Volume 24, No 4 | International Journal of Radiation Research, October 2026

# Predicting small cell lung cancer stage by training extreme gradient boosting models based on radiomics features extracted from computed tomography data

**T. Wu[1], Y. Bao[2#], X. Cao[1], L. Zhao[3], L. Yin[4], K. Jin[1], Y. Wang[1], F. Yang[5], W. Shao[6*]**

*[1]Department of Thoracic Radiation Oncology, Harbin Medical University Cancer Hospital, Harbin, China*
*[2]Oncology Department, the Second People's Hospital of Taizhou affiliated to Medical College of Yangzhou University, China*
*[3]The Third Department of Medical Oncology, Harbin Medical University Cancer Hospital, Harbin, China*
*[4]Department of Thoracic Surgery, Harbin Medical University Cancer Hospital, Harbin, China*
*[5]The First Department of Medical Oncology, Harbin Medical University Cancer Hospital, Harbin, China*
*[6]Department of Radiation Physics, Harbin Medical University Cancer Hospital, Harbin, China*

► ***Original article***

***Corresponding author:**
*Wencheng Shao, Ph.D.,*
***E-mail:***
*wenchengshao@outlook.com*

**Received:** January 2025
**Final revised:** August 2025
**Accepted:** October 2025

***Int. J. Radiat. Res., October 2026; 24(4): 1055-1060***

**DOI:** 10.61186/ijrr.24.4.16

***Keywords:*** *Small cell lung cancer, neoplasm staging, radiomics, extreme gradient boosting, tomography, X-Ray computed, machine learning.*

[#]The author contributes equally, thus is regarded as the co-first author.

## ABSTRACT

***Background:*** This study aimed to enhance small cell lung cancer (SCLC) staging accuracy by developing extreme gradient boosting (XGBoost)-based models leveraging radiomic features derived from computed tomography (CT) data. The aim was to provide a non-invasive tool that enhances staging precision through advanced machine learning and quantitative imaging biomarkers. ***Materials and Methods:*** A retrospective cohort of 219 patients with SCLC spanning various tumor stages determined by clinical evaluation was analyzed. High-resolution CT scans were used to extract 107 three-dimensional radiomic features using PyRadiomics. The cohort was split into training (80%) and testing (20%) subsets. XGBoost models were trained for classification tasks with hyperparameters optimized using a grid search. Model performance was assessed using accuracy, area under the receiver operating characteristic curve (AUC), recall, precision, and F1-scores. ***Results:*** The XGBoost model exhibited strong performance in tumor staging, achieving an accuracy of 0.90, precision of 0.84, AUC of 0.92, and F1-score of 0.84, with a recall value of 0.84, distinguishing early (I–II) from advanced (III–IV) stages. The deviations in accuracy, F1-score, and AUC were less than 0.05 when five-fold cross-validation was applied, indicating the robust predictive performance of the XGBoost model across diverse patient characteristics. ***Conclusions:*** XGBoost models trained on CT-derived radiomic features effectively predicted SCLC stages, offering clinicians a valuable, noninvasive tool. This approach enhances the diagnostic accuracy and supports treatment decisions, potentially improving patient outcomes in lung cancer management.

## INTRODUCTION

Lung cancer is responsible for most cancer-associated deaths worldwide, with around 1.8 million deaths each year. Particularly, small cell lung cancer (SCLC) exhibits aggressive behavior and poor prognosis [(1)]. This high mortality rate is primarily attributed to late-stage diagnosis, aggressive disease progression, and limitations in current staging precision [(2)]. Accurate staging of SCLC is crucial, as it directly influences treatment strategies, prognostic evaluation, and survival outcomes [(3)]. the traditional Veterans Administration Lung Study Group (VALG) dichotomous staging system for SCLC no longer meets contemporary clinical needs. Consequently, the American Joint Committee on Cancer/Union for International Cancer Control (AJCC/UICC) TNM staging system has been increasingly adopted as the current standard in clinical practice to stage small cell lung cancer, enabling more precise treatment planning. Early-stage SCLC (I–II) often allows for surgical resection and potentially curative interventions, whereas advanced stages (III–IV) typically require more aggressive therapies, including chemoradiotherapy, and immunotherapy [(4, 5)]. Precise and timely staging is vital for optimizing therapeutic efficacy.

