REVIEW ARTICLE   OPEN

Check for updates

Machine learning in attention-deficit/hyperactivity disorder: new approaches toward understanding the neural mechanisms

Meng Cao[1], Elizabeth Martin[2] and Xiaobo Li [1]✉

Attention-deficit/hyperactivity disorder (ADHD) is a highly prevalent and heterogeneous neurodevelopmental disorder in children and has a high chance of persisting in adulthood. The development of individualized, efficient, and reliable treatment strategies is limited by the lack of understanding of the underlying neural mechanisms. Diverging and inconsistent findings from existing studies suggest that ADHD may be simultaneously associated with multivariate factors across cognitive, genetic, and biological domains. Machine learning algorithms are more capable of detecting complex interactions between multiple variables than conventional statistical methods. Here we present a narrative review of the existing machine learning studies that have contributed to understanding mechanisms underlying ADHD with a focus on behavioral and neurocognitive problems, neurobiological measures including genetic data, structural magnetic resonance imaging (MRI), task-based and resting-state functional MRI (fMRI), electroencephalogram, and functional near-infrared spectroscopy, and prevention and treatment strategies. Implications of machine learning models in ADHD research are discussed. Although increasing evidence suggests that machine learning has potential in studying ADHD, extra precautions are still required when designing machine learning strategies considering the limitations of interpretability and generalization.

INTRODUCTION

Attention-deficit/hyperactivity disorder (ADHD) is one of the most prevalent neurodevelopmental disorders, affecting ~5–8% of children worldwide [1, 2]. For about 60% children with ADHD, the symptoms persist into adulthood [3, 4]. Individuals with ADHD have poorer educational and social outcomes, increased injury incidences during daily activities [5, 6], and an elevated risk of developing more severe mental disorders [7–9]. ADHD is a highly heterogeneous disorder [10]. For example, sex, genetic, and environmental factors have been implicated in the presentation of ADHD [11–13]. There is also diverging evidence regarding the developmental trajectories and comorbidities of individuals with ADHD [14, 15]. Considering the high prevalence and life-long consequences of ADHD, early detection, accurate diagnosis, and efficient treatments are highly desired. However, the field currently lacks a comprehensive understanding of the relevant neural mechanisms and is far from reaching an agreement regarding efficient treatment strategies.

Extensive studies have attempted to characterize ADHD in terms of neuropsychological performance, brain anatomy and functional responses, and genetic risk factors. Cognitive deficits in executive function, reaction time, vigilance, inhibition control, sustained attention, and working memory have been reported in ADHD [16–18]. Neuroimaging studies using T1-weighted magnetic resonance imaging (MRI), functional MRI (fMRI), resting-state fMRI (rs-fMRI), and electroencephalogram (EEG) have reported widespread and inconsistent anatomical and functional alterations in children with ADHD, including frontal lobe, parietal lobe, temporal lobe, thalamus [19–22]. Genome-wide association studies have also revealed several variants associated with ADHD [23–25]. In addition, the treatment of ADHD shows inconsistent results, with evidence suggesting that 30% of ADHD patients respond poorly to the most common ADHD medication [26, 27]. The existing evidence suggests that ADHD may not have a single etiological source but rather a combined effect of multiple subtle anomalies. Such a complex etiology is difficult to detect using parametric statistical methods, and interactions between widespread alterations have not been successfully translated into clinical practice due to the limited capacities of conventional analytical methods.

The increasing accessibility of machine learning models has led to increased interest in applying such models to investigate psychiatric disorders. Generally, machine learning models are mathematical models that learn complex patterns in an existing dataset. These learned patterns can then be used for prediction in a novel dataset (e.g., patient vs control participant, symptom scores), as well as to highlight the most important variables in creating this prediction. Machine learning models have proved effective in capturing the complex interactions between discrete alterations in schizophrenia, Alzheimer's disease, and autism spectrum disorder (ASD) [28, 29]. Most psychiatric studies have developed models that differentiate patient groups and controls using classification algorithms like SVM, random forest and linear discriminant analysis (LDA). Others predict symptom severity or behavioral performance using regression algorithms, for example,

---

[1]Department of Biomedical Engineering, New Jersey Institute of Technology, Newark, NJ, USA. [2]Icahn School of Medicine at Mount Sinai, New York, NY, USA.
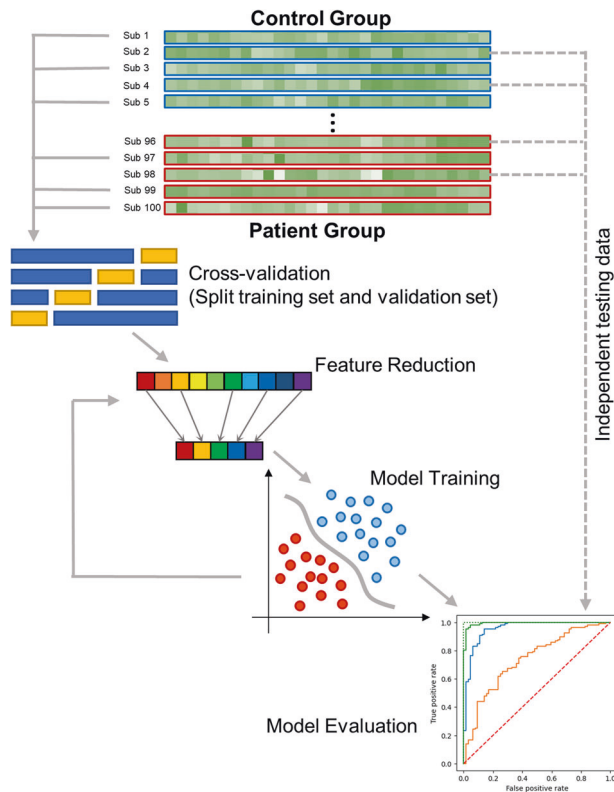✉email: xli.aecom@gmail.com

Fig. 1 Overview of machine learning steps. Typical steps, including data splitting, feature reduction, and model training, used in machine learning studies of patients with psychiatric conditions.

random forest regression, support vector regressor, and elastic net regression. The general steps involve data splitting, feature reduction, and model training, as shown in Fig. 1. The original data is first split into training set (for features selection and training machine learning models), validation set (for validating and tuning parameters of trained models), and testing set (for evaluating the model performance). Before using the training set to train the model, feature reduction is usually performed using feature selection or feature fusion to increase the efficiency of the training process and reduce the chance of overfitting. During the training process, adjustments are made to the model based on the model performance in the validation set. Finally, the effectiveness of the classification models is evaluated in the independent testing set using accuracy, specificity, sensitivity, or area-under-the-curve (AUC), and the performance of the regression models is evaluated by the mean square error or the correlation [30]. However, many studies to date report model performance based on the results of cross-validation processes without having independent set testing. Such studies can yield less reliable or less generalizable results, and therefore, extra precautions are needed when interpreting the findings of these studies. Using this general process, most machine learning studies have been able to differentiate patients with psychiatric disorders and controls with AUC from 60 to 90% [29]. For diagnostic purposes, models with an AUC of less than 60% were considered to have bad performance, while models with an AUC of more than 80% were considered as having very good performance [31].

Increasing evidence suggests that machine learning techniques are beneficial in improving ADHD diagnosis, understanding neurobiological substrates, and evaluating treatment strategies (for reviews [32–35]). For example, current diagnoses of ADHD require extensive interviews of parents and teachers (for child-hood ADHD) and of patients (for adult ADHD) on observations of

current and past ADHD symptoms and subsequent impairment of daily functioning. Machine learning studies can learn from sufficient samples to sort and select the most relevant interview questions for accurate diagnoses. Therefore, machine learning has the potential to facilitate the development of more efficient diagnostic procedures for ADHD. Additionally, the ability to predict treatment outcomes using machine learning may contribute to the emergence of precision medicine (for reviews on these topics, see [32–35]). The abundance of machine learning investigations in neuroimaging studies has partially been enabled by the public release of the ADHD-200 dataset [36], which has allowed exploration in automating ADHD diagnosis [35]. Recent public datasets like Adolescent Brain Cognitive Development (ABCD) dataset also boosted the machine learning research in ADHD [37].

The majority of existing machine learning studies in ADHD focus on developing classification algorithms between ADHD patients and controls or patients with other comorbid disorders. Undoubtedly, machine learning algorithms were best suited for predictive purposes. However, the sample size is often a study limitation. Large and high-quality datasets are difficult to obtain due to the substantial efforts required in the collection and maintenance processes. Studies in small-size samples tend to get over positive results and less generalizable models without including appropriate validation processes (e.g., lack of independent validation set or leakage between the validation set and training set) [38, 39]. For example, leakage between the training set and validation/testing set, or feature selection/reduction before data splitting, can lead to substantial bias in the machine learning models [40]. Despite these issues, there are merits of studies with smaller sample sizes. Relative to large sample studies, small sample studies may be able to recruit more homogeneous groups. In these homogeneous samples, after proper validation and evaluation process, the inference of important features, rather than building accurate classification models, is more beneficial to the research in ADHD. Due to the ability of machine learning models to extensively learn the complex patterns in a dataset, they can also be used to compare different modalities or identify important features. Be advised that current methods for calculating feature importance still have limitations. For example, the most advanced or complex models don't inherently rank feature importance; and generalized feature importance scores may not describe the true relationship that is utilized in the models [41, 42]. Additionally, many studies (despite reporting feature importance scores) aim at demonstrating machine learning models with high classification accuracy rather than understanding the most representative biological information.

Existing reviews have summarized the effectiveness of different machine learning models in differentiating subjects with ADHD from control subjects or subjects with other disorders [32–35]. However, in addition to building classification models to aid diagnosis, machine learning is advantageous in studying mechanisms underlying ADHD due to its ability to describe the vast heterogeneity in the etiology of ADHD. The purpose of the present narrative review is to summarize the current literature regarding the applications and benefits of using machine learning algorithms to understand the underlying neural mechanisms of ADHD, as well as ongoing issues and future research directions. Although the aforementioned limitations of the interpretation of feature importance exist, exploring the possible applications may lead to the development of feature-focused and explainable machine learning models in ADHD. Studies were included if they met the criteria of (1) using machine learning algorithms, (2) having a total sample size of at least 40, (3) applying a cross-validation step, (4) model evaluation is independent of model training, (5) reporting comparisons between features (e.g., feature importance, or performances of different measuring modalities). A full list of search terms used in this review can be found in the

Supplementary document. An overview of the studies that applied machine learning algorithms in investigating ADHD is presented in Supplementary Fig. 1. The detailed methodology and key findings of the included studies can be found in Supplementary Tables 1–4.

