Appendix 1 CT protocol and evaluation

Protocols for CT scanning were established with the following specifications: the tube voltage was set at 120 kVp and the tube current was regulated by automatic exposure control systems. The pitch of the beam ranged from 0.75 to 1.5, while collimation varied between 0.625 and 1.5 mm. The field of view was determined to be 200 by 200 with a pixel matrix configured at 512 by 512. Chest scans without contrast were conducted, encompassing the region from the thoracic inlet down to the adrenal gland while patients were positioned supine and instructed to inhale fully.

The CT findings were evaluated in the lung window (width, 1500 HU; level, -600 HU) and the mediastinal window (width, 400 HU; level, 40 HU). Analysis of the CT findings for each nodule was based on the following parameters: (I) location; (II) pleural indentation; (III) vessel convergence; (IV) margin; (V) lobulated sharp; (VI) spiculated sign; (VII) air bronchogram sign; (VIII) vacuole sign. The nodules' diameter was calculated as the average of their length and width, rounded to the nearest tenth of a millimeter. A regular margin was deemed to be present when the lesion surface showed neither indentations nor irregularities. Any shape not conforming to a round, oval, or polygonal form was categorized as irregular. The lobulated sharp was identified when a section of the lesion's surface exhibited a wavy or scalloped configuration. Spiculated sign was defined as the presence of strands extending from the nodule margin into the lung parenchyma without reaching the pleural surface. The air bronchogram sign is considered to be present when air-filled bronchi are observed within nodules. The distance from the pleura was assessed by drawing a perpendicular line between the pleura and nodules, and measuring the shortest distance. Pleural indentation was defined as linear areas of high attenuation originating from the nodules, extending peripherally to contact the pleura. When nodules were attached to the pleura, any distortions of the pleural architecture caused by the nodules were also defined as pleural indentation, and further categorized as concurrent pleural attachment with indentation. Vessel convergence appears on the slices as one or more vessels around the pulmonary nodule that touch, cut, or pass through the placeholder at its edge. The vacuole sign refers to a small round, oval, or strip-shaped area of reduced attenuation within the dense mass or nodule. Its diameter, less than 5 mm, can be distinguished from a pulmonary carcinoma cavity (greater than 5 mm is called a cavity), and the vacuole sign can be single or multiple (24,26-30).

Appendix 2 Image acquisition, preprocessing and segmentation

The most recent preoperative chest CT scans in Digital Imaging and Communications in Medicine (DICOM) format were downloaded from the hospital Picture Archiving and Communication System (PACS). All the CT images were then resampled, and the voxel units were standardized to 1×1×1. Subsequently, the window width and level were standardized to 1500 and –600, respectively. Two experienced thoracic surgeons (S.L. and X.G.) with 12 and 11 years of experience in thoracic oncology, respectively, and one experienced radiologist team (led by Dr. Jingzhi Wu with 15 years of experience in medical imaging, and including two other radiologists with 5 and 6 years of experience), performed a fully manual segmentation of the ROI. This was accomplished using the open-source software ITK-SNAP (http://www.itksnap.org/pmwiki/pmwiki.php). Each feature was analyzed for variations during manual segmentation between researchers. The intraclass correlation coefficient (ICC) was calculated for each feature, and only those with high stability (ICC >0.8) entered following feature selection and modeling process.

Appendix 3 Feature extraction and selection

Radiomics model

Pyradiomics in Python (version 3.7) was used to extract radiomics features from CT images, including first-order features, shape features, gray level features gray level cooccurrence matrix (GLCM), gray level size zone matrix (GLSZM), gray level run length matrix (GLRLM), neighbouring grey tone difference matrix (NGTDM) and gray level dependence matrix (GLDM). After conducting Mann-Whitney U test and Spearman's rank correlation analysis, a total of 202 features were retained. Among them, 19 were derived from exponential transformation, 36 from Laplacian of Gaussian filter, 15 from logarithm transformation, 14 from the original image, and 118 from wavelet transformation. Subsequently, a total of 16 radiomics features were retained through the absolute shrinkage and selection operator (LASSO) regression model.

