Distinguishing benign and malignant breast tumors: preliminary comparison of kinetic modeling approaches using multi-institutional dynamic contrast-enhanced MRI data from the International Breast MR Consortium 6883 trial

Anna G. Sorace
Savannah C. Partridge
Xia Li
Jack Virostko
Stephanie L. Barnes
Daniel S. Hippe
Wei Huang
Thomas E. Yankeelov

Distinguishing benign and malignant breast tumors: preliminary comparison of kinetic modeling approaches using multi-institutional dynamic contrast-enhanced MRI data from the International Breast MR Consortium 6883 trial

Anna G. Sorace,^{a,b,c} Savannah C. Partridge,^d Xia Li,^e Jack Virostko,^{a,b} Stephanie L. Barnes,^{c,f} Daniel S. Hippe,^d Wei Huang,^{g,h} and Thomas E. Yankeelov^{a,b,c,f,*}

^aUniversity of Texas at Austin, Department of Diagnostic Medicine, Austin, Texas, United States

^bUniversity of Texas at Austin, Livestrong Cancer Institutes, Austin, Texas, United States

^cUniversity of Texas at Austin, Department of Biomedical Engineering, Austin, Texas, United States

^dUniversity of Washington, Department of Radiology, Seattle, Washington, United States

^eGE Global Research, Niskayuna, New York, United States

^fUniversity of Texas at Austin, Institute for Computational and Engineering Sciences, Austin, Texas, United States

^gOregon Health and Science University, Advanced Imaging Research Center, Portland, Oregon, United States

^hOregon Health and Science University, Knight Cancer Institute, Portland, Oregon, United States

Abstract. Comparative preliminary analysis of dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) data collected in the International Breast MR Consortium 6883 multicenter trial was performed to distinguish benign and malignant breast tumors. Prebiopsy DCE-MRI data from 45 patients with suspicious breast lesions were obtained. Semiquantitative mean signal-enhancement ratio (SE_{mean}) was calculated for all lesions, and quantitative pharmacokinetic parameters K^{trans} , k_{ep} , and v_e , were calculated for the subset with available T_1 maps ($n = 35$). Diagnostic performance was estimated for DCE-MRI parameters and compared to standard clinical MRI assessment. Quantitative and semiquantitative metrics discriminated benign and malignant lesions, with receiver operating characteristic area under the curve (AUC) values of 0.71, 0.70, and 0.82 for K^{trans} , k_{ep} , and SE_{mean} , respectively ($p < 0.05$). At equal 94% sensitivity, the specificity and positive predictive value of SE_{mean} (53% and 63%, respectively) and K^{trans} (42% and 58%) were higher than clinical MRI assessment (32% and 54%). A multivariable model combining SE_{mean} and clinical MRI assessment had an AUC value of 0.87. Quantitative pharmacokinetic and semiquantitative analyses of DCE-MRI improves discrimination of benign and malignant breast tumors, with our findings suggesting higher diagnostic accuracy using SE_{mean} . SE_{mean} has potential to help reduce unnecessary biopsies resulting from routine breast imaging. © 2018 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JMI.5.1.011019]

Keywords: kinetic analysis; K^{trans} ; k_{ep} ; signal-enhancement ratio; breast cancer; dynamic contrast-enhanced MRI.

Paper 17182SSRR received Jun. 19, 2017; accepted for publication Dec. 18, 2017; published online Jan. 22, 2018.

1 Introduction

Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) is a noninvasive imaging method performed by serial acquisition of heavily T_1 -weighted images before, during, and after an intravenous injection of a contrast agent.¹ As contrast agent perfuses into the tissue of interest (TOI), it affects the native longitudinal relaxation times, and thus signal intensity, to a degree determined by the delivery, accumulation, and washout of the contrast agent. DCE-MRI is an integral part of a standard-of-care breast MRI exam. In the clinical setting, greater emphasis is placed on acquiring high spatial resolution data to maximize ability to assess anatomic and morphologic detail, at the expense of temporal resolution.² Under this constraint, the breast lesion signal intensity-time course [at either the voxel

or region of interest (ROI) level] is generally assessed only qualitatively or semiquantitatively.^{3–5} However, if high temporal resolution data are acquired, quantitative analysis^{6,7} using pharmacokinetic models^{8,9} can also be performed to assess specific tissue properties, including blood flow, vascular permeability, and tissue extracellular volume fractions.¹⁰

The American College of Radiology (ACR) MRI Breast Imaging Reporting and Data System (BI-RADS) lexicon has become the standardized methodology to assess the likelihood of malignancy for MRI-detected lesions. BI-RADS combines morphologic feature analysis, including lesion type, shape, and margin, and internal enhancement pattern, with qualitative assessment of enhancement kinetics of initial uptake and delayed enhancement curve shape.² Contrast-enhanced MRI of the breast has a high sensitivity for breast cancer detection and is the most sensitive technique for screening high-risk women^{11–14}

*Address all correspondence to: Thomas E. Yankeelov, E-mail: thomas.yankeelov@utexas.edu

and detecting contralateral¹⁵ or multifocal disease¹⁶ in patients with recently diagnosed breast cancer. However, overlap in BI-RADS clinical MRI assessment of benign and malignant breast lesions using DCE-MRI produces high false positive rates and leads to unnecessary biopsies of many benign lesions.

It is well-established that assessment of DCE-MRI enhancement kinetics can help differentiate between benign and malignant breast lesions.^{6,17,18} Semiquantitative analysis of DCE-MRI signal intensity over time (e.g., peak enhancement and washout slope) has been adopted in many commercially available computer-aided diagnosis systems and often used for clinical reading of breast MRI exams. More rigorous quantitative pharmacokinetic analysis techniques have shown potential to further improve the ability to distinguish benign from malignant breast lesions.^{8,9} Comparison of the diagnostic accuracy between quantitative and semiquantitative analyses of breast DCE-MRI data is needed to justify clinical implementation of quantitative DCE-MRI, which requires high temporal resolution data acquisition in addition to more complex data analysis. However, this comparison has not previously been performed on the same datasets from the same patient cohort, primarily because most published studies utilized clinical DCE-MRI data acquired with low temporal resolution, which is not suitable for pharmacokinetic analysis. In this study, we performed preliminary retrospective analysis of high temporal resolution breast DCE-MRI data collected from women with suspicious breast lesions enrolled in the International Breast MR Consortium (IBMC) multi-institutional 6883 trial.^{17,19} This dataset provides the unique opportunity to compare quantitative pharmacokinetic and semiquantitative analyses of DCE-MRI data collected over multiple sites and imaging platforms and allows for further comparison with clinical MRI assessment for breast cancer diagnostic accuracy. Specifically, a semiquantitative analysis based on signal-enhancement ratio (SER)²⁰ was compared to quantitative Tofts model¹⁰—based on pharmacokinetic analysis of the same datasets to evaluate their respective abilities to distinguish between benign and malignant breast lesions.

