A precision health approach to medication management in neurodivergence: a model development and validation study using four international cohorts

Marlee M. Vandewouw[1,2,3], Kamran Niroomand[1], Harshit Bokadia[1], Sophia Lenz[1], Jesiqua Rapley[1], Alfredo Arias[4], Jennifer Crosbie[5,6], Elisabetta Trinari[7], Elizabeth Kelley[8], Robert Nicolson[9], Russell J. Schachar[5,6], Paul D. Arnold[10], Alana Iaboni[1], Jason P. Lerch[11,12,13], Melanie Penner[1], Danielle Baribeau[1,5,6,14,15], Evdokia Anagnostou[1,13,14,16,17], Azadeh Kushki[1,18]

1. Autism Research Center, Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital, Toronto, Canada
2. Center for Precision Psychiatry, Massachusetts General Hospital, Boston, United States of America
3. Department of Psychiatry, Harvard Medical School, Boston, United States of America
4. Province of Ontario Neurodevelopmental Network, Toronto, Canada
5. Department of Psychiatry, The Hospital for Sick Children, Toronto, Canada
6. Department of Psychiatry, University of Toronto, Toronto, Canada
7. Faculty of Health Sciences, McMaster University, Hamilton, Canada
8. Departments of Psychology and Psychiatry, Queen's University, Kingston, Canada
9. Department of Psychiatry, University of Western Ontario, London, Canada
10. Department of Psychiatry, Cumming School of Medicine, University of Calgary, Calgary, Canada
11. Wellcome Centre for Integrative Neuroimaging, FMRIB, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, United Kingdom
12. Department of Medical Biophysics, University of Toronto, Toronto, Canada
13. Program in Neurosciences & Mental Health, The Hospital for Sick Children, Toronto, Canada
14. Institute of Medical Sciences, University of Toronto, Toronto, Canada
15. Institute of Health Policy, Management, and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, Canada
16. Department of Paediatrics, University of Toronto, Toronto, Canada
17. Department of Psychiatry, Icahn School of Medicine, Mount Sinai, New York, United States of America
18. Institute of Biomedical Engineering, The University of Toronto, Toronto, Canada

Corresponding author:
Dr. Azadeh Kushki
Bloorview Research Institute
Holland Bloorview Kids Rehabilitation Hospital
150 Kilgour Road Toronto, Ontario M4G 1R8
Tel: 416-425-6220, extension 3563

Keywords: artificial intelligence, medication management, neurodivergence, neurodevelopmental conditions, autism

Summary

Background: Psychotropic medications are commonly used for neurodivergent children, but their effectiveness varies, making prescribing challenging and potentially exposing individuals to multiple medication trials. We developed artificial intelligence (AI) models to predict medication success for stimulants, anti-depressants, and anti-psychotics. We first demonstrate feasibility using cross-sectional data from three research cohorts, then use a cohort of patients from a pharmacology clinic to predict medication choice by class, longitudinally, from electronic medical records (EMRs).

Methods: Models were built to predict cross-sectional medication usage from the Child Behaviour Checklist. Data from the Province of Ontario Neurodevelopmental (POND) network (*N*=598) trained and tested the models, while data from the Healthy Brain Network (HBN; *N*=1,764) and Adolescent Brain Cognitive Development (ABCD; *N*=2,396) studies were used for external validation. For the EMR cohort, data from the Psychopharmacology Program (PPP; *N*=312) at Holland Bloorview Kids Rehabilitation Hospital were used to predict longitudinal success. Stacked ensemble models were built separately for each medication class, and area under the receiving operating characteristic curve (AU-ROC) evaluated performance.

Findings: The research cohorts demonstrated feasibility, with internal testing (POND) achieving an AU-ROC (mean [95% CI]) of 0.72 [0.71,0.74] for stimulants, 0.83 [0.80,0.85] for anti-depressants, and 0.79 [0.76,0.82] for anti-psychotics. Performance in external testing sets (HBN and ABCD) confirmed generalizability. In the EMR cohort (PPP), AU-ROC were high: 0.90 [0.88,0.91] for anti-psychotics, 0.82 [0.92,0.83] for stimulants and 0.82 [0.80,0.84] for anti-depressants.

Interpretation: This study demonstrates the feasibility of using AI to enhance medication management for neurodivergent children, with expert clinician decisions learned with high accuracy. These findings support the potential for AI decision aids in community settings, promoting faster access to personalized care while highlighting the complexity of clinical and sociodemographic factors influencing medication decisions.

Funding: Funding was provided by the Canadian Institutes of Health Research (Operating Grant #527447) and Ontario Brain Institute.

Research in context

Evidence before this study: Current medication management practices for neurodevelopmental conditions, particularly for children who have not responded to first-line options, are based clinical best guess approaches that can have negative effects on children, their caregivers, and the health system. Precision health tools using artificial intelligence, which are suitable for community use, have the potential to improve the health system's capacity for providing timely and impactful care. We searched PubMed on October 17, 2024, for studies published in English evaluating artificial intelligence approaches for medication management in neurodivergence using the terms ("autism" OR "neurodevelopmental condition" OR "neurodevelopmental disorder" OR "neurodivergence") AND ("artificial intelligence" OR "machine learning" OR "deep learning" OR "prediction") AND ("medication management" OR "medication response" OR "medication appropriateness" OR "medication decision support") also including terms where "drug" is substituted for "medication". We did not find any relevant studies.

Added value of this study: This study demonstrates the feasibility of using AI to assist in medication management for neurodivergent children, with a strong ability to learn expert clinical decisions. These findings show AI may be able to support access to faster, more personalized treatment decisions regarding psychotropic medications. We also identified the relevant clinical and demographic features to the model's medication recommendations, as well as several biases with respect to sociodemographic factors, highlighting the complexity of factors that contribute to clinical decision-making.

Implications of all the available evidence: This work highlights AI's potential to improve medication management for neurodivergent children, by offering personalized treatment recommendations. However, identified biases underscore the need for addressing existing inequities. Future research should focus on prospective validation, integration into clinical practice, and bias mitigation to ensure equitable access to effective treatments for all children.

Introduction

Neurodevelopmental conditions, such as autism spectrum disorder (autism), intellectual disability (ID), and attention-deficit/hyperactivity disorder (ADHD), are highly heterogeneous and overlapping in their biology and phenotypic presentation[1,2], which includes hundreds of associated genes[3], and significant variability in neurobiological correlates[2]. Despite decades of research, no markers exist to indicate which biological pathway may be implicated for an individual child to guide individualized care[4,5], and current approaches to care provision remain reliant on broad diagnostic labels. This lack of precision is particularly challenging for pharmacological interventions in neurodevelopmental conditions. These interventions are highly prevalent with greater than 50% of children with neurodevelopmental conditions prescribed psychotropic medications to address behavioural and/or mental health concerns (e.g., anxiety, hyperactivity, aggressive or self-injurious behavior), and high rates of polypharmacy during childhood (~30%)[6].

Despite the prevalent use of psychotropic medications in neurodevelopmental conditions, there is tremendous variability in individual-level responses to psychotropic treatment for co-occurring mental health and behavioural condition[7], reflecting the heterogeneity of biological pathways impacted, diverse clinical presentations, as well as sociodemographic complexities that influence the acceptability and uptake of medications. This variability poses a significant challenge in determining the appropriate medication for an individual child, especially in cases where co-occurring conditions (e.g., autism, ADHD, anxiety) can interact and increase the complexity of presentation and response. In these cases, prescribing practices in the children and adolescents remain almost entirely off-label[6,8–12] and based on trial-and-error approaches that can expose children to multiple medication trials with the potential for severe and life-long adverse effects for some classes of medication[13]. This highlights the critical need for precision health tools that can provide medication recommendations based on an individual's unique presentation.

