Original Article Magnetic resonance imaging semantic and quantitative features analyses: an additional diagnostic tool for breast phyllodes tumors

Wenjuan Ma^{1,2,3,4*}, Xinpeng Guo^{1,2,3*}, Liangsheng Liu^{1,2,3*}, Lisha Qi^{1,2}, Peifang Liu^{1,2,3}, Ying Zhu^{1,2,3}, Xiqi Jian⁴, Guijun Xu⁵, Xin Wang⁶, Hong Lu^{1,2,3}, Chao Zhang⁵

¹Department of Breast Imaging, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Huanhuxi Road, Hexi District, Tianjin 300060, P. R. China; ²Key Laboratory of Breast Cancer Prevention and Therapy, Huanhuxi Road, Hexi District, Tianjin 300060, P. R. China; ³Tianjin's Clinical Research Center for Cancer, Huanhuxi Road, Hexi District, Tianjin 300060, P. R. China; ⁴Department of Biomedical and Engineering, Tianjin Medical University, 22 Qixiangtai Road, Heping District, Tianjin 300070, P. R. China; ⁵Department of Bone and Soft Tissue Tumors, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin's Clinical Research Center for Cancer, Tianjin, P. R. China; ⁶Department of Epidemiology and Biostatistics, First Affiliated Hospital, Army Medical University, 30 Gaotanyan Street Shapingba District, Chongqing 400038, P. R. China. *Equal contributors.

Received October 11, 2019; Accepted April 24, 2020; Epub May 15, 2020; Published May 30, 2020

Abstract: Objective: This study aimed to differentiate benign and non-benign (borderline/malignant) phyllodes tumors of the breast by the semantic and quantitative features in magnetic resonance imaging (MRI). Methods: The female patients, diagnosed with phyllodes tumors by MRI and pathological test, were retrospectively selected from December, 2006 to April, 2019. The MRI of benign, borderline and malignant phyllodes tumors was analyzed using 8 semantic features and 20 computed quantitative features from diffuse contrast-enhanced magnetic resonance imaging (DCE-MRI). The semantic features were analyzed by univariate analysis. The least absolute shrinkage and selection operator (LASSO) method was used to identify the optimal subset of MRI quantitative features. According to the results from multivariate logistic regression for the semantic and quantitative features, the model was constructed to differentiate benign and non-benign (borderline/malignant) phyllodes tumors. Results: Thirty-two benign (58.18%), 13 borderline (23.64%) and 10 malignant (18.18%) phyllodes tumors were identified in 54 patients. Five semantic features were proved to be significantly correlated with pathologic grade, including size, the T1 weighted image signal intensity, fat-saturated T2-weighted image signal intensity, enhanced signal intensity, and kinetic curve pattern. With the analysis of LASSO method, three quantitative texture features with significant predictive ability were selected. The model combining both the semantic and quantitative features was proved to have good performance in differentiation on phyllodes tumors, vielding an area under receiver operating characteristic curve, accuracy, sensitivity and specificity of 0.893, 0.933, 1.000, and 0.818, respectively. Conclusion: The constructed model based on the semantic and quantitative features of DCE-MRI can significantly improve the differential diagnosis of phyllodes tumors in breast.

Keywords: Phyllodes tumor, breast, semantic, quantitative, MRI

Introduction

Phyllodes tumors (PTs) of the breast are rare fibroepithelial neoplasms that have been reported to account for 0.3-1% of all breast neoplasms [1]. These tumors commonly occur in middle-aged women, and their course is often unpredictable [2]. Histologically, PTs can be classified as benign, borderline, and malignant. Local recurrences and distant metastases can occur regardless of the histological type [3]. As reported in a meta-analysis incorporating 54 studies, the pooled local recurrence rates for benign, borderline, and malignant PTs are 8%, 13%, and 18%, respectively [4]. Another study reported similar increasing recurrences from benign to malignant tumors of 3.6%, 14.1%, and 42.3%, respectively [5]. The 5-year disease-free survival rates of patients with benign, border, and malignant PTs are reported to be 94%, 91%, and 67%, respectively [6]. Thus, to decrease the PT recurrence rate and to improve patient survival, it is essential to preoperatively improve the accuracy of diagnosis of patients with malignant breast PTs.

The diagnosis of PTs is difficult without incisional or aspiration biopsy. Previous reports show that fine-needle aspiration offers limited information for differentiating a diagnosis between benign and malignant PTs because PTs are composed of various characteristics [7, 8]. Radiological evaluation is one of the most important methods for the diagnosis. Mammographic texture analysis has also been used to differentiate the benign and borderline/malignant PTs of breast [9]. Currently, magnetic resonance imaging (MRI) has been accepted as the most sensitive imaging technique for breast cancer [10-12].

