Using Brain Structural Neuroimaging Measures to Predict Psychosis Onset for Individuals at Clinical High-Risk

Shinsuke Koike (✉ skoike-tky@umin.ac.jp)
 University of Tokyo   https://orcid.org/0000-0002-3375-236X
Yinghan Zhu
Norihide Maikusa
 https://orcid.org/0000-0003-0943-4684
Joaquim Radua
 Consorci Institut D'Investigacions Biomediques August Pi I Sunyer
Philipp Sämann
 Max Planck Institute of Psychiatry   https://orcid.org/0000-0001-8523-3628
Paolo Fusar-Poli
 IoPPN   https://orcid.org/0000-0003-3582-6788

Abstract

Machine learning approaches using structural magnetic resonance imaging (sMRI) can be informative for disease classification, although their ability to predict psychosis is largely unknown. We created a model with individuals at CHR who developed psychosis later (CHR-PS+) from healthy controls (HCs) that can differentiate each other. We also evaluated whether we could distinguish CHR-PS + individuals from those who did not develop psychosis later (CHR-PS-) and those with uncertain follow-up status (CHR-UNK).

T1-weighted structural brain MRI scans from 1,165 individuals at CHR (CHR-PS+, n = 144; CHR-PS-, n = 793; and CHR-UNK, n = 228), and 1,029 HCs, were obtained from 21 sites. We used ComBat to harmonize measures of subcortical volume, cortical thickness and surface area data and corrected for non-linear effects of age and sex using a general additive model. CHR-PS+ (n = 120) and HC (n = 799) data from 20 sites served as a training dataset, which we used to build a classifier. The remaining samples were used external validation datasets to evaluate classifier performance (test, independent confirmatory, and independent group [CHR-PS- and CHR-UNK] datasets).

The accuracy of the classifier on the training and independent confirmatory datasets was 85% and 73% respectively. Regional cortical surface area measures-includingthose from the right superior frontal, right superior temporal, and bilateral insular cortices strongly contributed to classifying CHR-PS + from HC. CHR-PS- and CHR-UNK individuals were more likely to be classified as HC compared to CHR-PS+ (classification rate to HC: CHR-PS+, 30%; CHR-PS-, 73%; CHR-UNK, 80%).

We used multisite sMRI to train a classifier to predict psychosis onset in CHR individuals, and it showed promise predicting CHR-PS + in an independent sample. The results suggest that when considering adolescent brain development, baseline MRI scans for CHR individuals may be helpful to identify their prognosis. Future prospective studies are required about whether the classifier could be actually helpful in the clinical settings.

Introduction

The clinical high risk (CHR) paradigm is widely used with the goal of improving early detection of and prevention of psychotic disorders.[1] Individuals are considered at CHR for psychosis if they meet criteria for attenuated positive symptom syndrome (APSS), brief intermittent (limited) psychotic syndrome (BLIPS), and/or genetic risk and deterioration syndrome (GRDS) based on semistructured interviews.[2-6] The CHR state is present in 1.7% of the general population and 19.2% of clinical samples.[7] CHR individuals have a higher risk of developing psychosis (0.15 at 1 year) comparing to healthy controls, the transition risk increased from 0.09 at half years to 0.27 at 4 years.[8] However,most CHR subjects who do not transition to psychosis will continue to meet CHR criteria or experience attenuated psychosis symptoms at follow-up and only 33% will eventually remit .[8,9]

The CHR state, is also associated with alterations in proxy measures of brain structure. Previous structural magnetic resonance imaging (MRI) studies reported a progressive decrease in gray matter volume in the medial and superior temporal and medial frontal cortex during the transition period among CHR individuals.[10-13] Gray matter volume continued to decrease several years after disease onset.[10,13,14] Widespread lower cortical thickness has also been identified in cross-sectional MRI data in individuals at CHR in a large-scale pooled analysis of the Enhancing Neuro Imaging Genetics through Meta-Analysis (ENIGMA) CHR Working Group[15,16], with CHR individuals exhibiting a pattern of lower cortical thickness similar to individuals with established schizophrenia.[17] Furthermore, longitudinal reductions of cortical thickness in the paracentral, superior temporal, and fusiform gyrus have been reported to be associated with psychosis conversion in those at CHR.[11,18] Recent work has indicated that whole-brain sMRI patterns of schizophrenia forecasted 2-year psychosocial impairments in individuals with CHR.[19], suggesting that alterations in brain structure may predict real-life outcomes.

Adolescent development is a crucial time window that is associated with brain-wide changes, including reductions in cortical thickness and volume.[20,21] Cortical characteristics such as gray matter volume, cortical surface area, and cortical thickness decline by about 10% during adolescence.[22] On the other hand, white matter volume was reported peaking in young adulthood.[22] Since the period from adolescence to early adulthood is a high risk time window for psychosis onset,[18] age-related anatomical deviations from typically-occuring declines may hold valuable information to predict later psychosis conversion, especially in frontal and temporal regions that have been implicated in CHR[15,18,23-25] and schizophrenia[26-31]. Further, greater brain age deviations were found to be associated with a higher risk for psychosis over time.[24,32] Importantly, these results suggest that the adolescent brain development pattern of CHR individuals may differ from that of HCs. Indeed, the ENIGMA CHR Working Group has reported that CHR compared to HC participants exhibit altered non-linear age associations with cortical thickness [15], suggesting that cross-sectional between-group differences in sMRI metrics may involve altered adolescent development, trait characteristics associated with psychosis liability, and/or progressive brain pathology around the onset of psychosis.[18,25,33]

An increasing number of studies have attempted to use (cross-sectional) sMRI data to predict outcome or case-control status. These prior studies show that machine learning approaches are informative for differentiating individuals with schizophrenia from HCs[34-39]. Similar findings were observed in different clinical stages of psychosis, including first episode schizophrenia and CHR individuals.[35,36] A major limitation, however, is the need for large and diverse sample sizes to establish a well-tuned classifier that also provides generalized predictive performance.[40,41] Since single sites cannot typically provide the necessary sample sizes,[36,42,43] multisite consortia data may be advantageous if site effects are adequately accounted for (e.g., via cross-site harmonization procedures).[36,42,44] For example, without harmonization, a prior study failed to build a useful model with multi site data.[24] In the current study, we aimed to investigate whether cross-sectional sMRI data can be used to build a classifier to differentiate the neuroanatomical developmental patterns of HCs relative to participants who later developed a psychotic disorder later (CHR-PS+) as biomarkers for future psychosis conversion. As altered developmental processes are implicated in psychosis risk, we considered the potential non-linear effects of age and sex to gain optimal predictive accuracy of trained classifiers.

Here, we combined data from 21 sites harmonized through the ENIGMA CHR Working Group using ComBat[45] to minimize differences related to site-,, scanner- and scanning protocols. Second, to model non-linear age effects, we fitted generalized additive models (GAMs)[46,47] to the HC data, and then applied the fitted GAMs to obtain non-linear age- and sex-corrected features for the entire sample.[48] More specifically, we estimated the model in HCs and applied it to individuals at CHR to capture deviations from the expected patterns of physiological aging. As for patients with early-onset psychosis[49] and schizophrenia[27] have been reported to have abnormally low ICV, all procedures were performed after adjusting the MRI features for effects of ICV. Third, we developed an XGBoost[50] classifier using only HCs and CHR-PS + to determine deviation in neuroanatomical developmental patterns as potential predictors of future psychosis conversion. Finally, we tested the predictive performance of the classifier with the left-out site data, to avoid the potential for information leakage between the training and test data.

We hypothesized that CHR-PS + individuals would be distinguishable from HCs based on features derived from structural MRI features, based on the assumption that those CHR individuals who are most likely to convert to psychosis would show the greatest baseline anatomical alterations. Second, we expected our classifier to label individuals at CHR who had *not* developed a psychotic disorder (CHR-PS-) at follow-up, and most of the CHR with unknown transition status (CHR-UNK, for unknown), as HCs. Third, we expected the classifiers to perform similarly in independent confirmatory datasets, and expected to find associations between classifications and symptom severity.

