Original Study

Radiologists Versus AI-Based Software: Predicting Lymph Node Metastasis and Prognosis in Lung Adenocarcinoma From CT Under Various Image Display Conditions

Junya Sato,[1,2] Masahiro Yanagawa,[1] Daiki Nishigaki,[1] Akinori Hata,[1] Yukinori Sakao,[3] Noriaki Sakakura,[4] Yasushi Yatabe,[5] Yasushi Shintani,[6] Shoji Kido,[7] Noriyuki Tomiyama[1]

Abstract

Accurate measurement of lung adenocarcinoma is crucial for determining treatment plan and predicting prognosis. However, interobserver variability and display conditions can affect these measurements. We compared tumor size measurements between radiologists and commercially available AI-based software using preoperative CT images from 307 cases and evaluated their predictive value for lymph node metastasis, disease-free survival, and overall survival. We concluded that while radiologist measurements showed inter-observer variability, AI-based software provided accurate and reproducible prognostic indicators.

Purpose: To compare the variability of quantitative values from lung adenocarcinoma CT images independently assessed by 2 radiologists and AI-based software under different display conditions, and to identify predictors of pathological lymph node metastasis (LNM), disease-free survival (DFS), and overall survival (OS). **Methods:** Preoperative CT images of 307 patients were displayed under 4 conditions: lung-1, lung-2, mediastinum-1, and mediastinum-2. Two radiologists (R1, R2) measured total diameter (tD) and the longest solid diameter (sD) under each condition. The AI-based software automatically detected lung nodules, providing tD, sD, total volume (tV), and solid volume (sV). **Results:** All measurements by R1 and R2 with AI-based software were identical. Four out of the 8 measurements showed significant variation between R1 and R2. For LNM, multivariate logistic regression identified significant indicators including sD at mediastinum-2 of R1, sD at mediastinum-1 and mediastinum-2 of R2, tV, and the proportion of sV to tV (sV/tV) of AI-based software. For DFS, multivariate Cox regression identified sD at lung-1 of R1, the proportions of sD to tD at lung-2 of R1, sD at lung-2 and mediastinum-1 of R2, tV, and sV/tV of AI-based software as significant. For OS, multivariate Cox regression identified sD at lung-1 and mediastinum-2 of R1, tD at lung-2 of R2, sD at mediastinum-1 of R2, sV, and sV/tV of AI-based software as significant. **Conclusion:** Radiologists' CT measurements were significant predictors of LNM and prognosis, but variability existed among radiologists and display conditions. AI-based software can provide accurate and reproducible indicators for predicting LNM and prognosis.

Clinical Lung Cancer, Vol. 26, No. 1, 58–71 © 2024 The Author(s). Published by Elsevier Inc.

Keywords: Artificial intelligence, Lung adenocarcinoma, Lymph node metastasis, Prognosis, Tumor measurement

Abbreviations: AI, artificial intelligence; CAD, computer-aided diagnosis; CT, computed tomography; GGN, ground-glass nodule; LNM, lymph node metastasis; DFS, disease-free survival; OS, overall survival; ROC, receiver operating characteristic; sD, solid diameter; sV, solid volume; tD, total diameter; tV, total volume; aOR, adjusted odds ratio; aHR, adjusted hazard ratio.

[1]Department of Radiology, Osaka University Graduate School of Medicine, Osaka, Japan
[2]Department of Artificial Intelligence in Diagnostic Radiology, Osaka University Graduate School of Medicine, Osaka, Japan
[3]Department of Surgery, Teikyo University School of Medicine, Tokyo, Japan
[4]Department of Thoracic Surgery, Aichi Cancer Center, Nagoya, Japan
[5]Department of Diagnostic Pathology, National Cancer Center Hospital, Tokyo, Japan
[6]Department of General Thoracic Surgery, Osaka University Graduate School of Medicine, Osaka, Japan
[7]Osaka University Institute for Radiation Science, Osaka University Graduate School of Medicine, Osaka, Japan

Submitted: Aug 24, 2024; Revised: Oct 28, 2024; Accepted: Oct 30, 2024; Epub: 7 November 2024

Address for correspondence: Masahiro Yanagawa, MD, PhD, Department of Radiology, Osaka University Graduate School of Medicine, 2-2, Yamadaoka, Suita, Osaka 565-0871, Japan
E-mail contact: m-yanagawa@radiol.med.osaka-u.ac.jp

https://doi.org/10.1016/j.cllc.2024.10.015

Introduction

According to the eighth edition of the TNM classification, the size of a lung tumor is defined as the size of the pathological invasiveness.[1] The size of the solid component of the tumor on CT images correlates closely with the size of the pathological invasiveness.[2] The proportion of the solid component also reflects prognosis.[3,4] Therefore, accurate measurement of the solid component and overall size of a lung tumor on CT images is necessary for staging and determining the appropriate treatment plan for the patient.

To ensure consistency across various clinical settings, measurement protocols should be standardized. The suggested conditions for accurate measurement include: (1) using thin CT slices of 1 mm or less, (2) employing lung-window settings, (3) measuring in centimeters to the first decimal place, and (4) measuring in the transverse view.[1] However, a consensus on measuring conditions is lacking in clinical practice. Previous studies have reported interobserver variability[5,6] as well as differences in the measurement of solid component size depending on image display conditions.[7] These variations can negatively impact surveillance recommendations and prognostic determinations.

Recent advances in artificial intelligence (AI) have enabled accurate measurement of tumor size and volume. Wang et al.[7] reported that the concordance rate between the segmentation results obtained with an AI-based model and by radiologists was comparable to the inter-radiologist concordance rate, with a difference of less than 2%. Ahn et al.[8] showed that the measurements of tumor size by an AI model had good agreement with the size of pathological invasiveness. AI assessment is thus reproducible and objective, making it potentially effective for accurate prognostic evaluation, particularly in preoperative imaging. Furthermore, AI-based measurements quickly provide diameter and volume data without burdening the clinical workflow, thus offering comprehensive information on lung adenocarcinoma.

We hypothesized that, even when following the guidelines, differences related to measurers and display conditions could lead to inconsistencies that affect diagnosis and treatment decisions, while AI-based measurements could provide more consistent results across settings. The aims of this study were thus as follows: (1) to compare variabilities in measurements of tumor diameter made independently by 2 independent radiologists and the AI-based software on CT images under various image display conditions; and (2) to investigate the impact of these measurements on the prediction of lymph node metastasis (LNM), disease-free survival (DFS), and overall survival (OS) in lung adenocarcinoma patients.

Methods

Study Design and Patients

This retrospective study was approved by all institutions' internal ethics review boards and was conducted following the Declaration of Helsinki. Informed consent was waived for the retrospective review of patient records and images. The original dataset consisted of 484 patients who underwent CT prior to surgery at Aichi Cancer Center Hospital from January 2014 to December 2017. Of these, 449 patients were considered for inclusion. The inclusion criteria were as follows: (1) CT examination prior to surgery, (2) no previous treatment for malignant tumors in the lungs or any other organ, (3) CT images with slice thickness of 2 mm or less, (4) Age 20 years or older. Of the 444 patients, 327 patients were histopathologically diagnosed with adenocarcinoma. Of the remaining 327 patients, 20 patients were excluded as their tumors were unmeasurable by AI-based software. Finally, 307 patients were included in this study. Patient selection is summarized in Figure 1.

CT Protocols

CT protocols vary because different multidetector row CT systems are involved: the scan mode was helical; tube voltage, 120 kVp; tube current, auto exposure control; and the scan field of view (FOV), 350-400 mm. Target reconstruction of tumors was performed with a 200-mm FOV in 253 cases.