2 Materials and Methods

2.1 Patients and Breast Lesions

This retrospective study used data from the IBMC 6883 trial^{17,19} obtained through the American College of Radiology Imaging Network (ACRIN). IBMC 6883 was a prospective multicenter investigation conducted at multiple university hospitals in both North America and Europe. The purpose of the trial was to determine the accuracy of breast MRI for detection of breast cancer in patients with suspicious findings from mammography or clinical exams. The study protocol was approved by the institutional review board or ethics committees of all participating institutions, and informed consent was obtained from all participants. To protect participant's privacy and data confidentiality, all study data were deidentified by ACRIN in accordance with the HIPAA Privacy Rule.

The 6883 trial enrolled women referred for breast biopsy due to an initial presentation of a suspicious or highly suggestive of malignancy mammographic and/or ultrasound finding, or suspicious clinical (palpation) evaluation. MRI examinations were performed prior to breast biopsy, or any procedural or therapeutic interventions. All women first underwent a high spatial resolution three-dimensional (3-D) MRI examination of the breast to assess the suspicious lesion, which included precontrast

T_2 - and pre- and postcontrast T_1 -weighted images.¹⁹ Clinical MRI assessment of the 3-D high spatial resolution images was performed by applying a BI-RADS equivalent categorization to evaluate the probability for malignancy (i.e., 1-negative, 2-benign, 3-probably benign, 4-suspicious, or 5-highly suggestive of malignancy). We note, though, the data collection and interpretation preceded publication of the ACR breast MRI BI-RADS lexicon.² Clinical interpretations were made at each site, blinded to pathology results. Patients with focal enhancing abnormalities were asked to return 18 h or more after the first MRI examination for a high temporal resolution DCE-MRI examination (also before biopsy). For this retrospective study, high temporal resolution DCE-MRI, clinical, and pathological data with histologic assessments of benign and malignant status were made available to the authors by ACRIN for a small set of 45 patients with suspicious lesions, which were collected from 3 of the 14 participating sites (see Table 1). Of the 45 patients, 35 examinations included the appropriate precontrast T_1 maps required for quantitative pharmacokinetic analysis. This unique dataset allows for undertaking of both quantitative and semiquantitative analysis of the same DCE-MRI datasets acquired

Table 1 Patient and lesion characteristics for study cohort ($n = 45$ women).

Characteristic	<i>N</i> (%) or median (range)
Patient age (years)	54 (18 to 80)
Lesion size (mm)	15 (5 to 45)
Histopathology	
Benign	21 (47%)
High-risk atypia	3 (7%)
DCIS	2 (4%)
Invasive	19 (42%)
Lesion type	
Mass	38 (84%)
Nonmass	6 (14%)
Not reported	1 (2%)
Detectable by	
Palpation	13 (30%)
Ultrasound	25 (61%)
Mammography	40 (91%)
Clinical MRI assessment	
1-Negative	0 (0%)
2-Benign	1 (2%)
3-Probably benign	6 (13%)
4-Suspicious	27 (60%)
5-Highly suggestive of malignancy	11 (25%)

Table 2 MRI scan parameters.

MRI parameters	T_1 -mapping	DCE-MRI
Scan sequence	T_1 -weighted 2-D spoiled gradient-echo	T_1 -weighted 2-D spoiled gradient-echo
Time (min)		5
Repetition time (ms)	100, 200, 400, and 1200	100
Echo time (ms)	4 to 5	4 to 5
Flip angle (deg)	90	90
Acquisition matrix	256×128	256×128
Slice thickness (mm)	4	4
Temporal resolution (s)	N/A	15

across multiple sites, with clinical and histology assessments available for comparison and correlation.

2.2 High Temporal Resolution Breast DCE-MRI Acquisition

As previously reported,¹⁷ all MRI examinations were performed at 1.5 T using a dedicated breast coil. The high temporal resolution DCE-MRI examination for characterization of a previously identified breast lesion included a precontrast T_1 mapping sequence using progressive saturation and high temporal resolution two-dimensional (2-D) DCE-MRI (see Table 2 for detailed scan parameters). For DCE-MRI, images were centered on the focal abnormality and initiated at the same time as the start of the intravenous administration of 0.1 mmol/kg of gadolinium chelate [Omniscan (GE Healthcare), Prohance (Bracco), or Magnevist (Berlex); administered over 10 s], followed by a 20-mL saline flush. Image frames of the DCE series were labeled at each time point as follows: $t_1 = 15$ s, $t_2 = 30$ s, $t_3 = 45$ s, ... $t_{20} = 300$ s. Figure 1 shows the representative images from a subject with a pathologically proven benign lesion

Fig. 1 (a) Signal intensity-time course from a breast lesion, with labeling of S_0 , $S_1(t_7)$ (vertical dashed line), S_2 , and contrast injection timing (t_0). The associated images reveals (b) a precontrast image at S_0 , (c) a postcontrast image at S_1 , and (d) a subtraction image ($S_1 - S_0$) showing the enhanced lesion.

(clinical MRI assessment score 4) and the associated signal intensity-time curve measured from the lesion.

2.3 DCE-MRI Data Analysis

In the original 6883 trial, lesion DCE-MRI kinetics was evaluated with maximum enhancement ratio and classification of shape of the signal intensity-time curve as persistent, plateau, or washout. For this study, we further calculated quantitative metrics derived from pharmacokinetic modeling and semiquantitative signal intensity metrics that have previously shown potential to improve lesion characterization.^{8,21,22}

The whole tumor volume was first segmented by manually drawing an ROI of the breast tissue, conservatively surrounding the enhancing tumor on each tumor slice. To refine the selection of the tumor voxels, the percent enhancement of each voxel was used, defined as

$$\text{Percent enhancement} = \left(\frac{S_{\text{post}} - S_{\text{pre}}}{S_{\text{pre}}} \right) \times 100\%, \quad (1)$$

where $S_{\text{pre}} (= S_0)$ and S_{post} were the precontrast (t_1) and averaged postcontrast (t_3 to the last DCE time point) signal intensity, respectively. Tumor voxels selected for subsequent analysis were those within the ROI with percent enhancement $>50\%$.

