ARTIFICIAL INTELLIGENCE

original reports abstract

Machine Learning Radiomics Model for Early Identification of Small-Cell Lung Cancer on Computed Tomography Scans

Rajesh P. Shah, MD[1,2]; Heather M. Selby, PhD[1,3]; Pritam Mukherjee, PhD[3]; Shefali Verma, BS[4]; Peiyi Xie, MD[3,5]; Qinmei Xu, BS[3,6]; Millie Das, MD[1,7]; Sachin Malik, MD[1,2]; Olivier Gevaert, PhD[3]; and Sandy Napel, PhD[2]

PURPOSE Small-cell lung cancer (SCLC) is the deadliest form of lung cancer, partly because of its short doubling time. Delays in imaging identification and diagnosis of nodules create a risk for stage migration. The purpose of our study was to determine if a machine learning radiomics model can detect SCLC on computed tomography (CT) among all nodules at least 1 cm in size.

MATERIALS AND METHODS Computed tomography scans from a single institution were selected and resampled to 1 × 1 × 1 mm. Studies were divided into SCLC and other scans comprising benign, adenocarcinoma, and squamous cell carcinoma that were segregated into group A (noncontrast scans) and group B (contrast-enhanced scans). Four machine learning classification models, support vector classifier, random forest (RF), XGBoost, and logistic regression, were used to generate radiomic models using 59 quantitative first-order and texture Imaging Biomarker Standardization Initiative compliant PyRadiomics features, which were found to be robust between two segmenters with minimum Redundancy Maximum Relevance feature selection within each leave-one-out-cross-validation to avoid overfitting. The performance was evaluated using a receiver operating characteristic curve. A final model was created using the RF classifier and aggregate minimum Redundancy Maximum Relevance to determine feature importance.

RESULTS A total of 103 studies were included in the analysis. The area under the receiver operating characteristic curve for RF, support vector classifier, XGBoost, and logistic regression was 0.81, 0.77, 0.84, and 0.84 in group A, and 0.88, 0.87, 0.85, and 0.81 in group B, respectively. Nine radiomic features in group A and 14 radiomic features in group B were predictive of SCLC. Six radiomic features overlapped between groups A and B.

CONCLUSION A machine learning radiomics model may help differentiate SCLC from other lung lesions.

JCO Clin Cancer Inform 5:746-757. © 2021 by American Society of Clinical Oncology

INTRODUCTION

Lung cancer is the second most commonly diagnosed cancer among men and women and the most common cause of cancer-related deaths in the United States, with over 140,000 estimated deaths to occur in 2020 alone.[1] The subtype with the worst prognosis is small-cell lung cancer (SCLC), which represents 13% of all lung cancers.[2] The 5-year overall survival for SCLC is 6.5% compared with 20.5% for non–small-cell lung cancer (NSCLC).[3] SCLC is typically divided into limited stage (LS) versus extensive stage (ES) disease, with ES representing disease that cannot be encompassed into a tolerable radiation field. Two thirds of patients present with ES-SCLC.[4] Notably, survival in SCLC varies markedly for LS versus ES, with 5-year survival at 27.2% for LS-SCLC and 2.9% for ES-SCLC.[3] SCLC has an aggressive doubling time as short as 25-54 days.[5] Therefore, early detection of SCLC is critical to expedite diagnosis and treatment. A delay of 1-2 months between initial imaging and diagnosis could be within the doubling time of the tumor and is not uncommon.[6] In a National Cancer Database study, 28% of patients with SCLC underwent treatment initiation more than 28 days after diagnosis.[7] Presumably, the time to treatment from initial imaging was significantly longer. These longer delays could result in stage migration whereby a patient with initially localized disease can develop metastatic disease, adversely affecting overall prognosis and survival.

Currently, apart from surgical or needle biopsy, no clear method exists to distinguish SCLC from other subtypes of lung cancers or from benign nodules when identified incidentally or on screening computed tomography (CT) examinations. A possible method to allow early distinction is through the use of radiomics, which uses combinations of quantitative imaging features for diagnostic or predictive purposes.[8,9] Radiomic feature analysis has been used to distinguish benign from malignant nodules, to determine tumor phenotypes and patient prognosis, and to

ASSOCIATED CONTENT

Appendix

Author affiliations and support information (if applicable) appear at the end of this article.

Accepted on June 8, 2021 and published at ascopubs.org/journal/cci on July 15, 2021: DOI https://doi.org/10.1200/CCI.21.00021

CONTEXT

Key Objective

Can a machine learning radiomics model be used to predict if a lung nodule on computed tomography (CT) is likely to be small-cell lung cancer (SCLC)?

Knowledge Generated

Several machine learning classifiers were able to accurately identify whether a nodule was SCLC on either noncontrast (NC) and/or contrast-enhanced (CE) CT scans. A random forest radiomic classifier performed well on both NC and CE CT scans with an area under the receiver operating characteristic curve of 0.81 and 0.88, respectively, with six radiomic features found in common for both NC and CE scans that were predictive of SCLC.

Relevance

A machine learning radiomics model could be used for early prediction of SCLC on CT and allow for more rapid triage to tissue diagnosis, thus reducing the risk of stage migration.

determine response to treatment.[10-15] Most existing studies have aimed to categorize nodules as benign or malignant or to classify the subtype of malignancy. For radiomics to potentially allow appropriate triage of patients with SCLC, it will need to correctly identify SCLC among all types of nodules, benign and malignant. The purpose of this study is to determine if a radiomics model can adequately distinguish SCLC among histology of other nodules for lesions that are 1 cm or larger within the lung, specifically those that are benign, adenocarcinoma, and squamous cell carcinoma (SCC), for both contrast and noncontrast (NC) CT scans. Our subject population comes from the Veterans Health Administration (VA), the largest integrated health care system in the United States with 170 medical centers and over 1,000 outpatient care sites. As the incidence of smoking, a known risk factor for lung cancer, is high within VA patients,[16-18] it provides an opportunity to study radiomics in a population at great risk of lung cancer.

