Health and Technology (2020) 10:969–978
https://doi.org/10.1007/s12553-020-00452-3

ORIGINAL PAPER

Diagnostic value of apparent diffusion coefficient lesion texture biomarkers in breast MRI

Marialena I. Tsarouchi[1] · Georgios F. Vlachopoulos[1] · Anna N. Karahaliou[1] · Lena I. Costaridou[1]

Received: 4 February 2020 / Accepted: 15 June 2020 / Published online: 31 July 2020

Abstract

Quantitative assessment of breast intra-lesion heterogeneity in terms of contrast agent free Magnetic Resonance Imaging (MRI) approaches hold potential in breast cancer diagnosis. This study focuses on an Apparent Diffusion Coefficient (ADC) based approach, investigating the diagnostic role of 1st and 2nd order ADC statistics features, in differentiating benign from malignant breast lesion status. A total of 67 patients with 78 histologically verified breast lesions (40 benign and 38 malignant) was analyzed. ADC maps were generated for a slice representative of lesion largest diameter, considering intra Diffusion Weighted Imaging (DWI) sequence non rigid registration scheme. Lesion segments were defined by semi-automated Fuzzy C-Means (FCM) segmentation on high b-value diffusion images and propagated on ADC maps. 27 (11 1st order statistics and 16 2nd order statistics (texture) features were derived. To avoid overfitting a stepwise feature selection method was employed, while the discriminating ability of features was evaluated with univariate and multivariate Logistic Regression classification. The classification performance of the diagnostic model was evaluated by means of the Area Under Receiver Operating Characteristic curve (Az index). A combination of two features, one from 1st order statistics (25th Percentile) and one from 2nd order statistics, (texture Entropy), achieved high classification performance ($Az = 0.965 \pm 0.024$), suggesting both the diagnostic significance of 1st order statistics and texture biomarkers of the ADC map representation.

Keywords ADC map · Breast cancer diagnosis · DWI · Imaging lesion heterogeneity · Texture analysis

1 Introduction

Breast cancer is the most frequently diagnosed type of cancer and the leading cause of cancer death, in female population worldwide. Incidence rates of female breast cancer remain stable in the last few years (24.6%), while mortality rates follow a decreasing trend (15%) [1]. Investigating the underlying complexity (heterogeneity) of breast lesions is a key point in the era of personalized medicine. Heterogeneity characterized as inter- and/or intra- lesion heterogeneity, is evaluated with different genotypic assays, while phenotypic imaging appearance is emerging [2].

Although biopsy-proven biomarkers have shown to assess well breast lesion heterogeneity, their invasive nature limits tumor sampling. Thus, a more systematic and quantitative appreciation of breast intra-lesion heterogeneity has emerged considering whole lesion sampling. Advances in medical imaging analysis and pattern recognition led to high-throughput extraction of quantitative imaging features, termed radiomics [3] aim to provide an objective and reproducible imaging biomarker, reflecting intrinsic tumor properties and aggressiveness [4]. A major confound of radiomics analysis is the quantification of medical image texture. Texture is defined as the repetition of image patterns of the same intensity values, reflecting local spatial distribution variations of gray levels in the image matrix [5], and is particularly suitable to describe lesion imaging heterogeneity (phenotype). Extraction of texture-based features is numerically described by using statistical or structural approaches. Statistical approaches of texture analysis are the most widely utilized ones based on nth order statistics. Specifically, 1st order statistics, describe the way that gray levels are distributed over the pixels of an image, while 2nd order statistics, (herein called texture features), describe the local or regional relationships among image voxels (Gray Level Co-occurrence Matrix, Gray Level Run Length Matrix, Gray Level Size Zone Matrix) [6]. Radiomics

✉ Lena I. Costaridou
costarid@upatras.gr

[1] Department of Medical Physics, School of Medicine, University of Patras, 26500 Patras, Greece

approaches based on texture features, when combined with other clinical data and mined with sophisticated bioinformatics methods, such as machine learning algorithms, develop models that may potentially improve diagnostic, prognostic and predictive accuracy for response to treatment and survival outcomes.

Regarding breast cancer, radiomics approaches have been mainly investigated with Magnetic Resonance Imaging (MRI), while mammography is the gold standard in routine breast screening [7]. However, MRI plays a key role in detecting mammographically occult breast lesions [8] and its excellent soft tissue contrast makes MRI ideal for screening women with dense breast tissue [9]. In addition, MRI is a non-ionizing imaging modality, providing anatomical and functional tissue information with high temporal and spatial resolution by means of advanced MRI sequences, such as Dynamic Contrast Enhanced (DCE) and Diffusion Weighted Imaging (DWI).

DCE offers morphological and functional information related with tissue angiogenesis and vascular permeability through qualitative and/or quantitative analysis of signal intensity time curves of paramagnetic contrast agent uptake [10]. DCE is a powerful imaging technique in classifying breast lesions with high sensitivity ranging from 86% to 99%, however with low specificity ranging from 67% to 72% [11]. In addition, DCE has revealed promising radiomics signatures for breast lesions classification into histological and/or molecular subtypes [12, 13], while it is also utilized for evaluating treatment response [14, 15].

DWI, has been well integrated in routine breast MRI protocols with multiple clinical applications and holds promise in breast screening cost-effectively, less time-consuming and contrast agent free [16]. DWI depicts the mobility of water molecules in vivo, mapping tissue cellularity and microstructure. Quantitative analysis of DWI is mainly exploited through mono-exponential Apparent Diffusion Coefficient (ADC), which provides a measure of the degree of water diffusivity and describes the average area occupied by a water molecule per unit time (mm^2/s). When ADC is calculated on a voxel-by-voxel approach, a parametric map (ADC map) is generated [17]. As ADC reflects different levels of lesion aggressiveness, malignant breast lesions due to increased cellularity, demonstrate lower mean ADC values compared with the benign ones [16, 17].

Previously reported multiparametric MRI studies have exploited ADC as a supplementary imaging biomarker to DCE, in differentiating benign from malignant breast lesions increasing overall MRI specificity [18, 19]. Recently reported controversies in DCE referring to gadolinium contrast agent brain deposition have led to exploiting the diagnostic value of contrast-agent free MRI sequences, such as DWI, although characterized by reduced spatial resolution [20–23]. Specifically, the mean lesion ADC value has reported high diagnostic performance, (Area Under Receiver Operating Characteristic Curve, Az index, 0.94 [24]), however with significant overlapping of mean ADC cut-off values (0.90–1.76×10^{-3} mm^2/s), mainly attributed to its sensitivity to different acquisition protocols and choice of optimal b-values [25].

