RESEARCH



Whole lung radiomic features are associated with overall survival in patients with locally advanced non-small cell lung cancer treated with definitive radiotherapy

Meng Yan¹, Zhen Zhang^{2,3*}, Jia Tian¹, Jiaqi Yu⁴, Andre Dekker³, Dirk de Ruysscher³, Leonard Wee³ and Lujun Zhao^{1*}

Abstract

Background Several studies have suggested that lung tissue heterogeneity is associated with overall survival (OS) in lung cancer. However, the quantitative relationship between the two remains unknown. The purpose of this study is to investigate the prognostic value of whole lung-based and tumor-based radiomics for OS in LA-NSCLC treated with definitive radiotherapy.

Methods A total of 661 patients with LA-NSCLC treated with definitive radiotherapy in combination with chemotherapy were enrolled in this study, with 292 patients in the training set, 57 patients from the same hospital from January to December 2017 as an independent test set (test-set-1), 83 patients from a multi-institutional prospective clinical trial data set (RTOG0617) as test-set-2, and 229 patients from a Dutch radiotherapy center as test-set-3. Tumor-based radiomic features and whole lung-based radiomic features were extracted from primary tumor and whole lungs (excluding the primary tumor) delineations in planning CT images. Feature selection of radiomic features was done by the least absolute shrinkage (LASSO) method embedded with a Cox proportional hazards (CPH) model with 5-fold cross-internal validation, with 1000 bootstrap samples. Radiomics prognostic scores (RS) were calculated by CPH regression based on selected features. Three models based on a tumor RS, and a lung RS separately and their combinations were constructed. The Harrell concordance index (C-index) and calibration curves were used to evaluate the discrimination and calibration performance. Patients were stratified into high and low risk groups based on median RS, and a log-rank test was performed.

Results The discrimination ability of lung- and tumor-based radiomics model was similar in terms of C-index, 0.69 vs. 0.68 in training set, 0.68 vs. 0.66 in test-set-1, 0.61 vs. 0.62 in test-set-2, 0.65 vs. 0.64 in test-set-3. The combination of tumor- and lung-based radiomics model performed best, with C-index of 0.71 in training set, 0.70 in test-set-1, 0.69 in test-set-2, and 0.68 in test-set-3. The calibration curve showed good agreement between predicted values and actual

*Correspondence: Zhen Zhang zhen.zhang@maastro.nl; radiologyzhangzhen@gmail.com Lujun Zhao zhaolujun@tjmuch.com

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article's Creative Commons licence and indicate otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

values. Patients were well stratified in training set, test-set-1 and test-set-3. In test-set-2, it was only whole lung-based RS that could stratify patients well and tumor-based RS performed bad.

Conclusion Lung- and tumor-based radiomic features have the power to predict OS in LA-NSCLC. The combination of tumor- and lung-based radiomic features can achieve optimal performance.

Keywords Radiomics, Radiotherapy, Lung cancer, Overall survival

Introduction

Identification of biomarkers to predict treatment outcomes is essential to achieve personalized treatment. Several studies have demonstrated that radiomics extracted from LA-NSCLC (locally advanced non-small-cell lung cancer) can predict prognosis, such as locoregional failure [1], distant metastasis [2], progression-free survival [3] and overall survival (OS) [4]. Throughout these studies reported in literature, the predictors/models were built based exclusively on tumor features. However, prognosis of NSCLC might not only depend on the tumor but also on the tumor micro-environment and on the host, and thus the interplay between these factors may have a significant role in treatment and prognosis [5–8]. Different cells derived from different organs and tissues may contribute to the heterogeneity of the tumor environment, which is thus related to prognosis of NSCLC [9]. In clinical studies, survival of NSCLC patients are associated with both intrinsic tumor aggressiveness and the environment into which the tumor grows [10]. Several studies have found that coexisting chronic obstructive pulmonary diseases (COPD) was associated with worse survival in patients with NSCLC [11-14] and coexistence of interstitial lung disease and NSCLC was also independent risk factors for shorter survival of NSCLC [15-18]. These studies suggest that the heterogeneity of lung tissue is associated with OS of NSCLC. Shuo W et al. [19] established a fully automated artificial intelligence system to mine whole-lung information from CT images to predict EGFR genotype and prognosis with EGFR-TKI treatment, and its performance was overall better than tumor-only based deep learning methods. This study demonstrated the potential of artificial intelligence to decode the lung tissue phenotype. Based on these studies, we hypothesized that whole-lung radiomic features may be correlated with OS of LA-NSCLC.

In this study, we extracted radiomic features of tumor and lung regions from the planning CT images of LA-NSCLC patients treated with definitive radiotherapy in combination with chemotherapy and developed OS prediction models. The predictive value based on tumor and lung tissue alone as well as the combination of both was investigated. Three independent validation sets with a large heterogeneity of patient were used to validate the performance of the proposed models. In addition, patients were stratified into different risk groups based on the radiomic score to assess the performance of the models.