Methods

Participants

We included data from a total of 1,165 CHR individuals (144 CHR-PS+, 793 CHR-PS−, and 228 CHR-UNK individuals) and 1,029 healthy controls (HCs) from 21 ENIGMA Clinical High Risk for Psychosis Working Group sites (Table 1). CHR individuals met criteria for high-risk status via the Comprehensive Assessment of At-Risk Mental States (CAARMS[51]; n = 650) or the Structured Interview for Prodromal Syndromes (SIPS[52,53]; n = 799).Site-specific inclusion and exclusion criteria, the available scale scores in premorbid IQ, symptom severity, global functioning, and antipsychotic use at scan are the same as in a prior publication (Supplementary Table S1).[15] All sites obtained local institutional review board approval prior to data collection. Written informed consent was obtained from every participant, or from the participant's guardian for participants younger than 18 years. All studies were conducted in accordance with the Declaration of Helsinki.[54]

We applied a two-step approach[36] to evaluate the performance of the models by dividing the data into four datasets: training, test, independent confirmatory, and independent group datasets (**Figure 1**). Test and independent confirmatory datasets were used as external validation datasets. First, the training and test datasets comprised the data from CHR-PS+ and HC from 20 sites except for Toyama, which was used as the independent confirmatory dataset. We chose this dataset because the Toyama site contributed the largest HC sample and excluding this dataset reduced sample imbalance between groups in building a machine learning classifier. Ninety percent of the data were randomly sorted as the training dataset, and the remaining 10% as the test dataset. A Kolmogorov−Smirnov test did not show any significant differences between training and test datasets in any structural features. The independent confirmatory dataset comprised the data from HCs and CHRs at the Toyama site; this data was completely excluded from the training partition, and was used to perform an independent first-step evaluation without site information leakage. To evaluate the classifier on unseen new data, we defined the CHR-PS− and CHR-UNK individuals in all sites as the independent group dataset to perform the second step.

MRI data acquisition and preprocessing

Image Acquisition and Processing

Participating sites contributed to T1-weighted MRI brain scans from 31 MRI scanners, including 29 3-T scanners and 2 1.5-T scanners (Supplementary Table S2). Detailed scan protocols and the number of scans for each protocol are described in the Supplementary Materials. After processing the data using FreeSurfer analysis software at each site[15], we extracted structural features from 153 regions of interest (ROI) including 68 regional measures of cortical thickness, 68 surface area (SA), 16 subcortical volume, and one intracranial volume according to the Desikan-Killiany atlas.[55] We implemented the ENIGMA consortium quality assessment pipeline[26,27,56,56−59] and 8 samples were excluded for lacking 20% of the ROIs data. Remaining missing values (1.20%) were imputed using a k-Nearest Neighbor (k = 3) approach.

ComBat harmonization

ComBat[45] is a harmonization method used to remove scanner and protocol effects based on an adjusted general linear model harmonization method. Based on recent work demonstrating that neuroComBat harmonization increases statistical power within a mega-analytic framework, primary analyses were conducted within a mega-analysis framework using data that were corrected for site and scanner associations using neuroComBat harmonization.[60] Further analyses were conducted using Python version 3.7.12. We applied the extracted cortical thickness, surface area, subcortical volume, and intracranial volume measures with participants' age and sex as covariates, along with protocol and site effects. To confirm that group status had no significant influence on the ComBat harmonization steps, we also conducted ComBat harmonization using the training dataset only (see Supplementary Materials).

Features engineering

First, we fitted a general linear model to regress out effects of intracranial volume. Next, we fitted GAMs to only the HC data to estimate non-linear effects of age and sex for every structural feature; then we applied the fitted GAMs to obtain non-linear age- and sex-corrected features. To verify the absence of information leakage and the stability of the GAMs, we also repeated this procedure 1,000 times on randomly sub-sampled HC data to estimate the GAMs (see Supplementary Materials).

XGBoost

XGBoost is a scalable tree boosting algorithm.[50] We applied standardization for the structural features to building a classifier. The use of input data standardization, optimization of the hyperparameters of the classifier (eta, min_child_weight, max_depth, subsample, colsample_bytree) were tuned using GridSearchCV implemented in the 'scikit-learn' module (version 1.0.2) in Python (https://scikit-learn.org/stable/auto_examples/release_highlights/plot_release_highlights_1_0_0.html).[61] We plotted the weights of the classifier to determine the importance of the features for generalization. The classifier was optimized using a 10-fold cross-validated grid search over a defined parameter grid. Data from the HC group were randomly downsampled to the same sample size as the CHR-PS+ group in each fold. To reduce downsampling bias, downsampling and grid search were repeated 1,000 times and stratified 10-fold for the training data. Then, we applied 10-fold cross-validation and 1,000 permutations to evaluate the significance of the cross-validation scores of the model with the best hyperparameters for the training dataset. The best cross-validation accuracy score was averaged across 1,000 repeats. Permutation tests were conducted by shuffling the labels in the training data, and the permutation-based p-value was calculated.[62] The final model with the best hyperparameters was trained using the entire training dataset. Finally, the trained classifier was applied to the test set and the independent confirmatory dataset with the best parameters tuned by grid search. The predict probability was calculated by the trained classifier for each sample. Predict probability ranges from 0 to 1, with smaller values indicating more likely classification as CHR-PS+. The cut-off point for the predictive performance was set to 0.5.

The predictive performance of the classifier was evaluated using an independent group dataset (CHR-PS− and CHR-UNK). We compared the classifiers built from four different feature sets: (i) only cortical thickness values, (ii) only surface area values, (iii) only subcortical volumes only, and (iv) all features. The classifier with the best predictive performance for the independent confirmatory dataset was used for subsequent analysis.

Statistical analysis

Evaluation metrics

First, the classifier was evaluated using the test, independent confirmatory, and independent group datasets by the given scores of the tuned classifier. We calculated the confusion matrix, macro, and weighted average accuracies to evaluate the classifier because the data used were imbalanced (see Supplementary Materials).[36]

Predictive performance of the classifier

The predictive performance of the classifier was defined as its performance on unseen data (in the independent confirmatory/group datasets) and was assessed using standard evaluation metrics. Chi-squared tests were applied to the classified labels of the test, independent confirmatory, and independent group datasets. Since we compared 6 pairs of groups, respectively, a Bonferroni's correction was applied to adjust for the multiple statistical comparisons (p<0.05/6 = 0.008). Predict probabilities generated by the XGBoost classifier were also tested using an analysis of variance for all samples. A GAM was used to assess non-linear relationships between age and the predictive performance of the classifier. Moreover, we conducted 4 comparisons (HCs vs CHR-PS+, CHR-PS+ vs CHR-PS-, CHR-PS+ vs CHR-UNK and CHR-PS+ vs CHR-PS- and CHR-UNK) of the decision curve analysis[63–65] using 'dcurves' package (version 0.4.0) in R software to estimate the classifier as well.

Relationship between predict probability and demographic and clinical characteristics

We tested the difference in the predictive performance with respect to sex and the existence of APSS, BLIPS, and GRDS using t tests ($p < 0.05/3 = 0.016$). Pearson's correlation analyses were also conducted between standardized IQ and the predict probability. Bonferroni's correction was applied to the subscores ($p < 0.05/4 = 0.0125$). To determine the relationship between the predict probability and symptom severity, Pearson's correlation analyses were performed using the SIPS and CAARMS subscores and global functioning score for CHR-PS+, CHR-PS-, and CHR-UNK groups. We tested z-score normalized positive, negative and general subscores of the SIPS and CAARMS using Pearson's correlation coefficients. Bonferroni's correction was applied to the SIPS or CAARMS subscores ($p < 0.05/4 = 0.0125$). To determine the potential effect of antipsychotic medication on the classification, we also tested the difference in predict probabilities between those with and without medication use for each CHR subgroup using a t-test.

Results

Model evaluation

A non-linear effect of age, sex, and age x sex interaction on SA was found in HCs, as shown in Fig. 2. The classifier using only non-linear fitted SA features (i.e., fit to HCs, applied to all) obtained the best performance in differentiating HC and CHR-PS + groups (Supplementary Table S3). For the SA model, the best cross-validation accuracy within the training dataset was 85% (± 0.00008). The permutation test showed that the classifier performed significantly better than chance level (50%, $p < 0.001$). The accuracies with the best estimator for the test and independent confirmatory datasets were 68% and 73% (Fig. 3B), respectively. Regions with the largest features weights were the superior temporal, insula, superior frontal, superior parietal, fusiform, isthmus of cingulate, and parahippocampal gyri to differentiate HC from CHR-PS + groups (Fig. 3A, **Supplementary Table S2**). For SA in the right superior temporal gyrus, which was the strongest contributing feature of the classifier, the ComBat harmonized feature showed no significant difference among the groups ($p > 0.05$), while ComBat harmonized and non-linear age- and sex-adjusted feature revealed a difference between CHR-PS + and CHR-PS- ($t = 2.14$, $p = 0.033$), and CHR-PS + and CHR-UNK ($t = 2.14$, $p = 0.033$; Fig. 4).