Thus, texture-based radiomics have been introduced aiming to assess intra-lesion DWI spatial heterogeneity analysis in an image intensity invariant approach. Previously reported studies have mainly exploited ADC histogram analysis suggesting that minimum and/or lower percentiles of ADC values hold promise in differentiating benign from malignant breast lesions [22, 26, 27]. Considering breast lesions classification into histological or molecular subtypes [28–31], as well as monitoring response to therapy [32, 33] several studies have exploited the added value of 2nd order (texture) and other radiomics features, although with controversial results.

The current study focuses on investigating if ADC-based 2nd order (texture) analysis is capable to contribute in diagnostic accuracy of differentiating benign from malignant breast lesions, in addition to 1st order ADC statistics analysis (Figure 1).

2 Methods and materials

2.1 Patient cohort

This retrospective study was approved by the Institutional Ethics Committee and informed consent was waived to patients undergoing routine clinical breast DWI-MRI examination, in a 3.0 T MRI scanner. The exclusion criteria in this study were: (i) non mass like lesions, (ii) unavailable lesions' histopathological confirmation/report, (iii) breast surgery or chemotherapy/radiotherapy treatment prior to MRI examination and (iv) poor image quality in DWI images. A total of 78 histologically verified breast lesions, originating from 67 female patients were analyzed in this study. 32 female patients (mean age ± standard deviation, range; 45.6 ± 13.5, 22–74 years) with 40 benign findings (mean size ± standard deviation, range; 111.1 ± 70.8, 39.0–394.0 mm^2) and 35 female patients (mean age ± standard deviation, range; 57.9 ± 12.3, 25–76 years) with 38 malignant breast lesions (mean size ± standard deviation, range; 175.3 ± 102.3, 46.0–446.0 mm^2) were included.

2.2 MRI acquisition

Breast MRI examinations were performed in a 3.0 T MRI scanner (Signa HDx; GE Healthcare, Milwaukee, WI, USA) using a dedicated bilateral four-element two-channel, phased array breast coil, with the patients in prone position. Breast holders were utilized in order to remove patients' motion

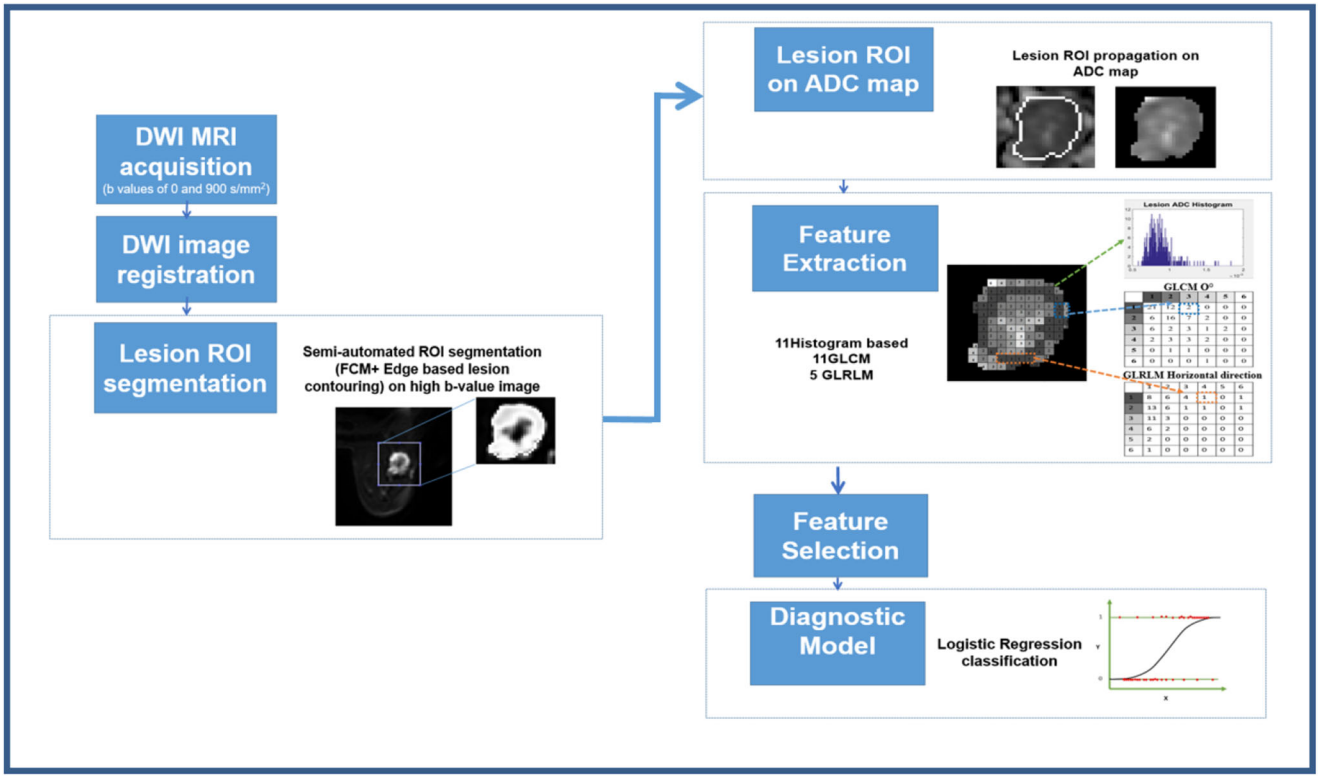

Fig. 1 Proposed radiomics analysis workflow.

artifacts. Breast imaging protocol included: (1) Axial T2-weighted fast spin echo (FSE) imaging sequence(T2–FSE, Time of Repetition (TR)/Time of Echo (TE), 3600/100 ms; slice thickness, 4 mm; spacing, 0 mm); (2) Axial short TI inversion recovery sequence (STIR) (TR/TE, 3875/90 ms; slice thickness, 4 mm; spacing, 0 mm); (3) Axial conventional diffusion-weighted echo-planar imaging sequence (DW-EPI, TR/TE, 6000/63.7 ms; slice thickness, 4 mm; spacing, 0 mm; matrix, 96 × 108; Field Of View (FOV), 360 mm × 360 mm). Sensitizing diffusion gradients were applied in three orthogonal directions (x, y, z) with b values of 0 and 900 s/mm^2, respectively. A three-dimensional (3D) fat suppressed (FS) T1-weighted VIBRANT dynamic sequence (flip angle, 10°; TR/TE, 5.8/2.1 ms; slice thickness, 1.2 mm; spacing, 1.2 mm; matrix, 512 × 512; FOV, 350 mm × 350 mm) was acquired once before and five times after intravenous injection of contrast medium (0.1 mmol/kg of gadopentate dimeglumine agent followed by a 20-mL flush of saline solution).

