Radiomics Nomograms for Predicting Different Patterns of Distant Metastases After Definitive Chemoradiotherapy for Locally Advanced Non-Small-Cell Lung Cancer

Xin Tong  
Shandong First Medical University  https://orcid.org/0000-0001-5248-4770

Xia Chen  
Shandong First Medical University

Yicong Chen  
Shandong University Cheeloo College of Medicine

Qingtao Qiu  
Shandong Cancer Hospital: Shandong Tumor Hospital and Institute

Xiaorong Sun  
Shandong Cancer Hospital: Shandong Tumor Hospital and Institute

Ligang Xing (xinglg@medmail.com.cn)  
Shandong Cancer Hospital: Shandong Tumor Hospital and Institute

Research

Keywords: Non-small-cell lung cancer, Distant metastases, Oligometastases, Radiomics, Chemoradiotherapy

DOI: https://doi.org/10.21203/rs.3.rs-84168/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

**Background:** To develop and validate radiomics-based nomograms for predicting different patterns of distant metastases (DMs), including oligometastases and polymetastases, after definitive chemoradiotherapy (CRT) for locally advanced non-small-cell lung cancer (LA-NSCLC).

**Methods:** In all, 139 LA-NSCLC patients treated with definitive CRT between January 2014 and December 2017 were analyzed retrospectively. Computed tomography (CT) radiomics features were quantitatively extracted by 3D Slicer software. The least absolute shrinkage and selection operator (LASSO) Cox regression model was applied for date dimension reduction, feature selection and radiomics signature development. Univariate and multivariate Cox regression analyses were utilized to identify independent predictors of DMs. Finally, a nomogram that incorporates both the radiomics signature and independent clinical risk factors was developed and assessed according to the Harrell concordance index (C-index), calibration curve and clinical usefulness.

**Results:** With a median follow-up of 20.6 months, DMs were detected by follow-up imaging in 47 patients (26 patients had oligometastatic disease). For those who experienced DM, the first metastases were most common in the lung (32.1%), followed by the brain (20.7%), and bone (15.1%). The median time to DMs was 13.5 months. Twelve radiomics features were found to be predictive for DMs. Multivariate Cox proportional hazards testing revealed that histology subtype and overall stage were independent clinical risk factors for DMs (p-value < 0.05). The radiomics nomogram combining the selected radiomics signature, histology subtype and overall stage predicted DMs better than either the radiomics signature or clinical factors alone (p-value < 0.001). The model showed good discrimination and good calibration. Furthermore, the radiomics signature consisting of nine selected features was significantly associated with oligometastases (p-value < 0.001). The radiomics-based nomogram showed strong discrimination ability for the prediction of oligometastatic disease. The C-index was 0.800 (95% CI= 0.710 to 0.890) for the primary cohort, and 0.828 (95% CI= 0.686 to 0.970) for the validation cohort.

**Conclusion:** The above-described radiomics nomograms can precisely predict DM and the oligometastatic state in patients with LA-NSCLC, which may constitute a useful clinical tool to guide subsequent personalized treatments.

Introduction

Definitive chemoradiotherapy (CRT) is considered the standard of care for patients presenting with inoperable locally advanced non-small-cell lung cancer (LA-NSCLC)[1]. Despite recent advances in radiation therapy technologies, approximately 30%-40% of NSCLCs metastasize to distant sites after definitive treatment[2]. There is a common belief that patients with metastatic NSCLC cannot be cured, and treatment consists mainly of systemic therapy with radiation reserved for palliation[3]. However, patients with limited metastatic burden, “oligometastatic” disease (OMD), may have a better prognosis than those with polymetastatic disease[4, 5]. Although there is no formal consensus on the definition of
oligometastatic disease, it is currently widely accepted that it constitutes up to five isolated lesions in up to three different organs that are potentially amenable to locally ablative therapy (surgical resection, radiotherapy or both)[6–8].

OMD was first coined by Hellman and Weichselbaum, who hypothesized an intermediate clinical state between locoregional and widely metastatic disease[9]. For patients with oligometastatic and oligorecurrent NSCLC, intensive local therapy, including surgical resection or definitive radiotherapy, is expected to achieve long-term survival and may even be curative[10–15]. However, the accurate classification of oligometastatic and polymetastatic patients and the identification of patients eligible for local radical therapy or systemic therapy remain of high importance. Lussier et al. found that molecular techniques using microRNAs can distinguish between oligometastases and polymetastases in lung cancer[16, 17]. Furthermore, some evidence has indicated that whole-body $^{18}$F-FDG-positron emission tomography (PET)/CT is available for identifying and following up patients with oligometastatic NSCLC[18, 19]. Consequently, biomarkers can be developed to identify and distinguish patients with distant metastases (DMs), oligometastases and polymetastases before treatment, which may help clinicians adopt personalized treatments to improve patient prognosis.

Radiomics, as an emerging field of quantitative imaging, has the capacity to capture tumor phenotypic differences by examining a large set of quantitative features based on noninvasive imaging, which provides a new basis for the accurate diagnosis and treatment of tumors. It shows great potential for assisting in the clinical diagnosis, treatment, and prognosis of cancer[20–22]. Recently, numerous studies have shown that imaging-based radiomics features can be used to quantify tumor heterogeneity and have potential application as clinical biomarkers for patient stratification[2, 23–25]. In particular, radiomics studies have shown that CT-derived imaging features may be novel prognostic indicators for DMs in stage III NSCLC patients treated with CRT[2, 20]. Nevertheless, there is still a lack of studies, particularly regarding imaging-based radiomics in the identification of oligometastatic and polymetastatic NSCLC.