Recently, several studies analyzed the MRI semantic features of PTs with different pathological grades [13-16]. In addition, quantitative radiomic methods with diffuse contrast-enhanced magnetic resonance imaging (DCE-MRI) have been extensively used [17-19]. All these studies provide the impetus to establish new methods to evaluate the pathological grade of PTs. Based on the development of these MRI techniques and quantitative analytics, we analyzed the semantic and quantitative features of PTs using MRI of the breast and constructed a model to assist in the differential diagnosis of the pathological grade of PTs.

Materials and methods

Patient selection and MRI technique

The institutional review board of Tianjin Medical University Cancer Institute and Hospital approved this retrospective study. The informed consent requirement was waived. The inclusion criteria were as follows: (1) histopathological diagnosis of PTs through surgically resected specimens; (2) MRI scanning before neoadjuvant therapy or surgical resection; (3) a time interval between MRI examination and surgery of less than 2 weeks. After screening, 55 female patients (56 lesions) with the pathological diagnosis of PTs and preoperative breast MRI participated in our study from December 2006 to April 2019.

Sixteen MRI scans were acquired with a 1.5 Tesla (T) scanner, and 38 MRI scans were acquired with a 3.0 T scanner. The specific scanning parameters for the 1.5 T scanner have been previously published [20]. The examinations performed with the 3.0 T scanner used a dedicated 8-channel phased-array breast coil (Discovery MR750, GE Medical Systems). The MRI protocols included axial T1-weighted, fatsaturated T2-weighted, and unilateral sagittal fat-saturated fast spin-echo (FSE) sequences. T2-weighted imaging of the affected breast was conducted before administration of the contrast agent. Diffusion-weighted imaging (DWI) was performed using a multi-section, spinecho, single-shot, echo-planar sequence bilaterally in the axial and sagittal planes of the affected breast. Sensitizing diffusion gradients were applied sequentially in the x, y, and z directions, with b values of 1,000 s/mm². Images were obtained by sagittal DCE-MRI using the volume imaging for breast assessment (VIBRANT) bilateral breast imaging technique. Sagittal data were obtained with the VIBRANT bilateral breast imaging technique, with TR = 6.1 ms, TE = 2.9 ms, flip angle = 15 degrees, matrix size = 256×128 , field of view = 26 cm \times 26 cm, NEX = 1, and slice thickness = 1.8 mm. Before injecting the contrast agent, serial mask images were obtained. Subsequently, the contrast agent (Gd-DTPA, 0.2 mL/kg body weight, flow rate 2.0 mL/s) was injected using an automatic MR-compatible power injector and then flushed with an equal volume of saline solution. The dynamic MRI acquisitions were performed immediately after the injection. The acquisition was repeated five times, and each phase took 90-100 s. Final axial 3-D fastspoiled gradient-recalled echo images were obtained after the dynamic study for all patients. Quantitative analysis was conducted at the first post-contrast time point of the MRI.

Semantic features analysis

Post processing of the images was performed on an Advantage Workstation (AW 4.2) using Functool II software (GE Healthcare). Two senior radiologists, who were experienced in breast MRIs, independently analyzed the MRI findings. They independently interpreted the MRIs using the 2013 MRI Breast Imaging Reporting and Data System (BI-RADS) tool of the American College of Radiology. Differences in interpretations were resolved by review of images and discussion according to the BI-RADS standard. The recorded data included morphology (size, shape, and margin), signal intensity on T1- and fat-saturated T2-weighted FSE images, enhancement patterns (compared with those of normal fibroglandular tissue of the breast), timesignal intensity curve (TIC) patterns, and the mean value of the apparent diffusion coefficient (ADC) on DWI. The signal intensity on T1-weighted FSE images was visually classified as either low or high. The signal intensity on fatsaturated T2-weighted FSE images was divided into heterogeneity or homogeneity. The TICs were classified as follows: type I, slow or rapid initial contrast enhancement with a persistent delayed phase; type II, rapid initial enhancement followed by a plateau of signal intensity: and type III, rapid initial enhancement followed by rapid washout [21]. We placed maximumsized circular regions of interest (ROIs) within the lesions on the ADC maps, avoiding apparent necrotic or cystic components by referring to other MRIs.