For a confirmatory analysis, machine learning classifiers using 152 sMRI raw brain characteristics showed poorer performance compared to the corresponding age- and sex-adjusted machine learning classifiers (Supplementary Materials). We also tried to build classifiers to differentiate CHR from HCs or CHR-PS + from CHR-PS-, however, those ones only showed approximate chance level (50%) accuracies.

Predictive performance of the classifier for the test, independent confirmatory, and independent group datasets

A chi-squared test showed a significant difference in the classified labels for the independent confirmatory, and independent group datasets, respectively (X^2(1, n = 151) = 6.34, p = 0.012 and X^2(1, n = 1021) = 4.39, p = 0.036). Further residual analysis showed that the HC group was significantly more likely to be classified as HCs than the CHR-PS + group (73% vs. 30%, *corrected p* = 0.004, Fig. 3B). For the independent group dataset, no difference between CHR-PS- and CHR-UNK groups was found (73% vs. 80%, *corrected p* = 0.029).

For the overall sample, a chi-square test showed a significant difference in the classified labels between the four groups (X^2(3, 1172) = 15.12, p = 0.002). Further residual analysis showed a significant difference in the classified labels between CHR-PS + and the other three groups (*Bonferroni corrected p's* < 0.05; Fig. 3B). For the predict probability, an ANOVA showed a significant difference between the four groups (F = 169.84, p < 0.001). Post-hoc comparisons showed that CHR-PS + group was different from all other groups and that the CHR-PS- group was in between CHR-PS + and HC and CHR-PS- groups (HC > CHR-PS- > CHR-PS+), while the predict probability did not differ between CHR-UNK and HCs (CHR-UNK > CHR-PS- > CHR-PS+; *Bonferroni corrected p*'s < 0.05; Fig. 3C). Although the classifier was built according to the features after controlling for non-linear age effect, a GAM analysis demonstrated that the predict probability was associated with age (F = 11.33, p = 0.003), and differed between CHR-PS + and HC (t = 20.72, p < 0.001), CHR-PS + and CHR-PS- (t = 17.83, p < 0.001), and CHR-PS + and CHR-UNK (t = 17.64, p < 0.001; Fig. 3D). No significant age × group interaction was found in the predict probability.

The estimated decision curve for all comparisons (HCs vs CHR-PS+, CHR-PS + vs CHR-PS-, CHR-PS + vs CHR-UNK and CHR-PS + vs CHR-PS- and CHR-UNK showed the that on the basis of getting MRI measurement to get a prediction from current classifier/model leads to higher benefit to discoverer transition of CHR (Fig. 3E).

Relationship between predict probability and demographic and clinical characteristics

We observed no effects of sex or APSS, BLIPS, or GRDS status, on the predict probability (p > .05). No significant correlations were found between standardized IQ and the predict probability for each group. No significant correlation was found between symptom severity scores and predict probability. No significant difference was found for the antipsychotics use was found among each CHR group (p > 0.05).

Discussion

To the best of our knowledge, the current study is the one of a few to apply a machine learning approach to discriminate HC and CHR-PS + groups in a large multisite sample.[40] To evaluate the classifier, we employed a two-step approach using an independent confirmatory dataset, obtained at a different site and using a different protocol from the ones used to build the classifier; we also used an independent group dataset including CHR-PS- and CHR-UNK groups. Although previous study reported 94% accuracy[40], we have achieved 85% accuracy in the 2-class classification in the training dataset using non-linear adjustment of SA features for age and sex. The patterns of neuroanatomical alterations were also useful in identifying CHR-PS- individuals. Specially, of the CHR groups, the CHR-UNK group was the most likely to be classified as HC by the classifier, than those in other CHR groups, showing no difference in the predict probability from HC.

In this study, we were able to differentiate HC from CHR-PS + group with 85% and 68% accuracy in the training and test sets, respectively. The performance accuracy achieved by the classifier on the independent confirmatory dataset was 73%. In contrast to prior studies,[24,40,41] we successfully built a model with promising predictive performance for new data. Our findings suggest that ComBat is not only useful to increase statistical power[43,45,60] but also crucial for improving the accuracy in building a machine learning model out of multisite data. As expected, the majority of CHR-PS- and CHR-UNK individuals were classified as HCs. Moreover, no significant associations were found between the predict probability and sex or IQ, or antipsychotics use for each CHR group. We suggest that a machine learning classifier trained to identify differences between CHR-PS + and healthy controls may be helpful to identify UHR individuals at risk for conversion.

In line with prior studies of cortical alterations in CHR,[23–25,29,32] we found that the pattern of surface area features, including the superior temporal, insula, superior frontal, superior parietal, fusiform, isthmus of cingulate, and parahippocampal gyri, contributed to identifying CHR-PS + from HCs (Fig. 2A, Fig. 4). CHR individuals who converted or presented with greater clinical symptom within a 2-year follow-up exhibited smaller SA in the rostral anterior cingulate, lateral and medial prefrontal regions, and parahippocampal gyrus.[32] SA is more closely related to volume than cortical thickness,[66] and the volume of the isthmus of cingulate gyrus has been reported to be different in resilient and non-resilient CHR individuals.[25] The neuroanatomical alteration/ deviance pattern of SA found in the current study between HCs and CHR-PS + groups are consistent with findings from other studies, which implicate the volume of superior temporal, frontal and fusiform regions in CHR transitions[15,16,24] and schizophrenia.[17,27,28] Although our initial ENIGMA CHR study showed the difference mainly in the cortical thickness between HC and CHR, the effect size of the group difference was equivalent to almost 8 years of healthy aging.[15] In addition, the age trajectory showed a non-linear pattern and differed by group.[15,67] It is possible that the current study engineered the features that made the differences in SA more prominent, by using GAM to estimate the brain age gap in a non-linear manner. Moreover, as the result of GAM eliminating the non-linear adolescent development of SA in differentiating HCs and CHR-PS+, our classifier achieved promising generalization of predictive performance.

The predict probability given by the classifier based on the neuroanatomical deviance showed significant differences among the HC or CHR-UNK, CHR-PS-, and CHR-PS + groups at baseline (HC, CHR-UNK > CHR-PS- > CHR-PS+; Fig. 2C). The results suggest that predict probability is a useful index allowing us to better

understand how neuroanatomical deviance is associated with psychosis conversion. This further implies that the neuroanatomical deviance was already observed at baseline in CHR-PS + group. Moreover, the predict probability was associated with age (Fig. 3D), in line with previous clinical follow-up studies that older individuals at CHR are with lower rate of psychosis conversion.[68,69] These results suggest that psychosis-related brain characteristics may decrease according to brain development which may effect on the onset of psychosis. Since this study focused on baseline brain features and future prognosis about the onset of psychosis, prospective investigation for follow-up MRI scans as well as clinical severity and social functioning may clarify the course of brain alteration in individuals with CHR along with adolescent development..

Our study has several limitations. First, to harmonize site effects, ComBat was applied to both HC and CHR subjects which by assuming a common covariate model (that is typically preserved by ComBat) might potentially lead to an information leak[70]. However, without traveling subject harmonization, ComBat was considered the most appropriate method for testing a classifier on individual samples from multi-site datasets.[42–44] Second, we could not test the effect of psychosis-by-age interaction on the predict probability as longitudinal MRI data were not available. Longitudinally tracking neuroanatomical changes around the onset of psychosis would offer more detailed information to understand the progressive brain pathology. Third, substance use of cannabis or alcohol was not available for the current study which is reported associated with increased risk of developing depression in young adulthood.[71] Fourth, mainly due to low(er) group sizes for CHR-PS-, we did not explicitly train our classifier to distinguish between CHR-PS + and CHR-PS- -even though this is the actual clinical scenario when CHR individuals presents at baseline and for whom a prediction is sought.

In conclusion, we successfully trained a 2-class XGBoost classifier (HC versus CHR-PS+) and showed promising predictive performance on a multi-site dataset after considering age and sex differences. This classifier successfully identified 73% of CHR-PS- individuals as HC, and further 80% of CHR individuals who were not follow-up for the onset. These results suggest that when considering adolescent brain development, baseline MRI scans for CHR individuals may be helpful to identify their prognosis. Future prospective studies are required about what and how the psychosis-related brain characteristics change according to the adolescent development, and whether the classifier could be helpful in the clinical settings.

