medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Predicting Mental and Neurological Illnesses Based on Cerebellar Normative Features

Milin Kim[1,2], Nitin Sharma[8,9], Esten H. Leonardsen[1,2], Saige Rutherford [5,6,7], Geir Selbæk[10,11,22], Karin Persson[10,11], Nils Eiel Steen[1,12,13], Olav B. Smeland[1], Torill Ueland[2,12], Geneviève Richard[1], Aikaterina Manoli[16], Sofie L. Valk[16, 17,18], Dag Alnæs[1,2], Christian F. Beckman[6,21], Andre F. Marquand[6,20], Ole A. Andreassen[1,3], Lars T. Westlye[1,2,3], Thomas Wolfers[1,8,9 **] & Torgeir Moberget[1,4,19 **].

Affiliations:

[1] Centre for Precision Psychiatry, Division of Mental Health and Addiction, University of Oslo and Oslo University Hospital, Oslo, Norway; [2] Department of Psychology, University of Oslo, Oslo, Norway; [3] KG Jebsen Centre for Neurodevelopmental Disorders, University of Oslo, Oslo, Norway; [4]Department of Behavioral Science, School of Health Sciences, Oslo Metropolitan University - OsloMet, Oslo, Norway; [5] Department of Cognitive Neuroscience, Radboud University Medical Centre, Nijmegen, Netherlands; [6]Donders Institute, Radboud University, Nijmegen, Netherlands; [7]Department of Psychiatry, University of Michigan, Ann Arbor, MI, United States; [8]Department of Psychiatry and Psychotherapy, University of Tübingen, Tübingen, Germany; [9]German Center for Mental Health (DZPG), Tübingen, Germany; [10]Department of Geriatric Medicine, Oslo University Hospital, Oslo, Norway; [11]The Norwegian National Centre for Ageing and Health, Vestfold Hospital Trust, Tønsberg, Norway; [12]Section for Clinical Psychosis Research, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway; [13]Division of Mental Health and Substance Abuse, Diakonhjemmet Hospital, Oslo, Norway; [14]Department of Neuroimaging, Centre of Neuroimaging Sciences, Institute of Psychiatry, King's College London, London, UK; [15]Centre for Functional MRI of the Brain, Nuffield Department of Clinical Neurosciences, Wellcome Centre for Integrative Neuroimaging, University of Oxford, Oxford, UK, [16]Otto Hahn Research

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Group for Cognitive Neurogenetics, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany; [17]Institute of Neurosciences and Medicine (INM-7), Research Centre Jülich, Jülich, Germany; [18]Institute of Systems Neuroscience, Heinrich Heine University Düsseldorf, Düsseldorf, Germany; [19]Department of Psychology, Pedagogy and Law, School of Health Sciences, Kristiania University College, Oslo, Norway, [20]Department of Neuroimaging, Centre of Neuroimaging Sciences, Institute of Psychiatry, King's College London, London, UK, [21]Centre for Functional MRI of the Brain, Nuffield Department of Clinical Neurosciences, Wellcome Centre for Integrative Neuroimaging, University of Oxford, Oxford, UK; [22]Institute for clinical medicine, University of Oslo, Oslo, Norway

**These authors contributed equally.

Corresponding authors

Milin Kim, Centre for Precision Psychiatry, University of Oslo and Oslo University Hospital, Oslo, Norway, E-mail:

milink@student.sv.uio.no

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract

Mental and neurological conditions have been linked to structural brain variations. However, aside from dementia, the value of brain structural characteristics derived from brain scans for prediction is relatively low. One reason for this limitation is the clinical and biological heterogeneity inherent to such conditions. Recent studies have implicated aberrations in the cerebellum – a relatively understudied brain region – in these clinical conditions. Here, we used machine learning to test the value of individual deviations from normative cerebellar development across the lifespan (based on trained data from >27k participants) for prediction of autism spectrum disorder (ASD) (n=317), bipolar disorder (BD) (n=238), schizophrenia (SZ) (n=195), mild cognitive impairment (MCI) (n=122), and Alzheimer's disease (AD) (n=116). We applied several atlases and derived median, variance, and percentages of extreme deviations within each region of interest. Our results show that lobular and voxel-wise cerebellar data can be used to discriminate healthy controls from ASD and SZ with moderate accuracy (the area under the receiver operating characteristic curves ranged from 0.56 to 0.64), The strongest contributions to these predictive models were from posterior regions of the cerebellum, which are more strongly linked to higher cognitive functions than to motor control.

Keywords: Cerebellum, Normative modelling, Magnetic Resonance Imaging, Mental Illnesses, & Neurological Diseases, Machine Learning

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Introduction

Clinical heterogeneity and complex pathobiological mechanisms impede the discovery of reliable biomarkers for many neurological and – especially – psychiatric disorders, thereby complicating precise clinical decision-making and treatments. Over the last two decades, there has been a trend in the development of neuroimaging-based tools and machine learning for prognosis and diagnosis of psychiatric disorders (1,2) and neurological illnesses (3). Neuroimaging-based prediction studies on autism spectrum disorder (ASD), bipolar disorder (BD), and schizophrenia (SZ) have reported a wide range of accuracies, underscoring the limitations associated with small samples, including poor generalization performance (4,5). Of note, prediction studies on dementias show greater promise for clinical usage in both Alzheimer's disease (AD) (3) and mild cognitive impairment (MCI).

Notably, the majority of these prediction studies (4–6) have focused on cerebral features, perhaps reflecting a "cortico-centric bias" in the literature (7). Nonetheless, disruptions in the cerebellum have been hypothesized to contribute to various clinical conditions, such as childhood psychiatric symptoms (8), AD (9), SZ (10), and ASD (11–13). Indeed, patient studies have shown that abnormalities in cerebellum can exert a significant influence on motor, cognitive, and emotional functions (14–16), yet, there is little exploration on the role of the cerebellum in predicting and classifying mental and neurological illnesses. Using a normative modelling approach, we recently demonstrated significant deviations from normal cerebellar developmental across the lifespan in ASD, MCI, AD, BD, and SZ (17). While these individual-level deviations revealed substantial cerebellar heterogeneity among individuals with the same disorder, the value of these cerebellar features with respect to classifying these disorders remain uncertain.

In this study, we addressed this gap by performing a set of predictions of ASD, MCI, AD, BD, and SZ, using MRI-based cerebellar features and cross-validated machine learning

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

classifiers. We applied lobular and voxel-wise normative models (17) and aggregated the median, variance and percentage of extreme deviations across atlases (18,19). Finally, for models that were able to meaningfully differentiate between patients and health controls, we identified cerebellar regions that contributed the most to the prediction.