To fill this gap, we propose an artificial intelligence (AI) model that uses an individual's characteristics to predict the medication choice among the commonly prescribed classes in neurodivergence: stimulants, anti-depressants, and anti-psychotics. We first establish the

84 feasibility of this approach by building models to predict medication usage using a research
85 cohort of neurodivergent children and youth, using two international research cohorts for external
86 validation. Then, to examine how this approaches generalizes to real-world clinical settings, we
87 leveraged a patient cohort to predict continuation of a medication class after 3 months from
88 electronic medical record (EMR) data of neurodivergent children and youth who visited a
89 specialized psychopharmacology clinic.

90 **Methods**

91 ***Study design and cohorts***

92 For the research cohorts, training and internal testing was conducted using data from 598
93 neurodivergent children and youth extracted from the Province of Ontario Neurodevelopmental
94 (POND) network (Ontario, Canada). External testing was performed using two international
95 cohorts: 1,764 neurodivergent children and youth from the Healthy Brain Network (HBN[14]; New
96 York, United States), and 2,396 neurodivergent children and youth from the Adolescent Brain
97 Cognitive Development study (ABCD[15]; United States).

98 For clinical validation, longitudinal EMRs from the Psychopharmacology Program (PPP) at
99 Holland Bloorview Kids Rehabilitation Hospital (Toronto, Ontario, Canada) were used. PPP sees
100 neurodivergent children who have tried at least one psychoactive medication without adequate
101 success. Data were extracted from electronic medical records of 312 children who visited the
102 PPP over a three-year period (2019 – 2022) and returned for a follow-up approximately 3 months
103 later.

104 Participant demographics are summarized in

Table 1. Further details on the cohorts, including eligibility and consent procedures, and a comparison of the research cohorts can be found in the Supplementary Material. This study was approved by the Research Ethics Board at Holland Bloorview Kids Rehabilitation Hospital.

Table 1. Cohort demographics and medication outcomes.

		Research cohorts			EMR cohort
		POND	**HBN**	**ABCD**	**PPP**
N		598	1,764	2,396	312
Age range (years)		5 – 19	5 – 17	8 – 11	4 – 17
Mean age [SD] (years)		11.4 [3.2]	10.1 [2.7]	9.9 [0.6]	10.4 [2.9]
Sex [%]	Male	424 [71]	1,213 [69]	1,523 [64]	262 [84]
	Female	174 [29]	5551 [31]	873 [36]	50 [16]
Autism [%]		338 [57]	325 [18]	0 [0]	292 [94]
ADHD [%]		193 [32]	1,381 [78]	1,751 [73]	181 [58]
OCD [%]		67 [11]	90 [5]	1,048 [44]	16 [5]
Intellectual disability [%]	Yes	84 [14]	86 [5]	149 [6]	105 [34]
	No	514 [86]	1,678 [95]	2,247 [94]	135 [43]
	Unknown	0 [0]	0 [0]	0 [0]	72 [23]
Race/ ethnicity [%]	White	327 [55]	871 [49]	1,257 [52]	NA
	Minoritized	156 [26]	866 [49]	1,137 [47]	NA
	Not collected	115 [19]	27 [2]	2 [1]	312 [100]
Annual household income [%]	Low	86 [14]	307 [17]	770 [32]	NA
	Medium	98 [16]	345 [20]	608 [25]	NA
	High	170 [28]	786 [45]	798 [33]	NA
	Not collected	244 [41]	326 [18]	220 [9]	312 [100]
Outcome [%]	Stimulant	171 [29]	177 [10]	551 [23]	114 [37]
	Anti-depressant	96 [16]	72 [4]	142 [6]	82 [26]
	Anti-psychotic	52 [9]	26 [1]	43 [2]	123 [39]

EMR: electronic medical record; PPP: Psychopharmacology Program; POND: Province of Ontario Neurodevelopmental network; HBN: Healthy Brain Network; ABCD: Adolescent Brain Cognitive Development study; SD: standard deviation; ADHD: attention-deficit/hyperactivity disorder; OCD: obsessive-compulsive disorder. For POND, primary diagnosis is reported; for ABCD, HBN, and PPP, refer to supplemental.

108 ***Predictors and outcome***

109 For the PPP cohort, the set of predictors were co-created with the clinicians at the PPP
110 program to reflect features relevant to the medication classes considered, and the availability of
111 variables in the electronic medical records at the first visit. These predictors were broadly
112 grouped into five categories: demographics, clinical, family-related, child-related, and care-
113 related. The predicted outcome was the medication decision at the follow-up visit (

Table 1).

For the research cohorts, demographics (age and sex), neurodivergent diagnoses (autism, ADHD, and OCD), IQ as a continuous score, and data from the Child Behaviour Checklist[16] were used as predictors. The predicted outcome was the medication class the child was on at the time of study (

Table 1). A full description of the predictors and outcome are presented in the Supplementary Material.

Model

AI models were built to predict the outcome from the set of predictors, separately for each medication class (Figure 1). For each model, 10 subsampling iterations were performed, each using under-sampling to balance the dataset. For each subsample, five-fold cross-validation was used to train and test the models. For the research cohorts, the models were trained and internally tested using the POND dataset, while HBN and ABCD were withheld for external testing. By validating with two external datasets, we assessed the generalizability and robustness of the models, ensuring results are not specific to a single dataset and can be replicated across diverse populations. For the EMR cohort (PPP), only internal testing was performed.

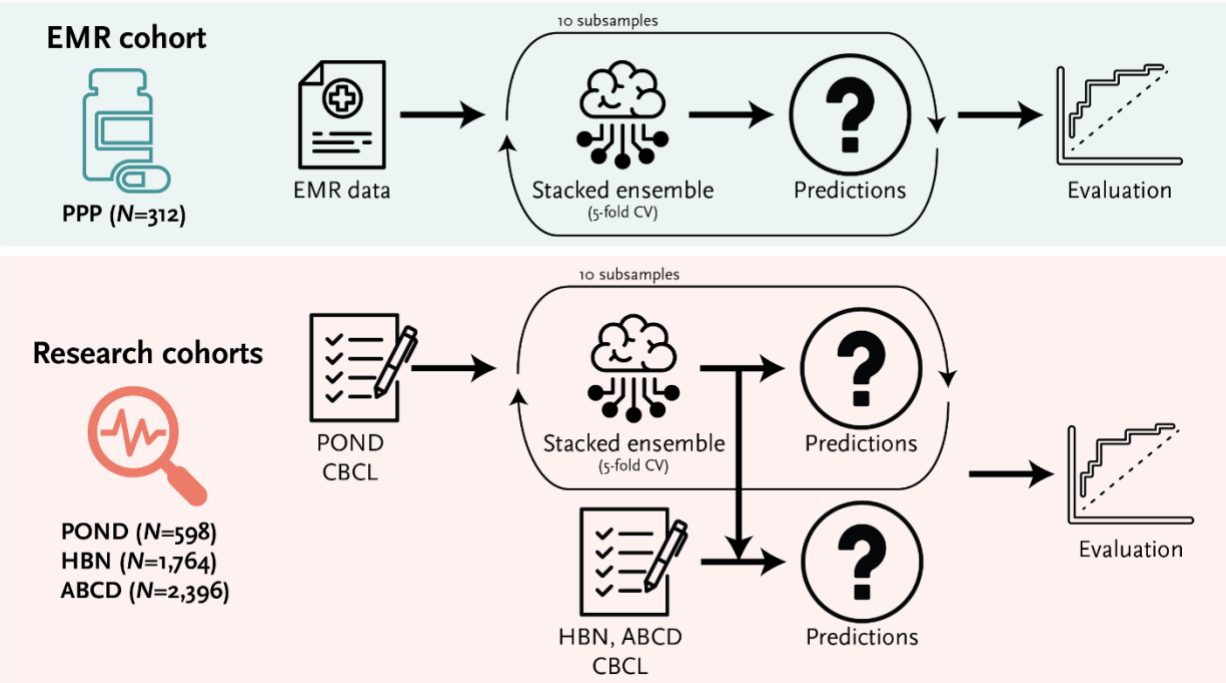

Figure 1. Study design.