MACHINE LEARNING IN CHARACTERIZING ADHD
The use of machine learning in aiding the diagnosis of ADHD has been covered extensively in existing reviews [32, 34, 35] and will, therefore, not be discussed in detail in this review. Briefly, some evidence suggests that machine learning algorithms have the potential to benefit the diagnosis of ADHD by either simplifying the diagnostic process in complex cases (e.g., achieving similar accuracy with less items, increasing accuracy in patients with comorbidities) [43–49] or increasing accuracy with additional neurobehavioral measures or activity records [50–55]. The contribution of the classification models can be limited by factors such as the sample sizes used, which often contribute toward inflated accuracies. Instead, by inspecting the features identified as most important in classification models, machine learning algorithms were able to identify the core characteristic of ADHD.

A recent nationwide study in Sweden applied multiple machine learning models, including random forest, elastic net, deep neural network, and gradient boosting, in identifying the significant predictors for ADHD based on family and medical histories from 238,696 individuals [56]. The best model achieved a sensitivity of 71.7% and a specificity of 65.0%, and the results showed that the top risk factors for ADHD in children are having parents with criminal convictions, male sex, having a relative with ADHD, academic difficulties, and learning disabilities. Another study investigated Conner's rating scale from both parents and teachers in differentiating children with ADHD and controls using a deep neural network [57]. The models demonstrated an accuracy of 89%. More interestingly, the study reported that teachers' ratings for the oppositional questions were more discriminative for ADHD than parents' ratings. In addition, questions directly describing the symptoms were more discriminative than the question worded metaphorically. Among adults with ADHD, one study with 1249 subjects reported that difficulty organizing, does not follow through, making careless mistakes, and difficulty engaging in leisure activities were key characteristics of adult ADHD [58]. This evidence from existing machine learning studies may expand the understanding of the characteristics of ADHD and provide guidance for developing more reliable and efficient diagnostic criteria.

Beyond allowing the classification of subjects into traditional diagnostic groups, research into machine learning-aided diagnosis of ADHD has contributed to the understanding of the clinical presentation and heterogeneity of ADHD by allowing the identification of novel subgroupings of participants, which can increase diagnostic accuracies [59, 60]. For example, Fair et al. evaluated the performance data during seven neuropsychological tasks, including inhibition, working memory, arousal, response variability, temporal information processing, memory span, and processing speed, in a cohort of 285 children with ADHD and 213 controls [61]. By implementing community detection methods, four subgroups in both the ADHD group and the control group were identified. Classification using SVM following this subgrouping led to a diagnostic accuracy as high as 84.1%, compared to a markedly lower classification accuracy of 65% without subgrouping. Similarly, Kleinman et al. regrouped healthy children and children with ADHD, bipolar disorder, or both into two groups based on continuous performance task (CPT) performance [62]. LDA was then used to build separate classification models on both the Diagnostic and Statistical Manual of Mental Disorder (DSM) IV-based groups and CPT-defined groups. CPT-defined groups had a markedly higher discriminative accuracy (95.2%) than the DSM IV-

defined groups (23.8%). A more recent study performed clustering analysis in a combined group of children with ADHD, children with ASD, and controls based on the behavioral measure from 12 domains [63]. Three executive function-defined groups were detected, including weakness in flexibility and emotion regulation, weakness in inhibitory control, and weakness in working memory, organization, and planning. SVM was used to validate the detected subtypes in an independent dataset and yield a classification accuracy of 88.9%. Within a subset of the subjects, the detected subgroups explained more between subject variance than the DSM-defined clinical groups. Such studies suggest that although existing clinical classifications may be sufficient to identify ADHD, they cannot comprehensively capture the heterogeneities.

In general, the accuracy of the classification models varies from 66 to 96% in the existing machine learning studies that investigated behavioral and cognitive performances in ADHD. The inconsistency was partially contributed by the differences in total sample size, percentage of the clinical group in the total sample, test or measurement selection, model selection, or validation methods. Therefore, extra precaution was required in designing reliable classification models. Furthermore, machine learning techniques that can explore the heterogeneities in ADHD (e.g., clustering analysis, regression analysis) may not only improve diagnosis but may contribute to improvements in future research investigating the underlying mechanisms by providing more appropriately defined samples.

MACHINE LEARNING IN INVESTIGATING BIOLOGICAL MECHANISMS OF ADHD
Neuroimaging studies
Structural MRI and diffusion tensor imaging. The neuroanatomy of ADHD has been investigated for decades. However, results are inconsistent [19, 64, 65]. A recent mega-analysis reported subtle alterations in surface area in various cortical regions in ADHD [20]. Studies using diffusion tensor imaging (DTI), a neuroimaging technique that measures microstructural changes, also reported white matter alterations in widespread regions [66]. This existing evidence suggests that ADHD might not be related to highly localized anatomical alterations but more diffuse changes [67–69]. Existing research may be limited by the use of conventional statistical methods, which lack sensitivity to subtle changes over multiple regions and the interactions between them.

Machine learning, on the other hand, can model a number of features simultaneously, making machine learning approaches particularly well-suited to understanding the widespread structural alterations underpinning ADHD. For example, Peng et al. reported results from an extreme learning machine-based classification model which differentiated children with ADHD and controls with an accuracy of 90.18% using sMRI data from ADHD-200 [70]. The model identified surface area, folding index, and volume in the parietal lobe, temporal lobe, and insula as the most important predictors of ADHD. Another study using SVM for classification showed that the white matter volume in the brain stem was the most important feature in differentiating boys with ADHD and controls [71]. Using LASSO regression, a recent DTI study reported that the tract strength between the substantia nigra/ventral tegmental area and the striatum was able to predict impulsivity with a Spearman's correlation of 0.17 in a group of 74 ADHD patients and controls [72]. In a large cohort (4183 subjects from 35 study sites), deep learning neural network revealed that sMRI was a good predictor of ADHD in children but not in adults, supporting the idea that structural alterations associated with ADHD lessen with age [73]. Studies using sMRI can also identify structural properties that distinguish ADHD from other common disorders. For example, Lim et al. investigated the discriminative power of structural properties in ADHD, ASD, and control
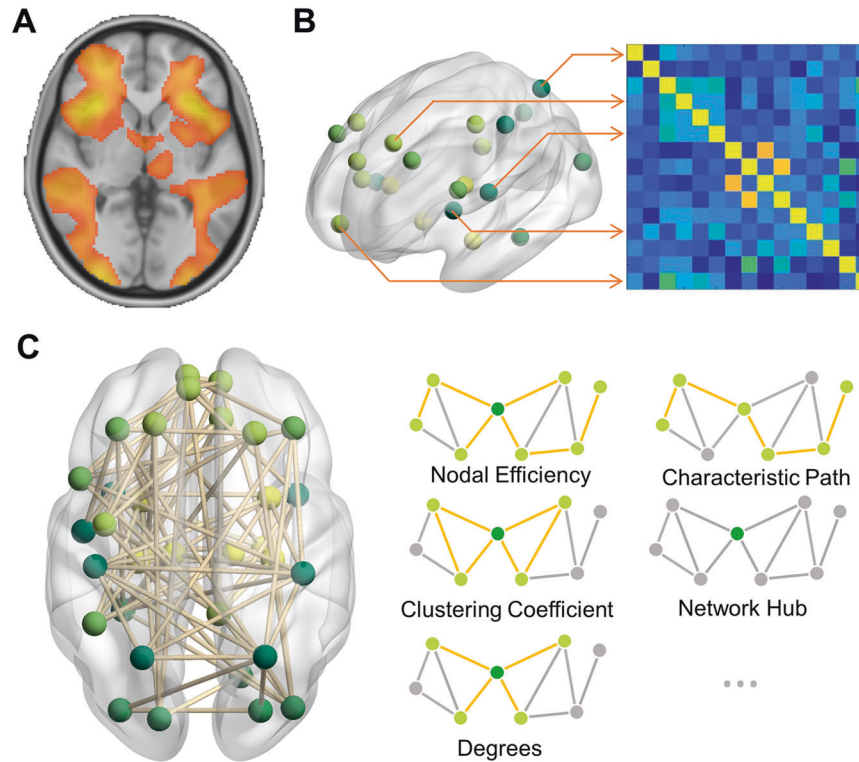
Fig. 2 **Functional neuroimaging features. A** Voxel-level features, for example, voxel-level activation in task-based fMRI or regional homogeneity in resting-state fMRI. **B** Functional connectivity. **C** Network topological properties in graph theoretical analysis. fMRI functional magnetic resonance imaging.

participants [74]. With voxel-level gray matter volume as features, a Gaussian process classification algorithm differentiated ADHD specifically (compared to ASD) from controls with 79.3% accuracy and highlighted several regions in which structural properties contributed highly to this classification. Those regions may be involved specifically in the pathophysiology of ADHD, as opposed to ASD. Despite these promising results, Oztekin et al. found that parent and teacher ratings of executive function in an SVM model resulted in an accuracy of 92.6%, while using sMRI data alone resulted in an accuracy of 61.2%, and adding anatomical features to a model containing neurocognitive measures had minimal benefit [75]. Therefore, in some cases, the additional benefit of sMRI measures for classification may be limited, although they can still contribute toward identifying underlying structural differences.

Machine learning can also be used to explore novel sMRI features, which may provide optimal discriminative power for ongoing research into ADHD. For example, Chang et al. generated novel morphological features based on the local binary patterns (an image texture categorization method) to differentiate data from 210 ADHD and 226 controls from the ADHD-200 dataset [76]. An SVM model applied to the generated features achieved an accuracy of 69.95% in detecting ADHD. Similarly, using volumetric features named Dissociated Dipoles, Igual et al. built an SVM-based classification model with an accuracy of 72.48%, a specificity of 85.93%, and a sensitivity of 60.07% [77]. Another team used a hybrid machine learning approach on novel interregional morphological connectivity features and reported a classification accuracy of 74.65% [78]. Although currently, these studies do not contribute to our understanding of anatomical alterations in ADHD per se, they contribute to the field by highlighting features that may be beneficial for improved diagnosis or sample classification.

