Original Article

CT-based radiomics analysis for differentiation between thymoma and thymic carcinoma

Ryosuke Ohira[1], Masahiro Yanagawa[1], Yuki Suzuki[2], Akinori Hata[1], Tomo Miyata[3], Noriko Kikuchi[1], Yuriko Yoshida[1], Kazuki Yamagata[2], Shuhei Doi[1], Keisuke Ninomiya[1], Noriyuki Tomiyama[1]

[1]Department of Radiology, Osaka University Graduate School of Medicine, Osaka, Japan; [2]Department of Artificial Intelligence Diagnostic Radiology, Osaka University Graduate School of Medicine, Osaka, Japan; [3]Department of Future Diagnostic Radiology, Osaka University Graduate School of Medicine, Osaka, Japan

Contributions: (I) Conception and design: R Ohira, M Yanagawa; (II) Administrative support: M Yanagawa; (III) Provision of study materials or patients: A Hata, T Miyata, N Kikuchi, Y Yoshida, K Yamagata, S Doi, K Ninomiya; (IV) Collection and assembly of data: R Ohira, M Yanagawa, A Hata, N Kikuchi; (V) Data analysis and interpretation: R Ohira, M Yanagawa, Y Suzuki; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Masahiro Yanagawa, MD, PhD. Department of Radiology, Osaka University Graduate School of Medicine, 2-2 Yamadaoka, Suita-city, Osaka 565-0871, Japan. Email: m-yanagawa@radiol.med.osaka-u.ac.jp.

Background: The purpose of our study was to differentiate between thymoma and thymic carcinoma using a radiomics analysis based on the computed tomography (CT) image features.

Methods: The CT images of 61 patients with thymic epithelial tumors (TETs) who underwent contrast-enhanced CT with slice thickness <1 mm were analyzed. Pathological examination of the surgical specimens revealed thymoma in 45 and thymic carcinoma in 16. Tumor volume and the ratio of major axis to minor axis were calculated using a computer-aided diagnostic software. Sixty-one different radiomics features in the segmented CT images were extracted, then filtered and minimized by least absolute shrinkage and selection operator (LASSO) regression to select the optimal radiomics features for predicting thymic carcinoma. The association between the quantitative values and a diagnosis of thymic carcinoma were analyzed with logistic regression models. Parameters identified as significant in univariate analysis were included in multiple analyses. Receiver-operating characteristic (ROC) curves were assessed to evaluate the diagnostic performance.

Results: Thymic carcinoma was significantly predominant in men (P=0.001). Optimal radiomics features for predicting thymic carcinoma were as follows: gray-level co-occurrence matrix (GLCM)-homogeneity, GLCM-energy, compactness, large zone high gray-level emphasis (LZHGE), solidity, size of minor axis, and kurtosis. Multiple logistic regression analysis of these features revealed solidity and GLCM-energy as independent indicators associated with thymic carcinoma [odds ratio, 14.7 and 14.3; 95% confidence interval (CI): 1.6–139.0 and 3.0–68.7; and P=0.045 and 0.002, respectively]. Area under the curve (AUC) for diagnosing thymic carcinoma was 0.882 (sensitivity, 81.2%; specificity, 91.1%). Multivariate analysis adjusted for sex similarly revealed two features (solidity and GLCM-energy) as independent indicators associated with thymic carcinoma (odds ratio, 14.6 and 23.9; 95% CI: 2.4–89.2 and 1.9–302.8; P=0.004 and 0.014, respectively). Adjusted AUC for diagnosing thymic carcinoma was 0.921 (95% CI: 0.82–0.97): sensitivity, 62.5% and specificity, 100%.

Conclusions: Two texture features (GLCM-energy and solidity) were significant predictors of thymic carcinoma.

Keywords: Computed tomography (CT); radiomics; thymic epithelial tumors (TETs); thymic carcinoma

Submitted Dec 15, 2021. Accepted for publication Apr 07, 2022.
doi: 10.21037/jtd-21-1948
View this article at: https://dx.doi.org/10.21037/jtd-21-1948

Introduction

Thymic epithelial tumors (TETs) account for 47% of anterior mediastinal tumors (1). According to the World Health Organization (WHO) classification (2-4), TETs are divided into three histological types: type A thymoma, consisting of oval and spindle-shaped tumor cells; type B thymoma, consisting of round and polygonal tumor cells; and type AB thymoma, consisting of a mixture of these cell types. Type B thymoma is further divided into B1, B2, and B3 subtypes according to the morphology of the tumor cells and the number of immature T lymphocytes associated with the tumor. Subtypes A, AB, and B1 have low risk of malignancy, whereas subtypes B2 and B3 have high risk of malignancy. Thymomas in the low-risk group are likely to be completely resected and require no additional therapy at stage III. However, thymic carcinomas and those in the high-risk groups require multidisciplinary treatment with a combination of surgical resection, radiation therapy, and chemotherapy. This may include preoperative chemotherapy for thymoma, and chemotherapy or chemoradiotherapy for thymic carcinoma.

Preoperative evaluation of TETs is generally performed by computed tomography (CT). Previous studies that have attempted to predict the TET subtype using CT have reported that the features of lobulated or irregular tumor contour, capsule, septum, and heterogenous enhancement were helpful in distinguishing low-risk thymoma from high-risk thymoma and carcinoma (5-7). However, these visual features were evaluated subjectively and there may have been variability among the observers.

