INDIVIDUALIZED PREDICTION MODELS IN ADHD:

A SYSTEMATIC REVIEW AND META-REGRESSION

Gonzalo Salazar de Pablo, MD, PhD[1,-3], Raquel Iniesta, PhD[4], Alessio Bellato, PhD[5], Arthur Caye, MD, PhD[6-8], Maja Dobrosavljevic, PhD[9], Valeria Parlatini[10,11], Miguel Garcia-Argibay, PhD[9,12], Lin Li, PhD[9,12], Anna Cabras, MD[13], Mian Haider Ali, MSc[4], Lucinda Archer, PhD[14], Alan J Meehan, PhD[15,16], Halima Suleiman, MD[17], Marco Solmi, MD, PhD[10,18-22], Paolo Fusar-Poli, MD, PhD[23-26], Zheng Chang, PhD[12], Stephen V Faraone, PhD[17], Henrik Larsson, PhD[9,12], Samuele Cortese, MD, PhD [10-11, 27-29]

[1]Department of Child and Adolescent Psychiatry, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London
[2]Child and Adolescent Mental Health Services, South London and Maudsley NHS Foundation Trust, London, UK
[3]Institute of Psychiatry and Mental Health. Department of Child and Adolescent Psychiatry, Hospital General Universitario Gregorio Marañón School of Medicine, Universidad Complutense, Instituto de Investigación Sanitaria Gregorio Marañón (IiSGM), CIBERSAM, Madrid, Spain
[4]Department of Biostatistics and Health Informatics, Institute of Psychiatry, Psychology and Neurosciences, King's College London, London, UK
[5]School of Psychology, University of Nottingham, Malaysia
[6] Post-Graduate Program of Psychiatry, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil
[7]National Center for Research and Innovation (CISM), University of São Paulo, Brazil.
[8]ADHD Outpatient Program, Hospital de Clínicas de Porto Alegre, Brazil
[9]School of Medical Sciences, Faculty of Medicine and Health, Örebro University, Örebro, Sweden
[10]Centre for Innovation in Mental Health - Developmental Lab, School of Psychology, University of Southampton
[11]Solent NHS Trust, Southampton, UK
[12]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
[13]Department of Neurology and Psychiatry, University of Rome La Sapienza, Rome, Italy
[14]Institute of Applied Health Research, University of Birmingham, UK
[15]Department of Psychology, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK
[16]Yale Child Study Center, Yale School of Medicine, New Haven, USA
[17]Departments of Psychiatry and of Neuroscience and Physiology, SUNY Upstate Medical University, Syracuse, NY, USA
[18]Department of Psychiatry, University of Ottawa, Ontario, Canada
[19]Department of Mental Health, The Ottawa Hospital, Ontario, Canada; Ottawa
[20]Hospital Research Institute (OHRI) Clinical Epidemiology Program University of Ottawa Ontario, Canada
[21]School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, Ottawa, Canada

[22]Department of Child and Adolescent Psychiatry, Charité Universitätsmedizin, Berlin, Germany.
[23]Early Psychosis: Interventions and Clinical-detection (EPIC) Lab, Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK
[24]Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy
[25]Outreach and Support in South-London (OASIS) service, South London and Maudsley NHS Foundation Trust, London, UK
[26]Department of Psychiatry and Psychotherapy, University Hospital, Ludwig-Maximilian-University (LMU), Munich, Germany
[27]Clinical and Experimental Sciences (CNS and Psychiatry), Faculty of Medicine, University of Southampton, Southampton, UK
[28]Hassenfeld Children's Hospital at NYU Langone, New York University Child Study Center, New York City, NY, USA
[29]DiMePRe-J-Department of Precision and Rigenerative Medicine-Jonic Area, University of Bari "Aldo Moro", Bari, Italy

Abstract: 297 words
Text: 3445 words
Tables: 2
Figures: 3
Supplemental text: 3:
Supplemental tables: 10
References: 123
Endnote reference file: ADHD_riskmodels.enl

Correspondence to: Prof Samuele Cortese MD PhD, Centre for Innovation in Mental Health - Developmental Lab, School of Psychology, University of Southampton, and Solent NHS Trust, Southampton, UK. Email: samuele.cortese@soton.ac.uk

ABSTRACT

There have been increasing efforts to develop prediction models supporting personalised detection, prediction, or treatment of ADHD. We overviewed the current status of prediction science in ADHD by: 1) systematically reviewing and appraising available prediction models; 2) quantitatively assessing factors impacting the performance of published models. We did a PRISMA/CHARMS/TRIPOD-compliant systematic review (PROSPERO: CRD42023387502), searching, until 20/12/2023, studies reporting internally and/or externally validated diagnostic/prognostic/treatment-response prediction models in ADHD. Using meta-regression, we explored the impact of factors affecting the area under the curve (AUC) of the models. We assessed the study risk of bias with the Prediction Model Risk of Bias Assessment Tool (PROBAST). From 7,764 identified records, 100 prediction models were included (88% diagnostic, 5% prognostic, and 7% treatment-response). Of these, 96% and 7% were internally and externally validated, respectively. None was implemented in clinical practice. Only 8% of the models were deemed at low risk of bias; 67% were considered at high risk of bias. Clinical, neuroimaging, and cognitive predictors were used in 35%, 31%, and 27% of the studies, respectively. The performance of ADHD prediction models was increased in those models including, compared to those models not including, clinical predictors (β=6.54, p=0.007). Type of validation, age range, type of model, number of predictors, study quality, and other type of predictors did not alter the AUC. Several prediction models have been developed to support the diagnosis of ADHD. However, efforts to predict outcomes or treatment response have been limited, and none of the available models is ready for implementation into clinical practice. The use of clinical predictors, which may be combined with other type of predictors, seems to improve the performance of the models. A new generation of research should address these gaps by conducting high quality, replicable, and externally validated models, followed by implementation research.

Key words: ADHD; Prediction models; Risk; Prediction; Evidence.

INTRODUCTION

Attention-Deficit/Hyperactivity Disorder (ADHD)[1] is a neurodevelopmental condition which is characterized by age-inappropriate and impairing inattention and/or hyperactivity/impulsivity. Over the past decades, neurobiological research has resulted in a shift in the understanding of the pathophysiology of ADHD, from theoretical views of isolated brain dysfunctions to more complex models reflecting the heterogeneity of the clinical manifestations of ADHD[2]. However, neurobiological findings have not yet impacted clinical practice and, currently, the diagnosis of ADHD is exclusively based on a clinical assessment, with no established objective tests being available as standalone tools to diagnose ADHD[3]. The exact factors that predict the persistence of ADHD beyond adolescence are currently unclear. Furthermore, while effective (at least in the short-term) treatments are available[4], there are no established evidence-based prediction models to inform individualized treatment strategies based on the patient's clinical, environmental, cognitive, genetic, or biological characteristics.

In the last decade, the new field of precision psychiatry has emerged, with the development of multivariable prediction models aimed at predicting the diagnosis, prognosis, or treatment response in relation to several mental health conditions[5, 6], considering individual variability in clinical characteristics, genes, environment, and lifestyle[7]. Advances in the field of prediction modelling have allowed the consolidation of an evidence-based science of precision medicine[8]. Prediction modelling studies investigate the development of such models, as well as their validation[9]. External validity is the extent to which predictions can be generalized to the data from other settings, while internal validity is the extent to which the predictions fit the derivation data[10].