2.3 Image analysis

2.3.1 Intra DWI registration

In order to reduce image spatial misalignment due to patient motion or Eddy-current distortions, a registration algorithm [34] was adopted to map high-b value (b = 900 s/mm^2) diffusion weighted images to corresponding low b-value (b = 0 s/mm^2) diffusion weighted images [35]. The applied three-level multi-resolution registration scheme combines two levels of rigid and a final resolution level of b-spline transform, using a Gaussian smoothing pyramid. The applied optimizer is a stochastic gradient descent, while the similarity metric is Mutual Information [34]. The Elastix 4.5 [36] software for intensity-based medical image registration based on the open source software Insight Toolkit (ITK) version 4.0 [37] was used for implementation of the registration scheme.

2.3.2 Lesion ROI segmentation

A two-step semi-automated segmentation algorithm, based on Fuzzy C-Means clustering (FCM) [38] and edge-based lesion contouring, was adopted for breast lesion delineation. High b-value diffusion images were utilized for breast lesion segmentation, characterized by improved lesion contour/border identification compared with ADC map [30, 31]. Initially, an experienced radiologist defined a loose rectangular region of interest (ROI), containing the whole lesion, in slices representative of lesion's largest diameter, on high b-value diffusion images. FCM was applied on the defined ROI, to build the likelihood membership map (cluster number, 2; weighting exponent, 2; stop criteria, 0.0005; max iteration, 100). Subsequently, an edge-based segmentation model was applied for binarization of the membership map, followed by

morphological operations to derive the final lesion contour (MatLabR2017b Math Works, Natick, MA).

The segmented lesion ROI was transferred to ADC map (Fig. 1), generated according to the mono-exponential fitting model [17], employing in-house code (MatLab R2017b Math Works, Natick, MA).

2.3.3 Feature extraction

A total of 27 features were extracted from ADC lesion maps, including two feature sets: the first one was consisted of 11 1st order statistics and the second one was consisted of 16 texture features (2nd order statistics).

Prior to 1st order statistics analysis, ADC lesion maps were not subjected to gray level normalization, in order to preserve the biophysical meaning of ADC map intensity values. In case of ADC 1st order statistics entropy, a fixed bin number discretization scheme (8 bits/pixel) of the lesion ADC maps was considered, as a means of removing inter-lesion intensity variabilities. Features were extracted with in-house code (MatLab R2017b Math Works, Natick, MA).

In case of texture feature (2nd order statistics) extraction, ADC lesion maps were gray-level normalized within $[\mu \pm 3\sigma]$, where μ is the mean of pixel intensity values within lesion ROI and σ is the standard deviation, and discretized to 6 bits/pixel in order to minimize contrast and brightness variation and reduce computational cost of texture features [39]. 11 features were calculated from Gray Level Co-occurrence Matrices-GLCM, considering a distance of one pixel ($d = 1$) and four directions (0°, 45°, 90° and 135°). Furthermore, the 5 features were calculated from Gray Level Run Length Matrices-GLRLM, considering four directions (Horizontally, vertically, 45°and 135°). Features extracted from GLCM and GLRLM were averaged over four directions to obtain rotationally invariant features. Texture features were extracted employing the publicly available MaZda package v 4.6 [40].

2.3.4 Statistical analysis

The Shapiro-Wilk test was utilized ($p < 0.05$) to test normality of features' distribution (IBM SPSS Statistics software 24.0, 2016) [41]. As the extracted features follow a non-normal distribution, the non-parametric Mann-Whitney U test was used in MatLab R2017b (Math Works, Natick, MA), to assess existence of statistically significant difference of individual features between benign and malignant lesion status. Bonferroni correction for multiple tests was considered to adjust the level of significance [42].

In order to reduce feature dimensionality and avoid redundant information, a stepwise feature selection method was adopted based on the minimization of Wilks-Λ [43]. The procedure employs stepwise feature entry and feature removal from the model according to predefined thresholds (probabilities). In this study, the probability-to-enter and the probability-to-remove were set to 0.05 and 0.10, respectively.

A machine learning classification model, Logistic Regression (LR), was exploited to further assess univariate diagnostic performance of each feature, as well as the performance of the combined features derived from the above feature selection procedure. The above machine learning model was implemented in the publicly available data mining software, Weka [44], using the 10-fold cross-validation training/testing methodology, accounting for a robust internal validation method in case of datasets of limited size [45, 46]. The classification performance was evaluated by means of Az index. In case of multivariate logistic regression, the bias of collinearity between features was evaluated using variance inflation factor (VIF) and the corresponding threshold was set at 3.

3 Results

Investigating the discriminative ability of individual DWI sequence, a total of 27 features were exploited from ADC map, for each breast lesion. Median and interquartile ADC values, as well as discriminating ability (in terms of Az index ± Standard Error (SE), of the univariate Logistic Regression classifier) of individual 1st statistics and 2nd order (texture) features, are provided on Tables 1 and 2, respectively.

Among 1st order statistics features, 25th Percentile achieved the highest classification performance ($Az = 0.943 \pm 0.031$), while among individual 2nd order (texture) features, Entropy achieved the highest classification performance ($Az = 0.738 \pm 0.052$).

Following stepwise feature selection method, the selected feature subset consisted of 2 quantitative imaging descriptors, one descriptive parameter of signal intensity distribution (25th Percentile) and one texture feature (Entropy 2nd order). The discriminating ability of the selected feature subset was further assessed by means of multivariate machine learning Logistic Regression classifier, achieving the highest classification performance ($Az = 0.965 \pm 0.024$), with sensitivity, specificity and accuracy of 92.3%, respectively.

Figures 2 and 3, illustrate a typical benign and a typical malignant breast lesion respectively. Corresponding histograms highlight the difference with respect to mean and 25th percentile of ADC values, as well as the value of 2nd order statistics (texture entropy) in discriminating benign from malignant breast lesions. Figure 4 illustrates a non-typical malignant breast lesion with its corresponding histogram. While the mean ADC value is capable of differentiating benign from a typical appearing malignant breast lesion, however, in case of equivocal (non-typical) breast lesions the reported high mean ADC value may result in misdiagnosis. In this case, the combination of the 25th percentile (1st order statistics) and entropy (2nd order statistics), with the latter texture measure reflecting lesion heterogeneity, resulted in high diagnostic accuracy (Figure 5).