Methods

Patients and data sets

This retrospective observational study is a Transparent Reporting of a multivariate prediction model for Individual Prognosis Or Diagnosis (TRIPOD) type 3 investigation. This study was approved by our institution Ethics Committee (IRB/bc2021135) and confirmation of informed consent by patients was waived. The workflow of the study is shown in Fig. 1. We retrieved 349 LA-NSCLC patients who underwent definitive radiotherapy in combination with chemotherapy from January 2015 to June 2020 in Tianjin Medical University Cancer Institute and Hospital, in which patients from January 2015 to December 2016 and January 2018 to June 2020 served as training set (292 patients) and patients from January to December 2017 formed the independent test-set-1 (57 patients). More details on the criteria of inclusion and exclusion are given in Supplementary Materials A.

In addition, a multi-institutional prospective RTOG clinical trial dataset (RTOG0617 [20, 21]) (test-set-2) and a data set from a Dutch radiotherapy center - MAASTRO (test-set-3) [20, 22, 23] were used as external validation. They comprised of 83 patients and 229 patients, respectively. The patient (in external test sets) inclusion criteria is provided in Supplementary Fig. S1.

The endpoint of the study was OS, which is defined as the time from the date of the final radiotherapy fraction until death for any cause or lost follow-up (right censor).

CT acquisition and ROI segmentation

All planning CT scans were acquired on a Brilliant (Philips Medical Systems; Best, The Netherlands) scanner in the training set and test-set-1. The scan parameters were set as follows: slice thickness 2–5 mm, tube voltage 120 kV, tube current 100 mAs, 512×512 pixels, vendor's default convolution kernel was used for reconstruction. To reduce the influence of slice thickness on feature extraction, all planning CT images were isotopically resampled as 1 mm cubes with linear interpolation before segmentation of regions of interest (ROIs). Two ROIs (primary tumor and whole lung excluding primary tumor) were segmented. Before radiotherapy, a junior physician manually delineated the primary tumor



Fig. 1 Workflow of the project. ROI segmentation: tumor region and whole-lung region segmented; Feature selection: 1000 unique bootstrap samples taken from all samples, features selected by correlation, least absolute shrinkage embedded with Cox proportional hazards model and Akaike information criterion for modeling; Model construction: tumor radiomics signature and lung radiomics signature for model construction; Model performance: model performance evaluated using discrimination, calibration and Kaplan-Meier analysis. Abbreviations: ROI: region of interest; Lasso-Cox: least absolute shrinkage embedded with Cox proportional hazards model; AIC: Akaike information criterion; T-RM: Cox proportional hazards model based on tumor signature; L-RM: Cox proportional hazards model based on lung signature; TL-RM: Cox proportional hazards model based on tumor signature; C-index: Harrell concordance index

and a senior radiation oncologist modified and reviewed this segmentation in the Pinnacle TPS (Philips Radiation Oncology Systems; Fitchburg, Wisconsin, United States), with image fusion against complementary imaging studies when possible (such as positron emission tomography). For the segmentation of lungs, in order to improve the efficiency and consistency of segmentation, a retrained deep-learning automatic contour model was used for lung segmentation based on a published model [24]. Then manual editing was performed by a radiation oncologists (author M.Y.), and the other experienced radiation oncologists (author Z.Z) independently reviewed the lung segmentations using 3D slicer software [25]. More details of the segmentation are available at Supplementary Material B.

Radiomic features extraction and feature selection

In total 824 radiomic features were extracted based on Pyradiomics v3.7, covering 17 intensity histogram features, 14 shape features, 73 textural features and 711 wavelet radiomic features. More details about the radiomic features extraction are described in previous publications [26, 27]. The feature extraction parameters are given in Supplementary Materials C. As shown in the Fig. 1, two steps were adopted for selection of radiomic features. First, pair-wise feature correlation reduction; we used the R caret library to recursively reduce the total number of pair-wise Pearson correlation between any two features with Spearman correlation in excess of 0.9. Then the radiomic features were screened using least absolute shrinkage (LASSO) embedded with COX proportional hazards model (COX) with 5-fold cross-internal validation by the R glmnet package, with 1000 unique bootstrap samples from the whole training set. From each of the 1000 bootstraps, we ranked each individual feature according to how frequently the feature was retained by the LASSO-COX. Then, we selected a cut-off frequency for the top most frequently appearing individual features

from the ranking table. Secondly, the radiomic features selected over the cut-off point were imported into the "feature pool" for further selection based on COX and stepwise backward Akaike information criterion (AIC) for the same 1000 bootstrap samples as in step 1.

Finally, the radiomic feature combination (signature) that ranked first in frequency was selected for model building. Prior to modeling, radiomic features were normalized by the Z-score. The same feature selection approach was used for tumor- and lung-based radiomics signatures.

Construction of the models

The tumor radiomic signature-based COX regression model (T-RM) was developed and the tumor radiomic prognostic score (T-RPS) was calculated. The same approach was performed for the COX regression model based on lung signature (L-RM) and lung radiomics prognostic score (L-RPS). A combined tumor and lungbased model (TL-RM) was constructed using T-RPS and L-RPS as covariates. The prognostic score (TL-RPS) was calculated using a Cox regression model, with weights for the covariates determined by the model's coefficients estimated during the fitting process.