Task-based fMRI. Task-based fMRI is a commonly used method to examine brain activation or functional connectivity during of engagement of a specific cognitive domain. Features like voxel-

level activation, functional connectivity between regions-of-interest (ROIs), or network topological properties (as shown in Fig. 2) can be used to build machine learning models. Several studies have applied machine learning techniques to fMRI data collected from participants with ADHD. For example, by applying various machine learning algorithms to the functional activations during time discrimination tasks [79], Flanker tasks [80], and stop-signal task [81], studies have highlighted that the task-related activations in frontal regions were important for the classification of ADHD, suggesting functional importance of frontal regions in ADHD.

Importantly, machine learning algorithms may be able to detect functional patterns (e.g., the collective contribution of multiple brain regions in differentiating ADHD and controls), which may otherwise be undetected when using traditional methods. For example, Wolfers et al. applied a Gaussian process classifier in differentiating subjects with ADHD, their unaffected siblings, and controls based on the fMRI data during stop-signal task [81]. The model was able to differentiate ADHD patients from their siblings with an AUC of 0.65 and from control participants with an AUC of 0.64. The results showed that the fronto-lateral and inferior parietal regions were highly discriminative features for ADHD. Hart et al. utilized a Gaussian process classifier to differentiate boys with ADHD from controls based on fMRI data recorded during a stop-signal task (used to measure response inhibition) [82]. Using voxel-level functional activation as feature, the classification accuracy reached 77%. Interestingly, voxels that showed no significant group differences using traditional univariate analysis demonstrated high discriminative power when using machine learning, suggesting that machine learning methods can tease out important discriminatory activations above and beyond traditional analysis methods.

Resting-state fMRI. The brain demonstrates intrinsic spontaneous activity that can be measured during rest. rs-fMRI measures such

activity, and the collected data can be used to generate machine learning features, such as regional homogeneity (ReHo), fractional amplitude of low-frequency fluctuation (fALFF), and network connectivity. As rs-fMRI data does not require the performance of a task, it is easy to implement in children with ADHD. Classification techniques have highlighted regions in which resting brain activity is of potential importance in ADHD. For example, studies using SVM have revealed that functional connectivity in default mode network, frontoparietal regions, cerebellum, precuneus/posterior cingulate cortex regions, and dorsal anterior cingulate cortex were important in differentiating ADHD [83, 84].

As previously mentioned, the ADHD-200 dataset has allowed numerous investigations into rs-fMRI correlates of ADHD using machine learning algorithms. Various rs-fMRI features have been explored, including ReHo, fALFF, power spectra, functional connectivity, and voxel- and ROI-level functional networks [85–87]. Eloyan et al. constructed a classification algorithm based on majority voting from four algorithms, including random forest on motor cortex connectivity, SVM on major clusters, gradient boosting method on decomposed functional connectivity, and gradient boosting on functional connectivity and motion parameters [88]. The final model achieved a specificity of 94% and a sensitivity of 21%, and connectivity within the motor network was most important in classifying ADHD participants. Several studies have utilized SVM to construct classification models and report that the frontal lobe, parietal lobe, and cerebellum are most discriminative between ADHD and controls and between ADHD inattentive subtype and ADHD combined subtype [89, 90]. Similarly, a graph convolutional neural network study identified the frontal, temporal, and occipital regions and the cerebellum as the most discriminative regions for ADHD and controls [91].

Despite the success of rs-fMRI-based machine learning models, it is possible that phenotypic information such as gender, age, and cognitive measures provide more discriminative power than rs-fMRI data [92, 93]. However, the addition of rs-fMRI features may be beneficial nonetheless. For example, Bohland et al. found that the addition of such features increased generalization to novel data [93]. Additionally, studies have suggested that rs-fMRI data are more predictive for inattentive symptoms rather than hyperactive/impulsive symptoms [94] and that classification accuracy increases when using an SVM trained separately for male and female subjects [95], reflecting that certain applications of such models can yield more accurate results. Such considerations may be useful in future rs-fMRI research.

EEG. Due to its high accessibility, low cost, and non-invasive nature, EEG has gained popularity in studying ADHD. Common features generated from EEG data are power in frequency bands at different locations and event-related potentials (ERPs), which are electrical responses that are time-locked to the occurrence of sensory or cognitive processes, as shown in Fig. 3. Several studies using machine learning have shown that features extracted from EEG data can be used to differentiate ADHD patients from controls and from other comorbid conditions with varied accuracy ranging between 69 and 91% [96–99]. Classification of specific diagnostic subtypes of ADHD based on EEG features is also possible, although with a lower classification accuracy of around 72% [100, 101].

Several studies have investigated the predictive power of specific features of EEG data. For example, using deep neural network, one study identified that ERPs within the time range from 100 to 200 ms post-stimulus are important in differentiating children with ADHD and controls during an interval-time task [99]. The model was able to differentiate the ADHD group and controls with an accuracy of 69%. In addition, several factors appear to contribute to the accuracy of EEG-based models. Several studies have assessed the optimum experimental paradigm for classification. For example, Chang et al. reported that the signal during the

transition period between the task and resting condition was more discriminative for ADHD than the signal during the task condition or resting condition [102]. Tenev et al. reported that a model combining multiple task conditions showed a significant increase in classification accuracy when compared with a single condition (82.3% vs 70%) [103]. Studies using the go/no-go task report inconsistent results regarding the most discriminative task conditions. For example, Mueller et al. reported that an ERP-based network for No-go had significantly higher predictive power than that during the Go condition for a visual sustained attention task [104]. However, Biederman et al. reported that an SVM-based model using the signal from the Go condition achieved higher AUC than the signal during the No-go condition (0.92 vs 0.84) [105]. Age may also be an important influence on classification. For example, splitting subjects into different age groups increased classification accuracy when applying SVM on EEG data [106].

Machine learning is also valuable in investigating novel EEG features. For example, Kim et al. used machine learning to validate mismatch negativity (a novel measure that contrasts activity during regular auditory stimuli and occasional novel stimuli) in differentiating adults with ADHD from controls [107]. The SVM-based model showed a classification accuracy of 81% and identified the frontal lobe, temporal lobe, and limbic lobe as the most important regions in the classification. Studies have also constructed machine learning models using other novel features, including various entropy-based features and fractal dimension-based features from chaotic theory [108–110]. As research continues to employ machine learning methods, it is likely that novel features to best classify individuals with ADHD will continue to be determined.

Functional near-infrared spectroscopy. Functional near-infrared spectroscopy (fNIRS) is a non-invasive and portable method to measure the hemodynamic response in the cortex. Relative to fMRI, fNIRS is less susceptible to the movements and is therefore well-suited to study ADHD, and machine learning has the potential to utilize the fNIRS's high temporal resolution while overcoming its low spatial resolution. One study applied SVM on fNIRS data from children with ADHD and controls during a working memory task [111]. The final model achieved an accuracy of 96% and highlighted the dorsal lateral prefrontal cortex, temporal cortex, medial prefrontal cortex, and posterior prefrontal cortex as the most discriminative in classifying ADHD and controls. Yasumura et al. applied an SVM-based model on fNIRS data from children with ADHD and controls collected during a reverse Stroop task [112]. The model achieved 86.25% accuracy with a sensitivity of 88.71% and a specificity of 83.78%. Splitting the sample into three age groups (<10 years, 10–12 years, >12 years) increased classification accuracy significantly.

Multimodal imaging. Given its ability to model several features simultaneously, machine learning is well-suited to multimodal investigations of neural markers of ADHD. For example, Zhou et al. combined rs-fMRI with sMRI and DTI data from the ABCD dataset and reported that the functional connectivity in frontal and temporal regions, cerebellum, thalamus, and anatomical regions in the basal ganglia were the most discriminative features for ADHD in children [113]. Luo et al. utilized multimodal imaging data, including fMRI data during a cued attention task, sMRI, and diffusion tensor imaging [114]. The algorithms combined a range of machine learning models and achieved an accuracy of 89% in differentiating adults with ADHD and controls and an accuracy of 90% in differentiating ADHD persisters and remitters. The results showed that functional connectivity in the frontal and parietal lobe and amygdala volume was important to differentiate ADHD with controls, while functional connectivity in the frontal lobe, parietal lobe, and putamen was important to differentiate ADHD persisters and remitters. Owens et al. combined task-based fMRI
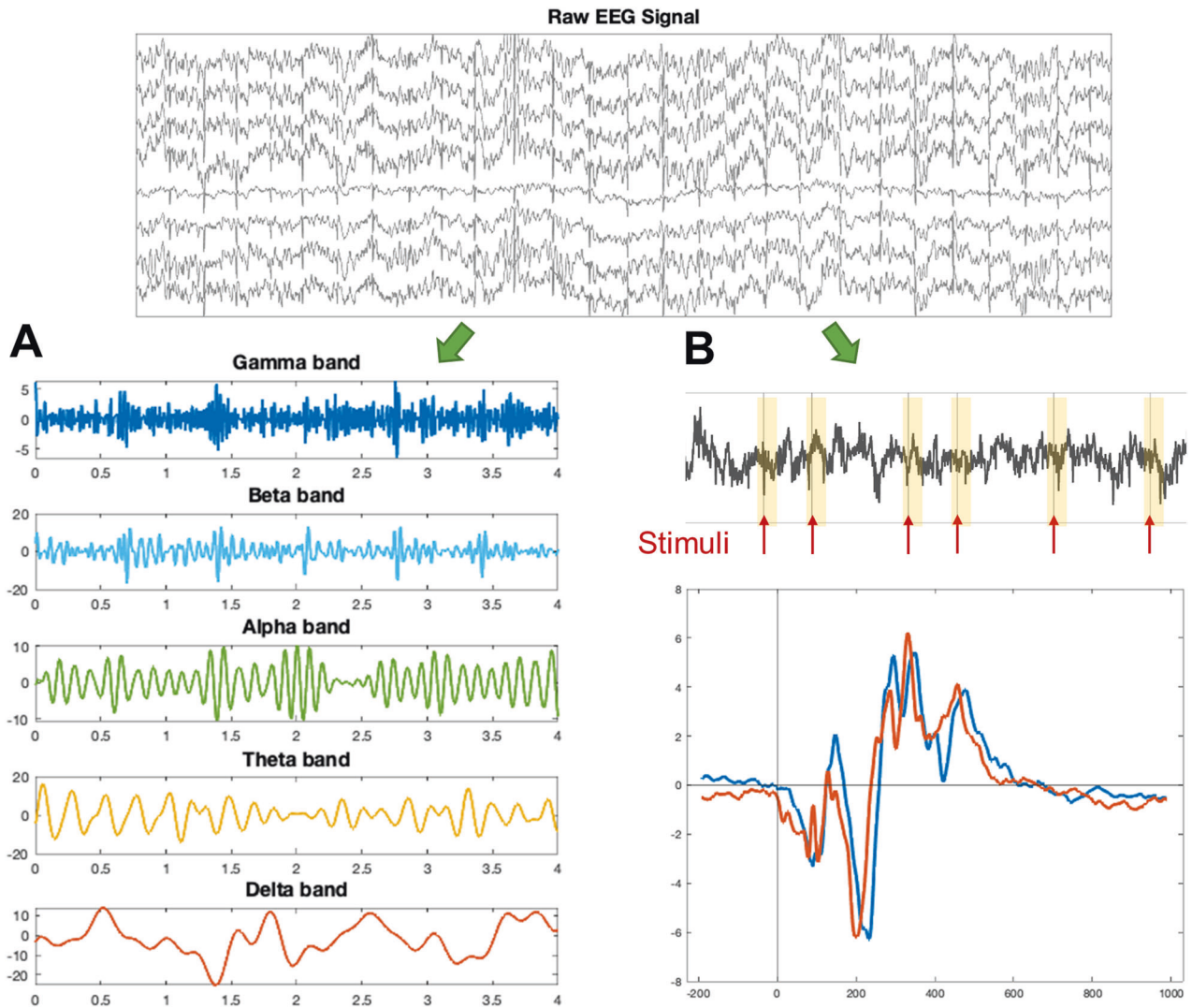
Fig. 3 Electroencephalogram features. A Frequency analysis, including power at different frequency bands. **B** Event-related potential analysis, including the time or amplitude of peaks after the target stimuli.

