Texture Analysis of Three-Dimensional MRI Images May Differentiate Borderline and Malignant Epithelial Ovarian Tumors

Rongping Ye, MD1*, Shuping Weng, MD2*, Yueming Li, MD1,3, Chuan Yan, MD1, Jianwei Chen, MD1, Yuemin Zhu, MD1, Liting Wen, MD1

1Department of Radiology, The First Affiliated Hospital of Fujian Medical University, Fuzhou, China; 2Department of Radiology, Fujian Maternity and Child Health Hospital, Fuzhou, China; 3Key Laboratory of Radiation Biology (Fujian Medical University), Fujian Province University, Fuzhou, China

Objective: To explore the value of magnetic resonance imaging (MRI)-based whole tumor texture analysis in differentiating borderline epithelial ovarian tumors (BEOTs) from FIGO stage I/II malignant epithelial ovarian tumors (MEOTs).

Materials and Methods: A total of 88 patients with histopathologically confirmed ovarian epithelial tumors after surgical resection, including 30 BEOT and 58 MEOT patients, were divided into a training group (n = 62) and a test group (n = 26). The clinical and conventional MRI features were retrospectively reviewed. The texture features of tumors, based on T2-weighted imaging, diffusion-weighted imaging, and contrast-enhanced T1-weighted imaging, were extracted using MaZda software and the three top weighted texture features were selected by using the Random Forest algorithm. A non-texture logistic regression model in the training group was built to include those clinical and conventional MRI variables with p value < 0.10. Subsequently, a combined model integrating non-texture information and texture features was built for the training group. The model, evaluated using patients in the training group, was then applied to patients in the test group. Finally, receiver operating characteristic (ROC) curves were used to assess the diagnostic performance of the models.

Results: The combined model showed superior performance in categorizing BEOTs and MEOTs (sensitivity, 92.5%; specificity, 86.4%; accuracy, 90.3%; area under the ROC curve [AUC], 0.962) than the non-texture model (sensitivity, 78.3%; specificity, 84.6%; accuracy, 82.3%; AUC, 0.818). The AUCs were statistically different (p value = 0.038). In the test group, the AUCs, sensitivity, specificity, and accuracy were 0.840, 73.3%, 90.1%, and 80.8% when the non-texture model was used and 0.896, 75.0%, 94.0%, and 88.5% when the combined model was used.

Conclusion: MRI-based texture features combined with clinical and conventional MRI features may assist in differentiating between BEOT and FIGO stage I/II MEOT patients.

Keywords: Borderline ovarian tumor; Malignant epithelial ovarian tumor; Magnetic resonance imaging; Texture analysis; Differential diagnosis

INTRODUCTION

Epithelial ovarian tumors are the most common type of ovarian tumors. They account for approximately 90% of all ovarian cancers and have been recognized as the leading cause of gynecological cancer death worldwide (1).
These tumors could be histologically classified into three groups, namely, benign (57–60%), borderline (4–15%), and malignant (21–33%) epithelial ovarian tumors (1). Borderline epithelial ovarian tumors (BEOTs) typically represent an intermediate grade of neoplasm between benign and malignant (1, 2) because of their specific histological features (low malignant potential and lack of stromal invasion, mitotic activity, and nuclear atypia). Due to their distinct therapeutic strategies and prognoses, it is of great importance to differentiate BEOTs from malignant epithelial ovarian tumors (MEOTs). Specifically, according to the National Comprehensive Cancer Network guidelines (Version 1.2020) (3), a conservative surgical approach with preserved fertility is recommended to treat young BEOT patients, whereas most patients with MEOTs should undergo radical surgery followed by adjuvant chemotherapy (3–5). Moreover, the prognosis of BEOT is far superior to MEOT, usually with a long-term (> 10 years) survival rate in 95% of cases (6, 7). Therefore, preoperative differential diagnosis of BEOTs and MEOTs may be crucial and helpful in determining treatment strategies and predicting the prognosis of patients with ovarian epithelial tumors.