Previous systematic reviews have identified a large number of prediction models across mental health conditions[10, 11]. Notably, in the last few years there has been a rapidly increasing interest in this field, and an emerging number of prediction models on ADHD have been rapidly

published, making an updated evaluation of the status of the field essential. Furthermore, to our knowledge, no study has comprehensively and specifically reviewed the status of validated prediction models in ADHD, systematically assessing factors that can affect their predictive performance. Therefore, our primary aim was to systematically review and critically appraise available prediction models that might be considered for clinical use in the identification or management of ADHD. Our secondary aim was to test potential moderating factors that could affect the performance of available models as measured by their area under the curve (AUC), the most reliable and most reported metric across studies.

METHODS

This study (pre-registered protocol: PROSPERO:CRD42023387502) was conducted and reported in accordance with the "Preferred Reporting Items for Systematic Reviews and Meta-analyses" (PRISMA) 2020 and the "Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis" (TRIPOD) statements and checklists (Tables S1-4, available online).

Search strategy and selection criteria

PubMed and Web of Science database including Web of Science Core Collection, BIOSIS Citation Index, KCI-Korean Journal Database, MEDLINE, Russian Science Citation Index, and SciELO Citation Index, Cochrane Central Register of Reviews, Embase and Ovid/PsycINFO databases, were searched from inception until 20/12/2023 with no language restrictions (search terms/syntax in Supplement 1, available online). The references of the included articles and those in previous relevant reviews were manually searched to identify any possible additional relevant studies. Titles and abstracts were screened, and, after the exclusion of those not relevant, the full texts were assessed against the inclusion and exclusion criteria by a group of researchers who worked independently in pairs on one third of the hits each (G.S.P., A.B., A.C., M.D., A.C., H.S., V.P.).

The inclusion criteria were: (a) original individual studies; (b) conducted in children and/or adults with ADHD according to established diagnostic criteria (DSM or ICD - any version); (c) reporting on multivariable internally and/or externally[12] validated prediction models; (d) providing diagnostic, prognostic, or treatment-response estimates at the individual subject level or in subgroups; (e) providing at least discrimination as per the AUC (i.e., the ability of the model to separate individuals who develop events from those who do not), accuracy (i.e., the degree of closeness of the measured value), or classification measures (sensitivity, specificity, or predictive values) (definitions in Table 1). The exclusion criteria were: (a) abstracts, conference proceedings, reviews, or meta-analyses; (b) prediction model studies that did not evaluate or report their internal or external validation; (c) predictor-finding studies that included one predictor only.

Descriptive measures and data extraction

Data extraction items (Supplement 2, available online) were based on the "Checklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies" (CHARMS) and the "Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis" (TRIPOD) statements. The model's ability to separate individuals with and without the outcome, e.g., AUC, was selected as the main outcome. Discriminative validity is usually considered 'acceptable' when AUC scores are between 0.7-0.8, 'good' between 0.8-0.9, and 'excellent' when >0.9[13]. We extracted information on the performance of each model assessed by other measures when reported. When more than one outcome per study was found in the same category, we extracted the information for the primary outcome, as defined in each article, unless the study reported multiple primary co-outcomes. We relied on what the individual authors reported as their primary outcome.

Quality assessment

Risk of bias was assessed for each of the included studies with a validated version - previously used in mental health research- of the Prediction Model Risk of Bias Assessment Tool (PROBAST v5/05/2019)[9] (Supplement 3, available online).

Strategy for Data Synthesis

Data from the included studies were first summarized in descriptive tables. The top 10% of the most commonly employed predictor types were shown in a bar chart. We then conducted meta-regressions to estimate the association, when data were available, between AUC and: (i) the type of validation (internal vs external); (ii) the age range (children and adolescents vs adults vs combined/not reported); (iii) the type of model (diagnostic vs prognostic vs treatment-response model); (iv) the number of predictors; (v) the type of predictors [clinical/sociodemographic vs any biomarker (neuroimaging, electroencephalography, magnetoencephalography, proteomic, genetic, cognitive, or a combination of modalities)][10]; (vi) the modality of predictors [unimodal, using only one type of predictor (e.g., clinical only) vs multimodal, using more than one type of predictor (e.g., clinical and biomarker)] and (vii) the quality of the studies (low risk vs unclear risk vs high risk). We used a random-effects model to allow for heterogeneity in underlying associations across studies. Number of studies permitting, we also planned sensitivity analyses to assess the impact of studies being at low risk of bias and without suboptimal validation. Suboptimal validation was appraised by two statisticians (R.I. and M.H.I.) with a focus on: 1) double dipping (i.e., performing feature selection or selection of tuning -or penalty- parameters on data samples from both the training and the test set)[14]; 2) reporting apparent/non-validated predictive performance instead of the validated predictive performance; 3) reporting the size and significance of apparent regression coefficients rather than the cross-validated performance measure, and 4) re-estimating regression coefficients in the test set, instead of applying the apparently validated model. The meta-regression was performed with Comprehensive Meta-Analysis Version 3. Statistical significance was set at $p<0.05$.

RESULTS

After removing duplicates, from an initial pool of 7764 references, we retained 100 eligible studies (Figure 1). None of the models reported in the included studies was implemented into clinical practice. 96 (96.0%) and seven (7.0%) models were internally and externally validated, respectively. None of the models was implemented into clinical practice. Among the eligible studies, 88.0% reported on diagnostic prediction models, 5.0% on prognostic models (with outcomes such as symptom change or development of substance use disorders), and 7.0% on treatment-response models. The retained studies most frequently used clinical (35.0%), neuroimaging (31.0%) and cognitive (27.0%), predictors (Figure 2). The total sample size was 323,554 individuals, ranging from 10 to 238,696 individuals per study. The average age was 15.7 years. The source of data encompassed case-control studies (73 studies, 73.0%), cohort studies (23 studies, 23.0%), and clinical trials (4 studies, 4.0%). AUC was the most commonly reported measure of model performance (61.0%), followed by accuracy (36.0%). Eight studies (8.0%) only reported the sensitivity and specificity of the models.

Predictors in prediction models

In the 88 diagnostic prediction models, studies used cognitive (K=5 studies)[15-19], clinical (K=13)[20-31287], neuroimaging (K=19)[32-50], EEG (K=15)[51-65], genetic (K=2)[66, 67], ECG (K=2)[68, 69], physical health (K=1)[70], EEG and cognitive (K=4)[71-74], sociodemographic and neuroimaging (K=4)[75-77294], clinical and cognitive (K=4)[78-81], sociodemographic and cognitive (K=2)[82, 83], cognitive and physical health (K=2)[84, 85], genetic and neuroimaging (K=2)[86, 87], sociodemographic and genetic (K=1)[88], clinical and sociodemographic (K=1)[89] EEG and EMG (K=1)[90], sociodemographic and neuroimaging (K=2)[91, 92], sociodemographic, clinical and cognitive (K=3)[93-95], cognitive, sociodemographic and neuroimaging (K=1)[96], clinical,

sociodemographic and neuroimaging (k=1)[97], clinical, cognitive and neuroimaging (k=1)[98] and sociodemographic, clinical, cognitive and physical health (K=2)[99, 100] predictors.

In the 5 prognostic prediction models, studies employed sociodemographic and clinical (K=2)[101, 102], physical health and clinical (K=1)[103], neuroimaging and genetic (K=1)[104], and clinical and genetic (K=1)[105] predictors.