data and structural MRI data from the ABCD dataset to investigate the relationship between ADHD symptoms and imaging measures [115]. Using the elastic net algorithm, results showed that, compared to other modalities, functional activation during a working memory task can predict ADHD symptoms with the best performance, which explained 2% of the variance with a small effect size. Combining multimodal data offers the opportunity to identify a range of biomarkers, which is particularly advantageous in ADHD, given its complex etiology.

Genetic studies
Genetic studies. Genetic and twin studies suggest that ADHD is highly heritable [116–119]. This heritability may be due to polygenic risk [120, 121]. Recent genome-wide association studies provide promising results in understanding the genetic associations with ADHD [25]. Machine learning handles multiple independent variables simultaneously, allowing the interactions between various risk factors to be assessed. In addition, it highlights risk factors that are statistically insignificant but may contribute to ADHD. These properties make machine learning a particularly valuable tool in studying genetic markers of ADHD.

van der Meer et al. used a random forest regression model to investigate the predictive power of 29 stress-related genes on ADHD severity in children with ADHD, subthreshold ADHD, and controls [122]. The model explained 12.5% of the variance in ADHD severity and indicated that, besides chronic stressors, the region that regulates the expression of telomerase reverse transcriptase was important in predicting ADHD severity. Other studies have used random forest and convolutional neural networks to study genetic predictors of ADHD and have revealed that the gene regions GRM1, GRM8, and EPHA5 are important predictors of ADHD [123, 124]. Using multiple machine learning algorithms, a recent study reported that age and sex were significant predictors in genetic information-based classification models [125]. In addition, gene regions SNAP25, ADGRL3, and DRD4 significantly contributed to the prediction of inattentive, hyperactive, or impulsive symptoms. SVM models have also shown that microRNA has high discriminative power for ADHD and can predict medication responses in ADHD patients [126].

Multi-omics studies. Machine learning algorithms allow the combination of genetic data with data such as cognitive and neuroimaging data. For example, using conditional random forests, Sudre et al. were able to predict ADHD severity with an AUC of 0.79 [127]. While cognitive measures were most important in the overall classification, genomics was important in detecting children with worsening ADHD, highlighting the utility of multimodal machine learning approaches. Yoo et al. combined

anatomical features from both sMRI and DTI, functional connectivity during rs-fMRI, and genetic data related to norepinephrine, dopamine, and glutamate to build a random forest-based classification model and regression model for ADHD [128]. The classification model using cortical thickness and volumes achieved the best performance with an accuracy of 85.1% and an AUC of 0.877 in differentiating ADHD participants and controls. Additionally, the regression model was able to explain 18% of the variance of the ADHD rating scale. Both models did not gain improvements when including genetic data. Future machine learning studies may be needed to further investigate the relations between genetic data, neurocognitive performance, behavioral problems, and neurobiological alterations in ADHD patients.

MACHINE LEARNING IN PREDICTING TREATMENT AND PROGNOSTIC OUTCOMES OF ADHD

Heterogeneity in ADHD imposes difficulties in developing effective and reliable treatment strategies. Methylphenidate (MPH) is one of the main pharmacological treatments for ADHD; however, 30% of patients are poor responders [26, 27]. Machine learning techniques are beneficial to predicting treatment outcomes as they have the ability to provide predictions from relatively little prior knowledge. Several studies have predicted response to MPH using SVM, with features including neuropsychological test performance and information on clinical information [129] and sMRI data [130]. Faraone et al. implemented lasso regression to predict the responses of adolescents to a novel non-stimulant medication (SPN-812) [131]. Responder status (with good responder defined as a >50% improvement in symptoms score) after 6 weeks was predicted with the response data (symptom score change from baseline) collected up to weeks 1, 2, and 3. The lasso regression model predicted the long-term result based on the outcome at 2 weeks with 75% accuracy.

Machine learning can also be utilized to predict adverse drug outcomes, which are common in ADHD treatment. For example, Yoo et al. predicted sleep side effects of using MPH treatment based on multiple variables and achieved an accuracy of 95.5% [132]. Based on findings using long short-term memory model, Fouladvand et al. reported that the initiation of ADHD medication during adolescence is a significant predictor for developing substance use disorder in a large cohort with 11,624 children with ADHD [133]. Zhang-James et al. also reported ADHD medication as one of the important predictors of substance use disorder, along with ADHD diagnosis before 12 years old and crime behaviors [134]. Given that finding the most suitable ADHD treatment is largely still dependent on trial-and-error of medications and the risk for adverse outcomes of drug treatment, the ability to predict treatment outcomes using machine learning models has the potential to reduce financial and medical burdens.

DISCUSSION

A growing number of studies are utilizing machine learning techniques to report interpretable results regarding neural mechanisms associated with ADHD, in addition to building accurate classification models. Such studies have already contributed to the literature regarding functional, structural, and physiological correlates of ADHD.

Performance of machine learning models

A particular benefit of classification models is the ability to label individuals. In addition to the detection of important features, machine learning can assist the development of individualized treatment plans for ADHD (e.g., [130, 135]). The idea of precision medicine has been introduced and practiced in many other diseases [136, 137]. Machine learning studies can accelerate this process in ADHD. This application will no doubt further benefit

from the increasing accessibility of large datasets (e.g., ABCD dataset, Human brain mapping dataset, and UK Biobank dataset), which can be used to train more reliable classification models. Groups with small cohorts can also benefit from collaborations with other groups like The Enhancing NeuroImaging Genetics through Meta-Analysis (ENIGMA) Consortium [138]. Alternatively, He et al. also proposed meta-matching methods to utilize information generated from large public datasets when working on independent datasets [139]. The construction of a reliable model using such data could dramatically reduce the workloads of clinicians, thereby increasing the capacities of the existing medical system and minimizing the burden on affected families and societies.

The existing classification models for ADHD reported largely inconsistent accuracy, with the majority varying from 60 to 90%. Several factors related to machine learning design may contribute to these discrepancies. First, the choice of machine learning algorithms may affect the performance of the classification model based on different datasets. Algorithms with very few or no trainable parameters were preferable for studies with small sample sizes, whereas studies with large datasets were able to explore the effectiveness of deep learning algorithms [56, 73, 133]. The second factor is the size balance between different groups. Most studies were able to recruit patient groups and control groups of relatively similar sizes. However, for clinical studies or population-based studies, the balance is hard to achieve [56, 58, 115]. This may potentially give overly positive results. For example, the same model may have a much higher AUC in a population sample than in a group-matched clinical sample (AUC: 0.86 vs 0.72) [140]. Lastly, the choice of validation-test strategy may contribute to inconsistencies in accuracy. A large independent testing set is the best choice for testing generalizability but is only affordable for studies with large datasets [115, 133]. Nested cross-validation may be an alternative, in which the inner cross-validation layer is responsible for training algorithm's parameter, and the outer cross-validation layer is solely responsible for performance evaluation [71, 113, 114]. However, more than half of the existing machine learning studies in ADHD have only reported results using only one cross-validation, which can cause overfitting of the features and reduce the generalizability of the results.

Identification of important features

A major benefit of machine learning techniques is that they always involve multivariate data, and some machine learning models like SVM and random forest can rank the contribution of features under the interaction of each other [141]. Therefore, the important clinical or biological features in identifying ADHD can be evaluated based on their contribution to the model. Accuracy (or AUC) can also be used to compare the effectiveness of different feature sets. The training process of a machine learning model extensively learns the information associated with the classification labels within the dataset. When training the same model with features from different modalities, accuracies can partially reflect the sensitivity of particular modalities in ADHD. This evidence can be used to guide experimental design in future hypothesis-driven research.