Therefore, the aim of this study was to develop CT-based radiomics nomograms to precisely predict different patterns of DMs, including oligometastases and polymetastases, in patients with LA-NSCLC and to provide useful information for adopting accurate and personalized treatments in the clinic.

**Materials And Methods**

**Patients**

Ethical approval was obtained for this study, and the necessity to obtain informed consent was waived, as the data were analyzed retrospectively and anonymously. We retrospectively reviewed the medical records of 403 patients with stage III (American Joint Committee on Cancer (AJCC) 8th edition) NSCLC after definitive CRT in Shandong Cancer Hospital between January 2014 and December 2017. Patients were excluded from the analysis if they met at least one of the following criteria: (I) surgery before CRT;
(II) radiotherapy dose less than 55 Gy; (III) no pretreatment CT images; (IV) no posttreatment CT images; (V) other causes of mortality; or (VI) loss to follow-up before the clinical endpoints.

In total, 139 patients were identified in our analysis: patients treated before May 4th, 2016, were included in the primary cohort (n = 100); other patients were evaluated to form an independent validation cohort (n = 39). At baseline, clinical features of primary NSCLC patients (age, gender, smoking history, tumor location, etc.), and the acquisition date of CT imaging were recorded.

**Clinical endpoints**

Patients were followed up every three to six months after treatment, and surveillance contrast-enhanced CT and/or PET/CT, brain magnetic resonance imaging (MRI), and whole-body bone scans were performed to assess treatment response or tumor progression based on the US National Comprehensive Cancer Network (NCCN) guidelines. The primary endpoint of this study was DMs, which was defined as progression of the disease to other organs as assessed in surveillance scans, and time to DMs was defined as the time interval between the start date of CRT and the first scan date of radiographically evident distant metastases or censoring (date of last negative scan).

Oligometastases was defined as 1–5 separate metastatic lesions in up to three different organs. Metastases to all organs were included, except diffuse serosal metastases (meningeal, pericardial, pleural, mesenteric) and bone marrow involvement, as these cannot be treated with radical intent. Other metastases were classified as polymetastases.

**Clinical variables**

The conventional clinical parameters considered for this study included age, gender, smoking status, histology subtype (1-squamous cell carcinoma (SCC), 2-adenocarcinoma (AC), 3-others), tumor location (peripheral or central), Eastern Cooperative Oncology Group (ECOG) performance status (PS), tumor-node-metastasis (TNM) stage per the AJCC staging system (8th edition), CT-based measurements commonly utilized in the clinic (e.g., maximal tumor diameter measured on a single axial slice), and treatment characteristics.

**Image acquisition**

All patients underwent pretreatment contrast-enhanced (reconstruction thickness of 5 mm) CT with a 64-row detector scanner (Somatom Definition AS, Siemens Healthineers, Germany). The acquisition parameters were as follows: tube voltage of 120 kV, tube current of 200 mAs, detector collimation of 64 × 0.625 mm, 1.5 beam pitch, and 512 × 512 matrix size. We retrieved the pretreatment CT images in DICOM format from the picture archiving and communication system (PACS; Carestream, Canada).
Tumor segmentation and radiomics feature extraction

All available pretreatment CT images were collected centrally and transferred to 3D Slicer (software version 4.8.1), an open-source image analysis platform for image registration, segmentation, 3D visualization, and feature extraction[26–29]. The regions of interest (ROI) was those that contained the entire primary lung tumor, and they were successfully segmented in 3D with a manual single-click ensemble segmentation approach by an experienced radiologist blinded to all clinical outcomes, running on the 3D-Slicer software. Then, primary tumor segmentation was confirmed by another senior radiologist.

In total, 724 quantitative radiomics features, including first-order features, shape, gray-level cooccurrence matrix (GLCM), gray-level run-length matrix (GLRLM), gray-level size zone matrix (GLSZM) and neighboring gray-tone difference matrix (NGTDM), were extracted from each patient's contrast-enhanced CT images. In addition, considering that a wavelet[30] provides a spatial and frequency representation of the signal, the aforementioned texture features were also extracted from the images that were preprocessed with the wavelet filter. Overall, those radiomics features were extracted from both filtered and unfiltered images.

Statistical analysis

Clinical variables Baseline continuous variables were compared between the primary and validation cohorts using the Mann-Whitney U-test, and categorical data were analyzed by the chi-square test or Fisher's exact test. Univariate and multivariate Cox proportional hazards regression modeling was utilized to evaluate clinical variables as predictors of DMs in SPSS 25 statistical software (IBM, Armonk, NY). P values < 0.05 were considered statistically significant, and all P values were two-sided.

Construction of the radiomics score-based signature The least absolute shrinkage and selection operator (LASSO) Cox regression model, which is suitable for the reduction of high-dimensional data, was applied to select the best predictive features to develop the radiomics signature[31, 32]. The discrimination of that signature was calculated by the area under the curve (AUC). A radiomics score (Rad-score) was calculated for each patient via a linear combination of selected features that were weighted by their respective coefficients.

Validation of the radiomics signature The patients were classified into high-risk or low-risk groups according to the Rad-score, the threshold of which was identified by using X-tile (X-tile software, version 3.6.1, Yale University School of Medicine, New Haven, Conn)[33]. Kaplan-Meier survival curves were generated to depict the association between the radiomics signature and clinical outcomes. It was first evaluated in the primary cohort and then verified in the validation cohort. Log-rank testing was performed to compare the difference in the survival curves between the high-risk and low-risk groups.

Model construction For the construction of the nomogram, we performed multivariate Cox analysis of clinical parameters, including age, gender, smoking status, histology subtype, tumor location, ECOG, TNM
stage, maximum tumor diameter (MTD) and treatment characteristics. The radiomics nomogram incorporated both the radiomics signature and the independent clinical risk factors.

