

学位論文（博士）

Investigation of the combination of
intratumoral and peritumoral radiomic
signatures for predicting epidermal growth
factor receptor mutation in lung adenocarcinoma
(腫瘍および腫瘍周囲のRadiomic特徴量を用いた肺腺
がんにおけるEGFR遺伝子変異の予測モデルの作成)

氏名 川添 優介

所属 山口大学大学院医学系研究科
医学専攻 放射線腫瘍学講座

令和5年 11月

所属 放射線腫瘍学講座

氏名 川添 優介

〔題名〕

Investigation of the combination of intratumoral and peritumoral radiomic signatures for predicting epidermal growth factor receptor mutation in lung adenocarcinoma

(腫瘍および腫瘍周囲の Radiomic 特徴量を用いた肺腺がんにおける EGFR 遺伝子変異の予測モデルの作成)

Journal of Applied Clinical Medical Physics Vol.26 Issue 6 e13980 (令和 5 年 4 月掲載)

〔研究背景〕

肺腺癌と診断された患者のうち、上皮成長因子受容体 (EGFR) 遺伝子変異を有する患者に対して、EGFR チロシンキナーゼ阻害剤 (EGFR-TKI) を用いた分子標的療法が行われる。EGFR-TKI による治療は、従来の化学療法よりも患者の生存率を向上させ、無増悪生存期間を延長させることが報告されている。さらに、EGFR 遺伝子変異は EGFR-TKI によって放射線感受性を増加させることが報告されている。そのため、EGFR 遺伝子変異の状態を正確に把握することは治療方針の決定において重要となる。

EGFR 遺伝子変異の検出には通常、生検や外科的な肺区域切除が必要であるが、これらは侵襲的な手技を伴う。また、EGFR 遺伝子変異を同定するには時間を要するため、治療開始が遅れることが問題となる。これまでに、放射線画像から抽出した高次元の定量的な特徴量 (Radiomic 特徴量) と機械学習や深層学習などを使用した EGFR 遺伝子変異の予測モデルが提案されている。予測モデルを構築することで、侵襲的な手技を伴うことなく短時間で EGFR 遺伝子変異の有無を識別することが期待される。しかし、これらの研究では腫瘍周囲領域の Radiomic 特徴量を考慮せず、腫瘍内部領域のみを評価している。腫瘍内部および周囲領域の Radiomic 特徴量は腫瘍の正常組織への広がりやの予測に有用であることが報告されている。腫瘍の正常組織への広がりや EGFR 遺伝子変異と関連しているため、腫瘍内部および周囲領域の両方が EGFR 遺伝子変異の予測に重要である。本研究では、EGFR 遺伝子変異の予測に最適な腫瘍周囲領域サイズを探索した。最適な腫瘍周囲領域から抽出した Radiomic 特徴量と機械学習アルゴリズムを用いて、EGFR 遺伝子変異の予測モデルを構築し、予測精度を調査した。

〔要旨〕

当院の倫理審査委員会で承認の得られた、肺腺癌を有する患者 164 名を後ろ向きに解析した。すべての患者は、腫瘍の生検、もしくは肺区域切除により EGFR 遺伝子変異の有無が確認されている。患者データは 7:3 の比率で学習データとテストデータに振り分けられた。本研究では、生検、もしくは肺区域切除を行う前に撮像された computed tomography (CT) 画像から腫瘍内領域の Radiomic 特徴量を抽出した。さらに腫瘍の境界から半径 3, 5, 7mm の放射状に伸びる領域を腫

瘍周囲領域とし、腫瘍内領域と大きさの異なる 3 つの腫瘍周囲領域を足し合わせた領域の Radiomic 特徴量をそれぞれ抽出した。分散分析およびラッソ回帰を用いて重要度の高い特徴量のみを選択した。最適な腫瘍周囲サイズは、Radiomic スコア (rad-score) により決定した。Rad-score は選択された特徴量とその重要度を掛け合わせることで算出し、Wilcoxon の順位和検定を用いて評価した。腫瘍内の Radiomic 特徴量と患者の臨床的特徴を合わせた intratumoral radiomic signature (IRS) を用いて、EGFR 遺伝子変異の予測モデルを構築した。同様に、腫瘍内領域と大きさの異なる 3 つの腫瘍周囲領域を足し合わせた領域の Radiomic 特徴量と臨床的特徴から構成される intratumoral and peritumoral radiomic signatures (それぞれ IPRS3, IPRS5, IPRS7) を用いて、予測モデルを構築した。予測モデルには、3 つの機械学習アルゴリズム (サポートベクターマシン (support vector machine: SVM), ロジスティック回帰 (logistic regression: LR), LightGBM) を用いて、それぞれのアルゴリズムのパラメータは 5 分割交差検証により最適化し、受信者動作特性曲線の曲線下面積 (area under the curve: AUC) を用いて評価した。学習モデルは独立したテストモデルを用いて評価した。また、Brier score (BS) によりモデルの正確性を、決定曲線解析 (decision curve analysis: DCA) によりモデルの臨床的有用性を評価した。

結果として、Wilcoxon の順位和検定によると、3mm サイズの腫瘍周囲領域が最適であることが明らかとなった。IRS から構築された SVM, LR, LightGBM のアルゴリズムを用いた予測モデルの AUC は、学習モデルでは 0.783 (95%信頼区間 : 0.602-0.956), 0.789 (0.654-0.927), 0.735 (0.613-0.958), テストモデルでは 0.791 (0.641-0.920), 0.781 (0.538-0.930), 0.734 (0.538-0.930) であった。一方、IPRS3 から構築された SVM, LR, lightGBM モデルの AUC は、学習モデルでは、0.831 (0.666-0.984), 0.804 (0.622-0.908), 0.769 (0.628-0.921) であり、テストモデルでは、0.765 (0.644-0.921), 0.783 (0.583-0.921), 0.796 (0.583-0.949) であった。IPRS3 から構築された LR および LightGBM のアルゴリズムを用いたモデルの BS および DCA は、IRS から構築されたモデルよりも良好であった。

EGFR 遺伝子変異を予測するための最適な腫瘍周囲領域は 3mm であり、IPRS3 は IPRS5 または IPRS7 よりも EGFR 遺伝子変異の有無をよく区別できることが示された。さらに、IPRS3 は IRS よりも EGFR 遺伝子変異の有無を高い精度で予測でき、正確性、臨床的有用性が高いことが明らかとなった。したがって、腫瘍内部だけでなく、腫瘍周囲領域の Radiomic 特徴量を考慮することが EGFR 遺伝子変異を予測するために重要であると考えられる。しかし、最適な腫瘍周囲領域は、腫瘍の輪郭抽出の精度によって変化しうる。腫瘍の輪郭抽出の精度は操作者によって異なるため、操作者間の再現性を評価する必要があると考えられる。EGFR 遺伝子変異を同定するためには、侵襲的な手技と時間を要する。我々が開発した CT 画像に基づく EGFR 遺伝子変異の予測モデルを用いることで、治療戦略に関する早期の意思決定を促進することができると考えられる。さらに、我々は、大きさの異なる 3 つの腫瘍周囲領域の Radiomic 特徴量と機械学習アルゴリズムを用いた予測モデルを構築した。IPRS3 を用いた場合、予測モデルは機械学習アルゴリズムの違いによらず高い精度を示したため、IPRS3 を用いた EGFR 遺伝子変異の予測モデルの頑健性は高いことが示唆された。

本研究では、EGFR 遺伝子変異の識別に最適な腫瘍周囲領域サイズを探索した。腫瘍内と 3mm サイズの腫瘍周囲領域の Radiomic 特徴量の組み合わせで構築した予測モデルは、EGFR 遺伝子変異の予測に有用であると考えられる。

Investigation of the combination of intratumoral and peritumoral radiomic signatures for predicting epidermal growth factor receptor mutation in lung adenocarcinoma

Yusuke Kawazoe¹ | Takehiro Shiinoki¹  | Koya Fujimoto¹ | Yuki Yuasa¹ |
Tsunahiko Hirano² | Kazuto Matsunaga² | Hidekazu Tanaka¹

¹Department of Radiation Oncology, Graduate School of Medicine, Yamaguchi University, Ube, Japan

²Department of Respiratory Medicine and Infectious Disease, Graduate School of Medicine, Yamaguchi University, Ube, Japan

Correspondence

Takehiro Shiinoki, Department of Radiation Oncology, Yamaguchi University, Graduate School of Medicine, 1-1-1 Minamikogushi, Ube, Yamaguchi 755–8505, Japan.
Email: shiinoki@yamaguchi-u.ac.jp

Funding information

Japan Society for the Promotion of Science (JSPS) KAKENHI, Grant/Award Number: 22K07667; Takeda Science Foundation

Abstract

Purpose: We investigated optimal peritumoral size and constructed predictive models for epidermal growth factor receptor (EGFR) mutation.