Table 1 First order statistics ADC features

1st order statistics ADC features	Benign lesions ($n = 40$) Median (interquartile range) values	Malignant lesions ($n = 38$)	Mann-Whitney U test p value	Logistic Regression Az ± SE
Mean ($\times 10^{-3}$ mm^2/s)	1.600 (0.872, 2.300)	0.888 (0.672, 2.100)	$4{,}69 \times 10^{-11}$*	0.924 ± 0.031
Standard Deviation ($\times 10^{-3}$ mm^2/s)	0.136 (0.045, 0.430)	0.164 (0.066, 0.685)	0.224	0.559 ± 0.056
Skewness	−0.236 (−1.386, 1.711)	0.668 (−0.975, 2.198)	1.36×10^{-3}*	0.748 ± 0.051
Kurtosis	3.225 (1.712, 8.348)	3.380 (1.665, 9.646)	0.315	0.551 ± 0.057
Entropy	5.220 (3.877, 6.257)	5.975 (5.099, 7.066)	6.93×10^{-7}*	0.821 ± 0.047
Minimum ($\times 10^{-3}$ mm^2/s)	1.200 (0.422, 2.000)	0.593 (0.240, 1.200)	1.16×10^{-9}*	0.888 ± 0.039
Maximum ($\times 10^{-3}$ mm^2/s)	1.900 (1.100, 2.500)	1.400 (0.980, 3.200)	2.38×10^{-7} *	0.831 ± 0.046
25th percentile ($\times 10^{-3}$ mm^2/s)	1.500 (0.766, 2.200)	0.785 (0.584, 1.900)	8.62×10^{-12} *	0.943 ± 0.031
50th percentile ($\times 10^{-3}$ mm^2/s)	1.600 (0.864, 2.400)	0.873 (0.646, 2.100)	2.35×10^{-11}*	0.926 ± 0.030
75th percentile ($\times 10^{-3}$ mm^2/s)	1.700 (0.969, 2.500)	0.988 (0.725, 2.500)	6.60×10^{-11} *	0.920 ± 0.033
Range ($\times 10^{-3}$ mm^2/s)	0.607 (0.016, 1.700)	0.733 (0.040, 2.300)	0.045	0.623 ± 0.055

*$p < 0.0045$ indicates statistically significant difference, Bonferroni correction for adjusting significance level a = 0.05/11)

4 Discussion

DCE or multiparametric studies have revealed radiomics signatures to assess breast lesions heterogeneity [10, 12–15, 18, 19], however controversies due to minimally invasive nature and contrast agent safety [20] have emerged, addressing the need of contrast agent free MRI approaches (i.e. DWI). Towards this direction, the current study investigates the

Table 2 Second order statistics (texture) ADC features

2nd order statistics ADC features	Benign lesions ($n = 40$) Median (interquartile range) values	Malignant lesions ($n = 38$)	Mann-Whitney U test p value	Logistic Regression Az ± SE
Angular Second Moment	0.009 (0.003, 0.023)	0.007 (0.004, 0.016)	2.30×10^{-3}*	0.737 ± 0.052
Contrast	40.702 (17.203, 82.864)	40.541 (10.332, 104.962)	0.822	0.647 ± 0.057
Correlation	0.797 (0.474, 0.911)	0.791 (0.381, 0.954)	1	0.550 ± 0.057
Sum Of Squares	95.552 (73.785, 115.205)	98.988 (76.234, 115.899)	0.058	0.612 ± 0.056
Inverse Difference Moment	0.186 (0.146, 0.290)	0.190 (0.135, 0.347)	0.525	0.524 ± 0.057
Sum Average	65.195 (60.229, 68.373)	63.776 (59.855, 66.094)	7.14×10^{-3} *	0.709 ± 0.053
Sum Variance	345.594 (235.742, 415.363)	346.615 (234.680, 438.581)	0.916	0.605 ± 0.057
Sum Entropy	1.608 (1.345, 1.813)	1.663 (1.396, 1.819)	2.75×10^{-2}*	0.691 ± 0.054
Entropy	2.073 (1.699, 2.557)	2.242 (1.817, 2.531)	1.81×10^{-3}*	0.738 ± 0.052
DifVariance	14.000 (6.275, 37.276)	16.822 (4.333, 45.917)	0.195	0.556 ± 0.056
DifEntropy	1.057 (0.879, 1.161)	1.080 (0.844, 1.191)	0.192	0.568 ± 0.056
RLNonUni	76.645 (33.529, 323.592)	119.518 (41.390, 314.063)	3.26×10^{-3}*	0.736 ± 0.053
GLevNonUni	3.659 (1.858, 12.414)	4.959 (2.036, 18.325)	3.34×10^{-2}*	0.689 ± 0.054
LngREmph	1.223 (1.113, 1.530)	1.230 (1.091, 1.629)	0.924	0.529 ± 0.057
ShrtREmph	0.951 (0.917, 0.972)	0.952 (0.899, 0.977)	0.760	0.525 ± 0.057
Fraction	0.935 (0.883, 0.945)	0.936 (0.858, 0.971)	0.760	0.521 ± 0.057

* $p < 0.0035$ indicates statistically significant difference, Bonferroni correction for adjusting significance level a = 0.05/16)