2.3.1 Semiquantitative SER analysis

The SER was calculated for each tumor voxel as

$$\text{SER} = \frac{S_1 - S_0}{S_2 - S_0}, \quad (2)$$

where S_1 is the initial peak signal intensity, S_0 is the precontrast signal intensity ($= S_{\text{pre}}$), and S_2 is the signal intensity at the last time point. Six parameters related to the SER were estimated for each lesion:^{21,23} SER_{mean} , N_{SERtotal} , $N_{\text{SERpartial}}$, $N_{\text{SERwashout}}$, $\% \text{SER}_{\text{washout}}$, and SER_{peak} (see Table 3 for definitions of these parameters).

S_1 was selected at time point 7 (i.e., t_7 , corresponding to 105 s after initiation of contrast injection) for calculation of SER metrics, as this timing most closely approximates the initial postcontrast time point utilized in standard-of-care clinical breast MRI protocols (typically centered 90 to 120 s after contrast injection).²⁴ That is, t_7 is approximately when clinical protocols typically acquire their postcontrast, high spatial resolution

image to display peak contrast enhancement for image interpretation. Further exploratory analysis was also performed to investigate the influence of the selection of the S_1 time point on diagnostic accuracy of the SER_{mean} metric by varying the S_1 timing between t_3 (45 s) and t_{11} (165 s).

2.3.2 Quantitative pharmacokinetic analysis

As precontrast T_1 maps were not available from 10 patients in this cohort, only 35 of the 45 patients were evaluated using pharmacokinetic analysis. To perform quantitative pharmacokinetic modeling analysis of DCE-MRI data, determination of the time course of the contrast concentration in the blood plasma C_p [arterial input function (AIF)] is required. While there are several commonly employed methods for obtaining the AIF,^{25–28} these were not amenable to this particular dataset because an individual or population-based AIF was not measurable from the DCE-MRI data. Thus, we employed a Tofts model-based reference region (RR) model^{29,30} to estimate the pharmacokinetic parameters. The RR model establishes a relationship between the concentrations of contrast agent in the TOI and the RR (denoted as C_{TOI} and C_{RR} , respectively) as follows:

$$C_{\text{TOI}}(T) = R \cdot C_{\text{RR}}(T) + R \left(\frac{K^{\text{trans,RR}}}{v_{e,\text{RR}}} - \frac{K^{\text{trans,TOI}}}{v_{e,\text{TOI}}} \right) \times \int_0^T C_{\text{RR}}(t) \times \exp \left(-\frac{K^{\text{trans,TOI}}}{v_{e,\text{TOI}}} \right) \times (T - t) dt, \quad (3)$$

where $R \equiv K^{\text{trans,TOI}}/K^{\text{trans,RR}}$, $K^{\text{trans,TOI}}$, and $K^{\text{trans,RR}}$ are the volume transfer rates for the TOI and RR, respectively, and $v_{e,\text{TOI}}$ and $v_{e,\text{RR}}$ are the extravascular extracellular volume fractions of the TOI and RR, respectively. In the present study, the RR ROI was drawn manually in the chest wall muscle to establish the C_{RR} time course. The values of $K^{\text{trans,RR}}$ and $v_{e,\text{RR}}$ were assumed to be 0.15 min^{-1} and 0.12, respectively.^{31–33} $C_{\text{TOI}}(t)$ and $C_{\text{RR}}(t)$, which were estimated directly from the DCE-MRI data using the linear relationship between contrast agent concentration and $(1/T_1)$, were then input into Eq. (3) to estimate the values of $K^{\text{trans,TOI}}$, $v_{e,\text{TOI}}$, and the efflux rate constant $k_{ep,\text{TOI}} (= K^{\text{trans,TOI}}/v_{e,\text{TOI}})$, using a standard curve-fitting procedure in MATLAB[®] (MathWorks, Natick, Massachusetts). The mean parameter value from all tumor voxels was quantified. For simplification of tumor parameters, $K^{\text{trans,TOI}}$, $k_{ep,\text{TOI}}$, and $v_{e,\text{TOI}}$ are denoted as K^{trans} , k_{ep} , and v_e , for the remainder of the paper.

2.4 Statistical Analysis

The nonparametric Wilcoxon rank sum test was used to assess differences in imaging parameters between benign and malignant lesions. Univariable and multivariable logistic regression modeling was performed to identify parameters that accurately diagnosed malignancy. Model performance was compared by receiver operating characteristic (ROC) curve analysis and calculation of areas under the ROC curves (AUCs). Cutoff points and corresponding sensitivities, specificities, positive predictive values (PPV), and negative predictive values (NPV) were evaluated for a minimum sensitivity requirement of 94%, corresponding to the observed sensitivity of the clinical MRI assessment benchmark (BI-RADS ≥ 4) in this 45-patient cohort. Comparison of AUCs was conducted with the DeLong–DeLong statistical test.³⁴ Statistical significance was defined as $p < 0.05$

Table 3 Definition of signal intensity-based parameters.

SER parameter	Calculation involving tumor voxels
SER_{mean}	Mean of voxels with $\text{SER} \geq 0$
N_{SERtotal}	Number of voxels with $\text{SER} \geq 0$
$N_{\text{SERpartial}}$	Number of voxels with $\text{SER} \geq 0.9$
$N_{\text{SERwashout}}$	Number of voxels with $\text{SER} > 1.3$
$\% \text{SER}_{\text{washout}}$	$(N_{\text{SERwashout}}/N_{\text{SERtotal}}) \times 100$
SER_{peak}	Highest average SER of five or more connected voxels within a $3 \times 3 \times 3$ cube

Fig. 2 Color SER_{mean} and pharmacokinetic parameter (K^{trans} , v_e , and k_{ep}) maps overlaid on an anatomical postcontrast MR image for two suspicious breast lesions in the study. (a) A malignant (invasive) lesion and (b) a benign lesion. Both lesions were assessed with standard clinical MRI interpretation and both given a clinical MRI assessment score of 4. It can be seen that $SER_{mean}(t_7)$, K^{trans} , and k_{ep} values of the malignant lesion were higher than the corresponding values of the benign lesion.

without adjustment for multiple comparisons. All statistical analyses were performed using JMP 11.0.0 (SAS Institute, Inc., Cary, North Carolina) or with MedCalc (MedCalc Software, Belgium).

