

CT Imaging-based Machine Learning Model: A Potential Modality for Predicting Low-Risk and High-Risk Groups of Thymoma “Impact of surgical modality choice”

Ayten KAYICANGIR (✉ cangir@medicine.ankara.edu.tr)

Ankara Üniversitesi Tıp Fakültesi <https://orcid.org/0000-0002-2052-1642>

Kaan ORHAN

Ankara University Faculty of Dentistry: Ankara Üniversitesi Dis Hekimliği Fakültesi

Yusuf KAHYA

Ankara Üniversitesi Tıp Fakültesi

Hilal ÖZAKINCI

Ankara University Faculty of Medicine: Ankara Üniversitesi Tıp Fakültesi

Betül Bahar KAZAK

Ankara University Faculty of Medicine: Ankara Üniversitesi Tıp Fakültesi

Buse Mine KONUK BALCI

Ankara University Faculty of Medicine, Department of Thoracic Surgery

Duru KARASOY

Hacette University Faculty of Science

Çağlar UZUN

Ankara Üniversitesi Tıp Fakültesi: Ankara Üniversitesi Tıp Fakültesi

Research

Keywords: Radiomics, machine learning, thymoma, minimally invasive surgery, diagnostic tool

Posted Date: January 5th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-139108/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Introduction

Radiomics has become a hot issue in the medical imaging field, particularly in cancer imaging. Radiomics methods are used to analyze various medical images, including computed tomography (CT), magnetic resonance, and positron emission tomography to provide information regarding the diagnosis, patient outcome, tumor phenotype, and the gene-protein signatures of various diseases.

This study evaluated the CT radiomics features of thymomas to discriminate between low- and high-risk thymoma groups.

Materials and Methods

In total, 83 patients with thymoma were included in this study between 2004 and 2019. We used the Radcloud platform (Huiying Medical Technology Co., Ltd.) to manage the imaging and clinical data and perform the radiomics statistical analysis. The training and validation datasets were separated by a random method with a ratio of 2:8 and 502 random seeds. The histopathological diagnosis was noted from the pathology report.

Results

Four machine learning radiomics features were identified to differentiate a low-risk thymoma group from a high-risk thymoma group. The radiomics feature names were Energy, Zone Entropy, Long Run Low Gray Level Emphasis, and Large Dependence Low Gray Level Emphasis.

Conclusions

The results demonstrated that a machine-learning model and a multilayer perceptron classifier analysis can be used on CT images to predict low- and high-risk thymomas. This combination could be a useful preoperative method to determine the surgical approach for thymoma.

Introduction

Radiomics is a rapidly growing field of mapping digital medical images to quantitative data, with the end goal of generating imaging biomarkers as clinical decision-making support tools. The fundamental premise of radiomics is that radiological images contain biological, prognostic, and predictive knowledge that is not revealed during visual inspection; thus, converting medical radiological images into high-dimensional data and the subsequent quantitative analysis of these data supports decision-making guidelines used in clinical practice [1–3]. Radiomics is intended to predict patient-specific results based on high throughput analysis and mining of sophisticated imaging biomarkers through machine learning algorithms.

Thymic epithelial tumors (TETs) originate from the thymus and consist of thymomas, thymic carcinomas, and thymic neuroendocrine tumors. Although rare (1.5 cases/million), thymomas are common primary tumors of the anterior mediastinum in adults. Thymic carcinomas are very rare. Thymic carcinomas often present with metastasis and involve a poorer prognosis compared to thymomas, so TETs are a heterogeneous group [4, 5]. Despite the good survival rate of patients with thymoma, the histological subtype affects tumor behavior and the prognosis. Thymomas are subdivided into type A, AB, B1, B2, and B3 according to the World Health Organization (WHO) classification. Thymomas can also be subdivided depending on the prognosis into low-risk (type A, AB, and B1) and high-risk (type B2 and B3) groups. The high-risk group of thymomas is more likely to invade locally than the low-risk group. Surgery is the main strategy to treat thymomas, and complete resection provides the best survival rate, but the thymoma subgroups are an important factor in determining the treatment approach. The possibility of complete surgical resection is very high in the low-risk group, and this is typically an adequate treatment without adjuvant or neoadjuvant chemotherapy. In contrast, the high-risk group of thymomas has less of an opportunity for complete surgical resection than the low-risk group and may require multimodal therapy. Histological classification can inform risk stratification for patients and personalize the surgical treatment course [6–10].

Robotic-assisted thoracoscopic surgery and video-assisted thoracoscopic surgery are considered minimally invasive surgeries (MISs) and have recently become common for many surgeries, including surgery for thymic neoplasms. Some studies have reported that the results obtained from MIS are equivalent to the median sternotomy approach in the surgical treatment of thymoma. However, the indication for MIS in thymoma surgery is controversial. Many surgeons, particularly those with open surgery experience, are reluctant to perform MIS to treat TETs because MIS techniques are associated with increased manipulation of the tumor during surgery and a corresponding risk for capsular disruption, pleural seeding with tumor fragments, and incomplete resection, particularly for thymic carcinomas and high-risk thymomas. Thus, MIS may cause higher local recurrence rates and lead to lower overall survival rates. The possibility of local recurrence and spread is less likely for low-risk thymomas, so MIS is an acceptable surgical treatment approach for these cases [11]. Knowing whether a thymoma is at high or low risk preoperatively would help inform the choice of surgical approach.

