Lack of Incremental Value of Three-Dimensional Measurement in Assessing Invasiveness for Lung Cancer

Meng-Min Wang¹, MD, Jia-Qi Li², PhD, Shi-Hua Dou, MD³, Hong-Ji Li⁴,¹, MD, Zhen-Bin Qiu¹,⁴, MD, PhD, Chao Zhang¹,⁴, MD, PhD, Xiong-Wen Yang¹,⁴, MD, PhD, Jia-Tao Zhang¹, MD, PhD, Xin-Hua Qiu¹, MS, Hong-Sheng Xie³, MD, Wen-Fang Tang⁵, MD, Mei-Ling Cheng¹, MD, Hong-Hong Yan¹, MD, Xue-Ning Yang¹, MD, PhD, Yi-Long Wu¹, MD, Xue-Gong Zhang²,⁶, PhD, Lin Yang, MD, PhD³, Wen-Zhao Zhong, MD, PhD¹,⁴

1. Guangdong Lung Cancer Institute, Guangdong Provincial People’s Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou 510080, China.
2. Bioinformatics Division, BNRIST and MOE Key Lab of Bioinformatics, Department of Automation, Tsinghua University, Beijing, China.
3. Department of Thoracic Surgery, The Second Clinical Medical College, Jinan University (Shenzhen People’s Hospital), Shenzhen, China.
4. School of Medicine, South China University of Technology, Guangzhou, 510006, China.
5. Department of Cardiothoracic Surgery, Zhongshan City People’s Hospital, Zhongshan, China.
6. School of Medicine, Tsinghua University, Beijing, China.

#: These authors contributed equally.
*Corresponding author:
Lin Yang, 13798314779@163.com; The Second Clinical Medical College, Jinan

© The Author(s) 2023. Published by Oxford University Press on behalf of the European Association for Cardio-Thoracic Surgery. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
Abstract

Objective To evaluate the performance of consolidation tumor ratio and the radiomic models in two- and three-dimensional modalities for assessing radiological invasiveness in early-stage lung adenocarcinoma.

Methods A retrospective analysis was conducted on patients with early-stage lung adenocarcinoma from Guangdong Provincial People's Hospital and Shenzhen People's Hospital. Manual delineation of pulmonary nodules along the boundary was performed on cross-sectional images to extract radiomic features. Clinicopathological characteristics and radiomic signatures were identified in both cohorts. Consolidation tumor ratio and radiomic score for every patient was calculated. The performance of consolidation tumor ratio and radiomic models were tested and validated in the respective cohorts.

Results A total of 818 patients from Guangdong Provincial People's Hospital were included in the primary cohort, while 474 patients from Shenzhen People's Hospital
constituted an independent validation cohort. Both consolidation tumor ratio and Rad-
score were identified as independent factors for predicting pathological invasiveness.
Consolidation tumor ratio in two- and three-dimensional modalities exhibited
comparable results with areas under the receiver operating characteristic curves, and
were demonstrated in the validation cohort (AUC: 0.807 vs. 0.826, $p = 0.059$)
Furthermore, both consolidation tumor ratio in two- and three-dimensional modalities
were able to stratify patients with significant relapse-free survival ($p < 0.000$ vs. $p <
0.000$) and overall survival ($p = 0.003$ vs. $p = 0.001$). The radiomics models in two- and
three-dimensional modalities demonstrated favorable discrimination and calibration in
independent cohorts ($p = 0.189$).

**Conclusion** Three-dimensional measurement provides no additional clinical benefit
compared to two-dimensional.

**Keywords** Early-stage Lung Adenocarcinoma, Surgery, Consolidation-to-tumor
Ratio, Radiomic Models.

**Lists of Abbreviations:**

- CTR: consolidation tumor ratio
- MTD: maximum tumor diameters
- MSCD: maximum solid component diameter
- TTV: total tumor volume
- SCV: solid component volume

**Introduction**
The use of low-dose helical computed tomography (LDCT) or high-resolution computed tomography (HRCT) has led to a higher number of early-stage lung cancer diagnoses [1]. Traditionally, lobectomy with systematic lymph node dissection has been the standard treatment for early non-small cell lung cancers (NSCLCs) [2, 3]. However, in recent years, there has been a growing trend toward limited surgical resections such as lung segmentectomy and wedge resection for early lung cancer patients [4, 5]. The Cancer and Leukemia Group B-140503 (CALGB-140503) study demonstrated that segmentectomy was comparable to pulmonary lobectomy in terms of survival for patients with early-stage NSCLCs [6]. Additionally, the JCOG0802/WJOG4607L clinical study emphasized the importance of accurate preoperative determination of invasiveness in lung cancer patients for prognosis and treatment decisions [7].