For the AI models, we used a stacked ensemble model implemented in Python (version 3.10) consisting of gradient boosting over decision trees (CatBoost[17]) and a two-layer deep neural network (Keras[18] (version 3.7.0) with TensorFlow[19] (version 2.18.0)). Both models were trained

134 independently, and prediction probabilities were extracted. Then, a logistic regression model,
135 acting as a meta-learner, was used to combine the prediction probabilities from both models into
136 a final prediction, optimizing the ensemble's performance by leveraging the strengths of both
137 base models. This stacked ensemble was found to outperform other, simpler models (e.g.,
138 random forest).

139 ***Performance evaluation***

140 Model performance was assessed using the area under the receiving operating
141 characteristic curve (AU-ROC), reporting means and 95% confidence intervals (CIs) across folds
142 and bootstrap subsamples. Additional metrics (area under the precision-recall curve (AU-PR),
143 sensitivity, and specificity) are provided in the Supplementary Material.
144 To interpret which features were most important for predicting longitudinal medication success,
145 Shapley Additive exPlanations (SHAP) values[20] were computed for the EMR (PPP) models. SHAP
146 values were computed for each of the base models and combined by taking their weighted
147 average using the coefficients from the logistic regression meta-learner.
148 The demographic parity ratio[21,22] was used to evaluate whether the model predictions
149 were independent of sensitive group status for sex, intellectual disability (ID), race/ethnicity, and
150 annual household income. The metric was computed at the default probability threshold (0.50)
151 across folds, bootstrap samples, and datasets for the research cohorts; the 80% rule was used
152 to determine any biases.

153 ***Role of the funding source***

154 The funders of this study had no role in the collection, analysis, and interpretation or data,
155 in the writing of the report, nor in the decision to submit the report for publication.

156 **Results**

157 ***Model performance***

158 The research cohorts demonstrated the feasibility of AI predicting treatment status cross-
159 sectionally for each medication class (Table 2 and Figure 2), with internal testing (POND)
160 demonstrating good-to-excellent AU-ROC that was significantly higher than chance for
161 stimulants (mean [95% CI]: 0.72 [0.71, 0.74]), anti-depressants (0.83 [0.80, 0.85]), and anti-
162 psychotics (0.79 [0.76, 0.82]). The performance of the external testing sets (HBN and ABCD)

163 indicated good generalizability, with the AU-ROCs slightly reduced but still in the good-to-

164 excellent range (except for anti-psychotics for ABCD).

Table 2. Model performance.

	Research cohorts			EMR cohort
	Internal testing	External testing		Internal testing
Mean AU-ROC [95% CI]	**POND**	**HBN**	**ABCD**	**PPP**
Stimulant	.72 [.71, .74]	.71 [.70, .71]	.70 [.70, .70]	.82 [.82, .83]
Anti-depressant	.83 [.80, .85]	.74 [.74, .75]	.66 [.65, .67]	.82 [.80, .84]
Anti-psychotic	.79 [.76, .82]	.77 [.76, .78]	.81 [.80, .82]	.90 [.88, .91]

EMR: electronic medical record; PPP: Psychopharmacology Program; POND: Province of Ontario Neurodevelopmental network; HBN: Healthy Brain Network; ABCD: Adolescent Brain Cognitive Development study; CI: confidence interval; AU-ROC: area under the receiving operating characteristic curve

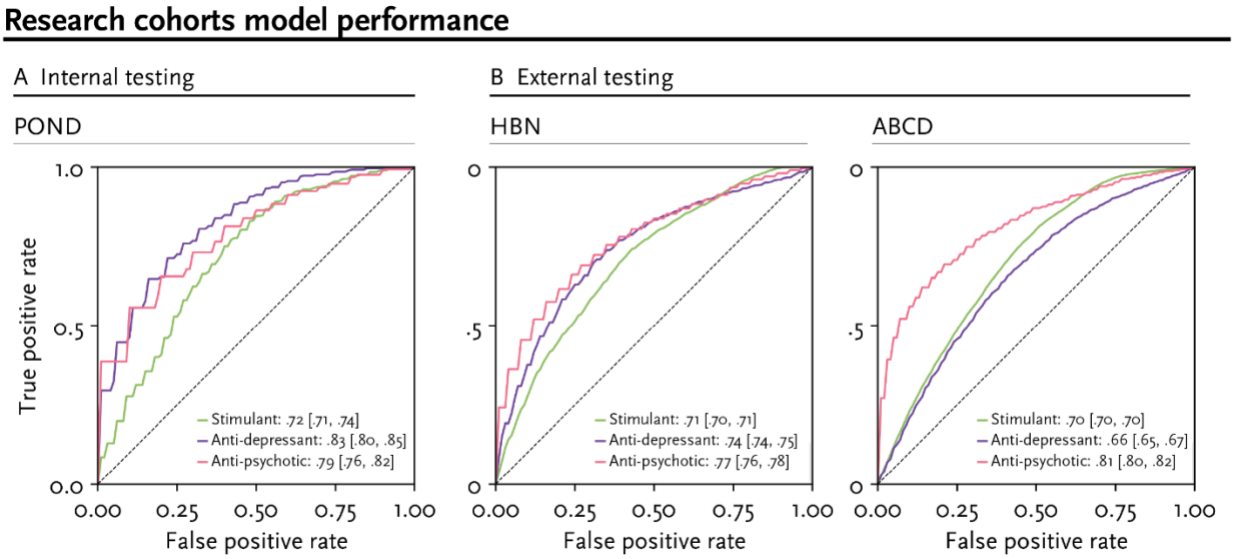

Figure 2. ROC curves for the research cohorts, for the internal testing (A; POND) and external testing (B; HBN and ABCD) sets for each of the stimulant classes (green: stimulants, purple: anti-depressants, pink: anti-psychotics); AU-ROCs are also presented.
POND: Province of Ontario Neurodevelopmental network; HBN: Healthy Brain Network; ABCD: Adolescent Brain Cognitive Development study; ROC: receiving operating characteristic curve; AU-ROC: area under the ROC

165 For the EMR cohort (PPP; Table 2 and Figure 3A), the AU-ROC across the cross-validation

166 folds was outstanding for anti-psychotics (0.90 [0.88, 0.91]), and excellent for both stimulants

167 (0.82 [0.92, 0.83]) and anti-depressants (0.82 [0.80, 0.84]).