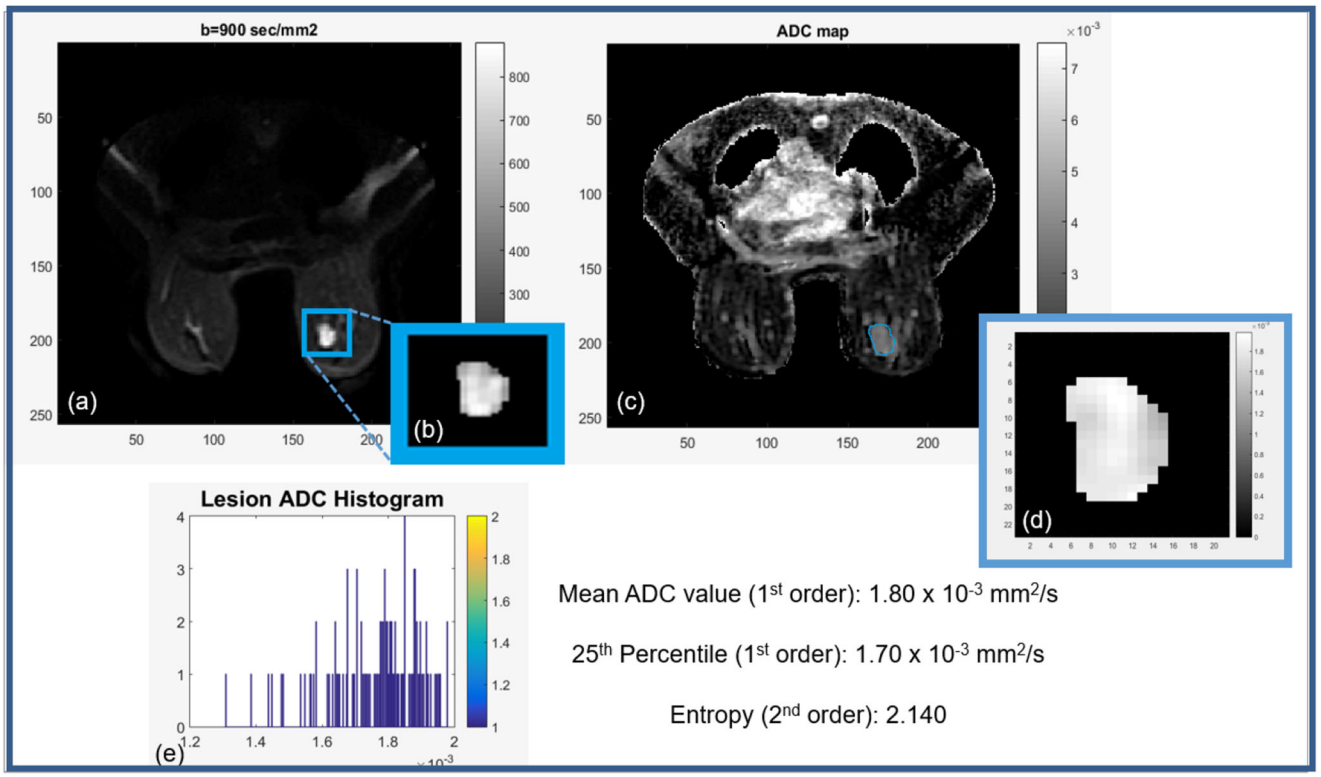

Fig. 2 Illustration of a typical benign breast lesion (**a**) High b-value (b = 900 s/mm^2) diffusion image, (**b**) lesion ROI on high b-value diffusion image, (**c**) ADC map, (**d**) lesion ROI propagation on ADC map and (**e**) lesion ADC histogram with mean ADC value of 1.80×10^{-3} mm^2/s, 25th percentile 1.70×10^{-3} mm^2/s, and 2nd order (texture) entropy 2.140

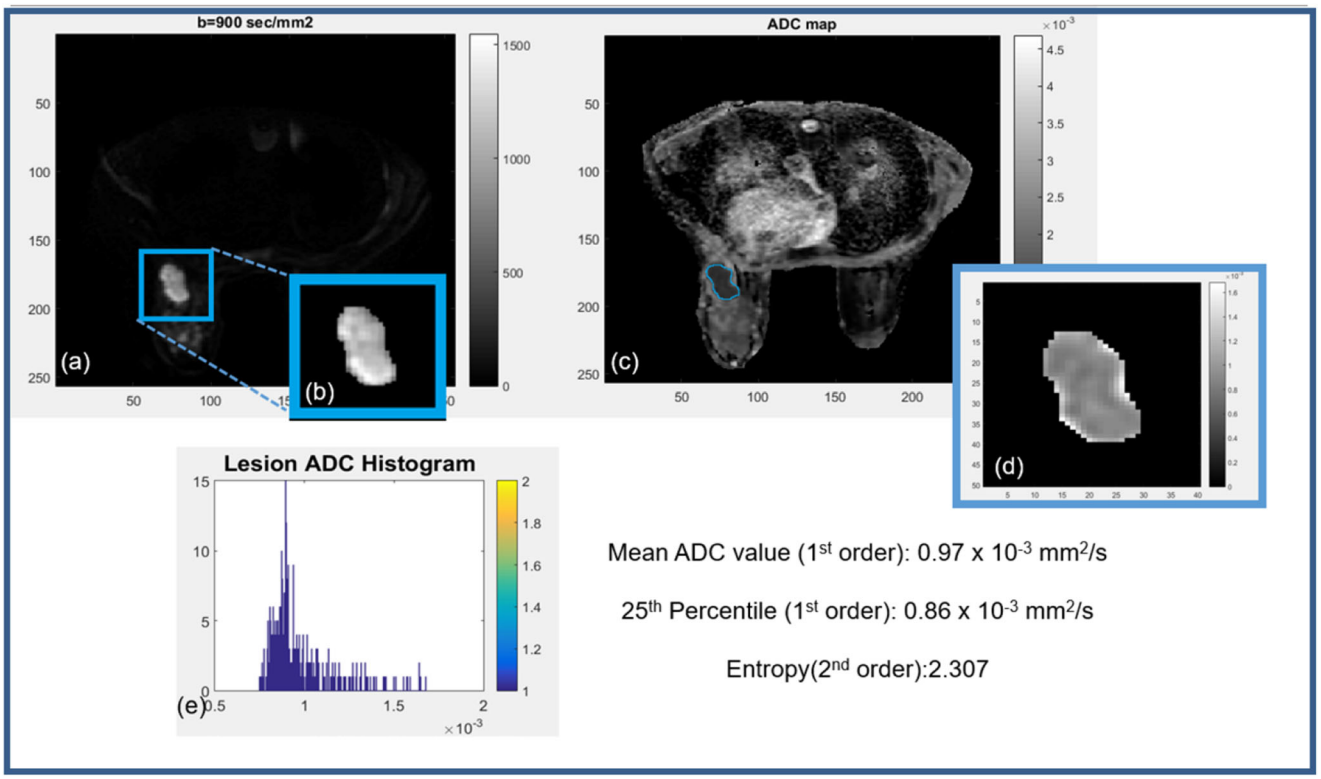

Fig. 3 Illustration of a typical malignant breast lesion (**a**) High b-value (b = 900 s/mm^2) diffusion image, (**b**) lesion ROI on high b-value diffusion image, (**c**) ADC map, (**d**) lesion ROI propagation on ADC map and (**e**) lesion ADC histogram with mean ADC value of 0.97×10^{-3} mm^2/s, 25th percentile 0.86×10^{-3} mm^2/s, and 2nd order (texture) entropy 2.307