MATERIALS AND METHODS

Patient Selection

This retrospective study was approved by the Institutional Review Board overseeing research at both the VA Palo Alto Health Care System and Stanford University. All images were obtained from a single VA medical center. We obtained chest CT studies exhibiting two types: cancerous and benign nodules. Figure 1 displays a flow diagram comparing the study groups. Studies from December 2015 to 2018 were included in the analysis and were obtained from the local tumor registry and from a review of radiology reports for benign nodules. For the cancer studies, the inclusion criteria were presence of SCLC, adenocarcinoma, or SCC. Exclusion criteria included (1) no CT chest scan available prior to diagnosis; (2) tumor indistinct from atelectasis, mediastinum, or consolidation; (3) tumor < 1 cm in greatest diameter; (4) tumor predominantly ground glass or cystic; or (5) no standard kernel CT series available. For benign studies, inclusion criteria were presence of a solid nodule 1 cm or larger in greatest diameter and confirmation of benign lesion. We considered a lesion benign if the patient had (1) a biopsy with adequate tissue confirming the lesion was noncancerous, (2) follow-up imaging demonstrating resolution of the lesion within 2 years of the initial study, (3) serial follow-up chest CT imaging for at least 2 years confirming no change in the lesion, or (4) prior imaging from a minimum of 2 years earlier showing no change. For benign lesions, we considered exclusion criteria as (1) low-dose technique that we defined as a dose length product < 200 mGy-cm, (2) series reconstruction without a standard soft-tissue kernel with contiguous axial images 2.5 mm or less in thickness, (3) ground glass nodules, and (4) known active metastatic cancer. In both cases, lesions were limited to 1 cm or larger in greatest diameter for both technical and clinical concerns. Clinically, the risk of cancer for a nodule 1-1.5 cm (12.2% at baseline and 13.8% at 2 years) is over twice that for nodules 8-10 mm in size (5.8% at baseline and 6.4% at 2 years) with the risk increasing substantially with size.[19] In addition, only 13% of SCLC is found at < 1 cm in size in a screening population because of its aggressive nature, and we could not identify any in our lung cancer registry.[20] From a technical standpoint, a prior study has shown a minimum volume, typically $> 1{,}000$ mm^3, which is recommended to calculate some radiomic features.[21]

After final study selection, the available series for each study were reviewed, and series with contiguous axial 1.25 mm thickness slices with a standard soft-tissue kernel were chosen for all studies. The 1.25 mm thickness slices were used as this was consistently present in all studies. We selected only scans performed before any initial diagnosis or treatment intervention and collected age, sex, smoking, and lesion centrality data for all patients. A lesion was considered central if it directly communicated with the hilum or mediastinum. Finally, the data were divided into two groups: group A was composed of all NC scans within the data set and group B was composed of all contrast-enhanced (CE) scans within the data set. Within each

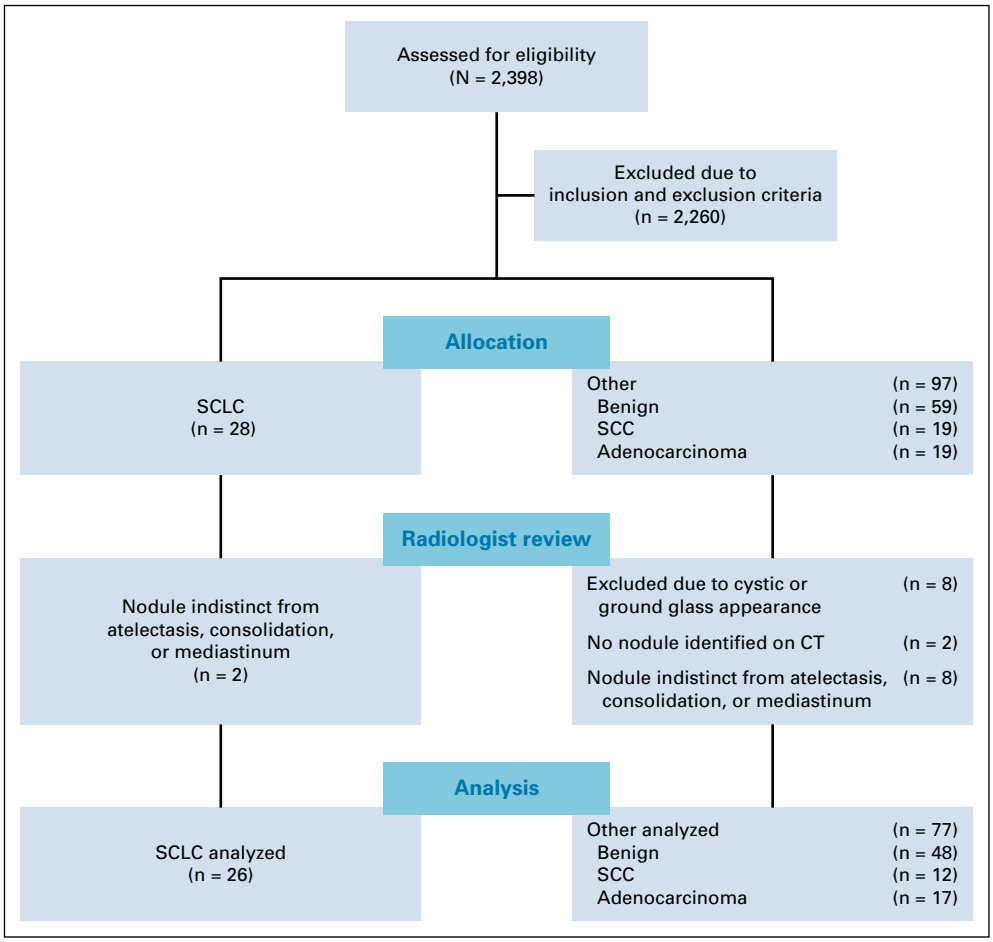

FIG 1. Patient selection flowchart for both cancer and benign nodules. CT chest scan reports were reviewed for presence of a nodule and then imaging selected on the basis of the listed exclusion criteria. CT, computed tomography; SCC, squamous cell cancer; SCLC, small-cell lung cancer.

group, studies were either classified as SCLC or other—with other composed of adenocarcinoma, SCC, or benign nodules.