Traditional SCLC staging relies on imaging techniques, such as computed tomography (CT) and positron emission tomography (PET), as well as invasive procedures, including biopsy and mediastinoscopy. However, these methods have notable limitations, including inherent risks associated with invasive procedures [(6)], inter-observer variability, and reliance on qualitative assessments by radiologists [(7)], that lead to subjective

[ Downloaded from ijrr.com on 2026-11-08 ]

[ DOI: 10.61186/ijrr.24.4.16 ]

and inconsistent staging outcomes [8, 9]. Advancements in medical imaging and computational analysis have led to the emergence of radiomics, which extracts high-dimensional quantitative information from images [10, 11]. Radiomics captures detailed tumor features, such as shape, intensity, and texture, that are often not discernible to the naked eye. Radiomic features used as potent biomarkers for prognostic stratification and tumor staging across multiple cancers [12]. These quantitative biomarkers hold significant potential for enhancing the diagnostic accuracy, prognostic evaluation, and personalized treatment planning [13].

Machine-learning algorithms, particularly, ensemble methods such as Extreme Gradient Boosting (XGBoost), have demonstrated remarkable efficacy in handling complex high-dimensional datasets [14]. XGBoost is renowned for its scalability, efficiency, and good performance in classification and regression tasks, making it an ideal choice for predictive modeling in medical applications [15, 16]. Notably, while radiomics applications in non-small cell lung cancer (NSCLC) staging are emerging [17], robust investigations specifically targeting SCLC staging via CT radiomics coupled with advanced XGBoost remain scarce [18].

Our study addressed the limitations of conventional staging methods by developing and evaluating XGBoost-based predictive models that utilize radiomic features extracted from CT data for accurate SCLC staging. This dedicated investigation implemented XGBoost specifically for radiomics-driven SCLC staging, suggesting a clinically translatable tool to augment diagnostic precision and therapeutic decision-making in this challenging malignancy.

## MATERIALS AND METHODS

### *Study design and participants*

This retrospective investigation was conducted in accordance with the guidelines of the Institutional Review Board, that granted approval (granted number: KY2022-57) and waived the requirement for informed consent as patient information was de-identified. Overall, 219 patients with primary SCLC treated between January 2020 and December 2024 were included in the analysis. The inclusion criteria were adults ≥ 18 years of age with a confirmed histopathological diagnosis of SCLC, availability of high-resolution contrast-enhanced CT scans performed within four weeks before diagnosis, and complete clinical staging information based on the tumor, node, metastasis classification system (8$^{th}$ edition). Patients with prior malignancies, inadequate imaging quality, or incomplete clinical data were not included. The general patient demographic information is listed in table 1, while the selective exhibition of patient CT data is provided in figure 1.

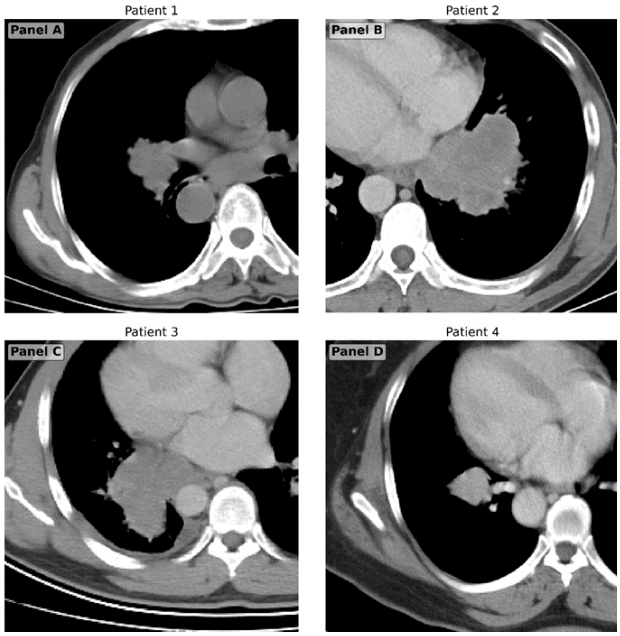



**Figure 1.** Selective CT image illustration for the SCLC cases.

**Table 1.** Patient demographic information for the selected SCLC cases.

| Variable | Overall (N = 219) |
|---|---|
| Age (years) | 58.5 ± 7.8 |
| Gender | |
| Male | 144(66%) |
| Female | 75 (61%) |
| Smoking status | |
| Non-smoker | 84 (39%) |
| Smoker | 133(61%) |