Model evaluation The discrimination performance of the model was evaluated by the Harrell concordance index (C-index). Calibration curves were adapted to compare the agreement between the actual clinical outcomes and the predicted outcomes[34]. Decision curve analysis (DCA) was conducted to determine the clinical usefulness of the nomogram by quantifying the net benefits at different threshold probabilities in the entire cohort[35, 36].

All statistical analyses are two-sided, with the statistical significance level set at 0.05. Statistical analysis was performed with “glmnet,” “rms,” “Hmisc,” “lattice,” “survival,” “Formula,” “ggplot2,” and “rmda” modules in R software (Version 3.6.1; http://www.Rproject.org).

Results

Clinical characteristics

The demographic and clinical characteristics of the primary and validation cohorts are listed in Table 1. In total, 139 LA-NSCLC patients treated with definitive CRT were analyzed in this study. There were 16 females and 123 males with a median age of 61 years (range: 39 – 84 years). The median follow-up time was 20.6 months (range: 3.7–58.8 months). The median time to DM was 13.5 months (range: 1.4–56.8 months), with 47 (34%) patients who developed DMs versus 92 (66%) who did not. For those who experienced DMs, the first metastases were most commonly in the lung (32.1%), followed by the brain (20.7%) and bone (15.1%). Table 2 lists sites of metastases and the number of sites of metastatic lesions.
| Characteristic                              | All patients | Primary Cohort | Validation Cohort | p-value |
|--------------------------------------------|--------------|----------------|-------------------|---------|
|                                            | N(%)         | N(%)           | N(%)              |         |
| Median age at diagnosis (range), y         | 61 (39–84)   | 61 (39–84)     | 62 (40–77)        | 0.974   |
| Gender                                     | 0.556        |                |                   |         |
| Male                                       | 123 (88.5)   | 87 (87.0)      | 36 (92.3)         |         |
| Female                                     | 16 (11.5)    | 13 (13.0)      | 3 (7.7)           |         |
| Smoking history                            | 0.499        |                |                   |         |
| Never                                      | 34 (24.5)    | 26 (26.0)      | 8 (20.5)          |         |
| Ever                                       | 105 (75.5)   | 74 (74.0)      | 31 (79.5)         |         |
| PS                                         | 0.190        |                |                   |         |
| 0–1                                        | 136 (97.8)   | 99 (99.0)      | 37 (94.9)         |         |
| 2                                          | 3 (2.2)      | 1 (1.0)        | 2 (5.1)           |         |
| Location                                   | 0.452        |                |                   |         |
| Central                                    | 112 (80.6)   | 79 (79.0)      | 33 (84.6)         |         |
| Peripheral                                 | 27 (19.4)    | 21 (21.0)      | 6 (15.4)          |         |
| Tumor diameter of CT                       | 0.601        |                |                   |         |
| Median (range)                             | 4.4 (0.7–12.3)| 4.3 (0.7–12.2) | 4.8 (1.5–12.3)    |         |
| Histology subtype                          | 0.404        |                |                   |         |
| SCC                                        | 93 (66.9)    | 65 (65.0)      | 28 (71.8)         |         |
| AC                                         | 36 (25.9)    | 26 (26.0)      | 10 (25.6)         |         |
| Other                                      | 10 (7.2)     | 9 (9.0)        | 1 (2.6)           |         |
| Overall stage                              | 0.750        |                |                   |         |
| IIIA                                       | 49 (35.2)    | 37 (37.0)      | 12 (30.8)         |         |
| IIIB                                       | 75 (54.0)    | 53 (53.0)      | 22 (56.4)         |         |
| IIIC                                       | 15 (10.8)    | 10 (10.0)      | 5 (12.8)          |         |
| T stage                                    | 0.986        |                |                   |         |
| Characteristic            | All patients | Primary Cohort | Validation Cohort | p-value |
|---------------------------|--------------|----------------|-------------------|---------|
|                           | N(%)         | N(%)           | N(%)              |         |
| T1                        | 11 (7.9)     | 8 (8.0)        | 3 (7.7)           | 0.718   |
| T2                        | 33 (23.8)    | 23 (23.0)      | 10 (25.6)         |         |
| T3                        | 27 (19.4)    | 20 (20.0)      | 7 (18.0)          |         |
| T4                        | 68 (48.9)    | 49 (49.0)      | 19 (48.7)         |         |
| N stage                   |              |                |                   |         |
| N0                        | 10 (7.2)     | 7 (7.0)        | 3 (7.7)           |         |
| N1                        | 12 (8.6)     | 9 (9.0)        | 3 (7.7)           |         |
| N2                        | 83 (59.7)    | 62 (62.0)      | 21 (53.8)         |         |
| N3                        | 34 (24.5)    | 22 (22.0)      | 12 (30.8)         |         |
| Radiotherapy technique    |              |                |                   | < 0.001 |
| 3D-CRT                    | 64 (46.0)    | 57 (57.0)      | 7 (18.0)          |         |
| IMRT                      | 56 (40.3)    | 29 (29.0)      | 27 (69.2)         |         |
| Unknown                   | 19 (13.7)    | 14 (14.0)      | 5 (12.8)          |         |
| Radiotherapy doze (Gy)    |              |                |                   | 0.073   |
| 55–60                     | 113 (81.3)   | 85 (85.0)      | 28 (71.8)         |         |
| >60                       | 26 (18.7)    | 15 (15.0)      | 11 (28.2)         |         |
| Chemotherapy sequence     |              |                |                   | 0.089   |
| concurrent                | 59 (42.4)    | 38 (38.0)      | 21 (53.8)         |         |
| sequential                | 80 (57.6)    | 62 (62.0)      | 18 (46.2)         |         |
Table 2
Patterns of distant metastases after definitive chemoradiotherapy

