Personalized prediction of brain fiber integrity in 396 young adults based on genotyping of multiple common genetic variants

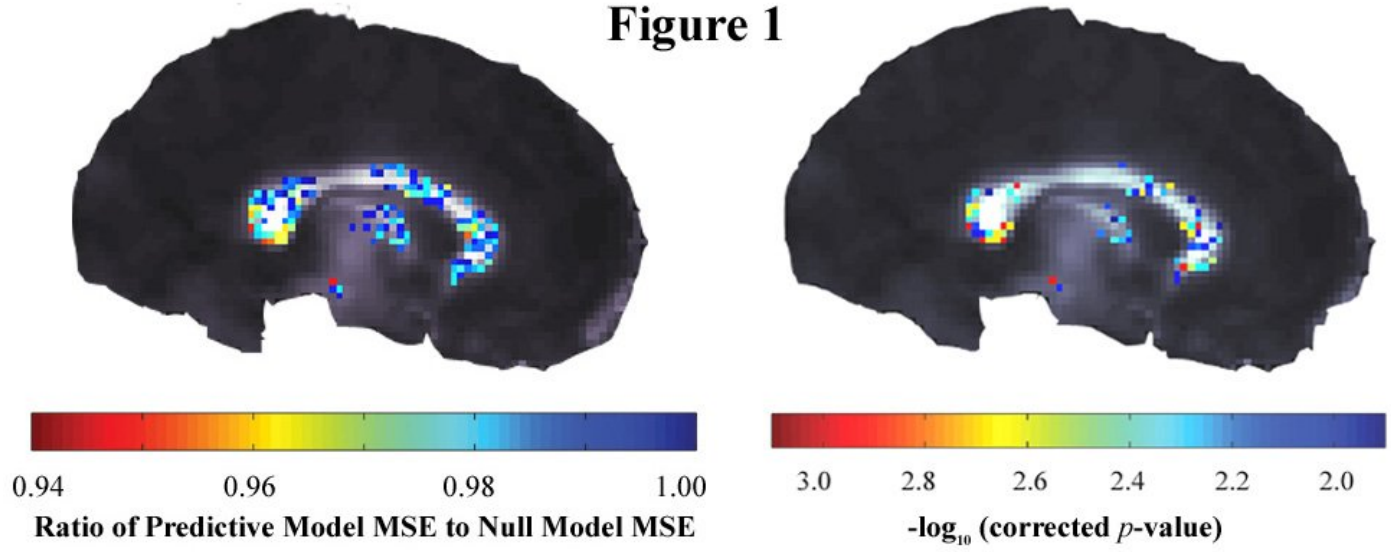
Omid Kohannim[1], Neda Jahanshad[1], Meredith N. Braskie[1], Jason L. Stein[1], Ming-Chang Chiang[1,2], April H. Reese[1], Arthur W. Toga[1], Katie L. McMahon[3], Greig I. de Zubicaray[4], Sarah E. Medland[5], Grant W. Montgomery[5], John B. Whitfield[5], Nicholas G. Martin[5], Margaret J. Wright[5], Paul M. Thompson[1]

[1]Laboratory of Neuro Imaging, Dept. of Neurology, UCLA School of Medicine, Los Angeles, CA, USA
[2]National Yang-Ming University, Taipei, Taiwan
[3]University of Queensland, Center for Advanced Imaging, Brisbane, Australia
[4]University of Queensland, School of Psychology, Brisbane, Australia
[5]Queensland Institute of Medical Research, Brisbane, Australia

We recently discovered several commonly carried genetic variants that individually show significant associations with white matter fiber integrity in the living brain, as measured by diffusion tensor imaging (DTI). We considered six candidate single nucleotide polymorphisms (SNPs) in the following genes, which we previously reported to influence brain white matter integrity: neurotrophic tyrosine kinase, receptor, type 1 (*NTRK1*), clusterin (*CLU*), brain-derived neurotrophic factor (*BDNF*), fat mass and obesity associated (*FTO*), raftlin, lipid raft linker 1 (*RFTN1*) and the hemochromatosis gene, *HFE*. Based on genotyping these SNPs, we set out to predict individual brain integrity as measured by fractional anisotropy (FA) in 4-Tesla DTI. In 396 healthy young twins and siblings (mean age: 23.7±2.2 years), we used stepwise multiple linear regression to reveal that a combination of *NTRK1*, *CLU*, *HFE*, and *RFTN1* SNPs yielded the regression model that explained most variance in brain integrity (R^2=0.067; p=4.91x10^{-5}) quantified using mean FA across all brain voxels. Regressions were adjusted for sex and age; kinship was accounted for with mixed effect modeling.

In a subset of 247 unrelated subjects (mean age: 23.8±2.3 years) we then investigated the voxelwise prediction of FA with the above 4-SNP combination using support vector regression. The numbers of SNP minor alleles were treated as training features and FA measures as outputs. Using voxelwise leave-one-out cross-validation, we computed mean squared errors (MSEs) for the FA predictions, which were better than those of a null predictor in widespread brain regions. In the midsagittal plane, for instance, the model was significantly predictive at 23% of the voxels (mostly callosal genu and splenium), with a corrected p-value threshold of 0.011 based on 1000 permutations, where FA measures were randomized at each voxel (Figure 1). Although they require further replication, such multi-gene, machine learning models to predict imaging-derived measures may be useful for early, personalized risk assessment of impaired brain integrity.

Figure 1


0.94
0.96
0.98
1.00
Ratio of Predictive Model MSE to Null Model MSE
3.0
2.8
2.6
2.4
2.2
2.0
-log10 (corrected p-value)