Contrast-enhanced chest computed tomography (CT) is the most common imaging modality to preoperatively assess thymomas. The value of qualitative CT features in determining thymic carcinoma versus thymoma or low- versus high-risk remains unclear [12, 13]. Also, both CT and magnetic resonance imaging (MRI) have limited value for predicting the histologic subtype of thymoma [14, 15]. Preoperative prediction of the histological subtype of thymoma may facilitate patient management. A pre-surgical needle biopsy is a reliable method for diagnosing thymoma, but a small biopsy sample may not always represent the entire tumor; a deep biopsy is an invasive procedure with a risk of complications, and a transpleural biopsy may cause tumor seeding [16, 17]. Also, CT-guided biopsy and a histopathological evaluation of specimens are expensive, destroy tissues, and take 10–14 days. Overall, no clear non-invasive preoperative criteria have been defined to help surgeons choose either an open thymectomy or a

minimally invasive approach. Therefore, an effective and objective surgical approach to preoperatively determine the thymoma subtype would be useful.

This study analyzed the textural features of thymomas using CT radiomics features to discriminate low-versus high-risk thymoma groups in a single center.

Materials And Methods

The study protocol was approved by the Institutional Review Board of Ankara University, Faculty of Medicine (IRB no: I7-426-20). The need for informed consent was waived because the study had a retrospective design. The initial study population included 221 consecutive patients who underwent surgical resection or biopsy between 2004 and 2019 in the Thoracic Surgery Department and were diagnosed with a thymic epithelial tumor in the Department of Pathology of Ankara University Medical Faculty. Of these patients, 158 were diagnosed with thymoma. The following inclusion criteria were applied: (1) histopathologically confirmed thymoma; (2) CT images since 2011, which was the beginning of the Ankara University Faculty of Medicine electronic database archive; (3) contrast-enhanced CT performed within 4 weeks before surgery or biopsy; (4) no history of resection for a thymic neoplasm or another malignant tumor; and (5) no history of chemotherapy or radiotherapy before the primary thoracic malignancy. After applying these selection criteria, 83 patients were included (Table 1).

Table 1
Clinical characteristics of the patients with the low-risk group and high-risk group

Patients (n = 83)	Low-risk group, n (%)	High-risk group, n (%)	P- value*
	51 (61)	32 (39)	
Sex, n (%)	25 (49)	20 (62)	0.23
Male	26 (51)	12 (38)	
Female			
Age, median (range) (years)	50 (21–73)	50 (24–66)	0.68
Smoking status, n (%)	34 (67)	18 (56)	0.34
Never	17 (33)	14 (44)	
Current or past smoker			
Previous malignancy or synchronous tumor, n (%)	48 (94)	27(52)	0.25
Absent	3 (6)	5(48)	
Present			
Clinical presentation, n (%)	16 (31)	10 (31)	0.99
Asymptomatic	35 (69)	22 (69)	
Symptomatic			
Myasthenia gravis, n (%)	43 (84)	22 (69)	0.09
Absent	8 (16)	10 (31)	
Present			
LDH, median (range) (U/L)	180 (113–338)	179 (89–310)	0.80
ALP, median (range) (U/L)	68 (29–124)	67 (31–167)	0.74
CRP, median (range) (mg/L)	4.1 (0.6–70.7)	3.5 (0.1–25.8)	0.67
HGB, median (range) (g/dL)	14 (9-16.6)	13.6 (9-16.3)	0.62
WBC count, median (range) (x10 ⁹ /L)	7.7 (3.3–16.8)	7.3 (3.9–13.9)	0.97

LDH: Lactate dehydrogenase, ALP: Alkaline phosphatase, CRP: C-reactive protein, HGB: Hemoglobin, WBC: White blood cell, LYMP: Lymphocyte, PLT: Platelet

*Differences were compared using the t-test/Mann-Whitney U test, Pearson Chi-square test/Fisher's Exact test.

Patients (n = 83)	Low-risk group, n (%)	High-risk group, n (%)	P-value*
	51 (61)	32 (39)	
LYMP count, median (range) (x10 ⁹ /L)	2.1 (0.01–6.1)	1.8 (0.5–3.8)	0.10
PLT count, median (range) (x10 ⁹ /L)	258 (121–454)	233 (122–404)	0.58
Type of treatment, n (%)	44 (86)	28 (87)	1.00
Surgery ± adjuvant treatment	7 (14)	4 (13)	
Definitive chemotherapy ± radiotherapy			
Largest dimension of tumor size on CT (mean ± SD) (range) in mm	66.8 ± 2.1(21–160)	63 ± 2.2 (20–134)	0.55
LDH: Lactate dehydrogenase, ALP: Alkaline phosphatase, CRP: C-reactive protein, HGB: Hemoglobin, WBC: White blood cell, LYMP: Lymphocyte, PLT: Platelet			
*Differences were compared using the t-test/Mann-Whitney U test, Pearson Chi-square test/Fisher's Exact test.			

Ct Protocol And Lesion Segmentation

All patients underwent contrast-enhanced CT before biopsy and/or surgery to evaluate suspected mediastinal tumors. Chest CT examinations were performed with either 320-row detector CT (Toshiba Aquilion ONE, Otawara-shi, Japan), 64-row detector CT (Toshiba Aquilion 64, Otawara-shi, Japan), or 16-row detector CT (Siemens Somatom Sensation 16, Forchheim, Germany) scanners. The acquisition parameters were: 0.5 mm, 0.5 mm, and 0.625 mm detector collimation, 120 kVp tube voltage, 0.5 s gantry rotation time, 1 mm, 1 mm, and 1.5 mm reconstructed section thickness, and 0.8 mm, 0.8 mm, and 1 mm reconstruction intervals. All examinations were performed after injecting a nonionic intravenous contrast agent via the antecubital vein. The area from the thoracic inlet caudally to include the adrenal glands was scanned. All images were reviewed by a senior radiologist (Ç.U) with more than 10 years of experience in thoracic imaging. She was blinded to the histopathological data to avoid bias. Multiplanar reformatted images were analyzed on a workstation (GE Healthcare, Waukesha, WI, USA).