Methods

Sample

The study consisted of healthy controls from the test set the cerebellar lifespan normative model (17) (n=26.985, 53% females), and the clinical samples (n=1.757, 30% females) (Figure 1A and Supplementary Table 2). Individuals without diagnoses were matched to the clinical datasets of patients with AD, ASD, BD, MCI, and SZ (Table 1) using nearest neighbor matching based on exact matches of sex and scanning site with age as implemented in *MatchIt* (20). The clinical datasets were obtained from the ABIDE, ADNI, AIBL**,** DEMGEN, and TOP cohorts. Information about each cohort and studies can be found in the corresponding publications (Supplementary Table 1). If participants were scanned at several timepoints, only baseline scans were chosen for this study. Individuals who withdrew from the studies or lacked essential demographic information and T1-weighted MRI data were excluded from the analyses.

Lobular-level processing

The T1-weighted images were skull-stripped using the FreeSurfer 5.3 auto-recon pipeline (21) and reoriented to the standard FSL orientation using the *fslreorient2std* (22). Linear registration was performed using *flirt* (23), which employed linear interpolation (with six degrees of freedom) and the default 1 mm FSL template (version 6.0). The borders were cropped at coordinates [6:173, 2:214, 0:160] to minimize their size without removing brain tissue. Finally,

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

126 the voxel intensity values of all brain images were normalized to the range of [0,1], adjusting
127 the intensity values of each voxel to a standardized scale.

128 To segment the cerebellum, we utilized the ACAPULCO algorithm (24), part of
129 ENIGMA Cerebellum Volumetric Pipeline, which is a cerebellum parcellation algorithm based
130 on convolutional neural networks. This algorithm delivers fast and precise quantitative in-vivo
131 regional assessment of the cerebellum. As part of the algorithm, the images were corrected for
132 inhomogeneity by N4 correction method (25) and registered to the 1mm isotropic ICBM 2009c
133 template in MNI space using the ANTs registration suite (26). The ACAPULCO algorithm is
134 based on 15 expert manual delineations of an adult cohort (27). It achieves per-voxel labelling
135 and employs post-processing of the parcellation to correct for mislabeling and for accurate
136 segmentation. ACAPULCO segments the cerebellum into 28 cerebellar lobules and computes
137 the volume (mm^3) for each region. These regions include bilateral Lobules I–VI; Crus I and II;
138 Lobules VIIB, VIIIA, VIIIB, and IX-X; Vermis VI, VII, VIII, IX, and X; and Corpus Medullare
139 (CM). To ensure data quality, participants with extreme outliers (2.698 s.d. above or below the
140 mean) (28) in more than two lobules based on automated quality control measures, were
141 excluded. We set the threshold at two lobules because the differences between one and two
142 lobules was not significant (see Supplementary Methods for detailed information for quality
143 control).
144
145 *Voxel-level processing*
146 We used SUIT version 3.4 (Spatially Unbiased Infratentorial Toolbox) (29) to segment
147 cerebellar grey and white matter voxel-based morphometry (VBM) maps. SUIT leverages the
148 outputs from ACAPULCO, an MNI-aligned T1 image (29,30), and an average mask derived
149 from a randomly selected group of 300 individuals without a diagnosis. After segmentation,
150 the grey matter maps were normalized for standardized comparison and re-sliced to align them

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

with the MNI152 template. Additionally, the grey matter maps were modulated by the Jacobian to preserve the value of each voxel in proportion to its original volume. This Jacobian modulation ensured that the values of the original volume were proportionally maintained.

Normative modelling

We used a publicly available cerebellar normative model, estimated using >27 participants (17) which is implemented in the PCNtoolkit package (version 0.24) (31,32). This normative model encompasses cerebellar volumes and voxel-wise intensity while including sex, age, and scanning-site as covariates (Figure 1B).

To analyze the data, we employed Bayesian Linear Regression (BLR) with the likelihood warping method (33), incorporating the 'sinarcsinsh' transformation (34,35), to handle non-linear basis functions and non-Gaussian predictive distributions for large datasets (34). Scanning-site was accounted for as a fixed effect (36,37). The normative model provides point estimates and evaluation metrics such as explained variance, mean squared log-loss, skew, and kurtosis (35). These evaluation metrics were calculated in the test set, which did not include clinical cohorts. Extreme deviations were defined as $|z| > 1.96$, corresponding to the most extreme 5% of cases in both directions in the reference cohort.

Feature engineering

Voxel-wise normative models were utilized to map deviation profiles onto existing atlases (see Supplementary Methods). Three existing atlases were selected: 28 cerebellar anatomical regions, 10 regions of interest (ROI) from the multi-domain task battery (MDTB) (19), and 17 ROI from resting-state connectivity (18,38). For each region of interest delineated by these atlases, we computed three key statistics: the median, variance and percentage of extreme deviations (Figure 1C). To quantify the extremes in deviation, we also calculated the

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

proportion of voxel-wise deviations that exceeded the established threshold of $|z| > 1.96$, denoting both extreme positive and negative deviations. This proportion was determined by dividing the count of such extreme deviations by the total voxel count within the corresponding region of interest. Variance has previously been used to examine the structural heterogeneity among patients in SZ (39,40). Unlike percentage of extreme deviation ($|z| > 1.96$) that has been used in past normative studies (41–43), variance assesses the dispersion within the region, capturing the regionally heterogeneous spread within patients.

Model training and evaluation

Machine learning models employing logistic regression (LR) were used to build prediction models (Figure 1D). In addition, results from the random forest (RF) algorithm from the scikit-learn library version 1.2.2 (44) and the eXtreme Gradient Boosting (XGBoost) library version 1.7.3 (45) can be found in the Supplementary Figure 3. RF is a non-parametric supervised learning method that addresses over-fitting by combining decision trees into a single outcome, effectively balancing the bias-variance trade-off. XGBoost is an open source library to implement advanced gradient boosting algorithms (45).

The features engineered from three atlases separately ran as inputs for the algorithm. We developed various machine learning models using deviations from the normative models and utilized their median, variance and percentage of extreme deviation onto the existing atlases as features. To evaluate the model's performance in held-out test data, we conducted a stratified five-fold cross-validation and used the area under the receiver operating characteristic curve (AUROC) as the primary performance metric. Additionally, we calculated precision, recall, sensitivity, specificity, balanced accuracy, and the area under the precision-recall curve.

Permutation testing

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

We used permutation testing to assess whether the AUROCs achieved by our model was different from chance level performance. To achieve this, we shuffled the diagnosis labels randomly 1000 times, for each permutation calculating an AUROC. For significance testing, the original AUROC was compared to the distribution of permuted AUROC values. If the original AUROC falls within the extreme ends of the permutation distribution ($p < 0.05$), it is considered statistically significant. We applied an identical approach for the lobular volume features. The comparison between models utilized an approach similar to that outlined in Supplementary Figure 2-3, wherein the previously calculated shuffled AUROC values were used. We calculated the difference in true AUROC scores, as well as the AUROC differences from 1000 permuted datasets, between the two models. Subsequently, we compared the true score and the permuted scores to assess statistical significance.

