


# A CT-based radiomics model to predict subsequent brain metastasis in patients with ALK-rearranged non-small cell lung cancer undergoing crizotinib treatment

Yongluo Jiang<sup>1,2,3</sup>  | Yixing Wang<sup>1,2,4</sup> | Sha Fu<sup>5</sup> | Tao Chen<sup>1,2,3</sup> | Yixin Zhou<sup>1,2,6</sup> | Xuanye Zhang<sup>1,2,4</sup> | Chen Chen<sup>1,2,7</sup> | Li-na He<sup>1,2,4</sup> | Wei Du<sup>1,2,4</sup> | Haifeng Li<sup>1,2,4</sup> | Zuan Lin<sup>1,2,8</sup> | Yuanyuan Zhao<sup>1,2,4</sup> | Yunpeng Yang<sup>1,2,4</sup> | Hongyun Zhao<sup>1,2,8</sup> | Wenfeng Fang<sup>1,2,4</sup> | Yan Huang<sup>1,2,4</sup> | Shaodong Hong<sup>1,2,4</sup> | Li Zhang<sup>1,2,4</sup>

<sup>1</sup>State Key Laboratory of Oncology in South China, Guangzhou, China

<sup>2</sup>Collaborative Innovation Center for Cancer Medicine, Guangzhou, China

<sup>3</sup>Department of Nuclear Medicine, Sun Yat-sen University Cancer Center, Guangzhou, China

<sup>4</sup>Department of Medical Oncology, Sun Yat-sen University Cancer Center, Guangzhou, China

<sup>5</sup>Cellular & Molecular Diagnostics Center, Guangdong Provincial Key Laboratory of Malignant Tumor Epigenetics and Gene Regulation, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China

<sup>6</sup>Department of VIP region, Sun Yat-sen University Cancer Center, Guangzhou, China

<sup>7</sup>Department of Radiation Oncology, Sun Yat-sen University Cancer Center, Guangzhou, China

<sup>8</sup>Department of Clinical Research, Sun Yat-sen University Cancer Center, Guangzhou, China

## Correspondence

Shaodong Hong, Sun Yat-sen University Cancer Center, 651 Dongfeng East Road, Guangzhou, Guangdong 510060, China  
Email: hongshd@sysucc.org.cn

Li Zhang, Sun Yat-sen University Cancer Center, 651 Dongfeng East Road, Guangzhou, Guangdong 510060, China  
Email: zhangli6@mail.sysu.edu.cn

## Funding information

This study was supported by the National Natural Science Foundation of China (81972898, 82172713, 81872499), the Natural Science Foundation of Guangdong Province (2019A1515011090), and the Outstanding Young Talents Program of Sun Yat-sen University Cancer Center (16zxc04).

## Abstract

**Background:** Brain metastasis (BM) comprises the most common reason for crizotinib failure in patients with anaplastic lymphoma kinase (ALK)-rearranged non-small cell lung cancer (NSCLC). We hypothesize that its occurrence could be predicted by a computed tomography (CT)-based radiomics model, therefore, allowing for selection of enriched patient populations for prevention therapies.

**Methods:** A total of 75 eligible patients were enrolled from Sun Yat-sen University Cancer Center between June 2014 and September 2019. The primary endpoint was brain metastasis-free survival (BMFS), estimated from the initiation of crizotinib to the date of the occurrence of BM. Patients were randomly divided into two cohorts for model training ( $n = 51$ ) and validation ( $n = 24$ ), respectively. A radiomics signature was constructed based on features extracted from chest CT before crizotinib treatment. Clinical model was developed using the Cox proportional hazards model. Log-rank test was performed to describe the difference of BMFS risk.

**Results:** Patients with low radiomics score had significantly longer BMFS than those with higher, both in the training cohort ( $p = 0.019$ ) and validation cohort ( $p = 0.048$ ). The nomogram combining smoking history and the radiomics signature showed good performance for the estimation of BMFS, both in the training (concordance index [C-index], 0.762; 95% confidence interval [CI], 0.663–0.861) and validation cohort (C-index, 0.724; 95% CI, 0.601–0.847).

**Conclusion:** We have developed a CT-based radiomics model to predict subsequent BM in patients with non-brain metastatic NSCLC undergoing crizotinib treatment. Selection of an enriched patient population at high BM risk will facilitate the design of clinical trials or strategies to prevent BM.

## KEYWORDS

ALK-positive, image biomarkers, lung cancer, response prediction, targeted therapy

Yongluo Jiang, Yixing Wang, Sha Fu, and Tao Chen have contributed equally to this work.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *Thoracic Cancer* published by China Lung Oncology Group and John Wiley & Sons Australia, Ltd.

## BACKGROUND

Lung cancer is one of the most common and lethal malignancies worldwide.<sup>1,2</sup> Among lung cancers, non-small cell lung cancer (NSCLC) is the most common histology. More than 60% of lung adenocarcinoma harbor mutations in specific driven genes, including epidermal growth factor receptor (EGFR), anaplastic lymphoma kinase (ALK), ROS proto-oncogene 1 receptor tyrosine kinase (ROS1), and so on.<sup>3</sup> The application of targeted therapies that block corresponding driver genes has significantly improved the survival of lung cancer patients.<sup>4</sup>

Approximately 5% of NSCLC has ALK rearrangement.<sup>5</sup> For patients with ALK-rearranged NSCLC, crizotinib is the first approved targeted agent that demonstrated survival benefit.<sup>6</sup> However, patients treated with crizotinib ultimately experience disease progression, with brain metastasis (BM) being the most common reason of treatment failure.<sup>7,8</sup> BM is a major cause of mortality among patients with NSCLC, with <5% of patients surviving over 5 years after BM diagnosis.<sup>9</sup> For patients with BM, psychological and physical impairments are common complications following brain-directed therapies and symptomatic therapies. Therefore, BM remains a critical challenge for maintaining patients' quality of life and prolonging patients' survival. On the other side, controversies still exist in terms of the frequency of BM screening in asymptomatic patients with ALK-rearranged lung cancer who were free of BM at baseline. Hence, it is clinically relevant to identify those who are more likely or less likely to develop BM to tailor patient's management.

Currently, the occurrence of BM from ALK-rearranged NSCLC cannot be accurately predicted in patients with non-brain metastatic disease undergoing crizotinib treatment. We hypothesized that a nomogram could be constructed by combining computed tomography (CT)-based radiomics signatures and selected clinical variables using a multivariate model to predict the likelihood of occurrence of BM. Radiomics is an attractive method for comprehensive assessment of the tumor and its microenvironment in entirety that can reflect tumor phenotypes, genotypes, microenvironment, and other biological features.<sup>10–12</sup> A number of studies have found that radiomics is a reliable tool in cancer diagnosis and prognosis prediction.<sup>13,14</sup> We, therefore, constructed and validated such a nomogram using retrospectively reviewed data from our institute. This prediction model could be used to develop radiologic screening or preventive treatment strategies for patients with NSCLC with non-brain metastatic disease treated with crizotinib.

## METHODS

### Patients

We searched the electronic Lung Tumor Database of the Sun Yat-sen University Cancer Center (Guangzhou, China) and retrospectively reviewed records for patients treated with crizotinib between June 2014 and September 2019. A

total of 75 consecutive patients were enrolled according to the inclusion and exclusion criteria as listed below.

The inclusion criteria: (1) patients pathologically diagnosed as NSCLC; (2) the tumor had ALK rearrangement as determined by established methods including next-generation sequencing (NGS), fluorescent in situ hybridization (FISH), immunohistochemistry (IHC), or reverse transcription-polymerase chain reaction (RT-PCR); (3) had available CT scan before crizotinib treatment that could provide lung tumor images for data mining; (4) without BM at the initiation of crizotinib as per brain magnetic resonance imaging (MRI); and (5) had subsequent brain MRI monitoring.

