Multivariate Analysis of 2D Correlated Breast Spectra

N. Wyckoff[1], N. Binesh[1], H. Chung[1], K. Yue[1], N. DeBruhl[1], M. A. Thomas[1]

[1]David Geffen School of Medicine at UCLA, Radiological Sciences, Los Angeles, CA, United States

Synopsis

The goal of this work was to differentiate the localized 2D COSY spectra from malignant and benign breast tumor patients and healthy controls, using linear discriminant analysis and logistic regression. Spectra were recorded using an L-COSY sequence optimized on a 1.5-T MRI scanner. The study sample included 23 healthy women, 12 malignant tumor patients and 2 benign tumor patients. Our preliminary results clearly indicate that an accuracy of 90% can be achieved using the stepwise logistic regression classifier, compared to 80% accuracy using the linear classifier, in distinguishing between healthy fatty tissue and invasive carcinoma.

Introduction

Magnetic resonance spectroscopy (MRS) allows the non-invasive measurement of choline, water, lipids, saturated and unsaturated fatty acids.[1-3] The purpose of this study was to use linear discriminant analysis and logistic regression to distinguish among the volume-localized two-dimensional L-COSY spectra[4] recorded from healthy breast tissue, benign breast tumors and malignant breast tumors.

Methods

The study sample included 12 women with invasive ductal carcinomas (average age 54 years), two women with benign fibroadenomas (average age 26.5 years) and 23 healthy women (average age 39.5 years). Each subject was scanned in a 1.5 T whole body MRI/MRS scanner (GE Medical Systems, Waukesha, WI). A body RF coil transmitted the RF pulses, and a dedicated phased-array breast coil (GE Medical Systems, Waukesha, WI) received the signals. L-COSY spectra were recorded using the following parameters: echo time (TE) of 30 ms, pulse repetition time (TR) of 2 seconds, 8 excitations per Δt_1, and 40 or 46 increments.

Spectra were acquired from localized volumes of interest (VOIs) in each subject's breast tissue. The size of each VOI was 15 x 15 x 15 mm^3 in breast cancer patients and 10 x 10 x 10 mm^3 in healthy controls. For each patient, at least one spectrum was acquired from a VOI within breast tumors. Whenever possible, another spectrum was acquired from a VOI in the unaffected contralateral breast. In healthy controls, spectra were recorded from both glandular and fatty tissue VOIs.

Raw two-dimensional MRS data files were processed using the Felix program (Accelerys Inc., San Diego, CA). The diagonal peaks due to water (4.77 ppm) and fat (1.4 ppm) were identified. Some spectra also had diagonal peaks due to choline (3.2-3.5 ppm) and purine-pyrimidine nucleotides (8.0 ppm and 8.2 ppm), in addition to the cross-peaks of unsaturated fatty acids ((F2=1.97 ppm, F1=5.25 ppm) and (F2=2.77 ppm, F1=5.30 ppm)).[3] The volumes of all peaks were integrated, and the ratios among various metabolite volumes were computed.

Statistical classifiers were used to make several distinctions between different tissue types: invasive carcinoma and healthy glandular tissue, invasive carcinoma and healthy fatty tissue, and fibroadenoma and invasive carcinoma. To make each distinction, the metabolite ratios with significant differences in their means were input as candidate features to two stepwise statistical classifiers: linear discriminant analysis and logistic regression. The linear discriminant classifier includes features to a linear classification function in a stepwise fashion, accounting for interactions between features. Stepwise logistic regression is a progressive method to construct a probability-based classification function.

Results

Samples of healthy fatty tissue and invasive carcinoma differed significantly in their mean values of water/fat ($p=0.001$) and water/unsaturated fatty acids ($p<0.02$ for the ratios of water to both unsaturated fatty acid peaks). The linear discriminant classifier selected one of these features, water/unsaturated fatty acids, for discrimination between healthy fatty tissue and carcinoma tissue. It classified 27/28 fatty voxels (96.4%) and 5/12 carcinoma voxels (41.7%) correctly (80.0% accuracy). Using a pair of features, water/fat and water/unsaturated fatty acids, the logistic regression classifier was able to identify 26/28 fatty tissue samples (92.9%) and 10/12 invasive carcinomas (83.3%) correctly (90.0% classification accuracy). These results are summarized in Table 1.

Healthy glandular tissue and invasive carcinoma tissue differed significantly only in their mean values of water/purine-pyrimidine ($p=0.036$). This ratio yielded 100% classification accuracy using linear discriminant analysis (2/2 voxels from glandular tissue and 4/4 invasive carcinomas). Logistic regression was not feasible for this classification problem, as only two glandular tissue voxels were available.

Between benign fibroadenomas and invasive carcinomas, only water/fat differed significantly ($p=0.002$). The linear classifier identified both benign fibroadenomas (100%) and 12/12 invasive carcinomas (100%) correctly using water/fat. Logistic regression was not feasible for this classification problem, as only two fibroadenomas were available.

Discussion

Two-dimensional L-COSY can identify the major metabolites found in human breast tissue, detecting both diagonal and cross peaks. From the L-COSY peak volumes, the stepwise logistic regression classifier achieved a greater accuracy (90.0%) than the linear classifier (80.0%) when distinguishing between healthy fatty tissue and invasive carcinoma. The ratio of water/fat, known to vary between healthy women and women with breast cancer[3] was found effective when combined with the ratio of water to unsaturated fatty acids. Linear discriminant analysis using L-COSY data may be a valuable method to distinguish benign breast tumors from breast cancers; however, more benign cases are needed to verify this claim.

References

1. Kvistad KA, Bakken IJ, Gribbestad IS, Ehrnholm RT, Lundgren S, Fj∅sne HE, Haraldseth O. Characterization of neoplastic and normal human breast tissues with in vivo ^{1}H MR spectroscopy. JMRI 1999; 10:159-164

2. Cecil KM, Schnall MD, Siegelman ES, Lenkinski RE. The evaluation of human breast lesions with magnetic resonance imaging and proton magnetic resonance spectroscopy. Breast Can Res Treat 2001; 68:45-54

3. Jagannathan NR, Singh M, Govindaraju V, Raghunathan P, Coshic O, Julka PK, Rath GK. Volume localized *in vivo* proton MR spectroscopy of breast carcinoma: variation of water-fat ratio in patients receiving chemotherapy. NMR Biomed 1998; 11:414-422

4. Thomas MA, Binesh N, Yue K, DeBruhl N. Volume-Localized Two-Dimensional Correlated Magnetic Resonance Spectroscopy of Human Breast Cancer. JMRI 2001; 14(2):181-186

	Healthy Fatty Tissue, Predicted	Invasive Carcinoma, Predicted	Total
Healthy Fatty Tissue, Actual	27 (96.4%)	1 (3.6%)	28 (100.0%)
Invasive Carcinoma, Actual	7 (58.3%)	5 (41.7%)	12 (100.0%)

Table 1a. Classification table showing the stepwise linear discriminant analysis classifier's diagnoses of healthy fatty tissue and invasive carcinoma tissue.

	Healthy Fatty Tissue, Predicted	Invasive Carcinoma, Predicted	Total
Healthy Fatty Tissue, Actual	26 (92.9%)	2 (7.1%)	28 (100.0%)
Invasive Carcinoma, Actual	2 (16.7%)	10 (83.3%)	12 (100.0%)

Table 1b. Classification table showing the stepwise logistic regression classifier's diagnoses of healthy fatty tissue and invasive carcinoma tissue.