3 Results

Pathologic assessment of biopsy specimens was available for all 45 lesions; of those, 24/45 (53%) were benign [3/24 (13%) were atypia and classified as benign] and 21/45 (47%) were malignant [2/21 (10%) were ductal carcinoma *in situ* (DCIS)] (Table 1). Across the three sites, site 1 contributed 34 lesions (13 malignant and 21 benign), site 2 contributed 2 lesions (both malignant), and site 3 contributed 9 lesions total (6 malignant and 3 benign). SER metrics and pharmacokinetic modeling results were obtained for the entire 45-patient cohort (with 21 malignant and 24 benign lesions) and a 35-patient subcohort (with 16 malignant and 19 benign lesions), respectively. Figure 2 shows the representative parametric maps of the signal intensity metrics of $SER_{mean}(t_7)$ and pharmacokinetic parameters of K^{trans} ,

k_{ep} , and v_e overlaid on the postcontrast images of a benign and malignant lesion.

3.1 SER Signal Intensity Assessment

For all SER calculations, S_1 was selected at time point 7. Of the six SER-related metrics, only SER_{mean} showed significant differentiation between benign (median = 0.64 and range = 0.16 to 1.34) and malignant lesions (median = 0.87, range = 0.61 to 1.43, and $p = 0.0001$), yielding an AUC of 0.82 (Table 4). Figure 3 shows the scatter $SER_{mean}(t_7)$ plots and mean values for the benign and malignant lesions [Fig. 3(a)] with the corresponding ROC analysis [Fig. 3(b)].

3.2 Influence of S_1 Timing for SER_{mean} Assessment

As a secondary analysis, we evaluated the influence of S_1 timing on the SER_{mean} value and found that $SER_{mean}(t_7)$ also provided the highest AUC, although differences were minimal for S_1 time point between t_5 (75 s) and t_{11} (165 s) with all resulting

Table 4 Signal intensity-based parameters (S_1 : t_7) for benign and malignant breast lesions ($n = 45$).

Variable	Benign ($n = 24$)	Malignant ($n = 21$)	p (Wilcoxon)	AUC
	Median (range)	Median (range)		
SER_{peak}	1.04 (0.40 to 6.02)	1.25 (0.73 to 7.20)	0.24	0.61
$SER_{mean}(t_7)$	0.64 (0.16 to 1.34)	0.87 (0.61 to 1.43)	0.0001	0.82
$N_{SERtotal}$	903.5 (1 to 20356)	157 (3 to 11737)	0.08	0.65
$N_{SERpartial}$	63.5 (0 to 4012)	77 (0 to 3450)	0.54	0.56
$N_{SERwashout}$	2 (0 to 2759)	7 (0 to 192)	0.48	0.56
$\%SER_{washout}$	0.26 (0 to 62.48)	4.02 (0 to 61.16)	0.33	0.59

Note: AUC, area under curve and SER, signal-enhancement ratio.

Fig. 3 Diagnostic performance of SER_{mean} for differentiating breast lesions in 45 patients: (a) scatter plots showing $SER_{mean}(t_7)$ values for the benign and malignant lesions and (b) the corresponding ROC curve with AUC = 0.82. The three bars in the scatter plot for each lesion group represent the mean and the 95% confidence intervals.

AUC values >0.75 (Table 5). This suggests that diagnostic performance of SER is optimal after the first minute of contrast injection ($S_1 > 60$ s). To further evaluate SER_{mean} , S_1 was assigned to be the peak intensity for each lesion between 90 (t_6) and 120 (t_8) s; this resulted in an AUC of 0.79. The diagnostic performances of the other SER metrics also did not improve by altering the S_1 timing.

Additionally, to mimic the temporal resolution of typical clinical protocols, we have also evaluated the SER_{mean} with the S_1 value set to be the average of t_6 to t_8 time points (simulating 45-s temporal resolution) and the average of t_3 to t_9 (simulating 90-s temporal resolution). The associated ROC analyses

comparing benign and malignant lesions resulted in AUC values of 0.78 ($p = 0.001$) and 0.79 ($p = 0.001$), respectively. These results justified the selection of $SER_{mean}(t_7)$ for further comparative and correlative analyses.

3.3 Quantitative Pharmacokinetics Assessment

A summary of the pharmacokinetic parameters for the 35 lesions where quantitative analysis was performed is shown in Table 6. Both K^{trans} and k_{ep} were significantly higher in malignant ($K^{trans} = 0.62 \text{ min}^{-1}$ and $k_{ep} = 1.34 \text{ min}^{-1}$) versus benign ($K^{trans} = 0.33 \text{ min}^{-1}$ and $k_{ep} = 0.87 \text{ min}^{-1}$) lesions ($p = 0.04$ and 0.04 , respectively). The two metrics provided comparable performance for predicting malignancy, with AUC = 0.71 for K^{trans} and 0.70 for k_{ep} . Conversely, v_e was a poor marker for discriminating benign (0.52, range = 0.13 to 0.82) and malignant (0.49, range = 0.22 to 0.88) lesions ($p = 0.96$ and AUC = 0.51). Figure 4 shows the scatter plots and the mean quantitative pharmacokinetic parameter values for the benign and malignant lesions.

Spearman's correlation analysis within this subcohort showed both K^{trans} and k_{ep} metrics to be significantly correlated with the $SER_{mean}(t_7)$ parameter (Table 7).