The exclusion criteria: (1) had a life-threatening medical disorder; and (2) had other invasive malignant disease within the past 5 years, except basal-cell skin carcinoma and cervical carcinoma in situ.

The particular main lung lesions were selected on CT images according to diagnostic reports, and pathological reports or follow-up data. The primary endpoint was brain metastasis-free survival (BMFS), which was defined as the time from the initiation of crizotinib to the date of the occurrence of BM. Patients were divided randomly into training cohort and validation cohorts by random numbers. The study flow diagram is shown in Figure 1.

The study was approved by the Ethics Committee of Sun Yat-sen University Cancer Center, and the informed consent requirement was waved because of the retrospective nature.

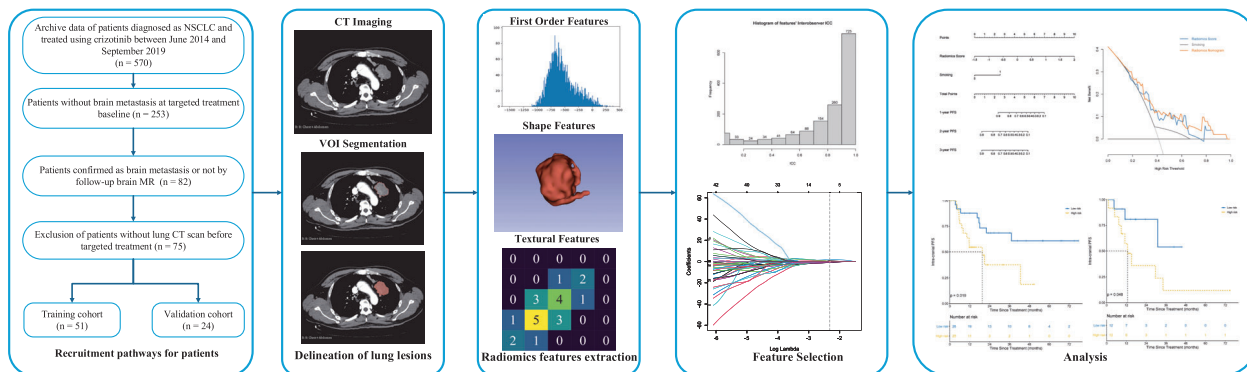
### CT image acquisition

All chest CT scans were performed by using the following scanners: performed with a 64-, 128-, or 256-detector row scanner CT machine (Aquilion TSX-101A, Toshiba Medical System; Discovery CT750 HD, GE System; and Brilliance iCT, Philips System) or a dual-source spiral CT system (SOMATOM Force, Siemens Medical System).

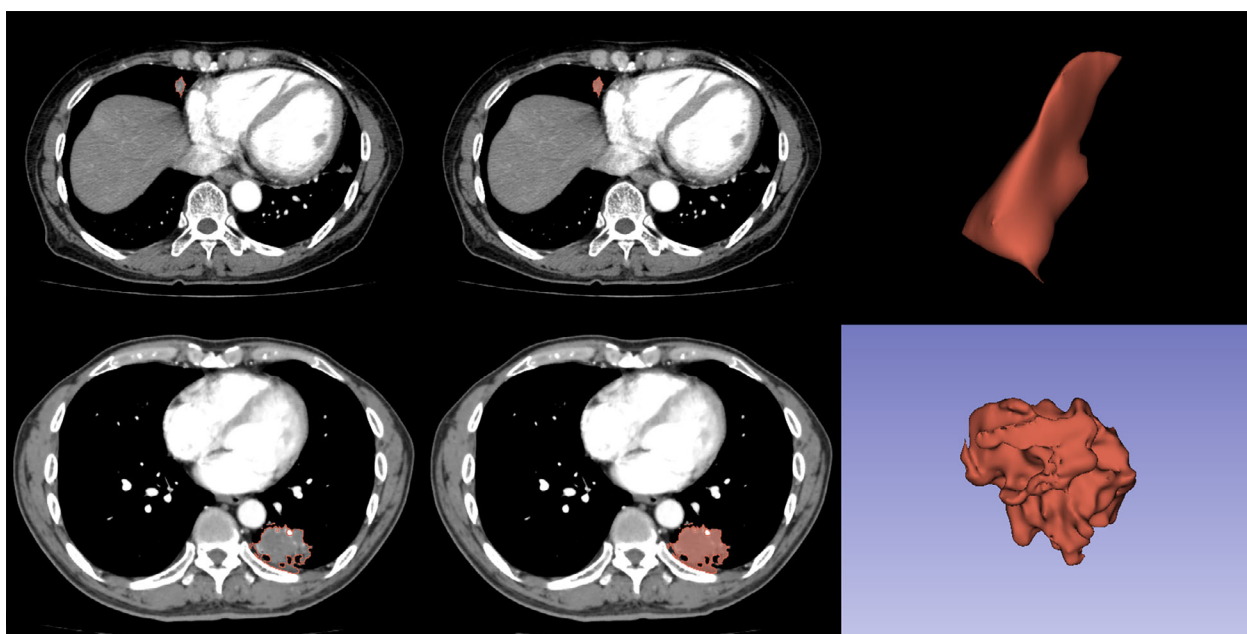
The scanning parameters were routinely set regardless of the type of devices as follows: tube voltage, 80–140 kVp; tube current, automatic tube current modulation (maximum, 450 mAs); pitch factor, 1.0; slice thickness, 1, 1.25, or 2 mm; and slice interval, 1 mm. Images were evaluated on a separate workstation using both mediastinal and lung visualization windows. Contrast-enhanced CT images were acquired after an intravenous bolus dose (1.5–2 mL/kg body weight) of non-ionic iodinated contrast agent (Ultravist 300; Bayer Healthcare Company), which was administered into the antecubital vein at a rate of 3.0 mL/s via a high-pressure syringe. All the CT images were reconstructed with a standard kernel and were retrieved from the picture archiving and communication system (PACS).

### Volume-of-interest segmentation

The pathological results and clinical information such as age and sex were referred and selected from our hospital's



**FIGURE 1** Workflow of necessary steps in current study. Patients were enrolled as the recruitment pathway. Tumors are semi-automatically segmented on axial arterial phase CT. Radiomic features are extracted from corresponding VOIs on CT images. Feature was selected by inter- and intra-observer reliability assessment and subsequent least absolute shrinkage and selection operator (LASSO) method. The radiomics signature is constructed by a linear combination of selected features. The performance of the prediction model is assessed by C-index and the calibration curve. A nomogram is built for individualized assessment, decision curve analysis, and survival prediction are then performed



**FIGURE 2** Representative clinical cases and related thoracic CT. This set of images showed the ability of brain metastasis risk stratification of the radiomics score. The upper images demonstrated a particular patient (patient Z) with brain metastasis 21.03 months after crizotinib treatment, whereas the lower images demonstrated one (patient J) without brain metastasis during 40.94 months' follow-up, both patients did not smoke. The selected lesions were marked by red and demonstrated with 3D visualization. Obviously, patient J's lesion was much larger than that of patient Z's, but brain metastasis occurred in patient Z. Volume of lesion did not work. The radiomics scores of these two patients were  $-0.368$  and  $-0.554$ , respectively. Patient Z was of high risk and patient J was of low risk, according to the cut-off point of  $-0.512$

medical records. Two radiologists experienced in lung CT interpretation reviewed all CT scans. The boundary between the chest and other tissues is mainly identified in the mediastinal window, and the mediastinal window is used for contouring (Figure 2).

The CT images were imported into the open-source software 3D Slicer (version 4.10.2, Brigham and Women's Hospital) using Digital Imaging and Communications in Medicine (DICOM) format for delineation. The delineation was performed meticulously to exclude the visible surrounding large vessels and bronchioles from the volume of the lesions as much as possible. The regions

of interest (ROIs) were delineated slice-by-slice semi-automatically on enhanced CT images by experienced radiologists, subsequently confirmed by a senior radiologist, and the corresponding ROIs were stacked up to construct volumes of interest (VOIs) of the primary lung lesions.