| Site of metastases                        | N(%)  |
|------------------------------------------|-------|
| Brain                                    | 11 (20.7) |
| Lung (ipsilateral or contralateral)      | 17 (32.1) |
| Bone                                     | 8 (15.1)   |
| Adrenal                                  | 5 (9.4)    |
| Liver                                    | 3 (5.7)    |
| Other                                    | 9 (17.0)   |

| Number of metastatic lesions            | N(%)  |
|-----------------------------------------|-------|
| 1                                       | 16 (34.0) |
| 2                                       | 4 (8.5)    |
| 3                                       | 4 (8.5)    |
| 4                                       | 2 (4.3)    |
| 5                                       | 0 (0)      |
| > 5                                     | 21 (44.7)  |

According to the time of enrollment, participants were divided into a primary cohort and a validation cohort. Except for the use of radiotherapy (P < 0.05), there were no significant differences in clinical characteristics. An outlier was observed among the patient characteristics: the proportion of elderly patients in the validation cohort was higher than that in the primary cohort. There were no significant differences in the probability of DMs (hazard ratio (HR) = 1.40, 95% CI = 0.75 to 2.61, P > 0.05) between the two cohorts.

Clinical risk factor selection

Univariate Cox proportional hazards testing revealed that histology subtype (P < 0.01) and overall stage (P < 0.05) were statistically significant predictors of DM. The multivariate Cox proportional hazards modeling results are presented in Table 3. Histology subtype appeared to be an independent clinical risk factor (P = 0.013) for metastasis. Overall stage was also independently associated with increased risk of distant metastases (HR = 2.37, 95% CI = 1.00 to 5.59, P < 0.05). The model that incorporated the above independent predictors was developed and presented as the nomogram.
Table 3
Univariate and multivariate cox proportional hazards regression modeling for distant metastases

| Variables                        | Univariate Analysis | Multivariate Analysis |
|----------------------------------|---------------------|-----------------------|
|                                  | Hazard ratio (95% CI) | p-value   | Hazard ratio (95% CI) | p-value   |
| Age, y                           | 0.98 (0.94 to 1.02)  | 0.322     |                       |           |
| Gender (female vs. male)         | 1.19 (0.46 to 3.10)  | 0.720     |                       |           |
| Smoking (ever vs. never)         | 1.01 (0.45 to 2.25)  | 0.978     |                       |           |
| PS (2 vs. 0–1)                   | 3.69 (0.50 to 27.46) | 0.203     |                       |           |
| Location (peripheral vs. central)| 2.05 (0.97 to 4.34)  | 0.060     |                       |           |
| Tumor diameter of CT             | 0.97 (0.82 to 1.15)  | 0.708     |                       |           |
| Histology subtype                | 0.005               |           | 0.013                 |           |
| AC vs. SCC                       | 1.83 (0.84 to 4.00)  | 0.127     | 1.99 (0.91 to 4.34)   | 0.086     |
| Other vs. SCC                    | 4.80 (1.83 to 12.56) | 0.001     | 4.04 (1.53 to 10.63)  | 0.005     |
| Overall stage                    |                     |           |                       |           |
| IIIB + IIIC vs. IIIA             | 2.52 (1.09 to 5.84)  | 0.032     | 2.37 (1.00 to 5.59)   | 0.049     |
| T stage                          |                     | 0.117     |                       |           |
| T2 vs. T1                        | 0.18 (0.03 to 0.96)  | 0.044     |                       |           |
| T3 vs. T1                        | 0.70 (0.20 to 2.50)  | 0.587     |                       |           |
| T4 vs. T1                        | 1.01 (0.34 to 2.97)  | 0.986     |                       |           |
| N stage                          | 0.821               |           |                       |           |
| N1 vs. N0                        | 0.71 (0.10 to 5.05)  | 0.731     |                       |           |
| N2 vs. N0                        | 1.01 (0.24 to 4.34)  | 0.988     |                       |           |
| N3 vs. N0                        | 1.38 (0.29 to 6.51)  | 0.685     |                       |           |
| Radiotherapy technique           |                     | 0.906     |                       |           |
| IMRT vs. 3D-CRT                  | 0.91 (0.41 to 2.00)  | 0.810     |                       |           |
| Unknown vs. 3D-CRT               | 0.77 (0.23 to 2.62)  | 0.679     |                       |           |
| Radiotherapy dose(Gy)            |                     |           |                       |           |
| > 60 vs. 55–60                   | 0.62 (0.19 to 2.03)  | 0.428     |                       |           |
| Chemotherapy sequence            |                     |           |                       |           |
| Variables                     | Univariate Analysis | Multivariate Analysis |
|-------------------------------|---------------------|-----------------------|
|                               | Hazard ratio (95% CI) | p-value | Hazard ratio (95% CI) | p-value |
| Sequential vs. concurrent     | 1.13 (0.54 to 2.34)  | 0.744     |                       |         |