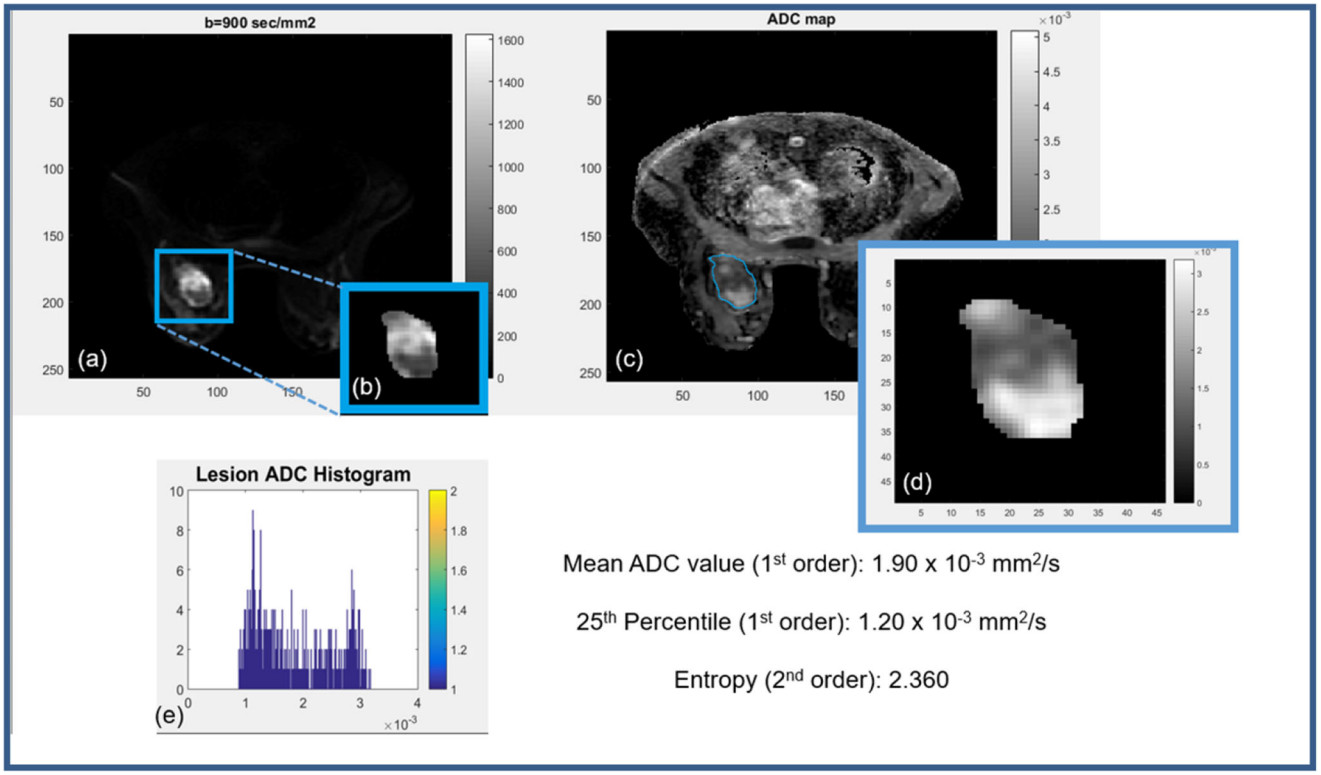

Fig. 4 Illustration of a non-typical malignant breast lesion (**a**) High b-value (b = 900 s/mm^2) diffusion image, (**b**) lesion ROI on high b-value diffusion image, (**c**) ADC map, (**d**) lesion ROI propagation on ADC map and (**e**) lesion ADC histogram with mean ADC value of 1.90 × 10^{-3} mm^2/s, 25th percentile 1.20 × 10^{-3} mm^2/s, and 2nd order (texture) entropy 2.360

ability of ADC texture-based biomarkers to capture spatially breast intralesion heterogeneity and to assess prediction of clinical diagnosis.

Regarding breast cancer diagnosis, previously reported studies have demonstrated the ability of radiomics in differentiating malignant from benign breast lesions with

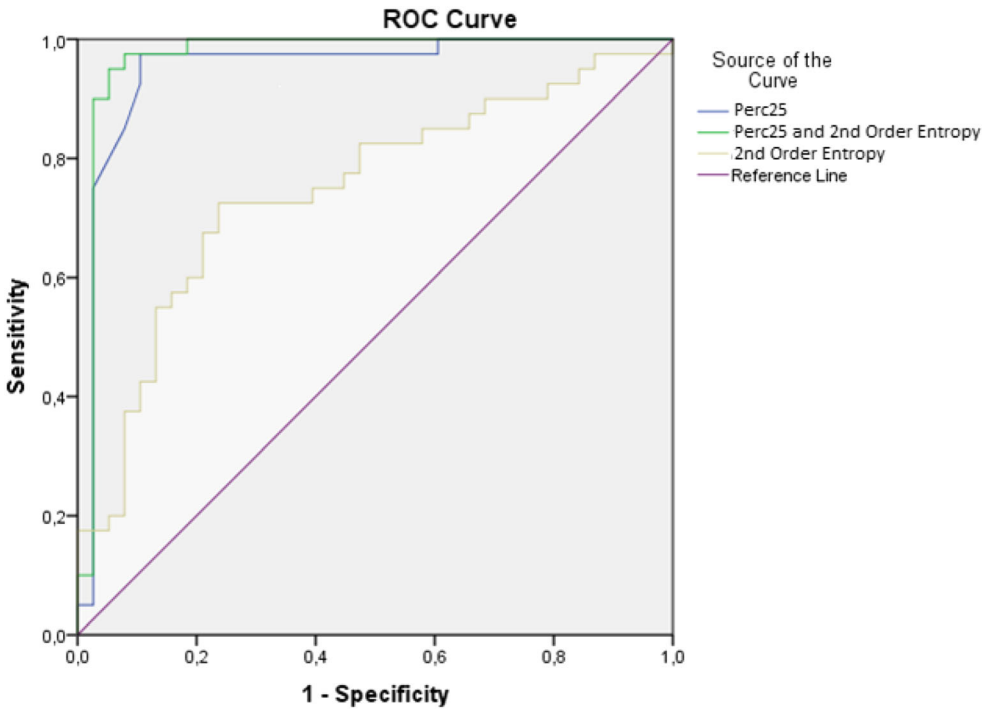

Fig. 5 ROC curves corresponding to the best individually performing 1st order statistics feature (25th percentile with Az = 0.943 ± 0.031), to the best individually performing 2nd order statistics texture feature (Entropy with Az = 0.738 ± 0.052) and the selected ADC feature subset (25th percentile [1st order] + Entropy [2nd order (texture)], with Az 0.965 ± 0.024)

contrast agent free MRI approaches [21–23, 28, 29]. Especially, Bickelhaupt et al. [22], investigated radiomics signatures incorporating T2-weighted, conventional DWI and DWI with background suppression (DWIBS) (b values of 0 and 1500 s/mm^2) and corresponding ADC maps, providing predictive models of breast cancer. Their results have demonstrated that radiomics, including 1st order statistics combined to 2nd order statistics (texture) analysis, extracted from high b-value DWIBS sequence yielded higher performance (Az 84.2%) than mean ADC individual feature (Az 77.4%). However, the T2 shine-through effect may introduce variabilities depending on the choice of high b-value.

Reported studies [26, 27] have mainly exploited ADC texture analysis, in an effort to minimize variabilities introduced by the choice of low and high b-values, utilized in different acquisition protocols. Specifically, Suo et al. [26] exploiting 1st order statistics analysis, reported that minimum and lower percentiles, as well as 1st order statistics entropy of ADC values, achieved high diagnostic performance (Az 0.893). The optimal/selected feature subset of the current study, including one 1st order statistics feature (25th percentile of lesion ADC map) and one 2nd order statistics (texture) feature (2nd order Entropy of lesion ADC map), demonstrating the highest classification performance (Az 0.965 ± 0.024) is consistent with Suo et al. [26] results, especially regarding with the 25th percentile of ADC (Az 0.943 ± 0.031), representing the most solid/aggressive part of the lesion.

