Identifying Bipolar patients from controls, using post- mortem cerebellum gene expression data and fully Automated Machine Learning.

Identifying bipolar patients using machine learning and cerebellar genetic data

1. Georgios V. Thomaidis, MD [1], 2. Sotirios Michos PhD [2] , 2. Konstantinos Papadimitriou MD [3], 3. Evangelos Chartampilas MD [4] , *Prof*. Ioannis Tsamardinos PhD [2,6].

1. Greek National Health System, Psychiatric Department, Katerini General Hospital, Katerini, Greece 2. Department of Computer Science, University of Crete, Heraclion, Greece 3. Greek National Health System, Psychiatric Hospital of Thessaloniki, Thessaloniki Greece , 4. Laboratory of Radiology, AHEPA General Hospital , University of Thessaloniki, Thessaloniki, Greece, 5. Gnosis DA , Heraclion, Greece.

Abstract

Objective: Complex machine learning classification algorithms using transcriptome data from post-mortem cerebellar tissue of bipolar patients and unaffected controls, have been recently included in pipelines for patient – control classification and identification of characteristic biomarkers. Transcriptomic profile differences between patients and controls, can provide useful information about the role of the cerebellum in the pathogenesis of bipolar disorder and mood deregulation and in normal mood regulation and physiology. User-friendly, fully automated machine learning algorithms, using data extracted from established repositories, could achieve extremely high classification scores and disease- related predictive biomarker identification, in very short time frames and scaled down to small datasets, thus facilitating research on mood disorders.

Method: An application of a fully automated machine learning platform, based on the most suitable algorithm selection and relevant set of hyper- parameter values, for classification between patients and controls and the production of models for biosignature selection, is presented. Transcriptome data used for the analysis were downloaded from the BioDataome preprocessed datasets database. The Dataome dataset, derived from the parent Gene Expression Omnibus GSE35974 (2013) and GSE35978 datasets, which have been originally produced from the cerebellar and parietal lobe tissue of deceased bipolar patients and unaffected controls, (from the Stanley Medical Research Institute's Neuropathology Consortium and Array collections), using Affymetrix Human Gene 1.0 ST Array. Patient and control groups were closely matched for age and sex .

Results: Bipolar patients have been identified from controls based on the cerebellar transcriptomic profile with AUC 0.929 and Average Precision 0.955. Patients and Controls have been classified in two separated groups with no close-to-the-boundary cases. Using 6 of the

characteristic features discovered during the selection process, 99,6% classification accuracy was achieved. The three biomarkers contributing most to the predictive power of the model (92,7%), are also deregulated in temporal lobe epilepsy.

Conclusion: The cerebellar transcriptome of bipolar patients has a discrete profile and can be used for further exploration of the role of this area in health and disease.

93% AUC and 96% Precision were achieved during classification between unaffected controls and patients with Bipolar Disorder.

Introduction

Bipolar disorder (BD) is a mood disorder characterized by unusual fluctuations of mood, thinking, activity and sleep patterns, classified in three major subtypes [1] and presented as a constellation of phenotypes, with a variety of cognitive and behavioral features [2] . It is a highly hereditary disease, running in families, with an early onset, unpredictable course and detrimental impact due to the great risk of fatal self destructive events, long term disability and great financial and social burden, despite existing pharmacological and psychotherapeutic treatment strategies [3]. For these reasons, the neuroanatomy [4] and neurobiology [5] of bipolar disorder are fields of intense research and of paramount importance for 45 million patients globally [6] .

The main functional role of the cerebellum is related to modulation of movement [7]. Beyond this established role, research has linked the cerebellum to emotional, cognitive and affective processing and their disruption in mood disorders [7] , [8]. Structural [9], [10], [11], [12], [13], [14] functional [7], [8], [14], [15], [16], [17], [18] , neurotransmission [19], [20] metabolic [21], [22], [23] and transcriptomic [24], [25], [26] alterations in the cerebellum in BD point to a discreterole in the affected brain networks.

Machine learning, is now gradually used in psychiatry, in order to optimize genetic analysis's results [27], [28], to highlight the most characteristic differences among groups of patients and normal controls and to confirm their importance for diagnostic classification between these groups. These complex classification algorithms, produce genetic signatures, both using living tissue, blood or saliva and postmortem (prefrontal cortex) [28] and SNP (5 studies) [27, Table 1.] and transcriptomic (2 studies) [28], [29] data. In this context, transcriptomic data analysis can contribute greatly in psychiatry [30], and data from the less explored area of the cerebellum, can add new and important signatures in the puzzle of bipolar disorder pathogenesis, progression and potentially treatment response and resistance. The current study is the first where specifically autoML and transcriptomic data from the cerebellum, were used for biomarker identification and patient classification.