In comparison to ultrasonography and computed tomography, magnetic resonance imaging (MRI) is a valuable tool for identifying the morphological characteristics of ovarian epithelial tumors (8). However, previous studies have revealed limitations of conventional MRI in distinguishing BEOT from MEOT, especially since these two types of tumors have very similar morphological features (9–12). Advanced MRI techniques, such as dynamic contrast-enhanced-MRI, diffusion-weighted imaging (DWI) with the measurement of the apparent diffusion coefficient (ADC), and enhanced T2*-weighted angioigraphy, have been previously used to distinguish between the two types of tumors (13–15). Most of these studies involved a pre-selected region of interest instead of the entire tumor. This technique limits the accuracy and heterogeneity of ovarian tumors due to its large size, rendering analysis susceptible to personal selection bias and sample bias.

Texture analysis is a quantitative technique that allows the evaluation of the gray-level intensity and the relationship among pixels, which are not discernible to human eye (16). Through the assessment of gray level distribution, coarseness, and regularity in a lesion, texture analysis may be able to evaluate tumor heterogeneity based on medical images obtained in daily clinical practice. Ng et al. (17) suggested that whole tumor texture analysis could provide a more representative evaluation of tumor heterogeneity than the largest cross-sectional area analysis. To date, no study has been performed to differentiate BEOTs from MEOTs using whole tumor texture analysis. Thus, this study aimed to explore the value of MR images-based whole tumor texture analysis in differentiating BEOTs from FIGO stage I/II MEOTs.

MATERIALS AND METHODS

Participants

Patients with known or suspected history of ovarian neoplasms between September 2014 and May 2019 were retrospectively reviewed in our hospital. Patients who met the following criteria were included in the study: 1) those who preoperatively underwent standard dedicated pelvic MRI including contrast-enhanced T1-weighted imaging (CE-T1WI) and DWI; 2) those with pathologically confirmed BEOTs or MEOTs after surgical resection; 3) those with FIGO stage I/II MEOTs (tumor limited to one or both ovaries or tumor involving one or both ovaries with pelvic extension below pelvic brim or primary peritoneal cancer); and 4) those in whom the serum cancer antigen-125 (CA-125) was measured preoperatively. Patients who met the following criteria were excluded: 1) those who underwent prior related treatment (n = 10); 2) those with poor image quality due to dielectric effect artifact (n = 1); and 3) those with ruptured cystic masses (n = 3). The clinical, MRI, and pathological data recorded were reviewed. Ultimately, a total of 88 consecutive patients were enrolled for analysis, including 30 BEOT and 58 MEOT patients. Patients were then randomly allocated to either the training (n = 62) or test (n = 26) group, with the distribution of BEOTs balanced between the two groups. A flow diagram showing the patient selection protocol and the inclusion and exclusion criteria is shown in Figure 1. In cases involving peritoneal carcinomatosis, which were confined to the pelvis (FIGO stage II), we evaluated the primary lesion. In bilateral lesions, lesions with a greater solid portion or a greater septa number were selected for analysis as these radiological features suggested a more aggressive histology.

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. This single-center retrospective cohort study was approved by the Institutional Review Board (No. [2016]118 and No. [2019]283), who determined the requirement for
informed consent could be waived.

The datasets generated and/or analyzed during the current study are not publicly available due to patient privacy protection, but are available from the corresponding author on reasonable request.

**MRI Technique**

All examinations were performed using a 3T MR scanner (MAGNETOM Verio, Siemens Healthineers). The standard dedicated pelvic MRI protocol consisted of the following sequences: transverse turbo spin echo (TSE)-T1WI, transverse and sagittal TSE-T2WI with fat-suppression, and DWI (b value = 50 and 800 s/mm²). For three phases, CE-T1WI and three-dimensional volumetric interpolated breath-hold examination with fat-suppression were performed in the transverse, sagittal, and coronal planes at 40–60 seconds, 90–110 seconds, and 170–190 seconds after intravenous injection of gadobenate dimeglumine (MultiHance, Bracco; 0.2 mmol/kg body weight; rate of 3.0 mL/s). The MRI protocol is shown in Table 1.

