

## Research Article

# Machine Learning-Based Multiomics Prediction Model for Radiation Pneumonitis

Lu Zhou , Yuefeng Wen , Guoqian Zhang , Linjing Wang , Shuyu Wu ,  
and Shuxu Zhang 

Department of Radiation Oncology, Affiliated Cancer Hospital and Institute of Guangzhou Medical University, Guangzhou, China

Correspondence should be addressed to Shuyu Wu; [wsyeasy@outlook.com](mailto:wsyeasy@outlook.com) and Shuxu Zhang; [gthzxs@163.com](mailto:gthzxs@163.com)

Received 18 November 2022; Revised 18 January 2023; Accepted 1 February 2023; Published 18 February 2023

Academic Editor: Thomas R. Chauncey

Copyright © 2023 Lu Zhou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Objective.** The study aims to establish and validate an effective CT-based radiation pneumonitis (RP) prediction model using the multiomics method of radiomics and EQD2-based dosiomics. **Materials and Methods.** The study performed a retrospective analysis on 91 nonsmall cell lung cancer patients who received radiotherapy from 2019 to 2021 in our hospital. The patients with RP grade  $\geq 1$  were labeled as 1, and those with RP grade  $< 1$  were labeled as 0. The whole lung excluding clinical target volume (lung-CTV) was used as the region of interest (ROI). The radiomic and dosiomic features were extracted from the lung-CTV area's image and dose distribution. Besides, the equivalent dose of the 2 Gy fractionated radiation (EQD<sub>2</sub>) model was used to convert the physical dose to the isoeffect dose, and then, the EQD2-based dosiomic (eqd-dosiomic) features were extracted from the isoeffect dose distribution. Four machine learning (ML) models, including DVH, radiomics combined with DVH (radio + DVH), radiomics combined with dosiomics (radio + dose), and radiomics combined with eqd-dosiomics (radio + eqdose), were established to construct the prediction model via eleven different classifiers. The fivefold cross-validation was used to complete the classification experiment. The area under the curve (AUC) of the receiver operating characteristics (ROC), accuracy, precision, recall, and F1-score were calculated to assess the performance level of the prediction models. **Results.** Compared with the DVH, radio + DVH, and radio + dose model, the value of the training AUC, accuracy, and F1-score of radio + eqdose was higher, and the difference was statistically significant ( $p < 0.05$ ). Besides, the average value of the precision and recall of radio + eqdose was higher, but the difference was not statistically significant ( $p > 0.05$ ). **Conclusion.** The performance of using the ML-based multiomics method of radiomics and eqd-dosiomics to predict RP is more efficient and effective.

## 1. Introduction

Radiotherapy is one of the most important treatment methods for lung cancer. However, radiation-induced pulmonary injury is the main limiting factor and the most common complication of thoracic tumor radiotherapy. An accurate prediction model is desired to clarify the risk factors of RP, guide the design of radiotherapy treatment plans, and prevent high-risk patients in advance.

In the clinical, the incidence of RP is correlated with the dose to the lung tissue. The lung volume within which the dose is greater than xGy (V<sub>x</sub>) from DVH, such as V<sub>30</sub>, V<sub>20</sub>, and V<sub>5</sub>, is widely used for RP prediction. Afterward, some studies have shown a strong correlation between these image

features (such as CT image density or 18F-fluorodeoxyglucose uptake in positron emission tomography) and tissue heterogeneity at the cellular level [1–3]. With the development of radiomics, the digital image processing and ML techniques can be applied in medical image analysis [4–10]. Cunliffe demonstrated the ability of radiomics to provide a quantitative, individualized measurement of patient lung tissue reaction to RT and assess RP development [11]. Based on these studies, the dosiomics method has been proposed, which attempts to extract the spatial features from dose distribution to construct a prediction model [12–16]. Liang et al. applied the framework of radiomics on dose distribution and demonstrated that the dosiomic features improve the prediction ability efficiently [17–19].

TABLE 1: Patient clinical and treatment characteristics.

Clinical and treatment characteristics	Median (range)/n (%)
Age	62.5 (31–85)
Sex	
Male	83 (91%)
Female	8 (9%)
Smoking history	
No	37 (40%)
Smoking < 30 years	8 (9%)
Smoking ≥ 30 years	46 (51%)
Stage	
1	0
2	3 (3%)
3	63 (69%)
4	25 (28%)
Prescription dose	60 (40–66) Gy
Prescription dose per fraction	2 (1.5–2.75) Gy
RP grade	
0	52 (57%)
1	24 (26%)
2	9 (10%)
3	6 (7%)
4	0

Most of the previous studies used the radiomic features of the whole lung, combined with the dose-volume factors and some clinical treatment characteristics, to train and obtain the prediction model of RP. Some studies showed that RP occurs outside the tumor, and lungs that receive low doses (5 or 20 Gy) are associated with RP [20, 21]. Thus, the region of interest was the lung-CTV in this study. The single dose characteristics only utilize partial information on the dose distribution without information regarding the spatial relationship of voxels. The same  $V_x$  value may be obtained from voxels that are spread throughout the OAR or connected voxels, so the  $V_x$ 's biological impact may differ [22–29]. The same absorbed dose with varying fractions of treatment, different time intervals, and different numbers of irradiation per day will represent different biological doses and result in other biological effects. The higher dose of each fraction, the greater the biological effect on tissue, especially for the late reaction tissues such as the lung [30, 31]. In the clinical, the technology of hyper-fractionated or large-fractionated may be adopted for some patients or changed to hyper-fractionated after a certain dose of conventional radiotherapy [30]. Besides, the radiation dose per fraction (Fx) has been shown to be a significant factor in RP [29]. Previous studies only used the dosiomic features from the physical dose distribution. In this study, the equivalent dose of 2 Gy fractionated radiation (EQD<sub>2</sub>) was used to solve the problem of equivalent dose calculation for different fractionated radiotherapy. The physical dose distribution of lung tissue was converted to the EQD<sub>2</sub> distribution, and the eqd-dosiomic features were extracted from the EQD<sub>2</sub> distribution using the framework of radiomics. Finally, the predictive performance of four ML models, including DVH, radio + DVH, radio + dose, and radio + eqdose, was compared.

