Statistical Metrics to Determine When Water Exchange Should be Incorporated into DCE-MRI Analysis: Simulations and Experimental Breast Cancer Results

T. Yankeelov[1], E. B. Welch[2], A. Chakravarthy[3], R. Lee[4], D. Freehardt[5], I. Mayer[6], I. Meszoely[7], M. Kelley[7], J. Means-Powell[7], and J. C. Gore[4]

[1]Institute of Imaging Science, Vanderbilt University, Nashville, Tennessee, United States, [2]Philips Medical Systems, United States, [3]Radiation Oncology, Vanderbilt University, United States, [4]Institute of Imaging Science, Vanderbilt University, United States, [5]Clinical Trials, Vanderbilt University, United States, [6]Oncology, Vanderbilt University, United States, [7]Surgery, Vanderbilt University, United States

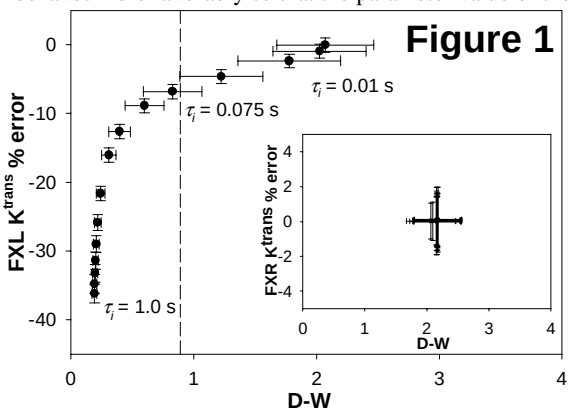
INTRODUCTION We apply three statistical measures to determine when water exchange effects should be incorporated in DCE-MRI data analysis. Simulated DCE-MRI data were analyzed with (fast exchange regime, FXR) and without (fast exchange limit, FXL) the incorporation of the effects of transcytolemmal water exchange to test the hypothesis that as water exchange effects become more pronounced, the χ^2, Durbin-Watson statistic, and the Akaike Information Criteria (AIC) all favor the FXR analysis. Additionally, DCE-MRI data were obtained at 3T from 12 breast cancer patients and analyzed with both models to test the predictions of the simulations. As DCE-MRI ultimately aims to aid diagnosis and assess treatment response, the choice of analytical model is of major importance.

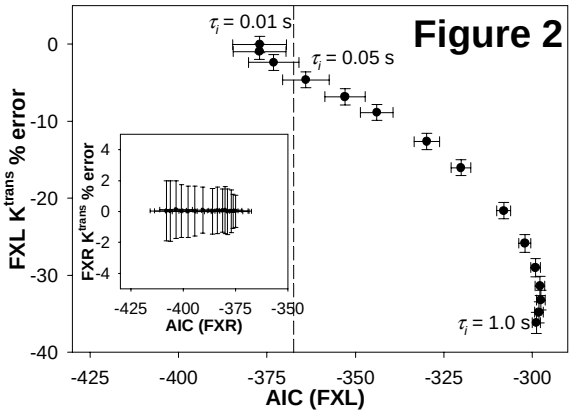
METHODS

Theory DCE-MRI models which include the effects of transcytolemmal water exchange (i.e., FXR models) return estimates of the volume transfer constant (K^{trans}), extravascular extracellular volume fraction (v_e), and the intracellular water lifetime (τ_i), whereas FXL models report only on K^{trans} and v_e. In simulations, we allow τ_i to vary over a large range while fixing K^{trans} and v_e to test the hypothesis that as τ_i increases the FXR model will provide more favorable values of χ^2, D-W, and AIC. The D-W statistic is a well-known test for detecting serial correlation in residuals (1) and the AIC is a method to select the model which best balances goodness of fit with number of free parameters (2). As τ_i increases, the FXR and FXL return different estimates of parameter values (3). The χ^2, D-W, and AIC determine which model behaves more favorably so that the parameter value of the statistically preferred model is chosen.