**Conventional MRI Features Analysis**

All MR images were reviewed independently by two radiologists (with 3 and 30 years of experience in abdominal MRI, respectively) who were blinded to the clinical information and the lesion type (either BEOTs or MEOTs). Similar to previous studies (18, 19), tumors were evaluated based on the following features: 1) morphology, classified as either lobulated (notches ≥ 2) or round/oval; 2) configuration, which was classified as either purely cystic, predominantly cystic (≤ 1/3 solid component), mixed cystic/solid (1/3 to 2/3 solid component), or predominantly-solid...
(≥ 2/3 solid component). Solid component was defined as a papillary projection (any enhancing solid projection into the cyst from the cyst wall with a height ≥ 3 mm) or a solid portion; 3) maximum diameter of lesion and solid component; 4) presence of thickened irregular septa defined as focal areas of septa thickening with a thickness ≥ 3 mm within a cyst (20); 5) number of loculi, classified as either unilocular, multilocular, or honeycomb loculi (grouped and innumerable sub-centimeter-sized locules) (21); 6) signal intensity, categorized as either mild, moderate, or high based on the signal of the myometrium in cystic components during T1WI and T2WI; 7) the presence hemorrhagic fluid in the loculi; 8) enhancement degree of the solid component or septa/wall, categorized as either mild, moderate, or significant based on the signal of the iliopsoas and myometrium; 9) the presence of ascites, graded as either limited to the Douglas pouch, moderate (limited to the pelvic cavity), or massive (beyond the pelvic); and 10) the presence of a peritoneal implant.

Image Segmentation

The sagittal T2WI, transverse DWI, and sagittal CE-T1WI images were retrieved from a picture archiving and communication system (PACS, Carestream) in Bitmap (BMP) format. The MR images were then loaded into the MaZda package for manual segmentation (open source software, http://www.eletel.pl.lodz.pl/mazda/) (22). This package included image processing, extraction, and selection of texture features. To minimize the influence of contrast and brightness variation, image gray-level intensity normalization was performed using a method that normalizes image intensities within a range (µ - 3δ, µ + 3δ; µ, mean gray level value; δ, standard deviation; both µ and δ are computed separately for every volume of interest [VOI]) in MaZda. The VOI, which encompassed the whole tumor was delineated on T2WI, DWI (b value = 800 s/mm²), and CE-T1WI images on each slice respectively segmented by the junior radiologist. ADC maps and DWI images with b values of 50 s/mm² were used for the segmentation of VOIs in DWI images with b values of 800 s/mm². The experienced radiologist validated all the segmented images. The procedure is presented in Figure 2.

Texture Feature Extraction and Selection

All 264 VOIs delineated by the junior radiologist were loaded into the MaZda package for texture analysis. The software automatically computed 314 texture features within each VOI, and generated a texture features analysis report. These texture features are shown in Table 2. Usually, only a limited number of features carry relevant information needed for texture discrimination. Therefore, texture feature reports for the training group (n = 186; 62 x 3) were loaded into the MaZda package to identify the most discriminative features for the classification of BEOTs and MEOTs. The MaZda package uses three feature selection methods: Fisher’s coefficient (Fisher), classification error probability combined with average correlation coefficient (POE + ACC), and mutual information (MI). These are supervised methods (23-25). To perform feature selection based on MaZda’s automatic techniques, we used a combination of Fisher, POE + ACC, and MI. Texture features with the lowest POE + ACC coefficients, the highest Fisher coefficients, and the highest individual MI were ultimately selected. Finally, thirty optimal discriminative power features from each sequence (T2WI, DWI, and CE-T1WI) of the training group were selected.

In order to solve the multicollinearity problem, the Random Forest Algorithm based on the Python Scikit-learn Library was used to select the top feature from the thirty optimal discriminative power features. It calculated the relative importance of each texture feature, and the most weighted texture features in each sequence were finally selected to construct a diagnostic model.

Statistical Analysis

Statistical analyses were performed using SPSS (version 25.0, IBM Corp.) and MedCalc (Version19.1.7, https://www.medcalc.org). The consistency of assessment between the two radiologists was evaluated by using intra-class correlation coefficients (ICC) and Cohen’s kappa statistics (κ) as follows: excellent (ICC/k > 0.80), good (ICC/k = 0.61–0.80), moderate (ICC/k = 0.41–0.60), and fair (ICC/k < 0.40). Numerical data were summarized using the median and interquartile range (IQR) when not normally distributed, and presented as mean and standard deviation (SD) when normally distributed. For normally distributed data, the independent sample t test was used, while the Mann-Whitney U-test was applied when the data are not normally distributed. Categorical variables were compared using either the Pearson’s chi-squared test or Fisher’s exact test.