Computational Analysis

A board-certified radiologist with 10 years of experience segmented all scans using the software program ITK-SNAP[22] using either autosegmentation or adaptive brush tools with manual adjustments in axial, coronal, and sagittal planes using both lung and abdomen viewing windows as needed. Eighteen scans, eight contrast and 10 NC, were randomly selected by an individual not performing segmentations and without viewing the scans. A second board-certified radiologist with 5 years of experience segmented these 18 scans using the same technique to measure interobserver agreement. We used the following rules for segmentation: (1) all solid areas of the tumor were included in the segmentation, (2) areas of tumor that were cystic were excluded from the segmentation, (3) on CE scans, lobar blood vessels were excluded from the segmentation, and (4) dense ground glass or nodularity contiguous with the lesion as visualized on lung windows were included in the segmentation. For image processing, resampling was performed to correct for voxel differences to a size of $1 \times 1 \times 1$ mm. We used PyRadiomics version 3.0[23] with default settings for feature extraction. Among the 120 features available, we selected only 105 features meeting the standards of the Image Biomarkers Standardization Initiative (IBSI).[24] Intraclass correlation coefficients (ICCs) were generated for the features using the segmentations for the 18 scans performed by two different radiologists. For feature reduction, only features with an excellent[25] ICC $\geq$ 0.75 were chosen for inclusion. *Z*-score normalization was performed across the features in the training set to generate a heatmap. Features from the following PyRadiomics classes were included: (1) first-order features, (2) shape features, (3) gray-level co-occurrence matrix, (4) gray-level run matrix, (5) neighboring gray-tone difference matrix, and (6) gray-level dependence matrix. The clinical characteristics in Table 1 were also included. Each classifier was trained using a leave-one-out-cross-validation (LOOCV). For each LOOCV, we applied (1) synthetic minority oversampling technique to overcome the class imbalance,[26-28] (2) *Z*-score normalization for scaling, and (3) minimum Redundancy Maximum Relevance for top-performing feature selection (n = 5).[29] The following four machine learning classifiers were used with the hyperparameters listed in Appendix Table A1: random forest (RF), support vector classifier (SVC), XGBoost (XGB), and logistic regression (LR). We applied the same method for

TABLE 1. Patient Demographics and Scanning Data of Included Studies

Clinical and Scanning Features	Benign	Adenocarcinoma	Squamous Cell	Benign plus NSCLC	Small Cell	*P*
Total	48	17	12	77	26	
Sex						1
Female	1	1	0	2	0	
Male	47	16	12	75	26	
Mean age (± SD), years	71.1 (9.9)	72.6 (10.5)	75.6 (6.9)	72.1 (9.7)	75.6 (7.9)	.062
Smoking history						.062
Yes	38	17	12	67	26	
No	10	0	0	10	0	
Central lesion						< .05
Yes	2	6	4	12	18	
No	46	11	8	65	8	
GE scanner model						.01
Discovery CT750 HD	48	17	12	77	23	
LightSpeed VCT	0	0	0	0	2	
Optima CT660	0	0	0	0	1	
CE						.0004
Yes	12	8	3	23	18	
No	36	9	9	54	8	
kVp						.559
100	1	0	0	1	0	
120	47	17	12	76	26	

Abbreviations: CE, contrast-enhanced; kVp, kilovolt peak; NSCLC, non–small-cell lung cancer; SD, standard deviation.

group A and group B. In addition, we created a clinical model using each classifier with sex, age, smoking, and lesion location features. Given prior work showing RF to have a high prognostic performance with high stability,[30] we created a final RF model using the intersection of all the top five features in the LOOCV. Feature importance was then determined from this final RF model.

Statistical Analysis

One-versus-rest modeling with a LOOCV was performed for SCLC versus other using each machine learning classifier for each group. Interobserver agreement was assessed for all 105 IBSI compliant features using ICC with a two-way random average measure. For each classifier, we calculated the mean area under the receiver operating characteristic curve (AUROC) and plotted the receiver operating characteristic (ROC) curve of the RF classifier with its 95% CI, calculated using 2,000 stratified bootstrap samples for each group. We used a two-sided DeLong's test[31] to compare the AUROC of the RF, SVC, XGB, and LR classification models and to compare clinical, radiomic, and clinical with radiomic models. We compared the clinical and scanning characteristics between the SCLC and other (benign and NSCLC) groups using an unpaired *t*-test for continuous variables and Fisher's exact test for categorical variables (Table 1). A P value $< .05$ was considered statistically significant for all comparisons. The final RF model was used to generate Shapley Additive Explanation (SHAP) plots demonstrating feature importance.[32] SHAP plots allow representation of the average contribution of a feature, either negative or positive, to the prediction when taking all possible combinations of features into account. Machine learning models were executed using Python 3.9 (Python Software Foundation, Beaverton, OR), and all statistical analyses were performed using R 4.0.2 (The R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

For the studies with cancerous nodules, a total of 28 patients with SCLC, 18 patients with adenocarcinoma, and 12 patients with SCC met the inclusion and exclusion criteria while a total of 48 patients met the inclusion and exclusion criteria for benign nodules (Fig 1). A total of 103 unique studies were included in the analysis with demographics and scanning data detailed in Table 1. Group A contained eight SCLC and 54 other NC scans, whereas group B contained 18 SCLC and 23 other CE scans. Using an ROC power test, group A had a power of 0.8 to detect an AUROC of 0.79 and group B had a power of 0.8 to detect an AUROC of 0.74.

A total of 59 IBSI-compliant radiomic features had an ICC ≥ 0.75 (Appendix Table A2) and were included in the classifier training. Figure 2 illustrates the heatmap generated after *Z*-score normalization before training. We trained the machine learning classifiers with and without inclusion of the clinical features; however, including them did not demonstrate a significant performance improvement ($P > .1$ for all classifiers in both groups; Appendix Table A3), and the results are reported without incorporation of these. The average AUROC for each classifier in group A and group B is listed in Table 2 for both the radiomics and clinical models. No classifier performed statistically better in either group ($P > .2$), except for LR, which performed better than SVC in group A ($P = .04$). In addition, the clinical model, despite having a lower AUROC across the board for all classifiers, did not have statistically lower performance except for LR and XGB classifiers in group A. Figure 3 shows the ROC curves for the RF classifier.

There were nine features that formed the union across all LOOCVs for group A and 14 features for group B. These features were used to generate a RF model, and the SHAP values for each feature are shown in Figure 4. The following six features overlapped between group A and B: firstorder_Range, firstorder_Skewness, glszm_LargeAreaLowGrayLevelEmphasis, glszm_SmallAreaEmphasis, glcm_Imc1, and glrlm_GrayLevelNonUniformityNormalized.

DISCUSSION

Our results demonstrate the possibility of using radiomics for distinguishing SCLC from other lung lesion histology. Lung cancer screening to date has demonstrated little impact in improving survival for patients with SCLC. The National Lung Screening Trial (NLST) found a significant benefit for low-dose screening CT in patients diagnosed with NSCLC, but no benefit was seen in patients diagnosed with SCLC, as the yearly time frame between screening examinations could allow for the growth and progression of tumor.[4,33] In the NLST trial, the 3-year survival was only 8% in those patients with ES-SCLC compared with 83% in those patients with LS-SCLC. Therefore, methods to detect SCLC earlier are needed, which is where radiomics may provide a solution.