## 2. Methods and Materials

**2.1. Patient Database.** In this study, 350 patients diagnosed with nonsmall cell lung cancer who received radiotherapy in our hospital from 2019 to 2021 were collected retrospectively, in which 91 patients met the selection criteria. At first, all patients were pathologically diagnosed with nonsmall cell lung cancer and had definite TNM stages of lung cancer (Stage I, II, III, IV). Second, each patient underwent two high-resolution diagnostic CT scans without surgical intervention. The first was before radiotherapy, and the second was within six months after radiotherapy. All patients were treated entirely with intensity-modulated radiation therapy (IMRT), and the plans were designed using the pinnacle treatment planning system (V9.10). The slice spacing of the planning CT image was 5 mm, and the grid spacing of dose calculation was 3 × 3 × 3 mm. The classification of RP was carried out according to the RTOG classification standard for acute radioactive lung injury. The endpoint of this study was grade ≥ 1 RP. The clinical and treatment characteristics of patients are shown in Table 1, and the details of 91 patients are available in the supplementary file 1.

**2.2. Data Acquisition and Dose Conversion.** In the study, the self-developed automatic radiotherapy plan analysis software was used to obtain the 3D volume (in Figures 1(a) and 1(b)) and the original 3D dose distribution (in Figure 1(c)) of the lung-CTV area from the Pinnacle treatment planning system (V9.10).

In the clinical, it is necessary to maintain the total dose required for equivalent biological effects when changing the conventional treatment plan. The EQD<sub>2</sub> is the dose that causes an equivalent biological effect to conventional 2 Gy fractionation. The problem of equivalent dose calculation for different fractionated radiotherapy was solved using the EQD<sub>2</sub> model [30–34]. The linear quadratic (L-Q) model is derived directly from the cell survival curve. The concept of the isoeffect dose and the mathematical derivation formula based on the L-Q model can be used to standardize and compare biological doses of radiotherapy with different fractions [35]. The formulation of EQD<sub>2</sub> was as follows:

$$\text{EQD}_2 = D \frac{d + (\alpha/\beta)}{2 + (\alpha/\beta)}, \quad (1)$$

where  $D$  represents the total dose, and  $d$  is the fractional dose for spatial points in the lung-CTV area. The  $(\alpha/\beta)$  reflects the radio-biological characteristics of tissue. The value of  $(\alpha/\beta)$  is larger for early-response tissue and smaller for late-response tissue. As the lung is late response tissue, the value of  $(\alpha/\beta)$  is set to 4.5 [31]. Then, formulation (1) was used to convert the physical dose into the EQD<sub>2</sub>. The EQD<sub>2</sub> distribution was shown in Figure 1(d).

**2.3. DVH Parameters.** In the clinical, the lung volume within which the dose is greater than xGy ( $V_x$ ) is widely used for RP prediction. In this study, the  $V_{30}$ ,  $V_{20}$ , and  $V_5$  were extracted as DVH features for all patients.

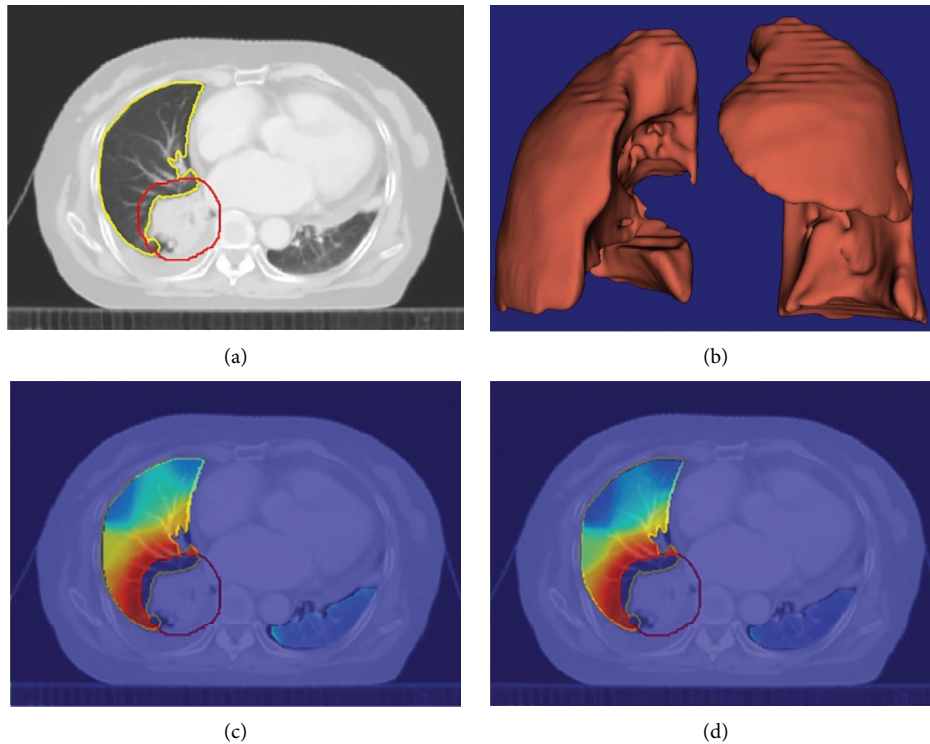


FIGURE 1: The results of data acquisition and dose conversion: (a) the delineated CTV and lung in CT image, the red is CTV, and the yellow is lung; (b) the 3D volume of lung-CTV; (c) the overlapping of physical dose distribution and CT image; (d) the overlapping of EQD<sub>2</sub> distribution and CT image.