Radiomics features extracted from CT and magnetic resonance (MR) image data have been used in several recent studies (8-10). Tumor size and the difference in contrast compared with surrounding tissue are useful in determining tumor staging and for distinguishing between benign and malignant tumors in diagnostic imaging. However, radiomics analysis can extract and quantify a large number of high-dimensional features that cannot be observed visually, leading to clinical benefits for diagnosis and prediction of prognosis (11). To date, few studies have evaluated TET using radiomics analysis (12,13).

Tissue biopsy is essential in modern medicine for examining the genetic status of a tumor. However, because tumor contents are non-uniform spatially and temporally, there remain problems with the accuracy and reliability of biopsy as only some of the tissue is sampled. In addition, only a limited number of facilities are capable of genetic testing, and the time and financial burdens on patients are large. In contrast, CT is a simple and non-invasive examination, and is widely available. In addition, medical imaging has a major advantage in enabling characterization and evaluation of the entire tumor.

We hypothesized that more TET information could be obtained with radiomics analysis than visual CT analysis, and that this additional information could improve diagnostic performance in patients with TET. The purpose of this study was to evaluate whether thymic carcinoma and thymoma can be differentiated preoperatively by texture analysis. We present the following article in accordance with the STARD reporting checklist (available at https://jtd.amegroups.com/article/view/10.21037/jtd-21-1948/rc).

Methods

Study population

All procedures performed were approved by the ethical committee of Osaka University Hospital (No. 18096-2) and were conducted in accordance with the principles of the Declaration of Helsinki (as revised in 2013). The need for informed consent was waived due to the retrospective review of patient records and images. We searched the records of 67 patients with suspected TETs (27 men, 40 women; mean age ± SD, 59±14 years; range, 27–81 years) from 196 consecutive patients who underwent CT for the mediastinal between January 2010 to December 2013. The inclusion criteria were as follows (*Figure 1*): (I) thymectomy (n=51) or biopsy (n=12) performed between January 2010 to December 2013 with the diagnosis proven histologically, and (II) availability of CT images that were obtained with slice thickness of <1 mm. Of the 67 patients excluding follow-up (n=129), 4 were excluded because of CT slice thickness ≥5 mm, for which the software program cannot calculate the TET volume; and 2 because the pathological slides were not available to confirm the diagnosis and classify the tumors based on Masaoka staging. A final total of 61 patients with TETs were included in our study (*Table 1*).

Histopathological data

Pathological specimens were stained with hematoxylin and eosin and classified by pathologists at our institution into one of six subtypes, including thymic carcinoma based on the 4th edition of the 2015 WHO classification (4). In addition, we classified all tumors into low-risk thymoma,

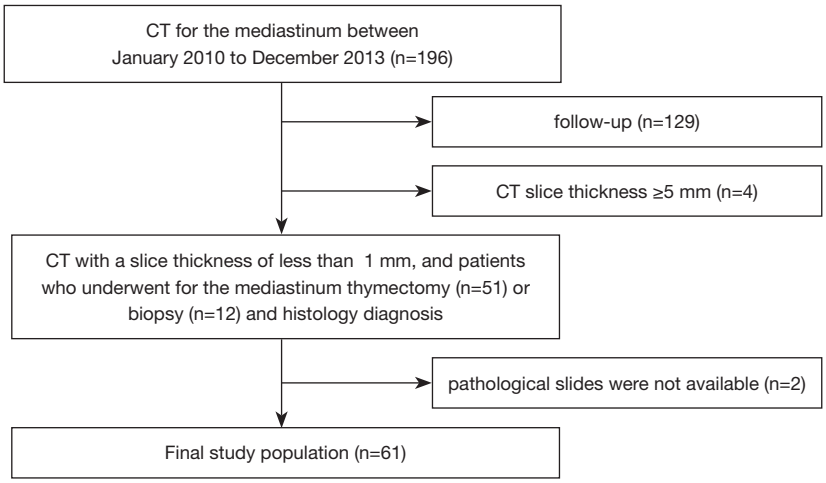

Figure 1 Flowchart of patient selection. CT, computed tomography.

Table 1 Patient demographics and tumor characteristics

Characteristic	Types	N	Age (years)	P value	Sex	P value
WHO classification						
Low risk thymoma	A	2	60 [27–76]	0.433*	Male 9; female 19	0.001#
	AB	13				
	B1	13				
High risk thymoma	B2	11	65 [31–81]		Male 3; female 14	
	B3	6				
Thymic carcinoma		16	62.5 [30–79]		Male 12; female 4	
Total		61	59 [27–81]		Male 24; female 37	
Comorbidities in patients with thymoma						
Myasthenia gravis		11				
Rheumatoid arthritis		1				
Systemic sclerosis		1				

Data are presented as the number or the mean [range]. *, there were no significant differences in the distribution of age between thymoma (low risk and high risk) and thymic carcinoma; #, there was a significant difference in the distribution of sex between thymoma (low risk and high risk) and thymic carcinoma. WHO, World Health Organization.

high-risk thymoma, and thymic carcinoma according to the prognostic value of WHO histologic classification, as described previously (4,14,15).