Patients And Dataset Management

Of the 83 patients included in analyses, 45 were male and 38 were female; the mean age was 49 ± 13.32 years (range 20–74 years). We used the Radcloud platform (Huiying Medical Technology Co., Ltd) to evaluate the data, and performed a radiomics statistical analysis. The training and validation datasets were separated by a random method 2:8 ratio with 502 random seeds. The histopathological diagnosis was noted from the pathology report.

Image Segmentation

Images were evaluated by two senior observers with 10 and 5 years of experience, respectively, in mediastinal surgery, and all disease lesions (VOIs) were annotated manually by observers who were blinded to the histopathological diagnoses of patients. Then, all images were re-evaluated by the senior radiologist. When there was $\geq 5\%$ discrepancy, the radiologist made the final decision on the tumor borders. Eventually, 79 VOIs were segmented from the scans of the 83 patients and used for subject analysis.

Feature Extraction

In total, 1409 features were extracted from the CT images using the Radcloud platform. These features were classified into three groups. Table 2 lists the details of the groups: **Group 1** (first-order statistics-126 descriptors); **Group 2** (shape- and size-based features-14 features); **Group 3** (textural features-525 textural) (Table 2).

Table 2
Radiomic features selected for quantifying the heterogeneity differences

Radiomic group	Associated filter	Number of features	Radiomic features
First-order statistics	None	126	Energy, total energy, entropy, minimum, 10 percentile, 90 percentile, maximum, mean, median, interquartile range, range, mean absolute deviation, robust mean absolute deviation, root mean square, standard deviation, skewness, kurtosis, variance
Shape	None	14	Volume, surface area, surface volume ratio, spherical disproportion, maximum 3D diameter, maximum 2D diameter column, maximum 2D diameter row, elongation
Texture features	GLCM	525	Autocorrelation, average intensity, cluster prominence, cluster shade, cluster tendency, contrast, difference average, difference entropy, difference variance, dissimilarity, entropy, sum average, sum entropy, sum variance, sum squares
Texture features	GLSZM		Large area emphasis, gray level non uniformity, size zone non uniformity, gray level variance, zone entropy, high gray level zone emphasis, small area high gray level emphasis,, large area high gray level emphasis
Texture features	GLRLM		Gray level non uniformity, run length non uniformity, gray level variance, run entropy, high gray level run emphasis, short run high gray level emphasis, long run high level emphasis
GLCM: Gray-level Co-occurrence Matrix, GLSZM: Gray-Level Size Zone Matrix, GLRLM: Gray Level Run Length Matrix			

Qualification

A large number of image features were measured. A dimensionality reduction was performed and task-specific features were selected to identify the appropriate features. To reduce the redundant features, selection methods included the variance threshold (variance threshold = 0.8), SelectKBest, and the least absolute shrinkage and selection operator (LASSO), which were used to detect significant differences between low- and high-risk groups. Eigenvalues of the variance < 0.8 were removed. The SelectKBest method was used with a p-value to analyze the relationship between the features and the classification results. All features with p-values < 0.05 were used. The L1 regularizer was used as the cost function in the LASSO model; the error value of the cross-validation was 5, and the maximum number of iterations was 1000.

Statistical analysis

Statistical analyses were performed in the Radcloud platform. The 1409 features identified were significantly correlated. The radiomics-based models were constructed with six classifiers: k-nearest neighbor (KNN), support vector machine (SVM), eXtreme Gradient Boosting (XGBoost), random forest (RF), logistic regression (LR), and decision tree (DT), and the validation method was used to improve the effectiveness of the model.

The following parameters were applied. For KNN: n_neighbors (5), weights (uniform). For SVM: kernel (rbf), C (1), gamma (auto), class_weight (balanced), decision_function_shape (ovr), random_state. For XGBoost: Eta (0.3), max_depth (6). For RF: n_estimators (10), class_weight (None). For LR: penalty (L2), C(1), solver (liblinear), class_weight (None), multi_class (ovr), random_state. For DT: splitter (best), criterion (gini).

The receiver operating characteristic (ROC) curve and the area under the curve (AUC) were used to assess the predictive performance of the training and validation datasets, respectively. The four indicators were P (precision = true positives/(true positives + false positives)), R (recall = true positives/(true positives + false negatives)), f1-score (f1-score = $P \times R \times 2 / (P + R)$), and support (total number in test set) to evaluate the performance of the classifier.

Results

Table 1 lists patient characteristics. The low-risk group included 51 patients (type A: 10, type AB: 19, type B1: 22) and the high-risk group included 32 patients (type B2: 14, type B3: 18). No significant differences were observed between the low-risk and high-risk groups in either the training cohort (n = 66 cases) or the validation cohort (n = 17) (Table 1).

First, 459 features were selected from the 1409 total features using the variance threshold method, and 30 features were determined with the SelectKBest method (Fig. 1). Finally, four optimal features were defined with the Lasso algorithm (Fig. 2).

Figure 3 presents the ROC curve analysis results for the training and validation datasets to differentiate between the low-risk and high-risk groups of thymomas.