Methods: A total of 164 patients with lung adenocarcinoma were retrospectively analyzed. Radiomic signatures for the intratumoral region and combinations of intratumoral and peritumoral regions (3, 5, and 7 mm) from computed tomography images were extracted using analysis of variance and least absolute shrinkage. The optimal peritumoral region was determined by radiomics score (rad-score). Intratumoral radiomic signatures with clinical features (IRS) were used to construct predictive models for EGFR mutation. Combinations of intratumoral and 3, 5, or 7 mm-peritumoral signatures with clinical features (IPRS3, IPRS5, and IPRS7, respectively) were also used to construct predictive models. Support vector machine (SVM), logistic regression (LR), and LightGBM models with five-fold cross-validation were constructed, and the receiver operating characteristics were evaluated. Area under the curve (AUC) of the training and test cohorts values were calculated. Brier scores (BS) and decision curve analysis (DCA) were used to evaluate the predictive models.

Results: The AUC values of the SVM, LR, and LightGBM models derived from IRS were 0.783 (95% confidence interval: 0.602–0.956), 0.789 (0.654–0.927), and 0.735 (0.613–0.958) for training, and 0.791 (0.641–0.920), 0.781 (0.538–0.930), and 0.734 (0.538–0.930) for test cohort, respectively. Rad-score confirmed that the 3 mm-peritumoral size was optimal (IPRS3), and AUCs values of SVM, LR, and lightGBM models derived from IPRS3 were 0.831 (0.666–0.984), 0.804 (0.622–0.908), and 0.769 (0.628–0.921) for training and 0.765 (0.644–0.921), 0.783 (0.583–0.921), and 0.796 (0.583–0.949) for test cohort, respectively. The BS and DCA of the LR and LightGBM models derived from IPRS3 were better than those from IRS.

Conclusion: Accordingly, the combination of intratumoral and 3 mm-peritumoral radiomic signatures may be helpful for predicting EGFR mutations.

KEYWORDS

EGFR mutation, machine learning, radiomic feature

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Journal of Applied Clinical Medical Physics* published by Wiley Periodicals, LLC on behalf of The American Association of Physicists in Medicine.

1 | INTRODUCTION

Mutational testing is the standard protocol for determining whether patients with non-small cell lung cancer (NSCLC) are likely to respond to targeted molecular therapy.¹ Lung adenocarcinoma is classified as an NSCLC.² Patients with lung adenocarcinoma with epidermal growth factor receptor (EGFR) mutations are treated with EGFR tyrosine kinase inhibitors (EGFR-TKIs).^{3,4} Treatment with EGFR-TKIs has given patients better survival rate and longer progression-free survival times than conventional chemotherapy.⁵ However, the regulation of EGFR mutations by EGFR-TKIs increases radiosensitivity.⁶ Therefore, identifying the EGFR mutation status is crucial for decision-making regarding treatment regimens.

Biopsies or surgical specimens are typically obtained for detecting EGFR mutations.⁷ However, these processes are time consuming, expensive, and invasive. Some researchers have proposed predictive models for EGFR mutations using radiomic features derived from computed tomography images.^{1,8–10} Radiomics can analyze tumor phenotypes by automatically extracting numerous quantitative features from medical images, such as CT and/or magnetic resonance images.¹¹ However, most studies have not considered the radiomic features derived from the peritumoral region and assessed the intratumoral region alone.^{7,9}

Some studies have reported the usefulness of radiomic features derived from intratumoral and peritumoral regions for predicting tumor spread in air space (STAS).^{12–14} STAS is also associated with EGFR mutations.¹⁴ Moreover, the predictive model by Wang et al. established that both the intratumoral and peritumoral regions are important for predicting EGFR mutation.¹⁵

Very few studies have used intratumoral and peritumoral radiomic features to predict EGFR mutation. Yamazaki et al. and Choe et al. reported the usefulness of peritumoral radiomic features in predicting EGFR mutation status.^{1,10} Their methods used a single setting with peritumoral sizes of 3 and 5 mm from the tumor border. Because the studies used different peritumoral size, optimal peritumoral size for predicting EGFR mutation status must be investigated. Therefore, this study explored the radiomic features of the optimal peritumoral size to determine EGFR mutation status and construct machine learning (ML) based predictive models for EGFR mutation status.

2 | MATERIALS AND METHODS

2.1 | Patient data

The Institutional Review Board of our institution approved this study. The inclusion criteria were as

follows: (a) pathologically confirmed lung adenocarcinoma, (b) confirmed EGFR mutation (EGFR+) or wild-type (EGFR–), (c) non-contrast enhanced chest CT images acquired before surgery or targeted molecular therapy or radiation therapy, and (d) only primary tumors. The exclusion criteria were as follows: (a) patients with tumors other than lung adenocarcinoma and, (b) patients who had previously undergone surgery or targeted molecular therapy. A total of 164 patients with NSCLC who had undergone biopsy or surgical specimens between 2016 and 2020 by our institution were randomly selected. These cases were divided into EGFR+ or EGFR– groups in both the training and test cohorts. The data were randomly divided into training and test cohorts with a ratio of 7:3. The clinical features included age, sex, location of the lung tumor, smoking status, and staging. The tumors were divided into five location categories: right upper, right middle, right lower, left upper, and left lower.^{2,16} The detailed characteristics of the patients in our study are shown in Table 1. The study workflow is shown in Figure 1.

2.2 | CT imaging

CT examinations were performed using five CT scanners: Aquilion Precision (Canon Medical Systems, Otawara, Japan), Optima CT 660 (GE Healthcare, Waukesha, WI, USA), SOMATOM Sensation 64, SOMATOM Force, and SOMATOM Drive (Siemens Healthcare, Forchheim, Germany). The scanning parameters were as follows: tube voltage, 70–120 kV; tube current, automatic exposure control; matrix size, 512 × 512; slice thickness, 1.00 or 1.25 mm; and field of view, 270–400 mm; rotation time of gantry, 0.5 s/rot. All CT images were acquired from patients in the supine position and deep inspiration breath-hold with both hands raised.

2.3 | Extraction of radiomic features and feature selection

The acquired images were converted to an isotropic volume (1.00 × 1.00 × 1.00 mm³) using linear interpolation. The intratumoral region of the lung tumor was segmented semi-automatically using the GrowCut module in the open-source software 3D Slicer (version 4.10.2, Brigham and Women's Hospital).^{17,18} The pathological usefulness of GrowCut segmentation for NSCLC has been reported.¹⁸ Two medical physicists observed CT images on the axial, coronal, and sagittal views using the mediastinum (width, 350 HU; level, 40 HU) and lung window (width, 1500 HU; level, –500 HU) settings and performed segmentation. These segmentations were confirmed by a radiation oncologist with over 16 years of experience in radiation therapy. The peritumoral region

TABLE 1 Patient characteristics in this study.