The consolidation tumor ratio (CTR) has shown potential as a parameter to assess prognosis and measure invasiveness in early lung adenocarcinoma [8, 9]. The Japan Clinical Oncology Group (JOCG) 0201 study demonstrated that invasive pulmonary nodules could be identified preoperatively by calculating CTR, which represents the proportion of solid component diameter within the entire tumor [10]. However, this was a planimetric method, and advancements in technology now allow for tumor volume extraction from CT images. Several studies have suggested that three-dimensional CTR (3D-CTR) provides greater accuracy than two-dimensional CTR (2D-CTR) for preoperative diagnosis [11, 12]. Unfortunately, these studies were limited by small sample sizes and a lack of external validation cohorts, thus preventing definitive
conclusions on the superior accuracy between the two methods.

With the rise in popularity of artificial intelligence, the concept of radiomics has emerged, referring to the process of extracting quantitative features from medical images to generate high-dimensional, mineable data. This may enhance predictive accuracy. These features can be derived from all layers involved (3D radiomic features) or the largest transverse area of the lesion (2D radiomic features) [13, 14]. While several studies have demonstrated the utility of texture features in predicting prognosis or diagnosis in various cancers, there is currently no definitive literature regarding the comparative performance of 2D and 3D radiomic signatures in predicting invasiveness specifically in lung adenocarcinoma [15-16].

In this study, we conducted a comparative analysis to evaluate the accuracy of 2D-CTR or 3D-CTR as well as 2D or 3D radiomic signature in quantifying the degree of invasiveness among patients with ground-glass nodules.

**Patients and Methods**

**Ethics Statement**

The study received approval from the Research Ethics Committee of Guangdong Provincial People’s Hospital, Guangdong Academy of Medical Sciences (No. GDREC2018115H, No. GDREC2019726H) on January 30th, 2018 and November 20th, 2019. Informed consent from patients was waived for these projects.
The primary cohort comprised patients who underwent curative surgical procedures at Guangdong Provincial People’s Hospital between January 2011 and October 2019. Patients from Shenzhen People’s Hospital between January 2018 and October 2021 were included into the independent validation cohort. Supplementary Figure S1 provides the inclusion and exclusion criteria. Clinicopathological baseline characteristics such as age, gender, smoking history, location, and preoperative clinical stage were obtained from medical records. CT imaging was also collected for these cases.

**CT Image Acquisition and Segmentation**

All preoperative CT images of the primary cohort were conducted using these four scanners: GE Discovery CT750 HD, 64-slice LightSpeed VCT (GE Medical Systems), SOMATOM Definition Flash, SOMATOM Sensation-64 (Siemens Medical Solutions), Ingenuity CT-64 (Philips), Ingenuity CT-256 (Philips). The scanning parameters were as follows: 120 kVp, 100–670 mAs, pitch 0.5–1.5, and collimation 0.60–1.25 mm. Reconstruction of all imaging data were performed using a medium sharp reconstruction algorithm with a thickness range of 0.60–1.25 mm. A region of interest (ROI) was delineated along the tumor borderline in each cross-sectional area, using ITK-SNAP software (Version 3.6.0, America), with the same lung window settings (window width, 1500HU; window level, -600HU) for extracting image features. Structures such as pulmonary vessels that overlapped with the pulmonary nodules were manually excluded when delineating the boundaries. Two radiologists, one with
3 years of experience (reader 1) and the other with 10 years of experience (reader 2) in pulmonary CT interpretation, independently performed the ROI segmentation in a blinded manner. Reader 1 also completed the residual images. The inter- and intra-observer agreements were assessed using inter- and intra-class correlation coefficients (ICCs) on 70 randomly ROI-based images. Features with ICCs above 0.75 were considered to demonstrate good agreement and were selected (see Supplementary Table S1).

