Accessible and fair machine learning models for risk prediction of schizophrenia spectrum disorders

Marina Camacho[1*], Polyxeni Gkontra [1], Angélica Atehortúa[1] and Karim Lekadir[1]

[1]*BCN-AIM laboratory. Mathematics and Computer Science Faculty. University of Barcelona. Gran Via de les Corts Catalanes, 585. 08007 Barcelona, Spain*
*Corresponding author: marinacamachosanz@ub.edu (M. Camacho)

Introduction

Machine learning (ML) has emerged as a promising tool for the improvement of mental health in areas such as prevention, diagnosis, treatment, research and administration[1,2]. Despite recent advanced in the field, there remains important gaps in the literature, particularly a lack of evaluation with large samples and external datasets, as well as concerns regard potential bias and discrimination. An example of a mental health domain that is faced with these limitations is the predictive modelling of schizophrenia spectrum disorders (SSZ) using machine learning.

SSZ affect more than 24 million individuals worldwide. They present an acute onset of psychotic symptoms such as delusions, hallucinations, perceptual disturbances, and severe disruption of ordinary behavior which affect the wellbeing of individuals[3]. Despite years of research, the mechanisms leading to the incidence of SSZ remains elusive[4]. A recent review on artificial intelligence-based algorithms for schizophrenia prediction has reported accuracies ranging from 67% to 93.9% (Cortes-Briones et al., 2022)[5]. However, those models were based on electronic health records, electroencephalograms, and genetic data, which are acquired in medical centres using expensive equipment, hence limiting widespread access to such tools by the general population. More recently, using the large longitudinal cohort of UK Biobank, a study on schizophrenia risk prediction based on support vector machines, environmental and genetic data was able to obtain an AUC-ROC of 0.71[6]. To the best of our knowledge, none of the aforementioned studies used accessible predictors, such as exposome factors, that are furthermore potential targets for preventive lifestyle changes. At the same time, none of these studies evaluated the fairness of these algorithms against potential sources of biases and their applicability across multiple population subgroups, including under-disadvantaged populations.

To tackle these limitations, we developed and validated a novel, accessible and fair ML model for risk prediction of SSZ. Based on easily acquired exposome variables, the model enables to identify individuals at risk of schizophrenia, schizotypal and delusional disorders, along with modifiable risk factors. Moreover, we evaluated the potential improvements offered by blood biochemistry and hematology data routinely acquired in clinical practice. We also compared different linear and non-linear methods and analysed the most important factors relevant for the prediction, including previously unknown exposome-related risk factors appropriate for beneficial lifestyle interventions. At the same time, we estimated and demonstrates the fairness of the proposed approach with respect to ethnicity, sex/gender, birth, education and socio-economics. The method was built using large data from the UK Biobank[7], then evaluated using internal and external validation cohorts originating from eighteen and four independent assessment centers, respectively.

Methods

An overview of the study design is provided in Figure 1. In brief, we selected readily available features from the UK Biobank cohort. We included exposome and hematological features such as maternal smoking and vitamin D. Subsequently, we performed data pre-processing, including selection of the study population, data cleaning and imputation. We then trained and evaluated our models using internal and external validation cohorts ensuring fairness for sensible parameters. Finally, we studied the most important features for the model decision to interpret the results and identify potential unknownmodifiable factors and biomarkers for risk assessment.

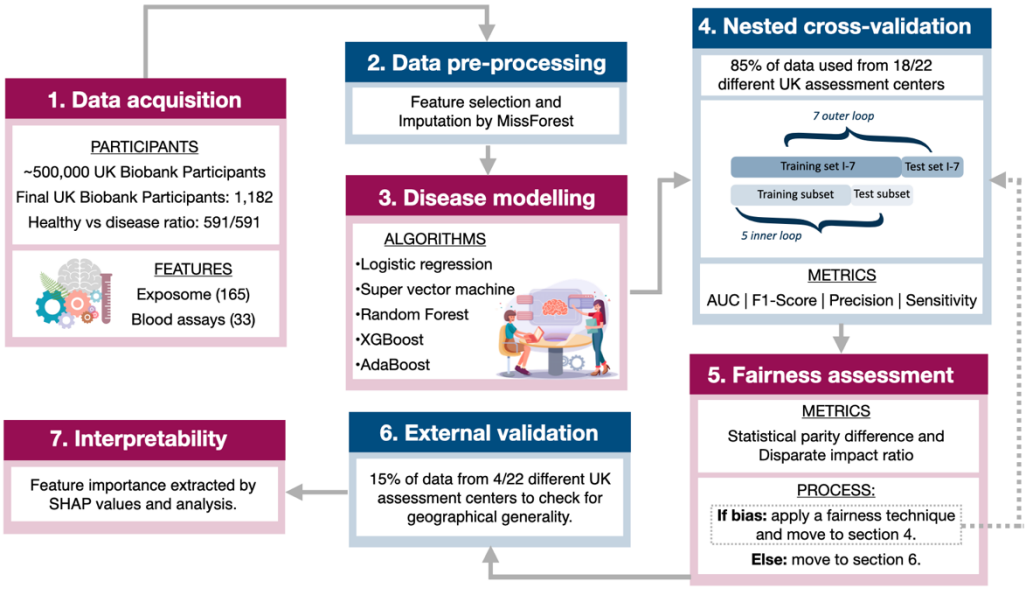

Fig. 1. Overview of our methodology for modeling the risk of schizophrenia, schizotypal and delusional disorders using data from the UK Biobank.