Our study separated both contrast and NC studies into different cohorts. A prior study demonstrated a distinct difference in feature selection and feature importance between CE and NC scans.[34] Interestingly, the heatmap illustrated in Figure 2 shows a radiomic signature that appears similar in both CE and NC cases for SCLC, yet different from the other group suggesting that a single model may be able to predict SCLC, regardless of contrast administration. CE, however, is a confounding factor in our study as 18 of 26 SCLC were CE. Separate radiomic NC (group A) and CE (group B) models were therefore developed.

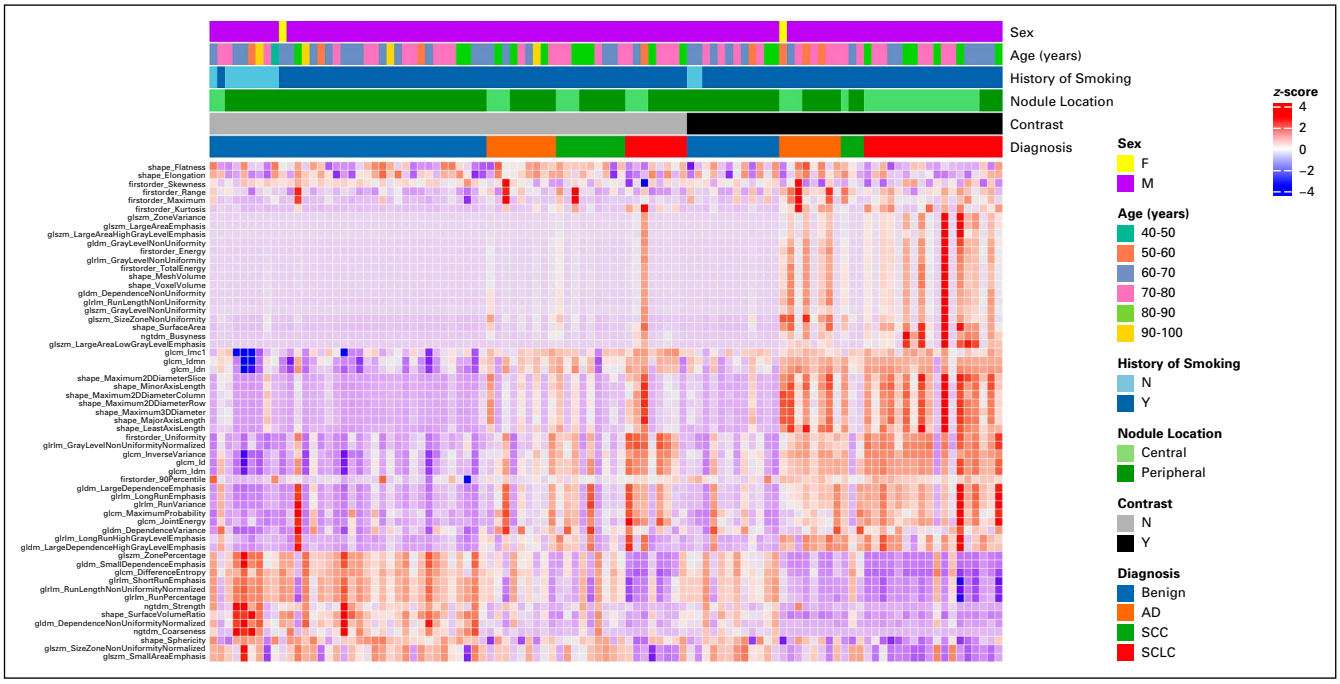

FIG 2. Heatmap demonstrating radiomic feature *Z*-score values for benign, SCLC, adenocarcinoma, and SCC across 59 features with a high intraclass correlation between two segmenters. Each column represents a unique patient. Contrast usage shows a different *Z*-score distribution for malignant tumors compared with benign lesions and NC scans. Therefore, separate models were generated for the use of contrast versus NC studies. AD, Adenocarcinoma; F, female; M, male; N, no; NC, noncontrast; SCC, squamous cell cancer; SCLC, small-cell lung cancer; Y, yes.

TABLE 2. AUROC (95% CI) for Different Machine Learning Classifiers

Classifier	Group A: Radiomic	Group A: Clinical	*P*	Group B: Radiomic	Group B: Clinical	*P*
RF	0.81 (0.59 to 0.99)	0.67 (0.46 to 0.86)	.14	0.88 (0.76 to 0.98)	0.82 (0.67 to 0.93)	.5
SVC	0.77 (0.5 to 0.98)	0.54 (0.29 to 0.78)	.2	0.87 (0.74 to 0.97)	0.73 (0.53 to 0.91)	.2
XGB	0.84 (0.66 to 0.99)	0.53 (0.24 to 0.82)	.02	0.85 (0.70 to 0.97)	0.82 (0.66 to 0.94)	.8
LR	0.84 (0.57 to 1.0)	0.47 (0.22 to 0.71)	.03	0.81 (0.64 to 0.95)	0.85 (0.71 to 0.96)	.7

Abbreviations: AUROC, area under the receiver operating characteristic curve; LR, logistic regression; RF, random forest; SVC, support vector classifier; XGB, XGBoost.

When comparing the clinical model with the radiomics model, it is not clear whether radiomics performs better than clinical features alone. The results for the clinical model for group A were underpowered, whereas the results for the clinical and radiomic models for group B were similar to one another. Lesion centrality is likely a driving factor on the clinical model, as SCLC is known to be more central in location,[35] and was the only statistically significant clinical factor between the SCLC and other cohorts. Regarding the combined model, a prior study showed no improvement in performance when clinical and radiomic features were combined, and our study had similar findings.[36] A larger sample size will likely help to determine the impact of clinical features on radiomics and is an area that our group is actively working on.

Six features were seen to overlap between the group A and group B RF models in predicting SCLC, and these may be feasible to use in creating a combined model. Given that the features were predominantly texture features, this may reflect the tumor heterogeneity of SCLC, as previously described.[34,37] Determining direct correlation between radiomic features and histologic features such as proliferation, necrosis, and vascularity is difficult with our small and heterogeneous data set but may be feasible if these features are found to be consistent in larger independent data sets.

Although this is the first published study demonstrating the efficacy of radiomics in distinguishing SCLC from other lung histology including benign lesions, other studies have found radiomics or deep learning to be effective in distinguish SCLC from other lung cancer histologies.[12,34,38,39] Similar to our experience, studies that have looked at SCLC versus NSCLC have generally performed well using a variety of classifiers, including deep learning, least absolute shrinkage and selection operator, support vector machine, and RF, but have struggled to separate SCLC from SCC.[12,34,38,39] Our findings add to the body of evidence that there may a unique radiomic signature for SCLC that can be used for earlier identification.