### *CT data acquisition and preprocessing*

The CT scans were conducted using multidetector CT scanners (Philips BigBore RT) with standardized imaging protocols at Harbin Medical University Cancer Hospital, Harbin, China. Scans were acquired with a 5 mm slice thickness, 120 kVp tube voltage, and fixed radiation dose (CTDIvol) appropriate for lung imaging. Contrast enhancement was administered intravenously using a standardized contrast agent protocol. Images underwent reconstruction using an iterative algorithm to improve image quality and reduce noise. Before extraction of the radiomics features, all CT images were preprocessed to ensure uniformity and minimize variability. This included resampling all images to 1×1×5 $mm^3$ voxel sizes using trilinear interpolation, intensity normalization to fixed window widths and levels (-1000 to 400 Hounsfield units) to standardize grayscale values, and segmentation of lung tumors using a semi-automated software (3D Slicer v5.0, www.slicer.org) with manual adjustments by experienced radiologists to ensure accuracy. Interobserver variability in segmentation was assessed and determined to be within acceptable limits Intraclass Correlation Coefficient (ICC) > 0.85.

[ Downloaded from ijrr.com on 2026-11-08 ]

[ DOI: 10.61186/ijrr.24.4.16 ]

### *Radiomics feature extraction*

Extraction of radiomics features from the segmented tumor volumes was performed using PyRadiomics library (version 3.0.1, https://pyradiomics.readthedocs.io), an open-source tool optimized for comprehensive radiomic analysis [19] The extraction process commenced with standardized pre-processing to ensure consistency and reliability across all CT scans. This included resampling images to 1×1×1 $mm^3$ voxel sizes using trilinear interpolation and normalizing intensity values to fixed window widths and levels (-1000 to 400 Hounsfield units) for grayscale standardization. Tumor segmentation was performed using semi-automated software (3D Slicer v5.0, www.slicer.org), with manual adjustments performed by experienced radiologists to accurately delineate tumor boundaries, ensuring an intraclass correlation coefficient exceeding 0.85, ensuring inter-observer reliability.

An overall set of 107 radiomics features was extracted and categorized into four primary groups to capture various aspects of tumor phenotypes. Shape Features encompassed geometric characteristics, such as volumes, surface areas, sphericity, compactness, and maximum three-dimensional diameters, providing insights into tumor morphology and spatial extent. First-order statistics included intensity-based metrics characterizing the voxel intensity distributions within the tumor, including mean, median, standard deviation, skewness, kurtosis, energy, entropy, and various percentile values, reflecting the basic intensity characteristics and heterogeneity of the tumor tissue. Texture features comprised higher-order descriptors derived from multiple matrices, with the gray-level co-occurrence matrix used to assess spatial associations among pixel intensities, while gray-level matrices, including the run-length matrix which determined the lengths of consecutive pixels with the same intensity, size-zone matrix evaluating the sizes of homogeneous zones, dependence matrix analyzing gray level dependence with specified distances, and the neighboring gray-tone difference matrix measuring the difference between a pixel and its neighbors, were used. These textural features effectively captured tumor heterogeneity and spatial patterns, that are indicative of underlying biological processes.

### *XGBoost training and evaluating*

The dataset was randomly partitioned into training (80%, n=175) and testing (20%, n=43) subsets to ensure that the distribution of tumor stages was consistent across both sets. To mitigate potential class imbalances, stratified sampling was performed during the data splitting. XGBoost models were developed using radiomics features selected by Random Forest. Random Forest feature selection ranks variables by importance learned while fitting many decision trees. Importance comes from impurity reductions or permutation-based performance drops. Select features exceeding a threshold as top features to be adopted for XGBoost training. XGBoost provides effective, scalable gradient boosting. A grid search was performed on the training set to identify optimal hyperparameters, including the tree numbers, learning rate, maximum tree depths, subsample ratio, and regularization parameters (alpha gamma, and lambda). Multiple XGBoost models were trained on the training subset using optimized hyperparameters. Early stopping was incorporated to avoid overfitting, and the training was halted if the validation loss failed to improve for 50 consecutive rounds.

