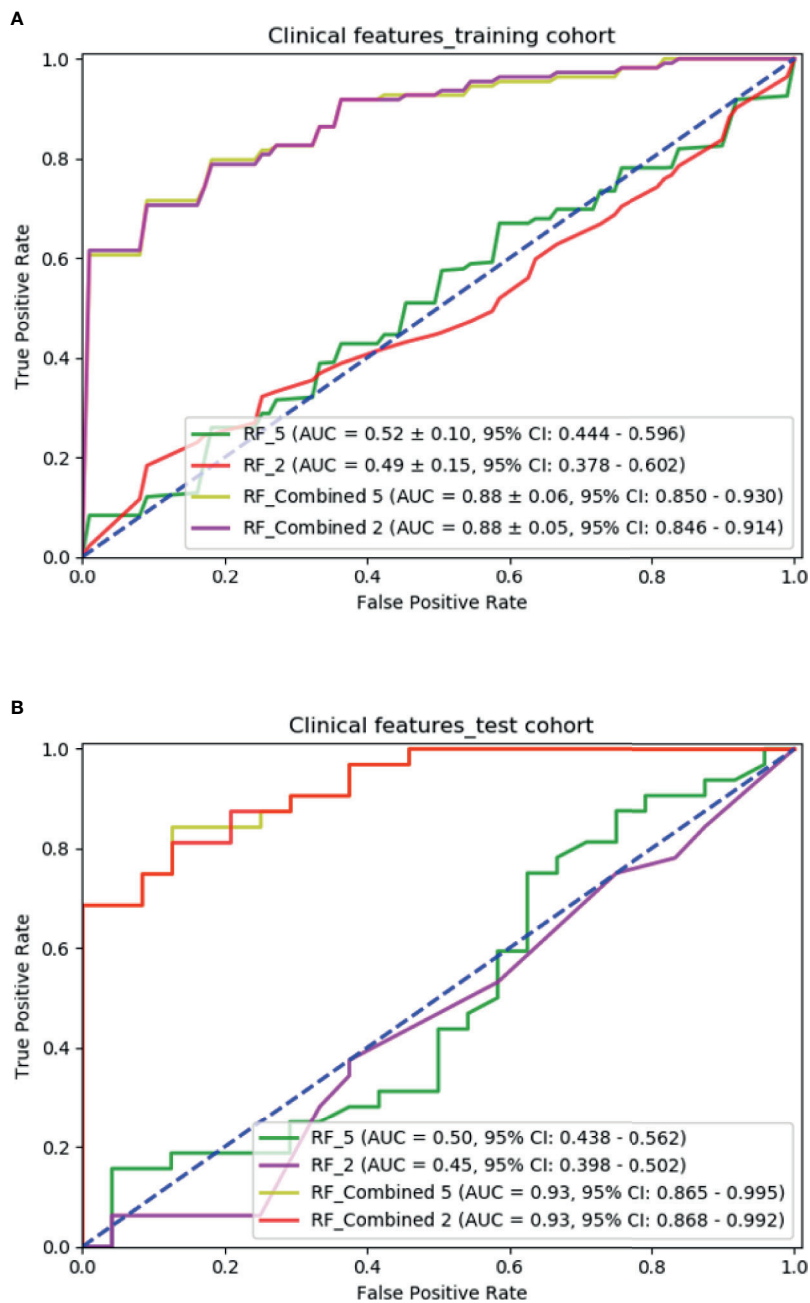


FIGURE 7 | Comparison of predictive models with different clinical features: **(A)** ROC curves for different models in the training cohort; **(B)** ROC curves for different models in the independent test cohort.

TABLE 4 | Prediction accuracy for different chemotherapy drugs and histological subtypes.

Chemotherapy drugs	Adenocarcinoma	Squamous cell carcinoma	Total
AP	85.2% (69/81)	80.0% (8/10)	84.6% (77/91)
GP	87.8% (43/49)	90.5% (19/21)	88.6% (62/70)
TP	70.0% (35/50)	54.5% (6/11)	67.2% (41/61)
DP	93.8% (45/48)	60.0% (6/10)	87.9% (51/58)
Total	84.2% (192/228)	75.0% (39/52)	-

than squamous cell carcinoma (84.2% vs. 75.0%). A possible reason was the smaller number of patients with squamous cell carcinoma ($n = 52$). For chemotherapy drugs AP, GP, TP, and DP, the accuracy was 84.6% (77/91), 88.6% (62/70), 67.2% (41/61), and 87.9% (51/58), respectively. Of all eight combinations, the accuracy was highest at 93.8% (45/48) for adenocarcinoma treated by DP. The lowest accuracy was 54.5% for squamous cell carcinoma treated with TP; similarly, there were only 11 instances of this combination, which might have influenced the prediction.