However, several factors limit the interpretability of the important features reported in existing machine learning studies. First, not all of the machine learning models have intrinsic operations to rank the input features for their importance during the learning procedures. Although generalized feature importance methods exist, such as permutation importance, these methods do not necessarily represent the covariate information used in the original models [41, 142]. On the other hand, for machine learning models that include feature ranking mechanisms, the reported results can be restricted by the ranking methods of the models. For example, feature ranking in a linear SVM only recognizes the high contribution features that show linear relationships with
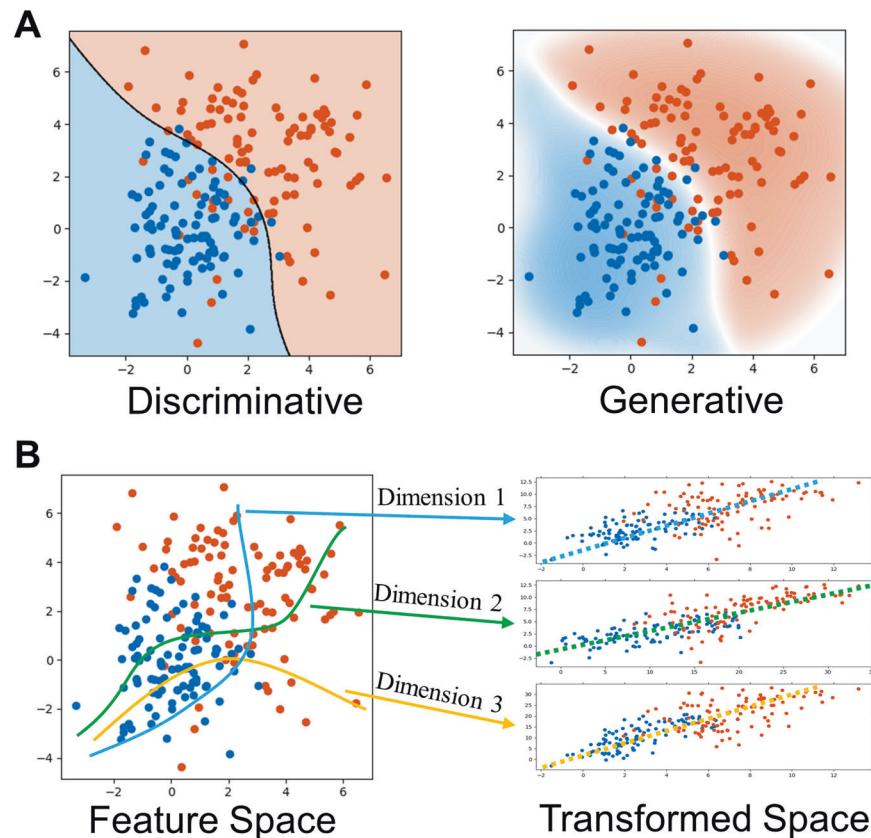
Fig. 4 Generative approaches and dimensional approaches for machine learning studies. A Discriminative model vs generative model for classification between two groups. Discriminative models aim to define the clear boundary between different groups, while generative models aim to characterize each group and classify based on probability or likelihood. **B** Dimensional approaches. Regression models are used to study the relationship between the target clinical measures and data in transformed dimensions.

ADHD. Additionally, the majority of the existing studies have only focused on reporting learning procedures that achieved high classification accuracies without giving enough consideration to the "biological meaningfulness" of the study features. Lastly, the field still lacks gold standards in evaluating the quality of a machine learning study. For example, the studies in this review have reported model performance based on different evaluation methods, like cross-validation, nested cross-validation, or independent testing set, using various metrics, including accuracy, AUC, specificity, and sensitivity, meaning the performance may not be comparable. Authors may choose the favorable metrics that do not represent the true performance, and interpretation of the feature importance of such overfitted or biased studies requires extra precaution.

Current challenges
Despite great promise, challenges are also present before machine learning can provide significant clinical benefits for ADHD due to its heterogeneity. First, machine learning algorithms currently lack interpretability. High-accuracy models are usually constructed with a collection of variables [91, 114, 122], with each variable contributing partial information in distinguishing subjects. The relationship between variables is hard to characterize. Currently, one can rely on the feature's importance score to provide future direction in investigating particular measures. Models that can translate complex interactions between objective measures are truly beneficial in understanding the neural mechanisms associated with ADHD. A second challenge is the limited generalizability of classification models trained on small samples. Although most studies reported here implemented cross-validation methods to combat overfitting and generalization

problems, the nature of the imbalances in the number of features vs the number of subjects in clinical studies and the high heterogeneity of study samples still impose limitations on generalizability [143]. Notably, classification accuracy can drop significantly when applying a trained model to new subjects [45, 46], highlighting the critical need to overcome the generalization problems when implementing machine learning.

Future directions
Machine learning techniques are still currently undergoing extensive development. Several directions have the potential to resolve the existing problems. One direction is taking a generative approach. Most existing machine learning studies have utilized discriminative models focused on finding the boundaries between known groups within a sample. On the contrary, generative models focus on characterizing groups and predicting group allocation based on probability, as shown in Fig. 4A. In addition, generative models can also characterize samples by identifying subgroups that cluster together. Considering that ADHD diagnosis is usually based on subjective measures and that comorbidities are frequently observed, a hard boundary in the classification process may not be an appropriate threshold for ADHD. Within the literature reviewed here, discriminative models were more effective in constructing accurate classification models (e.g., [75, 111, 114].). This is likely due to the more homogeneous sample using carefully selected inclusion and exclusion criteria [39]. Compared to discriminative models, generative models are less vulnerable to the bias induced in the dataset and therefore can generalize well. In addition, due to the heterogeneous nature of mental disorders, there might be multiple etiological sources or various clinical profiles. Generative unsupervised learning models

can detect the homogeneous subtypes otherwise hidden to traditional statistical methods [144]. This property opens the opportunities to capture the heterogeneities embedded in ADHD. More importantly, more homogeneous subgroups expand the interpretability of important features.

Another direction is taking a dimensional approach. The categorical definition of ADHD may not be sufficient to describe the ADHD symptoms in a comprehensive way. Fair et al. reported distinct ADHD subgroups based on cognitive performance, suggesting that neurobiological properties of ADHD might need to be characterized using multiple cognitive measures in addition to the DSM-based symptom measures [61]. Using regression-based machine learning algorithms to associate biological features to multiple clinical dimensions simultaneously can link heterogeneities in both clinical presentations and biological properties of ADHD, therefore increasing the interpretability, for example, in Fig. 4B. This dimensional direction is in line with the National Institute of Mental Health Research Diagnostic Criteria (RDoC) project, which introduced a framework to eliminate diagnosis-imposed boundaries [145]. Notably, several studies have defined the clinical groups based on cognitive or behavioral profiles, and all yield more distinctive groupings than traditional DSM clinical groupings [62, 63, 146–148]. As deficits can be explained along several dimensions (for example, attention, cognitive control, or perception constructs in the RDoC matrix), it may therefore be easier to link to the related biological substrates. In addition, this brings opportunities to explore the phenotypes or endophenotypes of ADHD and explain the heterogeneities in current findings.

In summary, early attempts to investigate ADHD using machine learning show promising results. In addition to seeking high classification accuracy, studies using machine learning to study ADHD can identify the importance of features and discriminative power of modalities, which provide clinical and research targets. Future studies focusing on increasing the interpretability and generalizability of models are highly desired.