To evaluate clinical and statistical significance, we included clinical and conventional MR variables with p < 0.10 to build a non-texture logistic regression model for patients
in the training group. Subsequently, a combined model integrating non-texture information and texture features was built for patients in the training group. The Hosmer-Lemeshow test was applied to assess the appropriateness of fit of the two models. The model assessed in the training group was applied to the test group. The diagnostic

Fig. 2. T2-weighted images of patients with BEOT (A) and MEOT (B), respectively. Manually defined volume of interest was drawn layer by layer in the tumor area by the MaZda software.

| Category                      | More Detailed Features                                                                 |
|-------------------------------|----------------------------------------------------------------------------------------|
| Histogram (n = 9)             | Mean, Variance, Skewness, Kurtosis, Perc.01%, Perc.10%, Perc.50%, Perc.90%, Perc.99% |
| Absolute gradient (n = 5)     | GrMean, GrVariance, GrSkewness, GrKurtosis, GrNonZeros (percentage of pixels with nonzero gradient) |
| Run-length matrix (n = 25)    | RLNonUni, GLevNonU, LngREmp, ShrtREmp, Fraction (features are computed for 5 various directions) |
| Co-occurrence matrix (n = 275) | AngScMom, Contrast, Correlat, SumOfSqs, InvDfMom, SumAverg, SumVarmc, SumEntrp, Entropy, DifVarmc, DifEntrp (features are computed for 5 between-pixels distances [1, 2, 3, 4, 5] and for 5 various directions) |

AngScMom = angular second moment, Correlat = correlation, DifEntrp = difference entropy, DifVarmc = difference variance, GLevNonU = grey level nonuniformity, GrKurtosis = absolute gradient kurtosis, GrMean = absolute gradient mean, GrNonZeros = percentage of pixels with nonzero gradient, GrSkewness = absolute gradient skewness, GrVariance = absolute gradient variance, InvDfMom = inverse difference moment, LngREmp = long run emphasis, n = total number of texture feature of each category extracted from MaZda, RLNonUni = run length nonuniformity, ShrtREmp = short run emphasis, SumAverg = sum average, SumEntrp = sum entropy, SumOfSqs = sum of suqares, SumVarmc = sum variance
performance of the models and single texture features that differentiated BEOTs from MEOTs were evaluated using receiver operating characteristic (ROC) curve analysis. The areas under the ROC curve (AUCs) were analyzed using the DeLong test. A $p < 0.05$ was considered as statistically significant. However, $p < 0.10$ was selected as the filter value for univariate and multivariate analyses.

**RESULTS**

**Clinical and Conventional MRI Characteristics**

This study included 30 patients with 36 BEOTs and 58 patients with 70 FIGO stage I/II MEOTs. The gynecological symptoms or findings of patients included accidentally discovered adnexal mass (35.2%, 31/88), abdominal distension (28.4%, 25/88), lower abdominal pain (27.2%, 24/88), and abnormal uterine bleeding (9.0%, 8/88). The BEOTs included 11 serous (three bilateral cases), 18 mucinous (one bilateral case), five seromucous (one bilateral case), and two Brenner tumors (one bilateral case); whilst the MEOTs included 46 serous (12 bilateral cases), 11 mucinous, four seromucous, and nine clear cell tumors. Among the 31 MEOT patients with pathologically confirmed pelvic extension below the pelvic brim or primary peritoneal