DISCUSSIONS

Clinical Characters

In this study, the progression group had a higher percentage of smokers than the response group, possibly indicating that NSCLC patients who smoke have a higher risk of progression during chemotherapy. Smoking is a high-risk factor for lung cancer (32, 33), and patients with lung cancer who continue to smoke after diagnosis can experience increased treatment-related toxicity and may have a decreased survival rate.

Another finding is that for the response group, a high percentage of those with adenocarcinoma were treated with AP and a high percentage of those with squamous cell carcinoma were treated with GP. This result agrees with the recommendation of AP for adenocarcinoma and GP for squamous cell carcinoma (34–37).

Heterogeneity of Tumors

One of our main findings is that NSCLC tumors in CT images are more heterogeneous in the response group than in the progression group. In the response group, the measures of heterogeneity (SDE, RLNU, and DNU) are overrepresented whereas the measure of homogeneity (DV) is underrepresented (**Figure 8**). This CT-driven textural heterogeneity may correlate with the tumor micro-environment heterogeneity, so the tumor growth rate, invasion ability, drug sensitivity, and prognosis will show differences in CT images (38). Imaging heterogeneity and micro-environment heterogeneity are important for therapeutic response, resistance, and clinical outcomes (39–41). NSCLC patients whose tumors have higher CT-driven textural heterogeneity have a longer overall mean survival (34.5 vs. 22.1 months) (42). Moreover, EGFR-mutated (EGFR+) lung adenocarcinoma is more heterogeneous than EGFR– in CT images (43).

For the cropped CT image cube used as the input in our study, the heterogeneity includes intra- and extra-tumor

components. The information in the extra-tumor region has been considered to be useful for predicting the treatment response and overall survival (8, 27, 28). Extra-tumor heterogeneity emphasizes the contour between the tumor lesion and its habitat. Therefore, here, the higher heterogeneity in the NSCLC response group indicates the higher combination of intra-tumor texture heterogeneity and extra-tumor heterogeneity (the complexity of the tumor contour or shape).

Advantage and Significance of Radiomic Model and Feature Selection

The RF model had an AUC of 0.941, and this test is simple, non-invasive, and quick. A predictive radiomic model could be used in the clinic before treatment to estimate the probability that a patient will respond to chemotherapy and high possibilities would give the oncologists more confidence in the chemotherapy, whereas otherwise other optional treatment plans should be considered.

We tried three different methods of feature selection to know which was suitable for our data and the RF method achieved the better performance than mRMR and relief. The possible reasons are given as follows: a) the mRMR algorithm does not provide a clear determination of the optimal amount of features and can thus still retain redundant features. b) relief is a filter-based feature selection method, but it is easy to ignore small samples and cannot reduce redundant features. We used RF feature selection method based on mean decrease accuracy strategy, it sorts the importance of features to find the most suitable feature subset (13, 44–46).

Clinical Features Do Not Help Prediction

Of the five clinical features, smoking status and number of treatment courses differed significantly between the response and progression groups. Histological type and chemotherapy drug may influence the response (36, 47), but these clinical features do not help to predict the chemotherapy response. Using only clinical features gives a prediction with an AUC of only 0.523 ± 0.09 and combining clinical features with radiomic