**Evaluation of 2D-CTR and 3D-CTR**

The solid component refers to an area of increasing opacification that progressively obscures the underlying vascular or bronchial structures, whereas the ground-glass opacities (GGO) are characterized by hazy density that does not obscure the underlying markings (see Supplementary Figure S2). Ishikawa et al [12] explained that the distinction between GGO and solid components is based on CT attenuation. The solid component is defined as the area with CT attenuation of −300HU or higher, while the GGO component is defined as the area with CT attenuation values ranging from −800HU to −300HU. The 2D-CTR is calculated as the ratio of the maximum solid component diameters (MSCD) to the maximum tumor diameters (MTD) in the largest horizontal section. On the other hand, the 3D-CTR is determined by the ratio of the volume of the solid parts (SCV) to the total tumor volume (TTV) (see Supplementary Figure S3).
Feature Extraction

The CT images were subjected to 2D/3D-CTR analysis and texture analysis using algorithms in Python programming (version 3.7, Netherlands). The 3D texture features included 7 feature classes, consisting of 14 shape-based features, 12 first-order statistics features, 24 Gray Level Co-occurrence Matrix (GLCM) features, 16 Gray Level Run Length Matrix (GLRLM) features, 16 Gray Level Size Zone Matrix (GLSZM) features, 14 Gray Level Dependence Matrix (GLDM) features, and 5 Neighboring Gray Tone Difference Matrix (NGTDM) features. The 2D texture, on the other hand, encompassed included 9 shape-based features, 18 first-order statistics features, 24 GLCM features, 16 GLRLM features, 16 GLSZM features, 14 GLDM features, and 5 NGTDM features. In total, all 172598 texture analyses were performed, comparing 2D and 3D textures with or without image filtration.

Model Development

The least absolute shrinkage and selection operator (LASSO) regression model identified the coefficients of specific features that were non-zero, and these coefficients were used to calculate the radiomic score (Rad-score) for each patient (see Supplementary Table S2). The 2D and 3D Rad-score models were developed using backward step-wise selection analysis in the internal training cohort (see Supplementary Figure S4).

Evaluation of Pathological Subtypes
The pathological findings for pulmonary adenocarcinoma in clinical IA stage patients were assessed based on the International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory (IASLC/ATS/ERS) classification. This classification system includes various subtypes such as atypical adenomatous hyperplasia, adenocarcinoma in situ, minimally invasive adenocarcinoma, lepidic-predominant adenocarcinoma, acinar-predominant adenocarcinoma, papillary-predominant adenocarcinoma, micropapillary-predominant adenocarcinoma, solid-predominant adenocarcinoma and others. The classification was based on the subtype with the highest proportion among the observed subtypes [17]. In this study, pulmonary nodules classified as AAH, AIS, and MIA were considered less-invasive pulmonary nodules, while the remaining subtypes were classified as invasive.

Follow-up Strategy

All patients were required to undergo a physical examination at regular intervals of no less than 6 months for a minimum duration of 3 years following the surgical procedures. Additional examinations were conducted as per the specific requirements of each patient's condition. Relapse-free survival (RFS) was defined as the period starting from the day of the operation until the occurrence of local or distant recurrence or death from any cause. Overall survival (OS) was measured from the time of the operation until either the date of death or the last follow-up day.
Statistical Analysis

The statistical analysis was conducted using SPSS software (version 26.0, USA), R (version 4.1.0, Austria) and GraphPad Prism (version 8.0.2, America). Continuous variables were reported as means and standard deviations (SD) or as median and interquartile ranges (IQR), while categorical variables were presented as frequencies and percentages. To compare the differences between patients with invasive and less-invasive nodules in both the primary and validation cohorts, Student's t-test or the Mann-Whitney U test was applied for continuous variables, while the Chi-squared test or Fisher's exact test was used for categorical variables.