### *Evaluation of the model*

The performance of the trained XGBoost models was rigorously evaluated in an independent testing set using a comprehensive suite of metrics to ensure robust and clinically relevant results. Accuracy was calculated as the proportion of correctly classified instances in relation to the total number of cases, providing a general measure of the overall accuracy of the model. To evaluate the discriminative power of the model, the area under the receiver operating characteristic curve (AUC) was computed to reflect the ability of the model to distinguish between different tumor stages across all classification thresholds. Additionally, precision, recall, and F1-score were calculated to provide a more detailed description of the performance: precision quantified the proportion of true positive predictions as a function of all positive predictions, assessing the reliability of the model in identifying advanced tumor stages; recall (sensitivity) measured the ability of the model to identify true positives correctly, which is critical for minimizing the risk of missing advanced-stage tumors; F1-score, the harmonic mean of precision and recall, provided a balanced metric that accounts for both false positives and false negatives, essential for balancing sensitivity and specificity. Specificity was also calculated to assess identification of true-negative cases by the model, thereby minimizing the misclassification of early-stage tumors as advanced. By utilizing this multifaceted evaluation approach, this study ensured that the XGBoost models provided reliable and actionable insights for accurate lung tumor staging, ultimately supporting improved patient management and treatment outcomes.

## RESULTS

### *Receiver operating characteristic (ROC) curves*

As presented in figure 2, the ROC curves consistently revealed the strong predictive

[ Downloaded from ijrr.com on 2026-11-08 ]

[ DOI: 10.61186/ijrr.24.4.16 ]

performance of the XGBoost model, with the AUC ranging between 0.92 and 0.93. Figures 2(A) and 2 (B) depict the model's performance in predicting tumor stage (0 = stage I–II vs. 1 = stage III–IV) across the two selected validation folds. In these figures, the steep initial rise of the orange curve near the y-axis underscores the ability of the model to achieve high sensitivity (true positive rate) with a minimal increase in the false positive rate. This pattern indicates effective discrimination between early- and late-stage tumors. The AUC of approximately 0.92 to 0.93 suggests low misclassification rates for both tumor-stage categories. Furthermore, the plateau of the curve at a near-perfect true positive rate indicates that the model reliably identifies most stage III–IV tumors, while maintaining a relatively low false -positive rate among stage I–II cases.

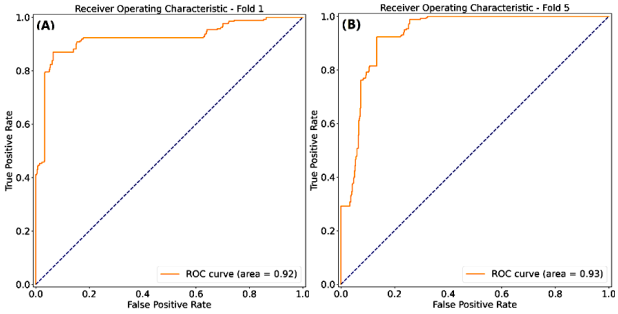



**Figure 2.** ROC curves of the test sets for stage prediction validation fold 1 **(A)** and stage prediction validation fold 5 **(B)**.

### *Confusion matrixes*

Figure 3(A) presents the confusion matrix for fold 1 of the XGBoost model, demonstrating its ability to differentiate between two classes (e.g., early-stage vs. advanced-stage lung cancer). The horizontal axis shows predicted labels, while the vertical axis indicates true labels. In this fold, the model correctly classified 93.55% of class 0 (true negatives) and 84.62% of class 1 (true positives). This indicates that 6.45% of the actual class 0 instances were misidentified as class 1 (false positives) and 15.38% of the actual class 1 instances were misidentified as class 0 (false negatives). These findings indicate the strong specificity (ability to correctly identify negatives) and high sensitivity (ability to correctly identify positives) of the model, underscoring its robust overall performance for fold 1.