In the 7 treatment-response prediction models, studies relied on neuroimaging (K=1)[106], genetic (k=1)[107], sociodemographic, clinical, cognitive, and physical health (K=1)[108], genetic, cognitive and physical health (K=1)[109], clinical, sociodemographic, service use and physical health (K=1)[110], sociodemographic and clinical predictors (K=1)[111], and sociodemographic, clinical and physical health (K=1) predictors[112].

Performance of prediction models

The performance of ADHD prediction models was highly variable, with AUC ranging from 0.50 to 0.99. AUC ranged from 0.50 to 0.96 in diagnostic models, from 0.73 to 0.87 in prognostic models, and from 0.72 to 0.99 in models for predicting treatment-response. Accuracy ranged from 0.53 to 1.0 (0.53 to 1.0 for diagnostic models, 0.73 to 0.87 for prognostic models, and 0.72 to 0.88 for treatment-response models) (Tables S5-7, available online). Model calibration was assessed in 6.0% of the studies.

Meta-regression results

The performance of ADHD prediction models was increased in those models including (K=26), as compared to those not including, clinical predictors (K=36) (β=6.540, p=0.007). No significant findings emerged when considering type of validation (internal K=58 vs external K=4), age range (children and adolescents K=33 vs adults K=11 vs combined/not reported K=18), type of prediction model (diagnostic K=52 vs prognostic K=3 vs treatment-response K=7) (p>0.05), number of predictors (K=34), other types of predictors comparisons

(clinical/service use/sociodemographic K=13 vs biomarkers K=33 vs combination K=17) or quality of the studies (low risk K=7 vs unclear risk K=11 vs high risk K=44) (all p>0.05) (Table 2), according to our meta-regression analyses.

Quality of prediction models

Sixty-seven (67.0%) of the included studies were deemed to be at high risk of bias according to the PROBAST tool. The results from the different domains were heterogeneous: 9 (9.0%) were at high risk of bias in the participants domain, 11 (11.0%) in the predictors domain, 19 (19.0%) in the outcomes domain, and 61 (61.0%) in the analysis domain. Only 8 (8.0%) of the included studies (seven diagnostic and one prognostic) were considered to be at overall low risk of bias; 86 (86.0%) of the studies were deemed at low risk of bias in the participants domain, 60 (60.0%) in the predictors domain, 57 (57.0%) in the outcomes domain, and 18 (18.0%) in the analysis domain (Table S8, available online; Figure 3). In 13 studies (13.0%) the performance was evaluated in development dataset only, resulting in suboptimal validation (Table S9, available online).

DISCUSSION

This is the first systematic review to quantitatively summarize the evidence regarding internally or externally validated diagnostic, prognostic, or treatment-response prediction models specifically in the field of ADHD, appraising the quality of the models and assessing possible factors affecting their performance in terms of AUC. Among the 100 prediction modelling studies included, 88% reported on diagnostic, 5% on prognostic, and 7% on treatment-response models. Furthermore, 35% of studies used clinical, 31.0% neuroimaging, and 27.0% cognitive predictors. Notably, only 7.0% of models were externally validated. The performance of ADHD prediction models was increased in those models including, compared to those models not including, clinical predictors. Meta-regressions did not detect any significant

changes in the AUC according to other evaluated variables. Also, 67.0% of included studies were found to be at high risk of bias according to PROBAST quality assessment.

Our review shows that the number of prediction models in the field of ADHD is increasing exponentially over the years, with a wide range of predictors that might potentially support the diagnosis of ADHD, and, to a lesser extent, the understanding of the clinical progression of the disorder or the factor influencing the response to interventions.

However, the discrimination and accuracy of the models, although good, may not be enough for implementation into clinical practice. This emerging body of research is limited by not only a small number of externally validated models, but also, and crucially, by lack of implementation research in real-world clinical practice. Our findings align with previous evidence related to other mental health conditions suggesting that external validation of prediction models is still infrequent in psychiatry/mental health[113]. A similar review exploring prediction models across any mental health condition found that only 20.1% of all prediction models were externally validated[11]. Another review found that 30.3% of all models were externally validated following strict validation criteria (4.6% of the total models)[10]. This is in contrast with the status of prediction science in other areas of medicine. For instance, several models have been externally validated between five to seventeen times in the field of chronic obstructive pulmonary disease[114]. Similar approaches may move the field of ADHD forward, ensuring generalizability of the model to clinical populations not used to develop the model.

Within the internally and externally validated models, there was no significant correlation seen between the internal and external performance measures. However, the number of models internally and externally validated (six studies) was low. Our findings may also reflect a suboptimal quality during the internal validation of the models, potentially leading to optimism in the reported performance measures and high risk of bias. In fact, 67.0% of the included studies were found to be at high risk of bias. The “analysis” domain, where 61.0% of the included studies were found to be at high risk of bias seems particularly problematic. Also,

calibration was assessed in 6.0% of the studies only. Future prediction models need to make sure that: 1) the sample size is appropriate and there is an appropriate number of participants developing the outcome (which may vary depending on the population and outcome of interest); 2) the number of predictors is appropriate, 3) the missing data is handled appropriately (of note, >80% studies did not report how they handled missing data, 7.0% deleted missing data and only 4.0% carried out multiple imputation techniques, being low adherence to ADHD treatment frequent); 4) complexities in the data (e.g. competing risks, sampling of controls) are accounted for appropriately, and 5) model overfitting and optimism in model performance are accounted for, among other key criteria to develop and validate prediction models[9]. We note that risk of bias was heterogeneous across the different PROBAST domains: only 9.0% of models were at high risk of bias in the participants domain and only 11.0% were at high risk of bias in the predictors domain. Given the strictness of PROBAST scoring thresholds, this highlights some strengths among published ADHD prediction models in the selection of participants and predictors.

In terms of the aim of the models, most of them were intended to support the prediction of ADHD diagnosis (88/100), reflecting an increasing interest in developing diagnostic prediction models in the ADHD field, alongside other mental health conditions such as depression[115], first episode psychosis[116], or bipolar disorder[117], following a similar route, likely due to the perception of the suboptimal nature of a "subjective" diagnosis. Notably, unlike other mental health conditions, where performance in diagnostic models has been found to be superior to that in prognostic and treatment-response models[10], we did not find this to be the case in the field of ADHD. The limited number of available treatment-response models points to a critical need for carefully designed experimentally controlled trials (or high quality observational studies) to identify biomarkers that index inter-individual variability and predict treatment response[118]. While studies on treatment-response models are complex to perform, mostly due to the intervention-related components (particularly randomized clinical trials), as well as to

ethical issues[10], observational studies relying on electronic health-care records on the long-term effectiveness and safety of the interventions could provide a meaningful alternative[119].

In terms of predictor types, a significant proportion of the reviewed prediction models included clinical predictors, followed closely by neuroimaging predictors and cognitive predictors. The performance of ADHD prediction models was higher in those models including clinical predictors, compared to those models not including clinical predictors. Thus, the use of clinical predictors, which may be combined with other type of predictors, may improve the performance of the models and their inclusion should be considered in prediction models[81]. However, it is important to note that further research is needed to validate these results across different populations, and including additional predictors[99]. While clinical predictors seem to be clearly predominant in other fields[10], in the ADHD field different biomarkers have commonly been used to aid the detection and correct characterization of ADHD. However, there is currently no biomarker in any neurodevelopmental condition, including ADHD, for which there is evidence from two or more studies from independent research groups, with results going into the same direction and of specificity and sensitivity of at least 80%[3]. This makes it difficult to recommend the use of any specific individual predictor in isolation, for future prediction models. Notably, we also found no evidence that multimodal prediction models achieved higher accuracy than unimodal models, arguing against the development of complex models with a wide variety of biomarkers and predictors (which would also be more difficult to apply and implement). In other words, we found no evidence that more complex prediction models encompassing biomarkers or a large number of predictors (which may be more prone to overfitting issues) outperformed less complex models. However, from a quality perspective, five of the six studies assessed at low risk of bias were multimodal, so caution is recommended in the interpretation of this finding.