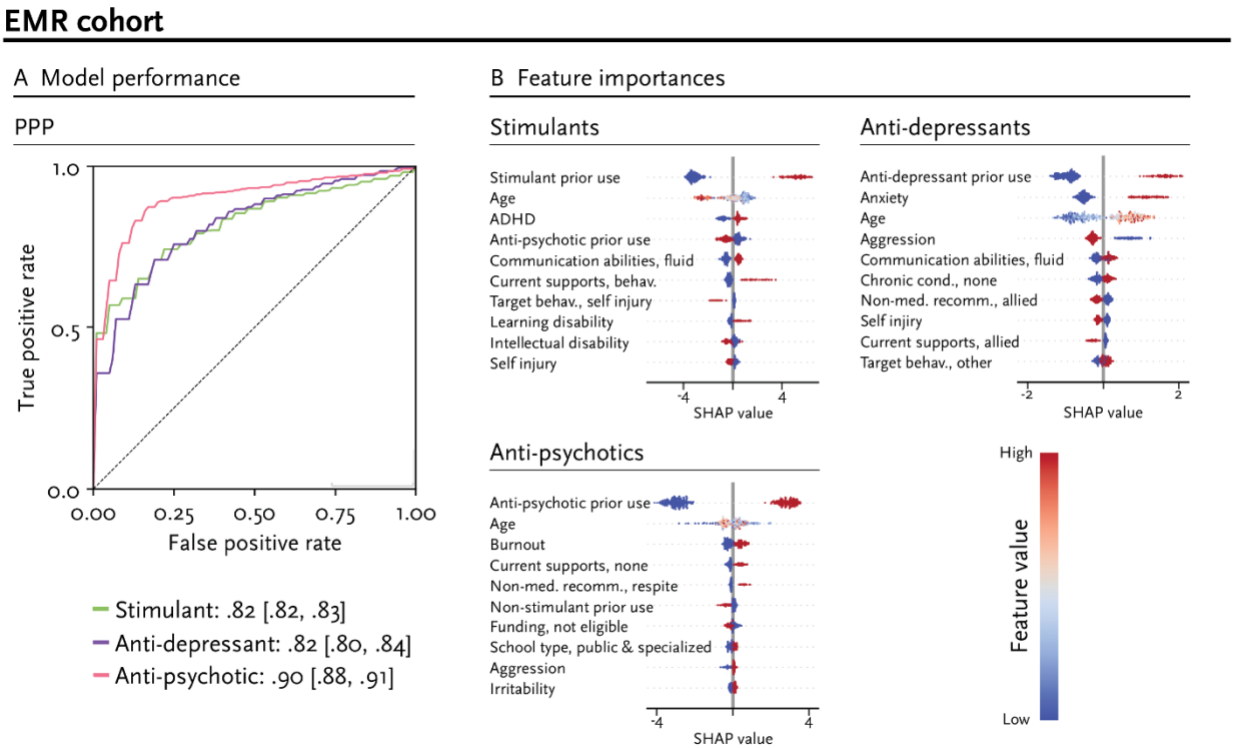

Figure 3. (A) ROC curves for the EMR cohort (PPP) for each of the stimulant classes (green: stimulants, purple: anti-depressants, pink: anti-psychotics); AU-ROCs are also presented. (B) Feature importance for the EMR cohort (PPP) assessed using SHAP values. The SHAP values are plotted for each participant for the 10 most important features for each medication class, sorted by importance (mean absolute SHAP value) from top to bottom. Colours represent the feature value (red for high values, blue for low values), indicating whether higher or lower values contribute positively or negatively to prediction scores.

EMR: electronic medical record; PPP: Psychopharmacology Program; ROC: receiving operating characteristic curve; AU-ROC: area under the ROC; SHAP: Shapley Additive exPlanations; ADHD: attention-deficit/hyperactivity disorder

Feature importance

SHAP values were computed for the EMR cohort (PPP), to identify which features were most important for predicting medication choice (Figure 3B). Prior use of a medication class was the most important feature across all four classes. In particular, prior use of a medication within the same class, increased the likelihood of a medication choice in that class. Age was also a salient feature: younger, middle, and older participants were more likely to be recommended stimulants, anti-psychotics, and anti-depressants, respectively. Other important features included ADHD and communication abilities for stimulants, anxiety and aggression for anti-depressants, and burnout (indicated by treating physician) and current supports for anti-psychotics.

178 ***Bias***

179 The demographic parity ratio was computed to evaluate the bias of the models (Table 3),
180 where values exceeding 0.80 indicate bias according to the four-fifths rule[23]. For the research
181 cohorts, parity was generally high, with some exceptions with respect to anti-depressants and
182 anti-psychotics. Individuals who were female, did not have an intellectual disability (ID), and who
183 were white were more likely to be recommended an anti-depressant (male, female: 34%, 46%;
184 ID, no-ID: 26%, 39%; white, minority: 43%, 30%). On the other hand, males, those with an ID, and
185 those with a lower annual household income were more likely to be recommended anti-
186 psychotics (male, female: 46%, 35%; ID, no-ID: 72%, 38%; low, middle, high income: 55%, 35%,
187 32%). These observations extended to the EMR cohort, where individuals without an ID were also
188 more likely to be recommended stimulants (ID, no-ID: 27%, 63%). This was in line with trends in
189 the data.

Table 3. Demographic parity ratios for sensitive groups.

		Research cohorts	EMR cohort
Sex	Stimulant	0.87	0.93
	Anti-depressant	0.74	0.77
	Anti-psychotic	0.76	0.91
Intellectual disability	Stimulant	0.79	0.44
	Anti-depressant	0.67	0.71
	Anti-psychotic	0.52	0.59
Race/ethnicity	Stimulant	0.88	–
	Anti-depressant	0.70	–
	Anti-psychotic	0.95	–
Annual household income	Stimulant	0.93	–
	Anti-depressant	0.78	–
	Anti-psychotic	0.58	–

EMR: electronic medical record

190 **Discussion**

191 Medication management for neurodivergent children remains a significant challenge,
192 particularly in community settings where subspecialty expertise may not be available. To
193 address this gap, this study aimed to develop and evaluate an AI algorithm to predict
194 medication choices in neurodivergent children. Ultimately, this type of tool can be used as

195 a decision-aid to support medication management where specialized expertise are not
196 available.

197 We showed that AI can learn expert clinician medication decisions with high accuracy. We
198 also demonstrated that the results generalize across independent, international datasets, which
199 is notable given the drastically different characteristics of these cohorts, in terms of
200 demographics, diagnosis, phenotype, and region-specific access to service and supports (see
201 Supplemental Table 1 and Supplemental Table 2). While the accuracy in these research cohorts
202 is lower compared to the EMR model, this is likely due to the cross-sectional nature of the
203 prediction in these cohorts. For example, use of stimulant medications to target ADHD
204 behaviours confounds the measures of ADHD traits when considered cross-sectionally, in turn
205 impacting predictive performance.

206 Overall, these results support the feasibility of using an AI-based decision aid tool. This can
207 be particularly advantageous for community practitioners, where subspecialty expertise required
208 for medication management is often not available. Ultimately, an AI-based decision aid can bring
209 precision to current prescribing practices in the community which may expose children to
210 multiple medication trials that may have severe and life-long adverse effects[24]. Future trials,
211 prospective trials are needed to further validate this type of algorithms in clinical settings.

212 Explainability and interpretability are critical features of responsible AI systems, contributing
213 to enhancing trust and transparency with both clinicians and patients[25]. To promote
214 explainability, we characterized the relative contributions of clinical and demographic features to
215 the AI algorithm's medication recommendations. For the EMR cohort, these contributions were
216 closely aligned with indications for each medication class (e.g., simulants for ADHD, SSRIs for
217 anxiety, antipsychotics for irritability and aggression). Interestingly, the top contributors for all
218 classes were prior use of a medication in that class. This finding reflects the fact that medication
219 choices often remain within the same class of medication but entail changes in the medication
220 type and dosing within that class (e.g., amphetamines, methylphenidate within the stimulant
221 class). Future work will focus on adding this type of resolution to the AI predictions. Other

222 contributing features to medication decision reflected sociodemographic and environmental
223 factors that may contribute to clinical decision making beyond primary indications. For example,
224 family burnout, lack of other supports, and need for respite were significant contributors to the
225 decision to recommend an antipsychotic. These finding highlight the complexity of the clinical
226 and sociodemographic factors that contribute to value-laden decisions around medication
227 management for neurodivergent children.