Histopathological Evaluation

All histopathological specimens were stained with hematoxylin–eosin and/or Elastica van Gieson and evaluated by 1 pathologist according to the multidisciplinary adenocarcinoma criteria.[9] The final histological diagnosis (adenocarcinoma in situ, minimally invasive adenocarcinoma, or invasive adenocarcinoma) and pathologic T and N descriptors, according to the eighth TNM classification, were recorded.[1]

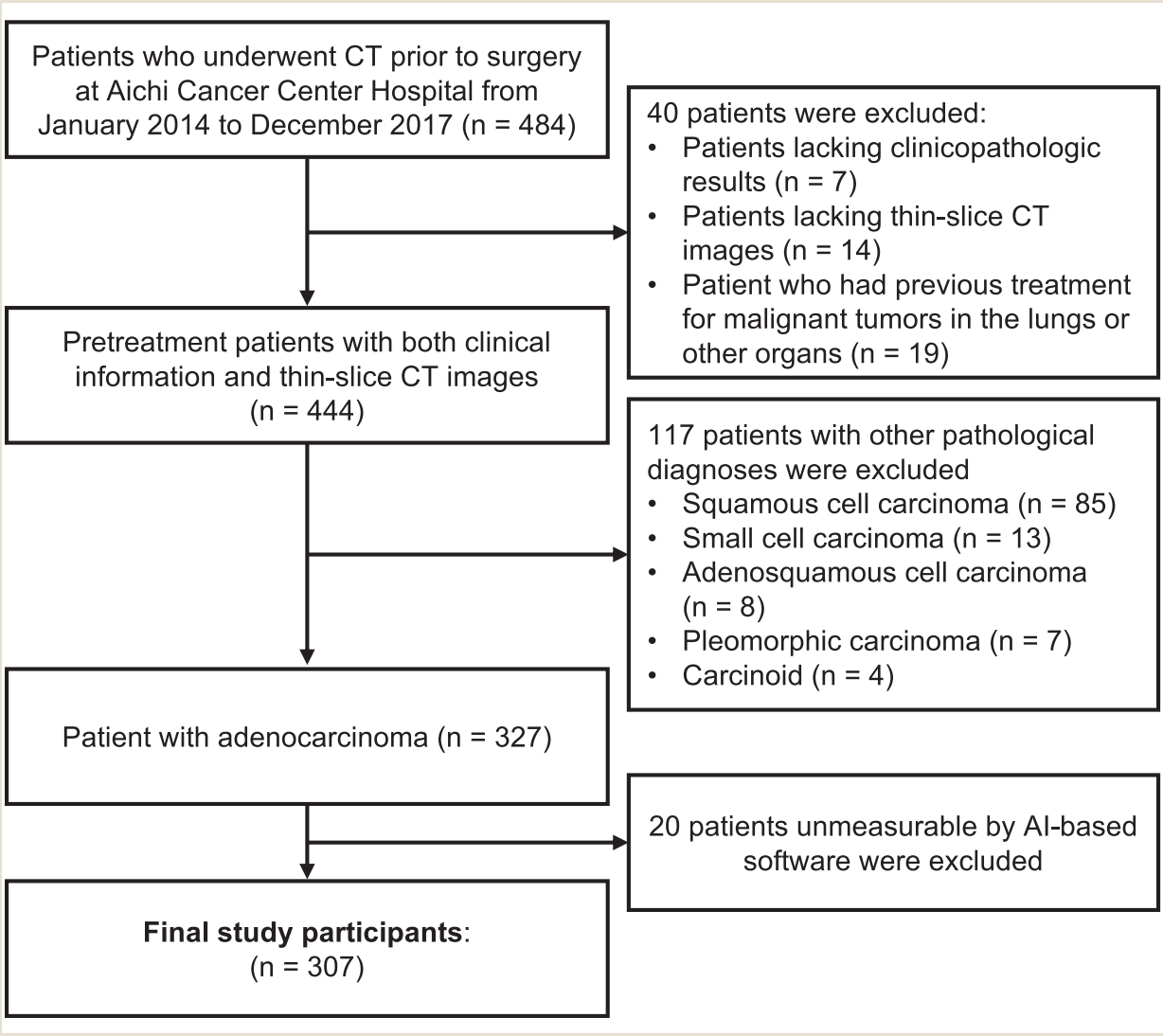
Radiologists' Measurements

All CT images were evaluated independently by 2 junior radiology residents (J.S. and D.N.) without knowledge of the patients' diagnoses and outcomes. These images were displayed using the following 4 conditions (window-width, window-level): lung-1 (1200 HU, -700 HU), lung-2 (1500 HU, -600 HU), mediastinum-1 (350 HU, 40 HU), and mediastinum-2 (1250 HU, 40 HU).[10] Representative examples of the 4 conditions are shown in Figure 2. The radiologists measured total diameter (tD) and the longest solid diameter (sD) in lung-1 and lung-2 conditions, and sD on mediastinum-1 and mediastinum-2 conditions in the transverse view. From these, the following eight features were finally generated: tD at lung-1, sD at lung-1, the proportion of the solid diameter (sD/tD) at lung-1, tD at lung-2, sD at lung-2, sD/tD at lung-2, sD at mediastinum-1, and sD at mediastinum-2.

AI-Based Software Measurements

After the evaluation by the radiologists alone, a similar evaluation was performed using AI-based software. Each radiologist used the commercially available AI-based computer aided detection/diagnosis (CAD) system integrated with SYNAPSE SAI Viewer V2.4 (FUJIFILM Corporation, Minato, Japan). This deep learning-based system identifies candidate tumor regions (Figure 3A) and automatically segments the total tumor and solid components (Figure 3B). The AI-based CAD system measured the diameter, area, and volume of the total tumor and of the solid components. Consequently, the following 6 features were generated: tD, sD, sD/tD, total volume (tV), solid volume (sV), and the

Figure 1 Flowchart of patient selection.
Abbreviations: CT = computed tomography.

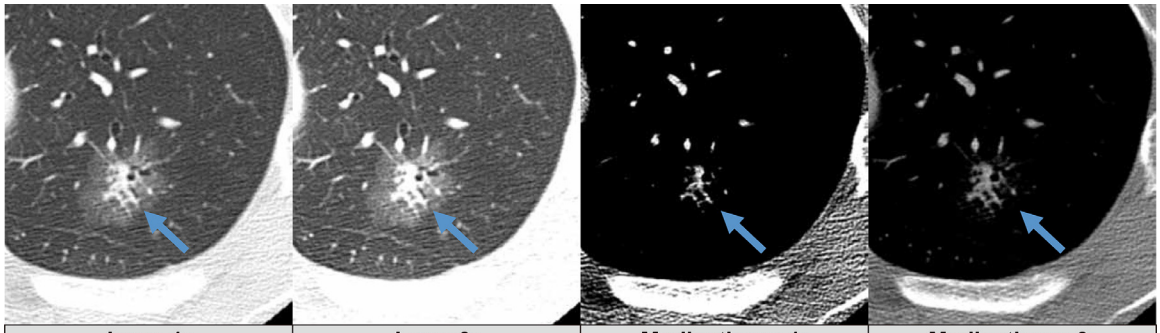


proportion of the solid volume (sV/tV). Automatic measurement with the AI-based CAD system did not depend on image display conditions. Radiologists were not permitted to change the results of their own measurements after using the AI-based CAD system.

Statistical Analysis

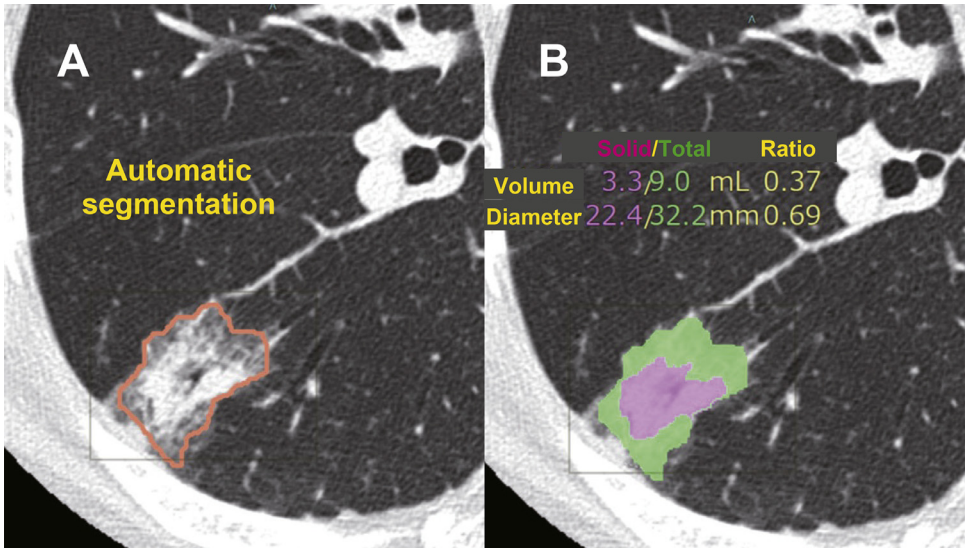
Interobserver agreement of each quantitative value was assessed by 2 radiologists (J.S. and D.N.) using the Bland–Altman method.[11] CT features in the radiologists and AI-based CAD system measurement were evaluated to examine associations with LNM, DFS, and OS. For each CT feature, the optimal cutoff values that distinguished best between patients with and without LNM and between patients who survived and died, were determined using the empirical receiver operating characteristic (ROC) method. All ROC analyses were univariate. Each variable's optimal cutoff value was identified separately. Subsequently, associations between prognostic factors and binary features, as defined by the cutoff values, were evaluated using univariate regression analyses with and without adjustment for age, sex, and the Brinkman index. Statistically significant parameters from the univariate analysis, adjusted for age, sex, and the Brinkman index were included in multivariable logistic regression and Cox proportional hazards regression analyses to identify predictors of LNM and prognosis, respectively (stepwise method; $P \leq .05$ or less was used for entry into the model and $P > .1$ was selected for removal). DFS and OS curves were generated by the Kaplan–Meier method, with comparisons performed by using the log-rank test. These

Figure 2 Four image display conditions in a 66-year-old male patient with the pathological diagnosis of invasive adenocarcinoma with an invasive focus diameter of 1.5 cm. From left to right: lung-1, lung-2, mediastinum-1, and mediastinum-2.