Statistical analysis

All statistical analyses were executed in R software (V4.4.3, https://www.R-project.org/) and SPSS27. Patients' baseline differences between training set and testing sets were analyzed by exact Fisher test for categorical variables, A two-side hypothesis test was applied, and a p-value less than 0.05 was considered statistically significant.

The Harrell concordance index (C-index) was used for evaluating the discrimination of models. Goodness of fit (the degree of concordance between the predicted and observed values) was assessed by Calibration Curve for 2 years (730days) with 1000 bootstrap resamples using the R rms package.

Based on the prognostic score, the fixed cut-off points, the median scores of training set in three prognostic scores (T-RPS, L-RPS, TL-RPS), were used in the test sets. The different risk groups were stratified by cut-off points. The Kaplan-Meier (K-M) method was used to calculate survival rates, and log-rank test was applied to compare survival distribution between different risk groups.

Results

Patient characteristics

Patients baseline characteristics are reported in Table 1 and a comparative analysis was conducted on the baseline characteristics among the three data sets. The median age of the training set, test-set-1 and test-set-2 was 61.5 years (range from 21 to 88 years), 53 years (range from 44 to 75 years) and 65 years (range from 39 to 82 years), respectively. In the whole cohort, a total of 433 patients reached the endpoint. Median follow-up time was 23.4 months (range of 1.4–85.8months). More details are listed in Supplementary Table S1.

Radiomic feature selection and construction of models

After feature selection, a tumor-based radiomics signature including 11 radiomic features and a lung-based radiomics signature including 8 radiomic features were constructed. More details and graphs of radiomic features selection together with the selected features for model building are provided in Supplementary Materials D.

The T-RPS and the L-RPS were calculated based on the coefficients weighted by COX. Similarly, TL-RPS were also calculated. The formulae for the construction of T-RPS, L-RPS, TL-RPS are provided in Supplementary Materials E. In the integrated model (TL-RM), the weights of the T-RPS, L-RPS was 0.801897 and 0.81187, respectively.

Performance of the models

The discrimination results are shown in Table 2. The integrated model TL-RM, which combined T-RPS and L-RPS, showed the best prediction power. The results of the tumor-based model (T-RM) and the lung-based model (L-RM) are similar (C-index 0.68 vs. 0.69 in the training set, 0.66 vs. 0.68 in the test-set-1, 0.62 vs. 0.61 in the test-set-2, 0.64 vs. 0.65 in the test-set-3). The calibration curve of 2 years OS with 1000 bootstrap resampling in training set is displayed in Supplementary Fig. S3, which shows good consistency between the predicted probabilities of 2 years OS.

The Kaplan-Meier survival curves of T-RM, L-RM and TL-RM are provided in Fig. 2. The fixed cut-off points (the median scores of the training set) for T-RM, L-RM and TL-RM were - 321.21, -52.96 and - 300.66, respectively. The patients were divided into high-risk group and low-risk group based on the models. As shown in Fig. 2, good stratification was observed in the T-RM, L-RM and TL-RM in the training set (log-rank test, T-RM: p < 0.001, L-RM: p < 0.001, TL-RM: p < 0.001) and test-set-3 (logrank test, T-RM: *p* < 0.001, L-RM: *p* = 0.0015, TL-RM: p < 0.001). In test-set-1, T-RM (log-rank test, p = 0.042), L-RM (log-rank test, p = 0.035) and TL-RM (log-rank test, p = 0.0027) showed significant stratification. In test-set-2, while the differences between the high-risk and low-risk groups in the T-RM (log-rank test, p = 0.27) and TL-RM (log-rank test, p = 0.76) were not statistically significant, the stratification of the L-RM showed a significant difference (log-rank test, p = 0.013) (Fig. 2). When cutoff points

Characteristic	All patients N=661	Training set N=292	Test-set = 1 N = 57	Test-set-2 N=83	Test-set-3 N=229	P-value
Gender						< 0.001*
Male	482(72.9%)	232(79.5%)	47(82.5%)	50(60.2%)	153(66.8%)	
Female	179(27.1%)	60(20.5%)	10(17.5%)	33(39.8%)	76(33.2%)	
Age						< 0.001*
≤60	257(38.9%)	135(46.2%)	41(71.9%)	27(32.5%)	54(23.6%)	
>60	400(60.5%)	157(53.8%)	16(28.1%)	56(67.5%)	171(74.7%)	
NA	4(0.6%)	0(0%)	0(0%)	0(0%)	4(1.7%)	
Histology						< 0.001*
SCC	336(50.8%)	159(54.5%)	41(71.9%)	30(36.1%)	106(46.3%)	
Non-SCC	325(49.2%)	133(45.5%)	16(28.1%)	53(63.9%)	123(53.7%)	
Clinical Stage						< 0.001*
IIIA	269(40.7%)	95(32.5%)	19(33.3%)	41(49.4%)	114(49.8%)	
IIIB	284(43.0%)	135(46.2%)	30(52.6%)	32(38.6%)	87(38.0%)	
IIIC	108(16.3%)	62(21.2%)	8(14.0%)	10(12.0%)	28(12.2%)	
Tumor						0.039*
location						
Central	286(66.2%)	189(64.7%)	46(80.7%)	51(61.4%)		
Peripheral	146(33.8%)	103(35.3%)	11(19.3%)	32(38.6%)		
Smoking						0.001*
No	74(18.2%)	63(21.6%)	6(10.5%)	5(6.0%)		
Yes	358(81.8%)	229(78.4%)	51(89.5%)	78(94.0%)		
CCRT						< 0.001*
No	236(54.6%)	194(66.4%)	42(73.7%)	0(0%)		
Yes	196(45.4%)	98(33.6%)	15(26.3%)	83(100%)		