228 Another aspect of responsible AI is potential algorithmic biases that may disproportionately
229 disadvantage subgroups of individuals. Our analysis revealed several biases including increased
230 recommendation of antipsychotics to children who were male, had intellectual disability, and
231 were from lower income households. On the other hand, anti-depressants recommendations
232 were increased for children who were female, did not have an intellectual disability, and who
233 were white. It is important to note that these biases are not related to any specific feature of the
234 AI algorithm, but mirror previously reported biases in prescription patterns which exist across the
235 system[26]. For example, decreased access to early diagnosis and behavioural interventions for
236 children from lower income household may contribute to a higher likelihood of crisis situations
237 that necessitate prescription of antipsychotic medications. Similarly, it is well established that
238 females are more vulnerable to anxiety and depression in adolescents[27], and depression can be
239 challenging to detect in individuals with an intellectual disability[28]. While computational
240 approaches to bias mitigation can be applied to equalize prescription rates for the AI algorithm,
241 these will not solve the underlying system-level inequities that impact children with marginalized
242 identities.

243 The results of this study should be interpreted considering several limitations. Firstly, for the
244 longitudinal prediction, we defined outcome as the medication choice at the follow-up visit.
245 Treatment success in neurodivergence is a complex, subjective, and value-laden construct. The
246 use of medication choice was co-created with clinicians as an outcome as it reflects the
247 aggregate effects of the multitude of factors contributing to a clinical decision to not stop a
248 medication (e.g., ineffectiveness, side-effects, acceptability or other factors). These children also
249 visited the specialized clinic after at least one psychoactive medication was tried without

250 success, and thus this may introduce biases when the models are used to make a first-line
251 medication decision. The models were trained using data exclusively from North America; future
252 work should investigate their translation to other global populations to account for potential
253 differences in healthcare systems, prescribing practices, and sociocultural influences on
254 mediation usage. Finally, this study exclusively focused on predicting medication classes; these
255 classes are broad, and future work should examine within-class predictability.

256 In summary, the present study focused on AI model development and evaluation,
257 demonstrating the feasibility of this approach for medication recommendation for neurodivergent
258 children. To enable effective clinical implication of this type of AI-tool, future studies are needed
259 to evaluate the algorithm in a prospective design at the point of care as well as acceptability, and
260 barriers and facilitators to uptake.

261 **Declaration of interests**

262 EA has received research funding from Roche and Anavex, consulting fees from Roche, Quadrant
263 Therapeutics, Ono, and Impel Pharmaceuticals, in-kind support from AMO Pharma and CRA-
264 Simons Foundation, royalties from APPI and Springer, an editorial honorarium from Wiley, and
265 has a patent for holly™. DB has received research funding from MapLight Therapeutics. MP has
266 received consulting fee from Addis & Associates/Roche, and MedCounsel. AK has a patent for
267 holly™ with royalties paid from Awake Labs and has received consulting fees from DNAStack and
268 Shaftesbury. The remaining authors have no conflicts to report.

269 **Acknowledgements**

270 Funding was provided by the Canadian Institutes of Health Research (Operating Grant #527447).
271 Acknowledgements for each of the research cohorts (POND, HBN, and ABCD) used in this study
272 are provided below.

273 **POND:** This research was conducted with the support of the Ontario Brain Institute (POND, PIs:
274 Dr. Evdokia Anagnostou and Dr. Jason P. Lerch), an independent non-profit corporation, funded
275 partially by the Ontario government. The opinions, results and conclusions are those of the
276 authors and no endorsement by the Ontario Brain Institute is intended or should be inferred.

277 **HBN:** This manuscript was prepared using a limited access dataset obtained from the Child Mind
278 Institute Biobank, the Healthy Brain Network (HBN). This manuscript reflects the views of the
279 authors and does not necessarily reflect the opinions or views of the Child Mind Institute.

280 **ABCD:** Data used in the preparation of this article were obtained from the Adolescent Brain
281 Cognitive DevelopmentSM (ABCD) Study (https://abcdstudy.org), held in the NIMH Data Archive
282 (NDA). This is a multisite, longitudinal study designed to recruit more than 10,000 children age 9-
283 10 and follow them over 10 years into early adulthood. The ABCD Study® is supported by the
284 National Institutes of Health and additional federal partners under award numbers
285 U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018, U01DA051037,
286 U01DA050987, U01DA041174, U01DA041106, U01DA041117, U01DA041028, U01DA041134,
287 U01DA050988, U01DA051039, U01DA041156, U01DA041025, U01DA041120, U01DA051038,
288 U01DA041148, U01DA041093, U01DA041089, U24DA041123, U24DA041147. A full list of
289 supporters is available at https://abcdstudy.org/federal-partners.html. A listing of participating

sites and a complete listing of the study investigators can be found at https://abcdstudy.org/consortium_members/. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in the analysis or writing of this report. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators. The ABCD data repository grows and changes over time. The ABCD data used in this report came from DOI: 10.15154/z563-zd24. DOIs can be found at https://nda.nih.gov/abcd.

Data sharing statement

Data from four cohorts were used in this study, each governed by their own data sharing restrictions. Full details can be found in the Supplementary Material.

Supplementary material

PPP

Dataset description

The clinical cohort used for this study was longitudinal electronic medical records from the Psychopharmacology Program (PPP) at Holland Bloorview Kids Rehabilitation Hospital (Toronto, Canada). The PPP is a program with provincial catchment in Ontario that provides medication management for neurodivergent children and youth. Children were eligible for this study if (1) their first visit was between 2019-2022, (2) they were at least four years of age at their first visit, and (3) they returned to the clinic for a follow-up visit. As the data was extracted from electronic medical records, no consent was obtained. Data from PPP is not publicly available due to ethics restrictions. Clinical data from all longitudinal clinical visits were extracted into a RedCap database from semi-structured medical record templates by trained chart reviewers. Two chart reviewers co-reviewed each medical record in sets of five and reviewed discrepancies until 85% agreement on all items was achieved. Reviewers proceeded independently after that, and any uncertainties were discussed with the study team.

Predictors

For PPP, the predictors were grouped into five categories: demographics, clinical, family-related, child-related, and care-related.

Demographic variables included age, sex (male or female), and region (urban (Toronto, General Toronto Area, or other Ontario city) or rural (rural or northern community)). Due to missing data, gender and race/ethnicity could not included as predictors. To capture socioeconomic status, whether a parent was receiving social assistance or means-tested disability support as a source of income was included as a predictor.

Clinical variables included neurodevelopmental diagnoses (attention-deficit/hyperactivity disorder (ADHD), autism, obsessive-compulsive disorder (OCD), intellectual disability, learning disorder, other psychiatric condition), mental health symptoms (anxiety, depression, irritability), and chronic conditions (epilepsy, cerebral palsy, asthma, migraines, gastroesophageal reflux disease, persisting constipation, obesity, and other).

Family-related factors included a family history of psychiatric or developmental diagnoses in a first degree relative, and family reports of conflict between family members in the household, caregiver burnout, and financial barriers to care, as reported by families and documented by the treating clinician in the child's chart.

Child-related factors included school type (public, mainstream classroom, private, mainstream classroom, public, specialized classroom, private, specialized classroom, homeschool, or other), school attendance (full days, part days or withdrawal, frequent calls home for pick-up, frequent school refusal or frequent absences, not attending or demitted from school, or other), communication abilities (speaks fluidly in sentences, speaks a few words or short phrases, non-speaking and uses communication technique, non-speaking and no other formal communication form, and other). History of trauma (children's aid involvement, witnessing violence, experiencing violence, physical abuse, sexual abuse, or neglect, other significant stressor as perceived by the child, or other stressor), police involvement, hospitalization, emergency department visits, intensive crisis service involvement due to behavioural or mental health concerns, aggression, self-injury, and suicidal thoughts were also included.