Chest CT protocol

CT of the chest was acquired using 64-detector row CT scanners: CT750HD (General Electric Medical Systems, Milwaukee, WI, USA), n=50, LightSpeed VCT (General Electric Medical Systems), n=8, and Aquilion ONE (Toshiba Medical Systems, Otawara, Japan), n=3. The following parameters were used: collimation, 0.5 or 0.625 mm; pitch, 0.828–1.375; rotation time, 0.4–0.5 s per rotation; field of view, 345 mm; tube voltage, 120 kVp with automatic tube current control. All CT images of 0.5–0.625 mm section thickness were reconstructed with a standard kernel using 30% adaptive statistical iterative reconstruction. All images were obtained 60 s after injection of 2 mL/kg bodyweight of

contrast material (IOHEXOL, 300 mg I/cc; Daiichi Sankyo Company, Limited, Tokyo, Japan).

Volumetric and radiomic feature extraction

We developed a texture analysis software by customizing a commercially available software (WatchinGGO; LISIT, Co., Ltd., Tokyo, Japan) (15). Our developed software enables extraction of radiomics features from selected two dimensional (2D) CT images as well as volumetric measurements. In terms of statistics-based features, we used Gray-Level, developed by Vallières *et al.* (16), in first-order statistics employing the following algorithm features: gray-level co-occurrence matrix (GLCM), neighborhood gray-tone-difference matrix (NGTDM), and grey-level run length matrix (GLRLM). Size and Shape Index features were calculated as morphology-based features. All feature classes installed in our developed software enabled calculation of both 2D and three dimensional (3D) computable indexes. Segmentation software was equipped with general segmentation tools such as manual drawing of region of interest (ROI), 2D ROI segmentation by the level set method, thresholding, and morphological opening and closing.

In this study, tumor volume and the ratio of major axis to minor axis were calculated by one chest radiologist using our developed software without being informed of any clinical information. The volume was calculated in the same way as in the previous study (15). The radiomics features were divided into the following seven classes: shape, size, intensity, histogram, GLCM, NGTDM, and GLRLM. A total of 61 radiomics features were calculated in the segmented CT images using the software (*Figure 2*). The image showing the maximal cross-sectional for each tumor was selected independently by two chest radiologists (MY and AH). In the case of difference between the radiologists, the final decision was made in consensus. Semi-automatic segmentation was performed in the maximum cross-sectional image of each tumor, which was as large as possible to minimize inter-tumor variability.

Statistical analysis

The relationship between age and sex distribution and WHO classification was analyzed by Mann-Whitney U test and chi-square test, respectively. Radiomics features in the low-risk thymoma, high-risk thymoma, and thymic carcinoma groups were filtered through Pearson correlation coefficient analysis to agglomerate redundant features and were subjected to least absolute shrinkage and selection operator (LASSO) logistic regression with a 100-time repeated 10-fold cross validation. The hyper-parameter of LASSO regression (weight parameter for the regularization term) was determined by the internal cross validation within each fold. Features with non-zero regression coefficients were considered important and importance of a feature was determined by counting how many times the feature was considered important through the repeated cross validation. We created the radiomics features via linear combination of the chosen features. A radiomics score was calculated for each patient, using the linear combination of the chosen features, weighted according to their coefficients. We selected the most relevant radiomics features in each of the three groups. Moreover, we added volume as one of the explanatory factors because tumor size has been reported to be one of the characteristic factors for histological subtypes of TETs (17-19). For each feature, the cutoff value that yielded the largest difference in numbers of patients with and without thymic carcinoma was determined using the receiver-operating characteristic (ROC) method. Optimal thresholds were determined for each variable separately using the Youden index (the highest sum of sensitivity and specificity). Associations between thymic carcinoma and each binary group designated by the cutoff value for the seven radiomics features were evaluated by univariate logistic regression analysis. Significant parameters identified by univariate analysis were included in multiple logistic regression (stepwise method; P value of 0.05 or less was used for entry into the model and P value greater than 0.1 was selected for removal). ROC curves were generated for prediction of thymic carcinoma: sensitivity, specificity, and area under the curve (AUC). P values <0.05 were considered significant.

Results

Patient data

Table 1 summarizes the patient characteristics. Of the 45 patients with thymoma, 28 had low-risk thymoma (type A, n=2; type AB, n=13; and type B1, n=13) and 17 had high-risk thymoma (type B2, n=11; and type B3, n=6). Of the 16 patients with thymic carcinoma, 15 had squamous cell carcinoma and 1 had mucoepidermoid carcinoma. There were no significant differences in the distribution of age between thymoma (low risk and high risk) and thymic carcinoma (P=0.433). Thymic carcinoma was significantly predominant in men (P=0.001). Thirteen patients with