Future studies should consider net benefit approaches for the evaluation of prediction models for ADHD, which were not used in any of the studies in this review. Net benefit approaches put the benefits and harms of using a prediction model on the same scale, to allow assessment of

the relative value associated with using prediction models to guide clinical decision making, over other patient management strategies[120], an approach which is currently lacking in the ADHD prediction literature. 74% of the studies were case-control studies which tried to differentiate individuals with ADHD and healthy controls. Future studies should also try to differentiate ADHD from other relevant syndromes such as the cognitive disengagement syndrome (CDS) -or sluggish cognitive tempo-. CDS is an emerging condition -as opposed to a transdiagnostic phenomenon- in the field of child, adolescent and adult psychiatry[1, 2]. The presence of CDS is particularly important as misdiagnosis of this condition may result in a poor response to first-line treatment with methylphenidate and unwanted side effects[3, 4]. Furthermore, among children crossing into adolescence with ADHD, CDS can result in poor physical activity and behavior[2].

Our study should be considered in the light of its limitations. Our study has several limitations that must be taken into consideration, mainly related to issues in the available studies rather than in our methods. The main limitation rests in the heterogeneity of the characteristics of prediction models developed in the included studies. The predictors used to develop the models varied considerably across studies. Therefore, in line with previous studies[10], we did not attempt to meta-analyse the categories of prediction models; rather, we presented only meta-regression analyses, stratifying the models for methodological features. We also could not conduct meta-regressions on the studies at low risk of bias and without suboptimal methodological strategies in regard to validation. The sample size and the quality of the studies was highly heterogeneous, with high risk of bias observed in 67.0% of included studies according to the PROBAST criteria, including 61.0% in the analysis domain. Final scores of the PROBAST should be taken with caution as the thresholds are stringent and an outcome is considered to be at high risk of bias when one or more of the questions is answered as not appropriate. We did not analyse the differences among validation measures, some of them being prone to data leakage and inflated accuracy or overfitting. We might have missed

relevant studies, particularly if not published. Finally, we could not provide data about calibration as this was rarely reported.

In conclusion, several validated prediction models have been proposed to support the diagnosis of ADHD. However, efforts to predict prognostic outcomes or treatment response to ADHD have been limited. Advances in the field are limited by lack of implementation research in real-world clinical practice. A new generation of research should address these gaps by conducting high quality, replicable and externally validated models. Once an evidence-based model is available, efforts to disseminate it and implement it into clinical practice are recommended.

DECLARATION OF INTEREST: Dr. Salazar de Pablo has received honoraria from Janssen Cilag, Lundbeck and Angelini. Dr Archer is supported by funding from the NIHR Birmingham Biomedical Research Centre at the University Hospitals Birmingham NHS Foundation Trust and the University of Birmingham. Dr. Solmi has received honoraria/has been a consultant for Abbvie, Angelini, Lundbeck, Otsuka. Prof. Fusar-Poli has received grant support from Lundbeck and honoraria fees from Angelini, Menarini, and Lundbeck. In the past year, Dr. Faraone received income, potential income, travel expenses continuing education support and/or research support from Aardvark, Aardwolf, AIMH, Tris, Otsuka, Ironshore, KemPharm/Corium, Akili, Supernus, Atentiv, Noven, Sky Therapeutics, Axsome and Genomind. With his institution, he has US patent US20130217707 A1 for the use of sodium-hydrogen exchange inhibitors in the treatment of ADHD. He also receives royalties from books published by Guilford Press: Straight Talk about Your Child's Mental Health, Oxford University Press: Schizophrenia: The Facts and Elsevier: ADHD: Non-Pharmacologic Interventions. He is Program Director of www.ADHDEvidence.org and www.ADHDinAdults.com. Prof. Cortese has declared reimbursement for travel and accommodation expenses from the Association for Child and Adolescent Central Health (ACAMH) in relation to lectures delivered for ACAMH,

the Canadian AADHD Alliance Resource, the British Association of Psychopharmacology, and from Healthcare Convention for educational activity on ADHD.

FUNDING: Prof. Fusar-Poli is supported by #NEXTGENERATIONEU (NGEU), funded by the Ministry of University and Research (MUR), National Recovery and Resilience Plan (NRRP), project MNESYS (PE0000006) – A Multiscale integrated approach to the study of the nervous system in health and disease (DN. 1553 11.10.2022). Samuele Cortese, NIHR Research Professor (NIHR303122) is funded by the NIHR for this research project. The views expressed in this publication are those of the author(s) and not necessarily those of the NIHR, NHS or the UK Department of Health and Social Care. Samuele Cortese is also supported by NIHR grants NIHR203684, NIHR203035, NIHR130077, NIHR128472, RP-PG-0618-20003 and by grant 101095568-HORIZONHLTH- 2022-DISEASE-07-03 from the European Research Executive Agency.

AUTHOR CONTRIBUTIONS:

GSP and SC conceived the study. GSP conducted the analyses and drafted the first version of the manuscript. RI provided statistical support. SC supervised the project and study. RI, AB, AC, MD, VP, MG-B, LL, AC, MHA and HS obtained the data. LA, AJM, MS, PF-P, ZC, SVF and HC revised the manuscript and provided a substantial conceptual contribution. All authors proofread and approved the final draft of the manuscript**.**

Table 1. Definitions of key terms in prediction science.

Term	Definition
Accuracy	The degree of closeness of the measured value
Area under the curve (AUC)	The area enclosed by the curve of a mathematical function and a reference axis
Calibration	The degree of adjustment of a measurement to account for the sources of variation
Classification	The degree of assignment of an outcome to the correct category
Discrimination	The ability of the model to separate individuals who develop events from those who do not
External validation	Process of evaluating the extent to which the predictions can be generalized to the data from other settings
Internal validation	Process of evaluating the extent to which the predictions fit the derivation data after controlling for overfitting and optimism
Optimism	Increase in the assigned performance values due to methodological bias
Overfitting	A modeling error consisting on a measure being too closely aligned to a previous set of data points
Performance	The degree of execution of a model in regards to discrimination and calibration aspects
Prediction modelling studies	Studies that use statistical procedures to predict the appearance of a condition or outcome.

Table 2. Meta-regressions exploring the possible moderating factors impacting the area under the curve (AUC).