3.4 Preliminary Comparison of Diagnostic Accuracy Among SER, Pharmacokinetic Parameters, and Clinical MRI Assessment

The diagnostic performances of $SER_{mean}(t_7)$ and pharmacokinetic parameters were estimated and compared with clinical MRI assessment within the $N = 35$ -patient subcohort. For clinical MRI assessment, the standard positive criteria of BI-RADS ≥ 4 were used, which produced a sensitivity of 94%. For comparison of diagnostic performance, cutoffs for the kinetics variables were selected to achieve the similar condition of 94% or greater sensitivity. At 94% sensitivity, the associated specificities of clinical MRI assessment, $SER_{mean}(t_7)$, K^{trans} , and k_{ep} were 32%, 53%, 42%, and 32%, respectively (Table 8). In this cohort, both the specificity and PPV of $SER_{mean}(t_7)$ (53% and 63%) and K^{trans} (42% and 58%) appeared to be higher than that of standard-of-care clinical MRI assessment (32% and

Table 5 SER_{mean} at various S_1 time points for benign and malignant breast lesions ($n = 45$).

S_1 time point (s)	Benign ($n = 24$)	Malignant ($n = 21$)	p (Wilcoxon)	AUC
	Median SER_{mean} (range)	Median SER_{mean} (range)		
t_3 (45)	0.24 (0.08 to 0.90)	0.49 (0.02 to 1.33)	0.07	0.66
t_4 (60)	0.40 (0.10 to 1.24)	0.71 (0.15 to 1.55)	0.02	0.70
t_5 (75)	0.52 (0.15 to 1.31)	0.75 (0.36 to 1.70)	0.001	0.78
t_6 (90)	0.57 (0.15 to 1.34)	0.80 (0.45 to 1.61)	<0.01	0.75
t_7 (105)	0.64 (0.16 to 1.34)	0.87 (0.61 to 1.43)	<0.001	0.82
t_8 (120)	0.67 (0.13 to 1.34)	0.92 (0.61 to 1.41)	<0.01	0.76
t_9 (135)	0.70 (0.34 to 1.35)	0.93 (0.66 to 1.35)	<0.01	0.77
t_{10} (150)	0.77 (0.19 to 1.54)	0.93 (0.72 to 1.26)	<0.01	0.77
t_{11} (165)	0.76 (0.29 to 1.50)	0.95 (0.72 to 1.31)	<0.01	0.77

Table 6 Pharmacokinetic parameters for benign and malignant breast lesions ($n = 35$).

MRI metrics	Benign ($n = 19$)	Malignant ($n = 16$)	p (Wilcoxon)	AUC
	Median (range)	Median (range)		
K^{trans}	0.33 (0.09 to 2.26)	0.62 (0.23 to 2.21)	0.04	0.71
v_e	0.52 (0.13 to 0.82)	0.49 (0.22 to 0.88)	0.96	0.51
k_{ep}	0.87 (0.37 to 3.69)	1.34 (0.59 to 3.85)	0.04	0.70

Note: AUC, area under curve.

Fig. 4 Scatter plots of lesion pharmacokinetic parameters (a) K^{trans} , (b) k_{ep} , and (c) v_e in the 35-patient subcohort for which pharmacokinetic modeling analysis of DCE-MRI data was performed. The three bars in the scatter plots for each lesion group represent the mean and the 95% confidence intervals.

54%). $\text{SER}_{\text{mean}}(t_7)$ had the highest NPV of 91%, followed closely by K^{trans} (89%). The ROC curves for $\text{SER}_{\text{mean}}(t_7)$, pharmacokinetic modeling parameters, and clinical MRI assessment are shown together for comparison in Fig. 5.

Interestingly, of the 13 false positives rendered by clinical MRI assessment, three (23%) were correctly identified as benign by both K^{trans} and SER_{mean} . Additionally, there were 4 of these 13 cases (31%), in which only SER_{mean} correctly

identified the lesions as benign, and one case (8%), in which only K^{trans} correctly identified the lesion as benign.

The diagnostic performance of $\text{SER}_{\text{mean}}(t_7)$ was further compared with clinical MRI assessment for the full $N = 45$ -patient cohort, with AUC values of 0.82 and 0.72, respectively. At 94% sensitivity, the specificity/PPV/NPV of clinical MRI assessment and $\text{SER}_{\text{mean}}(t_7)$ was 25%/53%/86% and 50%/63%/92%, respectively. Although the AUC for $\text{SER}_{\text{mean}}(t_7)$ was higher

Table 7 Spearman's correlations between SER_{mean} and pharmacokinetic metrics ($n = 35$).

MRI metrics	K^{trans} rho (p)	v_e rho (p)	k_{ep} rho (p)	$\text{SER}_{\text{mean}}(t_7)$ rho (p)
K^{trans}	1	0.64 (<0.0001)	0.80 (<0.0001)	0.48 (0.004)
v_e	—	1	0.16 (0.33)	0.17 (0.33)
k_{ep}	—	—	1	0.53 (0.001)
$\text{SER}_{\text{mean}}(t_7)$	—	—	—	1

Table 8 Comparison of diagnostic performances.

Diagnostic parameter	N	AUC (95% CI)	Optimal cutoff value ^a	Sensitivity (%)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
Clinical MRI Assessment	35	0.73 (0.56 to 0.90)	$\geq 4^b$	94	32% (9% to 56%)	54% (46% to 64%)	86% (59% to 91%)
$SER_{mean}(t_7)$	35	0.81 (0.66 to 0.95)	≥ 0.67	94	53% (26% to 79%)	63% (52% to 79%)	91% (84% to 94%)
K^{trans}	35	0.71 (0.53 to 0.89)	≥ 0.25	94	42% (11% to 74%)	58% (47% to 75%)	89% (64% to 93%)
k_{ep}	35	0.70 (0.52 to 0.88)	≥ 0.65	94	32% (11% to 74%)	54% (47% to 75%)	86% (64% to 93%)

Note: AUC, area under curve; PPV, positive predictive value (%); NPV, negative predictive value (%); and CI, confidence interval.

^aSelected to provide equal or greater sensitivity versus clinical MRI interpretation.

^bSelected as per standard of care (BI-RADS 4 or 5).

than that of clinical MRI assessment, the difference was not statistically significant ($p = 0.26$).