Care-related factors included receiving government funding for care (specifically, the Ontario Assistance for Children with Severe Disabilities Program), which was coded as "yes", "has not applied", "not eligible", or "other". Data on current, past, and recommended supports were collected using the following categories: (1) behavioural therapy, Applied Behaviour Analysis (ABA), or parent coaching by a behaviour therapist, (2) child therapy (e.g., Cognitive Behavioural Therapy (CBT) or group therapy), (3) case management, (4) allied health, (5) respite, (6) other, or (7) none. Prior use of five commonly prescribed medication classes (stimulants, non-stimulants, hypnotics, anti-depressants, and anti-psychotics) was determined as whether a child was taking a medication at the time of the first visit to the clinic. The behaviour that was targeted for treatment at the first visit was collected using the following categories: physical aggression, swearing/verbal threats, elopement, self-injury, verbal refusal, repetitive behaviours, property destruction, tantrums/meltdowns, and/or other. Data on side effects reported at the first visit were categorized as: headache, stomach-ache, constipation, nausea/vomiting, weight gain,

357 weight loss, irritability, drowsy/fatigue/sedation, insomnia, appetite decrease, appetite increase,
358 other, and/or none. Whether adherence problems for medications were discussed at the first
359 visit was also included as a predictor.

360 *Outcome*

361 The predicted outcome was the medication class choice at the follow-up visit (typically
362 scheduled 3 months after the first visit, with nursing support in between). While treatment
363 success in neurodivergence is a complex, subjective, and value-laden construct, the decision to
364 use this as an indicator of treatment success was co-created with PPP clinicians as it reflects the
365 aggregate effects of the multitude of factors contributing to a clinical decision to not stop a
366 medication (e.g., ineffectiveness, side effects, and acceptability).

367 *Bias*

368 Bias was assessed on sex (male versus female) and intellectual disability (yes versus no).

369 ***POND***

370 *Dataset description*

371 The POND network is a research network spanning several institutions across Ontario,
372 Canada that collects demographic, phenotypic, and biological data from children and youth with
373 and without neurodevelopmental conditions. Participants were eligible for the current study if
374 they received a diagnosis of autism, ADHD, or OCD, if they were between 5 and 19 years of age,
375 and if diagnostic, medication usage, IQ, and the Child Behaviour Checklist (CBCL[16]) was
376 available. Children and youth provided written informed consent if they were able; otherwise,
377 they provided verbal assent, and their caregiver provided written consent. Data from POND is
378 managed by the Ontario Brain Institute and is available through a controlled data release. All
379 requests for access should be directed to the Brain-CODE platform
380 (https://www.braincode.ca/content/controlled-data-releases#dr006).

381 *Predictors*

382 Predictors for POND included demographics (age and sex), diagnosis (autism, ADHD, or
383 OCD), IQ, and the CBCL. Diagnoses were confirmed with expert clinical judgement alongside the
384 Autism Diagnostic Observation Schedule, 2 (ADOS-2[29]) and Autism Diagnostic Interview, Revised
385 (ADI-R[30]) for autism, the Parent Interview for Child Symptoms (PI/orCS[31]) and/or the Kiddie

Schedule for Affective Disorders and Schizophrenia (K-SADS[32]) for ADHD, and the Children's Yale-Brown Obsessive-Compulsive Scale (YBOCS[33]) for OCD. Full-scale IQ was collected using age-appropriate Wechsler[34–38] or Standford-Binet[39] scales. All the individual questions, syndromic scale *T*-scores (anxious/depressed, withdrawn/depressed, somatic complaints, social problems, thought problems, attention problems, rule-breaking behaviour, and aggressive behaviour), and summary *T*-scores (internalizing problems, externalizing problems, and total problems) from the CBCL were included as predictors.

Outcome

Caregivers were asked to provide a list of medication their child was taking. These medications were classified as stimulants, non-stimulants, anti-depressants, and anti-psychotics with assistance from clinicians. The predicted outcome was the medication class reported.

Bias

Bias was be assessed for sex, intellectual disability, race/ethnicity, and socioeconomic status. For sex, biological sex (male versus female) was used. Intellectual disability (yes versus no) was determined using the full-scale IQ cutoff of 70. For race/ethnicity, participants/caregivers were asked to select all groups that applied: Aboriginal, Arab, Black, Chinese, East Asian, Filipino, Japanese, Korean, Latin American/Hispanic, South Asian, Southeast Asian, West Asian, White, other; this data was used to categorize individuals as white or minoritized. For socioeconomic status, we used annual household income categorized as low (<$50,000), middle ($50,000-$99,999), and high (≥$100,000).

HBN

Dataset description

HBN is a mental health study of children and adolescents across the New York city (United States) area that has collected demographic, phenotypic, and biological data; families who had psychiatric concerns about their child were recruited for the study. Participants were eligible for the current study if they received a diagnosis of autism, ADHD, or OCD, if they were between 5 and 19 years of age, and if diagnostic, medication usage, IQ, and the CBCL[16] was available. Participants at least 18 years of age provided written informed consent; otherwise, they

415 provided written assent, and their caregiver provided written consent. Access to the data used in
416 this study is protected by a Data Usage Agreement (DUA; see
417 https://fcon_1000.projects.nitrc.org/indi/cmi_healthy_brain_network for more details), which
418 was awarded to AK for this study.

419 *Predictors*

420 Aligning with POND, predictors for HBN included demographics (age and sex), diagnosis
421 (autism, ADHD, and/or OCD), full-scale IQ, and the CBCL. For diagnosis, the consensus DSM-5[40]
422 diagnoses provided by the HBN clinical team were used. Full-scale IQ was assessed using
423 Wechsler scales[37,38] or the Kaufman Brief Intelligence Test[41].

424 *Outcome*

425 Caregivers were asked to provide a list of medication their child had taken the day of the
426 visit. These medications were classified as stimulants, non-stimulants, anti-depressants, and
427 anti-psychotics with assistance from clinicians. The predicted outcome was the medication
428 class(es) the child was on.

429 *Bias*

430 Bias will be assessed on sex, intellectual disability, race/ethnicity, and socioeconomic
431 status. For sex, biological sex (male versus female) was used. Intellectual disability (yes versus
432 no) was determined using the full-scale IQ cutoff of 70. For race/ethnicity,
433 participants/caregivers indicated their ethnicity (not Hispanic or Latino, Hispanic or Latino,
434 decline to specify, or unknown) and race (White/Caucasian, Black/African American, Hispanic,
435 Asian, Indian, Native American Indian, American Indian/Alaskan Native, Native Hawaiian/Other
436 Pacific Islander, two or more races, or other race), which were used to categorize individuals as
437 white or minoritized. For socioeconomic status, we used annual household income categorized
438 as low (<$50,000), middle ($50,000-$99,999), and high (≥$100,000).

439 ***ABCD***

440 *Dataset description*

441 The ABCD study[42] is a longitudinal study of adolescents from the general population
442 across the United States that has collected demographic, phenotypic, and biological data. For
443 the purposes of this study, only the first time point (collected at approximately 9-11 years of age)

444 was used. Participants were eligible for the current study if medication usage, IQ, the Kiddie
445 Schedule for Affective Disorders and Schizophrenia (K-SADS), and the CBCL was collected.
446 Caregivers provided written informed consent, and participants provided written assent. Access
447 to the data used in this study is protected by an NIMH Data Archive DUA for the Adolescent Brain
448 Cognitive Development (ABCD) permission group (see
449 https://nda.nih.gov/user/dashboard/data_permissions.html for more details), which was
450 awarded under DUA #15619 for this study.

451 *Predictors*

452 Aligning with POND, predictors for ABCD included demographics (age and sex), diagnosis
453 (autism, ADHD, and/or OCD), IQ, and the CBCL. Diagnoses were assigned using a computerized
454 version of the K-SADS[43] completed by parents and/or youth. IQ was assessed using the NIH
455 Toolbox[44] Cognition Battery's age-corrected total cognition composite score, which shows a high
456 degree of alignment with the Wechsler scales used in POND[45].