The AUC of XGBoost, RF, and DT machine learning methods was the highest at 0.998–1 for the training data while KNN and LR were highest for the validation data. Table 3 lists the results of the machine learning classifiers of the validation set. The KNN scores were (AUC = 0.943) for the low-risk group and (AUC = 0.943) for the high-risk group. The LR scores were the same (AUC = 0.943) for both groups of thymomas. The KNN and LR classifiers were the best methods in the validation dataset in terms of differentiating between the low- and high-risk groups. Table 4 lists the diagnostic performance according to the four indicators. The ranges for the low-risk group were: precision (0.5–0.9), recall (0.3–0.9), F1-score (0.37–0.9), and support (10), while the ranges for the high-risk group were: precision (0.36–0.86), recall (0.57–0.86), F1-score (0.44–0.86), and support (7). The highest scores were achieved with the KNN machine learning method.

Table 3
ROC results with six machine learning classifiers of validation set

Risk Groups	Statistical measures	KNN	SVM	XGB oost	RF	LR	DT
Low	AUC	0.943	0.857	0.8	0.693	0.943	0.436
	95% CI	0.85–1	0.66–1	0.58–1	0.45–0.93	0.74–1	0.18–0.69
	Sensitivity	0.9	0.8	0.7	0.8	0.8	0.3
	Specificity	0.86	0.86	0.71	0.43	0.86	0.57
High	AUC	0.943	0.857	0.8	0.693	0.943	0.436
	95% CI	0.85–1	0.66–1	0.58–1	0.45–0.93	0.74–1.00	0.18–0.69
	Sensitivity	0.86	0.86	0.71	0.43	0.86	0.57
	Specificity	0.9	0.8	0.7	0.8	0.8	0.3

KNN: k-nearest neighbor, SVM: Support Vector Machine, XGBoost: eXtreme Gradient Boosting, RF: Random Forest, LR: Logistic Regression, DT: Decision tree

Table 4
The results of four indicators- Precision, Recall, F1-score, Support in validation set

Risk groups	Indicators	KNN	SVM	XGBoost	RF	LR	DT
Low	Precision	0.9	0.89	0.78	0.67	0.89	0.5
	Recall	0.9	0.8	0.7	0.8	0.8	0.3
	F1-score	0.9	0.84	0.74	0.73	0.84	0.37
	Support	10	10	10	10	10	10
High	Precision	0.86	0.75	0.62	0.6	0.75	0.36
	Recall	0.86	0.86	0.71	0.43	0.86	0.57
	F1-score	0.86	0.8	0.67	0.5	0.8	0.44
	Support	7	7	7	7	7	7

KNN: k-nearest neighbor, SVM: Support Vector Machine, XGBoost: eXtreme Gradient Boosting, RF: Random Forest, LR: Logistic Regression, DT: Decision tree

Four radiomics features were identified that differentiated the low-risk group from the high-risk group of thymomas using machine learning, including Energy, Zone Entropy, Long Run Low Gray Level Emphasis, and Large Dependence Low Gray Level Emphasis (Fig. 2).

Table 5 lists the details of the confusion matrix in the low-risk and high-risk thymoma groups using the best MLP learning classifier (KNN).

Table 5
The details of Confusion Matrix in low-risk and high-risk thymoma groups

KNN			
Risk groups	0	1	Accuracy (%)
Low	9	1	100
High	1	6	88
Accuracy (%)			94.3

KNN: k-nearest neighbor,

Discussion

Radiomics has the potential to detect specific characteristics of a disease that cannot be visualized by current medical imaging modalities by quantitatively analyzing digital images. Recent studies have reported promising radiomics results in oncological practice. This method may supplement traditional

imaging analysis and assist in providing personalized medicine for patients. Publications on applications of thoracic tumors have increased in recent years. In the present study, a radiomics platform was used to analyze both imaging and clinical data, and to perform a statistical analysis. Radiomics platforms have the potential to reveal distinct imaging algorithms that can be used to quantify the status of a disease, providing valuable knowledge for personalized medicine. They can also measure features in an imaging examination; shape, intensity, texture, wavelet and Laplacian of Gaussian (LoG) features can be used to build predictive or prognostic non-invasive biomarkers or imaging modalities [18]. This kind of platform can be used to extract radiomics features from two-dimensional (2D) and/or three dimensional (3D) images and dual masks on different imaging modalities, such as CT and MRI, which is why it was preferred for this study.

Thoracic oncology surgical information obtained from standard imaging modalities such as CT, MRI, and positron emission tomography scans usually refers to simple traits, such as gross shape, contrast enhancement, and size. However, imaging information is now much richer, and increased resolution quality has led to 3D image acquisitions containing millions of voxels available for analysis, making the development of radiomics a natural progression. Soon, data obtained from radiomics studies will be used to inform the diagnosis and treatment algorithms of thoracic malignancies.

Radiomics models have already been used to stage tumors and predict lymph node metastasis and prognosis [19–22]. A few studies have used radiomics models to predict the pathological invasiveness of TETs [23]. Although some previous studies have demonstrated that a textural analysis based on CT images can be used to differentiate high-risk TETs from low-risk TETs, they only analyzed 2D textural features, and their sample sizes were small [24, 25]. TETs are a heterogeneous group with different radiological appearances, histopathological features, and prognoses. They include thymomas, thymic carcinomas, and thymic neuroendocrine tumors, with a wide variety of histological features. Thymomas are the most common TETs and are subdivided into five groups (A, AB, B1, B2, and B3). These can resemble spindle cell tumors, lymphomas, or carcinomas depending on the tumor type. Interestingly, some recent studies have included thymic carcinomas as type C thymomas, which has not been used in the WHO classification since 2004, although thymic carcinomas have a heterogeneous morphology. Type A, AB, and B1 thymomas have a thymus-like architecture, whereas thymic carcinomas exhibit features that are encountered by other organs and are a heterogeneous group of tumors, such as squamous cell carcinoma, adenocarcinoma, and undifferentiated carcinoma [8]. For these reasons, thymic carcinomas and thymomas should not be analyzed together in a radiomics study.