3.5 Multivariate Modeling

In the multivariable logistic regression analysis within the 35-patient cohort where all kinetic parameters were available, only $SER_{mean}(t_7)$ and clinical MRI assessment were independently predictive of malignancy ($p = 0.0014$ and 0.002 , respectively). ROC analysis showed that a model combining the two parameters yielded an optimal performance to distinguish benign and malignant tumors, with an AUC of 0.87. At 94% sensitivity, the specificity/PPV/NPV of the combined model was 58%/65%/92%. Similar diagnostic performance was achieved when data from the full 45-patient cohort were used to generate the two-parameter model (AUC = 0.87). Furthermore, the multivariable assessment showed significant improvement compared to clinical assessment alone when comparing the two ROC curves ($p = 0.002$).

4 Discussion

Improving the diagnostic accuracy of MRI has the potential to reduce unnecessary biopsies of benign breast lesions and, thereby, decreases costs and morbidity associated with these invasive biopsy procedures. Numerous single-center studies have reported improved differentiation of benign or malignant lesions using DCE-MRI kinetic information.^{5-9,35} In this

preliminary study, we report both semiquantitative signal intensity-based analysis and pharmacokinetic model-based analysis of DCE-MRI data collected in the multisite IBMC 6883 trial from women with suspicious breast tumors found by screening mammography, ultrasound, or clinical examination. Specifically, the ability to distinguish benign and malignant breast lesions was evaluated and compared for the semiquantitative SER metric, model-based pharmacokinetic parameters, and clinical MRI assessment (based on the conventional high spatial, but low temporal resolution DCE-MRI scans) within the same cohort.

The parameter SER_{mean} provided the highest AUC among all the metrics for distinguishing benign versus malignant tumors. Moreover, at 94% sensitivity, SER_{mean} provided substantially higher specificity than clinical MRI assessment, K^{trans} , and k_{ep} . This approach provides a promising method for DCE-MRI evaluation of suspicious breast lesions and, compared to the pharmacokinetic analysis approach, is simpler and more straight-forward for implementation in routine clinical practice. While the high temporal resolution IBMC 6883 datasets allowed investigating a wide range of semiquantitative SER features and timings, our results suggested that standardizing the SER_{mean} calculation with S_1 obtained at postcontrast timing comparable to that of the first postcontrast image frame in standard clinical protocols (~105 s after initiation of contrast injection) could achieve optimal diagnostic performance. Although the available pool of semiquantitative features has higher dimensionality than would be available clinically, as analyses of multiple SER parameters and various S_1 time points were performed in this study, the robust performance of SER_{mean} across a range of time points selected for S_1 quantification supports its potential for future clinical translation. Furthermore, only modest performance differences were observed when varying the timing of S_1 between 75 and 165 s, further supporting the robustness and translatability of SER_{mean} as a diagnostic marker across a variety of clinical imaging protocols. To facilitate the use of the SER_{mean} method in clinical practice, where breast DCE-MRI is generally performed with high spatial but low temporal resolution, future work is needed to directly compare SER_{mean} diagnostic performance from high temporal resolution DCE-MRI data with that from low temporal resolution data.

This preliminary study corroborates other studies that have shown the signal intensity parameter, SER, to be significantly different between benign and malignant lesions.^{36,37} Abe et al.³⁸ and Jansen et al.³⁷ each reported SER to achieve diagnostic performance comparable to clinical assessment. Additionally,

Fig. 5 Diagnostic performances of quantitative pharmacokinetic modeling parameters, $SER_{mean}(t_7)$, and clinical MRI assessment for differentiating benign and malignant breast lesions in the 35-patient cohort. ROC curves are shown with AUCs = 0.81, 0.71, 0.70, 0.51, and 0.73 for SER_{mean} , K^{trans} , k_{ep} , v_e , and clinical MRI assessment, respectively.

Karahaliou et al.³⁹ demonstrated that the predictive value of the SER map can be further improved by extracting texture information.

Our analysis of the multisite IBMC trial data is also consistent with previous single-center studies^{8,9} showing that DCE-MRI pharmacokinetic parameters (K^{trans} and k_{ep}) extracted from DCE-MRI hold diagnostic value as markers for breast cancer, while v_e does not. Huang et al.⁸ performed pharmacokinetic quantitative modeling of high temporal resolution DCE-MRI data from 92 mammography occult but MRI-visible lesions (from 89 high-risk patients) and the diagnostic specificity reached 98.6% at 100% sensitivity using pharmacokinetic parameters as diagnostic markers. Our study also provides support for the use of the RR method as a valid approach for quantitatively modeling of DCE-MRI data that were collected with protocols not specifically designed for pharmacokinetic analysis, such as acquisitions lacking coverage of a visible artery. Our results indicate that DCE-MRI parameters of K^{trans} and SER_{mean} reflect unique underlying biology in tissue microvascular properties, as they primarily measure the uptake and delayed phases of the contrast kinetics, respectively, and were only moderately correlated.

All 45 lesions included in this study were found suspicious by mammographic and/or sonographic assessments, or clinical exams, and referred for biopsies as per standard of care. However, histological assessment found only 21 were malignant, demonstrating a relatively low PPV of 47% (21/45) for these standard diagnostic techniques. The complete datasets from the IBMC 6883 trial showed that breast MRI can improve breast cancer diagnostic accuracy compared to standard-of-care mammography, ultrasound, or clinical exam.¹⁹ Our preliminary findings from a small subset of the IBMC 6883 data further show that both semiquantitative and quantitative pharmacokinetic analysis of high temporal resolution DCE-MRI data can further improve the specificity and PPV over conventional clinical breast MRI assessment. Reducing false positives from conventional breast MRI interpretation and consequently the number of unnecessary biopsies could potentially make MRI cost-effective for screening women diagnosed with suspicious lesions and referred for biopsies based on standard imaging methods and/or clinical exams.

One limitation of our study is its exploratory nature. As such, we did not incorporate adjustments for multiple comparisons or use an independent dataset to confirm the primary study findings. However, we note that the p -value associated with SER_{mean} for discriminating between benign and malignant lesions ($p = 0.0001$) would remain statistically significant after a Bonferroni correction. Regardless, our preliminary findings require confirmation in a larger cohort. Conversely, this work does provide some unique insights into the relative diagnostic performance and robustness of the different DCE-MRI metrics to guide future studies. The limited sample size did not allow for comparison of kinetics features within histologic subtypes, direct statistical comparisons of diagnostic accuracy among different kinetics features, and the site effect could not be adjusted. Therefore, it is possible that site variations weakened, obscured, or exaggerated some associations. Additionally, precontrast T_1 maps were acquired for only 35 of the 45 datasets available, constraining the comparison between signal intensity metrics and pharmacokinetic modeling approaches to only the 35 subject subcohort. Image acquisition was not uniform across all MRI datasets as examinations were performed across different scanner platforms, which could cause variations in analyses.