Declarations

Conflict of Interest

Acknowledgement

References

1. Fusar-Poli P, Correll CU, Arango C, Berk M, Patel V, Ioannidis JPA. Preventive psychiatry: a blueprint for improving the mental health of young people. *World Psychiatry*. 2021;20(2):200-221. doi:10.1002/wps.20869

2. Miller TJ, McGlashan TH, Rosen JL, et al. Prospective Diagnosis of the Initial Prodrome for Schizophrenia Based on the Structured Interview for Prodromal Syndromes: Preliminary Evidence of Interrater Reliability and Predictive Validity. *Am J Psychiatry*. 2002;159(5):863-865. doi:10.1176/appi.ajp.159.5.863

3. Yung AR, Nelson B, Stanford C, et al. Validation of "prodromal" criteria to detect individuals at ultra high risk of psychosis: 2 year follow-up. *Schizophr Res*. 2008;105(1-3):10-17. doi:10.1016/j.schres.2008.07.012

4. Seiler N, Nguyen T, Yung A, O'Donoghue B. Terminology and assessment tools of psychosis: A systematic narrative review. *Psychiatry Clin Neurosci*. 2020;74(4):226-246. doi:10.1111/pcn.12966

5. Fusar-Poli P, Salazar De Pablo G, Correll CU, et al. Prevention of Psychosis: Advances in Detection, Prognosis, and Intervention. *JAMA Psychiatry*. 2020;77(7):755. doi:10.1001/jamapsychiatry.2019.4779

6. Catalan A, Salazar De Pablo G, Vaquerizo Serrano J, et al. Annual Research Review: Prevention of psychosis in adolescents – systematic review and meta-analysis of advances in detection, prognosis and intervention. *J Child Psychol Psychiatry*. 2021;62(5):657-673. doi:10.1111/jcpp.13322

7. Salazar De Pablo G, Woods SW, Drymonitou G, De Diego H, Fusar-Poli P. Prevalence of Individuals at Clinical High-Risk of Psychosis in the General Population and Clinical Samples: Systematic Review and Meta-Analysis. *Brain Sci*. 2021;11(11):1544. doi:10.3390/brainsci11111544

8. Salazar De Pablo G, Radua J, Pereira J, et al. Probability of Transition to Psychosis in Individuals at Clinical High Risk: An Updated Meta-analysis. *JAMA Psychiatry*. 2021;78(9):970. doi:10.1001/jamapsychiatry.2021.0830

9. Salazar De Pablo G, Soardo L, Cabras A, et al. Clinical outcomes in individuals at clinical high risk of psychosis who do not transition to psychosis: a meta-analysis. *Epidemiol Psychiatr Sci*. 2022;31:e9. doi:10.1017/S2045796021000639

10. Takahashi T, Wood SJ, Yung AR, et al. Progressive Gray Matter Reduction of the Superior Temporal Gyrus During Transition to Psychosis. *Arch Gen Psychiatry*. 2009;66(4):366. doi:10.1001/archgenpsychiatry.2009.12

11. Pantelis C, Velakoulis D, McGorry PD, et al. Neuroanatomical abnormalities before and after onset of psychosis: a cross-sectional and longitudinal MRI comparison. *The Lancet*. 2003;361(9354):281-288. doi:10.1016/S0140-6736(03)12323-9

12. Pantelis C, Yücel M, Bora E, et al. Neurobiological Markers of Illness Onset in Psychosis and Schizophrenia: The Search for a Moving Target. *Neuropsychol Rev*. 2009;19(3):385-398. doi:10.1007/s11065-009-9114-1

13. Takahashi T, Wood SJ, Yung AR, et al. Insular cortex gray matter changes in individuals at ultra-high-risk of developing psychosis. *Schizophr Res*. 2009;111(1-3):94-102. doi:10.1016/j.schres.2009.03.024

14. Kasai K, Shenton ME, Salisbury DF, et al. Progressive Decrease of Left Superior Temporal Gyrus Gray Matter Volume in Patients With First-Episode Schizophrenia. *Am J Psychiatry*. 2003;160(1):156-164. doi:10.1176/appi.ajp.160.1.156

15. ENIGMA Clinical High Risk for Psychosis Working Group, Jalbrzikowski M, Hayes RA, et al. Association of Structural Magnetic Resonance Imaging Measures With Psychosis Onset in Individuals at Clinical High Risk for Developing Psychosis: An ENIGMA Working Group Mega-analysis. *JAMA Psychiatry*. 2021;78(7):753. doi:10.1001/jamapsychiatry.2021.0638

16. Jalbrzikowski M, Hayes RA, Wood SJ, et al. *Thinner Cortex Is Associated with Psychosis Onset in Individuals at Clinical High Risk for Developing Psychosis: An ENIGMA Working Group Mega-Analysis.* Psychiatry and Clinical Psychology; 2021. doi:10.1101/2021.01.05.20248768

17. Zhao Y, Zhang Q, Shah C, et al. Cortical Thickness Abnormalities at Different Stages of the Illness Course in Schizophrenia: A Systematic Review and Meta-analysis. *JAMA Psychiatry*. Published online April 27, 2022. doi:10.1001/jamapsychiatry.2022.0799

18. Cannon TD, Chung Y, He G, et al. Progressive Reduction in Cortical Thickness as Psychosis Develops: A Multisite Longitudinal Neuroimaging Study of Youth at Elevated Clinical Risk. *Biol Psychiatry*. 2015;77(2):147-157. doi:10.1016/j.biopsych.2014.05.023

19. Koutsouleris N, Pantelis C, Velakoulis D, et al. Exploring Links Between Psychosis and Frontotemporal Dementia Using Multimodal Machine Learning: Dementia Praecox Revisited. *JAMA Psychiatry*. 2022;79(9):907. doi:10.1001/jamapsychiatry.2022.2075

20. Dima D, Modabbernia A, Papachristou E, et al. Subcortical volumes across the lifespan: Data from 18,605 healthy individuals aged 3–90 years. *Hum Brain Mapp*. 2022;43(1):452-469. doi:10.1002/hbm.25320

21. Frangou S, Modabbernia A, Williams SCR, et al. Cortical thickness across the lifespan: Data from 17,075 healthy individuals aged 3–90 years. *Hum Brain Mapp*. 2022;43(1):431-451. doi:10.1002/hbm.25364

22. Bethlehem RAI, Seidlitz J, White SR, et al. Brain charts for the human lifespan. *Nature*. 2022;604(7906):525-533. doi:10.1038/s41586-022-04554-y

23. Chung Y, Addington J, Bearden CE, et al. Adding a neuroanatomical biomarker to an individualized risk calculator for psychosis: A proof-of-concept study. *Schizophr Res*. 2019;208:41-43. doi:10.1016/j.schres.2019.01.026

24. Chung Y, Addington J, Bearden CE, et al. Use of Machine Learning to Determine Deviance in Neuroanatomical Maturity Associated With Future Psychosis in Youths at Clinically High Risk. Published online 2018:9.