**2.4. Feature Preprocessing.** At first, the radiomic, dosiomic, and eqd-dosiomic features were extracted from the CT image, the physical dose, and the EQD<sub>2</sub> distribution of the lung-CTV area. All of the features were extracted by Pyradiomics 3.0.1 [36]. The wavelet filter was used for radiomics calculations. Eight feature groups (a total of 863 radiomics features), including first-order statistics (18 features), shape-based (14 features), gray level co-occurrence matrix (GLCM, 24 features), gray level run length matrix (GLRLM, 16 features), gray level size zone matrix (GLSZM, 16 features), gray level dependence matrix (GLDM, 14 features), neighboring gray-tone difference matrix (NGTDM, five features), and wavelet features (744 features) were extracted, respectively.

Secondly, the data were randomly divided into two sets, with 70% for the training set and 30% for the testing set. The synthetic minority oversampling technique (SMOTE) was used to prevent overfitting by the unbalanced ratio in the training set.

Thirdly, the random forest algorithm was used to select the features. The random forest model performs well because of the sample randomness, feature randomness, and integration strategy. After data preprocessing, the random forest model was performed to calculate the importance of independent variables and filter out the redundant information [37].

At last, the fivefold cross-validation was used to reduce the overfitting in nonlinear regression in the study. The training data were randomly divided into five folder sets, iteratively performed in each model, and the metric

of interest was calculated on each validation set. Then, the five values of the metric were averaged to get the training AUC. The testing set was used to independently evaluate the validity of the prediction model by the testing AUC, accuracy, precision, recall, and F1-score.

**2.5. Prediction Model.** This study established four ML models, including DVH, radio + DVH, radio + dose, and radio + eqdose. In addition, eleven classifiers, including logistic regression, ridge, SVM, perceptron, decision tree, random forest, KNeighbors, passive aggressive, GaussianNB, multinomialNB, and Adaboost, were used to establish the multiomics prediction model of RP. The hyperparameters for the eleven classifiers are available in supplementary file 2. Python 3.7.6 and scikit-learn 0.24.2 were performed for feature extraction, data preprocessing, and ML modeling.

### 3. Result

**3.1. Feature Importance.** Six features were selected from the radiomic, dosiomic, and eqd-dosiomic features, respectively. As shown in Figure 2, most of the selected features were lung texture features, which represent periodic changes and structural rules in the spatial domain of images and can fully reflect the heterogeneity of tissues [38, 39]. The gray level size zone matrix (GLSZM) features quantize the region of continuous pixels in the image, including the features describing the distribution of small or large areas and low or

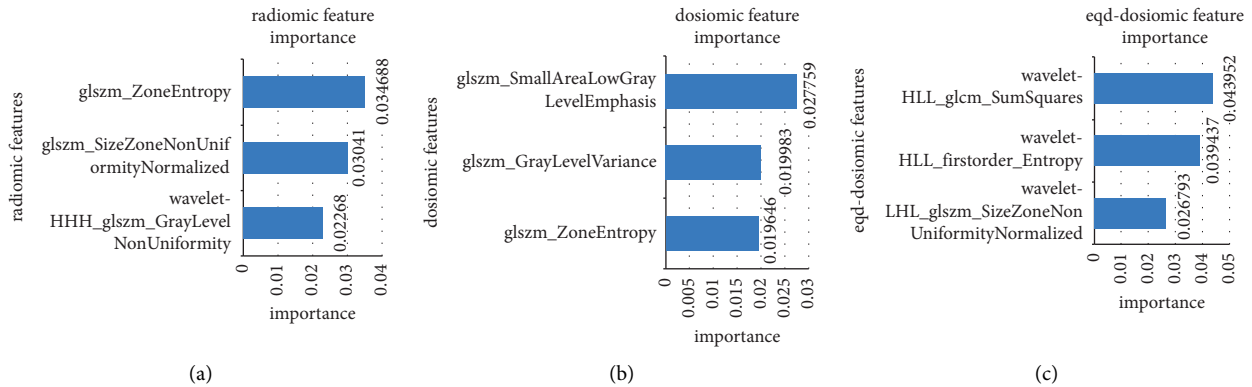


FIGURE 2: The three most important features: (a) the radiomic features, (b) the dosiomic features, and (c) the eqd-dosiomic features.

high gray regions. It significantly affects the representation of texture consistency, nonperiodic, or speckled texture and has better performance in image texture analysis [40]. The gray level co-occurrence matrix (GLCM) describes the joint distribution of the grayscale of two pixels with a specific spatial position relationship, which reflects the comprehensive information of the direction, adjacent interval, and variation amplitude of image gray [36].

Figure 3 shows the GLSZM derived from the 3D physical dose distribution and EQD<sub>2</sub> distribution for an example patient. The information on GLSZM for EQD<sub>2</sub> is more abundant than the physical dose.

**3.2. Model Performance.** The statistical analysis of the performance of four ML models under 11 classifiers is shown in Figure 4 and Table 2. The evaluation indicators of each classifier can be found in supplementary file 3. The statistical analysis was done to prove that the proposed method's performance in this study was not affected by different classifiers. The training AUC is the mean value of five folds for the training set. The testing AUC, accuracy, precision, recall, and F1-score are values for the testing set. The Student's *t*-test was used to calculate *p* values.

Compared with the DVH, radio+DVH, and radio+dose models, the training AUC value of radio+eqdose was higher, and the difference was statistically significant ( $p < 0.05$ ). Besides, for four models, the testing AUC average values were almost 0.8, and the difference was less than 1%. The results showed that the four ML models could predict RP well, among which the radio+eqdose performed better. Compared with the other three models, the accuracy of radio+eqdose was higher, and the difference was statistically significant ( $p < 0.05$ ), representing that the radio+eqdose model could more accurately predict whether patients will suffer from RP. The precision rate reflects the model's ability to distinguish negative samples, the recall rate demonstrates the model's ability to identify positive examples, and the F1-score considers both the classification model's precision and

recall. Compared with the other three models, the F1-score of radio+eqdose was higher, and the difference was statistically significant ( $p < 0.05$ ). Therefore, the multiomics method of radiomics and eqd-dosiomics can improve the predictive performance of RP.