The current study highlights the added value offered by 2nd order statistics (texture) entropy (Az 0.738 ± 0.052) in improving classification performance to (Az 0.965 ± 0.024), possibly representing local spatial intralesion heterogeneity of the most solid/aggressive part of the lesion. The role of 2nd order (texture) entropy is mainly observed in case of breast cancer prognosis, including molecular subtypes and hormonal receptor classification [29–31] and response to therapy studies [32, 33].

While results of the current study are encouraging towards the role of texture analysis of ADC in capturing whole intralesion heterogeneity, DWI acquisition protocols and choice of b-values need to be further validated and standardized according to QIBA (The Radiology Society of North America's Quantitative Imaging Biomarkers Alliance) [47], prior to its clinical application. Further acknowledging the importance of pre-processing in the workflow of radiomics analysis [48], future efforts will also focus on comprehensive investigation of the effect of normalization (i.e. histogram matching, z-score, deep learning) and discretization (i.e. fixed bin size, fixed bin width) [49] on the performance of the proposed analysis scheme. The incorporation of image normalization is necessary for ensuring reproducibility of the reported results on heterogeneous datasets, by removing variations induced by inter-scanner and intra-scanner technical specifications and imaging acquisition protocols [45, 47, 50, 51].

The current study is limited by a relatively small patient cohort, as well as variabilities introduced mainly by the semi-automated segmentation step, while the accuracy of image registration must be quantitatively assessed. Although, this study addresses the need of contrast agent free MRI approaches, focusing on assessing the discriminating ability of ADC based 1st order statistics and texture analysis in breast lesion diagnosis, further prospective steps should consider the contribution of DWI in the frame of multiparametric breast MRI.

5 Conclusion

Quantitative assessment of breast intra-tumor heterogeneity, in terms of ADC 1st order statistics combined to ADC texture analysis, demonstrated that a multivariable diagnostic model, consisting of two informative features (ADC 25th Percentile and ADC texture Entropy), improves diagnostic accuracy with respect to reported state of the art. Further studies are needed to warrant such contrast agent free MRI approaches into breast imaging clinical practice.

Acknowledgments Special thanks to the Department of Radiology, University Hospital of Larissa, University of Thessaly, Greece, for contributing in this work.

Funding information Support by Operational Program "Human Resources Development, Education and Lifelong Learning" and is co-financed by the European Union (European Social Fund) and Greek national funds (MIS: 5005772).

Compliance with ethical standards

Conflict of interest None of the authors have any conflicts of interest to declare.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel R, Torre L, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68:394–424.
2. JPB O'C, Rose CJ, Waterton CJ, RAD C, GJM P, Jackson A. Imaging intratumor heterogeneity: role in therapy response, resistance, and clinical outcome. Clin Cancer Res. 2015;21(2):249–57.
3. Aerts HJ, Velazquez ER, Leijenaar RTH, Parmar C, Grossman P, Carvalho S, et al. Decoding tumor phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun. 2014;3(1):4006.
4. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol. 2017;14(12):749–62.