Simulations Using a standardized arterial input function (REF), we generated R_1 time courses *via* the FXR model. Reasonable K^{trans} and v_e values (0.5 min^{-1} and 0.3, respectively) were selected and the τ_i value was allowed to range from 0.01 to 1.0s. The statistical metrics were applied to both models for each parameter set as τ_i increased to assess which model is statistically superior.

Data Acquisition Twelve patients with locally advanced breast cancer were enrolled in our clinical study. Patients underwent DCE-MRI on a Philips 3.0 T Achieva MR scanner prior to neoadjuvant chemotherapy. Data for a T_1 map were acquired *via* a 3D GRE multi-flip angle approach with a *TR**TE* of 7.9\1.3 ms and ten flip angles of 2,4,…20°. The acquisition matrix was 192×192×25 over a FOV of $(22\ cm)^2$ with slice thickness of 5 mm, NEX=2 and SENSE=2. A catheter placed within an antecubital vein delivered 0.1 mmol/kg of Magnevist over 20 seconds for the DCE study which used *TR**TE*\α= 7.9ms\1.3ms\20°. The χ^2, D-W, and AIC were computed for each tumor pixel and whole tumor ROI.





RESULTS For brevity we present only the D-W and AIC results. Simulation results are presented in **Figure 1** in which the D-W are plotted versus K^{trans} error as estimated for each model. It is evident that as τ_i increases (from right to left) the D-W decreases and there is an increase in the percent error of the returned FXL estimate of K^{trans}. The vertical dashed line indicates the D-W of 0.98; below this value there is positive serial correlation in the residuals of the best fit line (at the $p<0.01$ level) and this indicates the model is non-optimal. For these parameter sets, that value is τ_i = 0.075s. In the FXR (inset) D-W and K^{trans} error is independent of τ_i as all values are clustered at ~0% error and a D-W value of ~2.0 indicating no serial correlation. **Figure 2** presents the AIC data which indicate that as τ_i increases (from left to right) the AIC decreases and there is an increase in the percent error of the returned FXL estimate of K^{trans}. The vertical line indicates the τ_i value (0.05s) at which the AIC favors the FXR model over the FXL. The FXR data are presented in the inset.

For the breast imaging data, the χ^2, D-W, and AIC all returned statistically different values for the models and these results are summarized in the **Table**. The FXR analyses of the DCE-MRI data resulted in a significant reduction in percentage of voxels showing serial correlation of residuals: 42.6% +/- 12.6% and 21.5% +/- 7.7% for the FXL and FXR, respectively ($P<0.001$). In >95% of voxels, the AIC indicated that the FXR model was superior ($P<0.001$). This translated into significant differences in parameters extracted: 0.28 min^{-1} +/- 0.33 min^{-1} (FXR) versus 0.17 min^{-1} +/- 0.17 min^{-1} (FXL) for K^{trans} ($P<0.05$), and 0.37 +/- 0.20 (FXR) versus 0.10 +/- 0.06 (FXL) for v_e ($P<0.001$).

DISCUSSION In the majority of voxels, the FXR model results in a better description of the DCE time courses than the FXL model as determined by the the χ^2, Durbin-Watson statistic, and the Akaike Information Criteria, and this results in parameter values that are statistically different between the two models. The simulations predict that after τ_i becomes above ~0.1 s, the FXR model will be preferred and this is what we see in the breast cancer data. We conclude that, at least in the case of human breast cancer, water exchange effects should be explicitly incorporated in the analysis of DCE-MRI data.

χ^2	D-W	AIC	K^{trans}	v_e
$P < 0.05$	$P < 0.001$	$P < 0.001$	$P < 0.05$	$P < 0.001$

Table: Significant differences between the statistical metrics and between parameters returned by each model.

REFERENCES 1. Draper and Smith. Applied Regression Analysis. 1998. 2. Akaike IEEE TAC 1974;19:716-723. 3. Yankeelov, et al. MRM 2003;50:1151-69.

ACKNOWLEDGEMENTS NIH U24 CA126588, 1K25 EB005936-01