Feature importance ranking

We assessed feature importance based on logistic regression coefficients to highlight their influence on the predictions. The coefficients from the model directly infer the relative importance of each feature, thus facilitating interpretation. The magnitude of the coefficient indicates the strength of the effect a feature has on the prediction, while the sign (positive or negative) indicates the direction of the effect. Figure 3 illustrates the summary plot of the standardized feature importance, emphasizing the key features that have the greatest influence (see Supplementary Figure 1 for all feature importance).

Results

We conducted a comprehensive analysis at the lobular and voxel-wise level employing a variety of models (Figure 1C). The voxel-wise model calculations included variance, median,

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

and percentage of deviations across 143k voxels, organized into 28 ROIs for the anatomical atlas, 10 ROIs for the task-based atlas, and 17 ROIs for the resting-state atlas.

Permutation testing revealed significant predictions for ASD and SZ (AUROC values ranging from 0.56 to 0.64), using various models based on deviations from the cerebellar normative model (Figure 2). Prediction performance for MCI, AD and BD were not above chance levels. For SZ, the most predictive models were those centered around median and variance measures summarized within ROIs for the voxel-wise models. In contrast, for ASD, models based on the lobular volumes and voxel-wise variance within ROIs were found to be the most predictive. No notable differences between models based on different parcellations were found (Supplementary Figure 2). Furthermore, AUROC scores from RF predictions showed similar patterns of the logistic regression, underscoring the consistency of these techniques (Supplementary Figure 3).

Figure 3 presents the feature importance weights in a logistic regression model used to analyze SZ and ASD. In SZ, significant negative deviation percentages were found in the vermis IX and Left IV regions. In task-based functional areas, regions associated with verbal fluency, word comprehension and mental arithmetic (region 9) and autobiographical recall, visual letter recognition, interference resolution (region 10) were notable. From the resting-state atlas, limbic A (region 10) and somatomotor A (region 3) emerged as important. For median in SZ, the anatomical regions Right I-III and vermis VIII were highlighted. Using task-based atlases, the top predictive regions were functionality linked to divided attention (region 5) and right-hand movement (region 2). Predictive models using an atlas based on resting-state atlas highlighted Visual B (region 2) and limbic A (region 10).

In ASD, predictive models based on variance (summarized within ROIs) revealed the most significant contributions from posterior cerebellar regions of Left VIIB and Left Crus II while models based on lobular volumes features point to Right VI and Left Crus II. Using the

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

task-based functional atlas, the most predictive regions were functionally linked to narrative, emotion, language processing (region 7) and right-hand movement, motor planning, divided attention (region 2). Complementary insights and detailed rankings of feature importance are available in Supplementary Tables 3-7.

Discussion

This study aimed to test the predictive power of deviations from normal cerebellar anatomy with respect to classifying mental and neurological disorders and yielded two main findings. First, we demonstrated that cerebellar features offer moderate power for prediction for ASD and SZ but did not reliably distinguish healthy controls from patients with BD, MCI or AD. Second, feature importance analyses showed that posterior regions of the cerebellum, known for its contributions in cognitive processes (15) were the most important predictors of ASD and SZ.

Our study reveals that features derived from lobular and voxel-wise normative model possess moderate predictive capabilities in ASD and SZ. This is in line with our previous study which reported small to medium case-control differences in normative cerebellar anatomy for both ASD and SZ (17). However, it is worth mentioning that the current analysis considering cerebellar features only yielded a moderate level of prediction accuracy**.** Including other key brain regions and employing a multimodal approach that integrates different types of brain imaging data may improve the prediction (46,47).

Feature importance analysis for the prediction of SZ highlighted contributions from both motor (48,49) and cognitive regions (15). The limbic vermis, specifically Vermis IX, a region with reported reductions in individuals with SZ (50–52), displayed the highest feature importance in percentage of extreme negative deviations when using the anatomical atlas. This may be interpreted in the context of limbic vermis's role in emotional processing, facial

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

276 expression recognition (53,54), and mentalizing, which is integral to understanding others'
277 mental states, also known as theory of mind (55,56). The involvement of the limbic vermis in
278 these processes is supported by evidence showing connections from this cerebellar region to
279 both cortical and subcortical limbic areas (57). A small study showed smaller inferior posterior
280 lobe in children and adolescents with childhood-onset SZ compared to healthy controls (51),
281 yet, contrasting findings indicate abnormalities may also exist in the anterior lobe (58). A recent
282 study examining a series of 17 individuals with SZ or undifferentiated psychosis (59) showed
283 posterior vermis-predominant cerebellar hypoplasia.

284 As functional topography does not consistently adhere to anatomical boundaries in the
285 cerebellum, we also examined task-based and resting state atlases. In general, there were both
286 slight discrepancies and shared areas when identifying the features of highest importance
287 across models based on different atlases. Indeed, no one atlas consistently emerged as better
288 than any other. However, we believe that moving between atlases significantly aids in
289 functional interpretation of our findings. For instance, when using the percentage of extreme
290 negative deviations (summarized within ROIs) to predict SZ, cerebellar regions functionally
291 linked to verbal fluency, word comprehension and mental arithmetic (region 9) and limbic A
292 (region 10) exhibited the highest feature importance. And when examining vermis IX in the
293 task-based atlas, it highlights the region of saccades, visual working memory and visual letter
294 recognition (19). Past studies literatures exhibit strong resting-state connectivity between
295 lobules I-VI and vermis VIIB-IX of the cerebellum and the visual network (60), and
296 oculomotor abnormalities are observed in SZ (61). For median, Right I-III showed the highest
297 feature importance in the anatomical regions followed by vermis VIII. On the other hand, a
298 study investigating functional connectivity reported hypoactivation in the vermis III, VI, VII,
299 and VIII, along with a negative correlation between the vermis and time processing abilities in

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

individuals with SZ (62). Moreover, divided attention (region 5) in task-based and Visual B (region 2) and limbic A (region 10) in resting state are highlighted for features in median.

Like SZ (63), ASD is a complex neurodevelopmental condition (12) involving a range of clinical characteristics, including repetitive behaviors, restricted interests, and difficulties in social interaction and communication (64). The substantial heterogeneity in clinical characteristics and severity, ranging from the highest functioning form of autism to those requiring substantial support in their everyday life, makes it challenging to demarcate a common neurobiological underpinning (65). Previous studies did not provide conclusive associations between cerebellar volume and ASD (66). Perhaps related to this, in this study, ASD was significantly associated with variance, i.e. the spread of the deviations within a region. Analyses of feature importance in significant ASD model highlighted Left VIIB and Left Crus II as well as the narrative, emotion and language processing region 7 in the task-based atlas. In addition to median and extreme deviations, variability within cerebellar sub-regions, especially those connected to higher cognitive areas, could thus be a relevant imaging-based marker of ASD. Left VIIB of the anatomical atlas and the region 7 of the task-based atlas overlaps in ASD. These regions of Crus I-II and lobules VIIB are densely connected to the prefrontal and parietal cortices for higher level processes through cerebello-thalamo-cortico-pontine cerebellar circuits (67).