Characteristic	Train cohort (n = 120)			Test cohort (n = 44)		
	EGFR-	EGFR+	p-value	EGFR-	EGFR+	p-value
Age (y, mean ± SD)	70.30 ± 9.90	74.15 ± 7.04	0.40	69.30 ± 8.73	67.19 ± 12.07	0.41
Sex, n			<0.001			<0.001
Male	45	20		16	8	
Female	15	40		7	13	
Tumor location, n			0.25			0.21
Right upper	26	16		11	8	
Middle	0	6		1	1	
Right lower	16	12		4	4	
Left upper	13	17		4	3	
Left lower	5	9		3	5	
Smoking, n			<0.001			<0.001
Yes	48	24		21	9	
No	12	36		2	12	
Staging, n			0.37			0.37
I	25	33		14	19	
II	5	9		1	1	
III	11	5		4	0	
IV	17	12		4	1	
N/A	3	1		0	0	

Abbreviation: N/A, not available.

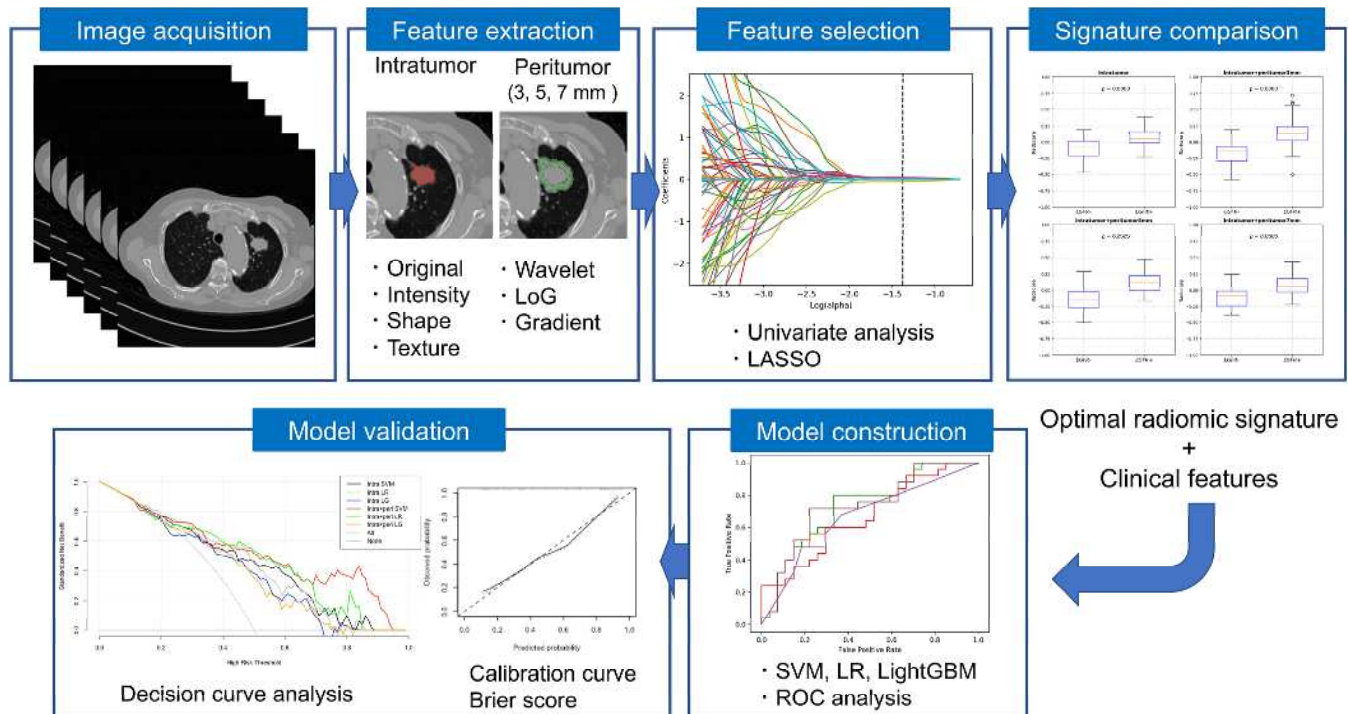


FIGURE 1 Study design and workflow.

was determined using quantitative morphologic operations as a radially extending region with 3, 5, and 7 mm radius from the intratumoral region of the tumor boundary.¹⁹ These peritumoral regions included air in the lungs, pulmonary vessels, and bronchi and did not include the thoracic wall and mediastinum.

Radiomic features were extracted from intratumoral and peritumoral regions using the open-source software Pyradiomics (version 3.7.1) in Python.²⁰ There were 1046 radiomic features extracted from each region, including first-order (14), shape (18), gray-level co-occurrence matrix (GLCM) (22), gray-level run length matrix (GLRLM) (16), gray-level size zone matrix (GLSZM) (16), gray-level dependence matrix (GLDM) (14), Laplacian of Gaussian filters (LoG) (2), gradient filter (1), and wavelet filters (8). The filtered features were acquired by multiplying the above filters by the first-order, GLCM, GLRLM, GLSZM, and GLDM features. Finally, 1046 radiomic features including first-order, shape, GLCM, GLRLM, GLSZM, and GLDM features (100) from original image and wavelet (688), LoG (172), and gradient (86) features from filtered image were extracted. The wavelet transform applies a wavelet filter to each CT image, which is then decomposed into low and high frequencies into eight different images.²¹ The major settings for radiomic features extraction were as follows: bin width of feature extraction parameters, 30²²; sigma size for the LoG filter, 1.0 or 3.0 mm; bin width of the wavelet filter, 10. ResamplePixelSpacing was set to none.¹⁹ In total, 1046 radiomic features were extracted from intratumoral and peritumoral regions using the above conditions. The combination of intratumoral and peritumoral regions included 2092 radiomic features.

All radiomic features were standardized using the StandardScaler method in the scikit-learn package.²³ For the 1046 radiomic features derived from the intratumoral region or 2092 radiomic features derived from the combination of intratumoral and peritumoral regions, the selectKbest method in the scikit-learn package based on analysis of variance and the least absolute shrinkage and selection operator were applied to training cohorts to reduce redundant features.^{9,24} The k value was set to 500 in the selectKbest method. Five-fold cross-validation was applied to the training cohort to determine the tuning parameter that regularized the magnitude of the penalization, and features with non-zero coefficients were selected. The radiomics score (rad-score) was calculated using a linear combination of selected features multiplied by their coefficients.^{9,17} The rad-scores calculated from radiomic features derived from intratumoral region and a combination of intratumoral and 3, 5, or 7 mm-peritumoral regions were evaluated using the Wilcoxon rank-sum test to determine optimal peritumoral size for distinguishing EGFR+ and EGFR-.

2.4 | Construction of machine learning based predictive models and performance evaluation

After comparing the rad-scores, the peritumoral size exhibiting the largest difference between the EGFR+ and EGFR- groups was determined as the optimal peritumoral radiomic signature. Then, combinations of intratumoral and 3, 5, or 7 mm-peritumoral radiomic signatures were combined with clinical features that showed significant differences, called intratumoral and peritumoral radiomic signatures with clinical features (IPRS3, IPRS5, and IPRS7, respectively). Similarly, we combined intratumoral radiomic signatures and clinical features which showed a significant difference, called intratumoral radiomic signatures with clinical features (IRS). Three ML predictive models (support vector machine [SVM], logistic regression [LR], and LightGBM) were constructed for EGFR mutation status using IRS and PRS. In the SVM model, a radial basis function was applied, and the grid search method with a five-fold CV was applied to optimize the hyperparameters. In the LightGBM model, to avoid overfitting, it was necessary to add a maximum depth limit; therefore, hyperparameters were optimized using random search in five-fold CV in the training cohorts.