Clinically, if validated, this study could result in earlier detection of SCLC which may allow for earlier treatment assuming that barriers, including time and expertise for tumor segmentation, can be overcome. Reproduction of this study is critical to independent confirmation of the results. Therefore,

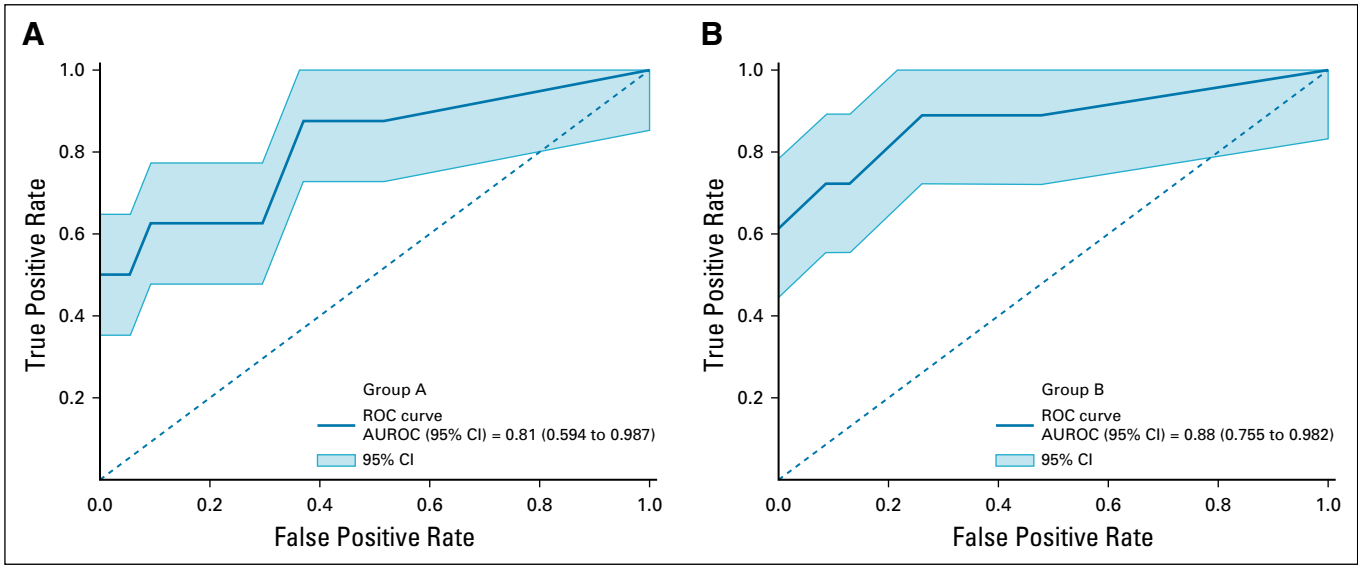

FIG 3. ROC for the RF classifier. The mean AUROC is obtained across all LOOCVs, and its performance in determining SCLC from all other lung nodules in (A) group A (NC) and (B) group B (CE) scans is shown. Note that the mean AUROC is > 0.8 in both groups, suggesting good discrimination of SCLC. AUROC, area under the receiver operating characteristic curve; CE, contrast-enhanced; LOOCV, leave-one-out-cross-validation; NC, noncontrast; RF, random forest; ROC, receiver operating characteristic curve; SCLC, small-cell lung cancer.

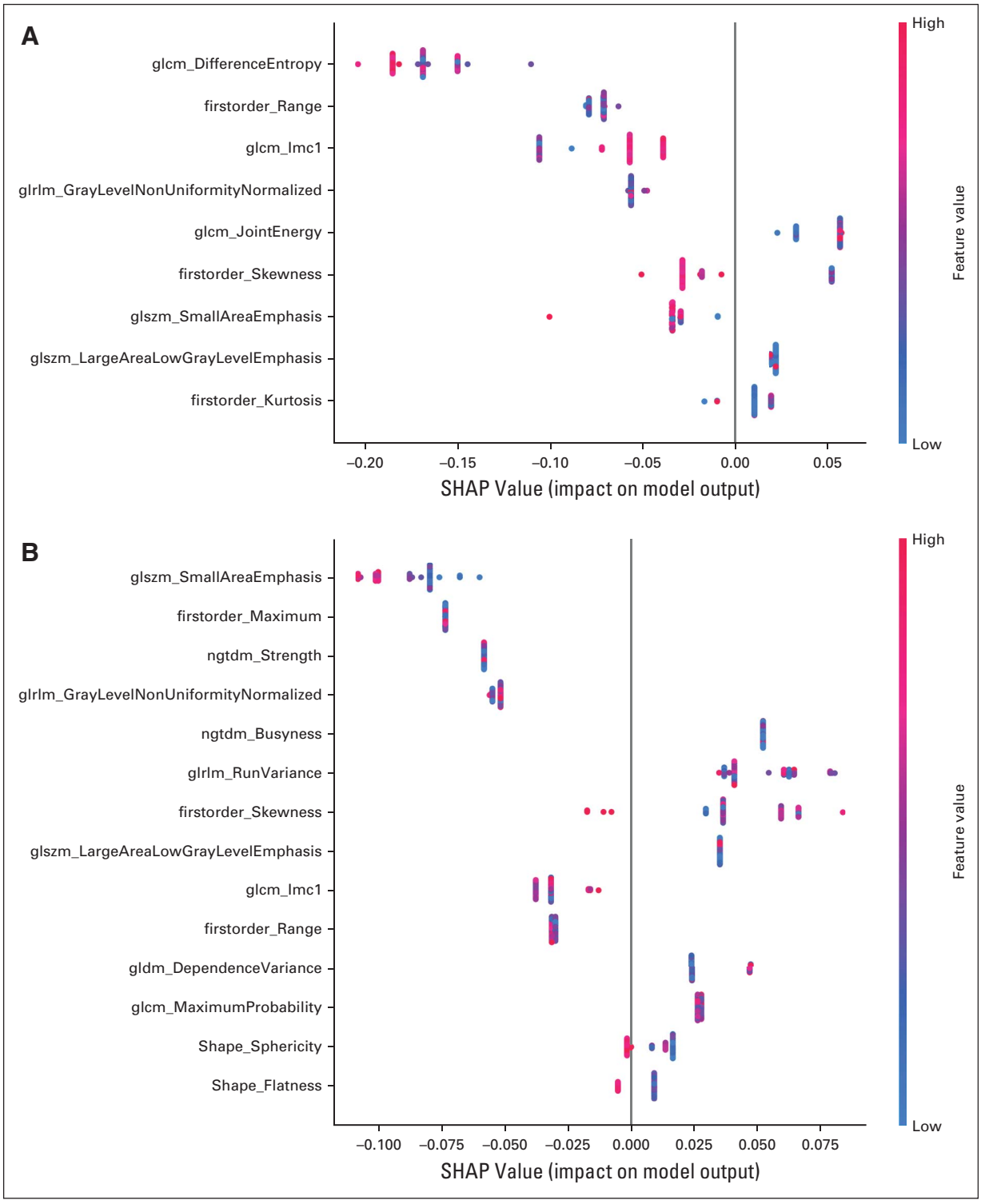

FIG 4. Feature importance on the basis of SHAP values as determined by the RF model for (A) group A and (B) group B scans. RF, random forest; SHAP, Shapley Additive Explanation.