Predictive modeling

We performed nested cross-validation (7 outer folds, 5 inner folds) using the internal validation cohort, and an additional external validation with several state-of-the art machine learning algorithms: Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), AdaBoost (ADA), and XGBoost (XGB). We compared the models' performance using three different sets of input features as predictors: 1) Exposome factors alone, 2) Blood factors alone, 3) Combined blood + exposome factors.

Models' analysis

We evaluated the model performance in the internal and external validation cohorts in terms of sensitivity (ratio of correct classifications), precision (positive predictive value), F1 score (harmonic mean between precision and sensitivity), and AUC (probability of ranking a randomly chosen positive instance higher than a randomly chosen negative one).

Moreover, we evaluated the model fairness by means of statistical parity difference, i.e. the difference of the rate of favourable outcomes received by the majority group, to the protected group, and disparate impact ratio, i.e. the ratio of the proportion of positive outcomes for the majority group over the proportion of positive outcomes for the minority group. For further details see Table 1.

Table. 1. Analyzed privilege and unprivileged groups characteristics.

Class	Ethnicity	Sex	Younger	Older	Education	Low socio-economics	High socio-economics
Group 1	White British (887)	Male (540)	Born before 1944 (246)	Born before 1959 (790)	Education (873)	< -3.241 (266)	> 2.462 (798)
Group 2	Non White British (177)	Female (524)	Born after 1944 (778)	Born after 1959 (242)	No education (191)	> -3.241 (798)	< 2.462 (266)

Results

Figure 2 provides the mean and standard deviation of the evaluation metrics (AUC, F1-score, Precision, Sensitivity) achieved by means of nested cross-validation by each of the selected algorithms: Logistic Regression, Support Vector Machine, Random Forest, AdaBoost and XGBoost.

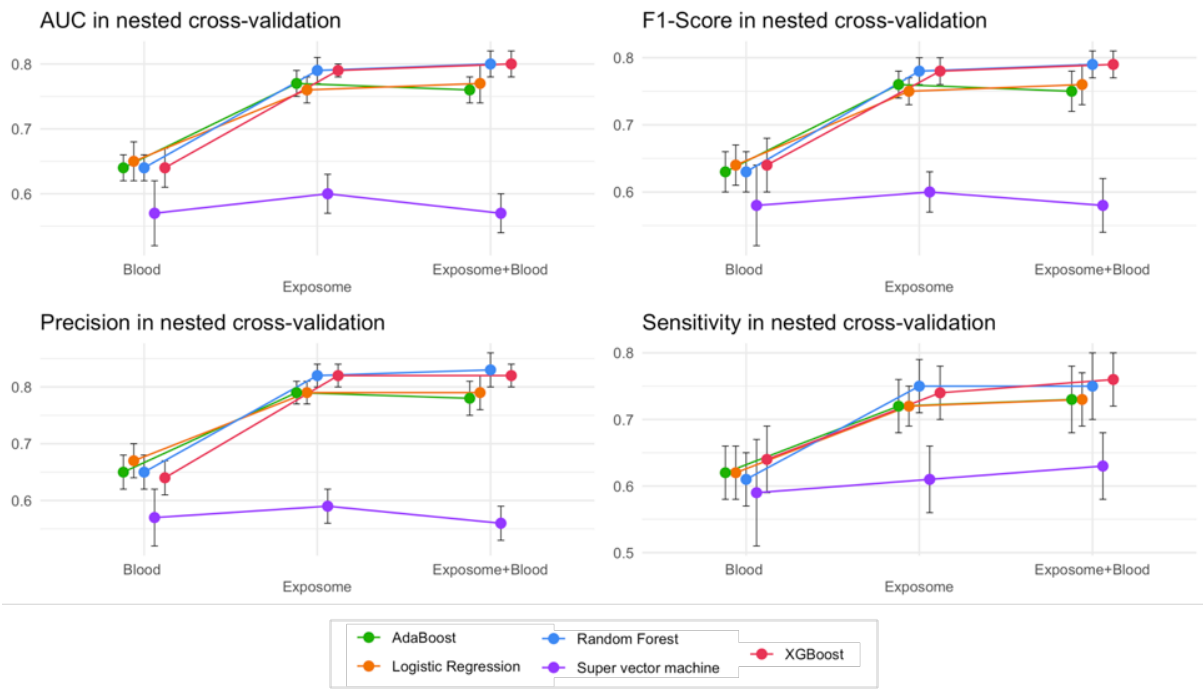

Figure. 2. Performance of all models by metric during nested cross-validation.