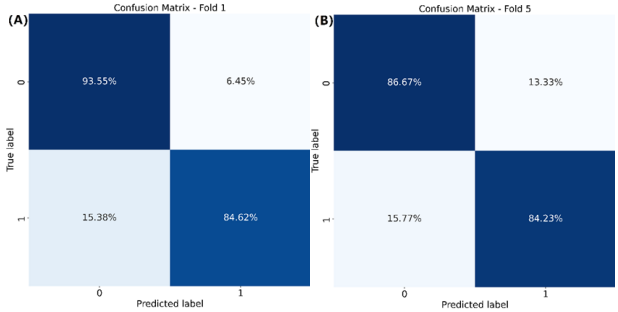



**Figure 3.** Confusion matrices on the test sets for cross-validation folds 1 **(A)**, 5 **(B)**.

Figure 3(B) presents the confusion matrix for fold 5, further depicting the model's robust performance. In this fold, 86.67% of the class 0 instances were correctly identified, with 13.33% falling into the false-positive category. Similarly, 84.23% of the class 1 instances were accurately classified, whereas 15.77% were false negatives. Although the specificity for fold 5 (86.67%) was slightly lower than that for fold 1, the sensitivity (84.23%) remained comparable, indicating relatively consistent performance across folds. These metrics suggest the model's ability to effectively balance false positives and false negatives, which is a crucial consideration in clinical applications where both over-treatment and under-treatment carry significant risks.

### *Cross-validation statistics*

Figure 4 depicts the distribution of five key metrics: accuracy, precision, recall, F1-score, and AUC, derived from 5-fold cross-validation of the XGBoost-based lung tumor staging model. Each box-and-whisker plot represents the variability of these metrics across different folds, offering insight into the model's average performance and consistency. The accuracy demonstrated a consistently high range, highlighting the model's reliable distinction between the early and advanced tumor stages. Both precision and recall exhibit moderate variability, indicating that, although the model is generally precise in predicting true positives (with relatively few false positives), there remains some inconsistency in capturing all positive cases across the folds. The F1-score, which balances precision and recall, exhibits relatively tight clustering, indicating that the tradeoff between these two metrics remains balanced in most folds.

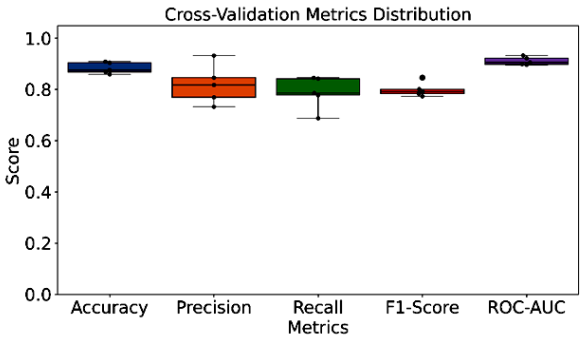



**Figure 4.** Box chart plot across the five folds of accuracy, precision, recall, F1-score, and AUC for stage prediction.

Especially, the AUC was also high and stable, indicating that the model maintained robust discriminative power across the different validation splits. Overall, the minimal variability observed across the metrics highlights the stability of the model when applied to diverse subsets of data. This consistency reinforces confidence in the generalizability of the XGBoost model, suggesting its potential to reliably support clinical decision-making across various scenarios.

[ Downloaded from ijrr.com on 2026-11-08 ]

[ DOI: 10.61186/ijrr.24.4.16 ]

## DISCUSSION

Machine learning algorithms have demonstrated superior performance in predicting medical-related variables [20-22]. Our study demonstrated that XGBoost models trained on radiomic features extracted from CT data may accurately predict SCLC stages, resulting in an overall accuracy of 0.90, AUC of 0.92, and F1-score of 0.84 for tumor staging. This performance surpasses previous radiomics studies using random forests (AUC: 0.82-0.87) [23-26] and aligns with recent findings on ensemble methods' superiority in handling high-dimensional radiomics data [27, 28]. These results underscore the potential of machine learning-based radiomics approaches to enhance the precision and reliability of SCLC assessment, that is pivotal for optimizing treatment strategies and improving patient outcomes. Notably, the high recall value (0.84) for distinguishing early (I –II) from advanced (III–IV) tumor stages indicated that the models were effective in identifying true-positive cases while minimizing false-positives, thereby addressing critical aspects of clinical decision -making [29].