### Feature extraction and feature selection

Texture analysis was performed at the same time using the PyRadiomics<sup>15</sup> platform implanted in the 3D Slicer

**TABLE 1** Characteristics of patients in the training and validation cohort

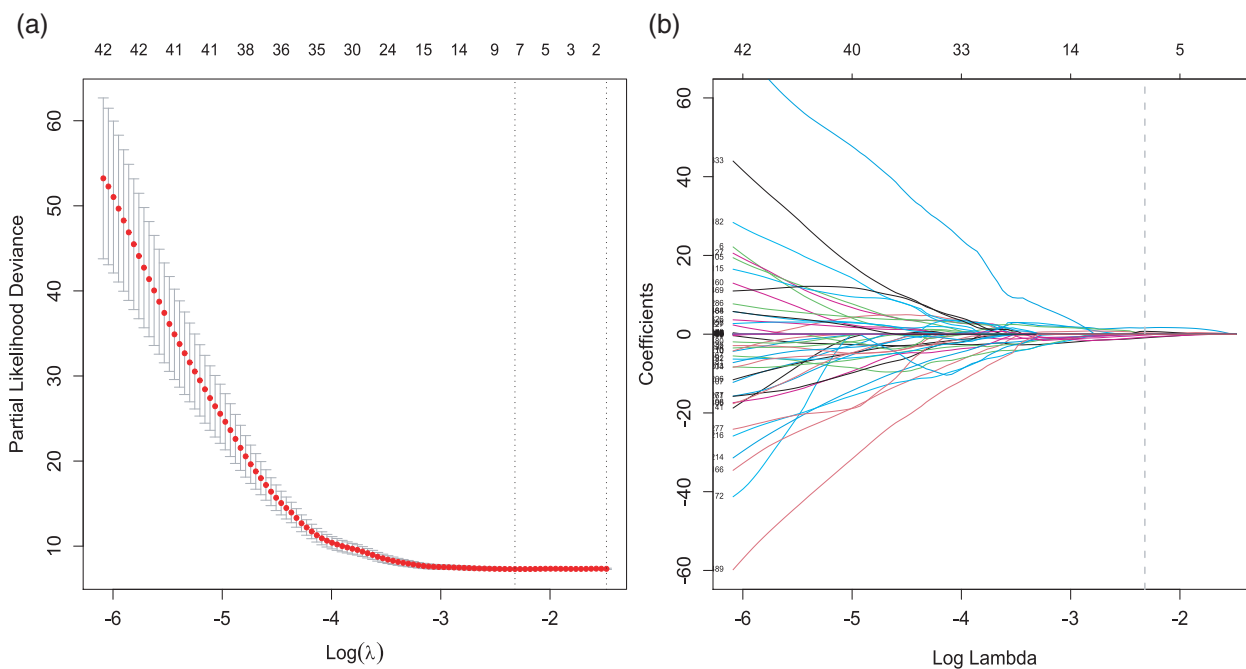
| Variable                                | Overall, <i>n</i> = 75 | Training cohort, <i>n</i> = 51 | Validation cohort, <i>n</i> = 24 |
|---|------------------------|--------------------------------|----------------------------------|
| Age                                     | 48 (29–76)             | 48 (29–76)                     | 49 (31–70)                       |
| Sex                                     |                        |                                |                                  |
| Male                                    | 36 (48%)               | 26 (51%)                       | 10 (42%)                         |
| Female                                  | 39 (52%)               | 25 (49%)                       | 14 (58%)                         |
| Smoking                                 | 8 (11%)                | 6 (12%)                        | 2 (8.3%)                         |
| ECOG PS                                 |                        |                                |                                  |
| 0                                       | 22                     | 18 (35%)                       | 7 (29%)                          |
| 1                                       | 49                     | 30 (59%)                       | 16 (67%)                         |
| 2                                       | 4                      | 3 (6%)                         | 1 (4)                            |
| Stage                                   |                        |                                |                                  |
| IIIB-IIIC/recurrence                    | 21 (28%)               | 15 (29%)                       | 6 (25%)                          |
| IV                                      | 54 (72%)               | 36 (71%)                       | 18 (75%)                         |
| Intrapulmonary metastasis               |                        |                                |                                  |
| Yes                                     | 35 (47%)               | 26 (51%)                       | 9 (38%)                          |
| No                                      | 40 (53%)               | 25 (49%)                       | 15 (62%)                         |
| Liver metastasis                        |                        |                                |                                  |
| Yes                                     | 6 (8%)                 | 5 (10%)                        | 1 (4%)                           |
| No                                      | 69 (92%)               | 46 (90%)                       | 23 (96%)                         |
| Bone metastasis                         |                        |                                |                                  |
| Yes                                     | 15 (20%)               | 7 (14%)                        | 8 (33%)                          |
| No                                      | 60 (80%)               | 44 (86%)                       | 16 (67%)                         |
| No. of metastatic sites                 |                        |                                |                                  |
| 0                                       | 21 (28%)               | 15 (29%)                       | 6 (25%)                          |
| 1                                       | 49 (65%)               | 32 (63%)                       | 17 (71%)                         |
| ≥2                                      | 5 (7%)                 | 4 (8%)                         | 1 (4%)                           |
| Treatment line                          |                        |                                |                                  |
| 1st                                     | 59 (79%)               | 41 (81%)                       | 18 (76%)                         |
| 2nd                                     | 7 (9%)                 | 4 (8%)                         | 3 (12%)                          |
| ≥3rd                                    | 9 (12%)                | 6 (11%)                        | 3 (12%)                          |
| Tumor CT Volume                         |                        |                                |                                  |
| Low volume (≤5643.75 mm <sup>3</sup> )  | 38 (51%)               | 25 (49%)                       | 13 (54%)                         |
| High volume (>5643.75 mm <sup>3</sup> ) | 37 (49%)               | 26 (51%)                       | 11 (46%)                         |
| Median BMFS (range)                     | 14 (8–30)              | 17 (7–31)                      | 13 (9–25)                        |
| BMFS status                             |                        |                                |                                  |
| Non-progression                         | 42 (56%)               | 30 (59%)                       | 12 (50%)                         |
| Progression                             | 33 (44%)               | 21 (41%)                       | 12 (50%)                         |
| Radiomics score                         | −0.51 (−0.68 to 0.10)  | −0.51 (−0.66 to 0.12)          | −0.52 (−0.76 to 0.10)            |

Abbreviations: CT, computed tomography; BMFS, brain metastasis-free survival; ECOG PS, Eastern Cooperative Oncology Group performance status

software. The code is open-source at <https://github.com/radiomics/pyradiomics>. A total of 1561 features were extracted. A total of 35 patients were randomly selected for VOI segmentation and feature extraction again by the same radiologist and a different one. The intraclass correlation coefficient (ICC) was used to assess the intra- and interobserver reproducibility of radiomics feature extraction. Features of ICC over 0.8 were selected for further feature selection.

### Prognosis-related feature selection and radiomics signature construction

Leave-one-out cross-validation was applied in the training cohort. The least absolute shrinkage and selection operator (LASSO) cox regression algorithm, which is suitable for the regression of high dimensional data, was applied in training dataset to select the most predictive features with nonzero coefficients from among the selected texture features with



**FIGURE 3** Feature selections using the LASSO Cox model. (a) The LASSO logistic regression model was used with tuning parameter ( $\lambda$ ) conducted by leave-one-out cross-validation based on the minimum criteria. The deviance was plotted versus  $\log(\lambda)$ . Dotted vertical lines were drawn at the optimal values, using the minimum criteria and the 1 standard error of the minimum criteria (the 1—standard error criteria). The optimal tuning parameter  $\lambda$  value of 0.098 with  $\log(\lambda) = -2.321$  was selected. (b) LASSO coefficient profiles of the 489 texture features. The dotted vertical line was plotted at the value selected using leave-one-out cross-validation in plot (a). The eight resulting features with nonzero coefficients are indicated in the plot

good reproducibility. A linear combination of the selected features weighted by their respective coefficients educed a radiomics score for each patient, reflecting the risk of BM.