Univariable and multivariable logistic regression analyses were conducted to identify independent predictive factors for the invasiveness of pulmonary nodules. The LASSO approach was utilized to identify potential predictive features from the primary cohort. A Rad-score was calculated combining the selected features with their respective coefficients. The backward step-wise selection was employed with the likelihood ratio test and Akaike’s information criterion. Discriminatory analysis was conducted using the area under the curve (AUC), and the difference between the operating characteristic curves (ROC) was assessed using De-Long’s test. Calibration curves were used to evaluate the calibration performance of the radiomic models. The relationship between related 2D and 3D radiomic parameters was assessed using Pearson’s test.

The associations of 2D-CTR and 3D-CTR with pathological invasiveness of pulmonary nodules were evaluated on restricted cubic splines (RCS) with four knots.
The reference points for these analyses were the CTR levels with the risk of pathological invasiveness equaling to one, with knots placed at 5th, 35th, 65th, 95th of each 2D-CTR and 3D-CTR distribution. Odds ratio (OR) was used to determine the degree of association. Survival outcomes, including RFS and OS, were analyzed using the Kaplan-Meier method and compared with Log-rank test. Cumulative incidence functions for the competing event were calculated by competing risks methodology. Analysis of the cumulative incidence of relapse considered death without recurrence as one competing risk, and Fine and Gray’s competing risks regressions were used to estimate the sub-hazard ratio (sHR) to evaluate the association between variables and risk of relapse. We included age, smoking history, sex and pathology in risk models. A p-value <0.05 was considered statistically significant.

Results

Baseline Characteristics for Enrolled Patients Between Primary Cohort and Validation Cohort

In the primary cohort, the internal training and testing sets were randomly divided in an 8:2 ratio using the R package. The baseline characteristics of both cohorts are presented in Supplementary Tables S3 and S4. The primary cohort consisted of 818 patients, while the independent validation cohort included 474 patients. Based on histological subtypes, the primary cohort was divided into two groups: the less-invasive adenocarcinoma group with 188 pulmonary nodules and the invasive group with 630 pulmonary nodules. In both cohorts, the patients in the invasive group were elder (p <
0.001), predominantly female (primary cohort: $p = 0.004$; independent cohort: $p = 0.003$), and mostly non-smokers (primary cohort: $p = 0.008$; independent cohort: $p = 0.002$) compared to the less-invasive group. Specifically, the less-invasive group had patients with IA1 stage pulmonary adenocarcinoma, while invasive group has more patients with IA2 stage lung adenocarcinoma ($p < 0.001$). In the primary cohort, lobectomy was the preferred treatment in the invasive group [452 (71.7%)], while wedge resection was more commonly used for less-invasive group [104 (55.3%)]. In the validation cohort, on the other hand, lobectomy was similar between the less-invasive and invasive groups [221 (84.0%) vs. 188 (89.1%)].

**Predictive Performance of CTR in the Primary Cohort**

Univariable and multivariable analyses were conducted to evaluate the clinicopathological characteristics as predictors of invasiveness in pulmonary nodules. In the univariable analysis of 2D-related characteristics, smoking status, age, gender, MTD, MSCD, and 2D-CTR were identified as significant predictors for distinguishing less-invasive nodules from invasive adenocarcinoma (see Supplementary Table S5). Similarly, in the univariable analysis of 3D-related characteristics, smoking status, age, gender, TTV, SCV, and 3D-CTR showed significant predictive value (see Supplementary Table S6). These predictors were further subjected to multivariable analysis, which revealed that MTD, 2D-CTR, TTV, SCV, and 3D-CTR were independent factors for predicting invasiveness.

The predictive performance of 2D-CTR and 3D-CTR for tumor invasiveness was
evaluated and presented in Figure 1. The AUC for 2D-CTR was 0.828 (95% CI: 0.798–0.858), and for 3D-CTR, it was 0.830 (95% CI: 0.800–0.859) in the entire primary cohort ($p = 0.83$). Specifically, their sensitivity and specificity were comparable (Figure 1A and Table 1). Further analysis was conducted in the internal training cohort, and there was no significant difference between 2D-CTR and 3D-CTR ($p = 0.94$) (Figure 1B and Table 1).