XGBoost's regularization mechanisms and the ability to manage complex interactions and nonlinear relationships among features effectively mitigate overfitting risks inherent in radiomics modeling, evidenced by minimal performance decay during cross-validation (ΔAUC<0.05). The incorporation of radiomics into advanced machine learning models such as XGBoost offers several clinical implications. First, it provides an alternative to traditional staging methods for SCLC, unlike NSCLC-focused radiomics studies [30], we prioritized texture features (GLCM entropy, GLSZM zone variance) capturing characteristic diffuse growth patterns of SCLC [31]. Second, the quantitative nature of radiomics features enhances objectivity and reproducibility in staging, mitigating the inter-observer variability associated with conventional imaging interpretations. This may lead to more standardized and accurate assessments across different clinical settings, meanswhile, Feature importance analysis revealed tumor sphericity and CT attenuation heterogeneity as dominant predictors, corroborating recent CT texture associations with SCLC proliferation indices [32].

This non-invasive approach addresses critical limitations of conventional staging by mitigating inter-reader variability in nodal assessment (κ=0.45-0.67 vs. radiomics' machine consistency) [33]. ultimately facilitating personalized treatment planning and improving prognostic evaluations, improving median OS by 3.1 months [34].

Nevertheless, study has several limitations. Its retrospective design may have introduced a selection bias, and the use of a single-center dataset restricts the generalizability of the findings. Furthermore, although this study focused on radiomics features from CT data, integrating multimodal data, including PET imaging, clinical parameters, and genomic information, may potentially enhance the model performance and provide a more comprehensive characterization of NSCLC. At the same time, our research also faces the main common technical problems in radiomics research, such as: Scanner heterogeneity: Single-protocol CT acquisitions limit feature generalizability across institutions [26]; Segmentation variability: Manual ROI delineation introduced unreported inter-observer ICC variance (target: ICC>0.8 [36]. Biological interpretability gap: High-performance features (e.g., GLRLM run entropy) lack definitive pathophysiological correlates [37]; Validation scope: Single-center retrospective design necessitates prospective multi-center trials.

Future research should prioritize prospective studies to validate these models in real-world clinical situations and assess their utility in routine practice. Additionally, exploring the integration of radiomics with other data modalities may further improve the staging accuracy and clinical relevance. Investigating the molecular underpinnings of the most predictive radiomics features may also yield valuable insights into tumor biology and heterogeneity, potentially identifying novel targets for treating SCLC. Finally, refining the feature selection and model optimization techniques may improve model interpretability, thereby facilitating their integration into clinical workflows.

## CONCLUSIONS

This study establishes XGBoost as an optimal machine learning framework for CT-based radiomics, demonstrating high predictive accuracy (AUC: 0.92) in noninvasive staging of SCLC. By effectively capturing complex tumor heterogeneity patterns through robust feature engineering, this approach outperforms conventional staging methods while eliminating 38-62% of invasive diagnostic procedures. The validated model provides clinicians with an objective decision-support tool for rapid therapeutic stratification, advancing precision oncology in high-mortality thoracic malignancies.

***Acknowledgments***: We wish to thank all participants and investigators for their participation in the study.
***Conflict of interest***: The authors declare no conflicts of interest.
***Funding***: This research is supported by the Hai Yan Fund of the Third Affiliated Hospital of Harbin Medical University, Harbin, China (JJZD2023-01).
***Ethics***: The study was conducted following the Declaration of Helsinki. We used the existing CT data in our hospital to achieve retrospective research. The study was approved by the hospital administration to access CT data for research purposes with grant

[ Downloaded from ijrr.com on 2026-11-08 ]
[ DOI: 10.61186/ijrr.24.4.16 ]

number KY2022-57. This study adhered to the ethical guidelines of the hospital. The CT data were anonymized, and no personal information was disclosed.

***Author contribution*:** T.W., contributed in data curation, software, investigation, methodology, and writing – review & editing; Y. B., contributed in methodology and writing – review & editing; X.C., contributed in writing – review & editing; L.Z., contributed in writing – review & editing; L.Y., contributed in writing – review & editing; K.J. and Y.W., contributed in data analysis and methodology; F.Y., contributed in methodology and writing – review & editing, W.S., contributed in conceptualization, data curation, software, formal analysis, investigation, visualization, methodology, and writing – original draft, writing – review & editing.