#### 4. Discussion

In the clinical, the doctors will determine the dose and time segmentation mode of radiotherapy according to the size and lethal dose of the tumor, the tolerated dose of normal tissue, and the normal tissue distribution around the target area. The technology of hyper-fractionated or large-fractionated may be adopted for some patients or changed to hyper-fractionated after a specific dose of conventional radiotherapy [30]. The radiation dose per fraction (Fx) has been shown to be a significant factor in RP. However, previous studies did not consider the effect of different fractional doses on dosiomics. The EQD<sub>2</sub> model was used to convert the physical dose to the isoeffect dose. The radiomic, dosiomic, and eqd-dosiomic features were extracted, respectively, from the image, the dose distribution, and the isoeffect dose distribution of the lung-CTV area.

In this study, most of the selected features are texture features extracted from the original image characteristics through some calculation and stored in an intermediate matrix. Then, a series of statistics are defined on this intermediate matrix as the texture features. In the GLCM, matrix elements *ij* represent the times a voxel with gray level *i* is a neighbor of voxels with gray level *j* [36]. The GLCM\_Sum Squares is a measure of the distribution of neighboring dose pairs about the mean dose, which reflects the information regarding the spatial relationship of voxels. The GLSZM features describe the regional consistency of the dose distribution textures, in which pixel *ij* is the number of times a 3D zone with dose *i* has a size *j* [40, 41]. The feature GLSZM\_Size Zone nonuniformity normalized measures the

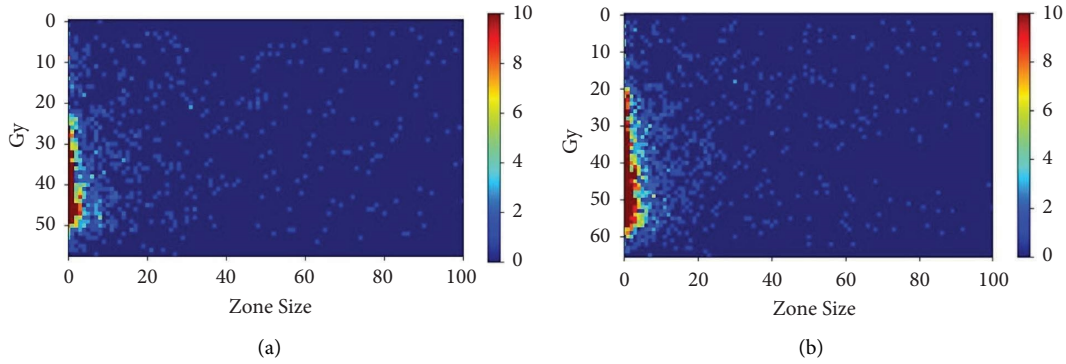


FIGURE 3: The gray level size zone matrix (GLSZM) of the 3D physical dose distribution (a) and the EQD<sub>2</sub> distribution (b).

TABLE 2: The evaluation indicators of four ML models under 11 classifiers (mean  $\pm$  standard deviation).

Models	Training AUC	Testing AUC	Accuracy	Precision	Recall	F1
DVH	0.804 $\pm$ 0.053	0.801 $\pm$ 0.06	0.751 $\pm$ 0.068	0.738 $\pm$ 0.067	0.783 $\pm$ 0.16	0.751 $\pm$ 0.095
Radio + DVH	0.856 $\pm$ 0.044	0.805 $\pm$ 0.046	0.809 $\pm$ 0.062	0.772 $\pm$ 0.066	0.886 $\pm$ 0.118	0.821 $\pm$ 0.067
Radio + dose	0.833 $\pm$ 0.041	0.801 $\pm$ 0.057	0.806 $\pm$ 0.048	0.756 $\pm$ 0.045	0.911 $\pm$ 0.128	0.821 $\pm$ 0.061
Radio + eqdose	0.900 $\pm$ 0.033	0.799 $\pm$ 0.050	0.856 $\pm$ 0.034	0.809 $\pm$ 0.070	0.950 $\pm$ 0.059	0.870 $\pm$ 0.024

variability of size zone volumes throughout the EQD<sub>2</sub> distribution, with a lower value indicating more homogeneity among zone size volumes. The first-order entropy specifies the uncertainty in the dose distribution. The GLSZM\_Zone entropy measures the uncertainty and randomness in the zone sizes and dose levels distribution. Lower zone entropy might indicate a more uniform dose in the region of interest. The GLSZM\_Gray level variance calculates the variance in dose values for the zones [36]. Therefore, we speculated that the occurrence of RP is related to the local dose variation in lung tissue. Besides, the GLSZM\_Small area low gray level emphasis measures the proportion of the joint distribution of smaller regions with lower dose values in the physical dose distribution. This feature indicates that the low-dose region of lung tissue is associated with the development of RP.

Bin Liang also showed that the higher the local dose variation within the ipsilateral lung and the greater the low-dose region of total lungs, the greater probability of RP incidence [17–19]. The DVH parameters, such as V5, V20, and V30, are just accumulated doses of a specific volume of lung tissue that cannot reflect the local dose variation. However, the texture features show information regarding the spatial relationship of dose distribution for radiotherapy response prediction. As shown in Figure 3, the information of GLSZM for EQD<sub>2</sub> distribution is more abundant, which is a benefit for revealing the hidden correlation with RP incidence. Besides, the radio + eqdose model uses the first order, GLCM, and GLSZM features. The accumulated doses, the local dose variation, and the comprehensive information of direction, adjacent interval, and variation amplitude of dose distribution are used to construct the prediction model, so the evaluation metrics of radio + eqdose are better than the other three models in the study.