Moderating factors	Number of studies	Meta-regression Coefficient	SE	Z value	P	95%CI
(i) Type of validation						
External vs Internal validation	62 (58;4)	2.319	-3.79	0.75	0.454	-3.790; 8.474
(ii) Age range						
Children&adolescents vs adults	44 (33;11)	-1.612	2.976	-0.54	0.588	-7.444; 4.221
Children&adolescents vs combined/not reported	51 (33;18)	0.482	2.696	0.18	0.858	-4.802; 5.766
Adults vs combined/not reported	29 (11;18)	2.093	3.344	0.63	0.531	-4.460; 8.647
(iii) Type of model						
Diagnostic vs prognostic	55 (52;3)	-2.073	5.786	-0.36	0.720	-13.413; 9.268
Diagnostic vs treatment-response	59 (52;7)	1.045	3.538	0.30	0.7677	-5.889; 7.979
Prognostic vs treatment-response	10 (3;7)	3.118	6.514	0.48	0.632	-9.649; 15.885
(iv) Number of predictors						
Number of predictors	34	-0.0047	0.003	-1.39	0.166	-0.011; 0.002
(v) Type of predictors						
Clinical or service use or sociodemographic vs biomarkers	46 (13; 33)	-1.249	4.337	-0.29	0.773	-9.750; 7.252
Clinical or service use or sociodemographic vs combination	30 (13;17)	-2.749	2.935	-0.94	0.349	-8.502; 3.003
Biomarkers vs combination	50 (33;17)	-0.404	3.786	-0.11	0.915	-7.825; 7.017
Including clinical vs not including clinical	62 (26;36)	-6.540	2.410	-2.72	**0.007**	-11.264; -1.824
(vi) Modality of predictors						
Unimodal vs Multimodal	62 (36;26)	-3.496	1.873	-1.87	0.062	-7.167; 0.174
(vii) Quality assessment						
Low risk vs unclear risk	18 (7;11)	3.629	5.058	0.72	0.473	-6.284; 13.542
Low risk vs high risk	51 (7; 44)	5.549	3.910	1.42	0.156	-6.284; 13.542
Unclear risk vs high risk	55 (11; 44)	1.920	3.999	0.48	0.631	-5.918; 9.759

Figure 1. **Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) flowchart outlining study selection process**.

Figure 2. Most frequently reported predictors across prediction model types.

Figure 3. Risk of bias of the retrieved studies, as assessed by the PROBAST tool.

REFERENCES

1. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders.*: Washington DC,, 2013.

2. Cortese S. The neurobiology and genetics of Attention-Deficit/Hyperactivity Disorder (ADHD): What every clinician should know. *European Journal of Paediatric Neurology* 2012; **16**(5)**:** 422-433.

3. Cortese S, Solmi M, Michelini G, Bellato A, Blanner C, Canozzi A *et al.* Candidate diagnostic biomarkers for neurodevelopmental disorders in children and adolescents: a systematic review. *World Psychiatry* 2023; **22**(1)**:** 129-149.

4. Cortese S. Pharmacologic Treatment of Attention Deficit-Hyperactivity Disorder. *N Engl J Med* 2020; **383**(11)**:** 1050-1056.

5. Fernandes BS, Williams LM, Steiner J, Leboyer M, Carvalho AF, Berk M. The new field of 'precision psychiatry'. *BMC Med* 2017; **15**(1)**:** 80.

6. Khanra S, Khess CRJ, Munda SK. "Precision psychiatry": A promising direction so far. *Indian J Psychiatry* 2018; **60**(3)**:** 373-374.

7. Terry SF. Obama's Precision Medicine Initiative. *Genet Test Mol Biomarkers* 2015; **19**(3)**:** 113-114.

8. Fusar-Poli P, Hijazi Z, Stahl D, Steyerberg EW. The Science of Prognosis in Psychiatry: A Review. *JAMA Psychiatry* 2018; **75**(12)**:** 1289-1297.

9. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS *et al.* PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. *Ann Intern Med* 2019; **170**(1)**:** 51-58.

10. Salazar de Pablo G, Studerus E, Vaquerizo-Serrano J, Irving J, Catalan A, Oliver D *et al.* Implementing Precision Psychiatry: A Systematic Review of Individualized Prediction Models for Clinical Practice. *Schizophr Bull* 2021; **47**(2)**:** 284-297.

11. Meehan AJ, Lewis SJ, Fazel S, Fusar-Poli P, Steyerberg EW, Stahl D *et al.* Clinical prediction models in psychiatry: a systematic review of two decades of progress and challenges. *Mol Psychiatry* 2022; **27**(6)**:** 2700-2708.

12. Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. *BMC Med Res Methodol* 2013; **13:** 33.

13. Safari S, Baratloo A, Elfil M, Negida A. Evidence Based Emergency Medicine; Part 5 Receiver Operating Curve and Area under the Curve. *Emerg (Tehran)* 2016; **4**(2)**:** 111-113.

14. Kassraian-Fard P, Matthis C, Balsters JH, Maathuis MH, Wenderoth N. Promises, Pitfalls, and Basic Guidelines for Applying Machine Learning Classifiers to Psychiatric Imaging Data, with Autism as an Example. *Front Psychiatry* 2016; **7:** 177.

15. Altinkaynak M, Dolu N, Guven A, Pektas F, Ozmen S, Demirci E *et al.* Diagnosis of Attention Deficit Hyperactivity Disorder with combined time and frequency features. *Biocybernetics and Biomedical Engineering* 2020; **40**(3)**:** 927-937.

16. 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**(1)**:** 204-215.

17. Lee HJ, Cho S, Shin MS. Supporting diagnosis of attention-deficit hyperactive disorder with novelty detection. *Artif Intell Med* 2008; **42**(3)**:** 199-212.

18. Predicting ADHD Risk from Touch Interaction Data. *Proceedings of the 20th ACM International Conference on Multimodal Interaction (ICMI)*; 2018 Oct 16-20 2018; Boulder, CO2018.

19. Slobodin O, Yahav I, Berger I. A Machine-Based Prediction Model of ADHD Using CPT Data. *Frontiers in Human Neuroscience* 2020; **14:** 560021.

20. Altun S, Alkan A, Altun H. Application of deep learning and classical machine learning methods in the diagnosis of attention deficit hyperactivity disorder according to temperament features. *Concurrency and Computation-Practice & Experience* 2022; **34**(13).

21. Bledsoe JC, Xiao C, Chaovalitwongse A, Mehta S, Grabowski TJ, Semrud-Clikeman M *et al.* Diagnostic Classification of ADHD Versus Control: Support Vector Machine Classification Using Brief Neuropsychological Assessment. *J Atten Disord* 2020; **24**(11)**:** 1547-1556.

22. Deserno MK, Bathelt J, Groenman AP, Geurts HM. Probing the overarching continuum theory: data-driven phenotypic clustering of children with ASD or ADHD. *Eur Child Adolesc Psychiatry* 2022.

23. Duda M, Ma R, Haber N, Wall DP. Use of machine learning for behavioral distinction of autism and ADHD. *Translational Psychiatry* 2016; **6**.

24. Duda M, Haber N, Daniels J, Wall DP. Crowdsourced validation of a machine-learning classification system for autism and ADHD. *Translational Psychiatry* 2017; **7**.

25. Kim S, Lee HK, Lee K. Can the MMPI Predict Adult ADHD? An Approach Using Machine Learning Methods. *Diagnostics (Basel)* 2021; **11**(6).

26. Shi Y, Schulte PJ, Hanson AC, Zaccariello MJ, Hu D, Crow S *et al.* Utility of medical record diagnostic codes to ascertain attention-deficit/hyperactivity disorder and learning disabilities in populations of children. *BMC Pediatr* 2020; **20**(1)**:** 510.

27. Silverstein M, Hironaka LK, Feinberg E, Sandler J, Pellicer M, Chen N *et al.* Using Clinical Data to Predict Accurate ADHD Diagnoses Among Urban Children. *Clin Pediatr (Phila)* 2016; **55**(4)**:** 326-332.

28. Chen T, Tachmazidis I, Batsakis S, Adamou M, Papadakis E, Antoniou G. Diagnosing attention-deficit hyperactivity disorder (ADHD) using artificial intelligence: a clinical study in the UK. *Front Psychiatry* 2023; **14:** 1164433.