Previous studies have consistently reported good classification of dementia based on imaging data (AUROC ranging from 0.904 to 0.920) (3). Thus, our lack of any significant predictive models for this condition was somewhat surprising (but note that effects of MCI and AD were also relatively small in our previous study (17)). While one must be careful in interpreting null-findings, this lack of any significant effects in a moderately large sample of baseline AD patients nonetheless suggests that the cerebellum is relatively spared (68,69). On the other hand, in both typical aging and AD, grey matter loss in the cerebellum's Crus I-II and

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

lobule VI is observed, with typical aging showing a bilateral decline and AD in the right hemisphere (70). The cerebellum remains under-studied, and we need to explore how aging and AD pathology contribute to cerebellar atrophy.

There are limitations to consider in our study. First, harmonizing behavioral, cognitive, genetic, phenotypic, lifestyle, symptomatology, and medical history data across various datasets poses significant challenges, especially when aiming for a large sample size essential for assessing generalizability. We assembled a group of participants for whom we had access to essential information such as diagnosis, sex, age, scanning site, and brain imaging data. General limitations in machine learning such as sample size (71) should be taken into consideration when interpreting the current findings. Next, accurately classifying complex clinical conditions is challenging due to the intrinsic heterogeneity of these conditions, which manifests as a wide array of symptoms and genetic variations. Some individuals may exhibit resilience due to genetic or lifestyle factors, which can complicate accurate predictions (72). Further, the existence of sub-groups within heterogeneous conditions, such as ASD, complicates the interpretation of performance metrics of prediction models. Neurodevelopmental changes raise concerns about the appropriateness of applying adult template space and atlases to younger children and adolescents (73). The cerebellum's distinct position within the skull and its intricate folding pattern also present challenges in obtaining precise MRI data. Finally, an AUROC value in the range of 0.7 to 0.8 can be deemed acceptable for certain clinical applications (74), indicating fair discrimination which includes the range of our model. However, for many clinical scenarios, this may not suffice, as values from 0.8 to 0.9 are generally regarded as appropriate (75). Future research efforts should aim to address these limitations and further enhance our understanding of predictive models.

Conclusion

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

This study tested the value of cerebellar-derived features for predictions of five mental and neurological conditions. The analysis revealed moderate prediction performance for ASD and SZ, with strongest contributions from posterior cerebellar regions.

Data availability

In this study, we used brain imaging from ABIDE, ADNI, AIBL**,** DEMGEN, and TOP. The cerebellar normative models from this work are available on via PCNportal (76): https://pcnportal.dccn.nl/.

Code availability

All code used in this work is available at FreeSurfer (https://surfer.nmr.mgh.harvard.edu), FSL (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FslInstallation), ACAPULCO (https://gitlab.com/shuohan/acapulco), and SUIT (https://github.com/jdiedrichsen/suit). Code for normative model is available as open-source python package, Predictive Clinical Neuroscience (PCN) toolkit (https://github.com/amarquand/PCNtoolkit). Further code is forked to or published on https://github.com/MHM-lab.

Ethics of the study

These are analyses of publicly and privately available data. Description of informed consent and other ethical procedures is extensively described in each study, referenced in the manuscript. The data were stored and analyzed using University of Oslo's secure platform, *Services for sensitive data* (TSD), in compliance with Norwegian privacy regulations.

Funding

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

The work was supported by the South-Eastern Norway Regional Health Authority (2021040, supporting M.K. & T.M.; 2018037, 2018076, 2019101, 2021070, 2023012, 500189), DFG Emmy Noether 513851350 (supporting T.W.), NordForsk (#164218), the Research Council of Norway (249795, 248238, 276082, 286838, 288083, 323951, 324499), Stiftelsen Kristian Gerhard Jebsen, ERA-Net Cofund through the ERA PerMed project IMPLEMENT, and the European Research Council under the European Union's Horizon 2020 research and Innovation program (ERC StG Grant No. 802998). A.F.M gratefully acknowledges funding from the European Research Council ('MENTALPRECISION' 101001118) and from the Raynor Foundation. The funders had no role in conception of the study as well as the analyses and/or interpretations of the results.

Conflict of Interest Disclosures

O.A.A. has received speaker fees from Lundbeck, Janssen, Otsuka, and Sunovion and is a consultant to Cortechs.ai and Precision-Health.ai. E.H.L. has received speaker fees from Lundbeck, and is the CTO and shareholder of baba.vision. L.T.W. and T.W. are shareholders of baba.vision. KP contributed to clinical trials for Roche (BN29553) and Novo Nordisk (NN6535-4730), outside the submitted work. The other authors report no competing interests.

Acknowledgements

We are grateful to all the individuals who participated in the studies and acknowledge the contributions of the clinicians and researchers involved in the recruitment and assessment of participants for making this work possible. We want to acknowledge the Norwegian registry of persons assessed for cognitive symptoms (NorCog), for providing access to patient data. We performed this work on the Services for sensitive data (TSD), University of Oslo, Norway,

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

with resources provided by UNINETT Sigma2 - the National Infrastructure for High-Performance Computing and Data Storage in Norway.

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. Also, data used in preparation of this article were obtained from the Australian Imaging Biomarkers and Lifestyle Study of Ageing (AIBL) databases (adni.loni.usc.edu).

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Authorship Contributions

T.M., T.W., and M.K. originally conceived of the project. M.K., N.S., T.W., T.M., and E.H.L. performed the analyses. M.K. T.W. T.M. wrote the initial draft of the manuscript. O.A., G.R., K.P., D.A., G.S., N.E.S., O.B.S., A.F.M., C.F.B., D.A., T.U., T.W. and L.W. contributed to data curation. All authors discussed the results and contributed to the final manuscript.

References

1. Chen J, Patil KR, Yeo BTT, Eickhoff SB (2023): Leveraging Machine Learning for Gaining Neurobiological and Nosological Insights in Psychiatric Research. *Biol Psychiatry* 93: 18–28.

2. Meehan AJ, Lewis SJ, Fazel S, Fusar-Poli P, Steyerberg EW, Stahl D, Danese A (2022): Clinical prediction models in psychiatry: a systematic review of two decades of progress and challenges [no. 6]. *Mol Psychiatry* 27: 2700–2708.

3. Leonardsen EH, Persson K, Grødem E, Dinsdale N, Schellhorn T, Roe JM, *et al.* (2024): Constructing personalized characterizations of structural brain aberrations in patients with dementia using explainable artificial intelligence. *Npj Digit Med* 7: 1–14.