Quantitative feature extraction

Quantitative features were extracted from the lesion area on each image. To extract the features, the first step was to outline the lesion area on each image. A radiologist manually outlined the contours of the tumor in each image using ImageJ (https://imagej.en.softonic. com/). Another radiologist reviewed the image segmentation results. To normalize the different image specifications from various MRI scanners, image resampling and gray-level normalization were performed before quantitative feature extraction. All image data were resampled at a 1 × 1 × 1-mm voxel space size. The guantitative features were extracted from threedimensional ROIs using an in-house software developed with MATLAB 2016a (MathWorks Inc.), which included eight morphological features, three histogram features, and nine graylevel co-occurrence matrix (GLCM) features [22]. Details of the procedures for quantitative features extraction are described in Supple-<u>mentary I</u>.

Statistical analysis

Continuous variables were presented as the mean (\pm standard deviation). Categorical variables were presented as the number (N) and percentage (%). Differences in continuous vari-

ables were analyzed with the Student's t-test or the Mann-Whitney U test. Differences in categorical variables were tested with the chisquare test or Fisher's exact test. A p-value of < 0.05 was considered be significantly different. The least absolute shrinkage and selection operator (LASSO) method was used to identify the optimal subset of MRI radiomic features [23, 24]. Multivariate logistic regression combining semantic and quantitative features was used to construct a model to distinguish benign and non-benign (borderline/malignant) PTs. The diagnostic sensitivity, specificity, accuracy, and area under the curve (AUC) of the receiveroperating characteristic (ROC) were calculated to evaluate the diagnostic accuracy.

All statistical analyses were performed with R software (R Core Team. R: A language and environment for statistical computing; R Foundation for Statistical Computing, Vienna, Austria; http://www.R-project.org, 2016). The LASSO regression was performed using the "penalized" package in R. Multivariate logistic regression analysis was performed using the "rms" package. The ROC was plotted using the "pROC" package.

Results

Patient characteristics

This cohort included 32 (58.18%) benign, 13 (23.64%) borderline, and 10 (18.18%) patients with malignant PTs. The patient ages ranged from 22 to 53 years (38.30 \pm 7.66; median 38 years) in the benign group, from 23 to 61 years (41.08 \pm 10.96; median 39 years) in the borderline group, and from 24 to 57 years (42.7 \pm 12.14, median 46 years) in the malignant group. There was no significant difference in age among the three groups (*P* = 0.401). The histogram in **Figure 1** shows the contribution of age to the histopathological grade.

Semantic features of MRI

The detailed MRI semantic features for benign, borderline, and malignant PTs are summarized in **Table 1**. There was a significant correlation between the size and the histologic grade of the PTs ($\chi^2 = 18.347$, P = 0.001). The percentages of tumor size > 5 cm for malignant, borderline, and benign tumors are 70%, 23.08%, and 6.06%, respectively. Tumor signal intensity



Figure 1. Histogram showing the contribution of age to the histopathological grade.

higher than normal breast tissue signal intensity on T1-weighted images was more common in the malignant (in 6 of 10 tumors) and borderline (in 2 of 13 tumors) tumor groups than in the benign group (in 1 of 33 tumors). These differences are statistically significant (χ^2 = 18.472, P < 0.001). On the fat-saturated T2-weighted FSE images and DCE-MRI malignant PTs show more heterogeneous patterns (χ^2 = 7.291/16.695. P < 0.001). As shown in the TIC analysis, malignant PTs are more likely to develop type II/III enhancement (χ^2 = 16.972, P < 0.001). There was no statistical difference among the three groups for two morphological features (shape and margin) and ADC values. Three typical MRI images on benign, borderline, and malignant PTs are shown in Figure 2.

Radiomics features selection

The LASSO regression method was used to select the most robust and non-redundant quantitative features from the features extracted. Among the quantitative features in the entire cohort, 20 were reduced to three potential predictors (**Figure 3**); namely, Compactness 1 (morphological feature), Correlation (GLCM feature), and SumAverage (GLCM feature).

Development of a prediction model

Correlations of all the features between benign and non-benign (borderline/malignant) PTs are shown in **Figure 4A**. There were moderate negative correlations between MRI semantic features (T1WI, T2WI, and Enhanced signal) and quantitative features (Contrast, SumAverage, and Autocorrelation). In contrast, there were moderate positive correlations between semantic features and quantitative features. Finally, a predictive model was developed to distinguish benign and nonbenign (borderline/malignant) PTs based on the semantic (P <0.05) and quantitative (LASSO selected) features and using multivariate logistic regression. A quantitative value to represent the predicted signature (Equation 1) included five semantic features and three quantitative features. This model was proven to have the

best overall performance, with significantly higher values than either semantic or quantitative features (Figure 4B; Table 2): AUC = 0.939 (95% CI: 0.871-0.994), accuracy = 0.892 (95% CI: 0.781-0.960), sensitivity = 1 (95% CI: 0.822-1), specificity = 0.818 (95% CI: 0.639-0.924). The heat map for the distribution of all the features is shown in Figure 4C.