Validation of CTR

Internal validation. Both 2D-CTR and 3D-CTR showed good performance with AUC of 0.827 [95% CI: 0.756–0.898] and 0.839 [95% CI: 0.776–0.902], respectively. The difference between the two methods was not statistically significant ($p = 0.60$) (Figure 1C). The specificity of the 2D-CTR method was slightly lower than that of 3D-CTR method, as reported in Table 1.

External validation. The performance of 2D-CTR and 3D-CTR with AUC was further confirmed in the independent set, which yielded a nonsignificant statistic ($p = 0.059$, Figure 1D). The specificity and sensitivity values are included in Table 1.

RCS and Survival Analysis

The associations between CTR values and risk of invasiveness were visualized in Figure 2A&B. We found that CTRs were significantly associated with the risk of invasiveness ($p <0.0001$ for overall association). No statistically significant non-linear association was found in 3D-CTR with the risk of invasiveness ($p =0.177$ for the non-
linear test). It was sharply increased when 2D-CTR > 0.83 and 3D-CTR > 0.34. Regarding to clinical backgrounds, patients were divided into low-risk or high-risk invasive groups based on 0.83 and 0.34, respectively.

The median follow-up time was 49.63 months. Thirty-one patients experienced relapse during follow-up time. Significant differences were observed in terms of RFS (2D-CTR, Figure 3A, \( p < 0.000 \); 3D-CTR, Figure 3C, \( p < 0.000 \)) and OS (2D-CTR, Figure 3B, \( p = 0.003 \); 3D-CTR, Figure 3D, \( p = 0.001 \)). A comparison of competing outcomes found significant changes in high-risk group and low-risk group. The risk of relapse was higher in the high-risk group (2D-CTR: sHR, 5.10, 95%CI, 1.87-14.0, \( p=0.002 \); 3D-CTR: sHR, 6.81, 95%CI, 2.21-20.9, \( p<0.001 \)) (see Supplementary Figure S5).

Performance of Radiomic Signature Models

Following the selection of relevant features, the Rad-score was calculated using six potential features from the 2D modality and seven potential features from the 3D modality (see Supplementary Table S2). The backward step-wise selection process identified the Rad-score as an independent predictor of pulmonary invasiveness, leading to development of separate models for 2D and 3D modalities. In our internal training cohort, there was no significant difference in the AUC between these two models (2D vs. 3D: 0.923 [95% CI: 0.901–0.945] vs. 0.926 [95% CI: 0.906–0.947], \( p = 0.38 \)) (Figure 4A and Table 2). Furthermore, the radiomic models were validated in the internal testing and independent cohort, yielding similar results as shown in Figure
4B&4C and Table 2. Calibration curves demonstrated the high consistency between the predicted and observed invasiveness probability in our models (Figure 5).

Discussion

The development of an accurate and convenient diagnostic method is crucial for improving surgeon confidence in predicting preoperative radiological invasiveness and determining appropriate treatment strategies. In this regard, both CTR and radiomic features have been gradually applied for evaluating radiological diagnoses in early lung cancer [19]. A series of studies conducted by the JCOG have shed light on the significant correlation between the 2D-CTR and different types of surgical interventions. The JCOG0201 study, for instance, revealed a favorable prognosis for tumors smaller than 3cm with a CTR of less than 0.5 when treated with sub-lobar resection or lobectomy [20]. Additionally, the JCOG0804 study reported a remarkable 5-year OS rate of approximately 100% for radiological non-invasive pulmonary nodules treated with sub-lobar resection [21]. Likewise, JCOG0802 study also emphasized the criticality of sub-lobar resection in the management of tumors ≤2.0 cm with ≤0.25 consolidation [7]. However, previous small-scale studies have suggested that 3D-CTR outperforms 2D-CTR in the radiological diagnosis of invasive lung adenocarcinoma [19, 22]. On the other hand, the convenience factor, which is important for surgeons in outpatient clinics, should also be considered. Therefore, before wide application, it is necessary to confirm the aforementioned conclusion. In this study, we established large-scale diagnostic models to validate the diagnostic performance of CTR, as well
Conventional metrics such as MSCD and SCV have traditionally been considered useful tools for distinguishing tumor invasiveness. However, when applied to the assessment of pulmonary nodules, MSCD and SCV often overlook irregular borders or protrusions that could indicate invasive features. As a result, these approaches may introduce observer variabilities and compromise the accuracy of invasiveness assessments. On the other hand, the performance of radiological features is also strongly associated with the accuracy of segmentation. Due to the lack of reliable automatic segmentation tools, manual input remains crucial in the automatic segmentation process [23]. In our study, two expert doctors played important roles in the extraction of image features. We aimed to achieve satisfactory reproducibility in the extraction of CTR-related and radiomic features between these two cohorts.