4. Rashid B, Calhoun V (2020): Towards a brain-based predictome of mental illness. *Hum Brain Mapp* 41: 3468–3535.

5. Wolfers T, Buitelaar JK, Beckmann CF, Franke B, Marquand AF (2015): From estimating activation locality to predicting disorder: A review of pattern recognition for neuroimaging-based psychiatric diagnostics. *Neurosci Biobehav Rev* 57: 328–349.

6. Zhu Y, Maikusa N, Radua J, Sämann PG, Fusar-Poli P, Agartz I, *et al.* (2024): Using brain structural neuroimaging measures to predict psychosis onset for individuals at clinical high-risk. *Mol Psychiatry* 29: 1465–1477.

7. Parvizi J (2009): Corticocentric myopia: old bias in new cognitive sciences. *Trends Cogn Sci* 13: 354–359.

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

8. Hughes DE, Kunitoki K, Elyounssi S, Luo M, Bazer OM, Hopkinson CE, *et al.* (2023): Genetic patterning for child psychopathology is distinct from that for adults and implicates fetal cerebellar development [no. 6]. *Nat Neurosci* 26: 959–969.

9. Schmahmann JD (2016): Cerebellum in Alzheimer's disease and frontotemporal dementia: not a silent bystander. *Brain* 139: 1314–1318.

10. Moberget T, Doan NT, Alnæs D, Kaufmann T, Córdova-Palomera A, Lagerberg TV, *et al.* (2018): Cerebellar volume and cerebellocerebral structural covariance in schizophrenia: a multisite mega-analysis of 983 patients and 1349 healthy controls [no. 6]. *Mol Psychiatry* 23: 1512–1520.

11. D'Mello AM, Stoodley CJ (2015): Cerebro-cerebellar circuits in autism spectrum disorder. *Front Neurosci* 9. Retrieved February 10, 2023, from https://www.frontiersin.org/articles/10.3389/fnins.2015.00408

12. Stoodley CJ (2016): The Cerebellum and Neurodevelopmental Disorders. *Cerebellum* 15: 34–37.

13. van der Heijden ME, Gill JS, Sillitoe RV (2021): Abnormal Cerebellar Development in Autism Spectrum Disorders. *Dev Neurosci* 43: 181–190.

14. Schmahmann JD (2004): Disorders of the Cerebellum: Ataxia, Dysmetria of Thought, and the Cerebellar Cognitive Affective Syndrome. *J Neuropsychiatry Clin Neurosci* 16: 367–378.

15. Schmahmann JD (2019): The cerebellum and cognition. *Neurosci Lett* 688: 62–75.

16. Schmahmann JD (2021): Chapter 6 - Emotional disorders and the cerebellum: Neurobiological substrates, neuropsychiatry, and therapeutic implications. In: Heilman KM, Nadeau SE, editors. *Handbook of Clinical Neurology*, vol. 183. Elsevier, pp 109–154.

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

17. Kim M, Leonardsen E, Rutherford S, Selbæk G, Persson K, Steen NE, *et al.* (2024): Mapping cerebellar anatomical heterogeneity in mental and neurological illnesses. *Nat Ment Health* 1–12.

18. Buckner RL, Krienen FM, Castellanos A, Diaz JC, Yeo BTT (2011): The organization of the human cerebellum estimated by intrinsic functional connectivity. *J Neurophysiol* 106: 2322–2345.

19. King M, Hernandez-Castillo CR, Poldrack RA, Ivry RB, Diedrichsen J (2019): Functional boundaries in the human cerebellum revealed by a multi-domain task battery. *Nat Neurosci* 22: 1371–1378.

20. Ho DE, Imai K, King G, Stuart EA (2011): MatchIt : Nonparametric Preprocessing for Parametric Causal Inference. *J Stat Softw* 42. https://doi.org/10.18637/jss.v042.i08

21. Ségonne F, Dale AM, Busa E, Glessner M, Salat D, Hahn HK, Fischl B (2004): A hybrid approach to the skull stripping problem in MRI. *NeuroImage* 22: 1060–1075.

22. Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM (2012): FSL. *NeuroImage* 62: 782–790.

23. Jenkinson M, Smith S (2001): A global optimisation method for robust affine registration of brain images. *Med Image Anal* 5: 143–156.

24. Han S, An Y, Carass A, Prince JL, Resnick SM (2020): Longitudinal analysis of regional cerebellum volumes during normal aging. *NeuroImage* 220: 117062.

25. Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, Gee JC (2010): N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging* 29: 1310–1320.

26. Fonov V, Evans AC, Botteron K, Almli CR, McKinstry RC, Collins DL, Brain Development Cooperative Group (2011): Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage* 54: 313–327.

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

498 27. Carass A, Cuzzocreo JL, Han S, Hernandez-Castillo CR, Rasser PE, Ganz M, *et al.*
499 (2018): Comparing fully automated state-of-the-art cerebellum parcellation from
500 magnetic resonance images. *NeuroImage* 183: 150–172.
501 28. Kerestes R, Han S, Balachander S, Hernandez-Castillo C, Prince JL, Diedrichsen J,
502 Harding IH (2022): A Standardized Pipeline for Examining Human Cerebellar Grey
503 Matter Morphometry using Structural Magnetic Resonance Imaging. *J Vis Exp JoVE*.
504 https://doi.org/10.3791/63340
505 29. Diedrichsen J (2006): A spatially unbiased atlas template of the human cerebellum.
506 *NeuroImage* 33: 127–138.
507 30. Diedrichsen J, Balsters JH, Flavell J, Cussans E, Ramnani N (2009): A probabilistic MR
508 atlas of the human cerebellum. *NeuroImage* 46: 39–46.
509 31. Marquand AF, Kia SM, Zabihi M, Wolfers T, Buitelaar JK, Beckmann CF (2019):
510 Conceptualizing mental disorders as deviations from normative functioning. *Mol*
511 *Psychiatry* 24: 1415–1424.
512 32. Rutherford S, Kia SM, Wolfers T, Fraza C, Zabihi M, Dinga R, *et al.* (2022): The
513 normative modeling framework for computational psychiatry. *Nat Protoc* 17: 1711–
514 1734.
515 33. Rios G, Tobar F (2019): Compositionally-warped Gaussian processes. *Neural Netw* 118:
516 235–246.
517 34. Fraza C, Dinga R, Beckmann CF, Marquand AF, Andre Marquand (2021): Warped
518 Bayesian Linear Regression for Normative Modelling of Big Data. *bioRxiv*.
519 https://doi.org/10.1101/2021.04.05.438429
520 35. Dinga R, Fraza CJ, Bayer JMM, Kia SM, Beckmann CF, Marquand AF (2021, June 14):
521 Normative modeling of neuroimaging data using generalized additive models of
522 location scale and shape. bioRxiv, p 2021.06.14.448106.