25. de Wit S, Wierenga LM, Oranje B, et al. Brain development in adolescents at ultra-high risk for psychosis: Longitudinal changes related to resilience. *NeuroImage Clin*. 2016;12:542-549. doi:10.1016/j.nicl.2016.08.013

26. van Erp TGM, Walton E, Hibar DP, et al. Cortical Brain Abnormalities in 4474 Individuals With Schizophrenia and 5098 Control Subjects via the Enhancing Neuro Imaging Genetics Through Meta Analysis (ENIGMA) Consortium. *Biol Psychiatry*. 2018;84(9):644-654. doi:10.1016/j.biopsych.2018.04.023

27. for the ENIGMA Schizophrenia Working Group, van Erp TGM, Hibar DP, et al. Subcortical brain volume abnormalities in 2028 individuals with schizophrenia and 2540 healthy controls via the ENIGMA consortium. *Mol Psychiatry*. 2016;21(4):547-553. doi:10.1038/mp.2015.63

28. for the ENIGMA Consortium, Thompson PM, Jahanshad N, et al. ENIGMA and global neuroscience: A decade of large-scale studies of the brain in health and disease across more than 40 countries. *Transl Psychiatry*. 2020;10(1):100. doi:10.1038/s41398-020-0705-1

29. Rimol LM, Nesvåg R, Hagler DJ, et al. Cortical Volume, Surface Area, and Thickness in Schizophrenia and Bipolar Disorder. *Biol Psychiatry*. 2012;71(6):552-560. doi:10.1016/j.biopsych.2011.11.026

30. Cheon E, Bearden CE, Sun D, et al. Cross disorder comparisons of brain structure in schizophrenia, bipolar disorder, major depressive disorder, and 22q11.2 deletion syndrome: A review of ENIGMA findings. *Psychiatry Clin Neurosci*. 2022;76(5):140-161. doi:10.1111/pcn.13337

31. Koike S, Uematsu A, Sasabayashi D, et al. Recent Advances and Future Directions in Brain MR Imaging Studies in Schizophrenia: Toward Elucidating Brain Pathology and Developing Clinical Tools. *Magn Reson Med Sci*. Published online 2021:rev.2021-0050. doi:10.2463/mrms.rev.2021-0050

32. Chung Y, Allswede D, Addington J, et al. Cortical abnormalities in youth at clinical high-risk for psychosis: Findings from the NAPLS2 cohort. *NeuroImage Clin*. 2019;23:101862. doi:10.1016/j.nicl.2019.101862

33. Shahab S, Mulsant BH, Levesque ML, et al. Brain structure, cognition, and brain age in schizophrenia, bipolar disorder, and healthy controls. *Neuropsychopharmacology*. 2019;44(5):898-906. doi:10.1038/s41386-018-0298-z

34. Koutsouleris N, Riecher-Rössler A, Meisenzahl EM, et al. Detecting the Psychosis Prodrome Across High-Risk Populations Using Neuroanatomical Biomarkers. *Schizophr Bull*. 2015;41(2):471-482. doi:10.1093/schbul/sbu078

35. Yassin W, Nakatani H, Zhu Y, et al. Machine-learning classification using neuroimaging data in schizophrenia, autism, ultra-high risk and first-episode psychosis. *Transl Psychiatry*. 2020;10(1):278. doi:10.1038/s41398-020-00965-5

36. Zhu Y, Nakatani H, Yassin W, et al. Application of a Machine Learning Algorithm for Structural Brain Images in Chronic Schizophrenia to Earlier Clinical Stages of Psychosis and Autism Spectrum Disorder: A Multiprotocol Imaging Dataset Study. *Schizophr Bull*. Published online March 30, 2022:sbac030. doi:10.1093/schbul/sbac030

37. Rozycki M, Satterthwaite TD, Koutsouleris N, et al. Multisite Machine Learning Analysis Provides a Robust Structural Imaging Signature of Schizophrenia Detectable Across Diverse Patient Populations and Within Individuals. *Schizophr Bull*. 2018;44(5):1035-1044. doi:10.1093/schbul/sbx137

38. Li A, Zalesky A, Yue W, et al. A neuroimaging biomarker for striatal dysfunction in schizophrenia. *Nat Med*. 2020;26(4):558-565. doi:10.1038/s41591-020-0793-8

39. Koutsouleris N, Worthington M, Dwyer DB, et al. Toward Generalizable and Transdiagnostic Tools for Psychosis Prediction: An Independent Validation and Improvement of the NAPLS-2 Risk Calculator in the Multisite PRONIA Cohort. *Biol Psychiatry*. 2021;90(9):632-642. doi:10.1016/j.biopsych.2021.06.023

40. Koutsouleris N, Meisenzahl EM, Davatzikos C, et al. Use of Neuroanatomical Pattern Classification to Identify Subjects in At-Risk Mental States of Psychosis and Predict Disease Transition. *Arch Gen Psychiatry*. 2009;66(7):700. doi:10.1001/archgenpsychiatry.2009.62

41. Zarogianni E, Storkey AJ, Borgwardt S, et al. Individualized prediction of psychosis in subjects with an at-risk mental state. *Schizophr Res*. 2019;214:18-23. doi:10.1016/j.schres.2017.08.061

42. Yamashita A, Yahata N, Itahashi T, et al. Harmonization of resting-state functional MRI data across multiple imaging sites via the separation of site differences into sampling bias and measurement bias. Macleod MR, ed. *PLOS Biol*. 2019;17(4):e3000042. doi:10.1371/journal.pbio.3000042

43. Maikusa N, Zhu Y, Uematsu A, et al. Comparison of traveling-subject and ComBat harmonization methods for assessing structural brain characteristics. *Hum Brain Mapp*. 2021;42(16):5278-5287. doi:10.1002/hbm.25615

44. Bayer JMM, Thompson PM, Ching CRK, et al. Site effects how-to & when: an overview of retrospective techniques to accommodate site effects in multi-site neuroimaging analyses.

45. Fortin JP, Cullen N, Sheline YI, et al. Harmonization of cortical thickness measurements across scanners and sites. *NeuroImage*. 2018;167:104-120. doi:10.1016/j.neuroimage.2017.11.024

46. Wood SN. *Generalized Additive Models: An Introduction with R, Second Edition*. Chapman and Hall/CRC; 2017. doi:10.1201/9781315370279

47. Hastie T, Tibshirani R. Generalized additive models. *Stat Sci*. 1986;1(3):297-310. doi:10.1214/ ss/1177013604

48. Pomponio R, Erus G, Habes M, et al. Harmonization of large MRI datasets for the analysis of brain imaging patterns throughout the lifespan. *NeuroImage*. 2020;208:116450. doi:10.1016/j.neuroimage.2019.116450

49. Gurholt TP, Lonning V, Nerland S, et al. Intracranial and subcortical volumes in adolescents with early-onset psychosis: A multisite mega-analysis from the ENIGMA consortium. *Hum Brain Mapp*. 2022;43(1):373-384. doi:10.1002/hbm.25212

50. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ; 2016:785-794. doi:10.1145/2939672.2939785

51. Yung AR, Yung AR, Yuen HP, et al. Mapping the Onset of Psychosis: The Comprehensive Assessment of At-Risk Mental States. :8.

52. Miller TJ, Woods SW, Corcoran CM, Davidson L. SYMPTOM ASSESSMENT IN SCHIZOPHRENIC PRODROMAL STATES. *Psychiatr Q*.:15.

53. Miller TJ, McGlashan TH, Rosen JL, et al. Prodromal Assessment With the Structured Interview for Prodromal Syndromes and the Scale of Prodromal Symptoms: Predictive Validity, Interrater Reliability, and Training to Reliability. *Schizophr Bull*. 2003;29(4):703-715. doi:10.1093/oxfordjournals.schbul.a007040

54. World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. *JAMA*. 2013;310(20):2191. doi:10.1001/jama.2013.281053

55. Desikan RS, Segonne F, Fischl B, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Published online 2006:13.

56. Ching CRK, Gutman BA, Sun D, et al. Mapping Subcortical Brain Alterations in 22q11.2 Deletion Syndrome: Effects of Deletion Size and Convergence With Idiopathic Neuropsychiatric Illness. *Am J Psychiatry*. 2020;177(7):589-600. doi:10.1176/appi.ajp.2019.19060583

57. Sun D, Ching CRK, Lin A, et al. Large-scale mapping of cortical alterations in 22q11.2 deletion syndrome: Convergence with idiopathic psychosis and effects of deletion size. *Mol Psychiatry*. 2020;25(8):1822-1834. doi:10.1038/s41380-018-0078-5

58. for the ENIGMA-Major Depressive Disorder Working Group, Schmaal L, Veltman DJ, et al. Subcortical brain alterations in major depressive disorder: findings from the ENIGMA Major Depressive Disorder working group. *Mol Psychiatry*. 2016;21(6):806-812. doi:10.1038/mp.2015.69

59. for the ENIGMA-Major Depressive Disorder Working Group, Schmaal L, Hibar DP, et al. Cortical abnormalities in adults and adolescents with major depression based on brain scans from 20 cohorts worldwide in the ENIGMA Major Depressive Disorder Working Group. *Mol Psychiatry*. 2017;22(6):900-909. doi:10.1038/mp.2016.60

60. Radua J, Vieta E, Shinohara R, et al. Increased power by harmonizing structural MRI site differences with the ComBat batch adjustment method in ENIGMA. *NeuroImage*. 2020;218:116956. doi:10.1016/j.neuroimage.2020.116956

61. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine Learning in Python. *Mach Learn PYTHON*.:6.