REFERENCES

1. Wolraich ML, Hagan JF, Jr., Allan C, Chan E, Davison D, Earls M, et al. Clinical practice guideline for the diagnosis, evaluation, and treatment of attention-deficit/hyperactivity disorder in children and adolescents. Pediatrics. 2019;144:e20192528.
2. Polanczyk GV, Salum GA, Sugaya LS, Caye A, Rohde LA. Annual research review: a meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. J Child Psychol Psychiatry. 2015;56:345–65.
3. Sibley MH, Swanson JM, Arnold LE, Hechtman LT, Owens EB, Stehli A, et al. Defining ADHD symptom persistence in adulthood: optimizing sensitivity and specificity. J Child Psychol Psychiatry. 2017;58:655–62.
4. Faraone SV, Biederman J, Mick E. The age-dependent decline of attention deficit hyperactivity disorder: a meta-analysis of follow-up studies. Psychol Med. 2006;36:159–65.
5. Chang Z, Lichtenstein P, D'Onofrio BM, Sjolander A, Larsson H. Serious transport accidents in adults with attention-deficit/hyperactivity disorder and the effect of medication: a population-based study. JAMA Psychiatry. 2014;71:319–25.
6. Dalsgaard S, Leckman JF, Mortensen PB, Nielsen HS, Simonsen M. Effect of drugs on the risk of injuries in children with attention deficit hyperactivity disorder: a prospective cohort study. Lancet Psychiatry. 2015;2:702–9.
7. Dalsgaard S, Mortensen PB, Frydenberg M, Thomsen PH. ADHD, stimulant treatment in childhood and subsequent substance abuse in adulthood—a naturalistic long-term follow-up study. Addict Behav. 2014;39:325–8.
8. Skirrow C, Asherson P. Emotional lability, comorbidity and impairment in adults with attention-deficit hyperactivity disorder. J Affect Disord. 2013;147:80–6.
9. Jacob CP, Romanos J, Dempfle A, Heine M, Windemuth-Kieselbach C, Kruse A, et al. Co-morbidity of adult attention-deficit/hyperactivity disorder with focus on personality traits and related disorders in a tertiary referral center. Eur Arch Psychiatry Clin Neurosci. 2007;257:309–17.
10. Luo Y, Weibman D, Halperin JM, Li X. A review of heterogeneity in attention deficit/hyperactivity disorder (ADHD). Front Hum Neurosci. 2019;13:42.
11. Arnett AB, Pennington BF, Willcutt EG, DeFries JC, Olson RK. Sex differences in ADHD symptom severity. J Child Psychol Psychiatry. 2015;56:632–9.
12. Thapar A. Discoveries on the genetics of ADHD in the 21st century: new findings and their implications. Am J Psychiatry. 2018;175:943–50.
13. Kim JH, Kim JY, Lee J, Jeong GH, Lee E, Lee S, et al. Environmental risk factors, protective factors, and peripheral biomarkers for ADHD: an umbrella review. Lancet Psychiatry. 2020;7:955–70.
14. Franke B, Michelini G, Asherson P, Banaschewski T, Bilbow A, Buitelaar JK, et al. Live fast, die young? A review on the developmental trajectories of ADHD across the lifespan. Eur Neuropsychopharmacol. 2018;28:1059–88.
15. Reale L, Bartoli B, Cartabia M, Zanetti M, Costantino MA, Canevini MP, et al. Comorbidity prevalence and treatment outcome in children and adolescents with ADHD. Eur Child Adolesc Psychiatry. 2017;26:1443–57.
16. Pievsky MA, McGrath RE. The neurocognitive profile of attention-deficit/hyperactivity disorder: a review of meta-analyses. Arch Clin Neuropsychol. 2018;33:143–57.
17. Willcutt EG, Doyle AE, Nigg JT, Faraone SV, Pennington BF. Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. Biol Psychiatry. 2005;57:1336–46.
18. Schoechlin C, Engel RR. Neuropsychological performance in adult attention-deficit hyperactivity disorder: meta-analysis of empirical data. Arch Clin Neuropsychol. 2005;20:727–44.
19. Norman LJ, Carlisi C, Lukito S, Hart H, Mataix-Cols D, Radua J, et al. Structural and functional brain abnormalities in attention-deficit/hyperactivity disorder and obsessive-compulsive disorder: a comparative meta-analysis. JAMA Psychiatry. 2016;73:815–25.
20. Hoogman M, Muetzel R, Guimaraes JP, Shumskaya E, Mennes M, Zwiers MP, et al. Brain imaging of the cortex in ADHD: a coordinated analysis of large-scale clinical and population-based samples. Am J Psychiatry. 2019;176:531–42.
21. Lukito S, Norman L, Carlisi C, Radua J, Hart H, Simonoff E, et al. Comparative meta-analyses of brain structural and functional abnormalities during cognitive control in attention-deficit/hyperactivity disorder and autism spectrum disorder. Psychol Med. 2020;50:894–919.
22. Hart H, Radua J, Nakao T, Mataix-Cols D, Rubia K. Meta-analysis of functional magnetic resonance imaging studies of inhibition and attention in attention-deficit/hyperactivity disorder: exploring task-specific, stimulant medication, and age effects. JAMA Psychiatry. 2013;70:185–98.
23. Demontis D, Walters RK, Martin J, Mattheisen M, Als TD, Agerbo E, et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. Nat Genet. 2019;51:63–75.
24. Sanchez-Mora C, Ramos-Quiroga JA, Bosch R, Corrales M, Garcia-Martinez I, Nogueira M, et al. Case-control genome-wide association study of persistent attention-deficit hyperactivity disorder identifies FBXO33 as a novel susceptibility gene for the disorder. Neuropsychopharmacology. 2015;40:915–26.
25. Demontis D, Walters GB, Athanasiadis G, Walters R, Therrien K, Farajzadeh L, et al. Genome-wide analyses of ADHD identify 27 risk loci, refine the genetic architecture and implicate several cognitive domains. Nat Genet. 2023;55:198–208.
26. Santosh PJ, Taylor E. Stimulant drugs. Eur Child Adolesc Psychiatry. 2000;9:I27–43.
27. Hodgkins P, Shaw M, Coghill D, Hechtman L. Amfetamine and methylphenidate medications for attention-deficit/hyperactivity disorder: complementary treatment options. Eur Child Adolesc Psychiatry. 2012;21:477–92.
28. Arbabshirani MR, Plis S, Sui J, Calhoun VD. Single subject prediction of brain disorders in neuroimaging: promises and pitfalls. Neuroimage. 2017;145:137–65.
29. Quaak M, van de Mortel L, Thomas RM, van Wingen G. Deep learning applications for the classification of psychiatric disorders using neuroimaging data: systematic review and meta-analysis. Neuroimage Clin. 2021;30:102584.
30. Scheinost D, Noble S, Horien C, Greene AS, Lake EM, Salehi M, et al. Ten simple rules for predictive modeling of individual differences in neuroimaging. Neuroimage. 2019;193:35–45.
31. Simundic AM. Measures of diagnostic accuracy: basic definitions. Electron J IFCC. 2009;19:203–11.
32. Uddin M, Wang Y, Woodbury-Smith M. Artificial intelligence for precision medicine in neurodevelopmental disorders. NPJ Digit Med. 2019;2:112.
33. Eslami T, Almuqhim F, Raiker JS, Saeed F. Machine learning methods for diagnosing autism spectrum disorder and attention-deficit/hyperactivity disorder using functional and structural MRI: a survey. Front Neuroinform. 2020;14:575999.
34. Montaleao Brum Alves R, Ferreira da Silva M, Assis Schmitz E, Juarez Alencar A. Trends, limits, and challenges of computer technologies in attention deficit hyperactivity disorder diagnosis and treatment. Cyberpsychol Behav Soc Netw. 2022;25:14–26.
35. Loh HW, Ooi CP, Barua PD, Palmer EE, Molinari F, Acharya UR. Automated detection of ADHD: current trends and future perspective. Comput Biol Med. 2022;146:105525.

36. The ADHD-200-Consortium. The ADHD-200 Consortium: a model to advance the translational potential of neuroimaging in clinical neuroscience. Front Syst Neurosci. 2012;6:62.

37. Casey BJ, Cannonier T, Conley MI, Cohen AO, Barch DM, Heitzeg MM, et al. The adolescent brain cognitive development (ABCD) study: imaging acquisition across 21 sites. Dev Cogn Neurosci. 2018;32:43–54.

38. Zhang-James Y, Razavi AS, Hoogman M, Franke B, Faraone SV. Machine learning and MRI-based diagnostic models for ADHD: are we there yet? J Atten Disord. 2023;27:335–53.

39. Pulini AA, Kerr WT, Loo SK, Lenartowicz A. Classification accuracy of neuroimaging biomarkers in attention-deficit/hyperactivity disorder: effects of sample size and circular analysis. Biol Psychiatry Cogn Neurosci Neuroimaging. 2019;4:108–20.

40. Schnack HG, Kahn RS. Detecting neuroimaging biomarkers for psychiatric disorders: sample size matters. Front Psychiatry. 2016;7:50.

41. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell. 2019;1:206–15.

42. Goodwin NL, Nilsson SRO, Choong JJ, Golden SA. Toward the explainability, transparency, and universality of machine learning for behavioral classification in neuroscience. Curr Opin Neurobiol. 2022;73:102544.

43. Kim S, Lee HK, Lee K. Can the MMPI predict adult ADHD? An approach using machine learning methods. Diagnostics. 2021;11:976.

44. Tachmazidis I, Chen T, Adamou M, Antoniou G. A hybrid AI approach for supporting clinical diagnosis of attention deficit hyperactivity disorder (ADHD) in adults. Health Inf Sci Syst. 2021;9:1.

45. Duda M, Ma R, Haber N, Wall DP. Use of machine learning for behavioral distinction of autism and ADHD. Transl Psychiatry. 2016;6:e732.

46. Duda M, Haber N, Daniels J, Wall DP. Crowdsourced validation of a machine-learning classification system for autism and ADHD. Transl Psychiatry. 2017;7:e1133.

47. Delavarian M, Towhidkhah F, Dibajnia P, Gharibzadeh S. Designing a decision support system for distinguishing ADHD from similar children behavioral disorders. J Med Syst. 2012;36:1335–43.

48. Christiansen H, Chavanon ML, Hirsch O, Schmidt MH, Meyer C, Muller A, et al. Use of machine learning to classify adult ADHD and other conditions based on the Conners' Adult ADHD Rating Scales. Sci Rep. 2020;10:18871.

49. Liu YS, Cao B, Chokka PR. Screening for adulthood ADHD and comorbidities in a tertiary mental health center using EarlyDetect: a machine learning-based pilot study. J Atten Disord. 2023;27:324–31.

50. Finch HW, Davis A, Dean RS. Identification of individuals with ADHD using the Dean-Woodcock sensory motor battery and a boosted tree algorithm. Behav Res Methods. 2015;47:204–15.

51. Slobodin O, Yahav I, Berger I. A machine-based prediction model of ADHD using CPT data. Front Hum Neurosci. 2020;14:560021.

52. Uluyagmur-Ozturk M, Arman AR, Yilmaz SS, Findik OTP, Genc HA, Carkaxhiu-Bulut G, et al. ADHD and ASD classification based on emotion recognition data. In: Proceedings of 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA). 2016:810–3.

53. Amado-Caballero P, Casaseca-de-la-Higuera P, Alberola-Lopez S, Andres-de-Llano JM, Villalobos JAL, Garmendia-Leiza JR, et al. Objective ADHD diagnosis using convolutional neural networks over daily-life activity records. IEEE J Biomed Health Inf. 2020;24:2690–700.

54. Heller MD, Roots K, Srivastava S, Schumann J, Srivastava J, Hale TS. A machine learning-based analysis of game data for attention deficit hyperactivity disorder assessment. Games Health J. 2013;2:291–8.

55. Das W, Khanna S. A robust machine learning based framework for the automated detection of ADHD using pupillometric biomarkers and time series analysis. Sci Rep. 2021;11:16370.

56. Garcia-Argibay M, Zhang-James Y, Cortese S, Lichtenstein P, Larsson H, Faraone SV. Predicting childhood and adolescent attention-deficit/hyperactivity disorder onset: a nationwide deep learning approach. Mol Psychiatry. 2023;28:1232–9.

57. Cheng CY, Tseng WL, Chang CF, Chang CH, Gau SS. A deep learning approach for missing data imputation of rating scales assessing attention-deficit hyperactivity disorder. Front Psychiatry. 2020;11:673.

58. Goh PK, Martel MM, Jones PJ, Bansal PS, Eng AG, Elkins AR, et al. Clarifying relations between ADHD and functional impairment in adulthood: utilization of network and machine learning approaches. Assessment. 2023;30:316–31.

59. Karalunas SL, Nigg JT. Heterogeneity and subtyping in attention-deficit/hyperactivity disorder-considerations for emerging research using person-centered computational approaches. Biol Psychiatry. 2020;88:103–10.

60. Nigg JT, Karalunas SL, Feczko E, Fair DA. Toward a revised nosology for attention-deficit/hyperactivity disorder heterogeneity. Biol Psychiatry Cogn Neurosci Neuroimaging. 2020;5:726–37.

61. Fair DA, Bathula D, Nikolas MA, Nigg JT. Distinct neuropsychological subgroups in typically developing youth inform heterogeneity in children with ADHD. Proc Natl Acad Sci USA. 2012;109:6769–74.

62. Kleinman A, Caetano SC, Brentani H, Rocca CC, dos Santos B, Andrade ER, et al. Attention-based classification pattern, a research domain criteria framework, in youths with bipolar disorder and attention-deficit/hyperactivity disorder. Aust N Z J Psychiatry. 2015;49:255–65.