The predictive performance of each ML model was evaluated using the area under the curve (AUC) of the receiver operating characteristic curve in five-fold CV. The training models were then evaluated using independent test cohorts. Furthermore, the calibration curve and the Brier score (BS) were used to evaluate the accuracy of ML models, and decision curve analysis (DCA) was used to evaluate the clinical applicability of the ML classifier models.²⁵ The BS is calculated by summing the squared difference between the probability of prediction and the real probability.²⁶ If the BS is 0, the model is considered to have perfect predictive accuracy; if the BS greater than 0.25, the model is considered to have no value.^{25,27} DCA was performed by calculating the net benefit. The net benefit = true positive rate—(false positive rate × weighting factor), where the weighting factor = the threshold/(1 - threshold). Differences were considered statistically significant at $p < 0.05$. All the procedures were performed using in-house programs (Python ver. 3.7.1, R ver. 4.1.1).

3 | RESULTS

Among 2092 features derived from combinations of the intratumoral region and 3, 5, and 7 mm-peritumoral regions, 22, 14, and 13 features were selected, respectively. Among the 1046 features derived from intratumoral features alone, 13 features were selected. Figure 2a shows the rad-score of the intratumoral

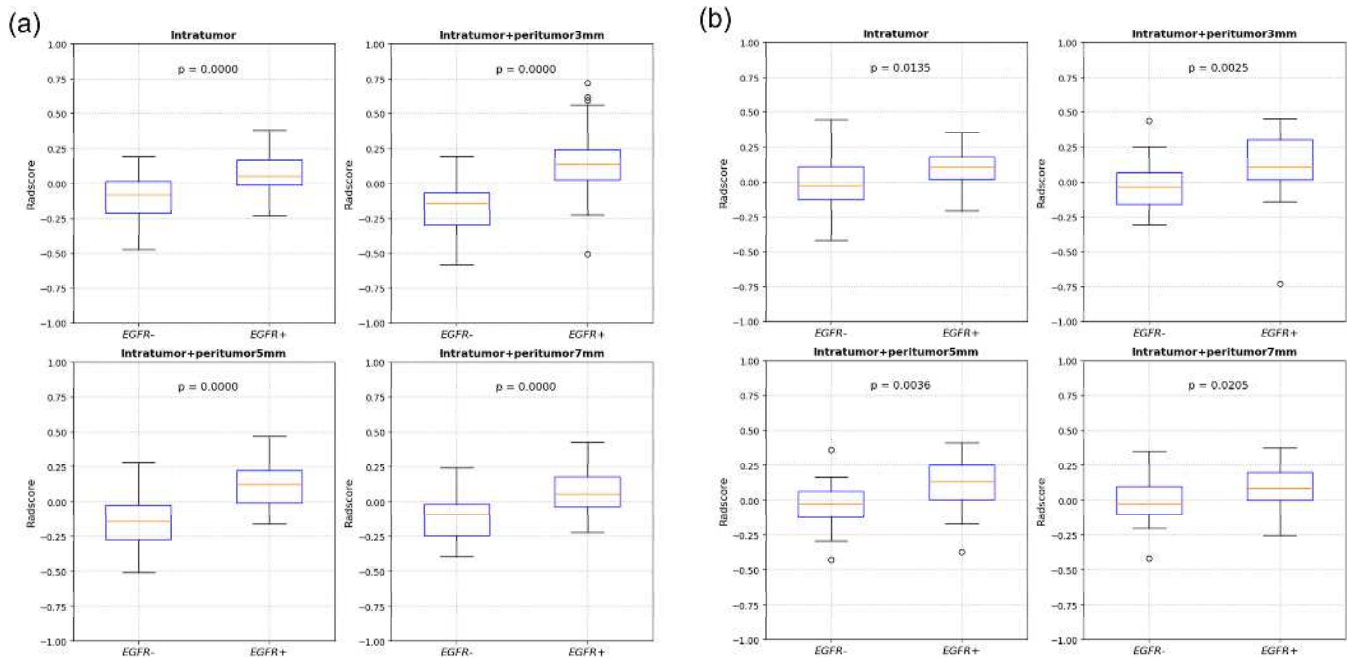


FIGURE 2 The rad-score of only intratumoral radiomic signatures and combinations of intratumoral and 3, 5, and 7 mm-peritumoral radiomic signatures between EGFR+ and EGFR– groups of the (a) training, and (b) test cohorts.

radiomic signatures alone and the combinations of intratumoral and 3, 5, and 7 mm-peritumoral radiomic signatures between the EGFR+ and EGFR– groups for the training, respectively. Figure 2b shows the rad-scores for the test cohort. The rad-score showed a significant difference between the EGFR+ and EGFR– groups in the intratumoral and peritumoral regions in both cohorts. In particular, the rad-score derived from the combination of intratumoral and 3 mm-peritumoral radiomic signatures showed the largest difference between the EGFR+ and EGFR– groups (training: $p = 0.0000$, test: $p = 0.0025$). Therefore, the optimal peritumoral size was determined to be 3 mm. Figure 3a shows the radiomic signatures for calculating the rad-score for the intratumoral region, the combinations of intratumoral region and (b) 3 mm-peritumoral, (c) 5 mm-peritumoral, and (d) 7 mm-peritumoral regions.

Differences in clinical features are shown in Table 1, with sex and smoking status being significantly different. Therefore, PRS were constructed using combinations of intratumoral and 3, 5, and 7 mm-peritumoral radiomic signatures with sex and smoking, namely IPRS3, IPRS5, and IPRS7, respectively, then ML models were constructed using these signatures. Similarly, IRS was constructed using intratumoral radiomic signatures, sex, and smoking status.

Table 2 shows the AUC values for the training and test cohorts for different ML models based on IRS, IPRS3, IPRS5, and IPRS7. For the training cohort, the AUC values in the SVM, LR, and LightGBM models derived from IRS were 0.783

(95% CI: 0.602–0.956), 0.789 (0.654–0.927), and 0.735 (0.613–0.958), respectively, and 0.831 (0.666–0.984), 0.804 (0.622–0.908), and 0.769 (0.628–0.921) derived from IPRS3, respectively. For the test cohort, these were 0.791 (95% CI: 0.641–0.920), 0.781 (0.538–0.930), and 0.734 (0.538–0.930) derived from IRS and 0.765 (0.644–0.921), 0.783 (0.583–0.949), and 0.796 (0.583–0.949) derived from IPRS3, respectively.

The calibration curves of the predictive models derived from IPRS3 are shown in Figure 4. The calibration curve evaluates the goodness of fit between the predicted probabilities and models with the actual outcomes of EGFR mutation, namely, predictive model accuracy, with the better model being closer to the actual outcome, as shown by the dashed line.²⁸ In the training cohort, the goodness of fit between the predicted probability and models with the actual outcomes of EGFR mutations appeared to be good in all models. In the test cohort, the goodness of fit LR and LightGBM models around 0.4 in predicted probability were not well. The BS of the SVM, LR, and LightGBM models in the training cohort were 0.189, 0.189, and 0.210, respectively, derived from IRS, and 0.165, 0.185, and 0.212, respectively, derived from IPRS3. In the test cohort, these were 0.196, 0.207, and 0.218 for IRS and 0.213, 0.205, and 0.202 for IPRS3, respectively.

