Can population registry data predict which children with ADHD are at risk of later substance use disorders?

By Dr. Jessica K Edwards

The first study to examine the potential of machine learning in early prediction of later substance use disorders (SUDs) in youth with ADHD has been published in the *Journal of Child Psychiatry and Psychology*. Funded by the Co-CA-project, Yanli Zhang-James and colleagues collected information on psychiatric and somatic diagnoses, family history of these disorders, socioeconomic status, and birth complications from Swedish national registers for >19,000 young people with ADHD. Using data available by age 17 to predict a diagnosis of a SUD by age 19, the researchers trained various machine learning models in a subset of the sample, and then tested the models' prediction performance in a different subset.

Zhang-James *et al*. found that their models adequately predicted later SUDs in these young people with ADHD. A longitudinal model was able to predict later SUD risk from as young as 2 years-of-age. The researchers identified that an early ADHD diagnosis was an important predictor of a lower risk for SUDs, which they explain might reflect benefits of early diagnosis and treatment of ADHD in children. They also found that lower socioeconomic status was an important predictor, which they believe reaffirms the importance of policies that alleviate poverty, deprivation and inequality.

These findings have important implications. "Being able to identify children who are at increased risk for developing substance use disorders later in adolescence or adulthood could help the school-education system and service providers to develop and deliver prevention programs for the targeted children at early ages", says Zhang-James. "Furthermore, the ability to identify risk predictors may also help to inform policy-makers, parents, teachers and clinicians and guide the development of more effective measures for prevention and risk-reduction".

Zhang-James *et al*. concede that there are limitations to their current model, explaining that additional features and a larger sample size are needed to improve the predictive accuracy. "Our models' sensitivity and specificity were moderate", explains Zhang-James. "As such, future studies are now needed to improve the predictive power, as well as the explainable mechanisms underlying the prediction models".