Feature selection and radiomics score-based signature construction

Twelve nonredundant predictors were extracted from the 724 features based on the primary cohort of the 100 cases (Fig. 1), and those features with nonzero coefficients were used in the LASSO Cox regression model. The radiomics signature, which was constructed based on the regression analysis with the radiomics score, was calculated for each patient (Table 4). The formula to calculate the radiomics score was: Score = Intercept + Coefficient × Radiomics features. The radiomics signature that we developed showed a significant capability to distinguish DM and non-DM in the primary cohort (AUC = 0.836, 95% CI = 0.749 to 0.903) and validation cohort (AUC = 0.661, 95% CI = 0.492 to 0.805).
Table 4
Radiomics features selected in LASSO regression analysis

| Intercept and radiomics features | Coefficients |
|----------------------------------|--------------|
| Intercept                        | 1.892        |
| original.glcm.Correlation        | -0.569       |
| wavelet.LHH.firstorder.Kurtosis  | 0.315        |
| wavelet.HHH.glcm.Idmn            | -1.399       |
| wavelet.HHH.glcm.Correlation     | -0.948       |
| wavelet.HHH.glcm.lmc1            | -1.146       |
| wavelet.HHH.firstorder.Mean      | 0.180        |
| wavelet.HHH.glrlm.LongRunLowGrayLevelEmphasis | 0.483 |
| wavelet.HHH.glszm.GrayLevelNonUniformityNormalized | 0.068 |
| wavelet.HHL.glszm.SmallAreaLowGrayLevelEmphasis | -0.785 |
| wavelet.LLL.firstorder.Kurtosis  | -0.082       |
| wavelet.LLL.glrlm.ShortRunLowGrayLevelEmphasis | -0.759 |
| wavelet.LLL.glrlm.LowGrayLevelRunEmphasis | -0.041 |

Twelve radiomics features with non-zero coefficients in the least absolute shrinkage and selection operator method (LASSO) cox regression model were selected. The radiomics signature was constructed based on the regression analysis with a radiomics score calculated for each patient. The formula to calculate the score of radiomics signature is \( \text{Score} = \text{Intercept} + \text{Coefficient} \times \text{Radiomics features} \).

Validation of the radiomics signature

The optimum cutoff generated by the X-tile plot was −2.45. Accordingly, patients were classified into a high-risk group (Rad-score ≥ -2.45) and a low-risk group (Rad-score < -2.45). Figure 2 demonstrates the potential stratification that can be achieved using the radiomics signature. The radiomics signature was associated with the metastasis-free probability in the primary cohort (HR = 20.11, 95% CI = 7.98 to 50.70, \( P < 0.001 \)). Using the cutoff value derived from the primary cohort, Kaplan-Meier analysis in the validation cohort showed a statistically significant difference (log-rank \( p < 0.05 \)) for metastasis-free probability estimates. The low-risk group showed an HR of 0.20 compared to the high-risk group.

Performance of the multimodality prediction model
A predictive nomogram (Fig. 3a) was generated on the basis of the selected radiomics signature, histology subtype, and overall stage. The C-index for the combined nomogram was 0.807 (95% CI = 0.748 to 0.866) for the primary cohort, which was confirmed to be 0.692 (95% CI = 0.556 to 0.828) by bootstrapping validation.

Further, the calibration curves of the nomogram for metastasis-free probability at 1, 2, or 3 years after definitive CRT are shown in Fig. 3b, and they demonstrated good agreement between the predicted and actual observations in the primary cohort.

The predictive performance of these models was assessed on the primary cohort and the validation cohort (Table 5). In the primary cohort, compared with either the radiomics signature (C-index = 0.763; 95% CI = 0.686 to 0.840) or the clinicopathologic model (C-index = 0.707; 95% CI = 0.629 to 0.785), the radiomics nomogram showed a better discrimination capability (P < 0.001). In the validation cohort, the radiomics signature and the radiomics nomogram had similar C-indexes (0.683 for the radiomics signature and 0.692 for the clinicopathologic model, P = .779).

| Model                      | Primary cohort          | Validation cohort       |
|----------------------------|-------------------------|-------------------------|
|                            | C-index  | 95% CI   | P-value | C-index  | 95% CI   | P-value |
| Histology subtype + Overall stage | 0.707   | 0.629–0.785 | <.001*  | 0.507   | 0.376–0.638 | <.001*  |
| Radiomics signature        | 0.763   | 0.686–0.840 | <.001** | 0.683   | 0.542–0.824 | 0.779** |
| Radiomics nomogram         | 0.807   | 0.748–0.866 | <.001** | 0.692   | 0.556–0.828 | 0.779** |

*The C-index was compared between the clinical nomogram and the radiomics signature.

**The C-index was compared between the radiomics signature and the radiomics nomogram.

Clinical use

The DCA for each model is presented in Fig. 4. Based on the threshold probability, the decision analysis curve was leveraged to evaluate the clinical application of the prediction model. In our study, a threshold probability < 0.38 or > 0.62 indicates that using the radiomics nomogram to predict DM improves the net benefit compared to that with standard treatment or no treatment. Within this range, the net benefit was comparable, with several overlaps, on the basis of the radiomics nomogram, the radiomics signature and the model with clinical risk factors integrated.
Subgroup analysis according to the number of metastatic lesions

Among the 47 NSCLC patients with DMs, 26 had oligometastatic disease. After excluding patients with polymetastatic NSCLC, we reclassified the remaining patients into the primary cohort (n = 96) and the validation cohort (22). In the oligometastatic NSCLC subgroup, multivariate Cox regression analysis proved that the histology subtype was an independent clinical risk factor (P < 0.05) (see Additional file 1: Tables S1). The radiomics signature, consisting of 9 features, was significantly associated with oligometastatic NSCLC (P < 0.001) (see Additional file 1: Tables S2). The AUC values of the radiomics signature in the primary and validation cohorts were 0.865 (95% CI = 0.781 to 0.926) and 0.659 (95% CI = 0.428 to 0.845), respectively. Consequently, we integrated some independent predictive factors (radiomics signature and histology subtype) to construct a nomogram specific to oligometastatic NSCLC (Fig. 3a), which has strong discrimination and calibration power (Fig. 3b). The C-index was 0.800 (95% CI = 0.710 to 0.890) for the primary cohort and 0.828 (95% CI = 0.686 to 0.970) for the validation cohort.