| Table 3. The Comparison of Clinical and Conventional MRI Characteristics between BEOTs and MEOTs |
|----------------------|----------------------|--------|
| Characteristics      | BEOT (n = 30)        | MEOT (n = 58) |
| Age, years $^1$      | 46 (28.0–60.3)       | 52 (46.0–57.3)  |
| Ca-125, U/mL$^1$     | 30.09 (17.07–55.60)  | 176.80 (36.22–1041.50) | $< 0.001$ |
| Morphology (%)       |                      |        |
| Lobulated            | 17/30 (56.7)         | 39/58 (67.2)  |
| Round/oval           | 13/30 (43.3)         | 19/58 (32.8)  |
| Configuration (%)    |                      |        |
| Purely cystic        | 1/30 (3.3)           | 0      |
| Predominantly cystic | 22/30 (73.3)         | 22/58 (37.9)  |
| Mixed cystic/solid   | 4/30 (13.3)          | 18/58 (31.0)  |
| Predominantly solid or solid | 3/30 (10.0) | 18/58 (31.0)  |
| Maximum diameter of lesions, cm$^2$ | 15.87 ± 7.67 | 12.22 ± 5.40 | 0.861 | 0.101 |
| Maximum diameter of solid component, cm$^2$ | 2.42 ± 3.68 | 5.27 ± 2.67 | 0.839 | $< 0.001$ |
| Thickened irregular septa (%) | 15/30 (50.0) | 40/58 (69.0) | 0.823 | 0.082 |
| Number of loculi (%) | 0.895 | 0.127 |
| Unilocular           | 6/30 (20.0)          | 8/58 (13.8)  |
| Multilocular         | 19/30 (63.3)         | 45/58 (77.6)  |
| Honeycomb loculi     | 5/30 (16.7)          | 3/58 (5.2)   |
| SI of cystic component on FS-T1WI (%) | 0.642 | 0.439 |
| Moderate to low      | 24/30 (80.0)         | 44/58 (75.9)  |
| High                 | 5/30 (16.7)          | 12/58 (20.7)  |
| SI of cystic component on T2WI (%) | 0.943 | 0.578 |
| Low                  | 1/30 (3.3)           | 3/58 (5.2)   |
| Moderate to high     | 28/30 (93.3)         | 53/58 (91.4)  |
| Enhancement (%)      | 0.865 | 0.691 |
| Mild                 | 6/30 (20.0)          | 8/58 (13.8)  |
| Moderate             | 7/30 (23.3)          | 17/58 (29.3)  |
| Significant          | 17/30 (56.7)         | 33/58 (56.9)  |
| Hemorrhagic fluid in the loculi (%) | 4/30 (13.3) | 9/58 (15.5) | 0.915 | 0.527 |
| Ascites (%)          | 0.871 | 0.983 |
| No/mild              | 11/30 (36.7)         | 22/58 (37.9)  |
| Moderate             | 15/30 (50.0)         | 29/58 (50.0)  |
| Massive              | 4/30 (13.3)          | 7/58 (12.1)   |

*Consistency of assessment was evaluated by ICC unless otherwise indicated. $^1$Abnormally distributed numerical data were presented as the median (interquartile range). $^2$Normally distributed numerical data were presented as the mean ± standard deviation. Consistency of assessment was evaluated by k. BEOT = borderline epithelial ovarian tumor, CA-125 = cancer antigen-125, FS = fat-suppressed, ICC = intra-class correlation coefficients, k = Cohen’s kappa statistics, MEOT = malignant epithelial ovarian tumor, SI = signal intensity.
cancer (FIGO stage II), only eight lesions (25.8%) were macroscopically observed by radiologists.

A comparison of clinical and MRI characteristics in BEOT and MEOT patients is presented in Table 3. Patients with MEOT (176.80 [36.22, 1041.50] U/mL) had higher concentrations of CA-125 (median [IQR]) than those with BEOT (30.09 [17.07, 55.60] U/mL) ($p < 0.001$). Furthermore, MEOTs had a greater solid component ($p < 0.001$) and more frequently exhibited a mixed cystic-solid or predominantly-solid configuration compared to BEOTs ($p = 0.003$) (Figs. 3, 4). Purely cystic features were only observed in BEOTs. However, there were no significant differences in other qualitative variables between the two groups.

**Texture Feature Extraction and Selection**

The 30 optimal texture features automatically selected by the MaZda software in the training group are shown in Table 4. Relative importance was calculated using the Random Forest Algorithm for each feature as outlined in Figure 5. The three most weighted texture features were the $S(0,1,0)$ difference entropy (DiffEntrp), extracted from T2WI images; the $S(0,0,4)$ sum average (SumAverg), extracted from DWI images; and the $S(0,0,5)$ SumAverg, extracted from CE-T1WI images.