Finally, the dynamic imaging data in this retrospective analysis were collected more than a decade ago. Recent advances in MRI technologies can now allow for simultaneous high temporal and high spatial DCE-MRI acquisitions, which would further facilitate clinical translation of quantitative kinetics analysis, but would not likely affect the results of SER calculations reported in this study. Despite these limitations, this preliminary study demonstrates the potential ability of semiquantitative and quantitative analysis of DCE-MRI data to distinguish benign and malignant breast lesions across multiple sites and imaging platforms.

5 Conclusions

In summary, this preliminary study investigated the comparative utility of quantitative pharmacokinetic modeling and semiquantitative signal intensity analysis of DCE-MRI for distinguishing benign and malignant breast lesions, using a retrospective dataset acquired across multiple institutions and scanner platforms. Our preliminary findings suggest that incorporating SER_{mean} into clinical breast MRI interpretations may hold potential to improve lesion diagnostic accuracy. This has potential in the future to reduce unnecessary biopsies of benign lesions while retaining the high sensitivity of breast MRI for detection of malignancy. Further work is needed to validate these results with a larger prospective cohort and/or by meta-analysis of data from smaller independent studies.

Disclosures

The authors have no conflicts of interest.

Acknowledgments

We thank the National Institutes of Health for funding through National Cancer Institute under Grant Nos. U01CA142565, R01CA151326, P50CA138293, U01CA154602, R44CA180425, P30CA069533, and U01CA174706. We thank the Cancer Prevention and Research Institute of Texas for funding through Grant No. RR160005. We thank ACRIN and its statistics core for providing clinical, imaging, and histopathological data. We thank funding support from Center for Women's Health Circle of Giving, Oregon Health and Science University.

References

1. T. E. Yankeelov and J. C. Gore, "Dynamic contrast enhanced magnetic resonance imaging in oncology: theory, data acquisition, analysis, and examples," *Curr. Med. Imaging Rev.* **3**(2), 91–107 (2007).
2. American College of Radiology, *Breast Imaging Reporting and Data System (BI-RADS): Breast Imaging Atlas*, American College of Radiology, Reston, Virginia (2003).
3. P. J. Moate et al., "A modified logistic model to describe gadolinium kinetics in breast tumors," *Magn. Reson. Imaging* **22**(4), 467–473 (2004).
4. L. D. Buadu et al., "Breast lesions: correlation of contrast medium enhancement patterns on MR images with histopathologic findings and tumor angiogenesis," *Radiology* **200**(3), 639–649 (1996).
5. O. Ikeda et al., "Characterization of breast masses by dynamic enhanced MR imaging. A logistic regression analysis," *Acta Radiol.* **40**(6), 585–592 (1999).
6. C. K. Kuhl et al., "Dynamic breast MR imaging: are signal intensity time course data useful for differential diagnosis of enhancing lesions?" *Radiology* **211**(1), 101–110 (1999).
7. T. C. Williams et al., "Breast MR imaging: computer-aided evaluation program for discriminating benign from malignant lesions," *Radiology* **244**(1), 94–103 (2007).
8. W. Huang et al., "Discrimination of benign and malignant breast lesions by using shutter-speed dynamic contrast-enhanced MR imaging," *Radiology* **261**(2), 394–403 (2011).