Lung-1:	Lung-2:	Mediastinum-1:	Mediastinum-2:
WW = 1200 HU,	WW = 1500 HU,	WW = 350 HU,	WW = 1250 HU,
WL = -700HU	WL = -600HU	WL = 40HU	WL = 40HU

Figure 3 Automatic segmentation by AI-based software in a 76-year-old man with histopathologically confirmed invasive adenocarcinoma with a pathological invasiveness size of 2.2 cm. The software automatically (A) identifies the tumor region, and (B) measures its diameter and volume, including the total tumor and solid components. Radiologist measurements: lung-1 (solid/total): R1 = 21.2/28.5 cm, R2 = 24.2/39.4 cm; lung-2: R1 = 21.4/29.8 cm, R2 = 21.7/39.2 cm; mediastinum-1: R1 = 13.7/13.7 cm, R2 = 21.0/21.0 cm; mediastinum-2: R1 = 13.2/13.2 cm, R2 = 20.8/20.8 cm. AI-based software measurements: 22.4/32.2 cm.



analyses were conducted separately for features identified by each radiologist and by the AI-based CAD system. All statistical analyses were performed using commercially available software (MedCalc Software; version 22.023-64 bit, Ostend, Belgium). All *P*-values were 2-sided, and *P*-values < .05 were considered significant.

Results

Clinical and Pathologic Characteristics

The final study population included 307 patients (mean age, 65.9 years; range 29-87 years) as shown in Table 1. Among them, 153 were men (mean age 66.7 years; range 29-84 years) and 154 were women (mean age 65.0; range 32-87 years). The 307 tumors

Table 1 Patient Characteristics

Characteristic			Value
Age			65.9 ± 10 (29-87)
Male sex			153 (50)
Lesion type at CT			
	pure GGN		36 (11.7)
	part-solid GGN		107 (34.9)
	solid		164 (53.4)
Histologic subtype			
	Adenocarcinoma in situ		13
	Minimally invasive adenocarcinoma		38
	Adenocarcinoma		
		Lepidic	27
		Acinar	77
		Papillary	95
		Solid	41
		Mucinous	7
		solid with mucin	7
		micropapilally	2
Pathological Invasiveness size (cm)			1.9 (0-8.5)
Interval from preoperative CT to surgery (d)			23 (1-159)
Pathologic T descriptor			
	Tis		13
	Tmi		38
	T1a		33
	T1b		80
	T1c		46
	T2a		69
	T2b		8
	T3		18
	T4		2

Data are n, *n* (%), means ± standard deviations, with range in parentheses.

consisted of 36 ground-glass nodules (GGNs, 11.7%), 107 part-solid GGNs (34.9%), and 164 solid nodules (53.4%). The median time between preoperative CT and surgery was 23 days (range, 1-159 days). At pathological evaluation, the mean invasiveness sizes of the GGNs, part-solid GGNs, and solid nodules were 0.60 cm ± 0.84 (range, 0-3.7 cm), 1.45 cm ± 1.00 (range, 0-4.5 cm), and 2.48 mm ± 1.41 (range, 0-8.5 cm), respectively. Pathological T descriptors included 13 pTis, 38 pTmi, 33 pT1a, 80 pT1b, 46 pT1c, 69 pT2a, 8 pT2b, 18 pT3, and 2 pT4 cases.

Interobserver Agreement for Each CT Feature

Nodule diameter measurements of the 2 radiologists is summarized in Table 2. Among the eight measurements made without the AI-based CAD system, 4 measurements (tD at lung-1, tD at lung-2, sD at lung-2, sD at mediastinum-2) showed significant variation between R1 and R2 on Bland–Altman plots (Fig. 4) ($P < .001$, $< .001$, $< .001$, and $= .002$, respectively). In contrast, all measurements generated by the AI-based CAD system were identical.

CT Features' Influence on LNM

Table 3 summarizes the relationship of LNM with CT measurements by radiologists (R1 and R2) and the AI-based CAD system. Univariable logistic regression analyses with and without adjustment for age, sex, and the Brinkman index revealed that all features were of significant use for predicting LNM. Multivariate logistic regression analysis adjusted for age, sex, and the Brinkman index revealed the following significant indicators of LNM: R1 sD at mediastinum-2 (adjusted odds ratio [aOR] 4.32; 95% confidence interval [CI]: 1.22, 15.3; $P = .024$), R2 sD at mediastinum-1 (aOR 3.27; 95% CI: 1.20, 8.93; $P = .021$) and sD at mediastinum-2 (aOR 4.51; 95% CI: 1.41, 14.4; $P = .011$), and AI tV (aOR 4.82; 95% CI: 2.41, 9.63; $P < .001$) and sV/tV (aOR 7.59; 95% CI: 2.82, 20.4; $P < .001$). sD at mediastinum-1 of R1 (aOR 2.68; 95% CI: 0.87, 8.29; $P = .086$) was included in the model, but no statistically significant association was found.

CT Features' Influence on DFS

Table 4 summarizes the relationship of DFS with CT measurements by radiologists (R1 and R2) and the AI-based CAD system.

Table 2 Tumor Measurement in Radiologists and AI-Based Software

Radiologist measurement conditions	tD at lung-1 (cm)	sD at lung-1 (cm)	sD/tD at lung-1	tD at lung-2 (cm)	sD at lung-2 (cm)	sD/tD at lung-2	sD at mediastinum-1 (cm)	sD at mediastinum-2 (cm)
R1	2.53±1.25	2.03±2.03	0.75±0.57	2.50±1.26	1.94±1.47	0.74±0.39	1.70±1.44	1.90±1.45
R2	2.67±1.32	2.12±1.44	0.76±0.30	2.61±1.31	2.05±1.44	0.75±0.31	1.74±1.46	1.98±1.43
AI-based software measurement								
conditions	tD (cm)	sD (cm)	sD/tD	tV (cm3)	sV (cm3)	sV/tV		
R1-AI	2.49±1.18	2.07±1.36	0.78±0.29	7.67±12.8	6.25±12.4	0.61±0.34		
R2-AI	2.49±1.18	2.07±1.36	0.78±0.29	7.67±12.8	6.25±12.4	0.61±0.34		

Data are means ± standard deviations.

Figure 4 Bland–Altman plots comparing 2 radiologists' measurements, with X-axes showing the mean of their measurements and Y-axes showing the differences between them. Shown are mean bias ± 1.96 standard deviation (SD), 95% confidence intervals, and limits of agreement. (A) Total diameter (tD) at lung-1 measured is −1.4 cm ± 10.1, −2.0 to -0.84 cm, and −11.5 to 8.7 cm, (B) tD at lung-2 measured is −1.1 cm ± 10.2, −1.6 to 0.51 cm, and −11.3 to 9.1 cm, (C) solid diameter (sD) at lung-2 measured is −1.1 cm ± 10.2, −1.6 to -0.53 cm, and −10.8 to 8.7 cm, and (D) sD at mediastinum-2 measured is −0.81 cm ± 9.0, −1.3 to -0.29 cm, and −9.8 to 8.2 cm.