Figure 5 shows the decision curves of the three ML models for the (a) training and (b) test cohorts. All ML models derived from IPRS3 in the test cohort had more net benefit than “treat all” and the “treat

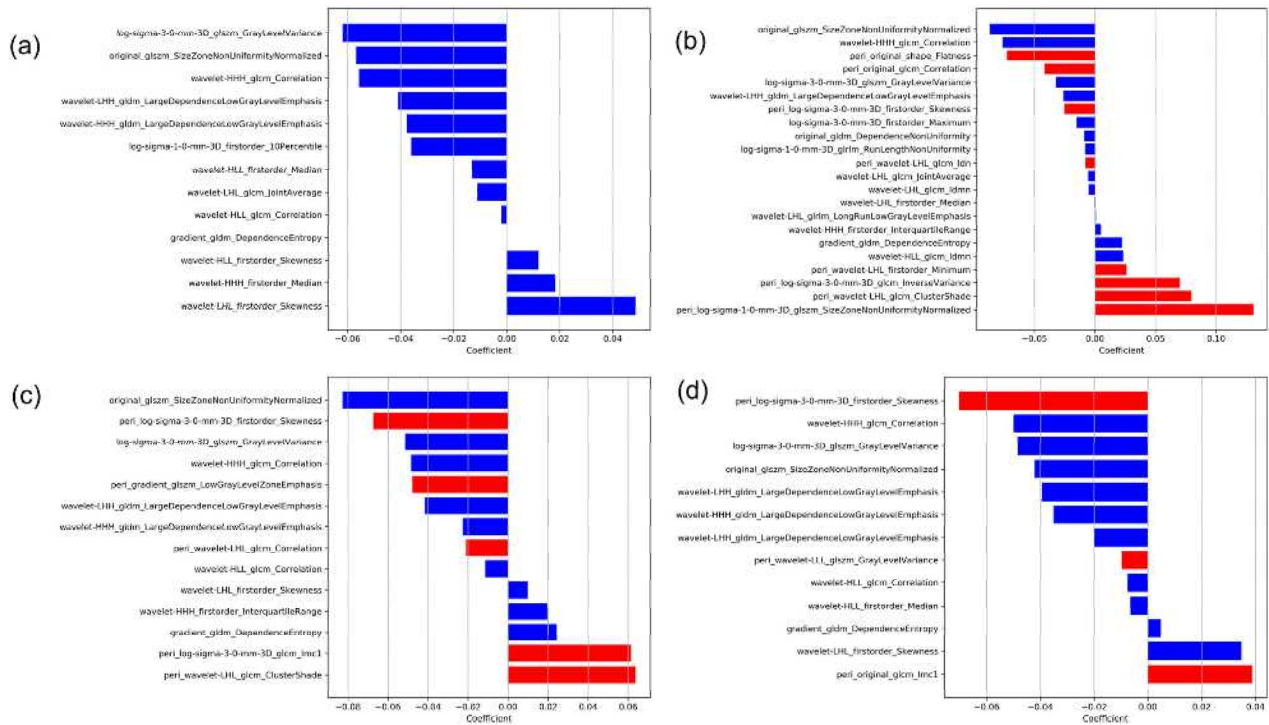


FIGURE 3 Selected radiomic signatures for calculating the rad-score of the (a) intratumoral region, (b) combination of intratumoral and 3 mm-peritumoral regions, (c) combination of intratumoral and 5 mm-peritumoral regions, and (d) combination of intratumoral and 7 mm-peritumoral regions.

none" with a threshold range over 0.3. Furthermore, in the LR model, compared with IRS, IPRS3 had more benefits in the threshold range from 0.45 to 0.60 and over 0.65 in the test cohort. In the LightGBM model, compared to IRS, IPRS3 in the test cohort had more benefits with the range from 0.05 to 0.55 in the test cohort.

4 | DISCUSSION

The peritumoral size for determining the EGFR mutation status was optimized. Our results demonstrate that radiomic signatures from the combination of intratumoral and 3 mm-peritumoral regions could distinguish EGFR+ and EGFR- groups better than 5 or 7 mm-peritumoral regions. We then constructed predictive models for EGFR mutation status using IPRS3. Our study showed that IPRS3 could better predict EGFR mutation status than IRS.

In terms of clinical features, sex, and smoking status were significantly different (Table 1). Because EGFR mutations are frequently observed in non-smokers and Asian females, this tendency was reasonable.³

As shown in Figure 3, both intratumoral and peritumoral features were selected for all combinations of intratumoral and peritumoral regions. Therefore, it is important to consider the peritumoral region to distin-

guish the EGFR mutation status, regardless of the size of the peritumoral region. In this study, the optimal peritumoral region was determined to be 3 mm based on the rad-score results. Our results showed that the ML models derived from IPRS3 performed well. Yamazaki et al. reported the usefulness of 3 mm-peritumoral radiomic features for predicting EGFR mutation status.¹⁰ Furthermore, Morales et al. reported that peritumoral lung parenchyma within 3 mm excluding the thoracic wall or mediastinum was correlated with overall survival in lung cancer.²⁹ Therefore, the use of IPRS3 is reasonable and meaningful for predicting the EGFR mutation status. However, the optimal peritumoral region can vary depending on the accuracy of intratumoral segmentation. We used GrowCut for tumor segmentation, which segments automatically from a given initial small set of label points in the algorithm.¹⁸ Therefore, it is expected that the difference in segmentation accuracy due to different operators being used may be reduced. However, because we did not evaluate the reproducibility for segmentation of interclass correlation coefficient, this will be validated.

Previous studies reported developed predictive models for the EGFR mutation status derived from intratumoral radiomic features alone. Zhao et al. reported an AUC value of 0.757 while Mei et al. reported an AUC value of 0.664.^{17,30} Moreover, Choe et al. developed a predictive model using both intratumoral

TABLE 2 AUC for training and test cohorts in different ML models based on intratumoral radiomic features with sex and smoking, and the combination of intratumoral and peritumoral radiomic features with sex and smoking.

Classifier	Signature	Training cohort			Test cohort			
		AUC	[95% CI]	Brier score	AUC	[95% CI]	Brier score	
EGFR mutation vs. wild-type	SVM	IRS	0.783 ± 0.086	[0.602–0.956]	0.189	0.791	[0.641–0.920]	0.196
		IPRS3	0.831 ± 0.059	[0.666–0.984]	0.165	0.765	[0.644–0.921]	0.213
		IPRS5	0.822 ± 0.102	[0.661–0.965]	0.171	0.776	[0.636–0.917]	0.219
		IPRS7	0.715 ± 0.109	[0.215–0.914]	0.219	0.687	[0.529–0.846]	0.233
	LR	IRS	0.789 ± 0.110	[0.650–0.927]	0.189	0.781	[0.538–0.930]	0.207
		IPRS3	0.804 ± 0.068	[0.622–0.908]	0.185	0.783	[0.583–0.949]	0.205
		IPRS5	0.785 ± 0.098	[0.605–0.955]	0.200	0.747	[0.600–0.895]	0.226
		IPRS7	0.779 ± 0.090	[0.597–0.955]	0.215	0.737	[0.588–0.887]	0.233
	LightGBM	IRS	0.735 ± 0.091	[0.613–0.958]	0.210	0.734	[0.538–0.930]	0.218
		IPRS3	0.769 ± 0.085	[0.628–0.921]	0.212	0.796	[0.583–0.949]	0.202
		IPRS5	0.736 ± 0.073	[0.537–0.934]	0.219	0.717	[0.537–0.934]	0.216
		IPRS7	0.802 ± 0.071	[0.626–0.973]	0.213	0.755	[0.626–0.973]	0.223

Abbreviations: AUC, area under the receiver operating characteristic curve; CI, confidence interval; EGFR, epidermal growth factor receptor; IPRS3, combination of intratumoral and 3 mm-peritumoral radiomic signature with clinical features; IPRS5, combination of intratumoral and 5 mm-peritumoral radiomic signature with clinical features; IPRS7, combination of intratumoral and 7 mm-peritumoral radiomic signature with clinical features; IRS, intratumoral radiomic signature with clinical features; LR, logistic regression; SVM, support vector machine.

and peritumoral radiomic features for EGFR mutations in lung adenocarcinoma, and the AUC value of their model was 0.64.¹ In contrast, the AUC value of our best model, LightGBM, demonstrated high performance in the test cohort (0.796). Although validation for a large number of cases is needed, our models derived from IPRS3 may be helpful for predicting EGFR mutation status.