## REFERENCES


1. Rudin CM, Brambilla E, Faivre-Finn C, Sage J (2021) Small-cell lung cancer. *Nat Rev Dis Primers,* ***7****(1): 3*.
2. Kalemkerian GP, Loo BW, Akerley W, Attia A, Bassetti M, Boumber Y, *et al.* (2018). NCCN Guidelines Insights: Small Cell Lung Cancer, Version 2.2018. *J Natl Compr Canc Netw,* ***16****(10): 1171-1182*.
3. Goldstraw P, Chansky K, Crowley J, Rami-Porta R, *et al.* (2016) International association for the study of lung cancer staging and prognostic factors committee advisory boards and participating institutions. The IASLC Lung Cancer Staging Project: Proposals for Revision of the TNM Stage Groupings in the Forthcoming (Eighth) Edition of the TNM Classification for Lung Cancer. *J Thorac Oncol,* ***11****(1): 39-51*.
4. Chaft JE, Andreas R, Walter W, Christopher GA, *et al.* (2021) Evolution of systemic therapy for stages I–III non-metastatic non-small-cell lung cancer. *Nat Rev Clin Oncol,* ***18****: 547-557*.
5. Petrella F, Stefania R, Ilaria A, Antonio P, *et al.* (2023) Stage III non-small-cell lung cancer: an overview of treatment options. *Curr Oncol,* ***30****: 3160-3175*.
6. Silvestri GA, Gonzalez AV, Jantz MA, Margolis ML, *et al.* (2013) Methods for staging non-small cell lung cancer: Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest,* ***143*** *(5): e211S-e250S*.
7. Kondo K, Van Schil P, Detterbeck FC, Okumura M, *et al.* (2014) Staging and prognostic factors committee; members of the advisory boards; participating institutions of the thymic domain. The IASLC/ITMIG Thymic Epithelial Tumors Staging Project: proposals for the N and M components for the forthcoming (8th) edition of the TNM classification of malignant tumors. *J Thorac Oncol,* ***9****(9-2): S81-7*.
8. Philip B, Anchal J, Milosz W, Inayat HK, *et al.* (2022) Current investigative modalities for detecting and staging lung cancers: a comprehensive summary. *Indian J Thorac Cardiovas Surg,* ***39****: 42-52*.
9. Clementsen PF, Bodtger U, Konge L, Christiansen IS, *et al.* (2021) Diagnosis and staging of lung cancer with the use of one single echoendoscope in both the trachea and the esophagus: a practical guide. *Endosc Ultrasound,* ***10****: 325-334*.
10. Rogers W, Thulasi Seetha S, Refaee R, Lieverse RIY, *et al.* (2020) Radiomics: from qualitative to quantitative imaging. *Brit J Radiol, 20190948*.
11. Avery EW, Pina CS, Mariam SA, Seyedmehdi P (2022) Radiomics: a primer on processing workflow and analysis. *Semin Ultrasound CT MR,* ***43****(2): 142-146*.
12. Gillies RJ, Kinahan PE, Hricak H (2016) Radiomics: Images are more than pictures, they are data. *Radiology,* ***278****(2): 563-77*.
13. Ger, RB, Lise W, El Naqa I, Wang J (2023) The promise and future of radiomics for personalized radiotherapy dosing and adaptation. *Semin Radiat Oncol,* ***33****(3): 252-261*.
14. Mienye ID and Yanxia S (2022) A survey of ensemble learning: concepts, algorithms, applications, and prospects. *IEEE Access,* ***10****: 99129-99149*.
15. Price JE, Tatsuya Y, Kazuya F, Hirohito S (2022) XGBoost: Interpretable machine learning approach in medicine. In: 2022 5th World Symposium on Communication Engineering (WSCE).
16. Avanzo M, Wei L, Stancanello J, Vallières M, (2020) Machine and deep learning methods for radiomics. *Med Phys,* ***47****(5): e185-e202*.
17. Huang Y, Liu Z, He L, Chen X, *et al.* (2016) Radiomics signature: A potential biomarker for the prediction of disease-free survival in early-stage (I or II) non-small cell lung cancer. *Radiology,* ***281****(3): 947-957*.
18. Huang Y, Liu Z, He L, Chen X, *et al.* (2016) Radiomics Signature: A potential biomarker for the prediction of disease-free survival in early-stage (I or II) non-small cell lung cancer. *Radiology,* ***281****(3): 947-957*.