### Clinical characteristics selection

The impact of potential risk factors on BMFS was evaluated using univariable and multivariable Cox proportional hazard regression analysis. Factors considered in the model included age, sex, smoking history, Eastern Cooperative Oncology Group performance status (ECOG PS), metastasis sites, treatment line, inflammatory markers, and tumor volume. For multivariable analysis, a variable selection method with selection criteria of  $p \leq 0.1$  was considered, and interactions among factors significant at 0.05 were included.

### Nomogram building, calibration, and validation

A radiomics nomogram was built by a multivariate Cox regression model to predict BMFS and tested in the validation cohort.

The predictive accuracy of the radiomics nomogram was quantified by concordance index (C-index) in both the training and validation cohort.

### Clinical use of the radiomics nomogram

Decision curve analysis (DCA) was performed by calculating the net benefits of different threshold probabilities in the model

and validation cohort combined, thereby estimating the clinical use of the nomogram. Net benefit can be calculated as net benefit = true positive rate - (false positive rate  $\times \frac{Pt}{1-Pt}$ ). Where Pt is the threshold probability given by predict model.

### Statistical analysis

A dedicated statistician using R statistical software (version 4.0.3; R Foundation for Statistical Computing) performed all statistical tests. The differences in age, sex, stage, smoking history, ECOG PS, metastatic sites, treatment line, and tumor volume between the training and validation data sets were assessed by using an independent samples  $\chi^2$  test or Wilcoxon test, where appropriate. The “glmnet” package was used to perform the LASSO Cox regression model analysis. Cox regression analyses were performed to find prognostic clinical and radiomics feature(s) for BM during follow-up and log-rank test was performed to describe the difference between two risk groups. Nomogram construction was performed using the “rms” package. DCA was performed using the “rmda” package. A two-sided  $p < 0.05$  was considered significant.

## RESULTS

### Patient characteristics

The baseline characteristics of all the 75 eligible patients are shown in Table 1. Among these patients, 36 were males; median age was 48 years (range, 29–76 years old); all had

TABLE 2 Correlation between characteristics of patients and radiomics score

| Variable                                | High radiomics score, <i>n</i> = 38 | Low radiomics score, <i>n</i> = 37 | <i>p</i> -value    |
|---|-------------------------------------|------------------------------------|--------------------|
| Age                                     |                                     |                                    | >0.999             |
| ≤60                                     | 33 (87%)                            | 32 (86%)                           |                    |
| >60                                     | 5 (13%)                             | 5 (14%)                            |                    |
| Sex                                     |                                     |                                    | >0.999             |
| Male                                    | 18 (47%)                            | 18 (49%)                           |                    |
| Female                                  | 20 (53%)                            | 19 (51%)                           |                    |
| Smoking                                 |                                     |                                    | >0.999             |
| Yes                                     | 4 (11%)                             | 4 (11%)                            |                    |
| No                                      | 34 (89%)                            | 33 (89%)                           |                    |
| ECOG PS                                 |                                     |                                    | 0.987              |
| 0                                       | 13 (34%)                            | 12 (32%)                           |                    |
| 1                                       | 23 (61%)                            | 23 (62%)                           |                    |
| 2                                       | 2 (5.3%)                            | 2 (5.4%)                           |                    |
| Stage                                   |                                     |                                    | 0.045 <sup>a</sup> |
| IIIB-IIIIC/recurrence                   | 2 (5.3%)                            | 9 (24%)                            |                    |
| IV                                      | 36 (95%)                            | 28 (76%)                           |                    |
| Intrapulmonary metastasis               |                                     |                                    | 0.904              |
| Yes                                     | 19 (50%)                            | 17 (46%)                           |                    |
| No                                      | 19 (50%)                            | 20 (54%)                           |                    |
| Liver metastasis                        |                                     |                                    | >0.999             |
| Yes                                     | 3 (7.9%)                            | 3 (8.1%)                           |                    |
| No                                      | 35 (92%)                            | 34 (92%)                           |                    |
| Bone metastasis                         |                                     |                                    | 0.603              |
| Yes                                     | 9 (24%)                             | 6 (16%)                            |                    |
| No                                      | 29 (76%)                            | 31 (84%)                           |                    |
| No. of metastatic sites                 |                                     |                                    | 0.305              |
| 0                                       | 8 (21%)                             | 13 (35%)                           |                    |
| 1                                       | 28 (74%)                            | 21 (57%)                           |                    |
| ≥2                                      | 2 (5.3%)                            | 3 (8.1%)                           |                    |
| Treatment line                          |                                     |                                    | 0.260              |
| 1st                                     | 27 (71%)                            | 32 (86%)                           |                    |
| 2nd                                     | 5 (13%)                             | 2 (5.4%)                           |                    |
| ≥3rd                                    | 6 (16%)                             | 3 (8.1%)                           |                    |
| Volume                                  |                                     |                                    | 0.002 <sup>a</sup> |
| Low volume (≤5643.75 mm <sup>3</sup> )  | 12 (32%)                            | 26 (70%)                           |                    |
| High volume (>5643.75 mm <sup>3</sup> ) | 26 (68%)                            | 11 (30%)                           |                    |

<sup>a</sup>Statistically significant.

Abbreviations: ECOG PS, Eastern Cooperative Oncology Group performance status.

histology of lung adenocarcinoma; most patients had metastatic disease (72%) and most patients received crizotinib as first-line treatment (79%). All the included patients were randomly divided into training cohort (*n* = 51) and validation cohort (*n* = 24). The characteristics were well balanced between the training and the validation cohorts. During the follow-up, 33 patients (44.0%) eventually had radiological evidence of brain metastases. The median BMFS was 14.39 months (95% confidence interval [CI], 16.92–25.84) in the entire cohort, 16.56 months (95% CI, 16.77–28.32) in

the testing cohort, and 13.21 months (95% CI, 11.80–25.99) in the validation cohort, respectively.

### Feature selection and radiomics signature construction

Totally 489 radiomics features with high reproducibility was selected and demonstrated with histograms (Figure S1). All outcomes were based on the measurements of the first radiologist.