Discussion

Considering the high incidence of DM, metastatic NSCLC is traditionally believed to be incurable, with very little information on the identification of oligometastases and polymetastases. To our knowledge, this study is the first to build CT-based radiomics prediction models for different patterns of DMs in patients with LA-NSCLC based on limited numbers of metastatic lesions. In this study, we extracted twelve radiomics features with nonzero coefficients in the LASSO Cox regression model from a total of 100 LA-NSCLC patients treated with definitive CRT, and we validated the radiomics-based signature on an independent validation dataset. The radiomics signature successfully stratified patients according to their risk of DM. Incorporating the radiomics signature and clinical risk factors into the radiomics-based nomogram resulted in better performance (P < 0.001) as well as better discrimination and calibration for the prediction of DM (C-index: 0.807 for the primary cohort and 0.692 for the validation cohort) than either the clinicopathologic nomogram or the radiomics signature. Furthermore, for the oligometastatic NSCLC subgroup, we developed and validated a diagnostic, radiomics-based nomogram specific to this subgroup for the individualized pretreatment prediction of oligometastatic disease. Since patients with oligometastatic disease are different from those with polymetastases in terms of treatment and prognosis, early identification of patients at different risk of developing DM would help clinicians stratify patients and may also provide useful information to adapt treatments to improve outcomes.

We evaluated the conventional clinical parameters associated with DM to determine its clinical risk factors in the context of all LA-NSCLC patients. The results of this study revealed that the histology subtype and overall stage were independent clinical risk factors. Predictors of DM after definitive treatment for NSCLC have been elucidated in a number of reports. Previous investigations suggested that the histologic subtype was associated with sites of DM\[37]. Cox et al. demonstrated nonsquamous histology to be associated with DM from a pooled analysis of four Radiation Therapy Oncology Group
(RTOG) trials including 1765 patients. Similar results were confirmed by Tang et al., who also identified nonsquamous histology and advanced initial disease stage predicted metastasis at any site[38, 39]. In addition, compared with SCC, the risk of DM is higher in AC[40]. However, there was no statistically significant difference in DM incidence between SCC vs. AC in our study (P = 0.086). One possible explanation for this phenomenon may be the limited number of patients with AC.

Preliminary evidence has suggested that CT-based radiomics features are associated with failure patterns in patients with LA-NSCLC. Fried et al. found that the combination of CT texture features and conventional prognostic factors (CPFs) may have better predictive value than either alone for local-regional control and distant metastasis in stage III NSCLC patients treated with definitive radiotherapy[20]. Another report also concluded that the CT-based radiomics signature could be used as a biomarker for predicting distant failure after CRT in patients with locally advanced lung AC[2]. Currently, there is still a lack of targeted studies about imaging-based radiomics for the identification of oligometastases and polymetastases. First, our study demonstrated that the integrative nomogram of CT-based radiomics and clinical features could be used to generate outcome models superior to those only considering clinical factors or the radiomics signature. We also built a radiomics-based model for the identification of oligometastatic NSCLC subgroups. By means of the above-described models, we can preliminarily stratify patients with DMs to select patients with oligometastatic disease suitable for local radical therapy and to select the remaining patients with polymetastatic disease suitable for systemic therapy, including targeted therapy and immunotherapy.

In this study, we found that approximately 32.1% of first metastases occurred in the lung, followed by the brain (20.7%) and bone (15.1%). Our observed sites of first metastases were consistent with those presented in other studies assessing treatment with definitive CRT. In the study of NSCLC patients, approximately 78% whom have stage III NSCLC, first metastases were most common in the lung (27%), brain (26%), and bone (18%)[37]. A recent analysis of this same patient group showed that first metastases were most common in the brain (30%), contralateral lung (27%), and bone (27%) in stage III NSCLC patients treated with definitive CRT[41]. The brain was also an organ with a high probability of extrapulmonary metastasis in stage IIIA NSCLC patients, especially in AC patients[40, 42]. In the future, we hope to conduct additional studies examining the use of radiomics analysis in tumor-specific metastatic sites such as the brain.

Various limitations of this analysis deserve mention. First, as this study was retrospective in nature, our analysis is limited by its inherent biases and the relatively small sample size. To build practical radiomic models, multiple larger databases and multicenter prospective studies are required. Second, manual delineation is most commonly used in these types of analyses but is far from perfect. Thresholding is a useful strategy to enhance contour reproducibility, particularly in lung tumors. However, although our study has some limitations, it is one of the largest attempts to identify a biomarker predictive of different patterns of DM, including oligometastases and polymetastases, in patients with LA-NSCLC.

Conclusion
In conclusion, the combined model of the radiomics signature and clinical risk factors results in better performance for the prediction of DM in LA-NSCLC compared to the radiomics signature or clinical factors alone. Herein, we reported a radiomics-based model for the identification of the oligometastatic NSCLC subgroup, which may aid in the stratification of patients with DM and the selection of personalized treatments.