Previous studies demonstrated that the radio + dose model performed better than the DVH parameters for RP

prediction [11–13]. In our study, the radio + dose model still performed well, and the radio + eqdose model was further investigated.

Fraction dose has been shown to be an essential factor in RP [29]. The impact of fraction dose can be illustrated using the well-known L-Q model. In this study, the fraction dose of the patients was from 1.5 to 2.75 Gy. It is defined as large-fractionated radiotherapy when the fractional dose exceeds 2.5 Gy. Compared with conventional fractionation radiation therapy, large-fractionated radiotherapy can increase the biological dose to improve tumor control and survival [42, 43]. The fraction dose of hyper-fractionated radiotherapy is generally lower than 2 Gy, more than once a day, and the interval time of irradiation is greater than 4 to 6 hours. The hyper-fractionated radiotherapy improves the local control rate of the tumor by increasing the total irradiation dose without significantly increasing the late reaction of normal tissue [44, 45]. When the total dose was 50 Gy, the EQD<sub>2</sub> and BED were 54 Gy and 78 Gy for the 2.5 Gy fraction dose, 50 Gy and 72 Gy for the 2 Gy fraction dose, and 46 Gy and 66 Gy for the 1.5 Gy fraction dose. The differences in radio biology were significant.

The EQD<sub>2</sub> model was used to solve the problem of equivalent dose calculation for different fractionated radiotherapy. Compared with the DVH, radio + DVH, and radio + dose model, the value of the training AUC, accuracy, and F1-score of radio + eqdose was higher, and the difference was statistically significant ( $p < 0.05$ ). Besides, the average value of the recall and precision of radio + eqdose under 11 classifiers was higher than the other three models. The F1-score is the harmonic mean of precision and recall, reflecting the model's robustness. The results demonstrated that the radio + eqdose model could effectively improve the prediction ability of RP. The difference in testing AUC value was less than 1%, and the radio + eqdose model showed no

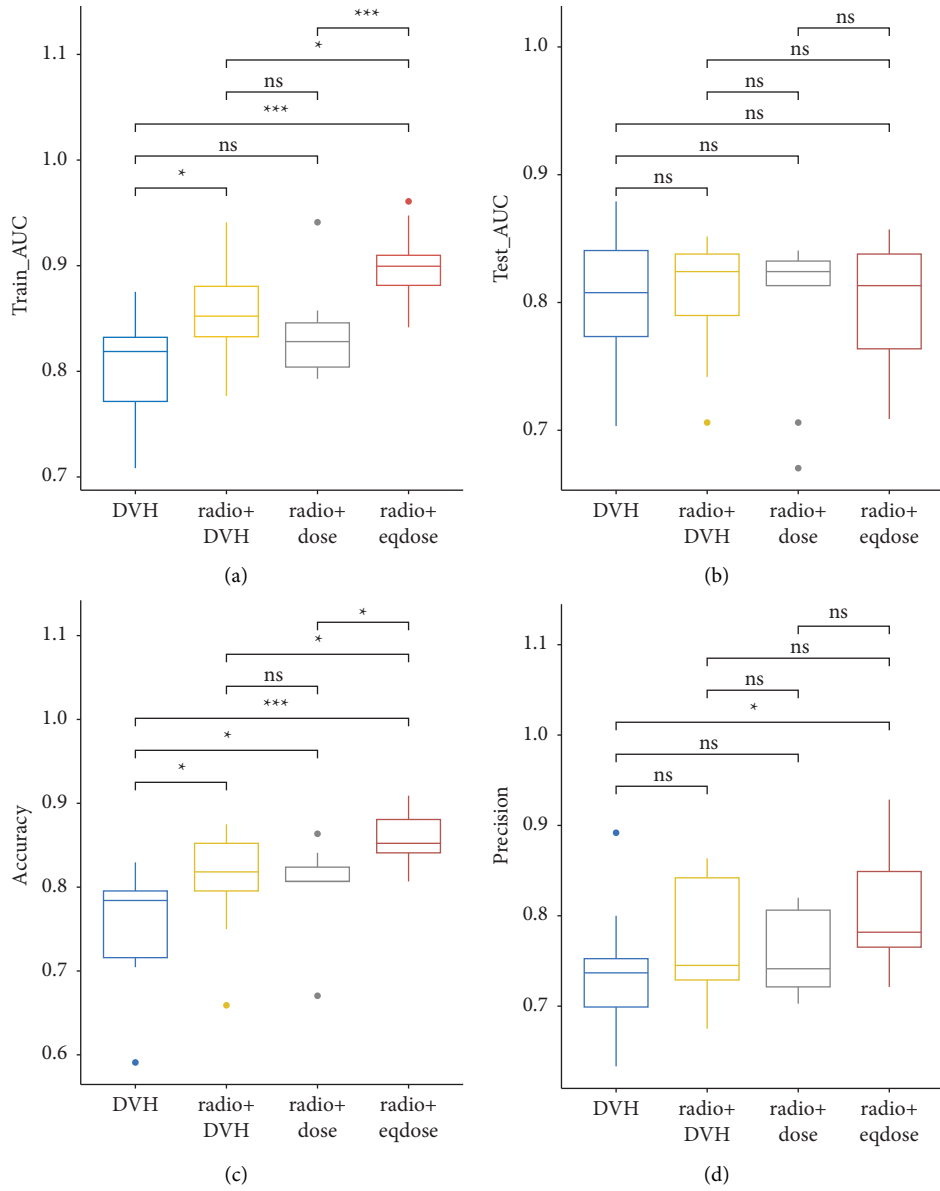


FIGURE 4: Continued.