TABLE 3 Differences of characteristics between BM group and non-BM group

| Variable                                | Non-BM group, <i>n</i> = 42 | BM group, <i>n</i> = 33 | <i>p</i> -value |
|---|-----------------------------|-------------------------|-----------------|
| Age                                     |                             |                         | >0.999          |
| ≤60                                     | 36 (86%)                    | 29 (88%)                |                 |
| >60                                     | 6 (14%)                     | 4 (12%)                 |                 |
| Sex                                     |                             |                         | 0.759           |
| Female                                  | 23 (55%)                    | 16 (48%)                |                 |
| Male                                    | 19 (45%)                    | 17 (52%)                |                 |
| Smoking                                 |                             |                         | 0.136           |
| Yes                                     | 2 (4.8%)                    | 6 (18%)                 |                 |
| No                                      | 40 (95%)                    | 27 (82%)                |                 |
| ECOG PS                                 |                             |                         | 0.193           |
| 0                                       | 17 (40%)                    | 8 (24%)                 |                 |
| 1                                       | 22 (52%)                    | 24 (73%)                |                 |
| 2                                       | 3 (7.1%)                    | 1 (3.0%)                |                 |
| Stage                                   |                             |                         | 0.378           |
| IIIB-IIIC/recurrence                    | 8 (19%)                     | 3 (9.1%)                |                 |
| IV                                      | 34 (81%)                    | 30 (91%)                |                 |
| Intrapulmonary metastasis               |                             |                         | 0.874           |
| Yes                                     | 21 (50%)                    | 15 (45%)                |                 |
| No                                      | 21 (50%)                    | 18 (55%)                |                 |
| Liver metastasis                        |                             |                         | 0.904           |
| Yes                                     | 4 (9.5%)                    | 2 (6.1%)                |                 |
| No                                      | 38 (90%)                    | 31 (94%)                |                 |
| Bone metastasis                         |                             |                         | 0.954           |
| Yes                                     | 9 (21%)                     | 6 (18%)                 |                 |
| No                                      | 33 (79%)                    | 27 (82%)                |                 |
| Treatment line                          |                             |                         | 0.781           |
| 1st                                     | 32 (76%)                    | 27 (82%)                |                 |
| 2nd                                     | 4 (9.5%)                    | 3 (9.1%)                |                 |
| ≥3rd                                    | 6 (14%)                     | 3 (9.1%)                |                 |
| No. of metastatic sites                 |                             |                         | 0.918           |
| 0                                       | 11 (26%)                    | 10 (30%)                |                 |
| 1                                       | 28 (67%)                    | 21 (64%)                |                 |
| ≥2                                      | 3 (7.1%)                    | 2 (6.1%)                |                 |
| Volume                                  |                             |                         | 0.918           |
| Low volume (≤5643.75 mm <sup>3</sup> )  | 22 (52%)                    | 16 (48%)                |                 |
| High volume (>5643.75 mm <sup>3</sup> ) | 20 (48%)                    | 17 (52%)                |                 |

Abbreviations: BM, brain metastasis; ECOG PS, Eastern Cooperative Oncology Group performance status.

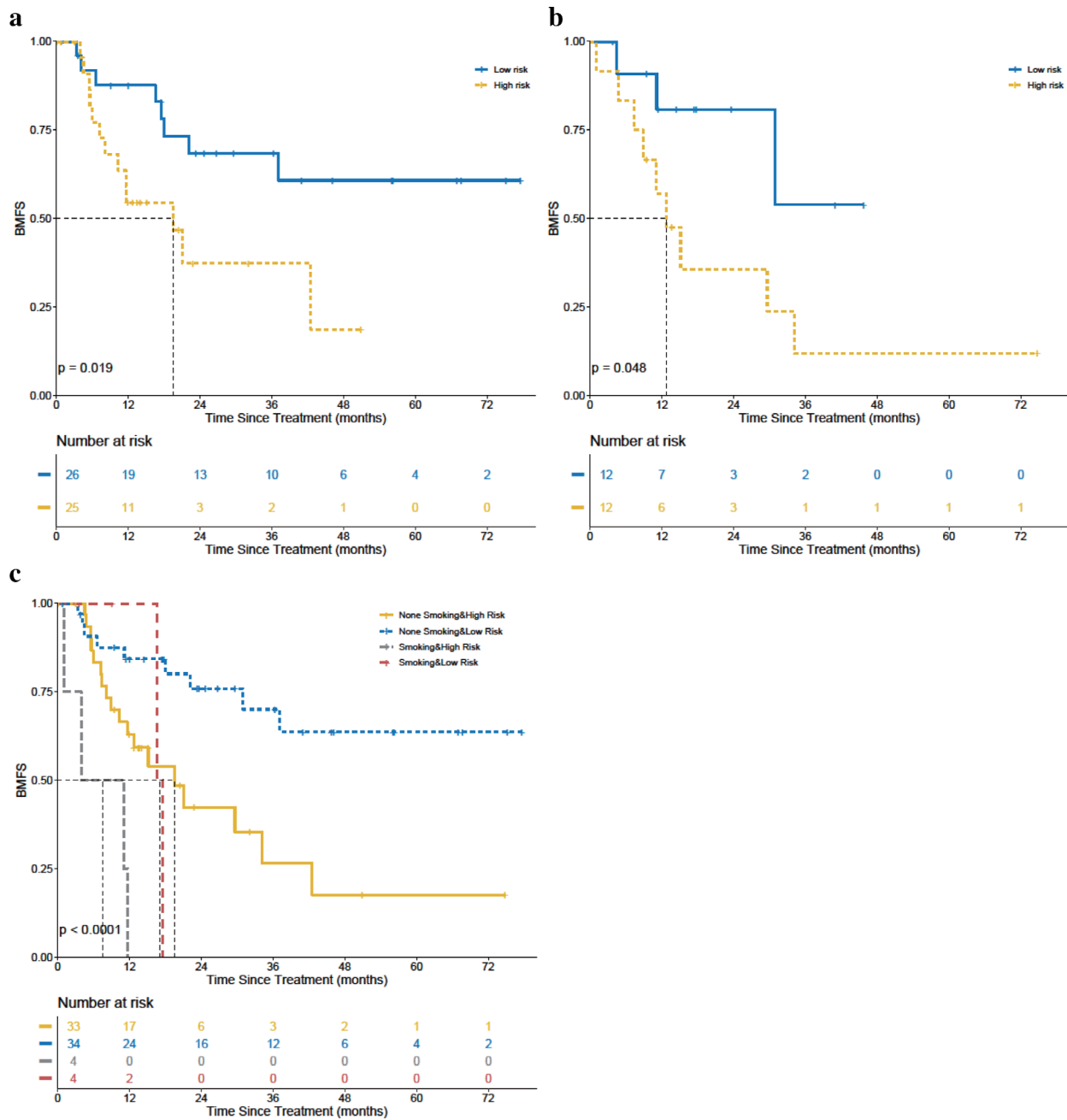
Eight texture features with nonzero coefficient in the LASSO Cox regression models were as follows (Figure 3):

original\_shape\_Elongation;  
 log.sigma.2.0.mm.3D\_glrmlm\_GrayLevelNonUniformity  
 Normalized;  
 log.sigma.3.0.mm.3D\_ngtdm\_Busyness;  
 wavelet.HLL\_firstorder\_Maximum;  
 wavelet.HLH\_gldm\_LargeDependenceHighGrayLevel  
 Emphasis;  
 wavelet.HLH\_ngtdm\_Busyness;  
 wavelet.HHH\_glcm\_DifferenceVariance; and

logarithm\_glszm\_HighGrayLevelZoneEmphasis.

The radiomics signature was constructed and represented as radiomics score calculated by the formula followed:

Radiomics Score =  
 -original\_shape\_Elongation·0.92777 +  
 log.sigma.2.0.mm.3D\_glrmlm\_GrayLevelNonUniformity  
 Normalized·1.596924 +  
 log.sigma.3.0.mm.3D\_ngtdm\_Busyness·0.341031 -  
 wavelet.HLL\_firstorder\_Maximum·0.06123 -  
 wavelet.HLH\_gldm\_LargeDependenceHighGrayLevel  
 Emphasis·0.23798 +



**FIGURE 4** Predictive capacity of radiomic signature. Kaplan–Meier curve shows that this radiomic signature could effectively discriminate patients with better BMFS survival from those with worse BMFS in training group (not reached [NR] vs. 19.6 months;  $p = 0.019$ ; HR = 2.09, 95% confidence interval [CI], 1.10–3.95) (Figure 4(a)) and validation group (NR vs. 12.8 months;  $p = 0.048$ ; HR, 2.41; 95% CI, 0.95–6.08) (Figure 4(b)). Patients with low radiomics score and non-smoking history had the longest BMFS than those with at least one of these two negative features (NR vs. 19.5 months vs. 17.1 months vs. 7.5 months;  $p < 0.001$ ) (Figure 4(c))

wavelet.HLH\_ngtdm\_Busyness-0.736794 -  
 wavelet.HHH\_glcm\_DifferenceVariance-0.75541 -  
 logarithm\_glszm\_HighGrayLevelZoneEmphasis-0.04781