457 *Outcome*

458 Caregivers were asked to provide a list of medication their child had taken in the two
459 weeks preceding the visit. These medications were classified as stimulants, non-stimulants, anti-
460 depressants, and anti-psychotics with assistance from clinicians. The predicted outcome was
461 the probability a child was using the medication class.

462 *Bias*

463 Bias will be assessed on sex, intellectual disability, race/ethnicity, and socioeconomic
464 status. For sex, biological sex (male versus female) was used. Intellectual disability (yes versus
465 no) was determined using the IQ cutoff of 70. For race/ethnicity, participants/caregivers indicated
466 their race/ethnicity (Hispanic/Latino, non-Hispanic Asian, non-Hispanic Black, non-Hispanic
467 White, and Multi-Racial/Other), which were used to categorize individuals as white or
468 minoritized. For socioeconomic status, we used annual household income categorized as low
469 (<$50,000), middle ($50,000-$99,999), and high (≥$100,000).

470 ***Comparison of research cohorts***

471 The demographic variables (age, sex, diagnosis, intellectual disability, race/ethnicity, and
472 socioeconomic status) were compared amongst the research cohorts (Supplemental Table 1).

Chi-squared tests were used for categorical variables, holding significance at $p<.050$ and reporting Cramer's *v* effect sizes. ANOVAs or Kruskal-Wallis tests were used for continuous variables, depending on normality, holding significance at $p<.050$, and reporting eta-squared effect sizes.

Supplemental Table 1. Demographic comparison of the research cohorts.

Measure	Test-statistic	*p*-value	Effect size	Post-hocs
Age	95.07	<.001	.05	ABCD > POND > HBN
Sex	18.49	<.001	.06	Females: ABCD > POND, HBN
Autism	1,319.21	<.001	.53	POND > HBN > ABCD
ADHD	472.72	<.001	.32	HBN > ABCD > POND
OCD	871.51	<.001	.43	ABCD > POND > HBN
Intellectual disability	61.88	<.001	.11	POND > HBN, ABCD
Race/ethnicity	47.78	<.001	.10	White: POND > HBN, ABCD
Annual household income	131.50	<.001	.13	HBN > POND > ABCD

POND: Province of Ontario Neurodevelopmental network; HBN: Healthy Brain Network; ABCD: Adolescent Brain Cognitive Development study; ADHD: attention-deficit/hyperactivity disorder; OCD: obsessive-compulsive disorder

The CBCL summary *T*-scores were also compared amongst the datasets (Supplemental Table 3), using ANOVAs or Kruskal-Wallis tests, depending on normality, holding significance at $p<.050$, and reporting eta-squared effect sizes.

Supplemental Table 2. CBCL comparison of the research cohorts.

	Median [IQR]			Test statistic	*p*-value	Effect size	Post-hocs
Measure	ABCD	HBN	POND				
Anxious/depressed	54 [50, 63]	55 [50, 64]	62 [53, 68]	120.64	<.001	.03	POND > HBN, ABCD
Withdrawn/depressed	54 [60, 52]	56 [50, 64]	60 [54, 68]	182.52	<.001	.04	POND > HBN > ABCD
Somatic complaints	57 [50, 64]	56 [50, 61]	57 [53, 64]	15.84	<.001	<.01	ABCD > POND, POND > HBN
Social problems	54 [51, 60]	58 [51, 64]	61 [54, 68]	244.41	<.001	.06	POND > HBN > ABCD
Thought problems	58 [51, 64]	58 [51, 67]	66 [58, 71]	216.91	<.001	.05	POND > HBN, ABCD
Attention problems	59 [53, 60]	64 [57, 69]	65 [59, 71]	264.73	<.001	.05	POND > HBN > ABCD
Rule-Breaking behaviour	53 [50, 60]	53 [51, 60]	54 [51, 62]	14.36	.001	<.01	ABCD > HBN, POND > ABCD
Aggressive behaviour	54 [50, 62]	57 [50, 66]	58 [51, 66]	73.53	<.001	.02	POND > HBN > ABCD
Internalizing problems	56 [48, 64]	57 [50, 65]	63 [55, 69]	143.27	<.001	.03	POND > HBN > ABCD
Externalizing problems	54 [46, 62]	56 [48, 65]	58 [50, 66]	53.60	<.001	.01	POND > HBN > ABCD
Total problems	56 [50, 64]	60 [52, 68]	64 [56, 70]	204.16	<.001	.04	POND > HBN > ABCD

CBCL: Child Behaviour Checklist; POND: Province of Ontario Neurodevelopmental network; HBN: Healthy Brain Network; ABCD: Adolescent Brain Cognitive Development study

Performance evaluation

Additional metrics (area under the precision-recall curve (AU-PR), sensitivity, and specificity) are provided in the Supplementary Material.

Supplemental Table 3. Model performance.

	Research cohorts		Clinical cohort
	Internal testing	External testing	Internal testing

	POND	HBN	ABCD	PPP
Mean AU-PR [95% CI]				
Stimulant	.70 [.67, .72]	.68 [.68, .69]	.65 [.65, .66]	.86 [.86, .87]
Anti-depressant	.82 [.79, .84]	.75 [.74, .76]	.66 [.65, .67]	.84 [.82, .85]
Anti-psychotic	.81 [.78, .84]	.79 [.77, .81]	.81 [.80, .82]	.90 [.89, .92]
Mean sensitivity [95% CI]				
Stimulant	.72 [.70, .73]	.82 [.80, .83]	.77 [.76, .79]	.70 [.68, .72]
Anti-depressant	.73 [.70, .76]	.51 [.49, .52]	.44 [.41, .46]	.69 [.66, .72]
Anti-psychotic	.71 [.66, .76]	.60 [.57, .64]	.58 [.54, .61]	.83 [.68, .72]
Mean specificity [95% CI]				
Stimulant	.63 [.61, .65]	.46 [.44, .47]	.53 [.51, .54]	.78 [.77, .80]
Anti-depressant	.76 [.73, .79]	.83 [.82, .85]	.76 [.73, .78]	.80 [.78, .83]
Anti-psychotic	.71 [.66, .76]	.80 [.77, .84]	.86 [.84, .89]	.90 [.89, .92]

PPP: Psychopharmacology Program; POND: Province of Ontario Neurodevelopmental network; HBN: Healthy Brain Network; ABCD: Adolescent Brain Cognitive Development study; CI: confidence interval; AU-ROC: area under the receiving operating characteristic curve

References

1 Astle DE, Holmes J, Kievit R, Gathercole SE. Annual Research Review: The transdiagnostic revolution in neurodevelopmental disorders. *Journal of Child Psychology and Psychiatry* 2021; **n/a**. DOI:10.1111/jcpp.13481.

2 Vandewouw MM, Brian J, Crosbie J, *et al.* Identifying Replicable Subgroups in Neurodevelopmental Conditions Using Resting-State Functional Magnetic Resonance Imaging Data. *JAMA Network Open* 2023; **6**: e232066.

3 Trost B, Thiruvahindrapuram B, Chan AJS, *et al.* Genomic architecture of Autism Spectrum Disorder from comprehensive whole-genome sequence annotation. Genetic and Genomic Medicine, 2022 DOI:10.1101/2022.05.05.22274031.

4 Persico AM, Ricciardello A, Lamberti M, *et al.* The pediatric psychopharmacology of autism spectrum disorder: A systematic review - Part I: The past and the present. *Prog Neuropsychopharmacol Biol Psychiatry* 2021; **110**: 110326.

5 Lord C, Charman T, Havdahl A, *et al.* The Lancet Commission on the future of care and clinical research in autism. *Lancet* 2022; **399**: 271–334.