63. Vaidya CJ, You X, Mostofsky S, Pereira F, Berl MM, Kenworthy L. Data-driven identification of subtypes of executive function across typical development, attention deficit hyperactivity disorder, and autism spectrum disorders. J Child Psychol Psychiatry. 2020;61:51–61.

64. Valera EM, Faraone SV, Murray KE, Seidman LJ. Meta-analysis of structural imaging findings in attention-deficit/hyperactivity disorder. Biol Psychiatry. 2007;61:1361–9.

65. Nakao T, Radua J, Rubia K, Mataix-Cols D. Gray matter volume abnormalities in ADHD: voxel-based meta-analysis exploring the effects of age and stimulant medication. Am J Psychiatry. 2011;168:1154–63.

66. Chen L, Hu X, Ouyang L, He N, Liao Y, Liu Q, et al. A systematic review and meta-analysis of tract-based spatial statistics studies regarding attention-deficit/hyperactivity disorder. Neurosci Biobehav Rev. 2016;68:838–47.

67. Wang B, Wang G, Wang X, Cao R, Xiang J, Yan T, et al. Rich-club analysis in adults with ADHD connectomes reveals an abnormal structural core network. J Atten Disord. 2021;25:1068–79.

68. Griffiths KR, Braund TA, Kohn MR, Clarke S, Williams LM, Korgaonkar MS. Structural brain network topology underpinning ADHD and response to methylphenidate treatment. Transl Psychiatry. 2021;11:150.

69. Beare R, Adamson C, Bellgrove MA, Vilgis V, Vance A, Seal ML, et al. Altered structural connectivity in ADHD: a network based analysis. Brain Imaging Behav. 2017;11:846–58.

70. Peng X, Lin P, Zhang T, Wang J. Extreme learning machine-based classification of ADHD using brain structural MRI data. PLoS ONE. 2013;8:e79476.

71. Johnston BA, Mwangi B, Matthews K, Coghill D, Konrad K, Steele JD. Brainstem abnormalities in attention deficit hyperactivity disorder support high accuracy individual diagnostic classification. Hum Brain Mapp. 2014;35:5179–90.

72. Elliott BL, D'Ardenne K, Mukherjee P, Schweitzer JB, McClure SM. Limbic and executive meso- and nigrostriatal tracts predict impulsivity differences in attention-deficit/hyperactivity disorder. Biol Psychiatry Cogn Neurosci Neuroimaging. 2022;7:415–23.

73. Zhang-James Y, Helminen EC, Liu J, Group E-AW, Franke B, Hoogman M, et al. Evidence for similar structural brain anomalies in youth and adult attention-deficit/hyperactivity disorder: a machine learning analysis. Transl Psychiatry. 2021;11:82.

74. Lim L, Marquand A, Cubillo AA, Smith AB, Chantiluke K, Simmons A, et al. Disorder-specific predictive classification of adolescents with attention deficit hyperactivity disorder (ADHD) relative to autism using structural magnetic resonance imaging. PLoS ONE. 2013;8:e63660.

75. Oztekin I, Finlayson MA, Graziano PA, Dick AS. Is there any incremental benefit to conducting neuroimaging and neurocognitive assessments in the diagnosis of ADHD in young children? A machine learning investigation. Dev Cogn Neurosci. 2021;49:100966.

76. Chang CW, Ho CC, Chen JH. ADHD classification by a texture analysis of anatomical brain MRI data. Front Syst Neurosci. 2012;6:66.

77. Igual L, Soliva JC, Escalera S, Gimeno R, Vilarroya O, Radeva P. Automatic brain caudate nuclei segmentation and classification in diagnostic of attention-deficit/hyperactivity disorder. Comput Med Imaging Graph. 2012;36:591–600.

78. Wang XH, Jiao Y, Li L. Diagnostic model for attention-deficit hyperactivity disorder based on interregional morphological connectivity. Neurosci Lett. 2018;685:30–4.

79. Hart H, Marquand AF, Smith A, Cubillo A, Simmons A, Brammer M, et al. Predictive neurofunctional markers of attention-deficit/hyperactivity disorder based on pattern classification of temporal processing. J Am Acad Child Adolesc Psychiatry. 2014;53:569–78.e1.

80. Iannaccone R, Hauser TU, Ball J, Brandeis D, Walitza S, Brem S. Classifying adolescent attention-deficit/hyperactivity disorder (ADHD) based on functional and structural imaging. Eur Child Adolesc Psychiatry. 2015;24:1279–89.

81. Wolfers T, van Rooij D, Oosterlaan J, Heslenfeld D, Hartman CA, Hoekstra PJ, et al. Quantifying patterns of brain activity: distinguishing unaffected siblings from participants with ADHD and healthy individuals. Neuroimage Clin. 2016;12:227–33.

82. Hart H, Chantiluke K, Cubillo AI, Smith AB, Simmons A, Brammer MJ, et al. Pattern classification of response inhibition in ADHD: toward the development of neurobiological markers for ADHD. Hum Brain Mapp. 2014;35:3083–94.

83. Sato JR, Hoexter MQ, Castellanos XF, Rohde LA. Abnormal brain connectivity patterns in adults with ADHD: a coherence study. PLoS ONE. 2012;7:e45671.

84. Sun Y, Zhao L, Lan Z, Jia XZ, Xue SW. Differentiating boys with ADHD from those with typical development based on whole-brain functional connections using a machine learning approach. Neuropsychiatr Dis Treat. 2020;16:691–702.

85. Colby JB, Rudie JD, Brown JA, Douglas PK, Cohen MS, Shehzad Z. Insights into multimodal imaging classification of ADHD. Front Syst Neurosci. 2012;6:59.

86. Olivetti E, Greiner S, Avesani P. ADHD diagnosis from multiple data sources with batch effects. Front Syst Neurosci. 2012;6:70.

87. Dey S, Rao AR, Shah M. Exploiting the brain's network structure in identifying ADHD subjects. Front Syst Neurosci. 2012;6:75.

88. Eloyan A, Muschelli J, Nebel MB, Liu H, Han F, Zhao T, et al. Automated diagnoses of attention deficit hyperactive disorder using magnetic resonance imaging. Front Syst Neurosci. 2012;6:61.

89. Cheng W, Ji X, Zhang J, Feng J. Individual classification of ADHD patients by integrating multiscale neuroimaging markers and advanced pattern recognition techniques. Front Syst Neurosci. 2012;6:58.

90. Fair DA, Nigg JT, Iyer S, Bathula D, Mills KL, Dosenbach NU, et al. Distinct neural signatures detected for ADHD subtypes after controlling for micro-movements in resting state functional connectivity MRI data. Front Syst Neurosci. 2012;6:80.

91. Zhao K, Duka B, Xie H, Oathes DJ, Calhoun V, Zhang Y. A dynamic graph convolutional neural network framework reveals new insights into connectome dysfunctions in ADHD. Neuroimage. 2022;246:118774.

92. Brown MR, Sidhu GS, Greiner R, Asgarian N, Bastani M, Silverstone PH, et al. ADHD-200 Global Competition: diagnosing ADHD using personal characteristic data can outperform resting state fMRI measurements. Front Syst Neurosci. 2012;6:69.

93. Bohland JW, Saperstein S, Pereira F, Rapin J, Grady L. Network, anatomical, and non-imaging measures for the prediction of ADHD diagnosis in individual subjects. Front Syst Neurosci. 2012;6:78.

94. Wang XH, Jiao Y, Li L. Predicting clinical symptoms of attention deficit hyperactivity disorder based on temporal patterns between and within intrinsic connectivity networks. Neuroscience. 2017;362:60–9.

95. Dey S, Rao AR, Shah M. Attributed graph distance measure for automatic detection of attention deficit hyperactive disordered subjects. Front Neural Circuits. 2014;8:64.

96. Sadatnezhad K, Boostani R, Ghanizadeh A. Classification of BMD and ADHD patients using their EEG signals. Expert Syst Appl. 2011;38:1956–63.

97. Nazhvani AD, Boostani R, Afrasiabi S, Sadatnezhad K. Classification of ADHD and BMD patients using visual evoked potential. Clin Neurol Neurosurg. 2013;115:2329–35.

98. Tor HT, Ooi CP, Lim-Ashworth NS, Wei JKE, Jahmunah V, Oh SL, et al. Automated detection of conduct disorder and attention deficit hyperactivity disorder using decomposition and nonlinear techniques with EEG signals. Comput Methods Prog Biomed. 2021;200:105941.

99. Vahid A, Bluschke A, Roessner V, Stober S, Beste C. Deep learning based on event-related EEG differentiates children with ADHD from healthy controls. J Clin Med. 2019;8:1055.

100. Pedrollo GR, Franco AR, Bagesteiro LB, Balbinot A. Spiking neural networks diagnosis of ADHD subtypes through EEG signals evaluation. Annu Int Conf IEEE Eng Med Biol Soc. 2022;2022:3166–9.

101. Luo N, Luo X, Zheng S, Yao D, Zhao M, Cui Y, et al. Aberrant brain dynamics and spectral power in children with ADHD and its subtypes. Eur Child Adolesc Psychiatry. 2022. https://doi.org/10.1007/s00787-022-02068-6.

102. Chang Y, Stevenson C, Chen IC, Lin DS, Ko LW. Neurological state changes indicative of ADHD in children learned via EEG-based LSTM networks. J Neural Eng. 2022;19:016021.

103. Tenev A, Markovska-Simoska S, Kocarev L, Pop-Jordanov J, Muller A, Candrian G. Machine learning approach for classification of ADHD adults. Int J Psychophysiol. 2014;93:162–6.

104. Mueller A, Candrian G, Grane VA, Kropotov JD, Ponomarev VA, Baschera GM. Discriminating between ADHD adults and controls using independent ERP components and a support vector machine: a validation study. Nonlinear Biomed Phys. 2011;5:5.

105. Biederman J, Hammerness P, Sadeh B, Peremen Z, Amit A, Or-Ly H, et al. Diagnostic utility of brain activity flow patterns analysis in attention deficit hyperactivity disorder. Psychol Med. 2017;47:1259–70.

106. Helgadottir H, Gudmundsson OO, Baldursson G, Magnusson P, Blin N, Brynjolfsdottir B, et al. Electroencephalography as a clinical tool for diagnosing and monitoring attention deficit hyperactivity disorder: a cross-sectional study. BMJ Open. 2015;5:e005500.