29. Chu KC, Huang HJ, Huang YS. Validity of Diagnostic Support Model for Attention Deficit Hyperactivity Disorder: A Machine Learning Approach. *J Pers Med* 2023; **13**(11).

30. Haque UM, Kabir E, Khanam R. Early detection of paediatric and adolescent obsessive-compulsive, separation anxiety and attention deficit hyperactivity disorder using machine learning algorithms. *Health Inf Sci Syst* 2023; **11**(1)**:** 31.

31. Kurokami T, Kobayashi H, Nakajima M, Mikami M, Koeda T. Establishment of an objective index for the diagnosis of attention deficit/hyperactivity disorder by the continuous performance test "MOGRAZ". *Brain Dev* 2022; **44**(10)**:** 664-671.

32. 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. *Frontiers in Systems Neuroscience* 2012; (DEC).

33. Chaim-Avancini TM, Doshi J, Zanetti MV, Erus G, Silva MA, Duran FLS *et al.* Neurobiological support to the diagnosis of ADHD in stimulant-naïve adults: pattern recognition analyses of MRI data. *Acta Psychiatr Scand* 2017; **136**(6)**:** 623-636.

34. Cheng W, Ji X, Zhang J, Feng J. Individual classification of ADHD patients by integrating multiscale neuroimaging markers and advanced pattern recognition techniques. *Frontiers in systems neuroscience* 2012; **6:** 58.

35. Dai D, Wang J, Hua J, He H. Classification of ADHD children through multimodal magnetic resonance imaging. *Front Syst Neurosci* 2012; **6:** 63.

36. 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. *Human brain mapping* 2014; **35**(7)**:** 3083-3094.

37. 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. *European child & adolescent psychiatry* 2015; **24**(10)**:** 1279-1289.

38. 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. *Human brain mapping* 2014; **35**(10)**:** 5179-5189.

39. Lan Z, Sun Y, Zhao L, Xiao Y, Kuai C, Xue SW. Aberrant Effective Connectivity of the Ventral Putamen in Boys With Attention-Deficit/Hyperactivity Disorder. *Psychiatry Investig* 2021; **18**(8)**:** 763-769.

40. Differentiation between Resting-State fMRI data from ADHD and Normal Subjects : Based on Functional Connectivity and Machine Learning. *Proceedings of the International Conference on Fuzzy Theory and Its Applications (iFUZZY)*; 2012 Nov 16-18 2012; Natl Chung Hsing Univ, Taichung, TAIWAN2012.

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

42. Mao Z, Su Y, Xu G, Wang X, Huang Y, Yue W *et al.* Spatio-temporal deep learning method for ADHD fMRI classification. *Information Sciences* 2019; **499:** 1-11.

43. McNorgan C, Judson C, Handzlik D, Holden JG. Linking ADHD and Behavioral Assessment Through Identification of Shared Diagnostic Task-Based Functional Connections. *Frontiers in physiology* 2020; **11:** 583005.

44. Sen B, Borle NC, Greiner R, Brown MRG. A general prediction model for the detection of ADHD and Autism using structural and functional MRI. *PLoS One* 2018; **13**(4)**:** e0194856.

45. Sun Y, Zhao L, Lan Z, Jia X-Z, Xue S-W. Differentiating Boys with ADHD from Those with Typical Development Based on Whole-Brain Functional Connections Using a Machine Learning Approach. *Neuropsychiatric Disease and Treatment* 2020; **16:** 691-702.

46. Chen M, Li H, Fan H, Dillman JR, Wang H, Altaye M *et al.* ConCeptCNN: A novel multi-filter convolutional neural network for the prediction of neurodevelopmental disorders using brain connectome. *Med Phys* 2022; **49**(5)**:** 3171-3184.

47. de Lacy N, Ramshaw MJ, McCauley E, Kerr KF, Kaufman J, Nathan Kutz J. Predicting individual cases of major adolescent psychiatric conditions with artificial intelligence. *Transl Psychiatry* 2023; **13**(1)**:** 314.

48. Gaus R, Pölsterl S, Greimel E, Schulte-Körne G, Wachinger C. Can we diagnose mental disorders in children? A large-scale assessment of machine learning on structural neuroimaging of 6916 children in the adolescent brain cognitive development study. *JCPP Adv* 2023; **3**(4)**:** e12184.

49. Saha P, Sarkar D. Characterization and Classification of ADHD Subtypes: An Approach Based on the Nodal Distribution of Eigenvector Centrality and Classification Tree Model. *Child Psychiatry Hum Dev* 2022.

50. Yang CM, Shin J, Kim JI, Lim YB, Park SH, Kim BN. Classifying Children with ADHD Based on Prefrontal Functional Near-infrared Spectroscopy Using Machine Learning. *Clin Psychopharmacol Neurosci* 2023; **21**(4)**:** 693-700.

51. Chen M, Li H, Wang J, Dillman JR, Parikh NA, He L. A Multichannel Deep Neural Network Model Analyzing Multiscale Functional Brain Connectome Data for Attention Deficit Hyperactivity Disorder Detection. *Radiol Artif Intell* 2019; **2**(1)**:** e190012.

52. Chen H, Song Y, Li X. A deep learning framework for identifying children with ADHD using an EEG-based brain network. *Neurocomputing* 2019; **356:** 83-96.

53. Dubreuil-Vall L, Ruffini G, Camprodon JA. Deep Learning Convolutional Neural Networks Discriminate Adult ADHD From Healthy Individuals on the Basis of Event-Related Spectral EEG. *Frontiers in neuroscience* 2020; **14:** 251.

54. An Auditory Brainstem Response-Based Expert System for ADHD Diagnosis Using Recurrence Qualification Analysis and Wavelet Support Vector Machine. *Proceedings of the 23rd Iranian Conference on Electrical Engineering*; 2015 May 10-14 2015; Sharif Univ of Technol, Tehran, IRAN2015.

55. Jahanshahloo HR, Shamsi M, Ghasemi E, Kouhi A. Automated and ERP-Based Diagnosis of Attention-Deficit Hyperactivity Disorder in Children. *J Med Signals Sens* 2017; **7**(1)**:** 26-32.

56. Kaur S, Singh S, Arun P, Kaur D, Bajaj M. Phase Space Reconstruction of EEG Signals for Classification of ADHD and Control Adults. *Clin EEG Neurosci* 2020; **51**(2)**:** 102-113.

57. Kiiski H, Bennett M, Rueda-Delgado LM, Farina FR, Knight R, Boyle R *et al.* EEG spectral power, but not theta/beta ratio, is a neuromarker for adult ADHD. *The European journal of neuroscience* 2020; **51**(10)**:** 2095-2109.

58. Maniruzzaman M, Shin J, Hasan MAM, Yasumura A. Efficient Feature Selection and Machine Learning Based ADHD Detection Using EEG Signal. *Cmc-Computers Materials & Continua* 2022; **72**(3)**:** 5179-5195.

59. Moghaddari M, Lighvan MZ, Danishvar S. Diagnose ADHD disorder in children using convolutional neural network based on continuous mental task EEG. *Comput Methods Programs Biomed* 2020; **197:** 105738.

60. Mueller A, Vetsch S, Pershin I, Candrian G, Baschera G-M, Kropotov JD *et al.* EEG/ERP-based biomarker/neuroalgorithms in adults with ADHD: Development, reliability, and application in clinical practice. *World Journal of Biological Psychiatry* 2020; **21**(3)**:** 172-182.