Numerous studies have demonstrated the association between CTR within part-solid nodules and increased risk in patients with early lung cancer. Survival analysis using the defined cutoff values of 0.83 for 2D-CTR and 0.34 for the 3D-CTR model in the primary cohort revealed that the high-risk group had similar rates of RFS and OS for both 3D-CTR and 2D-CTR. These results indicate that both 3D-CTR and 2D-CTR are useful for the predicting of invasiveness and have potential clinical significance. It is important to note that the diagnosis of intraoperative frozen sections and pathological sections relies on two-dimensional diagnostic approaches that focus on the maximum cross-section of the tumor to identify normal and atypical cells [24]. Considering this context, the significant outcomes achieved by both 2D-CTR and 3D-
CTR in distinguishing invasiveness reinforce the meaningfulness of our findings.

Based on our understanding, the contradictory findings regarding the use of 2D and 3D radiomic features in other types of cancer cannot be readily generalized to lung adenocarcinoma [25, 26]. Therefore, we conducted Rad-score models in 2D and 3D modalities. Subsequently, we compared the performance of these two models in different cohorts and found that the 3D radiomic model exhibited similar predictive capabilities as the 2D radiomic model for invasiveness in patients with early lung cancer.

We observed a similar number of radiomic features in both 2D and 3D modalities, with a considerable overlap between the feature families. Through a Pearson test, we determined that the majority of these features exhibited strong correlations between the 2D and 3D signatures (see Figure 6, Supplementary Table S7 and Supplementary Figure S6). In other words, texture-based features and texture filter image conversion and shape features play an important role in constructing radiomic features, which provide valuable directional information. These findings align with similar studies [27-29]. For instance, Huiping Zhao et al demonstrated comparable performance between 2D and 3D radiomic nomogram (0.919 vs. 0.945, respectively; \( p = 0.41 \)) in gastric cancer, suggesting the clinical applicability of the 2D radiomic model [30].

Our study is subject to several limitations that should be acknowledged. Firstly, being a retrospective study, certain biases such as sampling biases and performance errors in CT image interpretation may have influenced our predictive models. Secondly, our findings may be specific to patients with infected pulmonary nodules or those with
significant pulmonary diseases such as fibrosis, and may not be readily generalized to these populations. Thirdly, the number of patients with lymphatic vessel invasion or vascular invasion was limited in our study. Therefore, additional analysis is required to investigate the relationship between CTR and lymphatic vessel invasion or vascular invasion, which could potentially reveal more interesting findings. These limitations highlight the need for further improvement in device performance and the development of more precise prediction models.

Conclusion

In conclusion, our study demonstrates that both 2D-CTR and 3D-CTR are highly accurate in distinguishing radiologically less-invasive lung adenocarcinoma. The same result was also observed in 2D and 3D radiomic models. Therefore, we conclude that 3D measurement does not offer any additional advantage in clinical practice for this particular application.

Acknowledgment:

This study thanks iDataMan for data collection and data management. iDataMan is a Guangdong Medical Association (Evidence-based medicine branch) and Qitai Health joint Project on improving Data Management (https://www.qitaijk.cn/index.php/cms/show-710.html).

Funding:
This work was supported by

Guangdong Basic and Applied Basic Research Foundation (No.2019B1515130002),

The National Science Foundation of China (Grant No. 81872510), High-level Hospital Construction Project (No. DFJH201801).

Conflict of interest:

None declared.

Data Availability Statement:

All relevant data are available in the manuscript.