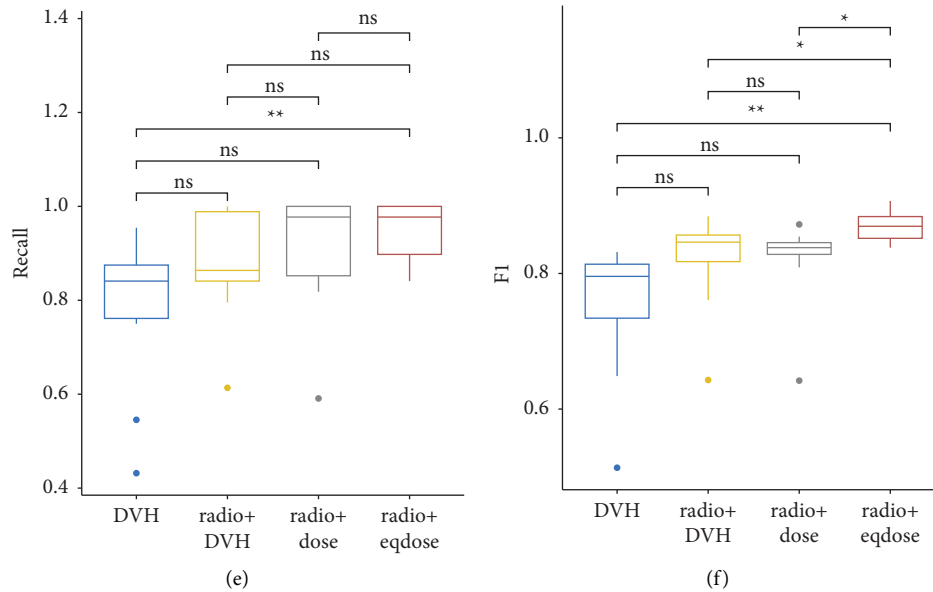


FIGURE 4: The performances of four ML models under 11 classifiers: (a) is the training AUC, (b) is the testing AUC, (c) is the accuracy, (d) is the precision, (e) is the recall, (f) is the F1-score. In the figure, \*indicates  $p < 0.05$ , \*\*indicates  $p < 0.01$ , \*\*\*indicates  $p < 0.005$ , and ns indicates  $p > 0.05$ .

obvious advantage. The possible reason is that the testing data are relatively small, so future studies will include more data to make the performance evaluation more objective. The single dose characteristics only utilize partial information on the dose distribution. Still, the multiomics method uses the info of CT image and dose distribution regarding the spatial relationship of voxels to train the prediction model. The advantages of the multi-omics method are apparent. Figure 3 showed that the information of GLSZM for EQD<sub>2</sub> was more abundant than the physical dose, and we could extract more information from the EQD<sub>2</sub> distribution for predicting RP.

The dose (absolute and relative dose) of the organ at risk can be obtained from DVH. However, the expression here is only the physical dose. Since the absorbed dose of the organ at risk is generally not 100%, the single dose is different from the conventional 2 Gy dose so that the EQD<sub>2</sub> formula can calculate the isoeffect dose of the organ at risk. Because the tolerance of each organ at risk is obtained under conventional irradiation, it should be converted to EQD<sub>2</sub> [32, 33]. Therefore, integrating the eqd-dosiomic features into training the model is reasonable. The experimental results also confirm the effectiveness of this method.

There are two main limitations of this study that are worth discussing. One is that the number of patients meeting the research requirements is relatively less. In order to avoid some errors due to equipment model, scanning conditions, and treatment planning, this study has only used the patients in our hospital. Besides, there are relatively few cases of RP  $\geq 2$  after radiotherapy in our hospital (15 patients). Therefore, the endpoint in this study was RP grade  $\geq 1$ , which might limit its usefulness in actual clinical practice. In the next step, enrolling more patients in our hospital and multiple centers is necessary to validate our proposed approach. The other is that only radiomic and dosiomic

features were considered in this study. However, RP is highly correlated with some clinical features, such as smoking history, disease stage, and tumor location. In future work, the clinical features will be added for further discussion. Finally, to prevent RP incidence, how to use the radiomic, dosiomic, and clinical characteristics to guide the design of radiotherapy treatment planning is the goal of our study.

## 5. Conclusion

In this study, we demonstrated that the multiomics model could improve the predictive performance of RP grade  $\geq 1$  compared to the DVH model. We also confirmed that the eqd-dosiomic features from the isoeffect dose distribution could improve the predictive model's performance with different fractionated radiotherapies. The multiomics method of radiomics and eqd-dosiomics could improve the predictive performance of RP. It is expected that further studies can be used to guide the design of radiotherapy treatment planning to realize individualized early intervention and treatment.

## Data Availability

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

## Ethical Approval

The studies involving human participants were reviewed and approved by the Ethics Committee of Affiliated Cancer Hospital and the Institute of Guangzhou Medical University. The patients provided their informed consent to participate in this study.

## Disclosure

Lu Zhou and Yuefeng Wen are cofirst authors.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Authors' Contributions

All authors contributed to the article and approved the submitted version. LZ designed the study, conducted most of experiments, analyzed the experimental results, and drafted the manuscript. YW designed the study, collected the patient data, and provided guidance for clinical data analysis. GZ and LW reviewed and revised the manuscript. SZ and SW designed the study, gave insight into experimental design, and reviewed and revised the manuscript.

## Acknowledgments

This study was supported by the Guangzhou Health Science and Technology Project of China (2021A011097), the Basic and Applied Basic Research Foundation of Guangdong Province, China (2020A1515110577), and the Basic and Applied Basic Research Foundation of Guangdong Province, China (2021A1515111200).