The best performance in terms of AUC in the internal validation cohort was obtained using Random Forest and XGBoost models. Both models had similar behaviors when using the three different combination of input features, as depicted in Tables 2 and 3, respectively).

Table. 2. Random Forest mean and standard deviation of the different performance metrics in the internal validation cohort using nested cross-validation.

Model	AUC	F1-Score	Precision	Sensitivity
1) Exposome	0.79±0.02	0.78±0.02	0.82±0.02	0.75±0.04
2) Blood	0.64±0.02	0.63±0.03	0.65±0.03	0.61±0.04
3) Blood+Exposome	0.80±0.02	0.79±0.02	0.83±0.03	0.75±0.05

Table. 3. XGBoost mean and standard deviation of the different performance metrics in the internal validation cohort using nested cross-validation.

Model	AUC	F1-Score	Precision	Sensitivity
1) Exposome	0.79±0.01	0.78±0.02	0.82±0.02	0.74±0.04
2) Blood	0.64±0.03	0.64±0.04	0.64±0.03	0.64±0.05
3) Blood+Exposome	0.80±0.02	0.79±0.02	0.82±0.02	0.76±0.04

For the RF and XGB best performing models, we computed the statistical parity difference and disparate impact ratio to control for potential bias in ethnicity, gender, birth, education, and Townsend deprivation index. For each variable, we report both metrics as the mean of the seven outer folds during nested-cross validation (Figure 3). The results demonstrated that RF and XGB were not discriminating against any of the assessed characteristics during the prediction. Hence, there was no need to apply fairness techniques such as re-weight or disparate impact remover[8].

After verifying algorithm's fairness, we decided to select the best final model considering both the performance in terms of fairness, as assessed by the statistical parity and disparate ratio metrics, and the AUC as a metric reporting the overall performance. According to these criteria, the best final model was the Boost model trained on the second fold during the nested cross-validation using a blend of exposome and haematology data. The model achieved an AUC of 0.822 and 0.796 in the internal and external validation cohorts, respectively.

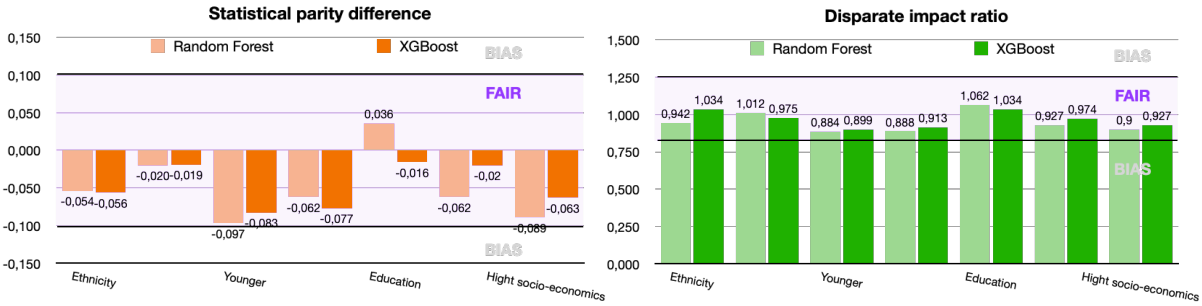

Figure. 3. Statistical parity difference and disparate impact ratio of sociodemographic factors in Random Forest and XGBoost.

The importance of the features of the best performance model is provided in Figure 4. The exposome variables of higher importance in the model were related to sleep patterns, sun protection, changes in diet and socioeconomics. Precisely, the most important one was a visit to the psychiatrist before dealing with SSD, which may be an indicator of other mental health comorbidities. Furthermore, hematological features as hemoglobin also appear to be in the top 20 of the model importance, which supports previous associations of these biomarker and the disorder[9].

