Crowdsourced validation of a machine-learning classification system for autism and ADHD.

Duda M, Haber N, Daniels J, Wall DP.

Transl Psychiatry. 2017 May 16;7(5):e1133. doi: 10.1038/tp.2017.86.

Abstract

Autism spectrum disorder (ASD) and attention deficit hyperactivity disorder (ADHD) together affect >10% of the children in the United States, but considerable behavioral overlaps between the two disorders can often complicate differential diagnosis. Currently, there is no screening test designed to differentiate between the two disorders, and with waiting times from initial suspicion to diagnosis upwards of a year, methods to quickly and accurately assess risk for these and other developmental disorders are desperately needed. In a previous study, we found that four machine-learning algorithms were able to accurately (area under the curve (AUC)>0.96) distinguish ASD from ADHD using only a small subset of items from the Social Responsiveness Scale (SRS). Here, we expand upon our prior work by including a novel crowdsourced data set of responses to our predefined top 15 SRS-derived questions from parents of children with ASD (n=248) or ADHD (n=174) to improve our model's capability to generalize to new, 'real-world' data. By mixing these novel survey data with our initial archival sample (n=3417) and performing repeated cross-validation with subsampling, we created a classification algorithm that performs with AUC=0.89±0.01 using only 15 questions.