Table 1 Patient characteristics

Abbreviations NA = Null value; SCC = squamous cell carcinoma; CCRT = concurrent chemoradiotherapy

*The differences in characteristics were evaluated by exact Fisher test for categorical variables

 Table 2
 C-index of all models in different sets

Model	Training set	Test-set-1	Test-set-2	Test-set-3	
	C-index	C-index	C-index	C-index	
	(95%Cl)	(95%Cl)	(95%Cl)	(95%Cl)	
T-RM	0.68	0.66	0.62	0.64	
	(0.63–0.72)	(0.60–0.73)	(0.54–0.78)	(0.59–0.68)	
L-RM	0.69	0.68	0.61	0.65	
	(0.64–0.73)	(0.62–0.75)	(0.48–0.74)	(0.55–0.74)	
TL-RM	0.71	0.70	0.69	0.68	
	(0.66–0.77)	(0.68–0.82)	(0.58–0.84)	(0.62–0.76)	

Abbreviations T-RM: COX proportional hazards model based on tumor signature; L-RM: COX proportional hazards model based on lung signature; TL-RM: COX proportional hazards model based on tumor signature and lung signature

were determined based on the optimal cutoffs of the training set, good stratification was observed in TL-RM (log-rank test, p = 0.012), (Supplementary Fig. S4).

As shown in Supplementary Table S4, the tumor location, and T stage were different between the high- and low-risk groups that stratified by T-RM. For the different risk groups based on L-RM, N stage and smoking showed statistical significance (Supplementary Table S4).

Discussion

Identification of patients with a poor prognosis and short OS is a vital clinical need on the basis of which appropriate treatment strategies can be adopted. In this study, we demonstrated that both tumor-based and lung-based radiomic features can predict OS of LA-NSCLC patients receiving definitive radiotherapy. In addition, the integrated model, combining tumor-based and lung-based radiomics signatures, can achieve an even higher predictive power.

Compared to the previous study [28], we employed a more rigorous and transparent method to select radiomic features and test the whole-lung radiomics model across three independent datasets. We believe that our approach enhances the reproducibility and generalizability of our findings. To explore the predictive power of lung radiomic features, the lung radiomics model was built and validated in three independent test sets, and patients were stratified into different risk group based on L-RM. This study showed that the predictive power of lungbased radiomics signature was similar to tumor-based radiomics signature in terms of C-index and risk group stratification. This may suggest to us that the predictive ability of latent information in the lungs is comparable to that of tumors. Moreover, according to the coefficients in



Fig. 2 Kaplan-Meier survival curves for the tumor radiomics model (A), lung radiomics model (B), and combined tumor and lung radiomics model (C) in the training set and three distinct test sets (test-set-1, test-set-2, and test-set-3)

tumor and lung signature combination model (TL-RSM), the importance of lung- and tumor-based radiomics signature was very close (0.81 vs. 0.80), which reminds us that lung radiomic features should not be neglected when predicting OS in LA-NSCLC.

We found that L-RM retains a good discrimination power across each of the training and test data sets (Fig. 2B). However, the stratification of T-RM and TL-R in test-set-2 performed poorly (Fig. 2A.2 C). This might be due to the fact that fewer patients were included in test-set-2. Also, this is a multi-institution North-American trial data set that might show greater heterogeneity in patient and data acquisition characteristics (i.e., patients are from different institutions). Also the lung cancer susceptibility genes and genomes are thought to exhibit variations between Chinese and Caucasian populations [29]. The training set consisted solely of Chinese patients, while the test-set-2 dataset included 84.3% Caucasian patients (Supplementary Table S1b). Even though, test-set-3 also consisted of almost exclusively Caucasian patients, it performed well perhaps because of relatively large sample size, while test-set-1 performed well perhaps because of it originated from the same data sources of training set. In addition, although we performed preprocessing such as resampling and normalization before extracting radiomic features to harmonize different datasets, there may still be image level differences due to scanning protocols, construction kernels, etc. Further indepth algorithms still need to be developed.

Other studies have demonstrated that pre- and postradiotherapy CT images of both have predictive ability for OS in LA-NSCLC [30–35]. In these studies, the radiomic features were derived from the primary tumor, while lung radiomic features may change after radiotherapy [36]. The main objective of this study was to explore the corrections between lung radiomic features and OS, so only pre-radiotherapy CT images were used in this study. In dataset division, selecting patients from January to December 2017 as test-set-1 minimizes enddata extreme value bias, ensures adequate follow-up, and prevents overfitting. Moreover, this period represents the IMRT to VMAT transition in the center, enhancing model generalization assessment.

