

# Classification Tree Approach to Validate and Improve Quantitative DCE-MRI Diagnosis of Breast Cancer: Analysis of Multicenter Data

Lian Wang<sup>1</sup>, Yiyi Chen<sup>2</sup>, Alina Tudorica<sup>2</sup>, Karen Oh<sup>2</sup>, Nicole Roy<sup>2</sup>, Mark Kettler<sup>2</sup>, Dongseok Choi<sup>2</sup>, and Wei Huang<sup>2</sup>

<sup>1</sup>Providence Health and Services, Portland, Oregon, United States, <sup>2</sup>Oregon Health & Science University, Portland, Oregon, United States

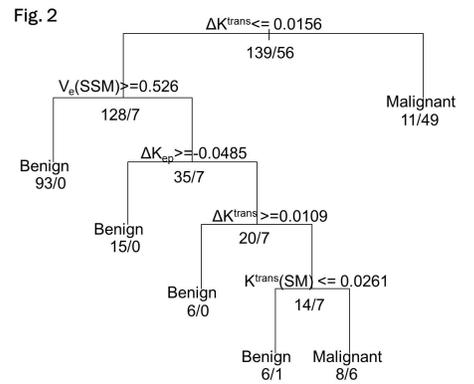
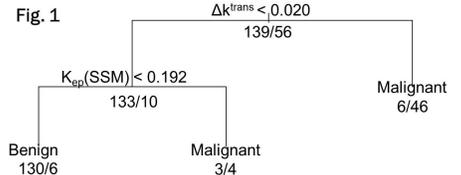
**Introduction:** Previous single-center studies (1,2) demonstrated that the new DCE-MRI biomarker,  $\Delta K^{trans}$ , is effective and consistent in discrimination of benign and malignant breast lesions in a high-risk pre-biopsy population, and outperforms the other DCE-MRI parameters derived from pharmacokinetic analysis of DCE-MRI time-course data.  $\Delta K^{trans}$  is defined as  $\Delta K^{trans} = K^{trans}(SSM) - K^{trans}(SM)$ , where SM stands for the standard Tofts model (3) and SSM the Shutter-Speed model (1,2) – the latter takes into account the finite intercompartmental water exchange kinetics.  $\Delta K^{trans}$  provides a measure of the exchange effects on  $K^{trans}$  estimation, which is generally stronger in malignant tumors compared to benign ones (1,2). In this study, we sought to validate the effectiveness of  $\Delta K^{trans}$  for breast cancer diagnosis using research breast DCE-MRI data from three institutions and to explore whether the diagnostic performance can be further improved by incorporating some of the other DCE-MRI parameters.

**Methods:** Research pre-biopsy breast DCE-MRI data were collected in three institutions from a total of 184 women who were referred for biopsies because of suspicious breast imaging findings. A total of 195 contrast-enhanced MRI lesions were identified. Among them, 95 were mammography-occult lesions initially detected by screening clinical MRI from a high-risk cohort at one institution. The other 100 lesions were found at the two other institutions through mammography and/or ultrasound screening of general population. There were considerable variations in data acquisition details among the 3 sites (4), such as scanner platform (1.5T Philips/Marconi, 1.5T GE, and 3T Siemens), pulse sequence [full-k-space sampling (2 sites) vs. k-space undersampling (1 site)], Gd contrast agent (Magnevist, Omniscan, and Prohance), etc.

All research DCE-MRI data were subjected to SM and SSM analyses (1,2) to extract the following mean lesion ROI pharmacokinetic parameters:  $K^{trans}(SSM)$ ,  $K^{trans}(SM)$ ,  $\Delta K^{trans}$ ,  $v_e(SSM)$ ,  $v_e(SM)$ ,  $\Delta v_e$ ,  $k_{ep}(SSM)$ ,  $k_{ep}(SM)$ ,  $\Delta k_{ep}$ , and  $\tau_1(SSM)$  only. These parameters, as well as the gold-standard biopsy pathology diagnoses of malignant/benign status, were supplied as inputs to construct classification trees that separate the lesions into benign and malignant groups. The classification tree was created with the rpart package in R, using recursive partitioning based on Gini index (5). The construction of the classification tree can be modified by adjusting the weights for different types of errors - in this case it is either false positive or false negative error. Either equal-weight (i.e., false negative and false positive errors are of equal consequence) or unequal-weight (with a false negative error being 4-10 times more serious than a false positive error) settings were implemented. In addition, there is a complexity parameter setting in constructing classification tree, which decides how complicated the classification tree is allowed to be. The typical range of 0.01-0.001 (the smaller the complexity parameter, the more complicated the tree) were used. The sensitivity, specificity, and overall accuracy for breast cancer diagnosis were calculated for each output classification tree.