62. Ojala M, Garriga GC. Permutation Tests for Studying Classifier Performance. In: *2009 Ninth IEEE International Conference on Data Mining*. IEEE; 2009:908-913. doi:10.1109/ICDM.2009.108

63. Fusar-Poli P, Rutigliano G, Stahl D, et al. Development and Validation of a Clinically Based Risk Calculator for the Transdiagnostic Prediction of Psychosis. *JAMA Psychiatry*. 2017;74(5):493. doi:10.1001/jamapsychiatry.2017.0284

64. Vickers AJ, Elkin EB. Decision Curve Analysis: A Novel Method for Evaluating Prediction Models. *Med Decis Making*. 2006;26(6):565-574. doi:10.1177/0272989X06295361

65. Pfeiffer RM, Gail MH. Estimating the decision curve and its precision from three study designs. *Biom J*. 2020;62(3):764-776. doi:10.1002/bimj.201800240

66. Winkler AM, Kochunov P, Blangero J, et al. Cortical thickness or grey matter volume? The importance of selecting the phenotype for imaging genetics studies. *NeuroImage*. 2010;53(3):1135-1146. doi:10.1016/j.neuroimage.2009.12.028

67. Wierenga LM, Langen M, Oranje B, Durston S. Unique developmental trajectories of cortical thickness and surface area. *NeuroImage*. 2014;87:120-126. doi:10.1016/j.neuroimage.2013.11.010

68. Fusar-Poli P, Bonoldi I, Yung AR, et al. Predicting Psychosis: Meta-analysis of Transition Outcomes in Individuals at High Clinical Risk. *ARCH GEN PSYCHIATRY*. 2012;69(3):10.

69. Allott K, Chanen A, Yuen HP. Attrition Bias in Longitudinal Research Involving Adolescent Psychiatric Outpatients: *J Nerv Ment Dis*. 2006;194(12):958-961. doi:10.1097/01.nmd.0000243761.52104.91

70. Solanes A, Gosling CJ, Fortea L, et al. Removing the effects of the site in brain imaging machine-learning – Measurement and extendable benchmark. *NeuroImage*. 2023;265:119800. doi:10.1016/j.neuroimage.2022.119800

71. Gobbi G, Atkin T, Zytynski T, et al. Association of Cannabis Use in Adolescence and Risk of Depression, Anxiety, and Suicidality in Young Adulthood: A Systematic Review and Meta-analysis. *JAMA Psychiatry*. 2019;76(4):426. doi:10.1001/jamapsychiatry.2018.4500

Table 1

Table 1. Demographic characteristics of study participants

		HC			CHR			CHR-PS+			CHR-PS-			CHR-UNK	
		N	Female, No. (%)	Age, mean (SD; range)	N	Female, No. (%)	Age, mean (SD; range)	N	Female, No. (%)	Age, mean (SD; range)	N	Female, No. (%)	Age, mean (SD; range)	N	Female, No. (%)
Total															
	Total	1029	438 (43)	22.48 (5.17; 11.30-39.87)	1165	535 (46)	20.78 (4.82; 10.30-39.00)	144	59 (41)	19.85 (4.60; 12.6-35.00)	793	373 (47)	20.83 (4.95; 10.30-39.00)	228	103 (45)
	Training	799	342 (43)	22.07 (5.20; 11.30-39.87)	120	52 (43)	19.76 (4.62; 12.6-35.00)	120	52 (43)	19.76 (4.62; 12.6-35.00)	0	NA	NA	0	NA
	Test	89	29 (33)	22.07 (5.08; 12.90-39.25)	14	4 (29)	19.96 (3.73; 14.00-26.00)	14	4 (29)	19.96 (3.73; 14.00-26.00)	0	NA	NA	0	NA
	Independent confirmatory	141	67 (48)	25.07 (4.22; 18.00-38.00)	10	3 (30)	20.76 (5.72; 14.90-31.40)	10	3 (30)	20.76 (5.72; 14.90-31.40)	0	NA	NA	0	NA
	Independent group	0	NA	NA	1021	476 (47)	20.91 (4.84; 10.30-39.00)	0	NA	NA	793	373 (47)	20.83 (4.95; 10.30-39.00)	228	103 (45)
Site															
	Columbia	9	2 (22)	24.35 (4.05; 19.58-33.09)	17	9 (53)	22.89 (4.96; 14.87-30.69)	3	2 (67)	19.40 (7.32; 14.87-27.84)	14	7 (50)	23.63 (4.32; 15.91-30.69)	0	NA
	Copenhagen	58	29 (50)	24.78 (3.30; 20.00-35.00)	163	86 (53)	24.18 (4.18; 18.00-38.00)	13	5 (38)	23.08 (3.40; 18.00-29.00)	95	54 (57)	24.51 (4.57; 18.00-38.00)	55	27 (49)
	CSU	59	25 (42)	21.49 (3.14; 15.00-30.00)	52	24 (46)	19.48 (5.00; 13.00-35.00)	21	12 (57)	19.43 (5.07; 13.00-35.00)	31	12 (39)	9.52 (5.04; 13.00-30.00)	0	NA
	Glasgow	46	30 (65)	22.80 (3.62; 18.00-32.00)	80	63 (79)	22.24 (4.79; 17.00-34.00)	6	5 (83)	18.50 (1.87; 17.00-22.00)	74	58 (78)	22.54 (4.83; 17.00-34.00)	0	NA
	Heidelberg	33	16 (48)	15.73 (0.91; 14.00-17.00)	22	13 (59)	15.14 (1.08; 14.00-17.00)	0	NA	NA	18	12 (67)	15.11 (1.02; 14.00-17.00)	4	1 (25)
	IDIBAPS	54	35 (65)	15.86 (1.63; 11.30-18.30)	74	49 (66)	15.26 (1.69; 10.30-18.10)	17	11 (65)	15.07 (1.38; 12.60-17.20)	42	28 (67)	15.49 (1.90; 10.30-18.10)	15	10 (67)
	ISMMS	12	5 (42)	27.94 (3.77; 22.79-34.83)	25	13 (52)	23.39 (5.43; 17.11-34.89)	1	0	26.31 (NA)	12	7 (58)	23.63 (5.46; 17.17-34.89)	12	6 (50)
	London	29	10 (34)	24.52 (4.73; 20.00-36.00)	81	25 (31)	22.61 (4.41; 18.00-38.00)	6	0	23.99 (4.41; 18.00-28.92)	62	20 (32)	22.26 (4.62; 18.00-38.00)	13	5 (38)
	Maastricht	38	12 (32)	25.61 (5.68; 18.35-39.25)	48	14 (29)	20.19 (4.10; 12.00-29.00)	6	2 (33)	20.50 (5.17; 15.00-26.00)	25	8 (32)	18.64 (3.44; 12.00-27.00)	17	4 (24)
	MHRC	51	0	22.21 (2.81; 16.07-27.07)	38	0	20.11 (2.50; 16.30-27.59)	3	0	19.02 (2.37; 16.30-20.60)	35	0	20.20 (2.52; 16.72-27.59)	0	NA
	MPRC	20	8 (40)	17.75 (4.31;	30	15 (50)	17.13 (3.25;	3	1 (33)	16.00 (4.36;	9	4 (44)	16.00 (2.78;	18	10 (56)