## Supplementary Materials

Supplementary File 1: the clinical and treatment characteristics of 91 patients. File 2: the hyper-parameters for the eleven classifiers. File 3: the evaluation indicators of four ML models under 11 classifiers. (*Supplementary Materials*)

## References

- [1] J. Ma, J. Zhang, S. Zhou et al., "Regional lung density changes after radiation therapy for tumors in and around thorax," *International Journal of Radiation Oncology, Biology, Physics*, vol. 76, no. 1, pp. 116–122, 2010.
- [2] D. A. Palma, J. V. S. D. Koste, W. Verbakel, A. Vincent, and S. Senan, "Lung density changes after stereotactic radiotherapy: a quantitative analysis in 50 patients," *International Journal of Radiation Oncology, Biology, Physics*, vol. 81, no. 4, pp. 974–978, 2011.
- [3] J. P. Hart, M. R. Mccurdy, M. Ezhil et al., "Radiation pneumonitis: correlation of toxicity with pulmonary metabolic radiation response," *International Journal of Radiation Oncology, Biology, Physics*, vol. 71, no. 4, pp. 967–971, 2008.
- [4] R. J. Gillies, P. E. Kinahan, and H. Hricak, "Radiomics: images are more than pictures. They are data," *Radiology*, vol. 278, no. 2, pp. 563–577, 2016.
- [5] P. Lambin, E. Rios Velazquez, R. Leijenaar et al., "Radiomics: extracting more information from medical images using advanced feature analysis," *European Journal of Cancer*, vol. 48, no. 4, pp. 441–446, 2012.
- [6] M. Vallières, C. R. Freeman, S. R. Skamene, and I. El Naqa, "A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities," *Physics in Medicine and Biology*, vol. 60, no. 14, pp. 5471–5496, 2015.
- [7] J. E. van Timmeren, R. T. H. Leijenaar, W. van Elmpt et al., "Survival prediction of non-small cell lung cancer patients using radiomics analyses of cone-beam CT images," *Radiotherapy & Oncology*, vol. 123, no. 3, pp. 363–369, 2017.
- [8] H. J. W. L. Aerts, E. R. Velazquez, R. T. H. Leijenaar et al., "Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach," *Nature Communications*, vol. 5, no. 1, p. 4006, 2014.
- [9] T. P. Coroller, P. Grossmann, Y. Hou et al., "CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma," *Radiotherapy and Oncology*, vol. 114, no. 3, pp. 345–350, 2015.
- [10] S. P. Krafft, A. Rao, F. Stingo et al., "The utility of quantitative CT radiomics features for improved prediction of radiation pneumonitis," *Medical Physics*, vol. 45, no. 11, pp. 5317–5324, 2018.
- [11] L. Rossi, R. Bijman, W. Schillemans et al., "Lung texture in serial thoracic computed tomography scans: correlation of radiomics-based features with radiation therapy dose and radiation pneumonitis development," *Radiotherapy & Oncology*, vol. 91, no. 3, pp. 1048–1056, 2015.
- [12] L. Rossi, R. Bijman, W. Schillemans et al., "Texture analysis of 3D dose distributions for predictive modeling of toxicity rates in radiotherapy," *Radiology and Oncology*, vol. 129, pp. 548–553, 2018.
- [13] H. S. Gabryś, F. Buettner, F. Sterzing, H. Hauswald, and M. Bangert, "Design and selection of machine learning methods using radiomics and dosiomics for normal tissue complication probability modeling of xerostomia," *Frontiers in Oncology*, vol. 8, p. 35, 2018.
- [14] X. Zhen, J. Chen, Z. Zhong et al., "Deep convolutional neural network with transfer learning for rectum toxicity prediction in cervical cancer radiotherapy: a feasibility study," *Physics in Medicine and Biology*, vol. 62, no. 21, pp. 8246–8263, 2017.
- [15] L. Placidi, E. Gioscio, C. Garibaldi et al., "A multicentre evaluation of dosiomics features reproducibility, stability and sensitivity," *Cancers*, vol. 13, no. 15, p. 3835, 2021.
- [16] Y. Murakami, T. Soyano, T. Kozuka et al., "Dose-based radiomic analysis dosiomics for intensity modulated radiation therapy in patients with prostate cancer: correlation between planned dose distribution and biochemical failure," *International Journal of Radiation Oncology, Biology, Physics*, vol. 112, no. 1, pp. 247–259, 2022.
- [17] B. Liang, H. Yan, Y. Tian et al., "Dosiomics: extracting 3D spatial features from dose distribution to predict incidence of radiation pneumonitis," *Frontiers in Oncology*, vol. 9, no. 269, p. 269, 2019.
- [18] B. Liang, Y. Tian, X. Y. Chen et al., "Prediction of radiation pneumonitis with dose distribution: a convolutional neural network (CNN) based model," *Frontiers in Oncology*, vol. 9, no. 1500, p. 1500, 2019.
- [19] L. Bin, T. Yuan, S. Zhaohui et al., "A deep learning-based dual-omics prediction model for radiation pneumonitis," *Medical Physics*, vol. 48, no. 10, pp. 6247–6256, 2021.
- [20] D. Kawahara, N. Imano, R. Nishioka et al., "Prediction of radiation pneumonitis after definitive radiotherapy for locally advanced non-small cell lung cancer using multi-region radiomics analysis," *Scientific Reports*, vol. 11, no. 1, Article ID 16232, 2021.
- [21] W. Jiang, Y. Song, Z. Sun, J. Qiu, and L. Shi, "Dosimetric factors and radiomics features within different regions of interest in planning CT images for improving the prediction of radiation pneumonitis," *International Journal of Radiation Oncology, Biology, Physics*, vol. 110, no. 4, pp. 1161–1170, 2021.