Score = 0.8223 - 0.0367 * Size + 0.1571 * T1WI - 0.0286 * T2WI + 0.2066 * Enhanced_ signal + 0.3117 * TIC - 0.0006 * Compactness_ 1 - 0.1289 * Correlation - 0.6081 * SumAverage Equation 1

Discussion

The preoperative identification between benign and non-benign PTs is challenging. Compared with ultrasound in the evaluation of PTs, MRI has high sensitivity and accuracy with regard to its ability to reveal malignant MRI features. Currently, these features cannot be identified through mammography or ultrasound. In the present study, we constructed a model based on the features of DCE-MRI to preoperatively predict tumor pathological grade in patients with breast PTs. The predictive performance was significantly improved by combining the semantic signatures with quantitative features in a calculation model, achieving an AUC of 0.939. The correlation analysis demonstrated the association between the MRI features and the quantitative texture features, suggesting that non-benign PTs (borderline/malignant) of the breast are more heterogeneous than benign PTs.

	Pa				
Semantic feature	Benign	Borderline	Malignant	χ^2 value	P value
	n=33	n=13	n=10		
Size				18.347	0.001*
<= 2 cm	10 (30.30%)	3 (23.08%)	0 (0%)		
>2 cm and <=5 cm	21 (63.64%)	7 (53.85%)	3 (30.00%)		
>5 cm	2 (6.06%)	3 (23.08%)	7 (70.00%)		
Shape of mass				1.846	0.764
Round/Oval	5 (15.15%)	2 (15.38%)	0 (0%)		
Lobular	21 (63.64%)	8 (61.54%)	7 (70.00%)		
Irregular	7 (21.21%)	3 (23.08%)	3 (30.00%)		
Margin				4.178	0.124
Well-defined	27 (81.82%)	10 (76.92%)	5 (50.00%)		
III-defined	6 (18.18%)	3 (23.08%)	5 (50.00%)		
High T_1 WI signal				18.472	<0.001*
Yes	1 (3.30%)	2 (15.38%)	6 (60.00%)		
No	32 (96.97%)	11 (84.62%)	4 (40.00%)		
Fat-saturated T_2 WI signal				7.291	0.026*
Homogeneous	19 (57.58%)	5 (38.46%)	1 (10.00%)		
Heterogeneous	14 (42.42%)	8 (61.54%)	9 (90.00%)		
Enhanced signal				16.695	<0.001*
Homogeneous	24 (72.73%)	6 (46.15%)	0 (0%)		
Heterogeneous	9 (27.27%)	7 (53.85%)	10 (100%)		
Kinetic curve pattern				16.972	<0.001*
Type II and III	14 (42.42%)	13 (100%)	9 (90.00%)		
Туре І	19 (57.58%)	0 (0%)	1 (10.00%)		
ADC value (×10 ⁻³ mm ² /s)	1.51±0.22	1.44±0.22	1.34±0.25	5.940	0.051

Table 1. Descriptive statistics of MRI semantic features of phyllodes tumors

*P<0.05 for benign, borderline and malignant phyllodes tumor of breast.

The age at diagnosis for PTs varied widely, but middle-aged individuals were the most frequently diagnosed individuals [13]. In our series, the mean age at diagnosis was 39.73 years. As shown in **Figure 1**, the highest incidence of PTs occurred in the 30-50-year age group, and patients under 30 years of age were more likely to develop benign tumors.

Tumor size is accepted as an important factor that can influence biological behavior and patient survival [25, 26]. Our results show that the malignancy rate increased significantly with increasing tumor size, especially those tumors with a diameter of > 5 cm. This finding reflects the high proliferative activity of high-grade PTs, though there has been no reported significant difference among the PTs [13]. PTs have been described to have the characteristic morphologic sign of well-defined margins, which show a round or lobulated shape [14, 15, 27]. However, our study suggested no significant difference among benign, borderline, and malignant PTs in terms of lesion shape. A study with a larger sample size is needed to shed light on these different results.