			(...;12.00-24.00)			(...;12.00-22.00)			(...;13.00-21.00)			(...;12.00-20.00)		
Oslo Region	62	23 (37)	19.90 (3.62; 15.20-29.40)	21	8 (38)	19.88 (3.63; 15.55-29.08)	2	1 (50)	20.01 (5.43; 16.18-23.85)	18	7 (39)	19.85 (3.71; 15.55-29.08)	1	0
Pitt	64	26 (41)	22.70 (5.56; 14.03-38.24)	26	14 (54)	20.75 (5.32; 12.39-35.84)	2	1 (50)	17.05 (0.72; 16.55-17.56)	11	7 (64)	19.52 (6.38; 12.39-35.84)	13	6 (46)
Singapore	53	25 (47)	21.96 (4.18; 14.51-29.84)	100	32 (32)	21.92 (3.57; 14.52-29.77)	11	3 (27)	20.08 (3.06; 14.76-26.50)	88	28 (32)	22.14 (3.60; 14.52-29.77)	1	1 (100)
SNUH	74	24 (32)	21.23 (2.49; 17.00-27.00)	74	19 (26)	20.70 (3.77; 15.00-34.00)	9	3 (33)	22.00 (4.97; 16.00-33.00)	46	11 (24)	20.37 (3.70; 15.00-34.00)	19	5 (26)
Toho	16	8 (50)	23.19 (2.86; 18.00-28.00)	40	28 (70)	23.73 (6.95; 13.00-39.00)	4	3 (75)	19.00 (4.40; 14.00-24.00)	36	25 (69)	24.25 (7.02; 13.00-39.00)	0	NA
Tokyo	25	12 (48)	22.08 (2.84; 16.00-25.00)	39	18 (46)	20.92 (3.50; 14.00-29.00)	3	0	22.67 (4.73; 19.00-28.00)	25	15 (60)	20.48 (3.10; 14.00-27.00)	11	3 (27)
Toronto	39	16 (41)	25.46 (5.21; 18.17-38.24)	27	12 (44)	20.77 (1.84; 18.12-26.65)	0	NA	NA	4	2 (50)	19.70 (1.09; 18.12-20.55)	23	10 (43)
Toyama	141	67 (48)	25.07 (4.22; 18.00-38.00)	78	35 (45)	18.47 (4.06; 12.60-31.40)	10	3 (30)	20.76 (5.72; 14.90-31.40)	68	32 (47)	18.13 (3.69; 12.60-29.80)	0	NA
UCSF	103	43 (42)	23.74 (7.55; 12.82-39.87)	70	33 (47)	19.39 (4.37; 12.39-32.36)	13	5 (38)	21.35 (4.40; 15.85-29.00)	46	22 (48)	19.04 (4.13; 12.39-28.75)	11	6 (55)
Zurich	43	22 (51)	22.23 (5.56; 13.00-36.00)	60	25 (42)	19.20 (4.90; 13.00-35.00)	11	2 (18)	18.55 (3.36; 14.00-24.00)	34	14 (41)	19.15 (5.43; 13.00-35.00)	15	9 (60)

Abbreviations: HC, healthy control; CHR, clinical high risk for psychosis; CHR-PS+, individuals at CHR who developed psychosis later individuals at CHR who developed psychosis later; CHR-PS-, individuals at CHR who did not develop psychosis later; CHR-UNK, could not follow up; SD, standard deviation. Site name abbreviations as follows: Columbia, New York State Psychiatric Institute, Columbia University, New York; Copenhagen, Mental Health Center Copenhagen and CINS, Mental Health Center Glostrup, University of Copenhagen, Copenhagen, Denmark; CSU, Central South University, Changsha, China; Glasgow, Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, Scotland; Heidelberg, Heidelberg University Hospital, Heidelberg, Germany; IDIBAPS, August Pi I Sunyer Biomedical Research Institute, Barcelona, Spain; ISMMS, Icahn School of Medicine at Mount Sinai, New York, New York; London, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom; Maastricht, Maastricht University, Maastricht, the Netherlands; MHRC, Mental Health Research Center Moscow, Moscow, Russia; MPRC, Maryland Psychiatric Research Center, University of Maryland School of Medicine, Baltimore; Oslo region, NORMENT, University of Oslo and Oslo University Hospital, Oslo, Norway; Pitt, University of Pittsburgh, Pittsburgh, Pennsylvania; Singapore, Institute of Mental Health and National University of Singapore, Singapore; SNUH, Seoul National University, Seoul, Republic of Korea; Toho, Department of Neuropsychiatry, Toho University School of Medicine, Tokyo, Japan; Tokyo, Department of Neuropsychiatry, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan; Toronto, Centre for Addiction and Mental Health, University of Toronto, Toronto, Ontario, Canada; Toyama, University of Toyama Graduate School of Medicine and Pharmaceutical Sciences, Toyama, Japan; UCSF, University of California, San Francisco; Zurich, Psychiatric Hospital, University of Zurich, Zurich, Switzerland. Additional site details can be found in the Supplement.
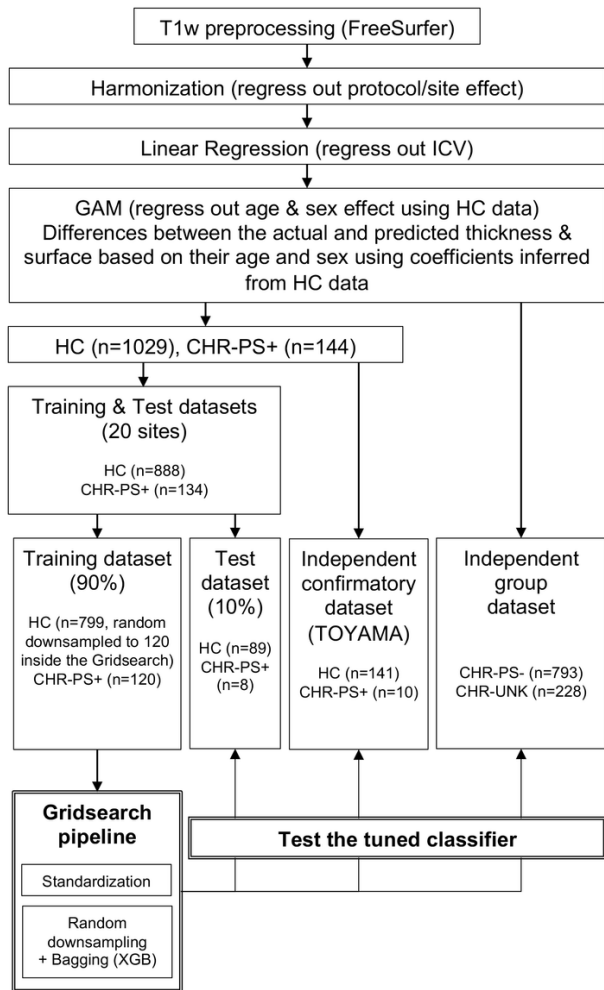
Figures

```
┌─────────────────────────────────────┐
│     T1w preprocessing (FreeSurfer)    │
└─────────────────────────────────────┘
                  ↓
┌─────────────────────────────────────────┐
│ Harmonization (regress out protocol/site effect) │
└─────────────────────────────────────────┘
                  ↓
┌─────────────────────────────────────┐
│   Linear Regression (regress out ICV)  │
└─────────────────────────────────────┘
                  ↓
┌──────────────────────────────────────────────────┐
│   GAM (regress out age & sex effect using HC data)  │
│   Differences between the actual and predicted thickness & │
│  surface based on their age and sex using coefficients inferred │
│                    from HC data                     │
└──────────────────────────────────────────────────┘
                  ↓
┌─────────────────────────────────────┐
│   HC (n=1029), CHR-PS+ (n=144)        │
└─────────────────────────────────────┘
                  ↓
┌─────────────────────────────┐
│   Training & Test datasets     │
│        (20 sites)              │
│                                │
│        HC (n=888)              │
│        CHR-PS+ (n=134)         │
└─────────────────────────────┘
```

Training dataset (90%) — HC (n=799, random downsampled to 120 inside the Gridsearch), CHR-PS+ (n=120)

Test dataset (10%) — HC (n=89), CHR-PS+ (n=8)

Independent confirmatory dataset (TOYAMA) — HC (n=141), CHR-PS+ (n=10)

Independent group dataset — CHR-PS- (n=793), CHR-UNK (n=228)

Gridsearch pipeline
- Standardization
- Random downsampling + Bagging (XGB)