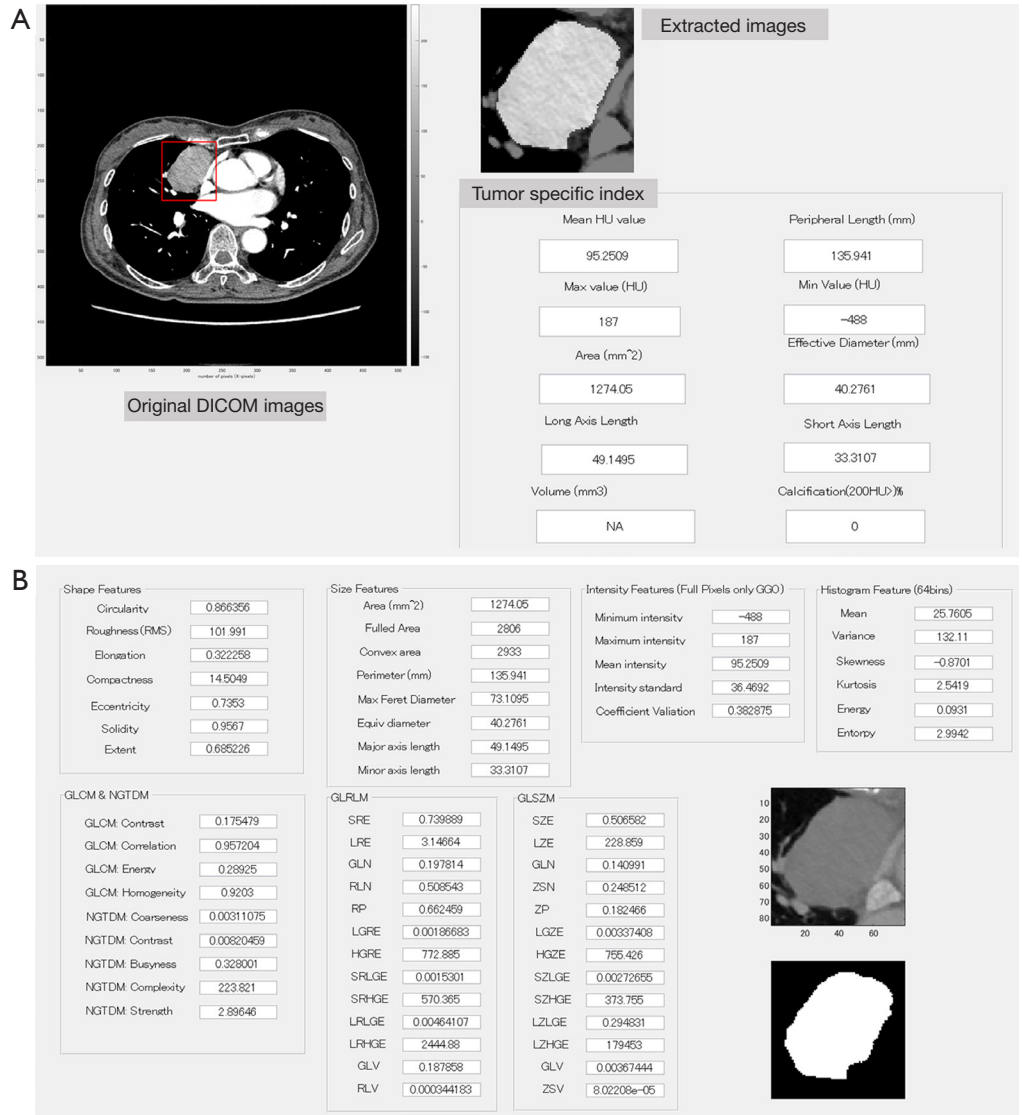

Figure 2 CT radiomics features extracted using our developed software. The area of interest including the mass is enclosed manually in a square, and the tumor is extracted automatically (A). The extent of the mass is then determined manually and the CT radiomics features are measured (B). DICOM, digital imaging and communications in medicine; HU, Hounsfield unit; NA, not applicable; RMS, root mean square; GGO, ground-glass opacity; GLCM, gray-level co-occurrence matrix; NGTDM, neighborhood gray-tone-difference matrix; GLRLM, grey-level run length matrix; GLSZM, gray-level size zone matrix; SRE, short run emphasis; LRE, long run emphasis; GLN, gray level non-uniformity; RLN, run-length nonuniformity; RP, run percentage; LGRE, low gray-level run emphasis; HGRE, high gray-level run emphasis; SRLGE, short run low gray-level emphasis; SRHGE, short run high gray-level emphasis; LRLGE, long run low gray-level emphasis; LRHGE, long run high gray-level emphasis; GLV, gray level variance; RLV, run length variance; SZE, small zone emphasis; LZE, large zone emphasis; ZSN, zone size non-uniformity; ZP, zone percentage; LGZE, low gray-level zone emphasis; HGZE, high gray-level zone emphasis; SZLGE, small zone low gray-level emphasis; SZHGE, small zone low gray-level emphasis; LZLGE, large zone low gray-level emphasis; LZHGE, large zone high gray-level emphasis; ZSV, zone-size variance; CT, computed tomography.

thymoma had comorbidities: myasthenia gravis (n=11), rheumatoid arthritis (n=1), and systemic sclerosis (n=1).

Radiomic features importance

The dimensionality of the 61 texture features was reduced using LASSO, and the radiomics features of most importance were identified (*Figure 3*). We extracted the seven most important radiomics features associated with the low-risk thymoma, high-risk thymoma, and thymic carcinoma groups: GLCM: homogeneity, GLCM: energy, compactness, LZHGE, solidity, minor axis length, and kurtosis.