For the calibration curve, all models derived from IPRS3 showed a better goodness of fit in the training cohort. However, in the LR and LightGBM models, the goodness of fit around 0.4 in predicted probability were poor in the test cohort (Figure 4). In the LR and LightGBM models, the BS derived from IPRS3 was slightly better than that derived from IRS. Because a lower BS indicates better model accuracy, these results indicate that the model accuracies of the LR and LightGBM models from IPRS3 are slightly better than that of IRS. Previously, no study has evaluated the model accuracy with BS for EGFR mutation status using 3 mm-peritumoral radiomic features; therefore, our results are considered to be valuable. However, the validity of these results must be evaluated. Moreover, it has been reported that the predictive model derived from intratumoral radiomic features had a low BS (0.162 in the SVM model)²⁵; therefore, the accuracy of our predictive model can be potentially improved.

For the DCA, the LR model derived from IPRS3 had more benefits with the threshold range from 0.45 to 0.60 and over 0.65 than that of IRS in the test cohort. The

LightGBM model derived from IPRS3 had more benefits with the threshold range from 0.05 to 0.55 than that of IRS in the test cohort (Figure 5). The threshold is where the expected benefit of treatment and the expected benefit of avoiding treatment are equal.³¹ Moreover, all models derived from IPRS3 showed at least more net benefit than “all treat” or “treat none” with a range over 0.3 in the test cohort. Therefore, clinicians can refer to our results to determine whether the EGFR mutation status based on our models will be useful or not.³² According to the results of Liu et al. and Zhang et al., net benefits vary depending on the predictive models.^{25,28} Therefore, validation of several predictive models is important for evaluating the net benefit. Although we validated three ML models, other predictive models need to be investigated.

The size zone non-uniformity normalized (SZNUN) feature extracted using GLSZM indicates the variation in volume, and a lower SZNUN indicates greater homogeneity. As shown in Figure 3b, the *glszm_SZNUN* coefficient was high in both intratumoral (*original_glszm_SZNUN*: -0.087) and 3 mm-peritumoral (*peri_log-sigma-1-0-mm-3D_glszm_SZNUN*: 0.131) features. Examples of the feature maps of these features in the test cohort are shown in Figure 6. The EGFR- and EGFR+ groups demonstrated different tendencies in the feature maps of the *original_glszm_SZNUN* and *peri_log-sigma-1-0-mm-3D_glszm_SZNUN*. Biopsy result showing EGFR- can include false negatives because of intratumor heterogeneity.¹⁵ Therefore, though further evaluation

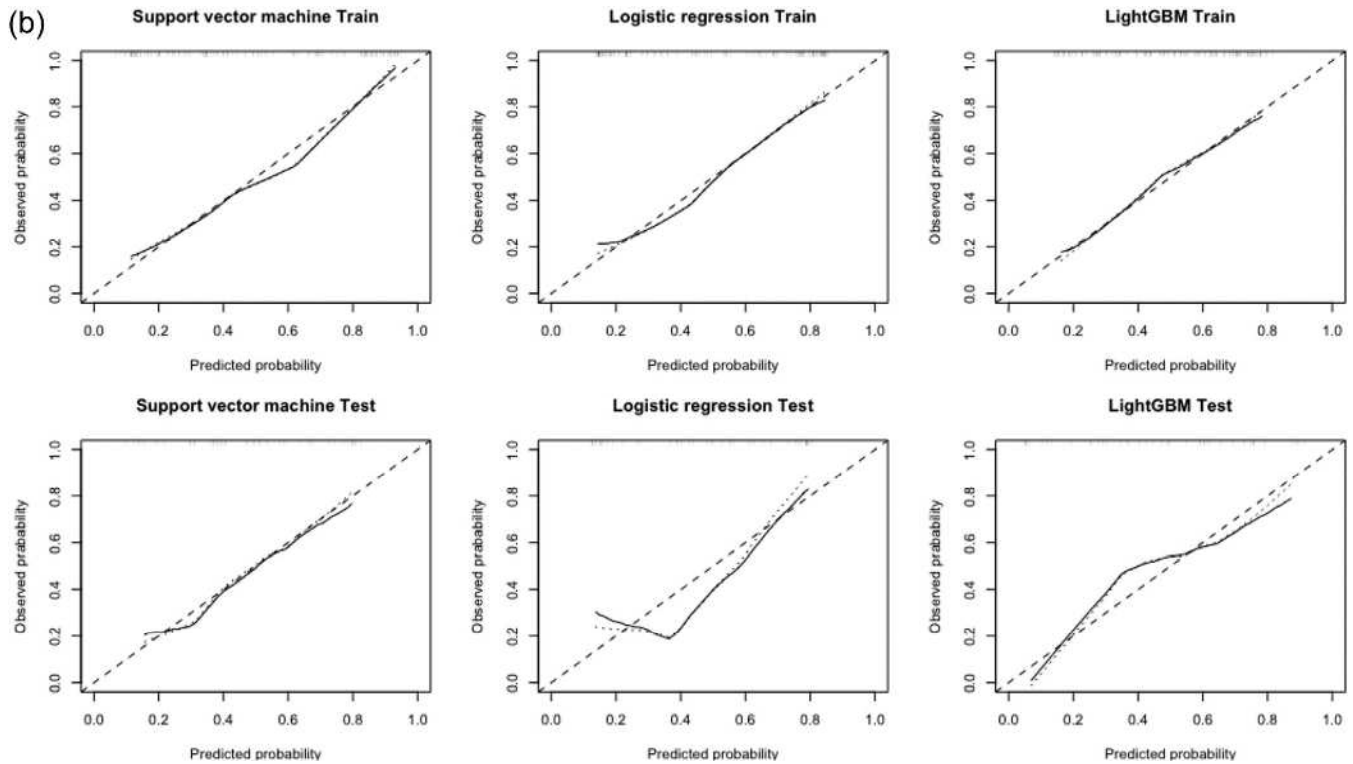


FIGURE 4 Calibration curves for each machine learning model derived using the combination of intratumoral and 3 mm-peritumoral radiomic signatures with clinical features. The dashed line shows the ideal model and the solid line shows actual model.