**Logistic Regression Models and ROC Curves**

The non-texture model built in the training group included CA-125, configuration, thickened irregular septa, and maximum diameter of solid component (all $p$ values < 0.10). The combined model built in the training group integrated non-texture information and the three most weighted texture features. Table 5 shows the performance...
of the models in differentiating BEOTs from MEOTs in the training and test groups. The Hosmer-Lemeshow test revealed that both models were predictive in differentiating BEOTs from MEOTs ($p = 0.348$ and $0.073$, respectively). In the training group, the combined model showed a superior diagnostic performance than the non-texture model, and the comparison of AUCs revealed that the difference in diagnostic performance was statistically significant ($p = 0.038$). In addition, there was no significant difference in the AUCs of the combined model between the training and test group ($p = 0.348$). In the test group, however, there was no difference ($p = 0.645$) between the combined model (AUC, 0.896; 95% confidence interval [CI], 0.713–0.980) and the non-texture model (AUC, 0.840; 95% CI, 0.644–0.953).

The single texture features revealed a statistically significant difference between BEOTs and MEOTs in the training group ($p < 0.001$, $p = 0.003$, $p = 0.015$ for T2WI, DWI, and CE-T1WI images, respectively). The performance of the models and the single texture features in the training group are described in Figure 6. A comparison of

| Category                          | BEOTs vs. MEOTs |
|-----------------------------------|-----------------|
| T2WI                              | 0               |
| DWI                               | 0               |
| CE-T1WI                           | 1/30            |

Table 4. The Frequency for Each Category of the 30 Optimal Texture Features

Fig. 4. A 45-year-old woman with bilateral low-grade serous carcinoma in the left fallopian tube (FIGO stage II).

Sagittal T2WI (B) shows cystic-predominant and mixed cystic-solid masses with thickened septas (white arrows, C). The solid component exhibited moderate signal intensity during T1WI (A) and restricted diffusion (black arrow, D, E). After contrast enhancing, it becomes markedly enhanced. The corresponding histopathological section (F) shows a solid component of cells with nuclear atypia (hematoxylin-eosin staining; magnification x 100).
Fig. 5. The relative importance of the 30 optimal texture features derived from T2WI (A), DWI (B), and CE-MRI images (C) were automatically calculated using the Random Forest model. AngScMom = angular second moment, Correlat = correlation, DiffEntrp = difference entropy, DiffVarnc = difference variance, GrKurtosis = absolute gradient kurtosis, GrNonZeros = percentage of pixels with nonzero gradient, GrSkewness = absolute gradient skewness, GrVariance = absolute gradient variance, Horzl_Fraction = horizontal_fraction of image in runs, Horzl_RLNonUni = horizontal_run length nonuniformity, InvDfMom = inverse difference moment, LngREmph = long run emphasis, SumAverg = sum average, SumOfSqs = sum of squares, SumVarnc = sum variance, Vertl_ShrtREmp = vertical_short run emphasis, 135dr_RLNonUni = 135 degree run length nonuniformity, 135dr_ShrtREmp = 135 degree_short run emphasis.
Texture Analysis of 3D MRI Images May Differentiate Epithelial Ovarian Tumors

Table 5. Diagnostic Performance of the Non-Texture Model and Combined Model in Differentiating BEOTs from MEOTs

| Model                  | Training Group (n = 62) | Test Group (n = 26) |
|------------------------|-------------------------|---------------------|
|                        | P*  | Se (%) | Sp (%) | Ac (%) | AUC     | Se (%) | Sp (%) | Ac (%) | AUC     |
| Non-texture model      | < 0.001 | 78.3   | 84.6   | 82.3   | 0.818   | 73.3    | 90.1    | 80.8    | 0.840   |
| Combined model         | < 0.001 | 92.5   | 86.4   | 90.3   | 0.962   | 75.0    | 94.0    | 88.5    | 0.896   |

*Omnibus Tests of Model Coefficients. Ac = accuracy, AUC = area under the receiver operating characteristic curve, Se = sensitivity, Sp = specificity.

Fig. 6. ROC curves of the logistic models and single texture features of BEOTs and MEOTs in the training group. The AUCs of the most weighted texture feature extracted from CE-T1WI, DWI, and T2WI images are 0.690, 0.728, and 0.818, respectively. The AUC of the non-texture model is 0.818 and that of the combined model is 0.962. AUC = area under the ROC curve, ROC = receiver operating characteristic.

DISCUSSION

This study assessed the diagnostic performance of texture analysis in the discrimination of BEOTs from FIGO stage I/II MEOTs. Our results showed that texture features derived from T2WI, DWI, and CE-T1WI images significantly differed between the BEOT and MEOT groups. In addition, diagnostic models that combine texture features and conventional variables could accurately differentiate BEOTs from MEOTs.