Predictive performance for thymic carcinoma using radiomics features

Table 2 shows the cutoff values obtained by ROC analysis for each of the seven radiomics features. Univariate logistic regression analysis of these characteristics and the semi-automatic tumor data (volume and ratio of major axis to minor axis), showed significant differences between thymoma and thymic carcinoma in all features except Ratio (major axis length/minor axis length) (*Table 3*). Multiple logistic regression analysis of these factors revealed two features (solidity and GLCM-energy) as independent indicators associated with thymic carcinoma [odds ratio, 14.7 and 14.3; 95% confidence interval (CI): 1.6–139.0 and 3.0–68.7; P=0.045 and 0.002, respectively] (*Table 3*). AUC for diagnosing thymic carcinoma was 0.882 (95% CI: 0.77–0.95): sensitivity, 81.2% and specificity, 91.1%. In a subanalysis comparing high-risk thymoma and thymic carcinoma, multiple logistic regression analysis revealed GLCM-energy as an independent indicator associated with thymic carcinoma (odds ratio, 69.3; 95% CI: 6.4–748.1; P=0.0005). AUC for diagnosing thymic carcinoma was 0.877 (95% CI: 0.72–0.97): sensitivity, 81.3% and specificity, 94.1%. Differentiation between low-risk and high-risk thymoma on the basis of the radiomic features did not reach statistical significance (AUC 0.499; 95% CI: 0.44–0.55).

Considering the predominance of men in patients with thymic carcinoma, multivariate analysis adjusted for sex similarly revealed two features (solidity and GLCM-energy) as independent indicators associated with thymic carcinoma (odds ratio, 14.6 and 23.9; 95% CI: 2.4–89.2 and 1.9–302.8; P=0.004 and 0.014, respectively). Adjusted AUC for diagnosing thymic carcinoma was 0.921 (95% CI: 0.82–0.97): sensitivity, 62.5% and specificity, 100%.

Discussion

Our study demonstrated that the CT radiomics features GLCM-energy and solidity were useful and had high specificity for predicting thymic carcinoma. Particularly, in comparing high-risk thymoma and thymic carcinoma, GLCM-energy was as an independent indicator associated with thymic carcinoma. The high specificity of our model suggests its suitability for use in tertiary hospitals to reduce unnecessary examinations. Radiomics analysis might be an effective tool in differentiating between thymoma and thymic carcinoma in clinical treatment planning, particularly in cases for which surgery or biopsy is not feasible. In addition, if thymic carcinoma is diagnosed using radiomics features, metastasis can be detected by positron emission tomography (PET)-CT, histological type can be determined by biopsy, and staging and an appropriate treatment plan can be determined.

The feature GLCM-energy extracted in this study is difficult to observe subjectively on CT images. GLCM-energy is to measure the degree of fluctuation in the space in the tumor. The feature GLCM-energy extracted in this study is difficult to observe subjectively on CT images. Generally, squamous cell carcinoma is the most frequent subtype of thymic carcinoma, which is composed of large polyhedral cells arranged in nests and cords, showing evidence of keratinization and/or intercellular bridges. Foci of spontaneous necrosis are frequently seen, as is the invasion of intratumoral blood vessels, resulting in the heterogeneity of the tumor (4). GLCM-energy might correlate to the heterogeneity of thymic carcinoma. 'Solidity' is the ratio of the number of pixels in the tumor region to the number of pixels in the convex hull of the tumor region, and takes a value of 1 at the maximum (20). A value of 1 signifies a solid object, and a value less than 1 will signify an object having an irregular boundary, or containing holes. Thymic carcinoma might indicate lower solidity than thymoma because of irregular margin and internal necrosis. Chen *et al.* reported that solidity had smaller values for high-risk tumors than for low-risk tumors, and considered that this finding was due to low-risk TETs being mostly well-differentiated with a complete capsule, whereas high-risk TETs were typically poorly differentiated without a capsule (12,21,22). This finding is consistent with the results of the present study.

Some previous studies have used CT radiomics to evaluate TETs. Chen *et al.* showed that radiomics features were independent predictors (AUC: 0.944, 95% CI:

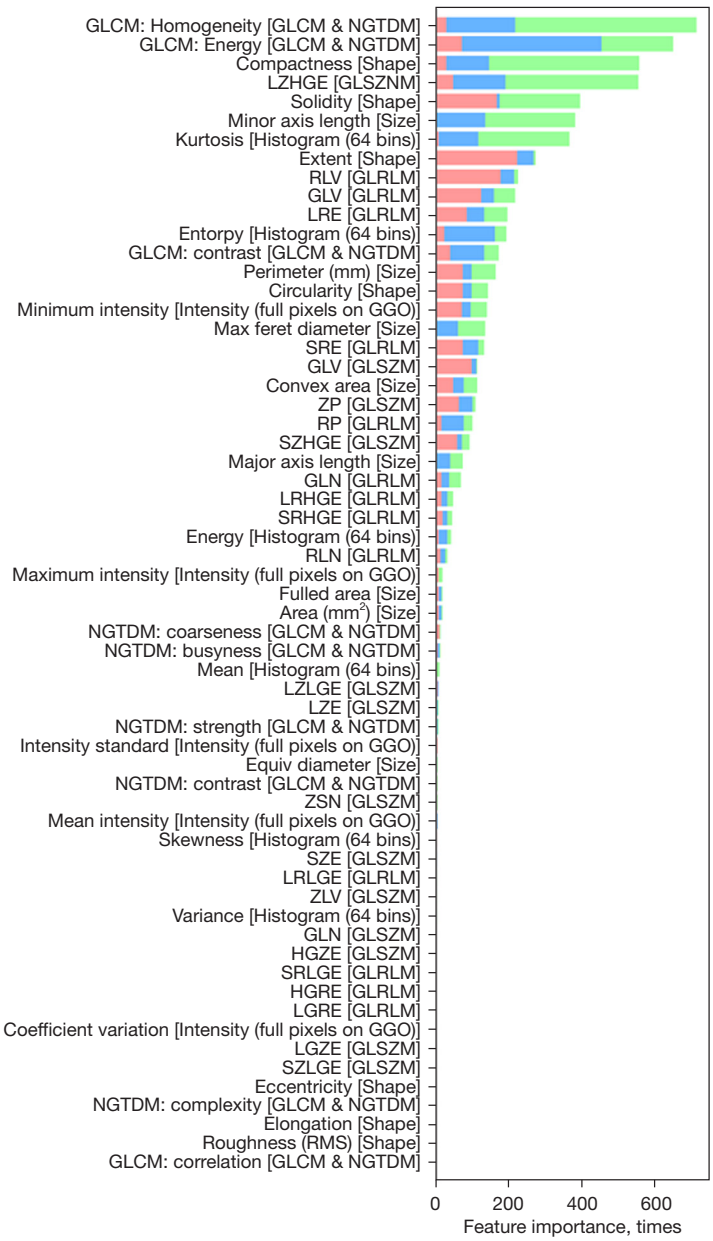

Figure 3 Important radiomics features associated with the three groups (red, low-risk thymoma; blue, high-risk thymoma; and green, thymic carcinoma). Feature importance was defined as the number of times for a feature to have non-zero LASSO regression coefficient over the repeated cross validation. For our 100-time repeated 10-fold cross validation, maximum possible importance was 1,000. GLCM, gray-level co-occurrence matrix; NGTDM, neighborhood gray-tone-difference matrix; LZHGE, large zone high gray-level emphasis; GLSZM, gray-level size zone matrix; RLV, run length variance; GLRLM, grey-level run length matrix; GLV, gray level variance; LRE, long run emphasis; GGO, ground-glass opacity; SRE, short run emphasis; ZP, zone percentage; RP, run percentage; SZHGE, small zone low gray-level emphasis; GLN, gray level non-uniformity; LRHGE, long run high gray-level emphasis; SRHGE, short run high gray-level emphasis; RLN, run-length nonuniformity; LZLGE, large zone low gray-level emphasis; LZE, large zone emphasis; ZSN, zone size non-uniformity; SZE, small zone emphasis; LRLGE, long run low gray-level emphasis; ZLV, zone level variance; GLN, gray level non-uniformity; HGZE, high gray-level zone emphasis; SRLGE, short run low gray-level emphasis; HGRE, high gray-level run emphasis; LGRE, low gray-level run emphasis; LGZE, low gray-level zone emphasis; SZLGE, small zone low gray-level emphasis; RMS, root mean square; LASSO, least absolute shrinkage and selection operator.

Table 2 Cutoff values of radiomics features for predicting thymic carcinoma and mean value in the binarized groups

Features	Cut-off value	Mean ± SD
Volume		
Score =0 (n=26)		10.1±7.9
Score =1 (n=35)	>23.6	114.8±116.7
Ratio (major axis length/minor axis length)		
Score =0 (n=26)		1.9±0.4
Score =1 (n=35)	<1.60	1.4±0.2
Compactness		
Score =0 (n=26)		14.6±1.0
Score =1 (n=35)	>16.1	21.5±6.3
Solidity		
Score =0 (n=26)		0.95±0.01
Score =1 (n=35)	<0.92	0.85±0.06
Minor axis length		
Score =0 (n=39)		22.8±7.8
Score =1 (n=22)	>34.4	47.6±8.8
GLCM-energy		
Score =0 (n=39)		0.23±0.04
Score =1 (n=22)	>0.29	0.36±0.04
GLCM-homogeneity		
Score =0 (n=33)		0.92±0.01
Score =1 (n=28)	>0.93	0.94±0.01
LZHGE		
Score =0 (n=34)		$4.78\times10^5\pm6.05\times10^5$
Score =1 (n=27)	$>1.71\times10^6$	$5.64\times10^6\pm3.76\times10^6$
Kurtosis		
Score =0 (n=37)		2.1±0.5
Score =1 (n=24)	>2.83	4.6±2.1

SD, standard deviation; GLCM, gray-level co-occurrence matrix; LZHGE, long-zone high-grey level emphasis.

0.874–0.981) to distinguish between high-risk and low-risk tumors, and that these features were significantly more accurate than subjective assessment by radiologists (AUC: 0.731, 95% CI: 0.627–0.819) (P<0.001) (12). However, there was no significant difference between evaluation that combined the radiomics features compared with subjective evaluation (P=0.266). Wang *et al.* reported that in discriminating between high- and low-risk thymomas, AUC values were 0.801 (95% CI: 0.740–0.863) for the radiomics signature based on non-enhanced CT images and 0.827 (95% CI: 0.771–0.884) for the radiomics signature based on enhanced CT images (13). There was no significant difference between non-enhanced and enhanced CT-based radiomics features (P=0.365). Both of these results were superior to subjective evaluation by radiologists (AUC: 0.731, 0.779, respectively), and the present study also showed high performance for diagnosing thymic carcinoma (AUC: 0.882) with high specificity. Higher specificity might suppress the tendency for overdiagnosis by radiologists, resulting in a positive effect on patient management and treatment strategies.