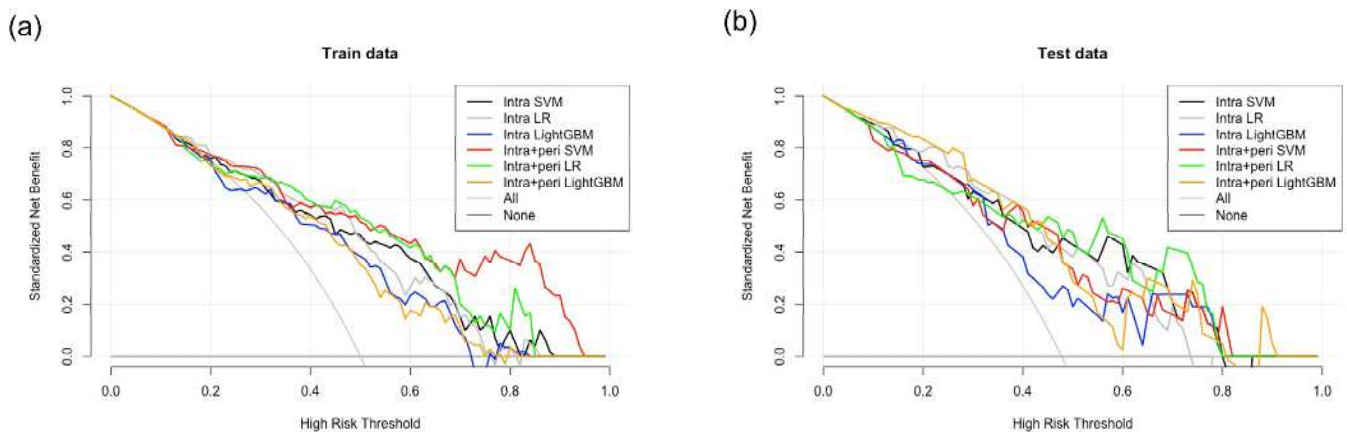


FIGURE 5 Decision curves of machine learning based predictive models in (a) training and (b) test cohorts.

should be performed, the feature map might be helpful for interpreting the heterogeneous areas of the tumor.

An invasive biopsy is required to confirm EGFR mutations patients. Our image-based method for EGFR mutation identification can eliminate this inconvenient procedure for patients and facilitate early decision-making regarding treatment strategies. Most studies for predicting EGFR mutation status focused on intratumoral features alone,^{8,9,13,16} while a few studies focused on peritumoral features. Previous studies used a single peritumoral region to construct a single ML model,^{1,10}

therefore, the robustness of radiomic signatures in different models is unknown. We compared the radiomic features of multiple peritumoral regions and constructed three ML models. LR and LightGBM models derived from IPRS3 showed similar AUCs and were better than those of IRS, indicating that IPRS3 has high robustness.

Our study has some limitations. First, the number of patients included in this study was limited. Therefore, a larger number of cases should be examined to further validate our results. In addition, we did not validate our predictive models with an external dataset; therefore, it

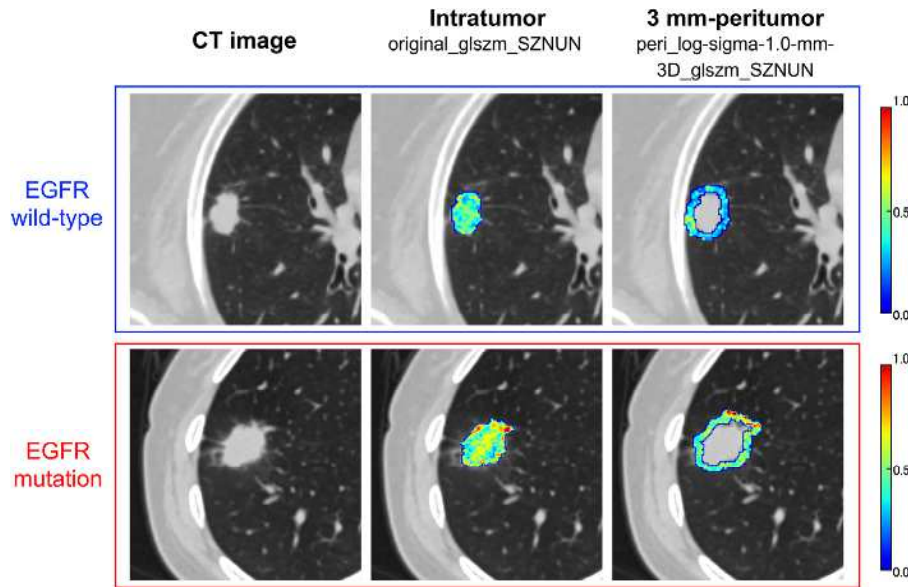


FIGURE 6 Feature maps generated by `glszm_SizeZoneNonUniformityNormalized` (center) and `peri_log-sigma-1-0-mm-3D_glszm_SizeZoneNonUniformityNormalized` (right) in the test cohort with color bar. Feature maps in the EGFR mutation group tended to show high value, whereas those in the wild-type group showed low value both in intratumoral and peritumoral regions.

is necessary to compare them with other models. Second, the variety of peritumoral regions was considered insufficient and multiple peritumoral regions to be evaluated in future works. Third, five different CT scanners were used in this study. The variability in the values of radiomics features from different CT scanners can be comparable to the variability in these features in CT images of NSCLC tumors.³³ Moreover, it is reported that imaging parameters affect the robustness of radiomic features.³⁴ Because a lot of facilities have multiple CT scanners, improving robustness of features by imaging parameters correction is necessary. The accuracy of our predictive model may be improved by imaging parameters correction. Furthermore, Zwanenburg et al. reported that image perturbation may be useful for assessing feature robustness.³⁵ As future works, we will evaluate feature robustness for extracted radiomic features.

5 | CONCLUSIONS

We determined the optimal peritumoral size and investigated radiomic features to construct predictive models for EGFR mutation status. The combination of intratumoral and 3 mm-peritumoral radiomic signatures could identify EGFR mutation status more accurately compared to combinations of 5 or 7 mm-peritumoral radiomic signatures. Furthermore, LR and LightGBM models derived from IPRS3 demonstrated better accuracy in predicting EGFR mutation status than those derived from IRS. Therefore, the combination of intratumoral and 3 mm-peritumoral radiomic signa-

tures can help accurately predict EGFR mutation status.

AUTHOR CONTRIBUTIONS

Yusuke Kawazoe and Takehiro Shiinoki designed this study. Yusuke Kawazoe and Koya Fujimoto carried out the experiment. Yusuke Kawazoe wrote the manuscript with support from Takehiro Shiinoki. Koya Fujimoto, Yuki Yuasa, Tsunahiko Hirano, Kazuto Matsunaga, and Hidekazu Tanaka supplied available data in terms of this study and discussed. All authors discussed the results and contributed to the final manuscript.

ACKNOWLEDGMENTS

The authors thank Medical Physics Research Unit in Yamaguchi University (<https://ds0n.cc.yamaguchi-u.ac.jp/~medphys/>) for providing support to accomplish this study. This study was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI, Grant number 22K07667 (T.S.), and the Takeda Science Foundation.


CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Takehiro Shiinoki 

<https://orcid.org/0000-0002-0246-8489>

REFERENCES

- Choe J, Lee SM, Kim W, et al. CT radiomics-based prediction of anaplastic lymphoma kinase and epidermal growth factor receptor mutations in lung adenocarcinoma. *Eur J Radiol.* 2021;139:109710. doi:10.1016/j.ejrad.2021.109710
- Rizzo S, Petrella F, Buscarino V, 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
- Mitsudomi T, Yatabe Y. Mutations of the epidermal growth factor receptor gene and related genes as determinants of epidermal growth factor receptor tyrosine kinase inhibitors sensitivity in lung cancer. *Cancer Sci.* 2007;98(12):1817-1824. doi:10.1111/j.1349-7006.2007.00607.x
- Jänne PA, Engelman JA, BE Johnson. Epidermal growth factor receptor mutations in non-small-cell lung cancer: implications for treatment and tumor biology. *J Clin Oncol.* 2005;23(14):3227-3234. doi:10.1200/JCO.2005.09.985
- Mitsudomi T, Morita S, Yatabe Y, et al. Gefitinib versus cisplatin plus docetaxel in patients with non-small-cell lung cancer harbouring mutations of the epidermal growth factor receptor (WJTOG3405): an open label, randomised phase 3 trial. *Lancet Oncol.* 2010;11(2):121-128. doi:10.1016/S1470-2045(09)70364-X
- Kriegs M, Gurtner K, Can Y, et al. Radiosensitization of NSCLC cells by EGFR inhibition is the result of an enhanced p53-dependent G1 arrest. *Radiother Oncol.* 2015;115(1):120-127. doi:10.1016/j.radonc.2015.02.018
- Liu Q, Sun D, Li N, et al. Predicting EGFR mutation subtypes in lung adenocarcinoma using 18F-FDG PET/CT radiomic features. *Transl Lung Cancer Res.* 2020;9(3):549-562. doi:10.21037/tlcr.2020.04.17
- Park H, Sholl LM, Hatabu H, Awad MM, Nishino M. Imaging of precision therapy for lung cancer: current state of the art. *Radiology.* 2019;293(1):15-29. doi:10.1148/radiol.2019190173
- Hong D, Xu K, Zhang L, Wan X, Guo Y. Radiomics signature as a predictive factor for EGFR mutations in advanced lung adenocarcinoma. *Front Oncol.* 2020;10:1-8. doi:10.3389/fonc.2020.00028
- Yamazaki M, Yagi T, Tominaga M, Minato K, Ishikawa H. Role of intratumoral and peritumoral CT radiomics for the prediction of EGFR gene mutation in primary lung cancer. *file:///Users/yusukekawazoe/Desktop/reference/Radiomics/peritumor/Peritumoral and intratumoral radiomic features predict survival outcomes amo. Br J Radiol.* 2022;9-12. doi:10.1259/bjr.20220374
- Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer.* 2012;48(4):441-446. doi:10.1016/j.ejca.2011.11.036
- Takehana K, Sakamoto R, Fujimoto K, et al. Peritumoral radiomics features on preoperative thin-slice CT images can predict the spread through air spaces of lung adenocarcinoma. *Sci Rep.* 2022;12(1):1-9. doi:10.1038/s41598-022-14400-w
- Zhuo Y, Feng M, Yang S, et al. Radiomics nomograms of tumors and peritumoral regions for the preoperative prediction of spread through air spaces in lung adenocarcinoma. *Transl Oncol.* 2020;13(10). doi:10.1016/j.tranon.2020.100820
- Liao G, Huang L, Wu S, et al. Preoperative CT-based peritumoral and tumoral radiomic features prediction for tumor spread through air spaces in clinical stage I lung adenocarcinoma. *Lung Cancer.* 2022;163:87-95. doi:10.1016/j.lungcan.2021.11.017
- Wang S, Shi J, Ye Z, et al. Predicting EGFR mutation status in lung adenocarcinoma on computed tomography image using deep learning. *Eur Respir J.* 2019;53(3). doi:10.1183/13993003.00986-2018
- Li S, Ding C, Zhang H, Song J, Wu L. Radiomics for the prediction of EGFR mutation subtypes in non-small cell lung cancer. *Med Phys.* 2019;46(10):4545-4552. doi:10.1002/mp.13747
- Zhao W, Wu Y, Xu Y, et al. The potential of radiomics nomogram in non-invasively prediction of epidermal growth factor receptor mutation status and subtypes in lung adenocarcinoma. *Front Oncol.* 2020;9:1485. doi:10.3389/fonc.2019.01485
- Velazquez ER, Parmar C, Jermoumi M, et al. Volumetric CT-based segmentation of NSCLC using 3D-slicer. *Sci Rep.* 2013;3:1-7. doi:10.1038/srep03529
- Shiinoki T, Fujimoto K, Kawazoe Y, et al. Predicting programmed death-ligand 1 expression level in non-small cell lung cancer using a combination of peritumoral and intratumoral radiomic features on computed tomography. *Biomed Phys Eng Express.* 2022;8(2):25008. doi:10.1088/2057-1976/ac4d43
- Van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* 2017;77(21):e104-e107. doi:10.1158/0008-5472.CAN-17-0339
- Morgado J, Pereira T, Silva F, et al. Machine learning and feature selection methods for egfr mutation status prediction in lung cancer. *Appl Sci.* 2021;11(7). doi:10.3390/app11073273
- Tixier F, Le Rest CC, Hatt M, et al. Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. *J Nucl Med.* 2011;52(3):369-378. doi:10.2967/jnumed.110.082404
- Xiong Z, Jiang Y, Tian D, et al. Radiomics for identifying lung adenocarcinomas with predominant lepidic growth manifesting as large pure ground-glass nodules on CT images. *PLoS One.* 2022;17:1-15. doi:10.1371/journal.pone.0269356
- Li S, Luo T, Ding C, Huang Q, Guan Z, Zhang H. Detailed identification of epidermal growth factor receptor mutations in lung adenocarcinoma: combining radiomics with machine learning. *Med Phys.* 2020;47(8):3458-3466. doi:10.1002/mp.14238
- Liu Y, Zhou J, Wu J, et al. Development and validation of machine learning models to predict epidermal growth factor receptor mutation in non-small cell lung cancer: a multi-center retrospective radiomics study. *Cancer Control.* 2022;29(168):1-8. doi:10.1177/10732748221092926
- Gao J, Chen X, Li X, et al. Differentiating TP53 mutation status in pancreatic ductal adenocarcinoma using multiparametric MRI-derived radiomics. *Front Oncol.* 2021;11(May):1-8. doi:10.3389/fonc.2021.632130
- Ji GW, Zhu FP, Xu Q, et al. Radiomic features at contrast-enhanced CT predict recurrence in early stage hepatocellular carcinoma: a multi-institutional study. *Radiology.* 2020;294(2):568-579. doi:10.1148/radiol.2020191470
- Zhang G, Cao Y, Zhang J, et al. Predicting EGFR mutation status in lung adenocarcinoma: development and validation of a computed tomography-based radiomics signature. *Am J Cancer Res.* 2021;11(2):546-560. <http://www.ncbi.nlm.nih.gov/pubmed/33575086%0A.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC7868761>
- Pérez-Morales J, Tunali I, Stringfield O, et al. Peritumoral and intratumoral radiomic features predict survival outcomes among patients diagnosed in lung cancer screening. *Sci Rep.* 2020;10(1):1-15. doi:10.1038/s41598-020-67378-8
- Mei D, Luo Y, Wang Y, Gong J. CT texture analysis of lung adenocarcinoma: can Radiomic features be surrogate biomarkers for EGFR mutation statuses. *Cancer Imaging.* 2018;18(1):1-9. doi:10.1186/s40644-018-0184-2
- Gu D, Hu Y, Ding H, et al. CT radiomics may predict the grade of pancreatic neuroendocrine tumors: a multicenter study. *Eur Radiol.* 2019;29(12):6880-6890. doi:10.1007/s00330-019-06176-x
- Zhang B, Liu Q, Zhang X, et al. Clinical utility of a nomogram for predicting 30-days poor outcome in hospitalized patients with

- COVID-19: multicenter external validation and decision curve analysis. *Front Med.* 2020;7:1-12. doi:10.3389/fmed.2020.590460
33. Mackin D, Fave X, Zhang L, et al. Measuring computed tomography scanner variability of radiomics features. *Invest Radiol.* 2015;50(11):757-765. doi:10.1097/RLI.000000000000180
34. Reiazi R, Abbas E, Famiyeh P, et al. The impact of the variation of imaging parameters on the robustness of Computed Tomography radiomic features: a review. *Comput Biol Med.* 2021;133:104400. doi:10.1016/j.combiomed.2021.104400
35. Zwanenburg A, Leger S, Agolli L, et al. Assessing robustness of radiomic features by image perturbation. *Sci Rep.* 2019;9(1):1-10. doi:10.1038/s41598-018-36938-4

How to cite this article: Kawazoe Y, Shiinoki T, Fujimoto K, et al. Investigation of the combination of intratumoral and peritumoral radiomic signatures for predicting epidermal growth factor receptor mutation in lung adenocarcinoma. *J Appl Clin Med Phys.* 2023;e13980. <https://doi.org/10.1002/acm2.13980>