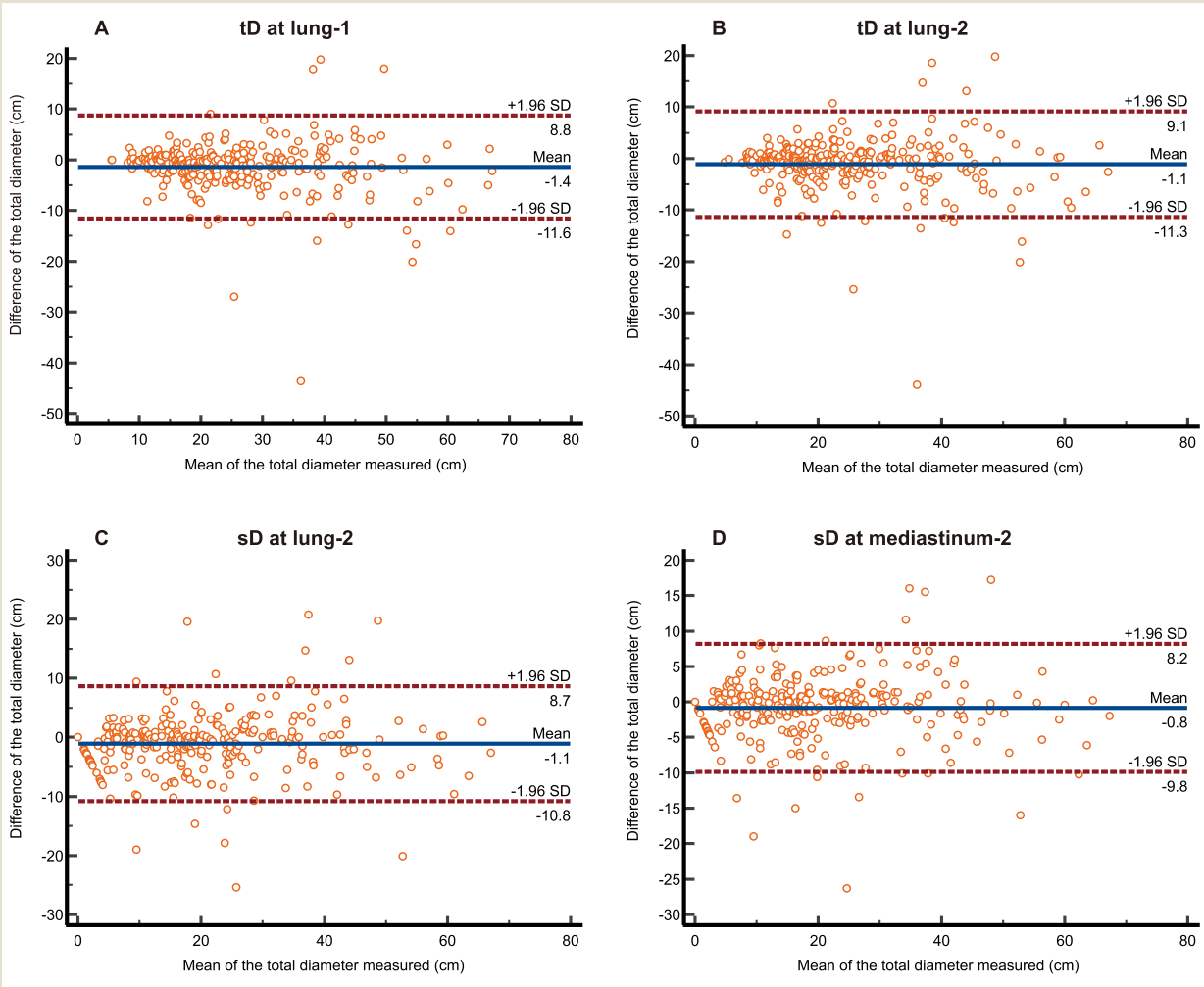


Table 3 Relationship of CT Findings in Radiologists and AI-Based Software With Lymph Node Metasitasis

CT Feature	Univariable Analysis		Univariable Analysis Adjusted for Age, Sex, and Brinkman Index		Multivariable Analysis Adjusted for Age, Sex, and Brinkman Index	
R1	**Odds Ratio**	***P* Value**	**Odds Ratio**	***P* Value**	**Odds Ratio**	***P* Value**
tD at lung-1						
Negative (≤ 2.23, *n* = 154)	5.46 (2.61, 11.4)	< .0001	7.04 (3.18, 15.6)	< .0001		
positive (> 2.23, *n* = 153)						
sD at lung-1						
negative (≤ 1.96, *n* = 181)	9.26 (4.40, 19.5)	< .0001	10.4 (4.81, 22.4)	< .0001		
positive (> 1.96, *n* = 126)						
sD/tD at lung-1						
negative (≤ 0.81, *n* = 126)	5.76 (2.49, 13.3)	< .0001	6.00 (2.56, 14.1)	< .0001		
positive (> 0.81, *n* = 181)						
tD at lung-2						
negative (≤ 2.22, *n* = 156)	5.67 (2.71, 11.9)	< .0001	7.30 (3.30, 16.1)	< .0001		
positive (> 2.22, *n* = 151)						
sD at lung-2						
negative (≤ 2.06, *n* = 189)	8.25 (4.08, 16.7)	< .0001	9.17 (4.41, 19.0)	< .0001		
positive (> 2.06, *n* = 118)						
sD/tD at lung-2						
negative (≤ 0.60, *n* = 101)	10.2 (3.08, 33.7)	.0001	10.4 (3.11, 34.6)	.0001		
positive (> 0.60, *n* = 206)						
sD at mediastinum-1						
negative (≤ 1.77, *n* = 190)	8.44 (4.17, 17.1)	< .0001	9.19 (4.45, 18.9)	< .0001	2.68 (0.87, 8.29)	.0858
positive (> 1.77, *n* = 117)						
sD at mediastinum-2						
negative (≤ 1.74, *n* = 170)	110.0 (4.49, 22.3)	< .0001	10.4 (4.65, 23.4)	< .0001	4.32 (1.22, 15.3)	.0235
positive (> 1.74, *n* = 137)						
R2	Odds ratio	*P* Value	Odds ratio	*P* Value	Odds ratio	*P* Value
tD at lung-1						
negative (≤ 2.42, *n* = 160)	5.90 (2.82, 12.3)	< .0001	7.22 (3.29, 15.9)	< .0001		
positive (> 2.42, *n* = 147)						
sD at lung-1						
negative (≤ 1.98, *n* = 165)	9.03 (4.06, 20.1)	< .0001	9.80 (4.32, 22.2)	< .0001		
positive (> 1.98, *n* = 142)						
sD/tD at lung-1						
negative (≤ 0.86, *n* = 150)	5.15 (2.46, 10.8)	< .0001	5.27 (2.48, 11.2)	< .0001		
positive (> 0.86, *n* = 157)						
tD at lung-2						
negative (≤ 2.40, *n* = 163)	5.45 (2.67, 11.2)	< .0001	6.51 (3.05, 13.9)	< .0001		
positive (> 2.40, *n* = 144)						
sD at lung-2						
negative (≤ 1.93, *n* = 168)	8.35 (3.87, 18.0)	< .0001	8.82 (4.04, 19.3)	< .0001		
positive (> 1.93, *n* = 139)						
sD/tD at lung-2						
negative (≤ 0.92, *n* = 153)	4.76 (2.33, 9.72)	< .0001	4.93 (2.36, 10.3)	< .0001		
positive (> 0.92, *n* = 154)						
sD at mediastinum-1						
negative (≤ 2.05, *n* = 206)	8.56 (4.33, 16.9)	< .0001	9.57 (4.69, 19.5)	< .0001	3.27 (1.20, 8.93)	.021
positive (> 2.05, *n* = 101)						
sD at mediastinum-2						

(*continued on next page*)

Table 3 (*continued*)

CT Feature	Univariable Analysis		Univariable Analysis Adjusted for Age, Sex, and Brinkman Index		Multivariable Analysis Adjusted for Age, Sex, and Brinkman Index	
R1	**Odds Ratio**	***P* Value**	**Odds Ratio**	***P* Value**	**Odds Ratio**	***P* Value**
negative (≤ 1.90, *n* = 173)	10.2 (4.58, 22.7)	< .0001	11.1 (4.89, 25.4)	< .0001	4.51 (1.41, 14.4)	.0109
positive (> 1.90, *n* = 134)						
AI-based software	Odds ratio	*P* Value	Odds ratio	*P* Value	Odds ratio	*P* Value
tD						
negative (≤ 2.70, *n* = 196)	6.28 (3.23, 12.2)	< .0001	7.75 (3.81, 15.8)	< .0001		
positive (> 2.70, *n* = 111)						
sD						
negative (≤ 2.21, *n* = 116)	7.44 (3.73, 14.8)	< .0001	8.99 (4.32, 18.7)	< .0001		
positive (> 2.21, *n* = 191)						
sD/tD						
negative (≤ 0.91, *n* = 153)	6.83 (3.18, 14.7)	< .0001	6.71 (3.10, 14.5)	< .0001		
positive (> 0.91, *n* = 154)						
tV						
negative (≤ 5.83, *n* = 213)	7.45 (3.86, 14.4)	< .0001	9.07 (4.49, 18.3)	< .0001	4.82 (2.41, 9.63)	< .0001
positive (> 5.83, *n* = 94)						
sV						
negative (≤ 2.35, *n* = 181)	10.1 (4.66, 21.8)	< .0001	11.6 (5.20, 25.8)	< .0001		
positive (> 2.35, *n* = 126)						
sV/tV						
negative (≤ 0.68, *n* = 140)	11.7 (4.47, 30.4)	< .0001	11.8 (4.49, 31.1)	< .0001	7.59 (2.82, 20.4)	.0001
positive (> 0.68, *n* = 167)						

Numbers in parentheses are the 95% limit of agreement.