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

523 36. Kia SM, Huijsdens H, Rutherford S, Dinga R, Wolfers T, Mennes M, *et al.* (2021, May
524 30): Federated Multi-Site Normative Modeling using Hierarchical Bayesian
525 Regression. bioRxiv, p 2021.05.28.446120.
526 37. Bayer JMM, Dinga R, Kia SM, Kottaram AR, Wolfers T, Lv J, *et al.* (2021, May 30):
527 Accommodating site variation in neuroimaging data using normative and hierarchical
528 Bayesian models. bioRxiv, p 2021.02.09.430363.
529 38. Yeo BTT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, *et al.*
530 (2011): The organization of the human cerebral cortex estimated by intrinsic
531 functional connectivity. *J Neurophysiol* 106: 1125–1165.
532 39. Alnæs D, Kaufmann T, van der Meer D, Córdova-Palomera A, Rokicki J, Moberget T, *et*
533 *al.* (2019): Brain Heterogeneity in Schizophrenia and Its Association With Polygenic
534 Risk. *JAMA Psychiatry* 76: 739–748.
535 40. Brugger SP, Howes OD (2017): Heterogeneity and Homogeneity of Regional Brain
536 Structure in Schizophrenia: A Meta-analysis. *JAMA Psychiatry* 74: 1104–1111.
537 41. Rutherford S, Barkema P, Tso IF, Sripada C, Beckmann CF, Ruhe HG, Marquand AF
538 (2023): Evidence for embracing normative modeling ((C. I. Baker, T. Constable, & O.
539 Esteban, editors)). *eLife* 12: e85082.
540 42. Wolfers T, Doan NT, Kaufmann T, Alnæs D, Moberget T, Agartz I, *et al.* (2018):
541 Mapping the Heterogeneous Phenotype of Schizophrenia and Bipolar Disorder Using
542 Normative Models. *JAMA Psychiatry* 75: 1146–1155.
543 43. Zabihi M, Oldehinkel M, Wolfers T, Frouin V, Goyard D, Loth E, *et al.* (2018):
544 Dissecting the Heterogeneous Cortical Anatomy of Autism Spectrum Disorder Using
545 Normative Models. *Biol Psychiatry Cogn Neurosci Neuroimaging* 4: 567–578.

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

546 44. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, *et al.* (2018, June
547 5): Scikit-learn: Machine Learning in Python [no. arXiv:1201.0490]. arXiv.
548 https://doi.org/10.48550/arXiv.1201.0490
549 45. Chen T, Guestrin C (2016, March 1): XGBoost: A Scalable Tree Boosting System. *arXiv*
550 *E-Prints*. https://doi.org/10.48550/arXiv.1603.02754
551 46. Rokicki J, Wolfers T, Nordhøy W, Tesli N, Quintana DS, Alnæs D, *et al.* (2021):
552 Multimodal imaging improves brain age prediction and reveals distinct abnormalities
553 in patients with psychiatric and neurological disorders. *Hum Brain Mapp* 42: 1714–
554 1726.
555 47. Tavares V, Vassos E, Marquand A, Stone J, Valli I, Barker GJ, *et al.* (2023): Prediction
556 of transition to psychosis from an at-risk mental state using structural neuroimaging,
557 genetic, and environmental data. *Front Psychiatry* 13: 1086038.
558 48. Walther S (2015): Psychomotor symptoms of schizophrenia map on the cerebral motor
559 circuit. *Psychiatry Res Neuroimaging* 233: 293–298.
560 49. Walther S, Strik W (2012): Motor Symptoms and Schizophrenia. *Neuropsychobiology* 66:
561 77–92.
562 50. Heath RG, Franklin DE, Walker CF, Keating JW (1982): Cerebellar vermal atrophy in
563 psychiatric patients. *Biol Psychiatry* 17: 569–583.
564 51. Jacobsen LK, Giedd JN, Berquin PC, Krain AL, Hamburger SD, Kumra S, Rapoport JL
565 (1997): Quantitative morphology of the cerebellum and fourth ventricle in childhood-
566 onset schizophrenia. *Am J Psychiatry* 154: 1663–1669.
567 52. Picard H, Amado I, Mouchet-Mages S, Olié J-P, Krebs M-O (2008): The Role of the
568 Cerebellum in Schizophrenia: an Update of Clinical, Cognitive, and Functional
569 Evidences. *Schizophr Bull* 34: 155–172.

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

570 53. Brady RO, Gonsalvez I, Lee I, Öngür D, Seidman LJ, Schmahmann JD, *et al.* (2019):
571 Cerebellar-Prefrontal Network Connectivity and Negative Symptoms in
572 Schizophrenia. *Am J Psychiatry* 176: 512–520.
573 54. Metoki A, Wang Y, Olson IR (2022): The Social Cerebellum: A Large-Scale
574 Investigation of Functional and Structural Specificity and Connectivity. *Cereb Cortex*
575 32: 987–1003.
576 55. Clausi S, Olivito G, Lupo M, Siciliano L, Bozzali M, Leggio M (2019): The Cerebellar
577 Predictions for Social Interactions: Theory of Mind Abilities in Patients With
578 Degenerative Cerebellar Atrophy. *Front Cell Neurosci* 12: 510.
579 56. Van Overwalle F, Manto M, Cattaneo Z, Clausi S, Ferrari C, Gabrieli JDE, *et al.* (2020):
580 Consensus Paper: Cerebellum and Social Cognition. *The Cerebellum* 19: 833–868.
581 57. Blatt GJ, Oblak AL, Schmahmann JD (2013): Cerebellar Connections with Limbic
582 Circuits: Anatomy and Functional Implications. In: Manto M, Schmahmann JD, Rossi
583 F, Gruol DL, Koibuchi N, editors. *Handbook of the Cerebellum and Cerebellar*
584 *Disorders*. Dordrecht: Springer Netherlands, pp 479–496.
585 58. Nopoulos PC, Ceilley JW, Gailis EA, Andreasen NC (1999): An MRI study of cerebellar
586 vermis morphology in patients with schizophrenia: evidence in support of the
587 cognitive dysmetria concept. *Biol Psychiatry* 46: 703–711.
588 59. Leslie AC, Ward MP, Dobyns WB (2024): Undifferentiated psychosis or schizophrenia
589 associated with vermis-predominant cerebellar hypoplasia. *Am J Med Genet A* 194:
590 e63416.
591 60. Sang L, Qin W, Liu Y, Han W, Zhang Y, Jiang T, Yu C (2012): Resting-state functional
592 connectivity of the vermal and hemispheric subregions of the cerebellum with both
593 the cerebral cortical networks and subcortical structures. *NeuroImage* 61: 1213–1225.

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

61. Hutton S, Kennard C (1998): Oculomotor abnormalities in schizophrenia. *Neurology* 50: 604–609.

62. Lošák J, Hüttlová J, Lipová P, Mareček R, Bareš M, Filip P, *et al.* (2016): Predictive Motor Timing and the Cerebellar Vermis in Schizophrenia: An fMRI Study. *Schizophr Bull* 42: 1517–1527.