61. Oztoprak H, Toycan M, Alp YK, Arikan O, Dogutepe E, Karakas S. Machine-based classification of ADHD and nonADHD participants using time/frequency features of event-related neuroelectric activity. *Clinical Neurophysiology* 2017; **128**(12)**:** 2400-2410.

62. Pereda E, García-Torres M, Melián-Batista B, Mañas S, Méndez L, González JJ. The blessing of Dimensionality: Feature Selection outperforms functional connectivity-based feature transformation to classify ADHD subjects from EEG patterns of phase synchronisation. *PLoS One* 2018; **13**(8)**:** e0201660.

63. Tenev A, Markovska-Simoska S, Kocarev L, Pop-Jordanov J, Mueller A, Candrian G. Machine learning approach for classification of ADHD adults. *International Journal of Psychophysiology* 2014; **93**(1)**:** 162-166.

64. Ahire N, Awale RN, Wagh A. Electroencephalogram (EEG) based prediction of attention deficit hyperactivity disorder (ADHD) using machine learning. *Appl Neuropsychol Adult* 2023**:** 1-12.

65. Chugh N, Aggarwal S, Balyan A. The Hybrid Deep Learning Model for Identification of Attention-Deficit/Hyperactivity Disorder Using EEG. *Clin EEG Neurosci* 2024; **55**(1)**:** 22-33.

66. Wang X-H, Jiao Y, Li L. Identifying individuals with attention deficit hyperactivity disorder based on temporal variability of dynamic functional connectivity. *Scientific reports* 2018; **8**(1)**:** 11789.

67. Zhu P, Pan J, Cai QQ, Zhang F, Peng M, Fan XL *et al.* MicroRNA profile as potential molecular signature for attention deficit hyperactivity disorder in children. *Biomarkers : biochemical indicators of exposure, response, and susceptibility to chemicals* 2022**:** 1-10.

68. 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* 2022; **140:** 105120.

69. Loh HW, Ooi CP, Oh SL, Barua PD, Tan YR, Molinari F *et al.* Deep neural network technique for automated detection of ADHD and CD using ECG signal. *Comput Methods Programs Biomed* 2023; **241:** 107775.

70. Lindhiem O, Goel M, Shaaban S, Mak KJ, Chikersal P, Feldman J *et al.* Objective Measurement of Hyperactivity Using Mobile Sensing and Machine Learning: Pilot Study. *JMIR Form Res* 2022; **6**(4)**:** e35803.

71. Abibullaev B, An J. Decision support algorithm for diagnosis of ADHD using electroencephalograms. *J Med Syst* 2012; **36**(4)**:** 2675-2688.

72. 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**(7)**:** 1259-1270.

73. Mueller A, Candrian G, Kropotov JD, Ponomarev VA, Baschera G-M. Classification of ADHD patients on the basis of independent ERP components using a machine learning system. *Nonlinear biomedical physics* 2010; **4 Suppl 1:** S1-S1.

74. Mueller A, Candrian G, Grane VA, Kropotov JD, Ponomarev VA, Baschera G-M. Discriminating between ADHD adults and controls using independent ERP components and a support vector machine: a validation study. *Nonlinear biomedical physics* 2011; **5:** 5-5.

75. 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. *Frontiers in systems neuroscience* 2012; **6:** 61.

76. Emser TS, Johnston BA, Steele JD, Kooij S, Thorell L, Christiansen H. Assessing ADHD symptoms in children and adults: evaluating the role of objective measures. *Behav Brain Funct* 2018; **14**(1)**:** 11.

77. Itani S, Rossignol M, Lecron F, Fortemps P. Towards interpretable machine learning models for diagnosis aid: A case study on attention deficit/hyperactivity disorder. *PLoS One* 2019; **14**(4)**:** e0215720.

78. Han D, Fang Y, Luo H. A Predictive Model Offor Attention Deficit Hyperactivity Disorder Based on Clinical Assessment Tools. *Neuropsychiatr Dis Treat* 2020; **16:** 1331-1337.

79. Öztekin 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.

80. Yeh S-C, Lin S-Y, Wu EH-K, Zhang K-F, Xiu X, Rizzo A *et al.* A Virtual-Reality System Integrated With Neuro-Behavior Sensing for Attention-Deficit/Hyperactivity Disorder Intelligent Assessment. *Ieee Transactions on Neural Systems and Rehabilitation Engineering* 2020; **28**(9)**:** 1899-1907.

81. Mooney MA, Neighbor C, Karalunas S, Dieckmann NF, Nikolas M, Nousen E *et al.* Prediction of Attention-Deficit/Hyperactivity Disorder Diagnosis Using Brief, Low-Cost Clinical Measures: A Competitive Model Evaluation. *Clin Psychol Sci* 2023; **11**(3)**:** 458-475.

82. Acosta-Lopez JE, Suarez I, Pineda DA, Cervantes-Henriquez ML, Martinez-Banfi ML, Lozano-Gutierrez SG *et al.* Impulsive and Omission Errors: Potential Temporal Processing Endophenotypes in ADHD. *Brain Sciences* 2021; **11**(9).

83. Cervantes-Henriquez ML, Acosta-Lopez JE, Martinez-Banfi ML, Velez JI, Mejia-Segura E, Lozano-Gutierrez SG *et al.* ADHD Endophenotypes in Caribbean Families. *Journal of attention disorders* 2020; **24**(14)**:** 2100-2114.

84. Crippa A, Salvatore C, Molteni E, Mauri M, Salandi A, Trabattoni S *et al.* The Utility of a Computerized Algorithm Based on a Multi-Domain Profile of Measures for the Diagnosis of Attention Deficit/Hyperactivity Disorder. *Front Psychiatry* 2017; **8:** 189.

85. Mwamba HM, Fourie PR, den Heever DV. PANDAS: Paediatric Attention-Deficit/Hyperactivity Disorder Application Software. *Annu Int Conf IEEE Eng Med Biol Soc* 2019; **2019:** 1444-1447.

86. Kautzky A, Vanicek T, Philippe C, Kranz GS, Wadsak W, Mitterhauser M *et al.* Machine learning classification of ADHD and HC by multimodal serotonergic data. *Transl Psychiatry* 2020; **10**(1)**:** 104.

87. 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 and Behavior* 2020; **14**(6)**:** 2132-2147.

88. Cervantes-Henríquez ML, Acosta-López JE, Martinez AF, Arcos-Burgos M, Puentes-Rozo PJ, Vélez JI. Machine Learning Prediction of ADHD Severity: Association and Linkage to ADGRL3, DRD4, and SNAP25. *J Atten Disord* 2022; **26**(4)**:** 587-605.

89. Christiansen H, Chavanon M-L, Hirsch O, Schmidt MH, Meyer C, Mueller A *et al.* Use of machine learning to classify adult ADHD and other conditions based on the Conners' Adult ADHD Rating Scales. *Scientific Reports* 2020; **10**(1).

90. Muthuraman M, Moliadze V, Boecher L, Siemann J, Freitag CM, Groppa S *et al.* Multimodal alterations of directed connectivity profiles in patients with attention-deficit/hyperactivity disorders. *Scientific reports* 2019; **9**(1)**:** 20028.

91. Brown MRG, 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. *Frontiers in systems neuroscience* 2012; **6:** 69-69.

92. Caye A, Agnew-Blais J, Arseneault L, Goncalves H, Kieling C, Langley K *et al.* A risk calculator to predict adult attention-deficit/hyperactivity disorder: generation and external validation in three birth cohorts and one clinical sample. *Epidemiology and Psychiatric Sciences* 2020; **29**.