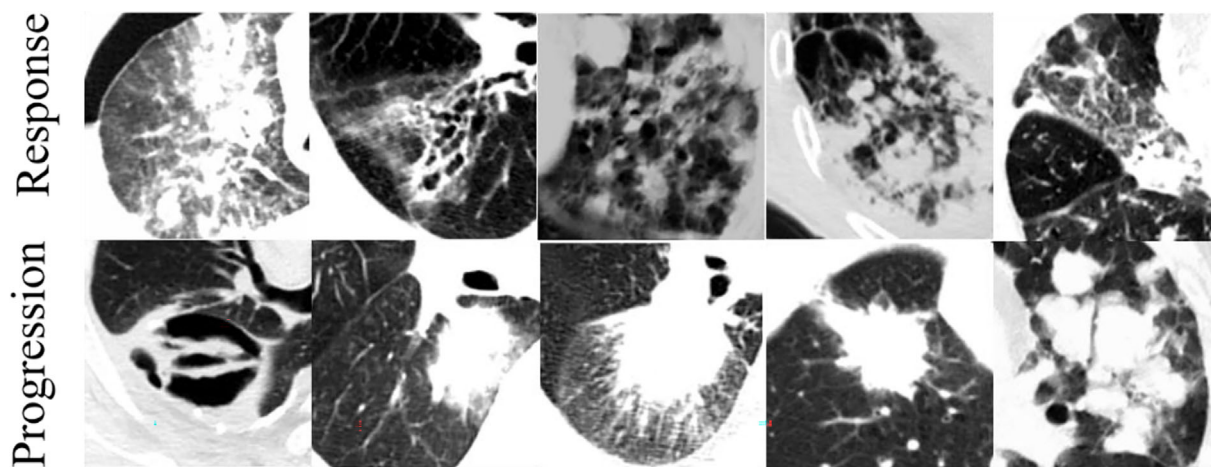


FIGURE 8 | Typical CT images for response and progression groups.

features does not improve the prediction as well. There are two possible reasons. First, the significance comparison is for groups, whereas response prediction involves individuals; features or parameters with significant differences are not necessarily discriminative nor do they always work for individual prediction (12). Second, the relation between clinical and radiomic features is more likely to be correlated than complementary; the radiomic features may represent the information underlying the clinical features and thus, make the latter redundant.

Whether combining clinical features with radiomic features improves the prediction is uncertain and specific to the task. For example, Velazquez et al. found that doing so substantially improved the predictive performance (AUC = 0.86) of EGFR mutation status, whereas using only clinical features gave a predictive model with an AUC of 0.81 (43). Moreover, Lou et al. found that models with both radiomic and clinical features were significantly better at predicting treatment failures than those with only radiomic features (21).

Limitations and Future Work

The present study has limitations. First, our data set comprises CT images and treatment records of only 280 patients. Although overfitting was controlled, the sample size was relatively small. Second, the numbers of patients were unbalanced between adenocarcinoma and squamous cell carcinoma (228 vs. 52). Third, the type and dose of chemotherapy drug were not accounted. Finally, none of the predictive models were constructed using either deep learning or the hybrid method of deep learning and machine learning.

As future work, we will use a predictive model with overall survival as the prognostic endpoint. A deep convolutional neural network will be used to improve the predictive performance and the radiomic nomogram will help facilitate clinical applications (48–50). For a given NSCLC histological type and choice of chemotherapy drug (AP, GP, TP, or DP), a predictive response model may help oncologists choose the correct chemotherapy drug according to the patient's histological type and pre-treatment CT images.

CONCLUSION

The chemotherapy response of NSCLC patients can be predicted by a radiomic model based on machine learning of pre-chemotherapy CT images. Several radiomic features differed

significantly between the response and progression groups and could be used as imaging biomarkers to predict the chemotherapy response. The NSCLC tumors were more heterogeneous in CT images in the response group than in the progression group. This radiomic model with these imaging biomarkers could help to stratify NSCLC patients and make better treatment decisions, simply, non-invasively, and inexpensively.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because they must be approved by the Ethics Committee of Shengjing Hospital of Chinese Medical University. Requests to access the datasets should be directed to SQ, (qisl@bmie.neu.edu.cn).

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the ethics committee of Shengjing Hospital of China Medical University. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

RC performed experiments and analyzed the data. SQ, XZ, and WQ proposed the idea, made discussions, composed the manuscript together with RC. YY collected and analyzed the data. JS directed the algorithm development and analyzed the data. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the National Natural Science Foundation of China (Grant number: 82072008, 81671773, 61672146) and the Fundamental Research Funds for the Central Universities (Grant number: N2124006-3).

REFERENCES

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* (2018) 68(6):394–424. doi: 10.3322/caac.21492
- Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2016. *CA Cancer J Clin* (2016) 66(1):7–30. doi: 10.3322/caac.21332
- 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
- Xu Y, Hosny A, Zeleznik R, Parmar C, Coroller T, Franco I, et al. Deep Learning Predicts Lung Cancer Treatment Response From Serial Medical Imaging. *Clin Cancer Res* (2019) 25(11):3266–75. doi: 10.1158/1078-0432.CCR-18-2495
- Ettinger DS, Akerley W, Borghaei H, Chang AC, Cheney RT, Chirieac LR, et al. Non-Small Cell Lung Cancer, Version 2.2013. *J Natl Compr Canc Netw* (2013) 11(6):645–53. doi: 10.6004/jnccn.2013.0084