Univariable logistic regression analyses with and without adjustment for age, sex, and the Brinkman index revealed that all features were of significant use for predicting DFS. Multivariate logistic regression analysis adjusted for age, sex, and the Brinkman index revealed the following specific indicators of DFS: R1 sD at lung-1 (adjusted hazard ratio [aHR] 3.93; 95% CI: 2.30, 6.71; $P < .001$) and sD/tD in lung-2 (aHR 2.41; 95% CI: 1.35, 4.31; $P = .0031$), R2 sD at lung-2 (aHR 2.76; 95% CI: 1.53, 4.98; $P < .001$) and sD in mediastinum-1 (aHR 2.59; 95% CI: 1.54, 4.33; $P < .001$), and AI tV (aHR 3.84; 95% CI: 2.17, 6.82; $P < .001$) and sV/tV (aHR 2.68; 95% CI: 1.68, 4.26; $P < .001$). Figure 5 shows that the 5-year DFS rate for patients with tV of 2.54 cm^3 or greater was 58.0% and that of patients with sV/tV of 84% or greater was 53.2% (all $P < .001$).

CT Features' Influence on OS

Table 5 summarizes the relationship of OS with CT measurements by radiologists (R1 and R2) and the AI-based CAD system. Univariable logistic regression analyses with and without adjustment for age, sex, and the Brinkman index revealed that all features were of significant use for predicting OS. Multivariate logistic regression analysis adjusted for age, sex, and the Brinkman index revealed the following specific indicators of OS: R1 sD at lung-1 (aHR 3.79; 95% CI: 1.54, 9.33; $P = .0037$) and sD in mediastinum-2 (aHR 2.25; 95% CI: 1.07, 4.75; $P = .033$), R2 tD at lung-2 (aHR 3.71; 95% CI: 1.22, 11.3; $P = .021$) and sD in mediastinum-1 (aHR 3.30; 95% CI: 1.70, 6.40; $P < .001$), and AI sV (aHR 3.54; 95% CI: 1.78, 7.01; $P < .001$) and sV/tV (aHR 2.81; 95% CI: 1.30, 6.11; $P = .009$). Figure 6 shows that the 5-year OS rate for patients with sV of 7.03 cm^3 or greater was 69.2% and that of patients with sV/tV of 85% or greater was 77.7% (all $P < .001$).

Discussion

In this study, radiologists measured the size of lung tumors using CT images under different display conditions, with a view to predicting LNM and prognosis. In addition to manual measurements, we also used an AI-based CAD system and compared the results. Four CT features significantly varied due to different image display conditions and interobserver variability, which resulted in differences in significant predictors of DFS and OS in multivariate analysis. In contrast, the AI-based CAD system could predict LNM using tV and sV/tV, DFS using tV and sV/tV, and OS using sV and sV/tV, with complete reproducibility, regardless of image display conditions or readers. Previous studies have shown that LNM was associated with sD,[12] tV,[13] and sV/tV,[14] and that prognosis was associated with sD,[15] tD,[16] sD/tD,[17] sV,[18] tV,[19] and sV/tV,[14] consistent with our findings. Our study is important because it supports the association between CT features and outcomes found in previous research and highlights the high reproducibility of features identified by AI.

Table 4 Relationship of CT Findings in Radiologists and AI-Based Software With Disease-Free Survival

CT Feature	Univariable Analysis		Univariable Analysis Adjusted for Age, Sex, and Brinkman Index		Multivariable Analysis Adjusted for Age, Sex, and Brinkman Index	
R1	**Hazard ratio**	***P* Value**	**Hazard ratio**	***P* Value**	**Hazard ratio**	***P* Value**
tD at lung-1						
Negative (≤ 2.23, *n* = 154)	3.68 (2.27, 5.96)	< .0001	3.43 (2.11, 5.57)	< .0001		
positive (> 2.23, *n* = 153)						
sD at lung-1						
negative (≤ 1.80, *n* = 169)	5.62 (3.14, 9.26)	< .0001	5.48 (3.32, 9.02)	< .0001	3.93 (2.30, 6.71)	< .0001
positive (> 1.80, *n* = 138)						
sD/tD at lung-1						
negative (≤ 0.90, *n* = 134)	4.44 (2.58, 7.63)	< .0001	4.10 (2.36, 7.10)	< .0001		
positive (> 0.90, *n* = 173)						
tD at lung-2						
negative (≤ 2.22, *n* = 156)	3.80 (2.35, 6.16)	< .0001	3.60 (2.21, 5.85)	< .0001		
positive (> 2.22, *n* = 151)						
sD at lung-2						
negative (≤ 1.80, *n* = 171)	5.38 (3.29, 8.78)	< .0001	5.18 (3.17, 8.47)	< .0001		
positive (> 1.80, *n* = 136)						
sD/tD at lung-2						
negative (≤ 0.95, *n* = 174)	4.38 (2.55, 7.53)	< .0001	4.06 (2.34, 7.03)	< .0001	2.41 (1.35, 4.31)	.0031
positive (> 0.95, *n* = 133)						
sD at mediastinum-1						
negative (≤ 1.36, *n* = 154)	5.05 (3.01, 8.48)	< .0001	4.76 (2.83, 8.00)	< .0001		
positive (> 1.36, *n* = 153)						
sD at mediastinum-2						
negative (≤ 1.81, *n* = 178)	4.54 (2.86, 7.22)	< .0001	4.41 (2.78, 7.01)	< .0001		
positive (> 1.81, *n* = 129)						
R2	Hazard ratio	*P* Value	Hazard ratio	*P* Value	Hazard ratio	*P* Value
tD at lung-1						
negative (≤ 2.49, *n* = 166)	3.64 (2.29, 5.78)	< .0001	3.35 (2.11, 5.35)	< .0001		
positive (> 2.49, *n* = 141)						
sD at lung-1						
negative (≤ 1.98, *n* = 165)	4.53 (2.79, 7.33)	< .0001	4.40 (2.71, 7.13)	< .0001		
positive (> 1.98, *n* = 142)						
sD/tD at lung-1						
negative (≤ 0.82, *n* = 143)	3.69 (2.24, 6.07)	< .0001	3.50 (2.11, 5.79)	< .0001		
positive (> 0.82, *n* = 164)						
tD at lung-2						
negative (≤ 2.51, *n* = 168)	3.75 (2.36, 5.95)	< .0001	3.42 (2.14, 5.45)	< .0001		
positive (> 2.51, *n* = 139)						
sD at lung-2						
negative (≤ 1.93, *n* = 168)	4.77 (2.95, 7.74)	< .0001	4.66 (2.88, 7.56)	< .0001	2.76 (1.53, 4.98)	.0008
positive (> 1.93, *n* = 139)						
sD/tD at lung-2						
negative (≤ 0.92, *n* = 153)	3.53 (2.19, 5.67)	< .0001	3.47 (2.13, 5.64)	< .0001		
positive (> 0.92, *n* = 154)						
sD at mediastinum-1						
negative (≤ 2.53, *n* = 237)	5.12 (3.37, 7.79)	< .0001	4.67 (3.05, 7.14)	< .0001	2.59 (1.54, 4.33)	.0003
positive (> 2.53, *n* = 70)						