Several steps were performed in this study to improve the robustness and reproducibility of the radiomic features. First, all CT images were reconstructed to the same slice thickness (1 mm), and segmentation was performed in the reconstructed images. Secondly, in terms of segmentation, manual-segmentation is considered time consuming, unstable and poorly repeatable [37]. In this study, a combination of automatic and manual segmentation method was applied to segment the lungs. Lastly, LASSO-COX was used for the selection of the radiomic features, which was suitable for high-dimensional radiomic features to reduce the effects of overfitting in small-sized data, and 1000 bootstrap method was used to reduce the risk of over-optimistic results.

Among the features selected for the tumor radiomics signature, nine radiomic features were wavelet-based

features. The wavelet filtered radiomic features play a key role in the tumor-based radiomic signature. Previous studies have found similar results [23, 38, 39], suggesting that wavelet-based features reflect tumor heterogeneity well and are closely related to OS in NSCLC. Chen [40] et al. established the radiomic signature for predicting OS of NSCLC, which included the same GLDM features we found in our study.

When comparing clinical characteristics between the high- and low- risk groups defined by the radiomics prognostic score, we noted that patients with centrally located tumors had high tumor-based prognostic scores and smokers had higher lung-based prognostic scores (Supplementary Table S4). The distribution of pathological types differs between central and peripheral lung cancer, and radiomics may have the potential to discriminate between these distinct subtypes [41, 42]. Moreover, even within the same pathological type, variations in clinicopathological factors exist between central and peripheral lung cancer [43, 44], which may contribute to different prognostic scores observed in these two subgroups. Weeden [45] et al. found that a specific T cells were highly enriched in smokers and there was a different inflammatory environment between the lungs of smokers and the lungs of non-smokers, which may explain the different lung-based prognostic scores between smokers and non-smokers. This suggests to us that radiomics may reflect tumor and lung conditions and thus predict OS.

Both peri-tumoral and whole-lung radiomic features exhibit prognostic value, reflecting distinct aspects of the disease. Previous studies [46–48] indicate that peritumoral features are associated with prognosis, likely due to tumor invasiveness and local microenvironment changes, whereas whole-lung features may reveal systemic pathological processes such as inflammation or interstitial lung disease. Future research should focus on integrating information from multiple regions to develop a comprehensive model, elucidating the interplay between local tumor characteristics and systemic lung pathology.

There are several limitations in the present study. Firstly, this study is retrospective and prospective data will be needed for future study to eliminate selection bias. Prospective studies also enable standardization of follow-up to ensure accuracy of endpoints. Secondly, most of the patients in the study were treated before the NCCN revised its guidelines based on the results of the PACIFIC trial [49], which was one of the most revolutionary updates in the treatment of LA-NSCLC, where immunotherapy was used for patients who did not progress after radiation therapy. However, immunotherapy is presently unaffordable for many patients. For these reasons, consolidation immunotherapy has not been routinely performed in China, which may have affected OS. Even in the immunotherapy era, identifying patients who are refractory to chemoradiotherapy is useful, for these patients should not be treated with this toxic combination. In the future, we need to conduct a radiomics study of OS prediction in patients treated with curative radiotherapy and consolidation immunotherapy. Thirdly, more attention is growing on the use of multi-omics, which has been proven to improve the performance of models [50]. Through analysis of genomics, radiomics may be able to be represented at the biological level. In future studies, clinical factors, dosiomics and genomics should be integrated. Fourthly, while this study has encouraging results, this is only a pilot study and further refinement is still needed to obtain better performance for clinical use.

Conclusion

A TRIPOD type 3 prediction model was developed and validated using external validation data. The selected lung radiomic signature has the power to predict OS in patients with LA-NSCLC treated with definitive radiotherapy in combination with chemotherapy. The model combining tumor radiomic features, and lung radiomic features has the best predictive power and can stratify patients well into different risk groups.

Supplementary Information

The online version contains supplementary material available at https://doi.or g/10.1186/s13014-025-02583-1.

Supplementary Material 1

Acknowledgements

This manuscript was prepared using data from datasets (RTOG-0617; NCT00533949-D1, D2, D3) from the NCTN/NCORP Data Archive of the National Cancer Institute's (NCI's) National Clinical Trials Network (NCTN). Data were originally collected from a clinical trial (identifier NCT00533949; "A Randomized Phase III Comparison of Standard-Dose (60 Gy) Versus High-Dose (74 Gy) Conformal Radiotherapy with Concurrent and Consolidation Carboplatin/Paclitaxel +/- Cetuximab (IND #103444) in Patients With Stage IIIA/IIIB Non-Small Cell Lung Cancer"). All analyses and conclusions in this manuscript are the sole responsibility of the authors and do not necessarily reflect the opinions or views of the clinical trial investigators, the NCTN, the NCORP or the NCI.