CT and MRI are common imaging modalities to preoperatively assess thymomas. However, they have limited value for predicting the histological subtypes of TETs [14]. Jeong et al. reported that the contour of the tumor, mediastinal fat, and large vessel invasion are useful CT features to distinguish between the WHO classification subgroups [12]. CT and/or MR imaging findings can help differentiate between low-risk and high-risk thymomas among thymic carcinomas, but they are insufficient to distinguish between the different histological subtypes of the WHO thymoma classification.

Few studies have used a machine learning system as an artificial intelligence approach with the application of radiomics features to analyze thymomas. This study addressed this research gap by developing a radiomics-based model incorporating machine learning to predict low- and high-risk thymomas.

Different machine learning strategies, such as KNN, SVM, or RF decision trees, can be applied to construct the map of a given training set and a given set of features. During training, the parameters that define the mapping (whose representation depends on the chosen learning strategy) are iteratively refined such that estimation performance is maximized on the training set itself. Then, the difference between the given 'ground truth' for each image and in the training set can be evaluated. The KNN classifier is a popular image classification algorithm that directly calculates image-to-image distances compared with other classifiers that need a training phase to calculate the distance between an image and a class [26]. RFs or random decision forests are ensemble learning methods for classification, regression, and other tasks that operate by constructing a multitude of decision trees at training time and outputting mode (classification) of the class or the mean prediction (regression) of the individual trees [27].

Wang et al. developed and compared the performance of radiomics signatures using textural features extracted from non-contrast-enhanced CT and contrast-enhanced CT scans [28]. They found that radiomics signatures performed better than radiologists with a high AUC, and that radiomics signatures based on a textural analysis extracted from a CT scan can be utilized as noninvasive biomarkers to differentiate high-risk thymomas from low-risk thymomas and advanced stage thymomas from early-stage thymomas. They concluded that as a quantitative method, a radiomics signature provides complementary diagnostic information and informs plans for personalized treatment for patients with thymomas. Yang et al. studied a preoperative staging tool that differentiates Masaoka-Koga (MK) stage I patients from stage II patients using CT images of thymoma patients [29]. They used an artificial neural network (ANN) deep learning model, namely the 3D-DenseNet model, to distinguish the MK stage I and stage II thymomas. They found that deep learning has great potential to preoperatively stage thymomas, which dramatically improves identification between MK stage I and stage II thymomas compared with visual observations. They concluded that deep learning models can help guide surgical treatment and improve outcomes compared to traditional methods. Our findings are consistent with the previous results: a deep learning-supported radiomics model, such as an ANN, can help distinguish between low- and high-risk group thymomas. Similar findings have been reported elsewhere [30, 31]. A previous study detected correlations between preoperative CT imaging features and the biological behavior of thymomas [32]. Similarly, we found that myasthenia gravis, lactate dehydrogenase, and the largest tumor dimension size on CT (mm) were predictors of the prognosis. A previous retrospective study developed a radiomics model using LR analysis and realized high diagnostic performance [25]; it reported that the AUCs for differentiating high-risk thymomas from low-risk thymomas were 0.89 for mean0c and 0.87 for a combination of mean0u and entropy. Similarly, we found that the AUC of the radiomics signatures was 0.943 for KNN, the best MLP classifier.

The International Association for the Study of Lung Cancer and the International Thymic Malignancy Interest Group concluded that the WHO histological classification, the completeness of tumor resection, the MK stage, and the 8th edition of the TNM staging system are independent prognostic factors for TETs [33–35]. Similar relationships have been reported between thymomas and histological classification, completeness of tumor resection, and staging, but these relationships are not as strong as in some solid tumors, so the optimal staging system for TETs has not been defined. Therefore, histological classification and resectability are more useful determinants than staging systems, for both treatment decisions and predicting prognosis. An important element in thymoma surgical treatment planning is knowing the preoperative risk group: whether the thymoma is in the low- or high-risk group may affect decisions about the surgical approach, which is one of the main determinants of the completeness of the resection. The low-risk group of thymomas is more likely to achieve complete resection with a MIS method; this may be less possible in the high-risk group. In the present study, we predicted thymoma risk groups by combining clinical and specific CT-based radiomics features with image variables, and this distinction may inform surgical treatment planning for thymomas. The most important advantage of this method is that it does not require a biopsy, which wastes time, and is costly, and can lead to complications.

Our study had some limitations. First, it was a retrospective study of thymomas from a single center, which may have caused selection bias. Second, it had a small sample size. A multicenter study with a larger sample size will be required to validate these results.

Conclusion

The results of this study demonstrated that a machine learning model and MLP classifier analysis can be used with CT images to predict low-risk and high-risk thymomas. The results also demonstrated that the combination of clinical and specific CT-based radiomics features and image variables can be used to predict thymoma risk groups. This method can be used as a preoperative technique to inform decisions about surgical approaches for treating thymoma.