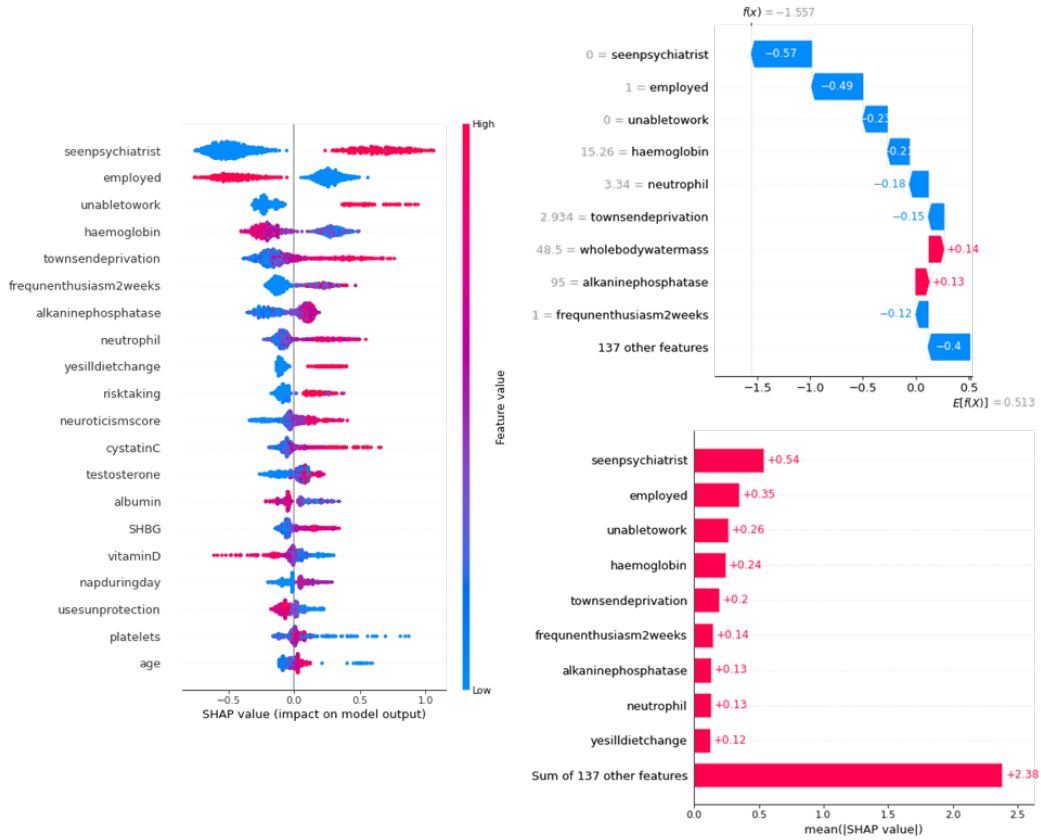

Figure. 4. Feature importance of the best model.

Conclusion

Our results demonstrate that machine learning models based on accessible exposome variables such as Townsend deprivation and diet, can reliably identify individuals at risk of schizophrenia, schizotypal and delusional disorders. Haematological data slightly improve the results in terms of accuracy. For the task at hand, XGBoost outperforms other models with the best fair model achieving an AUC of 0.822 and 0.796 in internal and external validation cohorts, respectively. These preliminary results show promise for further investigation of accessible and fair ML models in mental health that will benefit the general population across various ethnic, sex, age and socio-economics groups.

References

1. Shatte, A. B. R., Hutchinson, D. M. & Teague, S. J. Machine learning in mental health: a scoping review of methods and applications. *Psychol Med* **49**, 1426–1448 (2019).
2. Thieme, A., Belgrave, D. & Doherty, G. Machine Learning in Mental Health. *ACM Transactions on Computer-Human Interaction* **27**, 1–53 (2020).
3. Gaebel, W., Kerst, A. & Stricker, J. CLASSIFICATION AND DIAGNOSIS OF SCHIZOPHRENIA OR OTHER PRIMARY PSYCHOTIC DISORDERS: CHANGES FROM ICD-10 TO ICD-11 AND IMPLEMENTATION IN CLINICAL PRACTICE. *Psychiatr Danub* **32**, 320–324 (2020).
4. Causes - Schizophrenia - NHS. https://www.nhs.uk/mental-health/conditions/schizophrenia/causes/.
5. Cortes-Briones, J. A., Tapia-Rivas, N. I., D'Souza, D. C. & Estevez, P. A. Going deep into schizophrenia with artificial intelligence. *Schizophr Res* **245**, 122–140 (2022).
6. Bracher-Smith, M. *et al*. Machine learning for prediction of schizophrenia using genetic and demographic factors in the UK biobank. *Schizophr Res* **246**, 156–164 (2022).
7. UK Biobank - UK Biobank. https://www.ukbiobank.ac.uk/.
8. Bellamy, R. K. E. *et al*. AI Fairness 360: An extensible toolkit for detecting and mitigating algorithmic bias. *IBM J Res Dev* **63**, 4:1-4:15 (2019).
9. Wysokinski, A. & Szczepocka, E. RED BLOOD CELLS PARAMETERS IN PATIENTS WITH ACUTE SCHIZOPHRENIA, UNIPOLAR DEPRESSION AND BIPOLAR DISORDER. *Psychiatr Danub* **30**, 323–330 (2108).