Test the tuned classifier

Figure 1

Diagram employed for the processing and analysis

Abbreviations: HC, healthy control; CHR, clinical high risk for psychosis; CHR-PS+, individuals at CHR who developed psychosis later; CHR-PS-, individuals at CHR who did not develop psychosis later; CHR-UNK, individuals at CHR who could not follow up; SD, standard deviation.
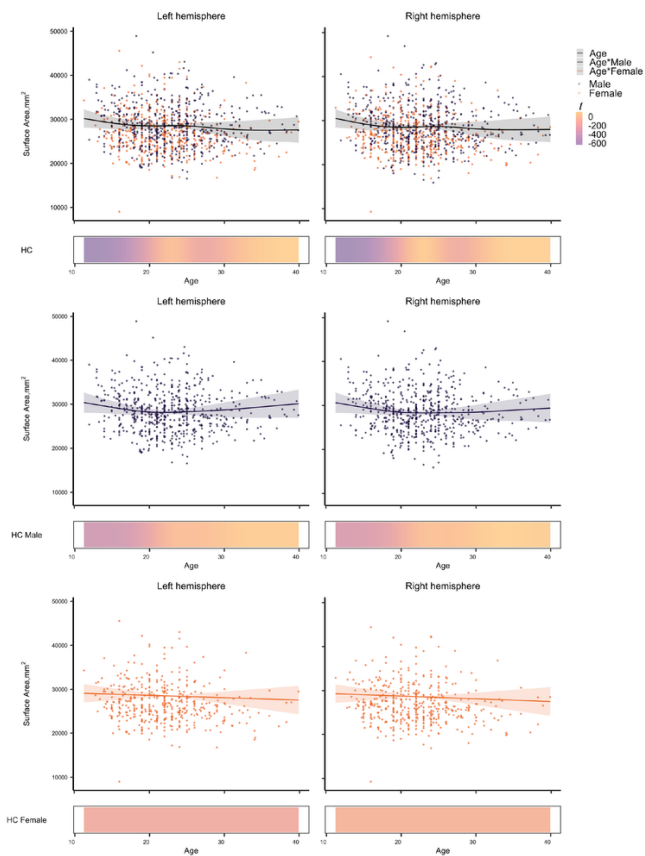
Figure 2

Non-linear age associations of the surface area in healthy controls

Each graph shows a partial effect of the best fit in GAMs. Shading around the line indicates the standard error. The bar underneath the age plots reflects the derivative of the slope.
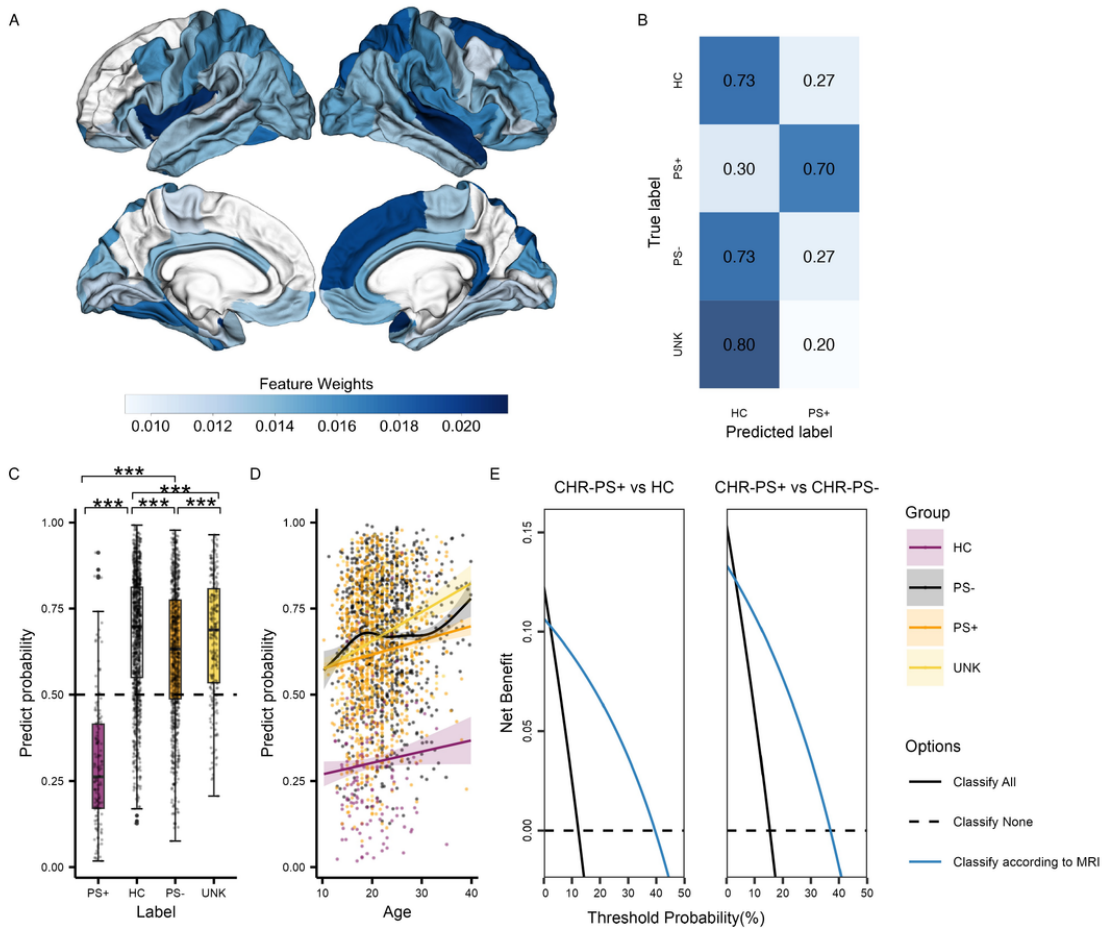
Figure 3

Surface area feature contributions and predictive performance comparisons of the XGBoost classifier

(A) Weighted surface area features of XGBoost classification in Desikan-Killiany atlas. (B) Predictive performance of HC and CHR-PS+ groups was evaluated using the independent confirmatory dataset, and CHR-PS- and CHR-UNK groups using the independent group dataset. (C) Box and scatter plot of predict probabilities of XGBoost. *P*-values of post hoc comparisons were corrected using a Bonferroni method (***$p < 0.001$, **$p < 0.01$, *$p < 0.05$). (D) Best fit for the association of age with the predict probability in a GAM. Shading around the line indicates the standard error. (E) Decision curve analysis showed the benefits of XGBoost predicting the risk of psychosis conversion according to MRI scan.
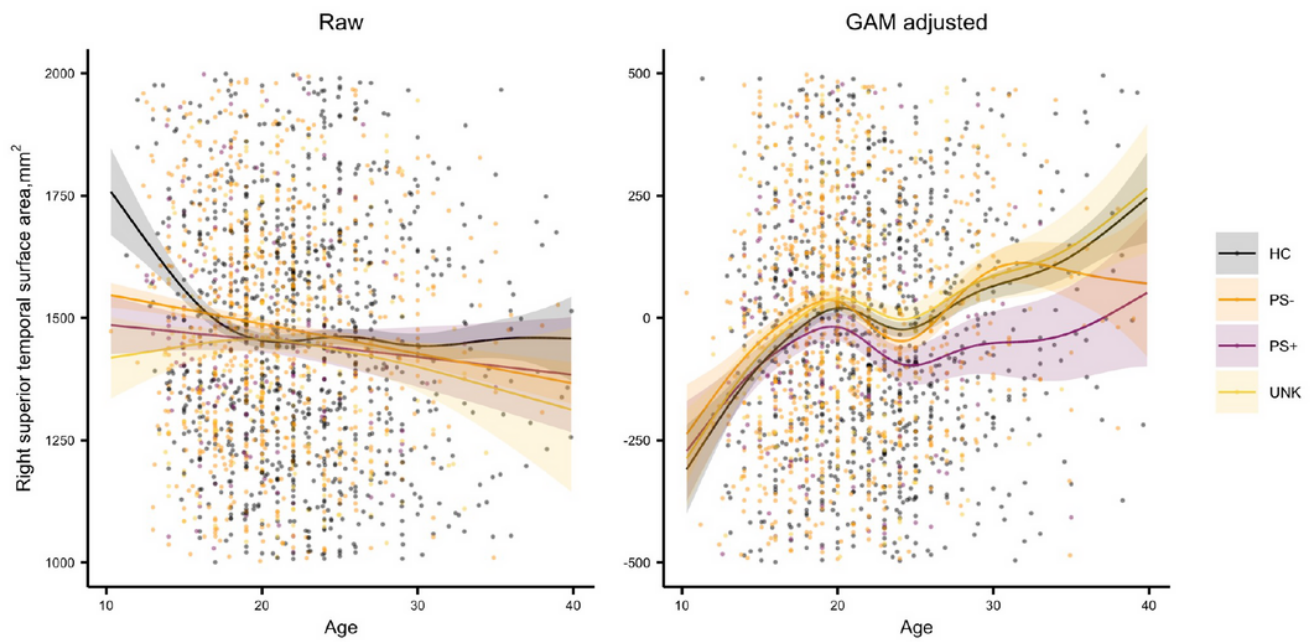
Figure 4

Age association of the surface area in the right superior temporal gyrus

Each graph shows a GAM fit of age, group, and age by group interaction. Shading around the line indicates the standard error.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementfinal.docx
- SupplementaryFigureS1.pdf