Our study revealed significant differences among benign, borderline, and malignant PTs in terms of T1WI features, T2WI features, enhanced signal intensities, and TICs. Nine of the 55 PTs showed high signal intensity on T1WIs, and most of these tumors were borderline and malignant tumors (malignant tumors were the most common). There were significant differences among the three groups, which is consistent with a previous study [13]. The high signal intensity on T1WIs may be related to hemorrhagic infarctions in PTs [13, 28]. The fat-saturated T2WIs in the present study showed that malignant PTs were more likely than benign and borderline PTs to have a heterogeneous signal



Figure 2. MR images of benign, borderline and malignant phyllodes tumors, respectively. A-D. Axial T1WI, axial fatsaturated T2WI, sagittal contrast-enhanced MR images about 1.5 min after contrast medium injection and time signal intensity curve of the benign phyllodes tumors. E-H. Axial T1WI, axial fat-saturated T2WI, sagittal contrastenhanced MR images about 1.5 min after contrast medium injection and time signal intensity curve of the borderline phyllodes tumors. I-L. Axial T1WI, axial fat-saturated T2WI, sagittal contrastenhanced MR images about 1.5 min after contrast medium injection and time signal intensity curve of the borderline phyllodes tumors. I-L. Axial T1WI, axial fat-saturated T2WI, sagittal contrast-enhanced MR images about 1.5 min after contrast medium injection and time signal intensity curve of the malignant phyllodes tumors.



Figure 3. Quantitative feature selection using the least absolute shrinkage and selection operator (LASSO) binary logistic regression model. A. Selection of tuning parameter lambda in the LASSO model used 10-fold cross-validation. The gray line in the figure is the partial likelihood estimate corresponding to the optimal value of lambda. The optimal lambda value of 5.237 was chosen. B. LASSO coefficient profiles of the twenty selected features. A vertical line was plotted at the optimal lambda value, which resulted in three features with nonzero coefficients.



Figure 4. Semantic and quantitative features of all phyllodes tumors. A. Correlation between predictors evaluated using Spearman rank coefficient. B. ROC curves differentiating benign from borderline/malignant PTs using multivariate logistic regression model and involved semantic and quantitative features. C. The heatmap for features shows the distribution of these features. Each column corresponds to one feature for all patients. Each row corresponds to all features for one patient.

intensity. This difference may be related to the presence of hemorrhages and cystic areas with internal separations.

Dynamic MRI patterns can guide differentiation of benign lesions from malignant lesions [19, 29]. Our study showed non-benign (borderline/ malignant) PTs were more heterogeneous than benign PTs. The enhanced signal intensity feature showed significant differences among the three groups. These differences may be related to intralesional hemorrhages, necrosis, and cystic degeneration. The results show that borderline and malignant PTs were more likely than benign PTs to have type II and III TIC curves patterns. A potential explanation for such a phenomenon is that there is a richer blood supply and more active tumor growth in borderline and malignant PTs than in benign PTs. One previous study with 23 benign and one malignant PTs found type I TICs in 16 tumors, type II TICs in three tumors, and type III TICs in five tumors [13], suggesting the inconsistency between the enhancement patterns/TIC patterns and the

Feature	AUC	Accuracy	Sensitivity	Specificity	Cut off value
Prediction model	0.939	0.893	1	0.818	0.320
Size	0.717	0.732	0.565	0.848	3.700
T1WI	0.659	0.714	0.348	0.970	0.5
T2WI	0.657	0.643	0.739	0.576	0.5
Enhanced signal	0.733	0.732	0.739	0.727	0.5
TIC	0.822	0.732	0.957	0.576	1.5
Compactness	0.675	0.696	0.435	0.879	93.883
Correlation	0.748	0.768	0.609	0.879	0.885
Sum Average	0.706	0.732	0.522	0.909	0.017

 Table 2. Classification performance of the prediction model and involved features

histological grade of PTs. This disagreement may be due to the limited sample size and different proportions of the histological types.

Several studies have revealed the limited ability of DWI in identifying benign and malignant tumors of the breast [28, 30]. The ADC values of breast tumors are inversely correlated with tumor cellularity. Most malignant breast tumors show lower ADC values than benign tumors [29, 31]. Yabuuchi and co-workers concluded that low ADC values are more prevalent in malignant PTs and that low values correspond to the histopathological finding of stromal hypercellularity [13]. In the present study, the borderline and malignant PTs had lower mean ADC values than benign PTs, but there were no significant differences among three different PT types. The diversity in results can be explained by the fact that DWI was performed in only 10 patients in the previously published study, leading to a weak correlation between ADC values and histologic grade.

In our study, three radiomic features were identified as significant predictors of histologic grade; namely, Compactness, SumAverage, and Correlation. The morphology feature "Compactness" was associated with shape, texture features were associated with "SumAverage", and overall image brightness and the smoothing gradient of the patterns were associated with the "Correlation" feature. Based on these interpretations, our results indicate that benign PT lesions can be irregular, brighter, and rougher than borderline and malignant PTs.

This complexity indicates one of the limitations of our study; namely, sample size. Due to the

limited sample size in this study, the ring enhancement of the mass lesions and cluster enhancement of non-mass lesions could not be analyzed separately. Our results need to be further evaluated in future studies with the larger sample size. Another limitation has to do with the knowledge that PTs have a high propensity to recur. Unfortunately, we did not perform long-term follow-up in our series to evaluate the relationship between the MRI features and the recurrence rate. In addition. the constructed model cannot be validated with a separate test. The

predictive ability of this model should be further validated externally.