9. M. C. Schabel et al., "Pharmacokinetic mapping for lesion classification in dynamic breast MRI," *J. Magn. Reson. Imaging* **31**(6), 1371–1378 (2010).
10. P. S. Tofts, B. Berkowitz, and M. D. Schnall, "Quantitative analysis of dynamic Gd-DTPA enhancement in breast tumors using a permeability model," *Magn. Reson. Med.* **33**(4), 564–568 (1995).
11. M. Kriege et al., "Efficacy of MRI and mammography for breast-cancer screening in women with a familial or genetic predisposition," *N. Engl. J. Med.* **351**(5), 427–437 (2004).
12. C. D. Lehman et al., "Cancer yield of mammography, MR, and US in high-risk women: prospective multi-institution breast cancer screening study," *Radiology* **244**(2), 381–388 (2007).
13. D. Saslow et al., "American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography," *CA Cancer J. Clin.* **57**(2), 75–89 (2007).
14. E. Warner et al., "Surveillance of BRCA1 and BRCA2 mutation carriers with magnetic resonance imaging, ultrasound, mammography, and clinical breast examination," *JAMA* **292**(11), 1317–1325 (2004).
15. C. D. Lehman et al., "MRI evaluation of the contralateral breast in women with recently diagnosed breast cancer," *N. Engl. J. Med.* **356**(13), 1295–1303 (2007).
16. M. D. Schnall et al., "MRI detection of distinct incidental cancer in women with primary breast cancer studied in IBMC 6883," *J. Surg. Oncol.* **92**(1), 32–38 (2005).
17. M. D. Schnall et al., "Diagnostic architectural and dynamic features at breast MR imaging: multicenter study," *Radiology* **238**(1), 42–53 (2006).
18. W. B. Demartini et al., "Probability of malignancy for lesions detected on breast MRI: a predictive model incorporating BI-RADS imaging features and patient characteristics," *Eur. Radiol.* **21**(8), 1609–1617 (2011).
19. D. A. Bluemke et al., "Magnetic resonance imaging of the breast prior to biopsy," *JAMA* **292**(22), 2735–2742 (2004).
20. L. Esserman et al., "Contrast-enhanced magnetic resonance imaging to assess tumor histopathology and angiogenesis in breast carcinoma," *Breast J.* **5**(1), 13–21 (1999).
21. V. A. Arasu et al., "Can signal enhancement ratio (SER) reduce the number of recommended biopsies without affecting cancer yield in occult MRI-detected lesions?" *Acad Radiol.* **18**(6), 716–721 (2011).
22. L. Esserman et al., "Utility of magnetic resonance imaging in the management of breast cancer: evidence for improved preoperative staging," *J. Clin. Oncol.* **17**(1), 110 (1999).
23. X. Li et al., "DCE-MRI analysis methods for predicting the response of breast cancer to neoadjuvant chemotherapy: pilot study findings," *Magn. Reson. Med.* **71**(4), 1592–1602 (2014).
24. R. L. Gutierrez et al., "Dynamic breast MRI: does lower temporal resolution negatively affect clinical kinetic analysis?" *Am. J. Roentgenol.* **199**(3), 703–708 (2012).
25. A. Shukla-Dave et al., "Average arterial input function for quantitative dynamic contrast enhanced magnetic resonance imaging of neck nodal metastases," *BMC Med. Phys.* **9**, 4 (2009).
26. Y. Wang et al., "Feasibility of using limited-population-based arterial input function for pharmacokinetic modeling of osteosarcoma dynamic contrast-enhanced MRI data," *Magn. Reson. Med.* **59**(5), 1183–1189 (2008).
27. X. Li et al., "A novel AIF tracking method and comparison of DCE-MRI parameters using individual and population-based AIFs in human breast cancer," *Phys. Med. Biol.* **56**(17), 5753–5769 (2011).
28. A. Fedorov et al., "A comparison of two methods for estimating DCE-MRI parameters via individual and cohort based AIFs in prostate cancer: a step towards practical implementation," *Magn. Reson. Imaging* **32**(4), 321–329 (2014).
29. T. E. Yankeelov et al., "Quantitative pharmacokinetic analysis of DCE-MRI data without an arterial input function: a reference region model," *Magn. Reson. Imaging* **23**(4), 519–529 (2005).
30. T. E. Yankeelov et al., "Comparison of a reference region model with direct measurement of an AIF in the analysis of DCE-MRI data," *Magn. Reson. Med.* **57**(2), 353–361 (2007).
31. C. R. Planey et al., "Temporal sampling requirements for reference region modeling of DCE-MRI data in human breast cancer," *J. Magn. Reson. Imaging* **30**(1), 121–134 (2009).
32. S. M. Galbraith et al., "Reproducibility of dynamic contrast-enhanced MRI in human muscle and tumours: comparison of quantitative and semi-quantitative analysis," *NMR Biomed.* **15**(2), 132–142 (2002).
33. A. R. Padhani, "Dynamic contrast-enhanced MRI in clinical oncology: current status and future directions," *J. Magn. Reson. Imaging* **16**(4), 407–422 (2002).
34. E. R. DeLong, D. M. DeLong, and D. L. Clarke-Pearson, "Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach," *Biometrics* **44**(3), 837–845 (1988).
35. L. C. Wang et al., "MRI-detected suspicious breast lesions: predictive values of kinetic features measured by computer-aided evaluation," *Am. J. Roentgenol.* **193**(3), 826–831 (2009).
36. A. Shimauchi et al., "Evaluation of kinetic entropy of breast masses initially found on MRI using whole-lesion curve distribution data: comparison with the standard kinetic analysis," *Eur. Radiol.* **25**(8), 2470–2478 (2015).
37. S. A. Jansen et al., "Differentiation between benign and malignant breast lesions detected by bilateral dynamic contrast-enhanced MRI: a sensitivity and specificity study," *Magn. Reson. Med.* **59**(4), 747–754 (2008).
38. H. Abe et al., "Kinetic analysis of benign and malignant breast lesions with ultrafast dynamic contrast-enhanced MRI: comparison with standard kinetic assessment," *Am. J. Roentgenol.* **207**, 1159–1166 (2016).
39. A. Karahalidou et al., "Assessing heterogeneity of lesion enhancement kinetics in dynamic contrast-enhanced MRI for breast cancer diagnosis," *Br. J. Radiol.* **83**(988), 296–309 (2010).

Anna G. Sorace is an assistant professor of Department of Diagnostic Medicine at the Dell Medical School, The University of Texas at Austin. Her research interests include developing and utilizing translational advanced multimodality imaging techniques to provide insight into the pathology of diseases, identifying imaging biomarkers for early response to neoadjuvant and adjuvant cancer treatment, and quantitative imaging to guide and enhance drug efficacy.

Savannah C. Partridge is a professor at the University of Washington, Department of Radiology. She leads the UW Quantitative Breast Imaging Laboratory, working to develop new quantitative imaging tools for clinical and translational breast magnetic resonance imaging (MRI) applications. She is also an active member of the ECOG-ACRIN Cancer Research Group and serves as chair and co-chair of two national breast MRI focused multicenter clinical trials (ACRIN 6702 and ACRIN 6698).

Xia Li is a lead scientist in the Artificial Intelligence and Image Analytics Group at GE Global Research Center. Prior to GE, she worked as an imaging research scientist at Vanderbilt University. Her expertise is in machine learning and deep learning, and medical image analysis in general. Her contribution in the field includes 36 peer-reviewed journal papers, more than 50 scientific proceeding papers, and 2 book chapters.

Jack Virostko is an assistant professor of Department of Diagnostic Medicine at the Dell Medical School, The University of Texas at Austin. He received his master of science and PhD in biomedical engineering from Vanderbilt University. He also received his master of science in clinical investigation from Vanderbilt University. He is interested in developing multimodal imaging techniques and integrating them into computational models of cancer development and progression.

Stephanie L. Barnes is an imaging research scientist at the University of Texas at Austin in the Department of Biomedical Engineering with over 10 years of experience in MRI and PET preclinical and clinical acquisition and analysis.

Daniel S. Hippe is a statistician and research scientist at the University of Washington, Department of Radiology. He received his master of science in statistics and his bachelor of science in electrical engineering from the University of Washington. He works on a wide variety of research projects in radiology, radiation oncology, and cardiology.

Wei Huang is an associate professor/scientist in the Advanced Imaging Research Center of Oregon Health and Science University.

He is an MRI physicist by training and has more than 25 years of experience in MRI and MR spectroscopy research. His current research focuses on imaging of underlying tumor biological functions using quantitative MRI methods for cancer detection and therapeutic monitoring.

Thomas E. Yankeelov is the W.A. "Tex" Moncrief professor of computational oncology and professor of biomedical engineering and

diagnostic medicine at The University of Texas at Austin. He also serves as the director of the Center for Computational Oncology, the Institute for Computational and Engineering Sciences. The overall goal of his research is to develop tumor forecasting methods by integrating advanced imaging technologies with predictive, multiscale models of tumor growth to optimize therapy.