the TRIPOD multivariable prediction model reporting guidelines were used.[40] On the basis of these guidelines, the current study is a TRIPOD 1b analysis type. Future studies should incorporate an independent test set to strengthen the model. Finally, the radiomic quality score (RQS)[41] for the study is a 5 (positive points for criteria 1, 2, 5, 6, 9, and 14; negative points for criteria 12). A recent literature review of 75 radiomics studies in NSCLC found the median RQS was 6 (interquartile range 2-12.25).[42] This study therefore has an RQS score in line with many existing studies.

There are several limitations to the existing study. The size of the existing study cohort is a major limitation, particularly because it prevented holding out some of the data for an independent validation. In addition, the imaging studies were from a single center, so no external validation was performed. Given the known challenges in reproducing radiomic results between institutions,[43] the existing results require further investigation with a multi-institutional cohort using correction algorithms to help mitigate differences between scanners and techniques.[44] This is particularly important since the data set used for this study did not have wide variations in scanning techniques aside from the use of contrast. Although this helps to minimize variables impacting a radiomics model creation, it may not be applicable to scans from other institutions. A final limitation is the gender imbalance of our existing data set. Given that the data were obtained from a VA hospital, the proportion of male patients is markedly larger than the proportion of female patients. As a result, it is uncertain whether the results will be applicable across genders.

In conclusion, radiomics may be an effective method to distinguish SCLC from other malignant and nonmalignant histology on both NC and CE CT. Further evaluation with a larger and multi-institutional data set is required to confirm the findings.

AFFILIATIONS

[1]Veterans Affairs Palo Alto Health Care System, Palo Alto, CA
[2]Department of Radiology, Stanford University, Stanford, CA
[3]Department of Medicine, Center for Biomedical Informatics Research (BMIR), Stanford University, Stanford, CA
[4]Palo Alto Veterans Institute for Research, Palo Alto, CA
[5]Present address: Department of Radiology, The Sixth Affiliated Hospital of Sun Yat-sen University, Guangzhou, Guangdong, China
[6]Present address: Department of Medical Imaging, Jinling Hospital, Nanjing University School of Medicine, Nanjing, Jiangsu, China
[7]Department of Medicine—Oncology, Stanford University, Stanford, CA

CORRESPONDING AUTHOR

Rajesh P. Shah, Veterans Affairs Palo Alto Health Care System, 3801 Miranda Ave, MC 114, Palo Alto, CA 94304; e-mail: rajshah@stanford.edu.

DISCLAIMER

The funding sources had no role in study design, in the collection, analysis and interpretation of data, in the writing of the report, or in the decision to submit the article for publication.

SUPPORT

Supported by an institutional grant from the Stanford University Artificial Intelligence in Medicine and Imaging Center; and by the National Institutes of Health grant number U24 CA180927. H.M.S. is supported by the Big Data Scientific Training Enhancement Program (BD-Step).

AUTHOR CONTRIBUTIONS

Conception and design: Rajesh P. Shah, Millie Das, Olivier Gevaert, Sandy Napel
Financial support: Olivier Gevaert, Rajesh P. Shah, Sandy Napel
Administrative support: Olivier Gevaert
Collection and assembly of data: Rajesh P. Shah, Shefali Verma, Peiyi Xie
Data analysis and interpretation: Rajesh P. Shah, Heather M. Selby, Pritam Mukherjee, Millie Das, Sachin Malik, Olivier Gevaert, Sandy Napel
Manuscript writing: All authors
Final approval of manuscript: All authors
Accountable for all aspects of the work: All authors

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The following represents disclosure information provided by the authors of this manuscript. All relationships are considered compensated unless otherwise noted. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/cci/author-center.
Open Payments is a public database containing information reported by companies about payments made to US-licensed physicians (Open Payments).

Rajesh P. Shah
Consulting or Advisory Role: Genentech, Intuitive Surgical, Histosonics Artio Medical, Inc
Research Funding: Merit Medical Systems

Millie Das
Honoraria: Plexus
Consulting or Advisory Role: Jazz Pharmaceuticals, AstraZeneca, BeiGene
Research Funding: Novartis, Verily, Varian Medical Systems, AbbVie, United Therapeutics, Cellsight Technologies

Olivier Gevaert
Research Funding: Nividien Inc, Lucence Diagnostics, Onc.AI

Sandy Napel
Consulting or Advisory Role: Fovia
Patents, Royalties, Other Intellectual Property: Several patents, none currently licensed

No other potential conflicts of interest were reported.

ACKNOWLEDGMENT

This material is the result of work supported with resources and the use of facilities at the VA Palo Alto Health Care System, Palo Alto, CA.

REFERENCES

1. Siegel RL, Miller KD, Jemal A: Cancer statistics, 2020. CA Cancer J Clin 70:7-30, 2020
2. American Cancer Society. *Cancer Facts & Figures 2020*. Atlanta: American Cancer Society; 2020
3. Howlader N, Noone AM, Krapcho M, et al: SEER Cancer Statistics Review, 1975-2017. https://seer.cancer.gov/csr/1975_2017/, 2020
4. National Comprehensive Cancer Network: Small Cell Lung Cancer (Version 1.2021). https://www.nccn.org/professionals/physician_gls/pdf/sclc.pdf