In conclusion, we constructed a predictive model based on the semantic and quantitative features of DCE-MRI. This model can be used as an additional tool in the differential diagnosis of PTs.

Acknowledgements

The present study was sponsored by Natural Science Foundation of China (81702161, 81-802508, 81903398, 81801781), Natural Science Foundation of Tianjin Science and Technology Committee China (17JCQNJC11000), the Doctor Start-up Grant of Tianjin Medical University Cancer Institute and Hospital (B1-711), Laboratory of Tumor Immunology and Pathology (Army Medical University), Ministry of Education (2017jszl01), Top talent training program of the first affiliated hospital of PLA Army Medical University (SWH2018BJKJ-12), Chongqing Natural Science Foundation Program (cstc2019jcyj-msxmX0466).

Disclosure of conflict of interest

None.

Address correspondence to: Chao Zhang, Department of Bone and Soft Tissue Tumors, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin's Clinical Research Center for Cancer, Tianjin, P. R. China. E-mail: drzhangchao@tmu.edu.cn

References

- Rowell MD, Perry RR, Hsiu JG and Barranco SC. Phyllodes tumors. Am J Surg 1993; 165: 376-379.
- [2] Guillot E, Couturaud B, Reyal F, Curnier A, Ravinet J, Lae M, Bollet M, Pierga JY, Salmon R and Fitoussi A. Management of phyllodes breast tumors. Breast J 2011; 17: 129-137.
- [3] Mangi AA, Smith BL, Gadd MA, Tanabe KK, Ott MJ and Souba WW. Surgical management of phyllodes tumors. Arch Surg 1999; 134: 487-492, 492-493.
- [4] Lu Y, Chen Y, Zhu L, Cartwright P, Song E, Jacobs L and Chen K. Local recurrence of benign, borderline, and malignant phyllodes tumors of the breast: a systematic review and meta-analysis. Ann Surg Oncol 2019; 26: 1263-1275.
- [5] Zhou ZR, Wang CC, Sun XJ, Yang ZZ, Chen XX, Shao ZM, Yu XL and Guo XM. Prognostic factors in breast phyllodes tumors: a nomogram based on a retrospective cohort study of 404 patients. Cancer Med 2018; 7: 1030-1042.
- [6] Rodrigues MF, Truong PT, McKevitt EC, Weir LM, Knowling MA and Wai ES. Phyllodes tumors of the breast: the British Columbia Cancer Agency experience. Cancer Radiother 2018; 22: 112-119.
- [7] Jayaram G and Sthaneshwar P. Fine-needle aspiration cytology of phyllodes tumors. Diagn Cytopathol 2002; 26: 222-227.
- [8] Tse GM, Ma TK, Pang LM and Cheung H. Fine needle aspiration cytologic features of mammary phyllodes tumors. Acta Cytol 2002; 46: 855-863.
- [9] Cui WJ, Wang C, Jia L, Ren S, Duan SF, Cui C, Chen X and Wang ZQ. Differentiation between G1 and G2/G3 phyllodes tumors of breast using mammography and mammographic texture analysis. Front Oncol 2019; 9: 433.
- [10] Kul S, Cansu A, Alhan E, Dinc H, Gunes G and Reis A. Contribution of diffusion-weighted imaging to dynamic contrast-enhanced MRI in the characterization of breast tumors. AJR Am J Roentgenol 2011; 196: 210-217.
- [11] Pereira FP, Martins G, Figueiredo E, Domingues MN, Domingues RC, Da FL and Gasparetto EL. Assessment of breast lesions with diffusionweighted MRI: comparing the use of different b values. AJR Am J Roentgenol 2009; 193: 1030-1035.
- [12] Angelini G, Marini C, Iacconi C, Mazzotta D, Moretti M, Picano E and Morganti R. Magnetic resonance (MR) features in triple negative breast cancer (TNBC) vs receptor positive cancer (nTNBC). Clin Imaging 2018; 49: 12-16.
- [13] Yabuuchi H, Soeda H, Matsuo Y, Okafuji T, Eguchi T, Sakai S, Kuroki S, Tokunaga E, Ohno S, Nishiyama K, Hatakenaka M and Honda H.

Phyllodes tumor of the breast: correlation between MR findings and histologic grade. Radiology 2006; 241: 702-709.

