Classification Accuracy of Neuroimaging Biomarkers in Attention-Deficit/Hyperactivity Disorder: Effects of Sample Size and Circular Analysis

Pulini AA, Kerr WT, Loo SK, Lenartowicz A.


Abstract

BACKGROUND:
Motivated by an inconsistency between reports of high diagnosis-classification accuracies and known heterogeneity in attention-deficit/hyperactivity disorder (ADHD), this study assessed classification accuracy in studies of ADHD as a function of methodological factors that can bias results. We hypothesized that high classification results in ADHD diagnosis are inflated by methodological factors.

METHODS:
We reviewed 69 studies (of 95 studies identified) that used neuroimaging features to predict ADHD diagnosis. Based on reported methods, we assessed the prevalence of circular analysis, which inflates classification accuracy, and evaluated the relationship between sample size and accuracy to test if small-sample models tend to report higher classification accuracy, also an indicator of bias.

RESULTS:
Circular analysis was detected in 15.9% of ADHD classification studies, lack of independent test set was noted in 13%, and insufficient methodological detail to establish its presence was noted in another 11.6%. Accuracy of classification ranged from 60% to 80% in the 59.4% of reviewed studies that met criteria for independence of feature selection, model construction, and test datasets. Moreover, there was a negative relationship between accuracy and sample size, implying additional bias contributing to reported accuracies at lower sample sizes.

CONCLUSIONS:
High classification accuracies in neuroimaging studies of ADHD appear to be inflated by circular analysis and small sample size. Accuracies on independent datasets were consistent with known heterogeneity of the disorder. Steps to resolve these issues, and a shift toward accounting for sample heterogeneity and prediction of future outcomes, will be crucial in future classification studies in ADHD.