5. Harris K, Khachaturova I, Azab B, et al: Small cell lung cancer doubling time and its effect on clinical presentation: A concise review. Clin Med Insights Oncol 6:CMO.S9633, 2012
6. Common JL, Mariathas HH, Parsons K, et al: Reducing wait time for lung cancer diagnosis and treatment: Impact of a multidisciplinary, centralized referral program. Can Assoc Radiol J 69:322-327, 2018
7. Bhandari S, Kumar R, Pham D, et al: Treatment timing in small cell lung cancer, a national cancer database analysis. Am J Clin Oncol 43:362-365, 2020
8. Napel S, Mu W, Jardim-Perassi BV, et al: Quantitative imaging of cancer in the postgenomic era: Radio(geno)mics, deep learning, and habitats. Cancer 124:4633-4649, 2018
9. Li R, Xing L, Napel S, et al: Radiomics and Radiogenomics: Technical Basis and Clinical Applications. Boca Raton, FL, CRC Press, 2019
10. Chen C-H, Chang C-K, Tu C-Y, et al: Radiomic features analysis in computed tomography images of lung nodule classification. PLoS One 13:e0192002, 2018
11. Zhu X, Dong D, Chen Z, et al: Radiomic signature as a diagnostic factor for histologic subtype classification of non-small cell lung cancer. Eur Radiol 28:2772-2778, 2018
12. E L, Lu L, Li L, et al: Radiomics for classification of lung cancer histological subtypes based on nonenhanced computed tomography. Acad Radiol 26:1245-1252, 2019
13. Aerts HJWL, Velazquez ER, Leijenaar RTH, et al: Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 5:4006, 2014
14. Choi W, Oh JH, Riyahi S, et al: Radiomics analysis of pulmonary nodules in low-dose CT for early detection of lung cancer. Med Phys 45:1537-1549, 2018
15. Zhou M, Leung A, Echegaray S, et al: Non–small cell lung cancer radiogenomics map identifies relationships between molecular and imaging phenotypes with prognostic implications. Radiology 286:307-315, 2018
16. Bergman BP, Mackay DF, Morrison D, et al: Smoking-related cancer in military veterans: Retrospective cohort study of 57,000 veterans and 173,000 matched non-veterans. BMC Cancer 16:311, 2016
17. Brown DW: Smoking prevalence among US veterans. J Gen Intern Med 25:147-149, 2010
18. Odani S, Agaku IT, Graffunder CM, et al: Tobacco product use among military veterans—United States, 2010–2015. Morb Mortal Wkly Rep 67:7, 2018
19. Pinsky PF, Gierada DS, Nath PH, et al: Lung cancer risk associated with new solid nodules in the national lung screening trial. Am J Roentgenol 209:1009-1014, 2017
20. Austin JHM, Yip R, D'Souza BM, et al: Small-cell carcinoma of the lung detected by CT screening: Stage distribution and curability. Lung Cancer 76:339-343, 2012
21. Lee S-H, Cho H, Lee HY, et al: Clinical impact of variability on CT radiomics and suggestions for suitable feature selection: A focus on lung cancer. Cancer Imaging 19:54, 2019
22. Yushkevich PA, Piven J, Hazlett HC, et al: User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. Neuroimage 31:1116-1128, 2006
23. van Griethuysen JJM, Fedorov A, Parmar C, et al: Computational radiomics system to decode the radiographic phenotype. Cancer Res 77:e104-e107, 2017
24. Zwanenburg A, Leger S, Vallières M, et al: Image biomarker standardisation initiative-feature definitions. ArXiv, 2016.Prepr ArXiv161207003
25. Cicchetti DV: Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess 6:284-290, 1994
26. Chawla NV, Bowyer KW, Hall LO, et al: SMOTE: Synthetic minority over-sampling technique. J Artif Intell Res 16:321-357, 2002
27. Tunali I, Gray JE, Qi J, et al: Novel clinical and radiomic predictors of rapid disease progression phenotypes among lung cancer patients treated with immunotherapy: An early report. Lung Cancer 129:75-–79, 2019
28. Wu L, Gao C, Xiang P, et al: CT-imaging based analysis of invasive lung adenocarcinoma presenting as ground glass nodules using peri- and intra-nodular radiomic features. Front Oncol 10:838, 2020
29. Auffarth B, López M, Cerquides J: Comparison of redundancy and relevance measures for feature selection in tissue classification of CT images, in: Perner P (ed): Advances in Data Mining. Applications and Theoretical Aspects. Berlin, Heidelberg, Germany, Springer, 2010. pp 248-262. http://link.springer.com/10.1007/978-3-642-14400-4_20
30. Parmar C, Grossmann P, Bussink J, et al: Machine learning methods for quantitative radiomic biomarkers. Sci Rep 5:13087, 2015
31. DeLong ER, DeLong DM, Clarke-Pearson DL: Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. Biometrics 44:837, 1988
32. Lundberg SM, Lee S-I: A Unified Approach to Interpreting Model Predictions. NIPS 2017
33. Thomas A, Pattanayak P, Szabo E, et al: Characteristics and outcomes of small cell lung cancer detected by CT screening. Chest 154:1284-1290, 2018
34. E L, Lu L, Li L, et al: Radiomics for classifying histological subtypes of lung cancer based on multiphasic contrast-enhanced computed tomography. J Comput Assist Tomogr 43:300-306, 2019
35. Carter BW, Glisson BS, Truong MT, et al: Small cell lung carcinoma: Staging, imaging, and treatment considerations. Radiographics 34:1707-1721, 2014
36. Grossmann P, Stringfield O, El-Hachem N, et al: Defining the biological basis of radiomic phenotypes in lung cancer. Elife 6:e23421, 2017
37. Chen BT, Chen Z, Ye N, et al: Differentiating peripherally-located small cell lung cancer from non-small cell lung cancer using a CT radiomic approach. Front Oncol 10:593, 2020
38. Liu S, Liu S, Zhang C, et al: 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, 2020
39. Guo Y, Song Q, Jiang M, et al: Histological subtypes classification of lung cancers on CT images using 3D deep learning and radiomics. Acad Radiol: S1076633220303603, 2020
40. Collins GS, Reitsma JB, Altman DG, et al: Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. Eur Urol 67:1142-1151, 2015
41. Lambin P, Leijenaar RTH, Deist TM, et al: Radiomics: The bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 14:749-762, 2017
42. Fornacon-Wood I, Faivre-Finn C, O'Connor JPB, et al: Radiomics as a personalized medicine tool in lung cancer: Separating the hope from the hype. Lung Cancer 146:197-208, 2020
43. Steiger P, Sood R: How can radiomics Be consistently applied across imagers and institutions? Radiology 291:60-61, 2019
44. Orlhac F, Frouin F, Nioche C, et al: Validation of a method to compensate multicenter effects affecting CT radiomics. Radiology 291:53-59, 2019

APPENDIX

TABLE A1. Hyperparameter Selections for Each Machine Learning Classifier

Classifier	Hyperparameter	Value
SVC	C	1
	kernel	rbf
RF	n_estimators	10
XGB	booster	gbtrees
	learning_rate	0.3
	max_depth	6
LR	C	1
	penalty	l2
	solver	lbfgs

Abbreviations: LR, logistic regression; RF, random forest; SVC, support vector classifier; XGB, XGBoost.