63. Moberget T, Alnæs D, Kaufmann T, Doan NT, Córdova-Palomera A, Norbom LB, *et al.* (2019): Cerebellar Gray Matter Volume Is Associated With Cognitive Function and Psychopathology in Adolescence. *Biol Psychiatry* 86: 65–75.

64. American Psychiatric Association (2013): *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition. American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596

65. Verhoeff B (2015): Fundamental challenges for autism research: the science–practice gap, demarcating autism and the unsuccessful search for the neurobiological basis of autism. *Med Health Care Philos* 18: 443–447.

66. Traut N, Beggiato A, Bourgeron T, Delorme R, Rondi-Reig L, Paradis A-L, Toro R (2018): Cerebellar Volume in Autism: Literature Meta-analysis and Analysis of the Autism Brain Imaging Data Exchange Cohort. *Biol Psychiatry* 83: 579–588.

67. Becker EBE, Stoodley CJ (2013): Chapter One - Autism Spectrum Disorder and the Cerebellum. In: Konopka G, editor. *International Review of Neurobiology*, vol. 113. Academic Press, pp 1–34.

68. Bernard JA, Seidler RD (2014): Moving forward: age effects on the cerebellum underlie cognitive and motor declines. *Neurosci Biobehav Rev* 42: 193–207.

69. Liang KJ, Carlson ES (2019): Resistance, Vulnerability and Resilience: A Review of the Cognitive Cerebellum in Aging and Neurodegenerative Diseases. *Neurobiol Learn Mem* S1074-7427(19)30005-X.

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

619 70. Gellersen HM, Guell X, Sami S (2021): Differential vulnerability of the cerebellum in
620 healthy ageing and Alzheimer's disease. *NeuroImage Clin* 30: 102605.
621 71. Arbabshirani MR, Plis S, Sui J, Calhoun VD (2017): Single subject prediction of brain
622 disorders in neuroimaging: Promises and pitfalls. *NeuroImage* 145: 137–165.
623 72. Venkatraghavan V, Voort SR van der, Bos D, Smits M, Barkhof F, Niessen WJ, *et al.*
624 (2023): Computer-Aided Diagnosis and Prediction in Brain Disorders. In: Colliot O,
625 editor. *Machine Learning for Brain Disorders*. New York, NY: Springer US, pp 459–
626 490.
627 73. Gaiser C, van der Vliet R, de Boer AAA, Donchin O, Berthet P, Devenyi GA, *et al.*
628 (2024): Population-wide cerebellar growth models of children and adolescents. *Nat*
629 *Commun* 15: 2351.
630 74. Mandrekar JN (2010): Receiver Operating Characteristic Curve in Diagnostic Test
631 Assessment. *J Thorac Oncol* 5: 1315–1316.
632 75. Zarogianni E, Storkey AJ, Borgwardt S, Smieskova R, Studerus E, Riecher-Rössler A,
633 Lawrie SM (2019): Individualized prediction of psychosis in subjects with an at-risk
634 mental state. *Schizophr Res* 214: 18–23.
635 76. Barkema P, Rutherford S, Lee H-C, Kia SM, Savage H, Beckmann C, Marquand A
636 (2023): Predictive Clinical Neuroscience Portal (PCNportal): instant online access to
637 research-grade normative models for clinical neuroscientists. *Wellcome Open Res* 8:
638 326.

639
640
641

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

642 **Table 1.** Matched sample description and demographics

		N (Participants	N (Scanners)	Age (Mean, S.D.)	Sex (%F: %M)
Matched HC	Alzheimer's Disease	116	13	71.72(7.12)	55:45
	ASD	317	25	15.96(7.45)	17:83
	Bipolar Disorder	238	3	33.06(10.50)	55:45
	Mild Cognitive Impairment	122	3	65.62(9.91)	42:58
	Schizophrenia	195	3	30.13 (8.15)	41:59
Clinical	Alzheimer's Disease	116	13	73.11(7.60)	55:45
	ASD	317	25	12.35(4.42)	17:83
	Bipolar Disorder	238	3	31.61(11.40)	55:45
	Mild Cognitive Impairment	122	3	67.25(9.27)	42:58
	Schizophrenia	195	3	28.29 (9.45)	41:59

643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Figure Captions

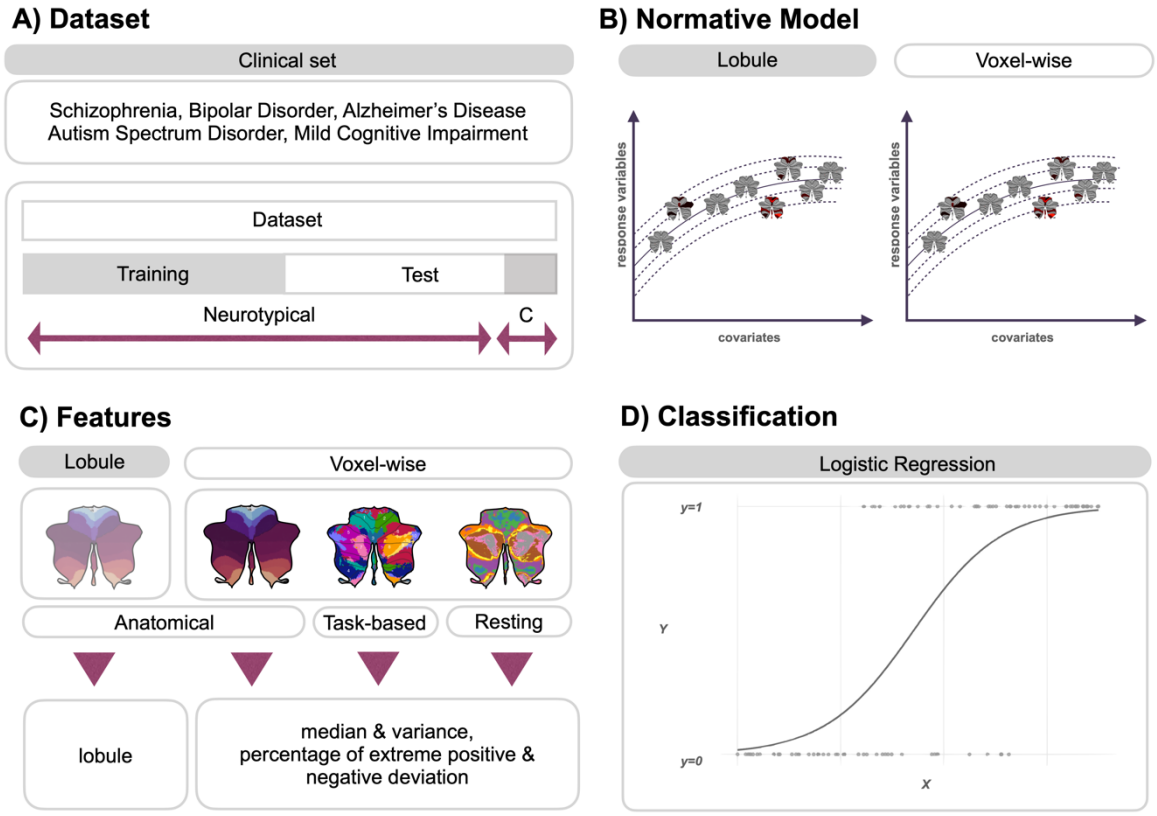


Figure 1. Overview of Predicting Mental and Neurological Illnesses. (A) The study incorporated five clinical datasets. (B) Individuals without a diagnosis were divided into training and testing sets to evaluate the cerebellar normative models, which were prepared in both lobular and voxel-wise features. (C) The analysis utilized six distinct types of features, median, variance, and percentage of extreme positive and negative deviation, alongside lobular volume. (D) Logistic regression algorithm was employed to determine the likelihood of a clinical diagnosis in an individual.

medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

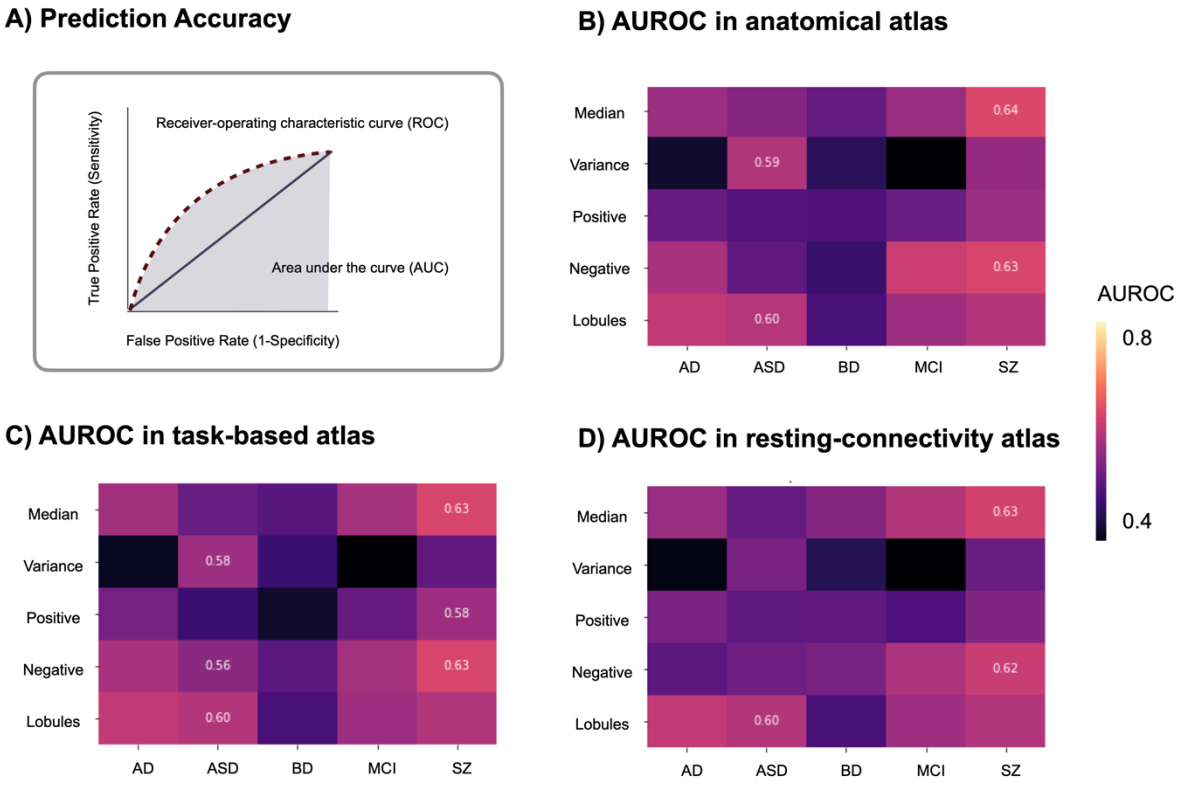
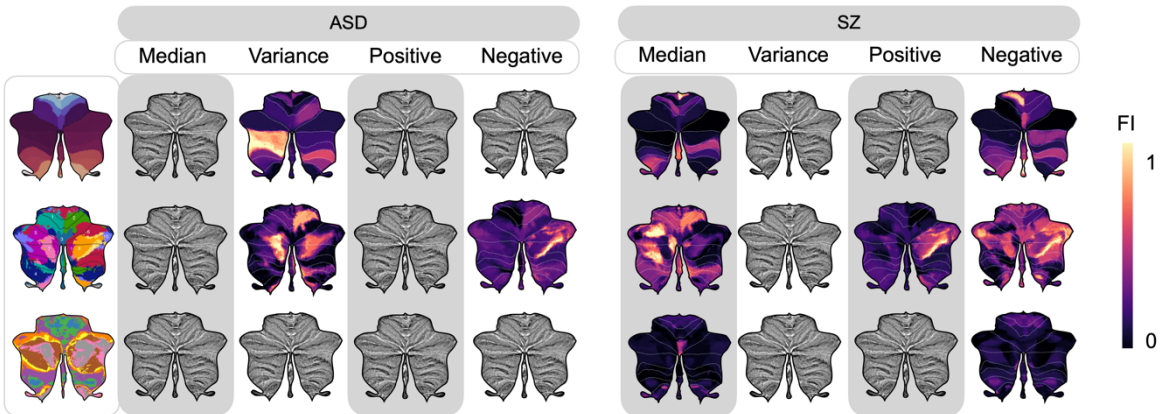


Figure 2. Cerebellar features moderately predict ASD and SZ. (A) Information from the anatomical atlas (28 regions), task-based (10 regions) or resting-state (17 regions) are compiled into features that were used as predictors by the logistic regression model to make predictions. The area under the receiver operating characters curve (AUROC) serves as an important measure in evaluating the performance of a binary classifier, representing a trade-off between the classifier's sensitivity (true positive rate) and specificity (true negative rate). The reliability and robustness of the AUROC were assessed by computing it over 1,000 permutations, which aids in determining whether the classifier's performance is statistically significant or due to random chance. (B-D) The values that survived multiple comparison are shown.



medRxiv preprint doi: https://doi.org/10.1101/2024.12.10.24318590; this version posted December 18, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 3. Different regions show distinct importance across atlases in ASD and SZ. The feature importance (FI) values derived from logistic regression reveal the contribution of each specific cerebellar region to predictive modelling, relative to average prediction outcomes. FI values accentuate distinct cerebellar regions with unique predictive capabilities as identified in lobules, anatomical, task-based, and resting-state atlases through voxel-wise analysis. Features that remained significant after adjustments for multiple comparisons of AUROC are shown.