93. Chen T, Antoniou G, Adamou M, Tachmazidis I, Su P. Automatic Diagnosis of Attention Deficit Hyperactivity Disorder Using Machine Learning. *Applied Artificial Intelligence* 2021; **35**(9)**:** 657-669.

94. Mikolas P, Vahid A, Bernardoni F, Süß M, Martini J, Beste C *et al.* Training a machine learning classifier to identify ADHD based on real-world clinical data from medical records. *Sci Rep* 2022; **12**(1)**:** 12934.

95. Altun S, Alkan A, Altun H. Automatic Diagnosis of Attention Deficit Hyperactivity Disorder with Continuous Wavelet Transform and Convolutional Neural Network. *Clin Psychopharmacol Neurosci* 2022; **20**(4)**:** 715-724.

96. 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. *Journal of attention disorders* 2020; **24**(14)**:** 2012-2020.

97. Lohani DC, Rana B. ADHD diagnosis using structural brain MRI and personal characteristic data with machine learning framework. *Psychiatry Res Neuroimaging* 2023; **334:** 111689.

98. Öztekin I, Garic D, Bayat M, Hernandez ML, Finlayson MA, Graziano PA *et al.* Structural and diffusion-weighted brain imaging predictors of attention-deficit/hyperactivity disorder and its symptomology in very young (4- to 7-year-old) children. *Eur J Neurosci* 2022; **56**(12)**:** 6239-6257.

99. 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* 2022.

100. Ehrig L, Wagner AC, Wolter H, Correll CU, Geisel O, Konigorski S. FASDetect as a machine learning-based screening app for FASD in youth with ADHD. *NPJ Digit Med* 2023; **6**(1)**:** 130.

101. Zhang Y, Sun Y, Yu Z, Sun Y, Chang X, Lu L *et al.* Risk factors and an early prediction model for persistent methamphetamine-related psychiatric symptoms. *Addict Biol* 2020; **25**(1)**:** e12709.

102. Lavigne JV, Hopkins J, Ballard RJ, Gouze KR, Ariza AJ, Martin CP. A Precision Mental Health Model for Predicting Stability of 4-year-olds' Attention Deficit/Hyperactivity Disorder Symptoms to Age 6 Diagnostic Status. *Acad Pediatr* 2023.

103. Franz AP, Caye A, Lacerda BC, Wagner F, Silveira RC, Procianoy RS *et al.* Development of a risk calculator to predict attention-deficit/hyperactivity disorder in very preterm/very low birth weight newborns. *J Child Psychol Psychiatry* 2022; **63**(8)**:** 929-938.

104. Suresh P, Ray B, Duan K, Chen J, Schoenmacker G, Franke B *et al.* Evaluating the Neuroimaging-Genetic Prediction of Symptom Changes in Individuals with ADHD. *Annu Int Conf IEEE Eng Med Biol Soc* 2021; **2021:** 1950-1956.

105. 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**(12)**:** 1370-1379.

106. Chang J-C, Lin H-Y, Lv J, Tseng W-YI, Gau SS-F. Regional brain volume predicts response to methylphenidate treatment in individuals with ADHD. *BMC psychiatry* 2021; **21**(1)**:** 26.

107. 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**(1)**:** 67.

108. 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.

109. Kim J-W, Sharma V, Ryan ND. Predicting Methylphenidate Response in ADHD Using Machine Learning Approaches. *The international journal of neuropsychopharmacology* 2015; **18**(11)**:** pyv052.

110. Morrow AS, Campos Vega AD, Zhao X, Liriano MM. Leveraging Machine Learning to Identify Predictors of Receiving Psychosocial Treatment for Attention Deficit/Hyperactivity Disorder. *Adm Policy Ment Health* 2020; **47**(5)**:** 680-692.

111. Setyawan J, Yang H, Cheng D, Cai X, Signorovitch J, Xie J *et al.* Developing a Risk Score to Guide Individualized Treatment Selection in Attention Deficit/Hyperactivity Disorder. *Value Health* 2015; **18**(6)**:** 824-831.

112. Wong HK, Tiffin PA, Chappell MJ, Nichols TE, Welsh PR, Doyle OM *et al.* Personalized Medication Response Prediction for Attention-Deficit Hyperactivity Disorder: Learning in the Model Space vs. Learning in the Data Space. *Front Physiol* 2017; **8:** 199.

113. Siontis GC, Tzoulaki I, Castaldi PJ, Ioannidis JP. External validation of new risk prediction models is infrequent and reveals worse prognostic discrimination. *J Clin Epidemiol* 2015; **68**(1)**:** 25-34.

114. Bellou V, Belbasis L, Konstantinidis AK, Tzoulaki I, Evangelou E. Prognostic models for outcome prediction in patients with chronic obstructive pulmonary disease: systematic review and critical appraisal. *BMJ* 2019; **367:** l5358.

115. Sajjadian M, Lam RW, Milev R, Rotzinger S, Frey BN, Soares CN *et al.* Machine learning in the prediction of depression treatment outcomes: a systematic review and meta-analysis. *Psychol Med* 2021; **51**(16)**:** 2742-2751.

116. Lee R, Leighton SP, Thomas L, Gkoutos GV, Wood SJ, Fenton SH *et al.* Prediction models in first-episode psychosis: systematic review and critical appraisal. *Br J Psychiatry* 2022; **220**(Spec Iss 4 Themed Iss Precision Medicine and Personalised Healthcare in Psychiatry)**:** 1-13.

117. Colombo F, Calesella F, Mazza MG, Melloni EMT, Morelli MJ, Scotti GM *et al.* Machine learning approaches for prediction of bipolar disorder based on biological, clinical and neuropsychological markers: A systematic review and meta-analysis. *Neuroscience and Biobehavioral Reviews* 2022; **135:** 104552.

118. Buitelaar J, Bölte S, Brandeis D, Caye A, Christmann N, Cortese S *et al.* Toward Precision Medicine in ADHD. *Front Behav Neurosci* 2022; **16:** 900981.

119. Wong ICK, Banaschewski T, Buitelaar J, Cortese S, Döpfner M, Simonoff E *et al.* Emerging challenges in pharmacotherapy research on attention-deficit hyperactivity disorder-outcome measures beyond symptom control and clinical trials. *Lancet Psychiatry* 2019; **6**(6)**:** 528-537.

120. Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ* 2016; **352:** i6.

121. Becker SP, Willcutt EG, Leopold DR, Fredrick JW, Smith ZR, Jacobson LA *et al.* Report of a Work Group on Sluggish Cognitive Tempo: Key Research Directions and a Consensus Change in Terminology to Cognitive Disengagement Syndrome. *J Am Acad Child Adolesc Psychiatry* 2023; **62**(6)**:** 629-645.

122. Fredrick JW, Becker SP. Cognitive Disengagement Syndrome (Sluggish Cognitive Tempo) and Social Withdrawal: Advancing a Conceptual Model to Guide Future Research. *J Atten Disord* 2023; **27**(1)**:** 38-45.

123. Froehlich TE, Becker SP, Nick TG, Brinkman WB, Stein MA, Peugh J *et al.* Sluggish Cognitive Tempo as a Possible Predictor of Methylphenidate Response in Children With ADHD: A Randomized Controlled Trial. *J Clin Psychiatry* 2018; **79**(2).

Figure 1. **Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) flowchart outlining study selection process**.

Figure 2. Most frequently reported predictors across prediction model types.

Figure 3. Risk of bias of the retrieved studies, as assessed by the PROBAST tool.