Author contributions

Meng Yan was responsible for conducting the experiments, collecting and organizing the data, and contributed to the writing of the initial draft.Jia Tian and Jiaqi Yu was responsible for collecting and organizing the data.Andre Dekker, Dirk de Ruysscher and Leonard wee provided the partial data for the experiments, participated in the design of the experiments, and assisted in revising the manuscript.Lujun Zhao and Zhen Zhang, as the corresponding author, was responsible for the conception of the entire research project, guided all aspects of the study, and participated in the finalization of the manuscript.All authors have reviewed and agreed to the final version of the manuscript.

Funding

This study was funded by the National Natural Science Foundation of China (no. 82303672) and Tianjin Key Medical Discipline (Specialty) Construction Project (No. TJYXZDXK-009 A).

Data availability

Research data are stored in an institutional repository and will be shared upon request to the corresponding author.

Declarations

Ethics approval

This study was approved by Tianjin Medical University Cancer Institute & Hospital Ethics Committee (IRB/bc2021135). This study is a retrospective analysis that does not involve any organizations, direct patient contact, or the disclosure of patient personal privacy. The research will not provide diagnostic reports to patients, and the experimental results are intended solely for comparative purposes and will not be used as clinical evidence. The informed consents were waived after review and approval by the ethics committee.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Radiation Oncology, Key Laboratory of Cancer Prevention and Therapy, Tianjin Medical University Cancer Institute & Hospital, National Clinical Research Center for Cancer, Tianjin's Clinical Research Center for Cancer, Tianjin 300060, China

²Zhejiang Cancer Hospital, Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences, Hangzhou, Zhejiang 310022, China

³Department of Radiation Oncology (Maastro), GROW Research Institute for Oncology and Reproduction, Maastricht University Medical Centre+, Maastricht, The Netherlands

⁴Department of Radiation Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China

Received: 30 April 2024 / Accepted: 3 January 2025 Published online: 17 January 2025

References

- Chen X, Tong X, Qiu Q, et al. Radiomics Nomogram for Predicting Locoregional failure in locally Advanced Non-small Cell Lung Cancer treated with definitive chemoradiotherapy. Acad Radiol. 2022;29(Suppl 2):S53–61. https:// doi.org/10.1016/j.acra.2020.11.018.
- Tao J, Lv R, Liang C, et al. Development and validation of a CT-Based signature for the prediction of distant metastasis before treatment of Non-small Cell Lung Cancer. Acad Radiol. 2022;29:S62–72. https://doi.org/10.1016/j.acra.202 0.12.007.
- Zhang N, Liang R, Gensheimer MF, et al. Early response evaluation using primary tumor and nodal imaging features to predict progression-free survival of locally advanced non-small cell lung cancer. Theranostics. 2020;10(25):11707–18. https://doi.org/10.7150/thno.50565.
- Luna JM, Barsky AR, Shinohara RT, et al. Radiomic phenotypes for improving early prediction of Survival in Stage III Non-small Cell Lung Cancer Adenocarcinoma after Chemoradiation. Cancers (Basel). 2022;14(3):700. https://doi.org /10.3390/cancers14030700.
- Wong KY, Cheung AHK, Chen B, et al. Cancer-associated fibroblasts in nonsmall cell lung cancer: from molecular mechanisms to clinical implications. Int J Cancer. 2022;151(8):1195–215. https://doi.org/10.1002/ijc.34127.
- Wen Y, Guo G, Yang L, et al. A tumor microenvironment gene set-based prognostic signature for non-small-cell lung cancer. Front Mol Biosci. 2022;9:849108. https://doi.org/10.3389/fmolb.2022.849108.
- Wood SL, Pernemalm M, Crosbie PA, Whetton AD. The role of the tumormicroenvironment in lung cancer-metastasis and its relationship to potential therapeutic targets. Cancer Treat Rev. 2014;40(4):558–66. https://doi.org/10.1 016/j.ctrv.2013.10.001.
- Lc D. The tumor organismal environment: role in tumor development and cancer immunotherapy. Sem Cancer Biol. 2020;65. https://doi.org/10.1016/j.s emcancer.2019.12.021.
- Suzuki J, Tsuboi M, Ishii G. Cancer-associated fibroblasts and the tumor microenvironment in non-small cell lung cancer. Expert Rev Anticancer Ther. 2022;22(2):169–82. https://doi.org/10.1080/14737140.2022.2019018.