6. Wang X, Duan H, Li X, Ye X, Huang G, Nie S. A Prognostic Analysis Method for Non-Small Cell Lung Cancer Based on the Computed Tomography Radiomics. *Phys Med Biol* (2020) 65(4):045006. doi: 10.1088/1361-6560/ab6e51
7. Schwartz LH, Litière S, de Vries E, Ford R, Seymour L. RECIST 1.1-Update and Clarification: From the RECIST Committee. *Eur J Cancer* (2016) 62:132–7. doi: 10.1016/j.ejca.2016.03.081
8. Khorrami M, Khunger M, Zagouras A, Patil P, Thawani R, Bera K, et al. Combination of Peri- and Intratumoral Radiomic Features on Baseline CT Scans Predicts Response to Chemotherapy in Lung Adenocarcinoma. *Radiol Artif Intell* (2019) 1(2):e180012. doi: 10.1148/ryai.2019180012
9. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: The Bridge Between Medical Imaging and Personalized Medicine. *Nat Rev Clin Oncol* (2017) 14(12):749–62. doi: 10.1038/nrclinonc.2017.141
10. Jamal-Hanjani M, Wilson GA, McGranahan N, Birkbak NJ, Watkins T, Veeriah S, et al. Tracking the Evolution of Non-Small-Cell Lung Cancer. *N Engl J Med* (2017) 376(22):2109–21. doi: 10.1056/NEJMoa1616288
11. 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
12. 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
13. Parmar C, Grossmann P, Bussink J, Lambin P, Aerts HJWL. Machine Learning Methods for Quantitative Radiomic Biomarkers. *Sci Rep* (2015) 5:13087. doi: 10.1038/srep13087
14. Avanzo M, Stancanello J, El Naqa I. Beyond Imaging: The Promise of Radiomics. *Phys Med* (2017) 38:122–39. doi: 10.1016/j.ejmp.2017.05.071
15. Lee G, Lee HY, Park H, Schiebler ML, van Beek E, Ohno Y, et al. Radiomics and Its Emerging Role in Lung Cancer Research, Imaging Biomarkers and Clinical Management: State of the Art. *Eur J Radiol* (2017) 86:297–307. doi: 10.1016/j.ejrad.2016.09.005
16. Grossmann P, Stringfield O, El-Hachem N, Bui MM, Rios Velazquez E, Parmar C, et al. Defining the Biological Basis of Radiomic Phenotypes in Lung Cancer. *Elife* (2017) 6:e23421. doi: 10.7554/eLife.23421
17. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial Intelligence in Radiology. *Nat Rev Cancer* (2018) 18(8):500–10. doi: 10.1038/s41568-018-0016-5
18. Hosny A, Parmar C, Coroller TP, Grossmann P, Zeleznik R, Kumar A, et al. Deep Learning for Lung Cancer Prognostication: A Retrospective Multi-Cohort Radiomics Study. *PLoS Med* (2018) 15(11):e1002711. doi: 10.1371/journal.pmed.1002711
19. Song J, Shi J, Dong D, Fang M, Zhong W, Wang K, et al. A New Approach to Predict Progression-Free Survival in Stage IV EGFR-Mutant NSCLC Patients With EGFR-TKI Therapy. *Clin Cancer Res* (2018) 24(15):3583–92. doi: 10.1158/1078-0432.CCR-17-2507
20. Paul R, Hawkins SH, Balagurunathan Y, Schabath MB, Gillies RJ, Hall LO, et al. Deep Feature Transfer Learning in Combination With Traditional Features Predicts Survival Among Patients With Lung Adenocarcinoma. *Tomography* (2016) 2(4):388–95. doi: 10.18383/j.tom.2016.00211
21. Lou B, Doken S, Zhuang T, Wingarter D, Gidwani M, Mistry N, et al. An Image-Based Deep Learning Framework for Individualizing Radiotherapy Dose. *Lancet Digit Health* (2019) 1(3):e136–47. doi: 10.1016/S2589-7500(19)30058-5
22. Chetan MR, Gleeson FV. Radiomics in Predicting Treatment Response in Non-Small-Cell Lung Cancer: Current Status, Challenges and Future Perspectives. *Eur. Radiol* (2021) 31(2):1049–58. doi: 10.1007/s00330-020-07141-9
23. Chen A, Saouaf J, Zhou B, Crawford R, Goldmacher G, Yuan J, et al. A Deep Learning-Facilitated Radiomics Solution for the Prediction of Lung Lesion Shrinkage in Non-Small Cell Lung Cancer Trials. In: *IEEE 17th International Symposium on Biomedical Imaging (ISBI)* (2020). p. 678–82. doi: 10.1109/ISBI45749.2020.9098561
24. Seki S, Fujisawa Y, Yui M, Kishida Y, Koyama H, Ohyu S, et al. Dynamic Contrast-Enhanced Area-Detector CT vs Dynamic Contrast-Enhanced Perfusion MRI vs FDG-PET/CT: Comparison of Utility for Quantitative Therapeutic Outcome Prediction for NSCLC Patients Undergoing Chemoradiotherapy. *Magn Reson Med Sci* (2020) 19:29–39. doi: 10.2463/mrms.mp.2018-0158
25. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin JC, Pujol S, et al. 3D Slicer as an Image Computing Platform for the Quantitative Imaging Network. *Magn Reson Imaging* (2012) 30(9):1323–41. doi: 10.1016/j.mri.2012.05.001
26. van Griethuysen JMM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res* (2017) 77(21):e104–7. doi: 10.1158/0008-5472.CAN-17-0339
27. Algohary A, Shiradkar R, Pahwa S, Purysko A, Verma S, Moses D, et al. Combination of Peri-Tumoral and Intra-Tumoral Radiomic Features on Bi-Parametric MRI Accurately Stratifies Prostate Cancer Risk: A Multi-Site Study. *Cancers (Basel)* (2020) 12(8):2200. doi: 10.3390/cancers12082200
28. Braman NM, Etesami M, Prasanna P, Dubchuk C, Gilmore H, Tiwari 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(1):57. doi: 10.1186/s13058-017-0846-1
29. Rigatti SJ. Random Forest. *J Insur Med* (2017) 47(1):31–9. doi: 10.17849/insm-47-01-31-39.1
30. Peng H, Long F, Ding C. Feature Selection Based on Mutual Information: Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy. *IEEE Trans Pattern Anal Mach Intell* (2005) 27(8):1226–38. doi: 10.1109/TPAMI.2005.159
31. Zhao Z, Morstatter F, Sharma S, Alelyani S, Anand A, Liu H, et al. Advancing Feature Selection Research. In: *ASU Feature Selection Repository*. Arizona State University (2010) 1–28. Available at: <http://www.public.asu.edu/~huanliu/papers/tr-10-007.pdf>.
32. Aberle DR, Berg CD, Black WC, Church TR, Fagerstrom RM, Galen B, et al. The National Lung Screening Trial: Overview and Study Design. *Radiology* (2011) 258(1):243–53. doi: 10.1148/radiol.10091808
33. Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, et al. Reduced Lung-Cancer Mortality With Low-Dose Computed Tomographic Screening. *N Engl J Med* (2011) 365(5):395–409. doi: 10.1056/NEJMoa1102873
34. Senan S, Brade A, Wang LH, Vansteenkiste J, Dakhil S, Biesma B, et al. PROCLAIM: Randomized Phase III Trial of Pemetrexed-Cisplatin or Etoposide-Cisplatin Plus Thoracic Radiation Therapy Followed by Consolidation Chemotherapy in Locally Advanced Nonsquamous Non-Small-Cell Lung Cancer. *J Clin Oncol* (2016) 34(9):953–62. doi: 10.1200/JCO.2015.64.8824
35. Choy H, Gerber DE, Bradley JD, Iyengar P, Monberg M, Treat J, et al. Concurrent Pemetrexed and Radiation Therapy in the Treatment of Patients With Inoperable Stage III Non-Small Cell Lung Cancer: A Systematic Review of Completed and Ongoing Studies. *Lung Cancer* (2015) 87(3):232–40. doi: 10.1016/j.lungcan.2014.12.003
36. Schiller JH, Harrington D, Belani CP, Langer C, Sandler A, Krook J, et al. Comparison of Four Chemotherapy Regimens for Advanced Non-Small-Cell Lung Cancer. *N Engl J Med* (2002) 346(2):92–8. doi: 10.1056/NEJMoa011954
37. Scagliotti GV, Parikh P, von Pawel J, Biesma B, Vansteenkiste J, Manegold C, et al. Phase III Study Comparing Cisplatin Plus Gemcitabine With Cisplatin Plus Pemetrexed in Chemotherapy-Naive Patients With Advanced-Stage Non-Small-Cell Lung Cancer. *J Clin Oncol* (2008) 26(21):3543–51. doi: 10.1200/JCO.2007.15.0375
38. Junttila MR, de Sauvage FJ. Influence of Tumour Micro-Environment Heterogeneity on Therapeutic Response. *Nature* (2013) 501(7467):346–54. doi: 10.1038/nature12626
39. O'Connor JP, Rose CJ, Waterton JC, Carano RA, Parker GJ, Jackson A. Imaging Intratumor Heterogeneity: Role in Therapy Response, Resistance, and Clinical Outcome. *Clin Cancer Res* (2015) 21(2):249–57. doi: 10.1158/1078-0432.CCR-14-0990
40. O'Connor JPB. Cancer Heterogeneity and Imaging. *Semin Cell Dev Biol* (2017) 64:48–57. doi: 10.1016/j.semcdb.2016.10.001
41. Bashir U, Siddique MM, Mclean E, Goh V, Cook GJ. Imaging Heterogeneity in Lung Cancer: Techniques, Applications, and Challenges. *AJR Am J Roentgenol.* (2016) 207(3):534–43. doi: 10.2214/AJR.15.15864
42. Win T, Miles KA, Janes SM, Ganeshan B, Shastry M, Endozo R, et al. Tumor Heterogeneity and Permeability as Measured on the CT Component of PET/CT Predict Survival in Patients With Non-Small Cell Lung Cancer. *Clin Cancer Res* (2013) 19(13):3591–9. doi: 10.1158/1078-0432.CCR-12-1307