- [14] Farria DM, Gorczyca DP, Barsky SH, Sinha S and Bassett LW. Benign phyllodes tumor of the breast: MR imaging features. AJR Am J Roentgenol 1996; 167: 187-189.
- [15] Kinoshita T, Fukutomi T and Kubochi K. Magnetic resonance imaging of benign phyllodes tumors of the breast. Breast J 2004; 10: 232-236.
- [16] Tan H, Zhang S, Liu H, Peng W, Li R, Gu Y, Wang X, Mao J and Shen X. Imaging findings in phyllodes tumors of the breast. Eur J Radiol 2012; 81: e62-e69.
- [17] Chang RF, Chen HH, Chang YC, Huang CS, Chen JH and Lo CM. Quantification of breast tumor heterogeneity for ER status, HER2 status, and TN molecular subtype evaluation on DCE-MRI. Magn Reson Imaging 2016; 34: 809-819.
- [18] Fan M, Li H, Wang S, Zheng B, Zhang J and Li L. Radiomic analysis reveals DCE-MRI features for prediction of molecular subtypes of breast cancer. PLoS One 2017; 12: e171683.
- [19] Li H, Zhu Y, Burnside ES, Drukker K, Hoadley KA, Fan C, Conzen SD, Whitman GJ, Sutton EJ, Net JM, Ganott M, Huang E, Morris EA, Perou CM, Ji Y and Giger ML. MR imaging radiomics signatures for predicting the risk of breast cancer recurrence as given by research versions of MammaPrint, Oncotype DX, and PAM50 gene assays. Radiology 2016; 281: 382-391.
- [20] Zhu Y, Zhang S, Liu P, Lu H, Xu Y and Yang WT. Solitary intraductal papillomas of the breast: MRI features and differentiation from small invasive ductal carcinomas. AJR Am J Roentgenol 2012; 199: 936-942.
- [21] El Khouli RH, Macura KJ, Kamel IR, Jacobs MA and Bluemke DA. 3-T dynamic contrast-enhanced MRI of the breast: pharmacokinetic parameters versus conventional kinetic curve analysis. AJR Am J Roentgenol 2011; 197: 1498-1505.
- [22] Haralick R M, Shanmugam K and Dinstein I. Texture parameters for image classification. IEEE Trans SMC 1973; 3: 610-621.
- [23] Huang YQ, Liang CH, He L, Tian J, Liang CS, Chen X, Ma ZL and Liu ZY. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. J Clin Oncol 2016; 34: 2157-2164.
- [24] Wu Y, Xu L, Yang P, Lin N, Huang X, Pan W, Li H, Lin P, Li B, Bunpetch V, Luo C, Jiang Y, Yang D, Huang M, Niu T and Ye Z. Survival prediction in high-grade osteosarcoma using radiomics of diagnostic computed tomography. EBioMedicine 2018; 34: 27-34.

- [25] Elkin EB, Hudis C, Begg CB and Schrag D. The effect of changes in tumor size on breast carcinoma survival in the U.S.: 1975-1999. Cancer 2005; 104: 1149-1157.
- [26] Carter CL, Allen C and Henson DE. Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases. Cancer 1989; 63: 181-187.
- [27] Wurdinger S, Herzog AB, Fischer DR, Marx C, Raabe G, Schneider A and Kaiser WA. Differentiation of phyllodes breast tumors from fibroadenomas on MRI. AJR Am J Roentgenol 2005; 185: 1317-1321.
- [28] Kim SG, Freed M, Leite A, Zhang J, Seuss C and Moy L. Separation of benign and malignant breast lesions using dynamic contrast enhanced MRI in a biopsy cohort. J Magn Reson Imaging 2017; 45: 1385-1393.
- [29] Cheng Z, Wu Z, Shi G, Yi Z, Xie M, Zeng W, Song C, Zheng C and Shen J. Discrimination between benign and malignant breast lesions using volumetric quantitative dynamic contrastenhanced MR imaging. Eur Radiol 2018; 28: 982-991.

- [30] Guo Y, Cai YQ, Cai ZL, Gao YG, An NY, Ma L, Mahankali S and Gao JH. Differentiation of clinically benign and malignant breast lesions using diffusion-weighted imaging. J Magn Reson Imaging 2002; 16: 172-178.
- [31] Suo S, Zhang K, Cao M, Suo X, Hua J, Geng X, Chen J, Zhuang Z, Ji X, Lu Q, Wang H and Xu J. Characterization of breast masses as benign or malignant at 3.0T MRI with whole-lesion histogram analysis of the apparent diffusion coefficient. J Magn Reson Imaging 2016; 43: 894-902.

Radiomics Features

Morphological features

1. Volume

The volume (V) of the tumor is determined by counting the number of pixels in the tumor region and multiplying this value by the voxel size.