(*continued on next page*)

Table 4 (*continued*)

CT Feature	Univariable Analysis		Univariable Analysis Adjusted for Age, Sex, and Brinkman Index		Multivariable Analysis Adjusted for Age, Sex, and Brinkman Index	
R1	**Hazard ratio**	***P* Value**	**Hazard ratio**	***P* Value**	**Hazard ratio**	***P* Value**
sD at mediastinum-2						
negative (≤ 1.90, *n* = 173)	4.11 (2.59, 6.53)	< .0001	4.07 (2.56, 6.47)	< .0001		
positive (> 1.90, *n* = 134)						
AI-based software	Hazard ratio	*P* Value	Hazard ratio	*P* Value	Hazard ratio	*P* Value
tD						
negative (≤ 2.10, *n* = 145)	4.74 (2.79, 8.05)	< .0001	4.44 (2.61, 7.55)	< .0001		
positive (> 2.10, *n* = 162)						
sD						
negative (≤ 2.21, *n* = 191)	4.38 (2.81, 6.83)	< .0001	4.18 (2.67, 6.53)	< .0001		
positive (> 2.21, *n* = 116)						
sD/tD						
negative (≤ 0.91, *n* = 153)	3.74 (2.31, 6.06)	< .0001	3.58 (2.20, 5.82)	< .0001		
positive (> 0.91, *n* = 154)						
tV						
negative (≤ 2.54, *n* = 139)	5.24 (3.00, 9.13)	< .0001	4.98 (2.85, 8.70)	< .0001	3.84 (2.17, 6.82)	< .0001
positive (> 2.54, *n* = 168)						
sV						
negative (≤ 2.35, *n* = 181)	4.73 (2.97, 7.51)	< .0001	4.42 (2.77, 7.04)	< .0001		
positive (> 2.35, *n* = 126)						
sV/tV						
negative (≤ 0.84, *n* = 188)	4.06 (2.61, 6.34)	< .0001	3.87 (2.47, 6.05)	< .0001	2.68 (1.68, 4.26)	< .0001
positive (> 0.84, *n* = 119)						

Numbers in parentheses are the 95% limit of agreement.

Evaluation of imaging findings is subject to considerable variability. Quantitative evaluation is generally considered to be highly objective and reproducible, but human measurements are subject to interobserver and intraobserver variability and are influenced by image display conditions and scan parameters. In this study, some differences in significant variables were found in the multivariate analysis of CT features measured by R1 and R2. This variability can reduce the accuracy and reproducibility of statistical results, thus negatively impacting the reliability of the study. The lack of definitive image display conditions that can achieve consensus in clinical practice for clinical staging could be addressed by supporting quantitative measurements with an AI-based CAD system.

One advantage of quantitative evaluation using AI-based CAD systems is the ease of volumetric measurement. Previous studies have shown that volume measurement is more accurate than diameter measurement for diagnosing malignancy[20] and for predicting prognosis.[21] The latest Fleischner Society lung nodule recommendations and the Lung CT Screening Reporting and Data System have also added volume measurement to the nodule assessment criteria.[22,23] In this study, volume-based features derived from AI were significant predictors of LNM and prognosis. Since AI-based CAD systems can automatically measure volume, without adding to clinicians' workload, volumetric measurement is likely to become standard practice in future.

This study had several limitations. First, this retrospective study may have included selection bias due to missing imaging or clinical information for some patients. Additionally, we used a commercially available AI for segmentation and measurement of tumor size and volume, which could have included errors, such as misidentifying other structures like blood vessels as tumors. Although AI can measure tumor volume with accuracy comparable to that of physicians,[7,8] systems that incorporate physician feedback to improve AI predictions are needed.[24] Furthermore, this study used only imaging information of the primary tumor. Previous reports have suggested that lymph node measurements also impact outcome,[25,26] indicating that incorporating information about lymph nodes and surrounding structures with primary tumor data could improve predictions. Lastly, the inclusion of slices thicker than 1 mm may have impacted predictions of lymph node metastasis and prognosis.

In conclusion, LNM and prognosis of patients with lung adenocarcinoma can be predicted based on radiologists' measurements of tumor parameters on CT images, but this is a subjective method that is also dependent on image display conditions. On the other hand, AI-based CAD systems can predict LNM and prognosis with complete reproducibility, regardless of image display conditions and independent of human readers. These results have the potential to provide a highly reproducible method for determining T descriptors in lung adenocarcinoma and may contribute to treatment decision-making and efficacy.

Table 5 Relationship of CT Findings in Radiologists and AI-based Software With Overall Survival

CT Feature	Univariable Analysis		Univariable Analysis Adjusted for Age, Sex, and Brinkman Index		Multivariable Analysis Adjusted for Age, Sex, and Brinkman Index	
R1	**Hazard Ratio**	***P* Value**	**Hazard Ratio**	***P* Value**	**Hazard Ratio**	***P* Value**
tD at lung-1						
negative ($\leq$ 2.23, $n = 154$)	4.74 (2.19, 10.2)	.0001	4.50 (2.07, 9.76)	.0001		
positive ($>$ 2.23, $n = 153$)						
sD at lung-1						
negative ($\leq$ 1.99, $n = 184$)	6.47 (3.10, 13.5)	$<$ .0001	6.21 (2.97, 13.0)	$<$ .0001	3.79 (1.54, 9.33)	.0037
positive ($>$ 1.99, $n = 123$)						
sD/tD at lung-1						
negative ($\leq$ 0.83, $n = 127$)	4.94 (2.08, 11.7)	.0003	4.21 (1.76, 10.1)	.0013		
positive ($>$ 0.83, $n = 180$)						
tD at lung-2						
negative ($\leq$ 2.17, $n = 150$)	5.27 (2.34, 11.9)	.0001	5.24 (2.32, 11.8)	.0001		
positive ($>$ 2.17, $n = 157$)						
sD at lung-2						
negative ($\leq$ 1.95, $n = 183$)	6.37 (3.05, 13.3)	$<$ .0001	5.99 (2.86, 12.5)	$<$ .0001		
positive ($>$ 1.95, $n = 124$)						
sD/tD at lung-2						
negative ($\leq$ 0.95, $n = 133$)	4.37 (1.94, 9.84)	.0004	3.72 (1.64, 8.47)	.0017		
positive ($>$ 0.95, $n = 174$)						
sD at mediastinum-1						
negative ($\leq$ 2.12, $n = 210$)	5.12 (2.69, 9.72)	$<$ .0001	4.86 (2.55, 9.24)	$<$ .0001		
positive ($>$ 2.12, $n = 97$)						
sD at mediastinum-2						
negative ($\leq$ 2.86, $n = 240$)	5.67 (3.08, 10.4)	$<$ .0001	5.10 (2.76, 9.42)	$<$ .0001	2.25 (1.07, 4.75)	.0327
positive ($>$ 2.86, $n = 67$)						
R2	Hazard ratio	*P* Value	Hazard ratio	*P* Value	Hazard ratio	*P* Value
tD at lung-1						
negative ($\leq$ 2.67, $n = 184$)	4.37 (2.24, 8.54)	$<$ .0001	3.93 (1.99, 7.75)	.0001		
positive ($>$ 2.67, $n = 123$)						
sD at lung-1						
negative ($\leq$ 2.83, $n = 230$)	5.42 (2.93, 10.1)	$<$ .0001	5.12 (2.75, 9.55)	$<$ .0001		
positive ($>$ 2.83, $n = 77$)						
sD/tD at lung-1						
negative ($\leq$ 0.86, $n = 150$)	3.49 (1.71, 7.10)	.0006	3.08 (1.49, 6.34)	.0023		
positive ($>$ 0.86, $n = 157$)						
tD at lung-2						
negative ($\leq$ 1.94, $n = 119$)	6.72 (2.40, 18.8)	.0003	6.50 (2.32, 18.2)	.0004	3.71 (1.22, 11.3)	.0208
positive ($>$ 1.94, $n = 188$)						
sD at lung-2						
negative ($\leq$ 2.80, $n = 236$)	5.52 (2.99, 10.2)	$<$ .0001	5.08 (2.74, 9.43)	$<$ .0001		
positive ($>$ 2.80, $n = 71$)						
sD/tD at lung-2						
negative ($\leq$ 0.92, $n = 153$)	3.62 (1.78, 7.38)	.0004	3.23 (1.56, 6.67)	.0016		
positive ($>$ 0.92, $n = 154$)						
sD at mediastinum-1						
negative ($\leq$ 2.53, $n = 237$)	5.60 (3.04, 10.3)	$<$ .0001	5.18 (2.80, 9.58)	$<$ .0001	3.30 (1.70, 6.40)	.0004
positive ($>$ 2.53, $n = 70$)						
sD at mediastinum-2						