6 Feroe AG, Uppal N, Gutiérrez-Sacristán A, *et al.* Medication Use in the Management of Comorbidities Among Individuals With Autism Spectrum Disorder From a Large Nationwide Insurance Database. *JAMA Pediatrics* 2021; **175**: 957–65.

7 Baribeau D, Vorstman J, Anagnostou E. Novel treatments in autism spectrum disorder. *Curr Opin Psychiatry* 2022; **35**: 101–10.

8 Murray ML, Hsia Y, Glaser K, *et al.* Pharmacological treatments prescribed to people with autism spectrum disorder (ASD) in primary health care. *Psychopharmacology (Berl)* 2014; **231**: 1011–21.

9 Lake JK, Weiss JA, Dergal J, Lunsky Y. Child, parent, and service predictors of psychotropic polypharmacy among adolescents and young adults with an autism spectrum disorder. *J Child Adolesc Psychopharmacol* 2014; **24**: 486–93.

10 Spencer D, Marshall J, Post B, *et al.* Psychotropic medication use and polypharmacy in children with autism spectrum disorders. *Pediatrics* 2013; **132**: 833–40.

11 Mellahn OJ, Knott R, Tiego J, *et al.* Understanding the Diversity of Pharmacotherapeutic Management of ADHD With Co-occurring Autism: An Australian Cross-Sectional Survey. *Front Psychiatry* 2022; **13**: 914668.

12 Ritter C, Hewitt K, McMorris CA. Psychotropic Polypharmacy Among Children and Youth with Autism: A Systematic Review. *J Child Adolesc Psychopharmacol* 2021; **31**: 244–58.

13 Howes OD, Rogdaki M, Findon JL, *et al.* Autism spectrum disorder: Consensus guidelines on assessment, treatment and research from the British Association for Psychopharmacology. *J Psychopharmacol* 2018; **32**: 3–29.

14 Alexander LM, Escalera J, Ai L, *et al.* An open resource for transdiagnostic research in pediatric mental health and learning disorders. *Sci Data* 2017; **4**: 170181.

15 Barch DM, Albaugh MD, Avenevoli S, *et al.* Demographic, physical and mental health assessments in the adolescent brain and cognitive development study: Rationale and description. *Dev Cogn Neurosci* 2018; **32**: 55–66.

16 Achenbach TM, Rescorla LA. Manual for the ASEBA school-age forms and profiles. Burlington, VT: University of Vermont Research Center for Children, Youth, and Families, 2001.

17 Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A. CatBoost: unbiased boosting with categorical features. In: Proceedings of the 32nd International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc., 2018: 6639–49.

18 Chollet F, others. Keras. 2015. https://keras.io.

19 Martín Abadi, Ashish Agarwal, Paul Barham, *et al.* TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. 2015. https://www.tensorflow.org/.

20 Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc., 2017: 4768–77.

21 Bird S, Dudík M, Edgar R, *et al.* Fairlearn: A toolkit for assessing and improving fairness in AI. Microsoft, 2020 https://www.microsoft.com/en-us/research/publication/fairlearn-a-toolkit-for-assessing-and-improving-fairness-in-ai/.

22 Agarwal A, Beygelzimer A, Dudik M, Langford J, Wallach H. A Reductions Approach to Fair Classification. In: Dy J, Krause A, eds. Proceedings of the 35th International Conference on Machine Learning. PMLR, 2018: 60–9.

23 Holzer H, Neumark D. Assessing Affirmative Action. *Journal of Economic Literature* 2000; **38**: 483–568.

24 Howes OD, Rogdaki M, Findon JL, *et al.* Autism spectrum disorder: Consensus guidelines on assessment, treatment and research from the British Association for Psychopharmacology. *J Psychopharmacol* 2018; **32**: 3–29.

25 Cutillo CM, Sharma KR, Foschini L, *et al.* Machine intelligence in healthcare-perspectives on trustworthiness, explainability, usability, and transparency. *NPJ Digit Med* 2020; **3**: 47.

26 Antoniou T, McCormack D, Kitchen S, *et al.* Geographic variation and sociodemographic correlates of prescription psychotropic drug use among children and youth in Ontario, Canada: a population-based study. *BMC Public Health* 2023; **23**: 85.

27 Shorey S, Ng ED, Wong CHJ. Global prevalence of depression and elevated depressive symptoms among adolescents: A systematic review and meta-analysis. *Br J Clin Psychol* 2022; **61**: 287–305.

28 Buckley N, Glasson EJ, Chen W, *et al.* Prevalence estimates of mental health problems in children and adolescents with intellectual disability: A systematic review and meta-analysis. *Aust N Z J Psychiatry* 2020; **54**: 970–84.

29 Lord C, Rutter M, DiLavore, P, Risi S, Gotham K, Bishop S. Autism diagnostic observation schedule: ADOS-2, 2nd edn. Los Angeles, CA: Western Psychological Services, 2012.

30 Rutter M, Le Couteur A, Lord C. Autism diagnostic interview revised. Los Angeles, CA: Western Psychological Services, 2003.

31 Ickowicz A, Schachar RJ, Sugarman R, Chen SX, Millette C, Cook L. The parent interview for child symptoms: a situation-specific clinical research interview for attention-deficit hyperactivity and related disorders. *Can J Psychiatry* 2006; **51**: 325–8.

32 Kaufman J, Birmaher B, Brent D, *et al.* Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): Initial reliability and validity data. *J Am Acad Child Adolesc Psychiatry* 1997; **36**: 980–8.

33 Scahill L, Riddle MA, McSwiggin-Hardin M, *et al.* Children's Yale-Brown Obsessive Compulsive Scale: Reliability and validity. *J Am Acad Child Adolesc Psychiatry* 1997; **36**: 844–52.

34 Wechsler D. Wechsler abbreviated scale of intelligence, 1st edn. San Antonio, TX: The Psychological Corporation, 1999.

35 Wechsler D. Wechsler preschool and primary scales of intelligence, 3rd edn. San Antonio, TX: The Psychological Corporation, 2002.

36 Wechsler D. Wechsler intelligence scale for children, 4th edn. San Antonio, TX: The Psychological Corporation, 2003.

37 Wechsler D. Wechsler abbreviated scale of intelligence, 2nd edn. San Antonio, TX: NCS Pearson, 2011.

38 Wechsler D. Wechsler intelligence scale for children, 5th edn. Bloomington, MN: Pearson, 2014.

39 Roid GH. Stanford-Binet intelligence scales, 5th edn. Itasca, IL: Riverside Publishing, 2003.

40 Diagnostic and statistical manual of mental disorders, 5th edn. Arlington, TX: American Psychiatric Association, 2013.

41 Kaufman AS, Kaufman NL. Kaufman Brief Intelligence Test, 2nd edn. Pearson.

42 Volkow ND, Koob GF, Croyle RT, *et al.* The conception of the ABCD study: From substance use to a broad NIH collaboration. *Dev Cogn Neurosci* 2018; **32**: 4–7.

43 Townsend L, Kobak K, Kearney C, *et al.* Development of Three Web-Based Computerized Versions of the Kiddie Schedule for Affective Disorders and Schizophrenia Child Psychiatric Diagnostic Interview: Preliminary Validity Data. *J Am Acad Child Adolesc Psychiatry* 2020; **59**: 309–25.

44 Weintraub S, Dikmen SS, Heaton RK, *et al.* Cognition assessment using the NIH Toolbox. *Neurology* 2013; **80**: S54-64.

45 Heaton RK, Akshoomoff N, Tulsky D, *et al.* Reliability and validity of composite scores from the NIH Toolbox Cognition Battery in adults. *J Int Neuropsychol Soc* 2014; **20**: 588–98.