2. Surface to Volume Ratio Surface Area

The surface area is calculated by triangulation (i.e. dividing the surface into connected triangles) and is defined as:

$$A = \sum_{i=1}^{N} \frac{1}{2} |a_i b_i \times a_i c_i|$$

Where is the total number of triangles covering the surface and, and are edge vectors of the triangles.

3. Eccentricity

The ellipsoid that best fits the tumor region is first computed using the framework of Li and Griffths (2004). The eccentricity is then given by $[1-a\times b/c^2]^{1/2}$, where c is the longest semi-principal axes of the ellipsoid, and a and b are the second and third longest semi-principal axes of the ellipsoid.

4. Solidity

Ratio of the number of voxels in the tumor region to the number of voxels in the 3D convex hull of the tumor region (smallest polyhedron containing the tumor region).

5. Surface to Volume Ratio

surface to volume ratio =
$$\frac{A}{V}$$

6. Compactness_1

compactness_1 =
$$\frac{V}{\sqrt{\pi} A^{\frac{2}{3}}}$$

7. Compactness_2

compactness_2 =
$$36\pi \frac{V^2}{A^3}$$

8. Sphericity

sphericity =
$$\frac{\pi^{\frac{1}{3}}(6V)^{\frac{2}{3}}}{A}$$

Histogram features

Histogram features calculated from the histogram of the tumor voxel intensities; Let X denote the threedimensional image matrix with N voxels. Their mathematical expressions are as below: 1. Skewness

$$H_{1} = \left(\sum_{i}^{N} \left(X(i) - \overline{X} \right) \right) / \sqrt{\frac{1}{N} \sum_{i}^{N} \left(X(i) - \overline{X} \right)^{2}} \right)^{3}$$

2. Kurtosis

$$H_{2} = \left(\sum_{i}^{N} \left(X(i) - \overline{X} \right) \right) / \sqrt{\frac{1}{N} \sum_{i}^{N} \left(X(i) - \overline{X} \right)^{2}} \right)^{4}$$

3. Variance

$$H_{3} = \sum_{i}^{N} (i - \overline{X})^{2} X(i)$$

Gray-Level Co-Occurrence Matrix based (GLCM) features:

A GLCM is defined as P(i, j; δ , α), a matrix with size N_g × N_g describing the second order joint probability function of an image, where the (i, j)th element represents the number of times the combination of intensity levels i and j occur in two pixels in the image, that are separated by a distance of δ pixels in direction α , and N_g is the number of discrete gray level intensities. In this article, prior to the computation of texture features, the full intensity range of the tumor region was quantized to a smaller number of gray levels 16. For angles = 0°, 45°, 90°, and 135°, we computed four values for each of the above texture measures. Each feature was computed using a distance of one pixel.

Let:

 $\mathsf{P}(\mathsf{i},\mathsf{j})$ be the co-occurrence matrix for an arbitrary δ and $\alpha,$

 μ be the mean of P(i, j),

 $\mu_{\rm v}$ be the mean of p,

 μ_v be the mean of p_v ,

- δ_x be the mean of δ_x ,
- δ_{v} be the mean of δ_{v} .

Their mathematical expressions are as below:

1. Energy

$$F_{1} = \sum_{i=1}^{N_{g}} \sum_{j=1}^{N_{g}} \left\{ P\left(i, j\right) \right\}^{2}$$

2. Contrast

$$F_{2} = \sum_{i=1}^{N_{g}} \sum_{j=1}^{N_{g}} |i - j|^{2} P(i j)$$

3. Correlation

$$F_{3} = \left(\sum_{i=1}^{N_{g}} \sum_{j=1}^{N_{g}} ijP(i, j) - \mu_{i}(i)\mu_{j}(j)\right) / \sigma_{x}(i)\sigma_{x}(j)$$

4. Entropy

$$F_{4} = -\sum_{i=1}^{N_{g}} \sum_{j=1}^{N_{g}} P(i, j) \log_{2} \left[P(i, j) \right]$$

5. Variance

$$F_5 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu)^2 P(i, j)$$

6. Homogeneity

$$F_{6} = \sum_{i=1}^{N_{g}} \sum_{j=1}^{N_{g}} \frac{1}{1 + \left|i - j\right|} P(i, j)$$

7. Autocorrelation

$$F_7 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} ijP(i, j)$$

8. Sum Average

$$F_{8} = -\sum_{i=2}^{2N_{g}} P_{x+y}(i) \log_{2} \left[P_{x+y}(i) \right]$$

9. Dissimilarity

$$F_{9} = \sum_{i=1}^{N_{g}} \sum_{j=1}^{N_{g}} |i - j| P(i, j)$$