It has become acceptable to perform conservative fertility-sparing surgery and to omit lymphadenectomy during the treatment of BEOTs due to their favorable prognosis, while conservative surgery is restricted to only stage I A serous, mucinous, or endometrioid MEOTs (18). Imaging analyses of stage III or higher MEOTs tend to reveal a more aggressive appearance (obvious lymphatic spread and distant metastasis), and an ovarian tumor with such features can be easily diagnosed as an MEOT. On the other hand, a considerable portion of stage II MEOTs with pathologically confirmed pelvic invasion cannot be seen with the naked eye. Thus, we recruited only stage I/II MEOTs patients for this study. Our results demonstrated that either a mixed cystic/solid or predominantly-solid configuration and a greater solid component were more frequently found in stage I/II MEOTs. These findings are consistent with previous observations (19, 26).

Texture analysis is a novel imaging technique that can extract extensive data from biomedical images and can be investigated with textural analysis tools. Texture analysis of MRI images is performed by analyzing the gray tone variations among image voxels. This method captures spatial and intensity information that identify pathological changes and microstructural heterogeneity within a tumor. Since the voxel size in the MR images in this study was 1 mm and the features of texture analysis were set to no less than 6 bits per voxel for texture extraction, the texture features detected subtle differences between MR images of BEOTs and MEOTs. Due to the heterogeneity of ovarian tumors, texture features based on the whole tumor contained more spatial information (along the extra-dimension) and higher sensitivity and specificity than traditional techniques which are based on a pre-selected region of interest.

In our study, single texture features extracted from T2WI, DWI, and CE-T1WI images were significantly different between BEOTs and MEOTs. Considering that the same acquisition conditions were used, this might be due to the essential differences in spatial and intensity characteristics.
among image voxels caused by pathological changes and microstructural heterogeneity within different tumors. However, the single texture feature of CE-T1WI (AUC, 0.690) and DWI (AUC, 0.728) images showed a relatively lower diagnostic performance in the training group compared to that of T2WI images (AUC, 0.818). On CE-T1WI images, the difference in radiologic features between BEOTs and MEOTs may be related to the limitation of CE images, which are not sufficient to identify changes in the microcirculation induced by neoangiogenesis. This is consistent with our conventional imaging findings ($p = 0.691$) and several previous studies (10, 11, 19). On DWI images, this may be associated with the lower signal-to-noise ratio and the lack of ADC mapping.

Our results also indicated that the combined model showed the highest diagnostic performance in differentiating BEOTs from MEOTs, when compared to the non-texture model and the single texture features. This combined model used in the training group (accuracy, 90.3%) was validated in the test group (accuracy, 88.5%) and there was no significant difference in the combined model between the two groups ($p = 0.348$). However, the superiority of the combined model compared to the non-texture model was not significant in the test group. The predictability of diagnostic models tend to be higher in the training group than in the test group (27). This may result in decreased differences between the combined model and non-texture model in a test group. This reduced predictability may also be due to the complexity of the combined model (three texture features were added) for the training group compared to the non-texture model, leading to the impaired generalization capability of the model and the decreased accuracy of the combined model in the test group. Moreover, the increased number of variables in the combined model are more likely to result in an over-fitting phenomenon due to the small sample size in the test group.

This study still had several limitations. First, due to its retrospective design and relatively small sample size, it was impossible to compare different histological tumor types, limiting the generalizability and statistical power of the study. Second, our study only separated samples into either the training or test group in order to perform internal validation; however, external validation is still required. Therefore, further multi-center studies with larger patient sample sizes are needed to validate the applicability of our findings. Finally, although this study described preliminary results on texture analysis of MR images for the differentiation of BEOTs from MEOTs, further studies are needed to investigate the pathological basis of our findings.

In conclusion, MRI-based texture analysis of the whole tumor may assist in characterizing the differences between BEOT and MEOT patients. These findings in turn may guide diagnostic protocols for future patients. Furthermore, MRI texture features combined with clinical and conventional MRI characteristics exhibit better diagnostic performance than texture features alone.

Conflicts of Interest
The authors have no potential conflicts of interest to disclose.

Acknowledgments
The authors thank the radiographers in our departments for their assistance in the experimental studies as well as data analysis.