Abbreviations

ANN

Artificial neural network

AUC

Area under curve

CT

Computed tomography

DT

Decision tree

IASLC

International Association for the Study of Lung Cancer

ITMIG
International Thymic Malignancy Interest Group
KNN
k-nearest neighbor
MG
Myasthenia gravis
MIS
Minimally invasive surgery
MK
Masaoka-Koga staging system
MLP
Multilayer perceptron
MRI
Magnetic resonance imaging
LASSO
Least absolute shrinkage and selection operator
LDH
Lactate dehydrogenase
LoG
Laplacian of Gaussian
LR
Logistic Regression
P
Precision
PET
Positron emission tomography scan
R
Recall
RATS
Robotic-assisted thoracoscopic surgery
RF
Random Forest
ROC
Receiver operating characteristic
SVM
Support Vector Machine
TETs
Thymic epithelial tumors
VATS
Video-assisted thoracoscopic surgery

VOI
Volume of interest
WHO
World Health Organization
XGBoost
eXtreme Gradient Boosting
2D
Two dimensional
3D
Three dimensional

Declarations

Disclosure of Funding: This research did not receive any specific grant funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of Interest Statement: The authors declare that they have no conflict of interest.

Ethics approval and Consent to participate: All procedures performed in studies involving human participants were in accordance with the ethical standards of the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards. For this type of study, ethical committee approval and formal consent is not required. Informed consent was obtained from all individual participants included in the study.

Consent for publication: Not applicable

Availability of data and material: The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Code availability: None.

Authors' contributions: PET/CT interpretation: AKC, KO, ÇU; Data documentation; YK, HÖ, BBK, BMKB; Data analysis and statistical analysis; KO, DK, Writing of the manuscript; AKC, KO, YK, HÖ; Supervising: AKC, ÇU

Article type: Original Article (Clinical Original)

References

1. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. 2016 Feb;278(2):563–77.
2. Kumar V, Gu Y, Basu S, et al. QIN Radiomics: the process and the challenges. *Magn Reson Imaging*. 2012 Nov;30(9):1234–48.

3. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. 2012 Mar;48(4):441–6.
4. National Comprehensive Cancer Network. NCCN Clinical practice guidelines in oncology. Thymomas and thymic carcinomas. Version 1.2020-November 27, 2019. https://www.nccn.org/professionals/physician_gls/pdf/thymic.pdf; 2019 Accessed 03 February 2020.
5. Ruffini E, Detterbeck F, Van Raemdonck D, et al. European Association of Thoracic Surgeons (ESTS) Thymic Working Group. Tumours of the thymus: a cohort study of prognostic factors from the European Society of Thoracic Surgeons database. *Eur J Cardiothorac Surg*. 2014 Sep;46(3):361–8.
6. Giaccone G, Wilmink H, Paul MA, et al. Systemic treatment of malignant thymoma. *Am J Clin Oncol*. 2006 Aug;29(4):336–44.
7. Okumura M, Miyoshi S, Fujii Y, et al. Clinical and functional significance of WHO classification on human thymic epithelial neoplasms: a study of 146 consecutive tumors. *Am J Surg Pathol*. 2001 Jan;25(1):103–10.
8. Marx A, Chan JK, Coindre JM, et al. The 2015 World Health Organization classification of tumors of the thymus: continuity and changes. *J Thorac Oncol*. 2015 Oct;10(10):1383–95.
9. Roden AC, Yi ES, Jenkins SM, et al. Reproducibility of 3 histologic classifications and 3 staging systems for thymic epithelial neoplasms and its effect on prognosis. *Am J Surg Pathol*. 2015 Apr;39(4):427–41.
10. Scorsetti M, Leo F, Trama A, et al. Thymoma and thymic carcinomas. *Crit Rev Oncol Hematol*. 2016 Mar;99:332–50.
11. Zhang X, Gu Z, Fang W, et al. **Minimally invasive surgery in thymic malignances: the new standard of care.** *J Thorac Dis*. 2018 Jun; 10(Suppl 14): S1666–S1670.
12. Jeong YL, Lee KS, Kim J, et al. Does CT of thymic epithelial tumors enable us to differentiate histologic subtypes and predict prognosis? *Am J Roentgenol*. 2004;183:283–9.
13. Yakushiji S, Tateishi U, Nagai S, et al. Computed tomographic findings and prognosis in thymic epithelial tumor patients. *J Comput Assist Tomogr Sep-Oct*. 2008;32(5):799–805.
14. 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 Oct;60(1):70–9.
15. Marom EM. Advances in thymoma imaging. *J Thorac Imaging*. 2013;28:69–83.
16. Fujii Y. Published guidelines for management of thymoma. *Thorac Surg Clin*. 2011 Feb;21(1):125–9.
17. Falkson CB, Bezjak A, Darling G, et al. The management of thymoma: a systematic review and practice guideline. *J Thorac Oncol*. 2009 Jul;4(7):911–9.
18. Bousquet O, Luxburg U, Rätsch G, editors. *Advanced Lectures on Machine Learning*. Berlin Heidelberg: Springer; 2003. pp. 21–40.