TABLE A2. Radiomic Features With an ICC ≥ 0.75

Feature	ICC	95% CI Lower Bound	95% CI Upper Bound
gldm_GrayLevelNonUniformity	0.999188742	0.998189735	0.999640426
glrlm_GrayLevelNonUniformity	0.998901332	0.997548832	0.999512997
firstorder_Energy	0.997293215	0.993967043	0.998799636
firstorder_TotalEnergy	0.99727419	0.993924709	0.998791193
shape_VoxelVolume	0.996601313	0.992428136	0.998492511
shape_MeshVolume	0.996598377	0.992421607	0.998491208
gldm_DependenceNonUniformity	0.995026989	0.988931446	0.997793251
shape_LeastAxisLength	0.994650069	0.988067409	0.997627892
glrlm_RunLengthNonUniformity	0.992732598	0.983847532	0.996773062
shape_Maximum2DDiameterColumn	0.991856351	0.98187693	0.996385629
shape_Maximum3DDiameter	0.988741408	0.9607486	0.995698627
glszm_GrayLevelNonUniformity	0.987466836	0.972233582	0.994426724
shape_Maximum2DDiameterRow	0.98686546	0.968561226	0.994328378
firstorder_90Percentile	0.984809538	0.86643516	0.995296782
shape_Maximum2DDiameterSlice	0.984416614	0.964259625	0.993161479
shape_MajorAxisLength	0.984381713	0.945137698	0.994047568
glszm_LargeAreaHighGrayLevelEmphasis	0.984225302	0.965121302	0.992978913
glszm_ZoneVariance	0.983973617	0.964499133	0.992871906
glszm_LargeAreaEmphasis	0.983936615	0.964418471	0.992855336
shape_MinorAxisLength	0.978243512	0.952070185	0.990300283
shape_SurfaceArea	0.977783174	0.951069749	0.990093773
glrlm_RunVariance	0.973735635	0.93238815	0.988950275
gldm_DependenceVariance	0.968065751	0.856747199	0.988630447
glrlm_LongRunEmphasis	0.960296924	0.900484194	0.983132391
shape_SurfaceVolumeRatio	0.959631869	0.637591807	0.987766254
gldm_LargeDependenceEmphasis	0.956473304	0.882500884	0.981988319
glszm_SizeZoneNonUniformity	0.954711069	0.901634893	0.97967486
ngtdm_Busyness	0.951793208	0.893325519	0.978521396
shape_Flatness	0.945318466	0.881899236	0.975394441
glrlm_RunPercentage	0.931808876	0.794752485	0.972820744
shape_Sphericity	0.928589451	0.847131163	0.967725723
glrlm_ShortRunEmphasis	0.916234097	0.77184861	0.965459957
glszm_SizeZoneNonUniformityNormalized	0.914024147	0.772784852	0.964162242
gldm_LargeDependenceHighGrayLevelEmphasis	0.913149763	0.815992153	0.960556215
glcm_MaximumProbability	0.911769291	0.712251687	0.965797929
glcm_Imc1	0.911426413	0.812539825	0.959752632
glszm_SmallAreaEmphasis	0.911182117	0.756510721	0.963465694
glrlm_RunLengthNonUniformityNormalized	0.906733497	0.718597442	0.962906643
glszm_LargeAreaLowGrayLevelEmphasis	0.904975992	0.798328312	0.95685351
ngtdm_Coarseness	0.901086031	0.763748166	0.957164926
firstorder_Kurtosis	0.892812127	0.775108075	0.951071872
glcm_JointEnergy	0.887567802	0.712427132	0.95257282
glcm_Idn	0.881480003	0.494935113	0.958089597
firstorder_Maximum	0.873315746	0.732526423	0.94226883

(Continued on following page)

TABLE A2. Radiomic Features With an ICC ≥ 0.75 (Continued)

Feature	ICC	95% CI Lower Bound	95% CI Upper Bound
glcm_Idmn	0.868286316	0.532345805	0.950674106
glszm_ZonePercentage	0.860478781	0.496005678	0.948149893
gldm_SmallDependenceEmphasis	0.856666675	0.497941244	0.946197705
glrlm_LongRunHighGrayLevelEmphasis	0.85445992	0.701792256	0.932785582
glcm_Idm	0.833055777	0.452270108	0.936129385
firstorder_Uniformity	0.82233652	0.554715165	0.924548289
shape_Elongation	0.821788585	0.641984304	0.916843846
firstorder_Skewness	0.819010799	0.583842287	0.920325396
glcm_Id	0.814893846	0.397330529	0.929253066
ngtdm_Strength	0.807968207	0.61526936	0.910201657
gldm_DependenceNonUniformityNormalized	0.805001214	0.305793779	0.928358904
glcm_InverseVariance	0.78754023	0.319286769	0.918852083
glrlm_GrayLevelNonUniformityNormalized	0.769461679	0.487740609	0.896963895
firstorder_Range	0.760137096	0.533343185	0.886116171
glcm_DifferenceEntropy	0.756938648	0.315450897	0.902442008

Abbreviation: ICC, intraclass correlation coefficient.

TABLE A3. AUROC of the Model Using Clinical and Radiomic Features Versus Radiomic Only Features

	Clinical and Radiomic Features		Radiomic Features Only		*P* Value Group A	*P* Value Group B
Classifier	Group A: AUROC (95% CI)	Group B: AUROC (95% CI)	Group A: AUROC (95% CI)	Group B: AUROC (95% CI)		
RF	0.82 (0.597 to 0.994)	0.93 (0.851 to 0.988)	0.81 (0.594 to 0.987)	0.88 (0.755 to 0.982)	0.368	0.224
SVC	0.77 (0.499 to 0.979)	0.88 (0.758 to 0.966)	0.77 (0.499 to 0.979)	0.87 (0.74 to 0.966)	1	0.832
XGB	0.84 (0.659 to 0.986)	0.87 (0.749 to 0.968)	0.84 (0.659 to 0.986)	0.85 (0.698 to 0.971)	1	0.699
LR	0.83 (0.554 to 0.991)	0.88 (0.751 to 0.978)	0.84 (0.565 to 1.0)	0.81 (0.64 to 0.954)	0.368	0.103

Abbreviations: AUROC, area under the receiver operating characteristic curve; LR, logistic regression; RF, random forest; SVC, support vector classifier; XGB, XGBoost.