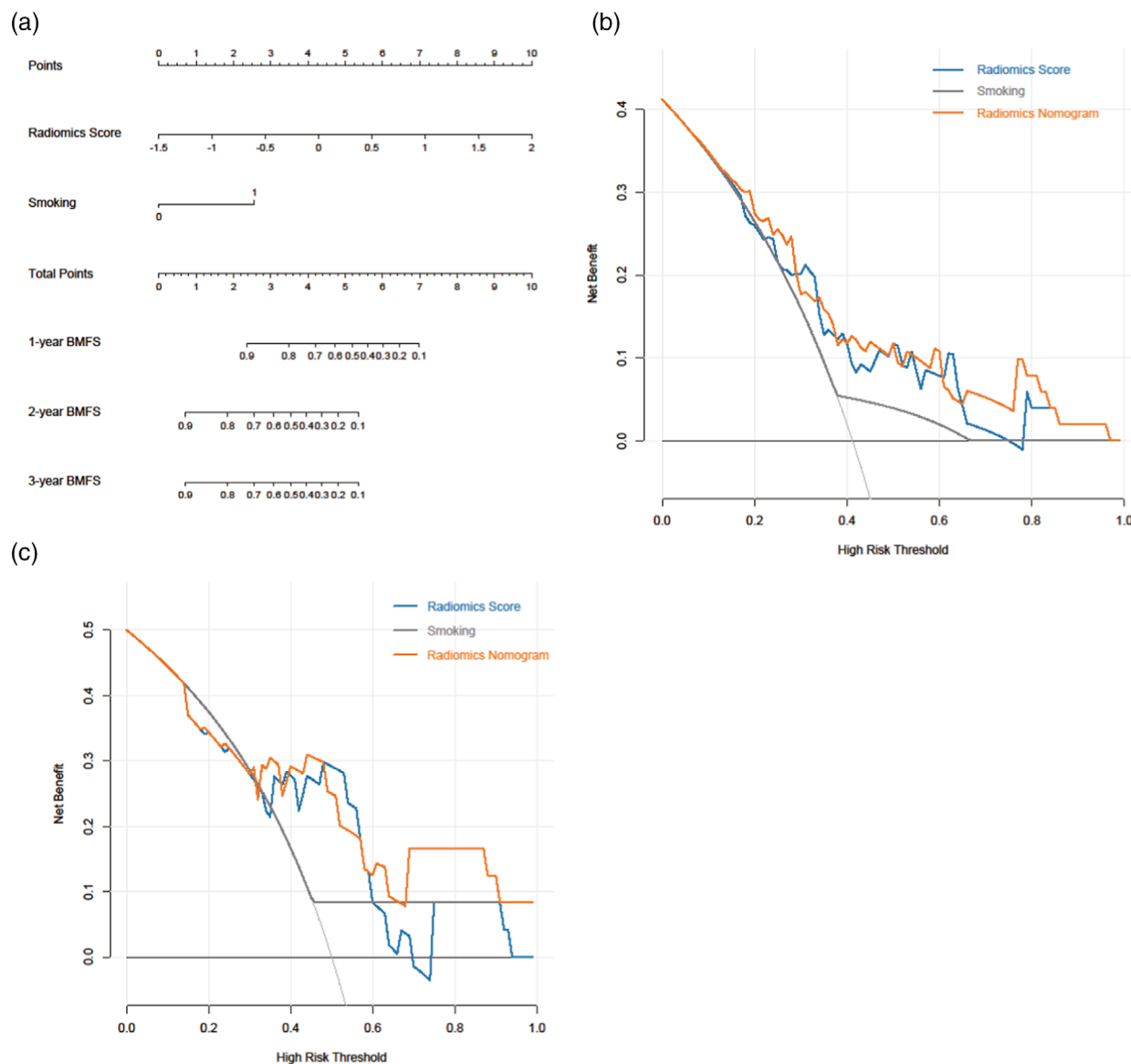
### Performance of multimodality prediction model

Patients were classified into a high-risk group and a low-risk group by the median radiomics scores in the training cohort ( $-0.512$ ) as cut-off. There are no correlations

between patient characteristics and radiomics score (Table 2), except that larger tumor volume was significantly associated with lower radiomics score ( $p = 0.002$ ) and higher stage ( $p = 0.045$ ). There is no difference of patient characteristics between BM group and those without BM (Table 3).

Patients with low radiomics score had significantly longer BMFS than those with high radiomics score in the training cohort (not reached [NR] vs. 19.6 months;  $p = 0.019$ ; HR, 2.09; 95% CI, 1.10–3.95) (Figure 4(a)). The findings





**FIGURE 5** (a) Nomogram integrated radiomics signature and smoking history predicting intracranial progression-free survival in ALK-positive NSCLC patients treated with crizotinib. Decision curve analysis for each model in the training (b) and validation (c) cohort. The y-axis measured by the net benefit is calculated by summing the benefit (true positive finding) and subtracting the harm (false positive finding). The latter is weighted by a factor related to the relative harm of brain metastasis compared to the harm of unnecessary treatment. Compared with the other characteristics and simple strategies such as follow-up of all patients (curve light gray line) or no patients (horizontal light gray line), the radiomics nomogram (the orange line) had the highest net benefit across most of threshold probabilities when a patient would choose to undergo imaging follow-up

were confirmed in the validation cohort (median BMFS between low and high radiomics score: NR vs. 12.8 months;  $p = 0.048$ ; HR, 2.41; 95% CI, 0.95–6.08) (Figure 4(b)).

Because of the limited sample size and the similarity between the training and validation set, we next conducted univariate and multivariate Cox regression analyses in the combined cohort (Table 4). In the univariate analysis, radiomics score (HR, 5.45; 95% CI, 2.49–11.92;  $p < 0.0001$ ), smoking history (HR, 3.52; 95% CI, 1.12–11.12;  $p = 0.032$ ), and platelet lymphocyte ratio (HR, 3.32; 95% CI, 1.58–6.96;  $p = 0.001$ ) was significantly associated with for BM occurrence. Multivariate Cox regression analysis revealed that radiomics score (HR, 5.07; 95% CI, 1.73–14.85;  $p = 0.003$ )

and smoking (HR, 5.84; 95% CI, 1.70–20.04;  $p = 0.005$ ) remained independent factors for BMFS. Subsequently, patients with low radiomics score and non-smoking history had the longest BMFS than those with at least one of these two negative features ( $p < 0.001$ ) (Figure 4(c)).

### Performance interpretation with nomogram

The radiomics signature built on eight selected features showed good discrimination, yielding a C-index of 0.733 (95% CI, 0.637–0.828) in the training cohort and 0.693 (95% CI, 0.569–0.818) in the validation cohorts.

TABLE 4 Univariate and multivariate COX analyses of brain metastases-free survival

| Characteristics           | Univariate cox |            |                     | Multivariate Cox |            |                    |
|---------------------------|----------------|------------|---------------------|------------------|------------|--------------------|
|                           | Hazard ratio   | 95% CI     | p-value             | Hazard ratio     | 95% CI     | p-value            |
| Age                       | 1.00           | 0.96–1.03  | 0.919               |                  |            |                    |
| Sex                       | 0.81           | 0.35–1.92  | 0.640               |                  |            |                    |
| Smoking                   | 3.52           | 1.11–11.12 | 0.032 <sup>a</sup>  | 5.84             | 1.70–20.04 | 0.005 <sup>a</sup> |
| ECOG PS                   | 1.79           | 0.70–4.60  | 0.226               |                  |            |                    |
| Bone metastasis           | 0.91           | 0.21–3.96  | 0.904               |                  |            |                    |
| Liver metastasis          | 1.39           | 0.32–6.03  | 0.66                |                  |            |                    |
| Intrapulmonary metastasis | 0.95           | 0.4–2.23   | 0.901               |                  |            |                    |
| No. of metastatic sites   | 0.73           | 0.28–1.87  | 0.506               |                  |            |                    |
| Treatment line            | 0.57           | 0.18–1.87  | 0.356               |                  |            |                    |
| Radiomics score           | 5.45           | 2.49–11.92 | <0.001 <sup>a</sup> | 5.07             | 1.73–14.85 | 0.003 <sup>a</sup> |
| Volume                    | 1.17           | 0.63–2.16  | 0.618               |                  |            |                    |
| Platelet                  | 1.00           | 1.00–1.01  | 0.135               |                  |            |                    |
| NLR                       | 1.02           | 0.85–1.22  | 0.506               |                  |            |                    |
| PLR                       | 3.32           | 1.58–6.96  | 0.001 <sup>a</sup>  | 1.55             | 0.52–4.63  | 0.430              |

<sup>a</sup>Statistically significant.