- [22] D. A. Palma, S. Senan, K. Tsujino et al., "Predicting radiation pneumonitis after chemoradiation therapy for lung cancer: an international individual patient data meta-analysis," *International Journal of Radiation Oncology, Biology, Physics*, vol. 85, no. 2, pp. 444–450, 2013.
- [23] H. Asakura, T. Hashimoto, and S. Zenda, "Analysis of dose-volume histogram parameters for radiation pneumonitis after definitive concurrent chemoradiotherapy for esophageal cancer," *Radiotherapy and Oncology*, vol. 95, no. 2, pp. 240–244, 2010.
- [24] B. Yao, Y. D. Wang, and Q. Z. Liu, "Radiation pneumonitis in non-small-cell lung cancer patients treated with helical tomotherapy," *Nigerian Journal of Clinical Practice*, vol. 19, no. 1, pp. 25–29, 2016.
- [25] K. Boonyawan, D. R. Gomez, R. Komaki et al., "Clinical and dosimetric factors predicting grade  $\geq 2$  radiation pneumonitis after postoperative radiotherapy for patients with non-small cell lung carcinoma," *International Journal of Radiation Oncology, Biology, Physics*, vol. 101, no. 4, pp. 919–926, 2018.
- [26] S. Shi, Z. C. Zeng, L. X. Ye, Y. Huang, and J. He, "Risk factors associated with symptomatic radiation pneumonitis after stereotactic body radiation therapy for stage I non-small cell lung cancer," *Technology in Cancer Research and Treatment*, vol. 16, no. 3, pp. 316–320, 2017.
- [27] S. Ramella, L. Trodella, T. C. Mineo et al., "Adding ipsilateral V20 and V30 to conventional dosimetric constraints predicts radiation pneumonitis in stage IIIA-B NSCLC treated with combined-modality therapy," *International Journal of Radiation Oncology, Biology, Physics*, vol. 76, no. 1, pp. 110–115, 2010.
- [28] T. Adachi, M. Nakamura, T. Shintani et al., "Multi-institutionaldose-segmented dosiomic analysis for predicting radiation pneumonitis after lung stereotactic body radiation therapy," *Medical Physics*, vol. 48, no. 4, pp. 1781–1791, 2021.
- [29] T. M. Briere, S. Krafft, Z. Liao, and M. K. Martel, "Lung size and the risk of radiation pneumonitis," *International Journal of Radiation Oncology, Biology, Physics*, vol. 94, no. 2, pp. 377–384, 2016.
- [30] G. G. Steel, *Basic Clinical Radiobiology*, Oxford University Press, London, UK, 3rd edition, 2002.
- [31] G. W. Barendsen, "Dose fractionation, dose rate, and isoeffect relationships for normal tissue responses," *International Journal of Radiation Oncology, Biology, Physics*, vol. 8, no. 11, pp. 1981–1997, 1982.
- [32] B. Maciejewski, J. M. Taylor, and H. Withers, "Alpha/beta value and the importance of size of dose per fraction for late complications in the supraglottic larynx," *Radiotherapy & Oncology*, vol. 7, no. 4, pp. 323–326, 1986.
- [33] K. K. Fu, T. F. Pajak, A. Trotti et al., "A radiation therapy oncology group (RTOG) phase III randomized study to compare hyperfractionation and two variants of accelerated fractionation to standard fractionation radiotherapy for head and neck squamous cell carcinomas: first report of RTOG 9003," *International Journal of Radiation Oncology, Biology, Physics*, vol. 48, no. 1, pp. 7–16, 2000.
- [34] J. H. Hendry, S. M. Bentzen, R. G. Dale et al., "A modelled comparison of the effects of using different ways to compensate for missed treatment days in radiotherapy," *Clinical Oncology*, vol. 8, no. 5, pp. 297–307, 1996.
- [35] J. F. Fowler, "The linear-quadratic formula and progress in fractionated radiotherapy," *British Journal of Radiology*, vol. 62, no. 740, pp. 679–694, 1989.
- [36] Pyradiomics, "Pyradiomics Radiomic Features," 2017, <https://pyradiomics.readthedocs.io/en/latest/features.html>.
- [37] M. Sandri, P. Zuccolotto, and C. Tuleau-Malot, "Variable selection using random forests," *Studies in Classification, Data Analysis, and Knowledge Organization*, Springer, no. 14, pp. 263–270, Berlin, Germany, 2010.
- [38] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 3, no. 6, pp. 610–621, 1973.
- [39] E. Huynh, T. P. Coroller, V. Narayan et al., "CT-based radiomic analysis of stereotactic body radiation therapy patients with lung cancer," *Radiotherapy & Oncology*, vol. 120, no. 2, pp. 258–266, 2016.
- [40] G. Thibault, B. Fertil, C. Navarro, S. Pereira, and J. L. Mari, "Texture Indexes and Gray Level Size Zone Matrix Application to Cell Nuclei Classification," in *Proceedings of the 10th International Conference on Pattern Recognition and Information Processing*, Minsk, Belarus, November 2009.
- [41] J. J. Van Griethuysen, A. Fedorov, C. Parmar et al., "Computational radiomics system to decode the radiographic phenotype," *Cancer Research*, vol. 77, no. 21, pp. e104–e107, 2017.
- [42] K. H. Cho, S. J. Ahn, H. R. Pyo et al., "A phase II study of synchronous three-dimensional conformal boost to the gross tumor volume for patients with unresectable stage III non-small-cell lung cancer: results of Korean radiation oncology group 0301 study," *International Journal of Radiation Oncology, Biology, Physics*, vol. 74, no. 5, pp. 1397–1404, 2009.
- [43] J. Maguire, N. O Rourke, and P. C. Soccar, "Sequential or concurrent chemotherapy and hypo-fractionated accelerated radiotherapy in inoperable stage I NSCLC," *Journal of Thoracic Oncology*, vol. 6, pp. 1561–1562, 2011.
- [44] A. S. Allal, S. Bieri, and B. A. Bründler, "Preoperative hyperfractionated radiotherapy for locally advanced rectal cancers: a phase I-II trial," *International Journal of Radiation Oncology Biological Physics*, vol. 54, no. 4, pp. 1076–1081, 2002.
- [45] M. Hatton and J. E. Martin, "Continuous hyperfractionated accelerated radiotherapy (CHART) and non-conventionally fractionated radiotherapy in the treatment of non-small cell lung cancer: a review and consideration of future directions," *Clinical Oncology*, vol. 22, no. 5, pp. 356–364, 2010.