Predicting the TET subtype is valuable clinically in the assessment and treatment of patients with thymoma (23). Compared with patients with low-risk thymoma, those with high-risk thymoma are more likely to require postoperative radiotherapy and chemotherapy (24) and have lower 5- and 10-year overall survival rates (25). Previous reports have reported that tumor volume can be larger for thymic carcinoma than thymoma, and that tumor volume can be larger for invasive TET than non-invasive TET (7,15,18,19). However, these studies could not distinguish between WHO subtypes or Masaoka stages. In the previous studies (7,15) also, univariate analysis revealed that tumor volume was useful for distinguishing between thymoma and thymic carcinoma, which was in accordance with the present study. But multivariate analysis using volume and radiomics features did not show that volume had a significant difference. In the present study, the CT radiomics features of GLCM-energy and solidity analyzed using our developed software could predict thymic carcinoma more accurately compared with volumetry alone. In general, thymic carcinoma has little gender difference (4), but in the present study, multivariate analysis adjusted for sex was also performed because of the predominant distribution in men. The result was almost the same: two texture features (GLCM-energy and solidity) were significant predictors of thymic carcinoma. Xiao *et al.* also reported that MRI-based radiomics signature is a noninvasive and reliable tool to differentiate between low- and high-risk thymomas preoperatively (26). Thus, radiomics analysis using imaging data is a more useful diagnostic tool than visual or volumetry evaluations for predicting histologic subtypes of TET. In clinical decision-making, evaluation of tumor invasion is important as well as the histological diagnosis.

Table 3 Relationship of radiomics features with prediction of thymic carcinoma

Radiomic feature	Univariate analysis					Multivariate analysis		
	Thymoma (n)	Thymic carcinoma (n)	Odds ratio	95% confidence interval	P value	Odds ratio	95% confidence interval	P value
Volume			8	1.6–39.4	0.011			
Score =0 (n=26)	24	2						
Score =1 (n=35)	21	14						
Ratio (major axis length/minor axis length)			2.9	0.8–10.3	0.105			
Score =0 (n=26)	22	4						
Score =1 (n=35)	23	1						
Compactness			8	1.6–9.4	0.011			
Score =0 (n=26)	24	2						
Score =1 (n=35)	21	14						
Solidity			18.8	2.3–154.3	0.006	14.7	1.6–139.0	0.045
Score =0 (n=26)	25	1						
Score =1 (n=35)	20	15						
Minor axis length			6.8	1.9–23.9	0.003			
Score =0 (n=39)	34	5						
Score =1 (n=22)	11	11						
GLCM-energy			17.3	4.1–74.1	0.0001	14.3	3.0–68.7	0.002
Score =0 (n=39)	36	3						
Score =1 (n=22)	9	13						
GLCM-homogeneity			15.5	3.1–77.6	0.001			
Score =0 (n=33)	31	2						
Score =1 (n=28)	14	14						
LZHGE			9.6	2.4–39.1	0.002			
Score =0 (n=34)	31	3						
Score =1 (n=27)	14	13						
Kurtosis			3.7	1.1–12.2	0.032			
Score =0 (n=37)	31	6						
Score =1 (n=24)	14	10						

GLCM, gray-level co-occurrence matrix; LZHGE, large zone high gray-level emphasis.

However, the pathological physiology of tumor invasion is complex, and further analysis will be required to elucidate the relationship between these pathological features and CT-based radiomics features in the future.

Our study had some limitations. First, as we performed the analysis with 2D texture analysis using maximum cross-sectional CT images, the possibility exists that the site characterizing the tumor was not contained. The further development and evaluation of software capable of 3D texture analysis is necessary in this regard. Second, the small number of patients in this study may have limited statistical detection. Class imbalance (45 thymomas and 16

thymic carcinomas) may potentially affect the investigation and induces some bias, what was not alleviated by proper strategies. Moreover, only 10-fold cross validation was performed because of the small number of cases in the present study. Validation using other cohort might have been a critical step in the workflow radiomics process. Third, minor discrepancies in volume measurements during semi-automatic tumor segmentation may have influenced the results of texture analyses. Development of fully automatic calculation software for mediastinal tumors is desirable. Finally, the radiomics analysis was performed using our developed model in a single institution. External validation using a larger cohort is needed to acquire high-level evidence for application to clinical practice.

In conclusion, CT-based radiomics features are useful for predicting thymic carcinoma. Two texture features (GLCM-energy and solidity) are significant predictors of thymic carcinoma, which may be helpful for determining the patient's treatment preoperatively. In particular, the high specificity of our radiomics model may reduce unnecessary examinations and contribute to clinical treatment planning and management of TET.

Acknowledgments

We thank Shuji Yamamoto (LISIT, Inc :Life Saving Imaging Technologies) for his technical support and the modifications made on the GGO® software. The scientific guarantor of this publication is Professor Noriyuki Tomiyama MD, PhD. No complex statistical methods were necessary for this paper. Study subjects or cohorts have been previously reported on *J Thorac Dis* 2018;10(10):5822-5832.
Funding: None.

Footnote

Reporting Checklist: The authors have completed the STARD reporting checklist Available at https://jtd.amegroups.com/article/view/10.21037/jtd-21-1948/rc

Data Sharing Statement: Available at https://jtd.amegroups.com/article/view/10.21037/jtd-21-1948/dss

Peer Review File: Available at https://jtd.amegroups.com/article/view/10.21037/jtd-21-1948/prf

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://jtd.amegroups.com/article/view/10.21037/jtd-21-1948/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. All procedures performed were approved by the ethics committee of Osaka University Hospital (No. 18096-2) and were conducted in accordance with the principles of the Declaration of Helsinki (as revised in 2013). The need for informed consent was waived due to the retrospective review of patient records and images.