**Results:** The biopsy pathology results revealed 56 malignant tumors out of the 195 suspicious lesions, indicating 28.7% diagnostic accuracy for clinical breast imaging. When false positive and false negative errors were equally weighted, the resulted classification tree is shown in Fig. 1. The tree construction was robust to adjustment in the complexity parameter values (0.01-0.001). Fig. 1 shows that the first split is based on the  $\Delta K^{trans}$  parameter with lesions fulfilling the criterion of  $\Delta K^{trans} < 0.020 \text{ min}^{-1}$  to be classified to the left (benign) and the others to the right (malignant). The number pairs under each node are the numbers of true benign and true malignant lesions (separated by a slash) at that node (before further splitting using the expression at that node). With unequal weighting of false negative being 4-10 times worse than false positive, the overall structure of the classification tree is shown in Fig. 2. Changing the complexity parameter only altered the number of splits (or decision points) from 5-split (complexity parameter = 0.001) down to 1-split (complexity parameter = 0.01). Table 1 summarizes the performance of each classification tree.

**Discussion:** Both classification tree structures have their first decision point/split based on the  $\Delta K^{trans}$  parameter, which confirms the previous single-center finding (1,2) that  $\Delta K^{trans}$  is the best single discriminator of benign and malignant breast lesions among all the DCE-MRI pharmacokinetic parameters. The fact that this validation was achieved using DCE-MRI data acquired in three institutions with different protocols and from populations of different screening backgrounds provides strong evidence for the use of quantitative imaging methods in future clinical breast MRI practice. The two one-split trees are equivalent to using  $\Delta K^{trans}$  as the only diagnostic marker. The unequally-weighted one-split tree slightly sacrificed the overall diagnostic accuracy (91.8% to 90.8%) and specificity (95.7% to 92.1%) to improve the sensitivity (82.1% to 87.5%) over the equally-weighted one-split tree. With either weighting scheme, there is always a trade-off between the sensitivity and specificity. Since a false negative diagnosis is far worse than a false positive one, the goal is to achieve the highest sensitivity possible with the least sacrifice in specificity. By just adding one more split, or another DCE-MRI parameter as diagnostic marker, the sensitivity improved from 82.1% to 89.3% for the equally-weighted tree, and from 87.5% to 100% for the unequally-weighted tree. For the latter, even though the specificity dropped dramatically from 92.1% to 66.9%, the overall accuracy of 76.4% still substantially outperformed the standard-of-care breast imaging methods (28.7%). By including more pharmacokinetic parameters for diagnosis, the three-, four-, and five-split classification trees with unequal-weighting (Fig. 2) maintained excellent sensitivities while incrementally improving specificity and overall accuracy.



**Table 1. Breast Cancer Diagnostic Performance**

Classification Tree		Sensitivity	Specificity	Overall Accuracy
Clinical breast imaging		N/A	N/A	28.7%
Equally-weighted (the consequence of false positive is the same as false negative)	One-split	82.1%	95.7%	91.8%
	Two-split	89.3%	93.5%	92.3%
Unequally-weighted (the consequence of false negative is 4-10 times worse than false positive)	One-split	87.5%	92.1%	90.8%
	Two-split	100%	66.9%	76.4%
	Three-split	100%	77.7%	84.1%
	Four-split	100%	82.0%	87.2%
	Five-split	98.2%	86.3%	89.7%

**Conclusion:** Using multicenter breast DCE-MRI data, this study has validated  $\Delta K^{trans}$  as the best single DCE-MRI marker for breast cancer diagnosis. The results from the classification tree approach suggest that incorporating other DCE-MRI pharmacokinetic parameters may further improve upon the diagnostic performance of  $\Delta K^{trans}$ . The decision to select the most effective classification scheme depends on both the clinical judgment and the practical considerations. Further evaluation with a larger data set and external validation of these classification schemes are warranted, which will help establish a final practical model.

**Grant Support:** NIH RO1-CA120861 and UO1-CA154602.

**References:** 1. Huang *et al.* *Radiology* 2011;261:394-403. 2. Huang *et al.* *PNAS* 2008;105:17943-48. 3. Tofts *et al.* *J Magn Reson Imaging* 1999;10:223-32. 4. Springer *et al.* *Proc Intl Soc Magn Reson Med* 2011;19:3097. 5. Breiman *et al.* *Classification and Regression Trees*. Wadsworth, 1984.