(*continued on next page*)

Table 5 (*continued*)

CT Feature	Univariable Analysis		Univariable Analysis Adjusted for Age, Sex, and Brinkman Index		Multivariable Analysis Adjusted for Age, Sex, and Brinkman Index	
R1	**Hazard Ratio**	***P* Value**	**Hazard Ratio**	***P* Value**	**Hazard Ratio**	***P* Value**
negative (≤ 2.69, *n* = 232)	5.16 (2.80, 9.51)	< .0001	4.78 (2.58, 8.85)	< .0001		
positive (> 2.69, *n* = 75)						
AI-based software	Hazard ratio	*P* Value	Hazard ratio	*P* Value	Hazard ratio	*P* Value
tD						
negative (≤ 3.24, *n* = 238)	5.67 (3.07, 10.5)	< .0001	5.18 (2.78, 9.64)	< .0001		
positive (> 3.24, *n* = 69)						
sD						
negative (≤ 2.44, *n* = 212)	5.27 (2.77, 10.0)	< .0001	5.04 (2.64, 9.59)	< .0001		
positive (> 2.44, *n* = 95)						
sD/tD						
negative (≤ 0.91, *n* = 153)	3.95 (1.89, 8.27)	.0003	3.67 (1.74, 7.70)	.0006		
positive (> 0.91, *n* = 154)						
tV						
negative (≤ 8.34, *n* = 235)	5.39 (2.92, 9.95)	< .0001	4.93 (2.65, 9.15)	< .0001		
positive (> 8.34, *n* = 72)						
sV						
negative (≤ 7.03, *n* = 242)	6.06 (3.28, 11.2)	< .0001	5.50 (2.97, 10.2)	< .0001	3.54 (1.78, 7.01)	.0003
positive (> 7.03, *n* = 65)						
sV/tV						
negative (≤ 0.85, *n* = 196)	5.51 (2.77, 11.0)	.0001	4.99 (2.49, 10.0)	< .0001	2.81 (1.30, 6.11)	.009
positive (> 0.85, *n* = 111)						

Numbers in parentheses are the 95% limit of agreement.

Figure 5 Kaplan–Meier survival curves show that (A) patients with a total tumor volume (tV) >2.54 cm^3 had a significantly lower probability of disease-free survival (DFS) (*P* < .001) than did patients with a percentage of tV ≤2.54 cm^3, and (B) patients with the proportion of the solid component volume (sV/tV) >0.84 had a significantly lower probability of DFS (*P* < .001) than did patients with percentage of sV/tV ≤0.84.

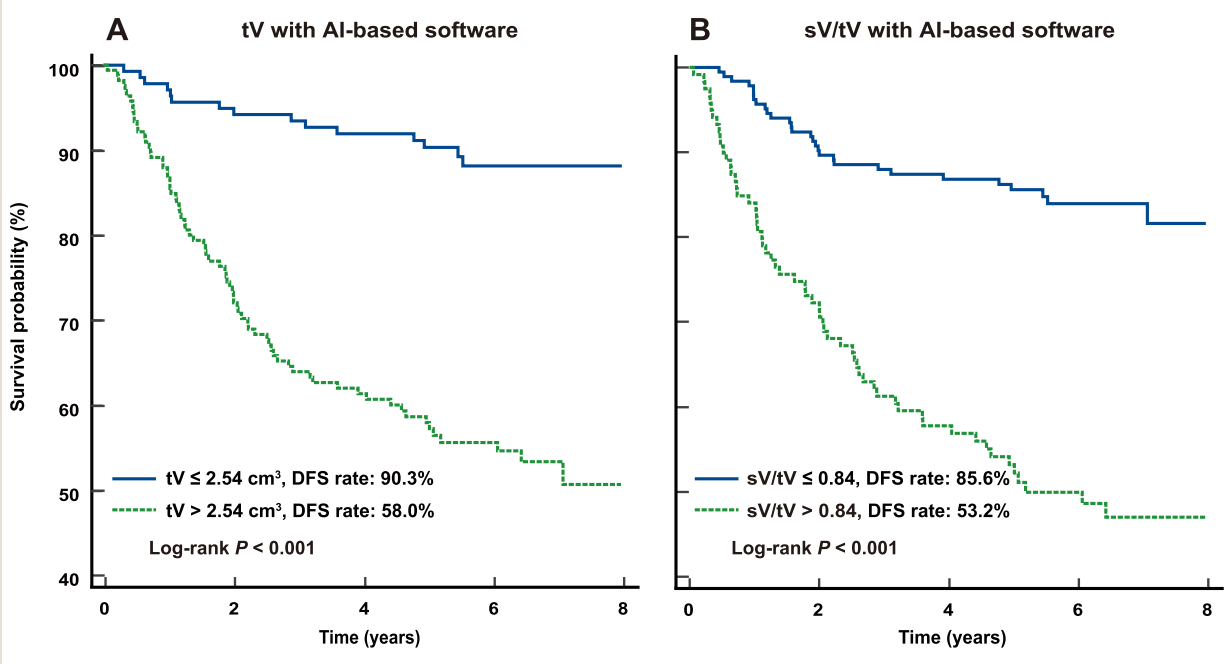
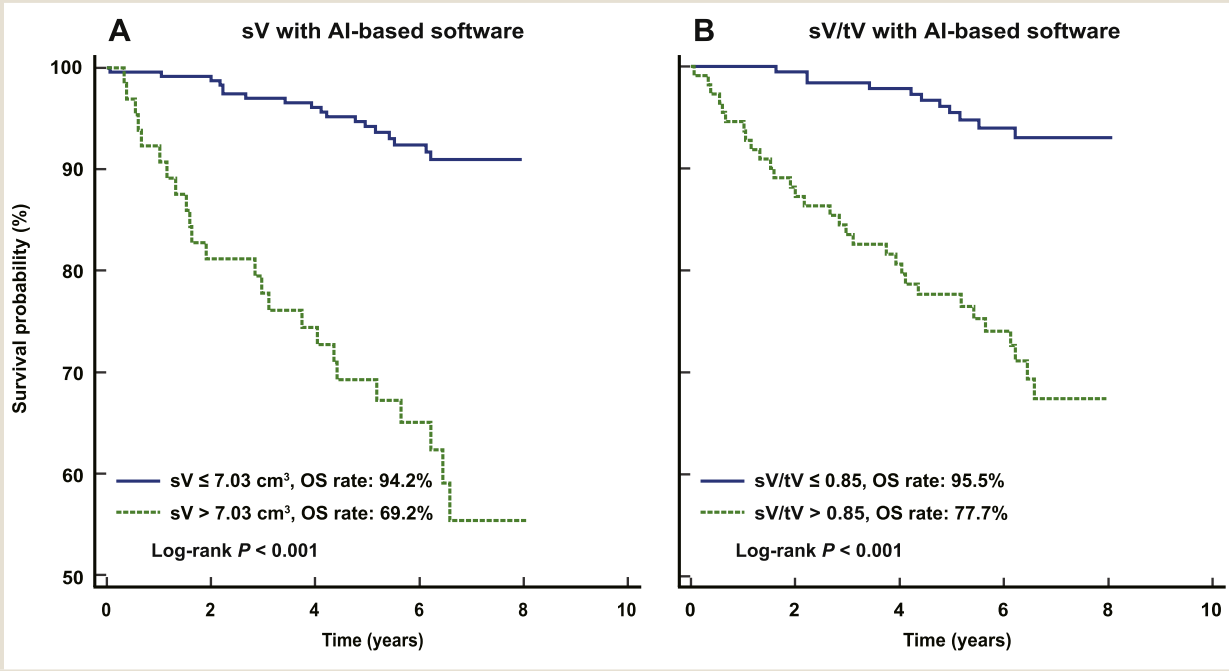


Figure 6 Kaplan–Meier survival curves show that (A) patients with a solid tumor volume (sV) >7.03 cm^3 had a significantly lower probability of overall survival (OS) (*P* < .001) than did patients with a percentage of sV ≤7.03 cm^3, and (B) patients with the proportion of the solid component volume (sV/tV) >0.85 had a significantly lower probability of OS (*P* < .001) than did patients with percentage of sV/tV ≤0.85.