https://dx.doi.org/10.21037/jtd-24-601

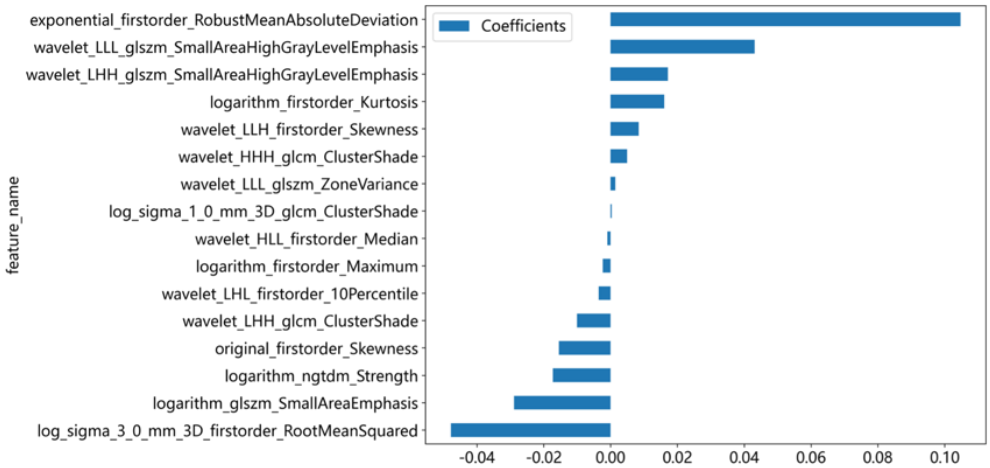

Figure S1 Radiomics feature coefficients after selection.

Radiomics feature coefficients are shown in Figure S1. Finally, a model was constructed using the radiomics signatures in the training cohort. Differences in the radiomics signature score (Rad score) were assessed using the Wilcoxon rank sum test.

Rad score = 0.2772775458602943
+0.104706 * exponential_firstorder_RobustMeanAbsoluteDeviation
+0.000347 * log_sigma_1_0_mm_3D_glcm_ClusterShade
−0.047752 * log_sigma_3_0_mm_3D_firstorder_RootMeanSquared
+0.016105 * logarithm_firstorder_Kurtosis
−0.002288 * logarithm_firstorder_Maximum
−0.028865 * logarithm_glszm_SmallAreaEmphasis
−0.017241 * logarithm_ngtdm_Strength
−0.015452 * original_firstorder_Skewness
+0.004964 * wavelet_HHH_glcm_ClusterShade
−0.000948 * wavelet_HLL_firstorder_Median
−0.010044 * wavelet_LHH_glcm_ClusterShade
+0.017165 * wavelet_LHH_glszm_SmallAreaHighGrayLevelEmphasis
−0.003542 * wavelet_LHL_firstorder_10Percentile
+0.008457 * wavelet_LLH_firstorder_Skewness
+0.043160 * wavelet_LLL_glszm_SmallAreaHighGrayLevelEmphasis
+0.001426 * wavelet_LLL_glszm_ZoneVariance

Deep learning model

The feature selection process was conducted in accordance with the aforementioned procedure, leading to the retention of a total of 32 deep learning features in 2D-ResNet50-ROI-only, 24 in 2D-ResNet50-ROI-rect, 52 in 2.5D-ResNet50-ROI-only, 39 in 2.5D-ResNet50-ROI-rect, 48 in 3D-ResNet50-ROI-only, and 36 in 3D-ResNet50-ROI-rect.

Model training process

At each iteration of the training process, binary cross entropy was used as the loss function to measure the information capacity included in the deep learning (DL) feature. If the minimum loss function was achieved, the intrinsic characteristics of the tumor were reflected by the DL feature, leading to the cessation of the training process. This process involved the

https://dx.doi.org/10.21037/jtd-24-601

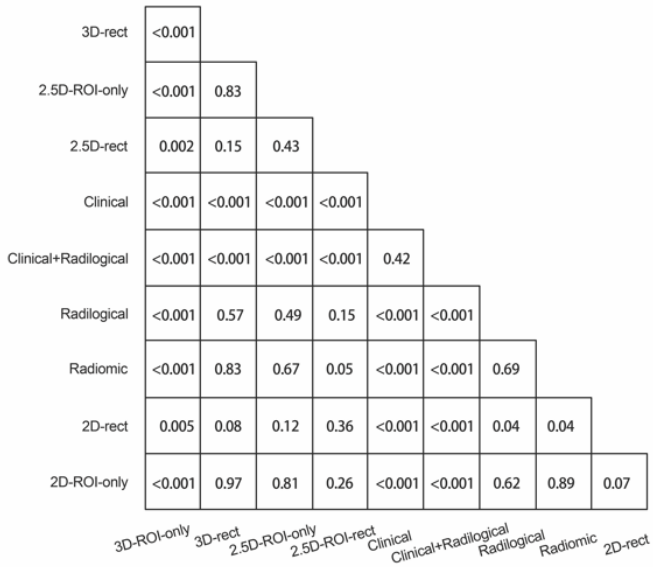

Figure S2 The results of Delong's test between the models.

iterative optimization of model parameters using stochastic gradient descent (SGD), with the learning rate and momentum set at 1e-4 and 0.7, respectively. The method was implemented in Python 2.7 and executed on a machine equipped with an Intel 11900K CPU, 64 GB memory, and NVIDIA GeForce GTX 3090 (24 GB on-board memory). A batch size of 32 was utilised for model training, and the training process converged after 50 epochs.