The only independent clinical risk factor was smoking history and was combined with the radiomics signature for nomogram construction (Figure 5(a)). The C-indexes of the radiomics-based nomogram incorporating radiomics signature and smoking history were 0.762 (95% CI, 0.663–0.861) and 0.724 (95% CI, 0.601–0.847), respectively.

Further, a decision curve analysis showed that the radiomics-based nomogram had a relatively higher overall net benefit than the clinical factor and radiomics signature only across the majority of the range of reasonable threshold probabilities (Figure 5(b),(c)). The calibration curves of the integrated nomograms for the probability of BM at 1, 2, or 3 years undergoing crizotinib treatment are shown in Figure S2, and these curves exhibit better agreement between the estimation and actual observation.

## DISCUSSION

We have developed and validated a robust, relatively simple radiomics-based nomogram that is able to predict the development of subsequent BM in patients with advanced NSCLC treated with crizotinib. This model can be used to predict which patients are at risk for BM, and therefore, can aid in selecting candidates for treatment with newer-generation ALK inhibitors or other preventive approaches. This tool is useful in deciding the monitoring strategies in patients with non-brain metastatic disease. To the best of our knowledge, this is the first predictive model for BM in this population. Extensive research efforts are currently under way at the molecular level, with the goal of optimizing predictive models incorporating clinical, radiological, pathological, and molecular data.

BM is an increasing challenge in the management of ALK-rearranged lung cancer.<sup>16</sup> It is associated with reduced

quality of life because of neurologic impairment following disease progression or brain local treatment. Therefore, predicting and clustering patients by the risk of developing BM would enrich populations for appropriate management to reduce or prevent occurrence of BM, and improve the clinical outcomes. Newer-generation ALK inhibitors, including alectinib, ceritinib, brigatinib, and lorlatinib, have better blood–brain barrier penetrating capacity and intracranial activity.<sup>17,18</sup> In this case, patients who are at higher risk of developing BM predicted by our proposed model should preferentially receive treatment with newer-generation ALK inhibitors. Currently, alectinib, brigatinib, and lorlatinib are recommended as preferred first-line options for ALK-rearranged lung cancer as per the National Comprehensive Cancer Network (NCCN) guidelines. However, there was no evidence of significant overall survival improvement with these agents, compared with crizotinib.<sup>19</sup> We proposed that crizotinib is useful in circumstances where patients are at lower risk of developing BM and it is better to retain newer-generation ALK inhibitors for subsequent therapies in low-risk population. Optionally, it remains unknown if high-risk patients should undergo prophylactic brain radiotherapy.

Regardless of baseline BM status, the brain remained the most common site of disease progression in ALK-rearranged NSCLC patient. We also found that the occurrence of BM was the most common reason for crizotinib failure, occurring in 44% of patients without baseline BM. The one underlying reason was the limited penetration of crizotinib into the central nervous system through the blood–brain barrier.<sup>20</sup> Another reason may be the neglect of tumor heterogeneity. Studies have reported molecular status, including EGFR, ALK, and KRAS, demonstrated significantly higher risk of BM than triple negative.<sup>21</sup> The potential mechanisms of promoting brain progression over time by ALK-rearrangement in lung cancer remain unclear. Yet, evaluating the risk of lung

cancer invasion and recurrence based on different genotypes and phenotypes will be the key to understand the tumor heterogeneity and implementation of individualized treatment.<sup>22</sup> Although more and more studies have tried to establish specific prediction models for ALK-positive NSCLC patients in terms of survival,<sup>23</sup> there are no effective and recognized methods for using radiomics information from primary lung tumor lesions to predict risk of developing BM in crizotinib-treated patients. Therefore, our research is enlightening and shows the potential for clinical practice and future research.

In this study, we analyze 489 radiomics features from pretreatment thoracic CT images of ALK-rearranged patients who were subsequently treated with crizotinib. In the process of building a radiomics model, eight signatures, including original\_shape\_Elongation, [log.sigma.2.0.mm.3D\\_glrlm\\_GrayLevelNonUniformityNormalized](#), [log.sigma.3.0.mm.3D\\_ngtdm\\_Busyness](#), wavelet.HLL\_firstorder\_Maximum, wavelet.HLH\_gldm\_LargeDependenceHighGrayLevelEmphasis, wavelet.HLH\_ngtdm\_Busyness, wavelet.HHH\_glcm\_Difference Variance, and logarithm\_glszm\_HighGrayLevelZoneEmphasis, were selected using LASSO Cox regression. The selected features obtained were combined as a radiomic signature, that demonstrated sufficient discrimination in both training and validation cohorts. Radiomics postulated that intratumor heterogeneity was exhibited on the spatial distribution of voxel intensity.<sup>24</sup> Findings of previous studies have supported the hypothesis that proteogenomic and phenotypic information of the tumor can be inferred from radiologic images.<sup>25</sup> Radiomics has been reported in predicting lymph node metastases in different kinds of tumor, including lung cancer,<sup>26</sup> colorectal cancer,<sup>27</sup> and so on. Moreover, it could decode different tumor genotype such as EGFR mutation, or ALK, ROS1, and RET fusion in lung cancer.<sup>28</sup> Most of these radiomic features are texture features that reflect image heterogeneity and were similar to those in other proposed signatures for NSCLC.<sup>29</sup> Our study focused on a minority of NSCLC patients with high incidence of BM and we aimed at finding possible model combined with radiomics marker and clinicopathological parameters for predicting subsequent BM following crizotinib treatment, which has significance for clinical treatment decision-making and prognosis evaluation.

In addition to the high-order radiomic features, multiple feature combinations can better reflect the complex heterogeneity of tumors, which will undoubtedly improve the accuracy of prediction. Risk factors for development of BM identified in our study only include smoking. Smoking was significant risk factor in the multivariate analysis. Cigarette smoking is a well-known environmental risk factor of lung carcinogenesis.<sup>30</sup> Tobacco smoke contains many mutagenic and carcinogenic chemicals, which may be related to mutations in tumor suppressor genes and oncogenes, such as p53<sup>31</sup> and K-Ras.<sup>32</sup> What is more, nicotine could promote BM by skewing the polarity of M2 microglia, which enhances metastatic tumor growth.<sup>33</sup> In this study, there was a significant correlation between the smoking and the brain metastases of patients using crizotinib. We found

smoking history plus lung tumor CT-based radiomics signature could produce a nomogram that better predicts the occurrence of BM. This could easily be incorporated into clinical practice by stratifying patients based on radiomics score and smoking history.

There are some limitations in this study. First, considering the general low incidence of ALK rearrangement among NSCLC, the sample size of this study is relatively small. However, by using LASSO Cox regression method, we were able to mine key features that together form an independent risk score. Additionally, the prediction model shows good and similar performance both in the training and validation cohorts. Prospective multicenter studies with larger sample size are needed to further validate the robustness and reproducibility of our prediction model. Second, changes in radiomic characteristics from one time point to the next (delta radiomics), such as changes before and after treatment, could dynamically respond to changes in tumor heterogeneity and may have a higher predictive value than a single time point extraction. Third, we did not analyze other potential genomic characteristics, which is undoubtedly a medical hotspot to promote the progress in the analysis.