- Choi N, Baumann M, Flentjie M, et al. Predictive factors in radiotherapy for non-small cell lung cancer: present status. Lung Cancer. 2001;31(1):43–56. https://doi.org/10.1016/s0169-5002(00)00156-2.
- Zhai R, Yu X, Shafer A, Wain JC, Christiani DC. The impact of coexisting COPD on survival of patients with early-stage non-small cell lung cancer undergoing surgical resection. Chest. 2014;145(2):346–53. https://doi.org/10.1378/che st.13-1176.
- 12. Lim JU, Yeo CD, Rhee CK, et al. Overall survival of driver mutation-negative non-small cell lung cancer patients with COPD under chemotherapy compared to non-COPD non-small cell lung cancer patients. Int J Chron Obstruct Pulmon Dis. 2018;13:2139–46. https://doi.org/10.2147/COPD.S167372.
- Wang P, Zhu M, Zhang D, et al. The relationship between chronic obstructive pulmonary disease and non-small cell lung cancer in the elderly. Cancer Med. 2019;8(9):4124–34. https://doi.org/10.1002/cam4.2333.
- Dong W, Zhu Y, Du Y, Wang L, Feng X, Ma S. Impact of severe-to-very severe chronic obstructive pulmonary disease on the prognosis of patients with non-small cell lung cancer who received chemotherapy. Clin Respir J. 2020;14(4):345–52. https://doi.org/10.1111/crj.13139.
- Schussler O, Bobbio A, Dermine H, et al. Twenty-year survival of patients operated on for non-small-cell Lung Cancer: the impact of Tumor Stage and patient-related parameters. Cancers (Basel). 2022;14(4):874. https://doi.org/10 .3390/cancers14040874.
- Ueda T, Aokage K, Mimaki S, et al. Characterization of the tumor immunemicroenvironment of lung adenocarcinoma associated with usual interstitial pneumonia. Lung Cancer. 2018;126:162–9. https://doi.org/10.1016/j.lungcan. 2018.11.006.
- Lee SJ, Lee J, Park YS, et al. Impact of chronic obstructive pulmonary disease on the mortality of patients with non-small-cell lung cancer. J Thorac Oncol. 2014;9(6):812–7. https://doi.org/10.1097/JTO.000000000000158.
- Sato T, Watanabe A, Kondo H, et al. Long-term results and predictors of survival after surgical resection of patients with lung cancer and interstitial lung diseases. J Thorac Cardiovasc Surg. 2015;149(1):64–9. https://doi.org/10.1016/j.jtcvs.2014.08.086. 2.
- Wang S, Yu H, Gan Y, et al. Mining whole-lung information by artificial intelligence for predicting EGFR genotype and targeted therapy response in lung cancer: a multicohort study. Lancet Digit Health. 2022;4(5):e309–19. https://d oi.org/10.1016/S2589-7500(22)00024-3.
- Clark K, Vendt B, Smith K, et al. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. J Digit Imaging. 2013;26(6):1045–57. https://doi.org/10.1007/s10278-013-9622-7.
- 21. Bradley JD, Forster K. Data from NSCLC-Cetuximab. *The Cancer Imaging Archive*. Published online. 2018. https://doi.org/10.7937/TCIA.2018.jze75u7v
- Aerts HJWL, Wee L, Rios Velazquez E et al. Data From NSCLC-Radiomics (version 4) [Data set]. *The Cancer Imaging Archive*. Published online. 2014. https://doi.org/10.7937/K9/TCIA.2015.PF0M9REI
- 23. Aerts HJWL, Velazquez ER, Leijenaar RTH, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun. 2014;5:4006. https://doi.org/10.1038/ncomms5006.
- Hofmanninger J, Prayer F, Pan J, Röhrich S, Prosch H, Langs G. Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem. Eur Radiol Experimental. 2020;4(1):50. https://doi.org/ 10.1186/s41747-020-00173-2.
- 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. https://doi.org/10.1016/j.mri.2012.05.001.
- Shi Z, Traverso A, Soest J, Dekker A, Wee L. Technical note: ontology-guided radiomics analysis workflow (O-RAW). Med Phys. 2019;46(12):5677–84. https:/ /doi.org/10.1002/mp.13844.
- Zhang Z, Wang Z, Yan M et al. Radiomics and dosiomics signature from whole lung predicts radiation pneumonitis: a model development study with prospective external validation and decision-curve analysis. Int J Radiation Oncology*Biology*Physics Published Online August 2022:S0360301622031893. https://doi.org/10.1016/j.ijrobp.2022.08.047
- Chen NB, Xiong M, Zhou R, et al. CT radiomics-based long-term survival prediction for locally advanced non-small cell lung cancer patients treated with concurrent chemoradiotherapy using features from tumor and tumor organismal environment. Radiat Oncol. 2022;17(1):184. https://doi.org/10.118 6/s13014-022-02136-w.
- Shen H, Zhu M, Wang C. Precision oncology of lung cancer: genetic and genomic differences in Chinese population. Npj Precision Oncol. 2019;3(1):1– 8. https://doi.org/10.1038/s41698-019-0086-1.