19. Huang Y, Liang C, He L, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. *Journal of Clinical Oncology*. 2016 Jun;34(18):2157–64.
20. Bianconi F, Fravolini ML, Bello-Cerezo R, et al. Evaluation of shape and textural features from CT as prognostic biomarkers in non-small cell lung cancer. *Anticancer Res*. 2018 April;38(4):2155–2160.
21. Junior JRF, Santos MK, Cipriano FG, et al. Radiomics-based features for pattern recognition of lung cancer histopathology and metastases. *Computer Methods Programs in Biomedicine*. 2018;159:23–30.
22. Xu X, Zhang X, Tian Q, et al. Quantitative identification of nonmuscle-invasive and muscle-invasive bladder carcinomas: a multiparametric MRI radiomics analysis. *J Magn Reson Imaging*. 2019 May;49(5):1489–98.
23. 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 Apr, **43(4): 1256–1266.**
24. Iannarelli A, Sacconi B, Tomei F, et al. Analysis of CT features and quantitative texture analysis in patients with thymic tumors: correlation with grading and staging. *Radiol Med (Torino)*. 2018;123:345–50.
25. Yasaka K, Akai H, Nojima M, et al. Quantitative computed tomography texture analysis for estimating histological subtypes of thymic epithelial tumors. *Eur J Radiol*. 2017 Jul;92:84–92.
26. Prince SJD. *Computer Vision: Models, Learning, and Inference.* **Cambridge:** Cambridge University Press; 2012.
27. Ho TK. The random subspace method for constructing decision forests. *IEEE Trans Pattern Anal Mach Intell*. 1998;20:832–44.
28. 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.**
29. Yang L, Cai W, Yang X, et al. Development of a deep learning model for classifying thymoma as Masaoka-Koga stage I or II via preoperative CT images. *Ann Transl Med*. 2020 Mar;8(6):287.
30. Zhao Y, Chen H, Shi J, et al. The correlation of morphological features of chest computed tomographic scans with clinical characteristics of thymoma. *Eur J Cardiothorac Surg*. 2015 Nov;48(5):698–704.
31. Ried M, Marx A, Götz A, et al. State of the art: diagnostic tools and innovative therapies for treatment of advanced thymoma and thymic carcinoma. *Eur J Cardiothorac Surg*. 2016 Jun;49(6):1545–52.
32. Ozawa Y, Hara M, Shimohira M, et al. Associations between computed tomography features of thymomas and their pathological classification. *Acta Radiol*. 2016 Nov;57(11):1318–25.
33. Weis CA, Yao X, Deng Y, et al. The impact of thymoma histotype on prognosis in a worldwide database. *J Thorac Oncol*. 2015 Feb;10(2):367–72.
34. Safieddine N, Liu G, Cuningham K, et al. Prognostic factors for cure, recurrence and long-term survival after surgical resection of thymoma. *J Thorac Oncol*. 2014 Jul;9(7):1018–22.

35. Detterbeck FC, Stratton K, Giroux D, et al. The IASLC/ITMIG Thymic Epithelial Tumors Staging Project: proposal for an evidence-based stage classification system for the forthcoming (8th) edition of the TNM classification of malignant tumors. J Thorac Oncol. 2014 Sep;9(9 Suppl 2):65–72.