## CONCLUSIONS

In conclusion, we have developed a radiomics-based nomogram that robustly predicts subsequent BM in patients with ALK-rearranged lung cancer with non-brain metastatic disease. Our model will allow selection of patients at higher risk for BM and therefore, will facilitate the design of prevention trials or development of drugs with better intra-cranial activity for the affected population.

## CONFLICT OF INTEREST

The authors declare no conflicts of interest.

## ORCID

Yongluo Jiang  <https://orcid.org/0000-0002-3456-0570>

## REFERENCES

- Chen W, Zheng R, Zhang S, Zeng H, Zuo T, Xia C, et al. Cancer incidence and mortality in China in 2013: an analysis based on urbanization level. *Chin J Cancer Res.* 2017;29:1–10.
- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA: a cancer journal for clinicians.* 2018;68:7–30.
- Arbour KC, Riely GJ. Systemic therapy for locally advanced and metastatic non-small cell lung cancer: a review. *JAMA.* 2019;322:764–74.
- Kris MG, Johnson BE, Berry LD, Kwiatkowski DJ, Iafrate AJ, Wistuba II, et al. Using multiplexed assays of oncogenic drivers in lung cancers to select targeted drugs. *JAMA.* 2014;311:1998–2006.
- Shaw AT, Engelman JA. ALK in lung cancer: past, present, and future. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology.* 2013;31:1105–11.
- Solomon BJ, Mok T, Kim D-W, Wu YL, Nakagawa K, Mekhail T, et al. First-line crizotinib versus chemotherapy in ALK-positive lung cancer. *N Engl J Med.* 2014;371:2167–77.

7. Costa DB, Shaw AT, Ou S-HI, Solomon BJ, Riely GJ, Ahn MJ, et al. Clinical experience with Crizotinib in patients with advanced ALK-rearranged non-small-cell lung cancer and brain metastases. *Am J Clin Oncol*. 2015;33:1881–8.
8. Yoshida T, Oya Y, Tanaka K, Shimizu J, Horio Y, Kuroda H, et al. Clinical impact of crizotinib on central nervous system progression in ALK-positive non-small lung cancer. *Lung cancer (Amsterdam, Netherlands)*. 2016;97:43–7.
9. Guérin A, Sasane M, Zhang J, Culver KW, Dea K, Nitulescu R, et al. Brain metastases in patients with ALK+ non-small cell lung cancer: clinical symptoms, treatment patterns and economic burden. *J Med Econ*. 2015;18:312–22.
10. Limkin EJ, Sun R, Dercle L, Zacharaki EI, Robert C, Reuzé S, et al. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. *Ann Oncol*. 2017;28:1191–206.
11. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RGPM, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. 1990;2012(48):441–6.
12. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. 2016;278:563–77.
13. Chen B, Zhang R, Gan Y, Yang L, Li W. Development and clinical application of radiomics in lung cancer. *Radiat Oncol (London, England)*. 2017;12:154.
14. Wu S, Zheng J, Li Y, Yu H, Shi S, Xie W, et al. A Radiomics nomogram for the preoperative prediction of lymph node metastasis in bladder cancer. *Clin Cancer Res*. 2017;23:6904–11.
15. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res*. 2017;77:e104–e7.
16. Johung KL, Yeh N, Desai NB, Williams TM, Lautenschlaeger T, Arvold ND, et al. Extended survival and prognostic factors for patients with ALK-rearranged non-small-cell lung cancer and brain metastasis. *Am J Clin Oncol*. 2016;34:123–9.
17. Hida T, Nokihara H, Kondo M, Kim YH, Azuma K, Seto T, et al. Alectinib versus crizotinib in patients with ALK-positive non-small-cell lung cancer (J-ALEX): an open-label, randomised phase 3 trial. *Lancet (London, England)*. 2017;390:29–39.
18. Shaw AT, Bauer TM, de Marinis F, Felip E, Goto Y, Liu G, et al. First-line lorlatinib or crizotinib in advanced ALK-positive lung cancer. *N Engl J Med*. 2020;383:2018–29.
19. Ito K, Yamanaka T, Hayashi H, Hattori Y, Nishino K, Kobayashi H, et al. Sequential therapy of crizotinib followed by alectinib for non-small cell lung cancer harbouring anaplastic lymphoma kinase rearrangement (WJOG9516L): a multicenter retrospective cohort study. *Eur J Cancer*. 1990;2021(145):183–93.
20. Shi W, Dicker AP. CNS metastases in patients with non-small-cell lung cancer and ALK gene rearrangement. *J Clin Oncol*. 2016;34:107–9.
21. Doebele RC, Lu X, Sumey C, Maxson DLA, Weickhardt AJ, Oton AB, et al. Oncogene status predicts patterns of metastatic spread in treatment-naïve non-small cell lung cancer. *Cancer*. 2012;118:4502–11.
22. Mak KS, Gainor JF, Niemierko A, Oh KS, Willers H, Choi NC, et al. Significance of targeted therapy and genetic alterations in EGFR, ALK, or KRAS on survival in patients with non-small cell lung cancer treated with radiotherapy for brain metastases. *Neuro Oncology*. 2015;17:296–302.
23. Song Z, Liu T, Shi L, Yu Z, Shen Q, Xu M, et al. The deep learning model combining CT image and clinicopathological information for predicting ALK fusion status and response to ALK-TKI therapy in non-small cell lung cancer patients. *Eur J Nucl Med Mol Imaging*. 2021;48:361–71.
24. Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun*. 2014;5:4006.
25. Rutman AM, Kuo MD. Radiogenomics: creating a link between molecular diagnostics and diagnostic imaging. *Eur J Radiol*. 2009;70:232–41.
26. Zhong Y, Yuan M, Zhang T, Zhang YD, Li H, Yu TF. Radiomics approach to prediction of occult mediastinal lymph node metastasis of lung adenocarcinoma. *AJR Am J Roentgenol*. 2018;211:109–13.
27. Huang Y-Q, Liang C-H, He L, Tian J, Liang CS, Chen X, et al. Development and validation of a Radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. *J Clin Oncol*. 2016;34:2157–64.
28. Yang X, Dong X, Wang J, Li W, Gu Z, Gao D, et al. Computed tomography-based Radiomics signature: a potential indicator of epidermal growth factor receptor mutation in pulmonary adenocarcinoma appearing as a subsolid nodule. *Oncologist*. 2019;24:e1156–e64.
29. Huang L, Chen J, Hu W, Xu X, Liu D, Wen J, et al. Assessment of a Radiomic signature developed in a general NSCLC cohort for predicting overall survival of ALK-positive patients with different treatment types. *Clin Lung Cancer*. 2019;20:e638–e51.
30. Karam-Hage M, Cinciripini PM, Gritz ER. Tobacco use and cessation for cancer survivors: an overview for clinicians. *Ca-Cancer J Clin*. 2014;64:272–90.
31. Denissenko MF, Pao A, Tang M, Pfeifer GP. Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in P53. *Science*. 1996;274:430–2.
32. Westra WH, Baas IO, Hruban RH, Askin FB, Wilson K, Offerhaus GJ, et al. K-ras oncogene activation in atypical alveolar hyperplasias of the human lung. *Cancer Res*. 1996;56:2224–8.
33. Wu S-Y, Xing F, Sharma S, Wu K, Tyagi A, Liu Y, et al. Nicotine promotes brain metastasis by polarizing microglia and suppressing innate immune function. *J Exp Med*. 2020;217:e20191131.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Jiang Y, Wang Y, Fu S, Chen T, Zhou Y, Zhang X, et al. A CT-based radiomics model to predict subsequent brain metastasis in patients with ALK-rearranged non-small cell lung cancer undergoing crizotinib treatment. *Thorac Cancer*. 2022;13:1558–69. <https://doi.org/10.1111/1759-7714.14386>