ORCID iDs
Rongping Ye  
https://orcid.org/0000-0001-7867-9752
Shuping Weng  
https://orcid.org/0000-0001-8271-934X
Yueming Li  
https://orcid.org/0000-0002-3669-568X
Chuan Yan  
https://orcid.org/0000-0003-4106-8995
Jianwei Chen  
https://orcid.org/0000-0002-5791-9865
Yuemin Zhu  
https://orcid.org/0000-0001-9630-0160
Liting Wen  
https://orcid.org/0000-0002-3009-8630

REFERENCES
1. Lalwani N, Prasad SR, Vikram R, Shanbhogue AK, Huettner PC, Fasih N. Histologic, molecular, and cytogenetic features of ovarian cancers: implications for diagnosis and treatment. Radiographics 2011;31:625-646
2. Kurman RJ, Carcangiu ML, Herrington CS, Young RH. WHO classification of tumors of female reproductive organs, 4th ed. Lyon: International Agency for Research on Cancer, 2014
3. National Comprehensive Cancer Network. Ovarian cancer including fallopian tube cancer and primary peritoneal cancer (Version 1.2020). Philadelphia: National Comprehensive Cancer Network, 2020
Texture Analysis of 3D MRI Images May Differentiate Epithelial Ovarian Tumors

15. Han X, Sun M, Wang M, Fan R, Chen D, Xie L, et al. The enhanced T2 star weighted angiography (ESWAN) value for differentiating borderline from malignant epithelial ovarian tumors. *Eur J Radiol* 2019;118:187-193

16. Incoronato M, Aiello M, Infante T, Cavaliere C, Grimaldi AM, Mirabelli P, et al. Radiogenomic analysis of oncological data: a technical survey. *Int J Mol Sci* 2017;18:805

17. Ng F, Kozarski R, Ganeshan B, Goh V. Assessment of tumor heterogeneity by CT texture analysis: can the largest cross-sectional area be used as an alternative to whole tumor analysis? *Eur J Radiol* 2013;82:342-348

18. Denewar FA, Takeuchi M, Urano M, Kamishima Y, Kawai T, Takahashi N, et al. Multiparametric MRI for differentiation of borderline ovarian tumors from stage I malignant epithelial ovarian tumors using multivariate logistic regression analysis. *Eur J Radiol* 2017;91:116-123

19. Li YA, Qiang JW, Ma FH, Li HM, Zhao SH. MRI features and score for differentiating borderline from malignant epithelial ovarian tumors. *Eur J Radiol* 2018;98:136-142

20. Thomassin-Naggara I, Aubert E, Rockall A, Jalgual-Coudray A, Rouzier R, Darai E, et al. Adnexal masses: development and preliminary validation of an MR imaging scoring system. *Radiology* 2013;267:432-443

21. Ma FH, Zhao SH, Qiang JW, Zhang GF, Wang XZ, Wang L. MRI appearances of mucinous borderline ovarian tumors: pathological correlation. *J Magn Reson Imaging* 2014;40:745-751

22. Szczypinski PM, Strzelecki M, Materka A, Klepaczko A. MaZda-A software package for image texture analysis. *Comput Methods Programs Biomed* 2009;94:66-76

23. Strzelecki M, Szczypinski P, Materka A, Klepaczko A. A software tool for automatic classification and segmentation of 2D/3D medical images. *Nucl Instrum Methods Phys Res A* 2013;702:137-140

24. Szczypinski PM, Strzelecki M, Materka A. MaZda-A software for texture analysis. 2007 International symposium on information technology convergence (ISITC 2007);2007 November 23-24; Jeonju, Korea

25. Tourassi GD, Frederick ED, Markey MK, Floyd CE Jr. Application of the mutual information criterion for feature selection in computer-aided diagnosis. *Med Phys* 2001;28:2394-2402

26. Kurata Y, Kido A, Moribata Y, Kameyama K, Himoto Y, Minamiguchi S, et al. Diagnostic performance of MRI imaging findings and quantitative values in the differentiation of serous/mucinous borderline tumour from endometriosis-related malignant ovarian tumour. *Eur Radiol* 2017;27:1695-1703

27. Iasonos A, Schrag D, Raj GV, Panageas KS. How to build and interpret a nomogram for cancer prognosis. *J Clin Oncol* 2008;26:1364-1370