Figures

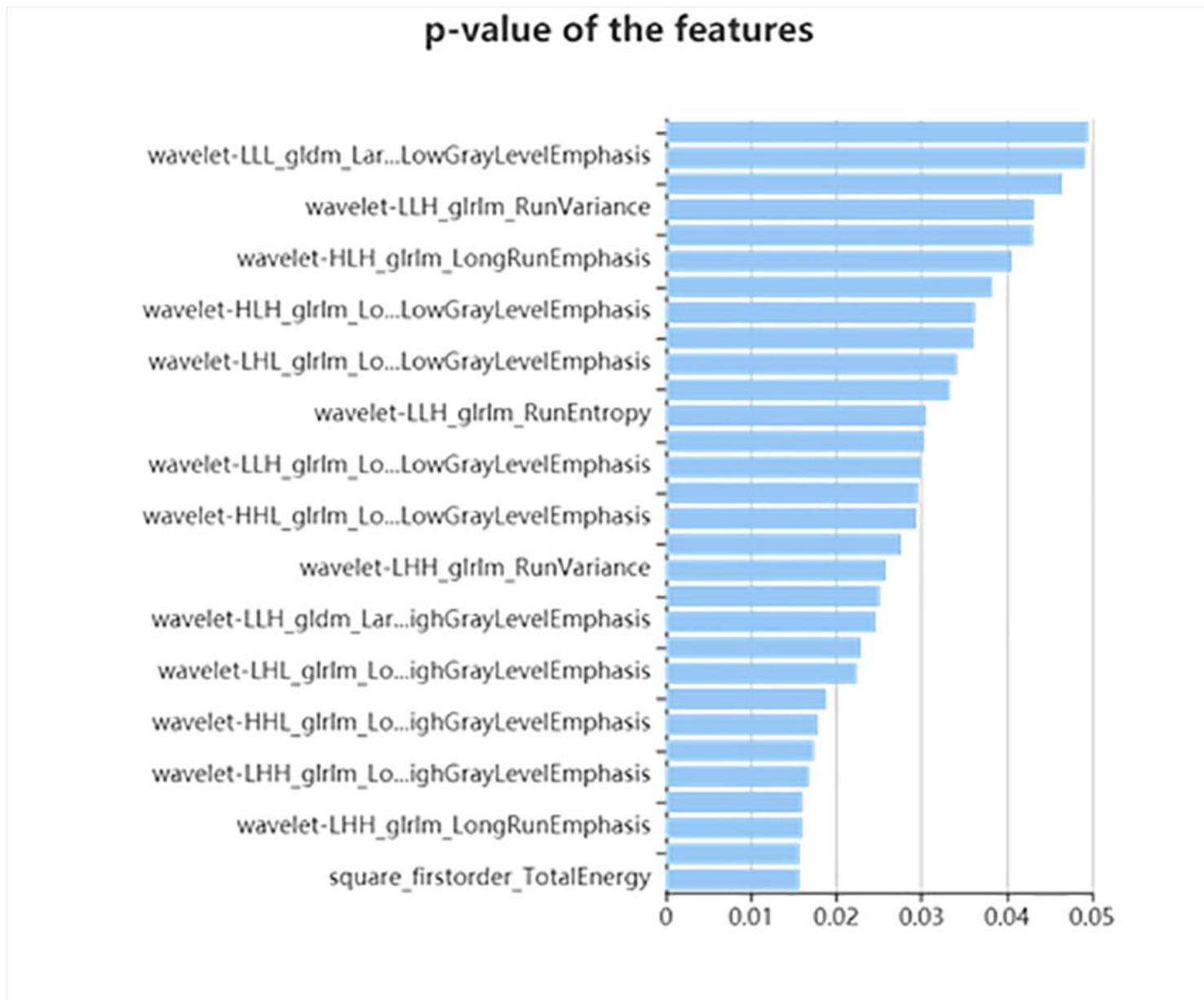


Figure 1

The Select K-best method was used to further select the radiomics features; 30 features were selected.

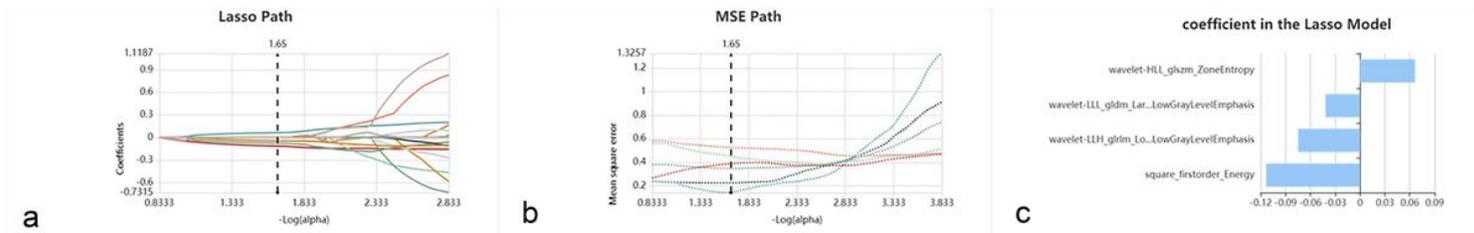


Figure 2

Lasso algorithm for feature selection. (a) Lasso path, (b) MSE path, (c) coefficients in the Lasso model. The Lasso model was used to select four features that correspond to the optimal alpha value.

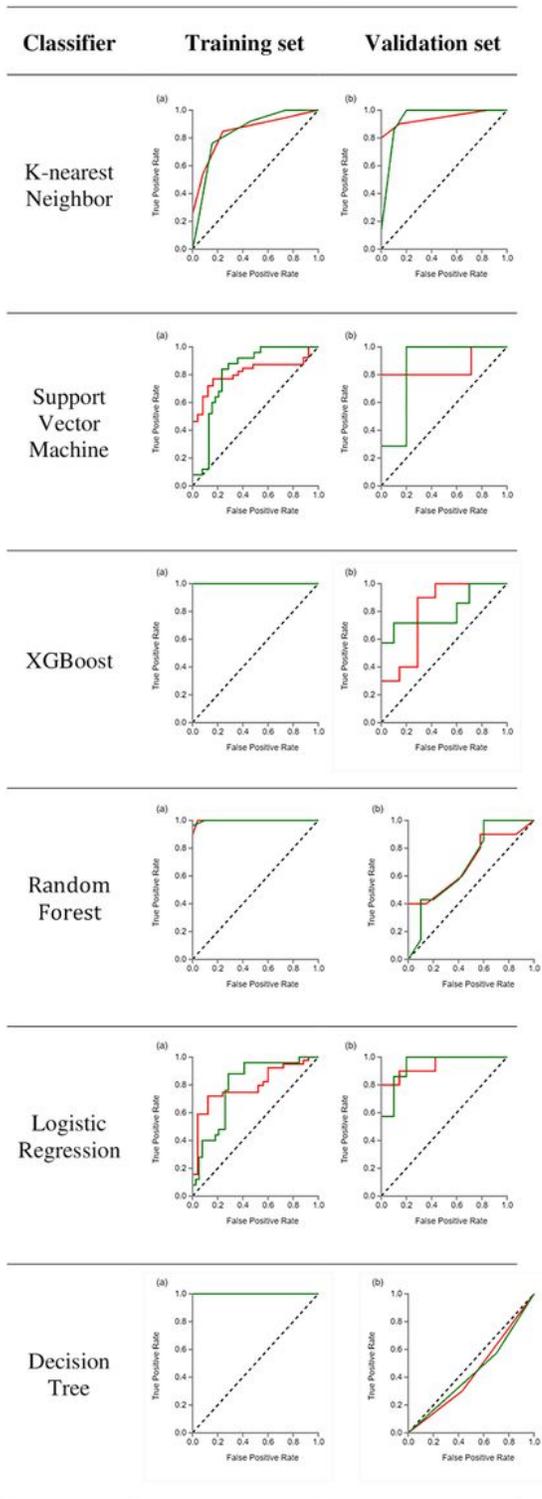


Figure 3

ROC curves of machine learning methods for classification. Green indicates low-risk and red indicates high-risk thymomas. (a) ROC curve of the training dataset, (b) ROC curve of the validation dataset.