Appendix 4 Cut-off value calculation

The cut-off value was determined through a comprehensive consideration of the accuracy, sensitivity, and specificity of the prediction effect. The gold standard (prognosis and follow-up time) and the predicted value probability for ROC analysis were specifically analyzed. Firstly, a one-to-one correspondence was established between the gold standard and the predicted results, which were then sorted in ascending order based on their probabilities. Secondly, the probability of each case was treated as a common threshold, leading to the classification of cases into high-risk or low-risk categories, and subsequently calculating a confusion matrix. Thirdly, multiple confusion matrices were computed along with their corresponding pairs of sensitivity and specificity. Finally, the best cut-off value, referred to as the "cut-off value" in the text, was identified when the sum of sensitivity and specificity reached its maximum for the i-th pair. The cut-off extraction process involved determining the value of i that maximized the expression {sensitivity(i) + specificity(i)}, and then finding the corresponding probability as the cut-off value.

Appendix 5 Delong's test

Figure S2 displays the results of Delong's test between the models. P values are represented by the numbers in the figure.

Appendix 6 Overall survival for all patients

Figure S3 displays the survival time curves for all patients. Out of the total number of patients, 34 succumbed to mortality. The average follow-up time was 47.57±0.50 months (95% CI: 36.581–38.555), with a median follow-up time of 48 months (95% CI: 36.346–38.654). The OS rates at 1, 3, and 5 years were 99.0%, 94.7%, and 88.7%, respectively. In the training cohort, 27 patients deceased, with the overall survival rates at 1, 3, and 5 years being 99.1%, 95.5%, and 87.2%. In the validation cohort,

https://dx.doi.org/10.21037/jtd-24-601

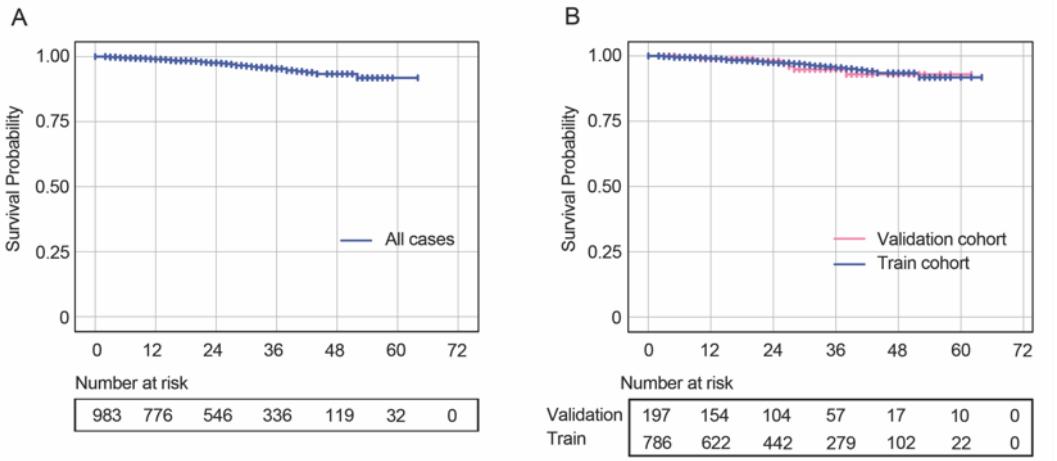

Figure S3 Comparison of Kaplan-Meier curves for prognosis in all patients.

7 patients died, and the overall survival rates at 1, 3, and 5 years were 98.8%, 93.8%, and 89.9%, respectively. Furthermore, there was no significant difference in the survival rates between the training and validation cohorts of patients (P=0.79).

References

26. Ko JP, Suh J, Ibidapo O, et al. Lung Adenocarcinoma: Correlation of Quantitative CT Findings with Pathologic Findings. Radiology 2016;280:931-9.
27. Zhang Y, Shen Y, Qiang JW, et al. HRCT features distinguishing pre-invasive from invasive pulmonary adenocarcinomas appearing as ground-glass nodules. Eur Radiol 2016;26:2921-8.
28. Lee SM, Park CM, Goo JM, et al. Invasive pulmonary adenocarcinomas versus preinvasive lesions appearing as ground-glass nodules: differentiation by using CT features. Radiology 2013;268:265-73.
29. Li P, Kong X, Li J, et al. A Dataset of Pulmonary Lesions With Multiple-Level Attributes and Fine Contours. Front Digit Health 2020;2:609349.
30. Gao B, McKinney A. Classic Imaging Signs. 1 ed. Springer Cham: 2020:103-75.

https://dx.doi.org/10.21037/jtd-24-601