43. Rios Velazquez E, Parmar C, Liu Y, Coroller TP, Cruz G, Stringfield O, et al. Somatic Mutations Drive Distinct Imaging Phenotypes in Lung Cancer. *Cancer Res* (2017) 77(14):3922–30. doi: 10.1158/0008-5472.CAN-17-0122
44. Zhou Q, Zhou H, Li T. Cost-Sensitive Feature Selection Using Random Forest: Selecting Low-Cost Subsets of Informative Features. *Knowledge-Based Systems* (2016) 95:1–11. doi: 10.1016/j.knsys.2015.11.010
45. Deviaene M, Testelmans D, Borzee P, Buyse B, Huffel SV, Varon C. (2019). Feature Selection Algorithm Based on Random Forest Applied to Sleep Apnea Detection. In: *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE (2019) p. 2580–3. doi: 10.1109/EMBC.2019.8856582
46. Zhou Q, Hao Z, Zhou Q, Fan Y, Luo L. Structure Damage Detection Based on Random Forest Recursive Feature Elimination. *Mech Syst Signal Proc* (2014) 46(1):82–90. doi: 10.1016/j.ymssp.2013.12.013
47. Paz-Ares LG, de Marinis F, Dediu M, Thomas M, Pujol JL, Bidoli P, et al. PARAMOUNT: Final Overall Survival Results of the Phase III Study of Maintenance Pemetrexed Versus Placebo Immediately After Induction Treatment With Pemetrexed Plus Cisplatin for Advanced Nonsquamous Non-Small-Cell Lung Cancer. *J Clin Oncol* (2013) 31(23):2895–902. doi: 10.1200/JCO.2012.47.1102
48. Zhao X, Liu L, Qi S, Teng Y, Li J, Qian W. Agile Convolutional Neural Network for Pulmonary Nodule Classification Using CT Images. *Int. J Comput Assist Radiol Surg* (2018) 13(4):585–95. doi: 10.1007/s11548-017-1696-0
49. Dong D, Fang MJ, Tang L, Shan XH, Gao JB, Giganti F, et al. Deep Learning Radiomic Nomogram Can Predict the Number of Lymph Node Metastasis in Locally Advanced Gastric Cancer: An International Multicenter Study. *Ann. Oncol* (2020) 31(7):912–20. doi: 10.1016/j.annonc.2020.04.003
50. Dong D, Tang L, Li ZY, Fang MJ, Gao JB, Shan XH, et al. Development and Validation of an Individualized Nomogram to Identify Occult Peritoneal Metastasis in Patients With Advanced Gastric Cancer. *Ann. Oncol* (2019) 30(3):431–8. doi: 10.1093/annonc/mdz001

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.

Copyright © 2021 Chang, Qi, Yue, Zhang, Song and Qian. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.