References

1. Fujii Y. Published guidelines for management of thymoma. Thorac Surg Clin 2011;21:125-9, viii.
2. Travis WD, Brambilla E, Müller-Hermelink HK,et al. Pathology and Genetics of Tumours of the Lung, Pleura, Thymus and Heart, 3rd ed. Lyon: IARC Press; 2004.
3. Suster S, Moran CA. Histologic classification of thymoma: the World Health Organization and beyond. Hematol Oncol Clin North Am 2008;22:381-92.
4. Travis WD, Brambilla E, Burke AP, et al. WHO Classification of Tumours of Lung, Pleura, Thymus & Heart, 4th ed. Lyon: IARC Press; 2015.
5. Sadohara J, Fujimoto K, Müller NL, et al. Thymic epithelial tumors: comparison of CT and MR imaging findings of low-risk thymomas, high-risk thymomas, and thymic carcinomas. Eur J Radiol 2006;60:70-9.
6. Marom EM, Milito MA, Moran CA, et al. Computed tomography findings predicting invasiveness of thymoma. J Thorac Oncol 2011;6:1274-81.
7. Tomiyama N, Johkoh T, Mihara N, et al. Using the World Health Organization Classification of thymic epithelial

neoplasms to describe CT findings. AJR Am J Roentgenol 2002;179:881-6.
8. Wilson R, Devaraj A. Radiomics of pulmonary nodules and lung cancer. Transl Lung Cancer Res 2017;6:86-91.
9. Tagliafico AS, Piana M, Schenone D, et al. Overview of radiomics in breast cancer diagnosis and prognostication. Breast 2020;49:74-80.
10. Sui H, Liu L, Li X, et al. CT-based radiomics features analysis for predicting the risk of anterior mediastinal lesions. J Thorac Dis 2019;11:1809-18.
11. Li GW, Xie XS. Central dogma at the single-molecule level in living cells. Nature 2011;475:308-15.
12. Chen X, Feng B, Li C, et al. A radiomics model to predict the invasiveness of thymic epithelial tumors based on contrast-enhanced computed tomography. Oncol Rep 2020;43:1256-66.
13. Wang X, Sun W, Liang H, et al. Radiomics Signatures of Computed Tomography Imaging for Predicting Risk Categorization and Clinical Stage of Thymomas. Biomed Res Int 2019;2019:3616852.
14. Jeong YJ, Lee KS, Kim J, et al. Does CT of thymic epithelial tumors enable us to differentiate histologic subtypes and predict prognosis? AJR Am J Roentgenol 2004;183:283-9.
15. Sato Y, Yanagawa M, Hata A, et al. Volumetric analysis of the thymic epithelial tumors: correlation of tumor volume with the WHO classification and Masaoka staging. J Thorac Dis 2018;10:5822-32.
16. Vallières M, Freeman CR, Skamene SR, et al. A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. Phys Med Biol. 2015;60:5471-96.
17. Chang S, Hur J, Im DJ, et al. Volume-based quantification using dual-energy computed tomography in the differentiation of thymic epithelial tumours: an initial experience. Eur Radiol 2017;27:1992-2001.
18. Liu GB, Qu YJ, Liao MY, et al. Relationship between computed tomography manifestations of thymic epithelial tumors and the WHO pathological classification. Asian Pac J Cancer Prev 2012;13:5581-5.
19. Blumberg D, Port JL, Weksler B, et al. Thymoma: a multivariate analysis of factors predicting survival. Ann Thorac Surg 1995;60:908-13; discussion 914.
20. Zwanenburg A, Vallières M, Abdalah MA, et al. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. Radiology 2020;295:328-38.
21. Detterbeck FC. Clinical value of the WHO classification system of thymoma. Ann Thorac Surg 2006;81:2328-34.
22. Qu YJ, Liu GB, Shi HS, et al. Preoperative CT findings of thymoma are correlated with postoperative Masaoka clinical stage. Acad Radiol 2013;20:66-72.
23. Okumura M, Ohta M, Tateyama H, et al. The World Health Organization histologic classification system reflects the oncologic behavior of thymoma: a clinical study of 273 patients. Cancer 2002;94:624-32.
24. Berghmans T, Durieux V, Holbrechts S, et al. Systemic treatments for thymoma and thymic carcinoma: A systematic review. Lung Cancer 2018;126:25-31.
25. Falkson CB, Bezjak A, Darling G, et al. The management of thymoma: a systematic review and practice guideline. J Thorac Oncol 2009;4:911-9.
26. Xiao G, Hu YC, Ren JL, et al. MR imaging of thymomas: a combined radiomics nomogram to predict histologic subtypes. Eur Radiol 2021;31:447-57.

Cite this article as: Ohira R, Yanagawa M, Suzuki Y, Hata A, Miyata T, Kikuchi N, Yoshida Y, Yamagata K, Doi S, Ninomiya K, Tomiyama N. CT-based radiomics analysis for differentiation between thymoma and thymic carcinoma. J Thorac Dis 2022;14(5):1342-1352. doi: 10.21037/jtd-21-1948