- Fave X, Zhang L, Yang J, et al. Delta-radiomics features for the prediction of patient outcomes in non-small cell lung cancer. Sci Rep. 2017;7(1):588. https:/ /doi.org/10.1038/s41598-017-00665-z.
- Chen W, Hou X, Hu Y, Huang G, Ye X, Nie S. A deep learning- and CT imagebased prognostic model for the prediction of survival in non-small cell lung cancer. Med Phys. 2021;48(12):7946–58. https://doi.org/10.1002/mp.15302.
- Wang L, Dong T, Xin B, et al. Integrative nomogram of CT imaging, clinical, and hematological features for survival prediction of patients with locally advanced non-small cell lung cancer. Eur Radiol. 2019;29(6):2958–67. https:// doi.org/10.1007/s00330-018-5949-2.
- Le VH, Kha QH, Hung TNK, Le NQK. Risk score generated from CT-Based Radiomics signatures for overall survival prediction in Non-small Cell Lung Cancer. Cancers (Basel). 2021;13(14):3616. https://doi.org/10.3390/cancers131 43616.
- Hou KY, Chen JR, Wang YC, et al. Radiomics-based deep learning prediction of overall survival in Non-small-cell Lung Cancer using contrast-enhanced computed Tomography. Cancers (Basel). 2022;14(15):3798. https://doi.org/10. 3390/cancers14153798.
- 35. Li R, Peng H, Xue T, et al. Prediction and verification of survival in patients with non-small-cell lung cancer based on an integrated radiomics nomogram. Clin Radiol. 2022;77(3):e222–30. https://doi.org/10.1016/j.crad.2021.12.002.
- Wang L, Gao Z, Li C, et al. Computed tomography-based Delta-Radiomics analysis for discriminating Radiation pneumonitis in patients with Esophageal Cancer after Radiation Therapy. Int J Radiat Oncol Biol Phys. 2021;111(2):443–55. https://doi.org/10.1016/j.ijrobp.2021.04.047.
- Liu Z, Wang S, Dong D, et al. The applications of Radiomics in Precision diagnosis and treatment of Oncology: opportunities and challenges. Theranostics. 2019;9(5):1303–22. https://doi.org/10.7150/thno.30309.
- Liu Y, Qi H, Wang C, et al. Predicting Chemo-Radiotherapy Sensitivity with Concordant Survival Benefit in Non-small Cell Lung Cancer via Computed Tomography Derived Radiomic features. Front Oncol. 2022;12:832343. https:/ /doi.org/10.3389/fonc.2022.832343.
- Yang L, Yang J, Zhou X, et al. Development of a radiomics nomogram based on the 2D and 3D CT features to predict the survival of non-small cell lung cancer patients. Eur Radiol. 2019;29(5):2196–206. https://doi.org/10.1007/s00 330-018-5770-y.
- Chen W, Qiao X, Yin S, Zhang X, Xu X. Integrating Radiomics with Genomics for Non-small Cell Lung Cancer Survival Analysis. J Oncol. 2022;2022. https:// doi.org/10.1155/2022/5131170.
- Gharraf HS, Mehana SM, ElNagar MA. Role of CT in differentiation between subtypes of lung cancer; is it possible? Egypt J Bronchol. 2020;14(1):28. https:/ /doi.org/10.1186/s43168-020-00027-w.

- Lu EL, Li L, Yang L, Schwartz H, Zhao LH. Radiomics for classification of Lung Cancer histological subtypes based on Nonenhanced Computed Tomography. Acad Radiol. 2019;26(9):1245–52. https://doi.org/10.1016/j.acra.2018.10.0 13.
- Tomashefski JF, Connors AF, Rosenthal ES, Hsiue IL. Peripheral vs central squamous cell carcinoma of the lung. A comparison of clinical features, histopathology, and survival. Arch Pathol Lab Med. 1990;114(5):468–74.
- Sung YE, Cho U, Lee KY. Peripheral type squamous cell carcinoma of the lung: clinicopathologic characteristics in comparison to the central type. J Pathol Translational Med. 2020;54(4):290–9. https://doi.org/10.4132/jptm.2020.05.04.
- Weeden CE, Gayevskiy V, Marceaux C, et al. Early immune pressure initiated by tissue-resident memory T cells sculpts tumor evolution in non-small cell lung cancer. Cancer Cell. 2023;41(5):837–e8526. https://doi.org/10.1016/j.ccel l.2023.03.019.
- Bove S, Fanizzi A, Fadda F, et al. A CT-based transfer learning approach to predict NSCLC recurrence: the added-value of peritumoral region. PLoS ONE. 2023;18(5):e0285188. https://doi.org/10.1371/journal.pone.0285188.
- Ma Y, Li Q. An integrated model combined intra- and peritumoral regions for predicting chemoradiation response of non small cell lung cancers based on radiomics and deep learning. Cancer/Radiothérapie. 2023;27(8):705–11. https://doi.org/10.1016/j.canrad.2023.05.005.
- Zhang X, Zhang G, Qiu X, et al. Optimizing the size of Peritumoral Region for assessing Non-small Cell Lung Cancer Heterogeneity using Radiomics. In: Li Y, Huang Z, Sharma M, Chen L, Zhou R, editors. Health Information Science. Lecture Notes in Computer Science. Springer Nature; 2023. pp. 309–20. https: //doi.org/10.1007/978-981-99-7108-4_26.
- Spigel DR, Faivre-Finn C, Gray JE, et al. Five-year survival outcomes from the PACIFIC Trial: Durvalumab after Chemoradiotherapy in Stage III Non-small-cell Lung Cancer. J Clin Oncol. 2022;40(12):1301–11. https://doi.org/10.1200/JCO. 21.01308.
- 50. Lam SK, Zhang Y, Zhang J, et al. Multi-organ Omics-based prediction for adaptive Radiation Therapy eligibility in nasopharyngeal carcinoma patients undergoing concurrent Chemoradiotherapy. Front Oncol. 2021;11:792024. https://doi.org/10.3389/fonc.2021.792024.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.