5. Castellano G, Bonilha L, Li LM, Cendes F. Texture analysis of medical images. Clin Radiol. 2004 Dec;59(12):1061–9.
6. Yip SSF, Aerts HJWL. Applications and limitations of radiomics. Phys Med Biol. 2016;61:R150–66.
7. Valdora F, Houssami N, Rossi F, Calabrese M, Tagliafico AS. Rapid review: radiomics and breast cancer. Breast Cancer Res and Treat. 2018;169(2):217–29.
8. Alcusky M, Philpotts L, Bonafede M, Clarke J, Skoufalos A. The patient burden of screening mammography recall. J Women's Health. 2014;23(S1):S11–9.
9. Bakker MF, De Lange SV, Pijnappel RM, Mann RM, Peeters PHM, Monninkhof EM, et al. Supplemental MRI screening for women with extremely dense breast tissue. N Engl J Med. 2019;381:2091–102.
10. Karahaliou A, Vassiou K, Arikidis NS, Skiadopoulos S, Kanavou T, Costaridou L. Assessing heterogeneity of lesion enhancement kinetics in dynamic contrast-enhanced MRI for breast cancer diagnosis. Br J Radiol. 2010;83(988):296–306.
11. Leithner D, Wengert GJ, Helbich TH, Thakur S, Ochoa-Albiztegui RE, Morris EA, et al. Clinical role of breast MRI now and going forward. Clin Radiol. 2018;73(8):700–14.
12. Holli-Helenius K, Salminen A, Rinta-Kiikka I, Koskivuo I, Brück N, Boström P, et al. MRI texture analysis in differentiating luminal a and luminal B breast cancer molecular subtypes-a feasibility study. BMC Med Imaging. 2017;17:69.
13. Jiang Z, Song L, Lu Hand Jiandong Y. The potential use of DCE-MRI texture analysis to predict HER2 2+ status. Front Oncol. 2019;9:242.
14. Cain EH, Saha A, Harowicz MR, Marks JR, Marcom PK, Mazurowski MA. Multivariate machine learning models for prediction of pathologic response to neoadjuvant therapy in breast cancer using MRI features: a study using independent validation set. Breast Cancer Res and Treat. 2019;173(2):455–63.
15. Jahani N, Cohen E, Kang Hsieh M, Weinstein SP, Pantalone L, Hylton N, et al. Prediction of treatment response to neoadjuvant chemotherapy for breast cancer via early changes in tumor heterogeneity captured by DCE-MRI registration. Scientific Reports, Nature. 2019;(9):12114.
16. Partridge SC, Nissan N, Rahbar H, Kitsch AE, Sigmund EE. Diffusion-weighted breast MRI: clinical applications and emerging techniques. J Magn Reson Imaging. 2017;45(2):337–55.
17. Camps-Herrero J.. Diffusion-weighted imaging of the breast: current status as an imaging biomarker and future role. 2019 BJR open, 1:20180049.
18. Jiang X, Xie F, Liu L, Peng Y, Cai H, Li L. Discrimination of malignant and benign breast masses using automatic segmentation and features extracted from dynamic contrast enhanced and diffusion weighted MRI. Oncol Lett. 2018;16(2):1521–8.
19. Zhang M, Horvat JV, Bernard-Davila B, Marino MA, Leithner D, Ochoa-Albiztegui RE, et al. Multiparametric MRI model with dynamic contrast-enhanced and diffusion-weighted imaging enables breast Cancer diagnosis with high accuracy. J Magn Reson Imaging. 2019;49(3):864–74.
20. McDonald RJ, McDonald JS, Kallmes DF, Jentoft ME, Murray DL, Thielen KR, et al. Intracranial gadolinium deposition after contrast-enhanced MR imaging. Radiology. 2015;275(3):772–82.
21. Baltzer PAT, Bickel H, Spick C, Wengert G, Woitek R, KapetasP CP, et al. Potential of noncontrast magnetic resonance imaging with diffusion-weighted imaging in characterization of breast lesions: intraindividual comparison with dynamic contrast-enhanced magnetic resonance imaging. Investig Radiol. 2018;53(4):229–35.
22. Bickelhaupt S, Paech D, Kickingereder P, Steudle F, Lederer W, Daniel H, et al. Prediction of malignancy by a radiomic signature from contrast agent-free diffusion MRI in suspicious breast lesions found on screening mammography. J Magn Reson Imaging. 2017;46(2):604–16.
23. Dietzel M, Ellman S, Schulz-Wendtland R, Clauser P, Wenkel E, Uder M, et al. Breast MRI in the era of diffusion weighted imaging: do we still need signal-intensity time curves? Eur Radiol. 2020;30(1):47–56.
24. Shi RY, Yao QY, WuLM XJR. Breast lesions: diagnosis using diffusion weighted imaging at 1.5 T and 3.0T-systematic review and meta-analysis. Clin Breast Cancer. 2018;18(3):e305–20.
25. Partridge SC, Amornsiripanitch N. DWI in the assessment of breast lesions. Top Magn Reson Imaging. 2017;26(5):201–9.
26. Suo S, Zhang K, Cao M, Suo X, Hua J, Geng X, et al. 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(4):894–902.
27. Liu C, Wang K, Li X, Zhang J, Ding J, Spuhler K, et al. Breast lesion characterization using whole-lesion histogram analysis with stretched-exponential diffusion model. J Magn Reson Imaging. 2018;47(6):1707–10.
28. Tsarouchi MI, Vlachopoulos GF, Karahaliou AN and Costaridou LI. Diffusion weighted magnetic resonance imaging texture biomarkers for breast cancer diagnosis. H. J. Neves N. De Carvalho (eds) XV mediterranean conference on medical and biological engineering and computing – MEDICON 2019. MEDICON 2019. IFMBE Proceedings, vol 76. Springer, Cham, 2019.
29. Parekh VS, Jacobs MA. Intergraded radiomic framework for breast cancer and tumor biology using advanced machine learning and multiparametric MRI. NPJ Breast Cancer. 2017;14(3):43.
30. Suo S, Cheng F, Cao M, Kang J, Wang M, Hua J, et al. Multiparametric diffusion-weighted imaging in breast lesions: association with pathologic diagnosis and prognostic factors. J Magn Reson Imaging. 2017;46(3):740–50.
31. Leithner D, Bernard-Davila B, Martinez DF, Horvat JV, Jochelson MS, Marino MA, Avendano D, Ochoa-Albiztegui RE, Sutton EJ, Morris EA, Thakur SB, Pinker K.. Radiomic signatures derived from diffusion-weighted imaging for the assessment of breast cancer receptor status and molecular subtypes. Mol Imaging Biol, 2019.
32. Li W, Newitt DC, Wilmes LJ, Jones EF, Arasu V, Gibbs J, et al. Additive value of diffusion-weighted MRI in the I-SPY 2 TRIAL. J Magn Reson Imaging. 2019;50(6):1742–53.
33. Kim Y, Kim SH, Lee HW, Song BJ, Kang BJ, Lee A, et al. Intravoxel incoherent motion diffusion weighted MRI for predicting response to neoadjuvant chemotherapy in breast cancer. Magn Reson Imaging. 2018;48:27–33.
34. Vlachopoulos G, KorfiatisP SS, KazantziA KC, Pratikakis I, Costaridou L. Selecting registration schemes in case of interstitial lung disease follow-up in CT. Med Phys. 2015;42(8):4511–25.
35. Arlinghaus LR, Welch EB, Chakravarthy AB, Xu L, Farley JS, Abramson VG, et al. Motion correction in diffusion-weighted MRI of the breast at 3.0T. J Magn Reason Imaging. 2011;33(5):1063–70.
36. Klein S, Staring M, Murphy K, Viergever M, Pluim J. Elastix: a toolbox for intensity-based medical image registration. IEEE Trans Med Imaging. 2010;29(1):196–205.
37. Ibanez L, Schroeder W, Ng L, Cates J. The ITK software guide, 2nd ed. (Kitware, Clifton Park, NY, 2005).
38. Chen W, Giger ML, Bick U. A fuzzy c-means (FCM)-based approach for computerized segmentation of breast lesions in dynamic contrast-enhanced MR images. Acad Radiol. 2006;13(1):63–72.
39. Collewet G, Strzelecki M. andMarriette F. influence of MRI acquisition protocols and image intensity normalization methods on texture classification. Magn Reson Imaging. 2004;22(1):81–91.
40. Szczpinski PM, Strzelecki M, Materka A, Klepaczko A. MaZda- a software package for image texture analysis. Comput Methods Prog Biomed. 2009;94(1):66–76.
41. Field AP. Discovering statistics using SPSS. Los Angeles: Sage Publications; 2009.

42. Wright SP. Adjusted p-values for simultaneous inference. Biometrics. 1992;43:1005–13.
43. Draper N, Smith H. Applied regression analysis. 2nd ed. New York: Wiley; 1981.
44. Frank E, Hall MA, and Witten IH. The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016.
45. Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med. 2015 Jan 6;162(1): W1–73.
46. Van Houwelingen HC, Sauerbrei W. Cross-validation, shrinkage and variable selection in linear regression revisited. Open J Statist. 2013;3:79–102.
47. Shukla-Dave A, Obuchowski NA, Chenevert TL, Jambawalikar S, Schwartz LH, Malyarenko D, et al. Quantitative imaging biomarkers alliance (QIBA) recommendations for improved precision of DWI and DCE-MRI derived biomarkers in multicenter oncology trials. J Magn Reson Imaging. 2019;49(7):e101–21.
48. Zwanenburg A, Vallieres M, Abdalah MA, Aerts HJWL, Andrearczyk V,Apte A, et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. Radiology, vol. 295, no.2, 2020.
49. Duron L, Balvay D, VandePerre S, Bouchouicha A, Savatovsky J, Sadik JC, et al. Gray-level discretization impacts reproducible MRI radiomics texture features. PLoS One. 2019;14(3):e0213459.
50. Zhu J, Zhang J, Gao JY, Li JN, Yang DW, Chen M, et al. Apparent diffusion coefficient normalization of normal liver: will it improve the reproducibility of diffusion-weighted imaging at different MR scanners as a new biomarker? Medicine (Baltimore). 2017;96(3): e5910.
51. Isaksson LJ, Raimondi S, Botta F, Pepa M, Gugliandolo SG, De Angelis SP, et al. Effects of MRI image normalization techniques in prostate cancer radiomics. Physica Medica. 2020;71.