**Abbreviations**

DM: distant metastasis; CRT: chemoradiotherapy; LA-NSCLC: non-small-cell lung cancer; CT: Computed tomography; LASSO: least absolute shrinkage and selection operator; C-index: concordance index; OMD: oligometastatic disease; PET/CT: Positron emission tomography/computed tomography; AJCC: American Joint Committee on Cancer; MRI: magnetic resonance imaging; NCCN: National Comprehensive Cancer Network; SCC: squamous cell carcinoma; AC: adenocarcinoma; ECOG: Eastern Cooperative Oncology Group; PS: performance status; TNM: tumor-node-metastasis; ROI: regions of interest; GLCM: gray-level cooccurrence matrix; GLRLM: gray-level run-length matrix; GLSZM: gray-level size zone matrix; NGTDM: neighboring gray-tone difference matrix; AUC: area under the curve; Rad-score: radiomics score; MTD: maximum tumor diameter; DCA: Decision curve analysis; HR: hazard ratio; RTOG: Radiation Therapy Oncology Group; CPFs: conventional prognostic factors.

**Declarations**

**Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Availability of data and materials**

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

**Competing interests**

The authors declare that they have no competing interests.

**Funding**

This work was financially supported by grants from the National Key Research and Development Project (2018YFC1313200), the National Natural Science Foundation of China (81572970), the Shandong
Natural Science Foundation (ZR2019LZL019), and the Jinan Scientific and Technology Development Project (201805005).

Authors’ contributions

LX and XS conceived of the present idea. XC and QQ delineated and confirmed the ROI. YC acquired the patient date. XT analyzed the date and drafted the manuscript. LX and XS edited and corrected the manuscript. All authors read and approved the final manuscript.

Acknowledgements

None.

References

1. Eberhardt We, De Ruysscher D, Weder W, et al. 2nd ESMO Consensus Conference in Lung Cancer: locally advanced stage III non-small-cell lung cancer. Ann. Oncol. 2015;26(8):1573-88.

2. Coroller T P, Grossmann P, Hou Y, et al. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. Radiother Oncol. 2015;114(3):345-50.

3. Schanne D H, Heitmann J, Guckenberger M, et al. Evolution of treatment strategies for oligometastatic NSCLC patients - A systematic review of the literature. Cancer Treat Rev. 2019;80:101892.

4. Parikh R B, Cronin A M, Kozono D E, et al. Definitive primary therapy in patients presenting with oligometastatic non-small cell lung cancer. Int J Radiat Oncol Biol Phys. 2014;89(4):880-7.

5. Petrelli F, Ghidini A, Cabiddu M, et al. Addition of radiotherapy to the primary tumour in oligometastatic NSCLC: A systematic review and meta-analysis. Lung Cancer. 2018;126:194-200.

6. Tree A C, Khoo V S, Eeles R A, et al. Stereotactic body radiotherapy for oligometastases. The Lancet Oncology. 2013;14(1):e28-e37.

7. Ashworth A, Rodrigues G, Boldt G, et al. Is there an oligometastatic state in non-small cell lung cancer? A systematic review of the literature. Lung Cancer. 2013;82(2):197-203.

8. Palma D A, Salama J K, Lo S S, et al. The oligometastatic state - separating truth from wishful thinking. Nat Rev Clin Oncol. 2014;11(9):549-57.

9. Hellman S, Weichselbaum R R. Oligometastases. J Clin Oncol. 1995;13(1):8-10.

10. Palma D A, Olson R, Harrow S, et al. Stereotactic ablative radiotherapy versus standard of care palliative treatment in patients with oligometastatic cancers (SABR-COMET): a randomised, phase 2, open-label trial. The Lancet. 2019;393(10185):2051-2058.

11. Gomez D R, Tang C, Zhang J, et al. Local Consolidative Therapy Vs. Maintenance Therapy or Observation for Patients With Oligometastatic Non-Small-Cell Lung Cancer: Long-Term Results of a Multi-Institutional, Phase II, Randomized Study. J Clin Oncol. 2019;37(18):1558-1565.
12. Iyengar P, Wardak Z, Gerber D E, et al. Consolidative Radiotherapy for Limited Metastatic Non-Small-Cell Lung Cancer: A Phase 2 Randomized Clinical Trial. JAMA Oncol. 2018;4(1):e173501.

13. Fleckenstein J, Petroff A, Schafers H J, et al. Long-term outcomes in radically treated synchronous vs. metachronous oligometastatic non-small-cell lung cancer. BMC Cancer. 2016;16: 348.

14. Buglione M, Jereczek-Fossa B A, Bonu M L, et al. Radiosurgery and fractionated stereotactic radiotherapy in oligometastatic/ oligoprogresive non-small cell lung cancer patients: Results of a multi-institutional series of 198 patients treated with "curative" intent. Lung Cancer. 2020;141:1-8.

15. Pfannschmidt J, Dienemann H. Surgical treatment of oligometastatic non-small cell lung cancer. Lung Cancer. 2010;69(3):251-8.

16. Lussier Y A, Xing H R, Salama J K, et al. MicroRNA expression characterizes oligometastasis(es). PLoS One. 2011;6(12):e28650.

17. Lussier Y A, Khodarev N N, Regan K, et al. Oligo- and polymetastatic progression in lung metastasis(es) patients is associated with specific microRNAs. PLoS One. 2012;7(12):e50141.

18. Desouza N M, Liu Y, Chiti A, et al. Strategies and technical challenges for imaging oligometastatic disease: Recommendations from the European Organisation for Research and Treatment of Cancer imaging group. Eur J Cancer. 2018;91:153-163.

19. Postmus P E, Kerr K M, Oudkerk M, et al. Early and locally advanced non-small-cell lung cancer (NSCLC): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of Oncology. 2017;28:iv1-iv21.

20. Friedman D V, Tucker S L, Zhou S, et al. Prognostic value and reproducibility of pretreatment CT texture features in stage III non-small cell lung cancer. Int J Radiat Oncol Biol Phys. 2014;90(4): 834-42.