Conclusion

Clinical Practice Points

- Tumor measurements by human readers are affected by CT display conditions.
- AI-based CAD systems are unaffected by CT display conditions and human readers.
- AI-based CAD systems can predict LNM, DFS, and OS with complete reproducibility.

Disclosure

The authors have stated that they have no conflicts of interest.

CRediT authorship contribution statement

Junya Sato: Writing – review & editing, Writing – original draft, Investigation, Funding acquisition, Formal analysis, Data curation. **Masahiro Yanagawa:** Writing – review & editing, Writing – original draft, Project administration, Methodology, Formal analysis, Conceptualization. **Daiki Nishigaki:** Writing – review & editing, Investigation, Data curation. **Akinori Hata:** Writing – review & editing, Investigation. **Yukinori Sakao:** Writing – review & editing, Resources, Investigation. **Noriaki Sakakura:** Writing – review & editing, Resources, Investigation. **Yasushi Yatabe:** Writing – review & editing, Investigation, Data curation. **Yasushi Shintani:** Writing – review & editing, Investigation. **Shoji Kido:** Writing – review & editing, Investigation. **Noriyuki Tomiyama:** Writing – review & editing, Investigation.

Acknowledgments

We express our gratitude to the financial support from JST SPRING for the professional proofreading of this work. This research was supported by JST SPRING (grant number JPMJSP2138).

References

1. Travis WD, Asamura H, Bankier AA, et al. The IASLC lung cancer staging project: proposals for coding T categories for subsolid nodules and assessment of tumor size in part-solid tumors in the forthcoming eighth edition of the TNM classification of lung cancer. *J Thorac Oncol*. 2016;11(8):1204–1223. doi:10.1016/j.jtho.2016.03.025.
2. Lee KH, Goo JM, Park SJ, et al. Correlation between the size of the solid component on thin-section CT and the invasive component on pathology in small lung adenocarcinomas manifesting as ground-glass nodules. *J Thorac Oncol*. 2014;9(1):74–82. doi:10.1097/JTO.0000000000000019.
3. Sun F, Huang Y, Yang X, et al. Solid component ratio influences prognosis of GGO-featured IA stage invasive lung adenocarcinoma. *Cancer Imaging*. 2020;20(1):87. doi:10.1186/s40644-020-00363-6.
4. Li J, You W, Zheng D, et al. A comprehensive evaluation of clinicopathologic characteristics, molecular features and prognosis in lung adenocarcinoma with solid component. *J Cancer Res Clin Oncol*. 2018;144(4):725–734. doi:10.1007/s00432-018-2588-6.
5. Ridge CA, Yildirim A, Boiselle PM, et al. Differentiating between subsolid and solid pulmonary nodules at CT: Inter- and intraobserver agreement between experienced thoracic radiologists. *Radiology*. 2016;278(3):888–896. doi:10.1148/radiol.2015150714.
6. Penn A, Ma M, Chou BB, Tseng JR, Phan P. Inter-reader variability when applying the 2013 Fleischner guidelines for potential solitary subsolid lung nodules. *Acta Radiol*. 2015;56(10):1180–1186. doi:10.1177/0284185114551975.

7. Wang S, Zhou M, Liu Z, et al. Central focused convolutional neural networks: developing a data-driven model for lung nodule segmentation. *Med Image Anal*. 2017;40:172–183. doi:10.1016/j.media.2017.06.014.
8. Ahn Y, Lee SM, Noh HN, et al. Use of a commercially available deep learning algorithm to measure the solid portions of lung cancer manifesting as subsolid lesions at CT: comparisons with radiologists and invasive component size at pathologic examination. *Radiology*. 2021;299(1):202–210. doi:10.1148/radiol.2021202803.
9. Travis WD, Brambilla E, Noguchi M, et al. International association for the study of lung cancer/American thoracic society/European respiratory society international multidisciplinary classification of lung adenocarcinoma. *J Thorac Oncol*. 2011;6(2):244–285. doi:10.1097/JTO.0b013e318206a221.
10. Mao H, Labh K, Han F, Jiang S, Yang Y, Sun X. Diagnosis of the invasiveness of lung adenocarcinoma manifesting as ground glass opacities on high-resolution computed tomography. *Thorac Cancer*. 2016;7(1):129–135. doi:10.1111/1759-7714.12269.
11. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307–310. doi:10.1016/S0140-6736(86)90837-8.
12. Seok Y, Yang HC, Kim TJ, et al. Frequency of lymph node metastasis according to the size of tumors in resected pulmonary adenocarcinoma with a size of 30 mm or smaller. *J Thorac Oncol*. 2014;9(6):818–824. doi:10.1097/JTO.0000000000000169.
13. Wang X, Zhao X, Li Q, et al. Can peritumoral radiomics increase the efficiency of the prediction for lymph node metastasis in clinical stage T1 lung adenocarcinoma on CT? *Eur Radiol*. 2019;29(11):6049–6058. doi:10.1007/s00330-019-06084-0.
14. Yanagawa M, Tanaka Y, Leung AN, et al. Prognostic importance of volumetric measurements in stage I lung adenocarcinoma. *Radiology*. 2014;272(2):557–567. doi:10.1148/radiol.14131903.
15. Xu S, Xi J, Jiang W, Lu S, Wang Q. Solid component and tumor size correlate with prognosis of stage IB lung adenocarcinoma. *Ann Thorac Surg*. 2015;99(3):961–967. doi:10.1016/j.athoracsur.2014.10.079.
16. Kim H, Goo JM, Suh YJ, Park CM, Kim YT. Implication of total tumor size on the prognosis of patients with clinical stage IA lung adenocarcinomas appearing as part-solid nodules: does only the solid portion size matter? *Eur Radiol*. 2019;29(3):1586–1594. doi:10.1007/s00330-018-5685-7.
17. Matsuguma H, Oki I, Nakahara R, et al. Comparison of three measurements on computed tomography for the prediction of less invasiveness in patients with clinical stage I non-small cell lung cancer. *Ann Thorac Surg*. 2013;95(6):1878–1884. doi:10.1016/j.athoracsur.2013.02.022.
18. Koo TR, Moon SH, Lim YJ, et al. The effect of tumor volume and its change on survival in stage III non-small cell lung cancer treated with definitive concurrent chemoradiotherapy. *Radiat Oncol*. 2014;9(1):283. doi:10.1186/s13014-014-0283-6.
19. Kamiya S, Iwano S, Umakoshi H, et al. Computer-aided volumetry of part-solid lung cancers by using CT: solid component size predicts prognosis. *Radiology*. 2018;287(3):1030–1040. doi:10.1148/radiol.2018172319.
20. Horeweg N, van Rosmalen J, Heuvelmans MA, et al. Lung cancer probability in patients with CT-detected pulmonary nodules: a prespecified analysis of data from the NELSON trial of low-dose CT screening. *Lancet Oncol*. 2014;15(12):1332–1341. doi:10.1016/S1470-2045(14)70389-4.
21. Xie HJ, Zhang X, Mo YX, Long H, Rong TH, Su XD. Tumor volume is better than diameter for predicting the prognosis of patients with early-stage non-small cell lung cancer. *Ann Surg Oncol*. 2019;26(8):2401–2408. doi:10.1245/s10434-019-07412-w.
22. Bankier AA, MacMahon H, Goo JM, Rubin GD, Schaefer-Prokop CM, Naidich DP. Recommendations for measuring pulmonary nodules at CT: a statement from the FLEISCHNER society. *Radiology*. 2017;285(2):584–600. doi:10.1148/radiol.2017162894.
23. Lung Rads. Accessed June 15, 2024. https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/Lung-Rads
24. Detterbeck FC, Boffa DJ, Kim AW, Tanoue LT. The eighth edition lung cancer stage classification. *Chest*. 2017;151(1):193–203. doi:10.1016/j.chest.2016.10.010.
25. Yu R, Jiang KW, Bao J, et al. PI-RADSAI: introducing a new human-in-the-loop AI model for prostate cancer diagnosis based on MRI. *Br J Cancer*. 2023;128(6):1019–1029. doi:10.1038/s41416-022-02137-2.
26. Bayanati H, E Thornhill R, Souza CA, et al. Quantitative CT texture and shape analysis: can it differentiate benign and malignant mediastinal lymph nodes in patients with primary lung cancer? *Eur Radiol*. 2015;25(2):480–487. doi:10.1007/s00330-014-3420-6.