Author contribution statement:


Figures legends

Central image: The performance of 2D and 3D measurement in predicting radiological invasiveness. 2D/3D: two-/three-dimensional.

Figure 1. Receiver operating characteristic curves (ROC) were performed in each
dataset. A. A ROC plot of the whole primary cohort. B. ROC curve for internal training cohort. C. A plot of internal validation cohort. D. A plot of independent cohort. Figure 2. The relation between predicted invasiveness and CTR values in primary cohort. Solid lines are OR, with light color area showing 95% confidence intervals derived from restricted cubic spline regressions with four knots. A. A plot of 2D-CTR. B. A plot of 3D-CTR. OR: odd ratios. Figure 3: Survival analysis between the low-risk and high-risk groups with 2D-CTR and 3D-CTR method in the primary cohort. A. a plot of RFS between low-risk and high-risk group with 2D-CTR. B. a K-M plot of OS in two groups with 2D-CTR. C. a plot of RFS between low-risk and high-risk group with 3D-CTR. D. a K-M plot of OS in two groups with 3D-CTR. CI: confidence interval; RFS: Relapse-free survival; OS: overall survival. Figure 4: The performance of Rad-score models in predicting the radiological invasiveness. A. A ROC curve in the internal training set. B. A ROC plot in the internal testing cohort. C. A plot of ROC curve independent cohort. ROC: receiver operating characteristic curve. Figure 5: Calibration curves of the Rad-score models. A. Calibration curve of the Rad-score models in the training cohort. B. Calibration curve of the Rad-score models in the testing cohort. C. Calibration curve of the Rad-score models in the independent validation cohort. Figure 6. Correlations between selected radiomic features for 2D and 3D using Pearson Test. The red boxes represent strong positive correlation and the blue boxes represent strong negative correlation between 2D and 3D features.
**Tables**

Table 1. Preoperatively Accuracy for Distinguishing Less-invasive from Invasive Adenocarcinoma with Radiological (CTR) Models in Different Cohorts.

<table>
<thead>
<tr>
<th>Cohorts</th>
<th>Parameter</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>AUC (95 CI%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Whole Primary Cohort</td>
<td>2D-CTR</td>
<td>64.10</td>
<td>86.20</td>
<td>0.828 (0.798-0.858)</td>
</tr>
<tr>
<td></td>
<td>3D-CTR</td>
<td>70.70</td>
<td>81.90</td>
<td>0.830 (0.800-0.859)</td>
</tr>
<tr>
<td>Internal Training Cohort</td>
<td>2D-CTR</td>
<td>73.60</td>
<td>78.80</td>
<td>0.828 (0.795-0.861)</td>
</tr>
<tr>
<td></td>
<td>3D-CTR</td>
<td>60.10</td>
<td>90.70</td>
<td>0.827 (0.794-0.861)</td>
</tr>
<tr>
<td>Internal Testing Cohort</td>
<td>2D-CTR</td>
<td>64.30</td>
<td>86.50</td>
<td>0.827 (0.756-0.898)</td>
</tr>
<tr>
<td></td>
<td>3D-CTR</td>
<td>65.90</td>
<td>91.90</td>
<td>0.839 (0.776-0.902)</td>
</tr>
<tr>
<td>External Validation Cohort</td>
<td>2D-CTR</td>
<td>82.00</td>
<td>73.80</td>
<td>0.807 (0.768-0.845)</td>
</tr>
<tr>
<td></td>
<td>3D-CTR</td>
<td>93.40</td>
<td>57.00</td>
<td>0.826 (0.790-0.863)</td>
</tr>
</tbody>
</table>

Abbreviations: CTR, consolidation tumor ratio; AUC, an area under the curve; CI, confidence interval.

Table 2. Preoperatively Accuracy for Distinguishing Less-invasive from Invasive Adenocarcinoma with Radiomic feature Models in Different Cohorts.