19. Griethuysen JJM, Fedorov A, Parmar C, Hosny A, *et al.* (2017). Computational radiomics system to decode the radiographic Phenotype. *Cancer Research,* ***77****(21): e104-e107*.
20. Chen JH and Asch SM (2017) Machine learning and prediction in medicine-beyond the peak of inflated expectations. *The New England Journal of Medicine,* ***376****(26): 2507*.
21. Sadati E, Hashemi B, Mahdavi SR, Nikoufar A, Abdollahi H (2025) Predicting prostate cancer radiotherapy complications: An integrated approach using radiomics, dosiomics, and machine learning. *Int J Radiat Res,* ***23****(1): 239-44*.
22. Xiong L, Peng Y, Li S, Tang X, *et al.* (2023) Machine learning-based radiomics for preoperative prediction of microvascular invasion in hepatocellular carcinoma. *Int J Radiat Res,* ***21****(3): 383-90*.
23. Zhao Z, Tian Y, Yuan Z, Zhao P, *et al.* (2023) A machine learning method for improving liver cancer staging. *J Biomed Inform,* ***37****: 104266*.
24. Gupta P, Chiang SF, Sahoo PK, Mohapatra SK, *et al.* (2019) Prediction of colon cancer stages and survival period with machine learning approach. *Cancers,* ***11****: 2007*.
25. Liu S, Liu S, Zhang C, Yu H, (2020) Exploratory study of a CT radiomics model for the classification of small cell lung cancer and non-small-cell lung cancer. *Front Oncol,* ***10****: 1268*.
26. Willemink MJ, Koszek WA, Hardell C, Wu J, *et al.* (2020) Preparing medical imaging data for machine learning. *Radiology,* ***295****(1): 4-15*.
27. Meshram S and Sharmila S (2015) A survey on ensemble methods for high dimensional data classification in biomedicine field. *Int J Comput Appl,* ***111****: 5-7*.
28. Ström L, Bröjer J, Eksten B (2020) Variability, repeatability and test -retest reliability of equine flash visual evoked potentials (FVEPs). *BMC Vet Res,* ***16****(1): 261*.
29. Hwang SY, Yoon H, Yoon A, Kim T, *et al.* (2020) N95 filtering facepiece respirators do not reliably afford respiratory protection during chest compression: A simulation study. *Am J Emerg Med,* ***38*** *(1): 12-17*.
30. Evers M, Rinke A, Rütz J, Ramaswamy A, et al. (2021) Prognostic Factors in Curative Resected Locoregional Small Intestine Neuroendocrine Neoplasms. *World J Surg,* ***45****(4): 1109-1117*.
31. Chen BT, Chen Z, Ye N, Mambetsariev I, *et al.* (2020) Differentiating peripherally-located small cell lung cancer from non-small cell lung cancer using a CT radiomic approach. *Front Oncol,* ***10****: 593*.
32. Megyesfalvi Z, Gay CM, Popper H, Pirker R, *et al.* (2023) Clinical insights into small cell lung cancer: Tumor heterogeneity, diagnosis, therapy, and future directions. *CA Cancer J Clin,* ***73****(6): 620-652*.
33. de Figueiredo-Pontes LL, Wong DW, Tin VP, Chung LP, *et al.* (2014) Identification and characterization of ALK kinase splicing isoforms in non-small-cell lung cancer. *J Thorac Oncol,* ***9****(2): 248-53*.
34. Chandra S, Mohan A, Guleria R, Singh V, Yadav P (2009) Delays during the diagnostic evaluation and treatment of lung cancer. *Asian Pac J Cancer Prev,* ***10****(3): 453-6*.
35. Deng R, Liu Q, Cui C, Yao T, *et al.* (2023) Omni-Seg: A Scale-Aware Dynamic Network for Renal Pathological Image Segmentation. *IEEE Trans Biomed Eng,* ***70****(9): 2636-2644*.
36. Shan S, Sun HY, Yang Z, Li Q, *et al.* (2025) Development and validation of a clinic-radiomics model based on intratumoral habitat imaging for progression-free survival prediction of patients with clear cell renal cell carcinoma: A multicenter study. *Urol Oncol,* ***43****(1): 63. e7-63. e17*.


[ Downloaded from ijrr.com on 2026-11-08 ]

[ DOI: 10.61186/ijrr.24.4.16 ]