Prediction of genetic risks in schizophrenia: A DTI based pattern classification study

M. Ingalhalikar[1], S. Kanterakis[1], D. Parker[1], R. E. Gur[2], R. C. Gur[2], and R. Verma[1]

[1]Section of Biomedical Image Analysis, Department of Radiology, University of Pennsylvania, Philadelphia, PA, United States, [2]Brain Behavior Laboratory, Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, United States

Background and Objective: Differences in brain structure patterns in asymptomatic family members (FM) of schizophrenia patients (SCZ) have demonstrated an effect of the genetic predisposition to the disease. Identification of the affected FM using behavioral markers is limited by the phenotypic heterogeneity of SCZ [1]. Prior neuroimaging studies on FM were primarily based on structural images and other modalities like EEG [2]. The goal of this work is to quantify the degree to which the brains of the asymptomatic biological relatives match the patients, by employing gender specific high dimensional classifiers based on DTI features of anisotropy and diffusivity computed from a collection of anatomically meaningful regions of interest (ROI) over the entire brain.

Method: Our data consisted of 58 controls (28 females and 30 males), 70 SCZ (38 females and 32 males) and 30 FM (20 females and 10 males). Images were acquired on Siemens 3T Trio™ scanner using a 12 channel head coil. Diffusion tensor imaging was performed using a single shot spin-echo, echo-planar sequence with the following parameters: TR/TE=6400/97 ms, b-value of 1000 s/mm^2 and 64 gradient directions. All the DTI images were spatially normalized to a standard atlas [3] that included 176 labeled ROIs. The average DTI features of fractional anisotropy (FA) and mean diffusivity (MD) computed from these ROIs, were used as the input feature set. Thus the feature vector for each subject consisted of 352 total features (176 FA and 176 MD). In the next step, feature selection was performed to remove the redundant features and maintain only the relevant features. We implemented recursive feature elimination (RFE) introduced by Guyon [4] for efficient classification and to increase the generalizability of the classifier. These features were then used to train a non-linear support vector machine (SVM) classifier using a Gaussian kernel. The classifier model was validated using the jackknife method. In the jackknife or leave-one-out (LOO) validation, one sample was chosen for testing, while other samples were used for feature selection and classifier modeling. The classifier was then evaluated based on the classification result of the left out subject. The classifier produced a probabilistic score that represented the level of abnormality in the subject. Finally, after validation, a classifier model over all the controls and patients was applied to the FM and an abnormality score referred as the genetic pattern score, was computed for each FM. The genetic pattern score has a range of -1 to +1 describing the extent to which the brain patterns of the FM match the patterns of patients or controls (where 0 to -1 is the patient spectrum and 0 to +1 is the control spectrum).

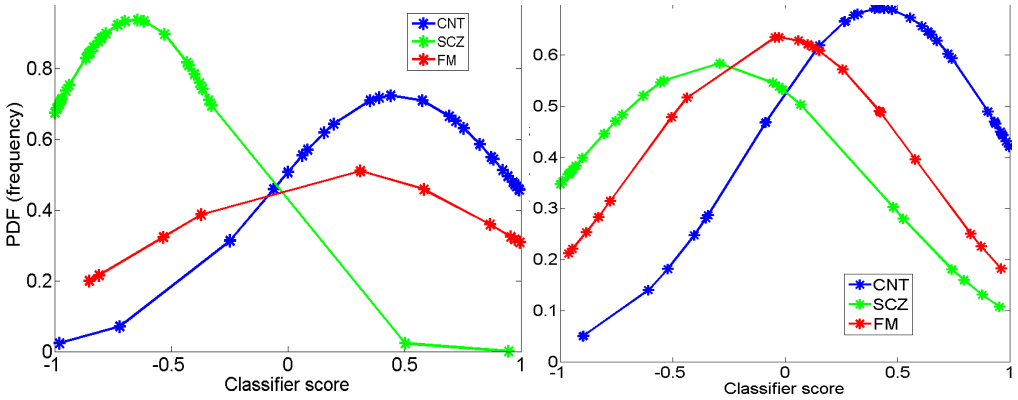

Fig 1. Plot showing LOO classification (abnormality) scores vs. frequency of occurrence of the score value (PDF) where green indicates patients and blue indicates controls. Each subject's score is represented with a star. The red stars represent the abnormality scores of the tested family members. (a) Classification result for males (b) females.

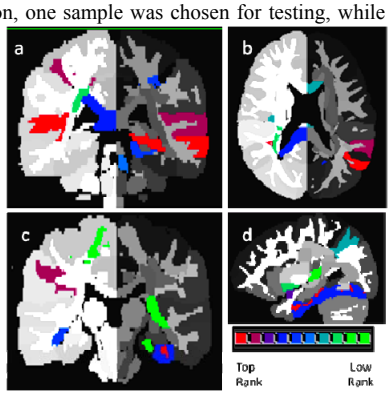

Fig.2 Top ranked ROI's that contributed largely towards classification based on mutual information ranking scheme (a-b) males (c-d) females.

Results: Fig.1 shows the normal probability density (PDF), which represents the likelihood of each LOO score plotted against the LOO score. The average LOO classification for males was 88.70% while for females it was 75.04 %. The FM test scores are plotted in red (Fig.1). Out of these, 40% of males and 50% of females were classified as patients. Fig.2 displays the ROIs that were selected frequently in the LOO validation. The regions were overlaid on the template ROI image and color-coded according to their rank. In males, the main discriminating regions with FA features included body of corpus callosum, splenium of the corpus callosum, anterior limb of internal capsule and superior temporal white matter while the MD features involved superior temporal gyrus and middle temporal gyrus. In females, regions with FA changes were mainly the cuneus white matter, uncinate fasciculus, inferior fronto-occipital fasciculus and retrolenticular part of internal capsule while MD features included fusiform gyrus, middle-central white matter and inferior longitudinal fasciculus.

Discussion: This study investigated the degree of brain pattern changes exhibited by the asymptomatic family members of patients with schizophrenia. A non-linear SVM classifier was constructed using FA and MD features computed from atlas-based anatomical ROIs of patients and controls. The FM were then tested over the learned classifier. Results suggest that FM demonstrate an endophenotypic trait since in both cases more than 40% of FM were classified as patients. Females displayed scores that were closer to the patient spectrum than males exhibiting increased genetic vulnerability to SCZ. Furthermore, as the classifier model was learned over the top ranked ROIs, the FM that were classified as patients were expected to have significant differences in these brain regions while the ones lying in between were potentially undergoing changes in these areas. Moreover, the genetic pattern score could locate the asymptomatic FM over the patient-control spectrum, making it a potential prognostic biomarker. In conclusion, we provide evidence that high-dimensional pattern classification can identify complex and subtle structural endophenotypes that are shared by SCZ and unaffected FM.

References: 1. Tsuang et al. Biological Psychiatry, 47(3), 2000.
2. Seidman et al. Biological Psychiatry 46(7) 1999.
3. Wakana et al. Radiology, 230(1), 2004.
4. Guyon et al. Machine learning, 46, 2002.

Acknowledgement: This work was supported by NIH grants MH060722, MH042191and MH064045.