21. Kumar V, Gu Y, Basu S, et al. Radiomics: the process and the challenges. Magn Reson Imaging. 2012;30(9):1234-48.

22. 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):441-6.

23. Coroller T P, Agrawal V, Huynh E, et al. Radiomic-Based Pathological Response Prediction from Primary Tumors and Lymph Nodes in NSCLC. J Thorac Oncol. 2017;12(3):467-476.

24. Coroller T P, Agrawal V, Narayan V, et al. Radiomic phenotype features predict pathological response in non-small cell lung cancer. Radiother Oncol. 2016;119(3):480-6.

25. Huynh E, Coroller T P, Narayan V, et al. CT-based radiomic analysis of stereotactic body radiation therapy patients with lung cancer. Radiother Oncol. 2016;120(2):258-66.

26. Fedorov A, Beichel R, Kalpathy-Cramer J, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magn Reson Imaging. 2012;30(9):1323-41.

27. Velazquez Er, Parmar C, Jermoumi M, et al. Volumetric CT-based segmentation of NSCLC using 3D-Slicer. Sci Rep. 2013;3:3529.

28. Yang M, Ren Y, She Y, et al. Imaging phenotype using radiomics to predict dry pleural dissemination in non-small cell lung cancer. Ann Transl Med. 2019;7(12):259.
29. Gu Y, She Y, Xie D, et al. A Texture Analysis-Based Prediction Model for Lymph Node Metastasis in Stage IA Lung Adenocarcinoma. Ann. Thorac. Surg. 2018;106(1):214-220.
30. Aerts Hj, Velazquez Er, Leijenaar Rt, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun. 2014;5:4006.
31. Tibshirani R. The lasso method for variable selection in the Cox model. Stat Med. 1997;16(4):385-95.
32. Gui J, Li H. Penalized Cox regression analysis in the high-dimensional and low-sample size settings, with applications to microarray gene expression data. Bioinformatics. 2005;21(13):3001-8.
33. Camp Rl, Dolled-Filhart M, Rimm Di, et al. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. Clin. Cancer Res. 2004;10(21):7252-9.
34. Huang Y, Liu Z, He L, et al. Radiomics Signature: A Potential Biomarker for the Prediction of Disease-Free Survival in Early-Stage (I or II) Non-Small Cell Lung Cancer. Radiology. 2016;281(3):947-957.
35. Huang Yq, Liang Ch, He L, et al. Development and Validation of a Radiomics Nomogram for Preoperative Prediction of Lymph Node Metastasis in Colorectal Cancer. J. Clin. Oncol. 2016;34(18):2157-64.
36. Vickers Aj, Cronin Am, Elkin Eb, et al. Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. BMC Med Inform Decis Mak. 2008;8:53.
37. Jensen Gl, Tang C, Hess Kr, et al. Patterns of metastatic progression after definitive radiation therapy for early-stage and locally advanced non-small cell lung cancer. Clin. Exp. Metastasis. 2017;34(5):315-322.
38. Cox Jd, Scott Cb, Byhardt Rw, et al. Addition of chemotherapy to radiation therapy alters failure patterns by cell type within non-small cell carcinoma of lung (NSCCL): analysis of radiation therapy oncology group (RTOG) trials. Int. J. Radiat. Oncol. Biol. Phys. 1999;43(3):505-9.
39. Tang C, Liao Z, Hess K, et al. Prognosis and predictors of site of first metastasis after definitive radiation therapy for non-small cell lung cancer. Acta Oncol. 2016;55(8):1022-8.
40. Nygård L, Vogelius I, Fischer B, et al. A Competing Risk Model of First Failure Site after Definitive Chemoradiation Therapy for Locally Advanced Non-Small Cell Lung Cancer. J Thorac Oncol. 2018;13(4):559-567.
41. Grass Gd, Naghavi Ao, Abuodeh Ya, et al. Analysis of Relapse Events After Definitive Chemoradiotherapy in Locally Advanced Non-Small-Cell Lung Cancer Patients. Clin Lung Cancer. 2019;20(1):e1-e7.
42. Gaspar Le, Chansky K, Albain Ks, et al. Time from treatment to subsequent diagnosis of brain metastases in stage III non-small-cell lung cancer: a retrospective review by the Southwest Oncology Group. J. Clin. Oncol. 2005;23(13):2955-61.
Figure 1

Radiomics features selection using the least absolute shrinkage and selection operator (LASSO) Cox regression model. (a) Tuning parameter ($\lambda$) selection in the LASSO model used 10-fold cross-validation via minimum criteria. The area under the receiver operating characteristic (AUC) curve was plotted versus log ($\lambda$). Dotted vertical lines were drawn at the optimal values by using the minimum criteria and 1 standard error of the minimum criteria (the 1-SE criteria). (a) LASSO coefficients profiles of the 724 radiomics features. A coefficient profile plot was produced against the log ($\lambda$) sequence. A vertical line was drawn at the value selected using 10-fold cross-validation, where optimal $\lambda$ resulted in 12 nonzero coefficients.
Figure 2

Graphs show results of Kaplan-Meier curves according to the radiomics signature for patients in the primary cohort (a) and those in the validation cohort (b).
Figure 3

Use of the constructed radiomics nomogram to estimate the risk of cancer metastases for LA-NSCLC, along with the assessment of the model calibration. (a) Nomogram predicting the risk of distant metastases (left) and oligometastases (right) at 1-year, 2-year and 3-year after definitive chemoradiotherapy in LA-NSCLC. (b) Calibration curves for predicting the risk of metastases-free and oligometastases-free at 1-year, 2-year and 3-year in the primary cohort.
**Figure 4**

Decision curve analysis for each model in the primary cohort (a) and the validation cohort (b).

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.docx