107. Kim S, Baek JH, Kwon YJ, Lee HY, Yoo JH, Shim SH, et al. Machine-learning-based diagnosis of drug-naive adult patients with attention-deficit hyperactivity disorder using mismatch negativity. Transl Psychiatry. 2021;11:484.

108. Boroujeni YK, Rastegari AA, Khodadadi H. Diagnosis of attention deficit hyperactivity disorder using non-linear analysis of the EEG signal. IET Syst Biol. 2019;13:260–6.

109. Mohammadi MR, Khaleghi A, Nasrabadi AM, Rafieivand S, Begol M, Zarafshan H. EEG classification of ADHD and normal children using non-linear features and neural network. Biomed Eng Lett. 2016;6:66–73.

110. Koh JEW, Ooi CP, Lim-Ashworth NS, Vicnesh J, Tor HT, Lih OS, et al. Automated classification of attention deficit hyperactivity disorder and conduct disorder using entropy features with ECG signals. Comput Biol Med. 2021;140:105120.

111. Gu Y, Miao S, Han J, Liang Z, Ouyang G, Yang J, et al. Identifying ADHD children using hemodynamic responses during a working memory task measured by functional near-infrared spectroscopy. J Neural Eng. 2018;15:035005.

112. Yasumura A, Omori M, Fukuda A, Takahashi J, Yasumura Y, Nakagawa E, et al. Applied machine learning method to predict children with ADHD using prefrontal cortex activity: a multicenter study in Japan. J Atten Disord. 2020;24:2012–20.

113. Zhou X, Lin Q, Gui Y, Wang Z, Liu M, Lu H. Multimodal MR images-based diagnosis of early adolescent attention-deficit/hyperactivity disorder using Multiple Kernel Learning. Front Neurosci. 2021;15:710133.

114. Luo Y, Alvarez TL, Halperin JM, Li X. Multimodal neuroimaging-based prediction of adult outcomes in childhood-onset ADHD using ensemble learning techniques. Neuroimage Clin. 2020;26:102238.

115. Owens MM, Allgaier N, Hahn S, Yuan D, Albaugh M, Adise S, et al. Multimethod investigation of the neurobiological basis of ADHD symptomatology in children aged 9–10: baseline data from the ABCD study. Transl Psychiatry. 2021;11:64.

116. Faraone SV, Larsson H. Genetics of attention deficit hyperactivity disorder. Mol Psychiatry. 2019;24:562–75.

117. Pingault JB, Viding E, Galera C, Greven CU, Zheng Y, Plomin R, et al. Genetic and environmental influences on the developmental course of attention-deficit/hyperactivity disorder symptoms from childhood to adolescence. JAMA Psychiatry. 2015;72:651–8.

118. Greven CU, Rijsdijk FV, Plomin R. A twin study of ADHD symptoms in early adolescence: hyperactivity-impulsivity and inattentiveness show substantial genetic overlap but also genetic specificity. J Abnorm Child Psychol. 2011;39:265–75.

119. Chang Z, Lichtenstein P, Asherson PJ, Larsson H. Developmental twin study of attention problems: high heritabilities throughout development. JAMA Psychiatry. 2013;70:311–8.

120. Riglin L, Collishaw S, Thapar AK, Dalsgaard S, Langley K, Smith GD, et al. Association of genetic risk variants with attention-deficit/hyperactivity disorder trajectories in the general population. JAMA Psychiatry. 2016;73:1285–92.

121. Hudziak JJ, Derks EM, Althoff RR, Rettew DC, Boomsma DI. The genetic and environmental contributions to attention deficit hyperactivity disorder as measured by the Conners' rating scales—revised. Am J Psychiatry. 2005;162:1614–20.

122. van der Meer D, Hoekstra PJ, van Donkelaar M, Bralten J, Oosterlaan J, Heslenfeld D, et al. Predicting attention-deficit/hyperactivity disorder severity from psychosocial stress and stress-response genes: a random forest regression approach. Transl Psychiatry. 2017;7:e1145.

123. Liu Y, Qu HQ, Chang X, Nguyen K, Qu J, Tian L, et al. Deep learning prediction of attention-deficit hyperactivity disorder in African Americans by copy number variation. Exp Biol Med. 2021;246:2317–23.

124. Liu L, Feng X, Li H, Cheng Li S, Qian Q, Wang Y. Deep learning model reveals potential risk genes for ADHD, especially Ephrin receptor gene EPHA5. Brief Bioinform. 2021;22:bbab207.

125. Cervantes-Henriquez ML, Acosta-Lopez JE, Martinez AF, Arcos-Burgos M, Puentes-Rozo PJ, Velez JI. Machine learning prediction of ADHD severity: association and linkage to ADGRL3, DRD4, and SNAP25. J Atten Disord. 2022;26:587–605.

126. Wang LJ, Kuo HC, Lee SY, Huang LH, Lin Y, Lin PH, et al. MicroRNAs serve as prediction and treatment-response biomarkers of attention-deficit/hyperactivity disorder and promote the differentiation of neuronal cells by repressing the apoptosis pathway. Transl Psychiatry. 2022;12:67.

127. Sudre G, Sharp W, Kundzicz P, Bouyssi-Kobar M, Norman L, Choudhury S, et al. Predicting the course of ADHD symptoms through the integration of childhood genomic, neural, and cognitive features. Mol Psychiatry. 2021;26:4046–54.

128. Yoo JH, Kim JI, Kim BN, Jeong B. Exploring characteristic features of attention-deficit/hyperactivity disorder: findings from multi-modal MRI and candidate genetic data. Brain Imaging Behav. 2020;14:2132–47.

129. Johnston BA, Coghill D, Matthews K, Steele JD. Predicting methylphenidate response in attention deficit hyperactivity disorder: a preliminary study. J Psychopharmacol. 2015;29:24–30.

130. Chang JC, Lin HY, Lv J, Tseng WI, Gau SS. Regional brain volume predicts response to methylphenidate treatment in individuals with ADHD. BMC Psychiatry. 2021;21:26.

131. Faraone SV, Gomeni R, Hull JT, Busse GD, Melyan Z, O'Neal W, et al. Early response to SPN-812 (viloxazine extended-release) can predict efficacy outcome

in pediatric subjects with ADHD: a machine learning post-hoc analysis of four randomized clinical trials. Psychiatry Res. 2021;296:113664.

132. Yoo JH, Sharma V, Kim JW, McMakin DL, Hong SB, Zalesky A, et al. Prediction of sleep side effects following methylphenidate treatment in ADHD youth. Neuroimage Clin. 2020;26:102030.

133. Fouladvand S, Hankosky ER, Bush H, Chen J, Dwoskin LP, Freeman PR, et al. Predicting substance use disorder using long-term attention deficit hyperactivity disorder medication records in Truven. Health Inform J. 2020;26:787–802.

134. Zhang-James Y, Chen Q, Kuja-Halkola R, Lichtenstein P, Larsson H, Faraone SV. Machine-learning prediction of comorbid substance use disorders in ADHD youth using Swedish registry data. J Child Psychol Psychiatry. 2020;61:1370–9.

135. Kim JW, Sharma V, Ryan ND. Predicting methylphenidate response in ADHD using machine learning approaches. Int J Neuropsychopharmacol. 2015;18:pyv052.

136. Peterson BS. Editorial: Biomarkers in precision medicine for mental illnesses. J Child Psychol Psychiatry. 2020;61:1279–81.

137. Lenze EJ, Rodebaugh TL, Nicol GE. A framework for advancing precision medicine in clinical trials for mental disorders. JAMA Psychiatry. 2020;77:663–4.

138. Thompson PM, Stein JL, Medland SE, Hibar DP, Vasquez AA, Renteria ME, et al. The ENIGMA Consortium: large-scale collaborative analyses of neuroimaging and genetic data. Brain Imaging Behav. 2014;8:153–82.

139. He T, An L, Chen P, Chen J, Feng J, Bzdok D, et al. Meta-matching as a simple framework to translate phenotypic predictive models from big to small data. Nat Neurosci. 2022;795–804.

140. Ter-Minassian L, Viani N, Wickersham A, Cross L, Stewart R, Velupillai S, et al. Assessing machine learning for fair prediction of ADHD in school pupils using a retrospective cohort study of linked education and healthcare data. BMJ Open. 2022;12:e058058.

141. Altmann A, Tolosi L, Sander O, Lengauer T. Permutation importance: a corrected feature importance measure. Bioinformatics. 2010;26:1340–7.

142. Fisher A, Rudin C, Dominici F. All models are wrong, but many are useful: learning a variable's importance by studying an entire class of prediction models simultaneously. J Mach Learn Res. 2019;20:1–81.

143. Janssen RJ, Mourao-Miranda J, Schnack HG. Making individual prognoses in psychiatry using neuroimaging and machine learning. Biol Psychiatry Cogn Neurosci Neuroimaging. 2018;3:798–808.

144. Stephan KE, Schlagenhauf F, Huys QJM, Raman S, Aponte EA, Brodersen KH, et al. Computational neuroimaging strategies for single patient predictions. Neuroimage. 2017;145:180–99.

145. Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. Am J Psychiatry. 2010;167:748–51.

146. Kushki A, Anagnostou E, Hammill C, Duez P, Brian J, Iaboni A, et al. Examining overlap and homogeneity in ASD, ADHD, and OCD: a data-driven, diagnosis-agnostic approach. Transl Psychiatry. 2019;9:318.

147. Cordova M, Shada K, Demeter DV, Doyle O, Miranda-Dominguez O, Perrone A, et al. Heterogeneity of executive function revealed by a functional random forest approach across ADHD and ASD. Neuroimage Clin. 2020;26:102245.

148. Jacobs GR, Voineskos AN, Hawco C, Stefanik L, Forde NJ, Dickie EW, et al. Integration of brain and behavior measures for identification of data-driven groups cutting across children with ASD, ADHD, or OCD. Neuropsychopharmacology. 2021;46:643–53.

AUTHOR CONTRIBUTIONS

This review was conceptualized by XL. The literature search, review of publications, and data extraction were completed by MC. Manuscript and supplementary materials were written and all figures were constructed by MC, EM, and XL.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41398-023-02536-w.

Correspondence and requests for materials should be addressed to Xiaobo Li.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.