<table>
<thead>
<tr>
<th>Cohorts</th>
<th>Parameter</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>AUC (95 CI%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal Training Cohort</td>
<td>2D Rad-score</td>
<td>82.9</td>
<td>90.0</td>
<td>0.923 (0.901-0.945)</td>
</tr>
<tr>
<td></td>
<td>3D Rad-score</td>
<td>84.7</td>
<td>88.0</td>
<td>0.926 (0.906-0.947)</td>
</tr>
<tr>
<td>Internal Testing Cohort</td>
<td>2D Rad-score</td>
<td>73.0</td>
<td>86.8</td>
<td>0.888 (0.834-0.941)</td>
</tr>
<tr>
<td></td>
<td>3D Rad-score</td>
<td>67.5</td>
<td>94.7</td>
<td>0.887 (0.834-0.940)</td>
</tr>
<tr>
<td>External Validation Cohort</td>
<td>2D Rad-score</td>
<td>79.4</td>
<td>83.8</td>
<td>0.890 (0.866-0.932)</td>
</tr>
<tr>
<td></td>
<td>3D Rad-score</td>
<td>81.7</td>
<td>81.1</td>
<td>0.902 (0.876-0.940)</td>
</tr>
</tbody>
</table>

Abbreviations: Rad-score, radiomic score; AUC, an area under the curve; CI, confidence interval.
Reference


Figure 2

A

OR (95% CI)

2D-CTR = 0.83

OR = 1

B

OR (95% CI)

3D-CTR = 0.34

OR = 1

Click here to access/download;Figure;Figure 2(R2 version).pdf

ACCEPTED MANUSCRIPT
Figure 3

A. Relapse-free Survival

Log-rank

$p < 0.000$

HR = 5.81 (95% CI: 2.44-13.8)

Number at risk

Low-risk: 408, 391, 186, 69, 20, 0
High-risk: 410, 379, 192, 74, 34, 2

Time (months)

B. Overall Survival

Log-rank

$p = 0.003$

HR = 3.89 (95% CI: 1.46-10.37)

Number at risk

Low-risk: 408, 390, 189, 70, 20, 0
High-risk: 410, 386, 201, 79, 36, 2

Time (months)

C. Relapse-free Survival

Log-rank

$p < 0.000$

HR = 6.99 (95% CI: 2.74-17.81)

Number at risk

Low-risk: 414, 398, 180, 60, 17, 1
High-risk: 404, 372, 198, 83, 37, 1

Time (months)

D. Overall Survival

Log-rank

$p = 0.001$

HR = 4.96 (95% CI: 1.70-14.46)

Number at risk

Low-risk: 414, 397, 183, 61, 17, 1
High-risk: 404, 379, 207, 88, 39, 1

Time (months)
Figure 4

(A) AUC: 0.926
AUC: 0.923

B

AUC: 0.887
AUC: 0.888

C

AUC: 0.902
AUC: 0.890

p = 0.38
p = 0.91
p = 0.189

3D Rad-score model
2D Rad-score model
Figure 5

A

B

C

Observed Probability

Predicted Probability

2D Rad-score model

3D Rad-score model

Observed Probability

Predicted Probability

2D Rad-score model

3D Rad-score model

Observed Probability

Predicted Probability

2D Rad-score model

3D Rad-score model
Figure 6

![Heatmap showing correlation values for various 2D and 3D shape and texture features]

- 3D\_gllum\_RunEntropy
- 3D\_gldm\_DependenceEntropy
- 3D\_shape\_SurfaceVolumeRatio
- 3D\_shape\_Maximum2DDiameterSlice
- 3D\_shape\_Maximum2DDiameterRow
- 3D\_shape\_LeastAxisLength
- 3D\_firstorder\_Median
- 2D\_shape2D\_MajorAxisLength
- 2D\_shape2D\_PerimeterSurfaceRatio
- 2D\_gllum\_Idn
- 2D\_gllum\_JointAverage
- 2D\_gllum\_RunEntropy
- 2D\_glszm\_ZoneEntropy

Value range: -0.5 to 0.5
Key question

Are the performance of measurement in 3D modality superior to that of 2D for predicting invasiveness in lung cancer?

Key findings

Both 2D and 3D measurements have good performance in predicting invasiveness for patients with early stage lung cancer.

Take-home message

3D measurement does not offer any additional advantage in clinical practice for this particular application.