We applied a fully automatic machine learning (autoML) platform Just Add Data Bio (JADBIO) platform [31] on public transcriptomic data from previous studies [25] , [26] , which analyzed the transcriptomic profiles of the cerebellum and parietal cortex of post mortem brain tissues.

Patient and control groups were homogenized by tissue sample location (cerebellum), psychiatric diagnosis, sex and age. The autoML system has a simple, user-friendly interface and has been created for direct application on low-sample, high-dimensional databases. It is automatically trained and evaluated (tested), in order to identify highly optimized predictive and classification models, using characteristic biosignature profiles. Provided a specific outcome (ex. diagnosis) and a well-defined set of features (ex. data from transcriptomic, biochemical, neuroimaging, psychometric or symptom intensity measurements), it can provide a minimal subset of predictors (biomarkers), selected from the features, leading to increased predictive power performance. In studies using binary classification (ex. between BD patients and controls, using transcriptomic data in the current study), the classification boundary is defined by the most statistically significant combination of biomarkers (the characteristic biosignature), which identifies patients from controls. System applicability has been tested for diagnostic classification and time to event prediction, producing robust classification, biomarker identification and prediction results (AUCs 85 -95%) using data from oncology, neurology and psychiatry and in international evaluations [32-36].

Aims of the study

Scope of the analysis, is the selection of characteristic transcriptomic biosignatures of bipolar disorder in the cerebellum -using the autoML platform for optimal performance- and a primary interpretation of the results. Information on the features identified, could facilitate the discovery of the genetic networks leading to BD or are altered during disease course and their importance at the local and global brain network level.

Materials and Methods

Data acquisition

For this analysis, publicly available data have been used, from the online BioDataome database [37], which is constructed by uniformly preprocessed, disease-annotated omics data from GEO and RECOUNT databases, based on a uniform preprocessing pipeline [38] . We analyzed the BioDataome csv. which corresponds to the GEO dataset GSE35978, a. containing expression data from the human cerebellum (produced from GSE35974) and parietal cortex, b. from post mortem brain tissue samples, c. extracted from unaffected subjects and schizophrenic, bipolar and depressed patients, d. from the Stanley Foundation Brain Collection [39]. The expression data were obtained by microarray analysis using the “[HuGene-1_0-st] Affymetrix Human Gene 1.0 ST Array [transcript (gene) version]”. The dataset was initially used at the analyses by Chen C et al [25] , [26] .

Data Processing

a. Dataset selection and homogenization

Data have been downloaded in .csv format from the BioDataome database. The preprocessed file includes data for 144 samples from the cerebellum and 168 samples from the parietal cortex. The 144 cerebellum samples include unaffected subjects and patients with bipolar disorder, schizophrenia and depression (SI, information on GSE35974 and GSE35978). From the cerebellum group, all 50 unaffected subjects and 37 bipolar disorder patients (sex: females/ males, age span: 20 - 70) were initially chosen (SI, Tables 1a, 1b). From the initial heterogeneous groups of affected / unaffected, a number of subjects were removed, and two new, smaller groups of affected / unaffected subjects were produced, matched for sex (female / male) and for age. In parallel, we aimed to exceed (as much as the samples allowed), the minimum threshold of 30 subjects per group, required for the machine learning analysis, without an impact on the age and sex between-sample matches and distributions (SI, Tables 2a, 2b). The final dataset includes these two groups, Group A with 35 bipolar patients (18 female +17 male) and Group B with 37 unaffected controls (19 female +18 male). The small size of available data excluded the possibility of a testing after the initial training; this is balanced by the extremely high AUCs produced during the initial (training) analysis. During the initial microarray analysis, a number of transcriptomes is used as controls [25] , [26] . These have been identified, removed from the csv. of the analysis and the final datasheet (Diagnosed Subjects x Features) was produced. The datasets are 2D matrices (features/ genes x diagnosis for any given subject, unaffected or patient).

b. Feature selection and biosignature construction

For the analysis, data were uploaded to JADBIO version 1.4.14 (April 2021) and the binary classification (categorical) functionality of the platform was employed. The classification process is based on the Statistically Equivalent biosignatures (SES) method, with Support Vector Machines, Random Forest, and Penalized Linear Models algorithms. [31] , [35]. For the given 2D matrices, the predicted outcome is diagnosis (Bipolar or Unaffected), and the metric chosen for optimization is the AUC. Preprocessing used Constant Removal Standardization.

Feature selection was performed using LASSO Feature Selection (penalty=0.0, lambda=5.509e-02). The analysis protocol followed has been a repeated 10-fold cross validation without dropping (max. repeats = 20), with 596 configurations, 5760 predictive models trained and 83440 predictive models omitted (total 89200). The chosen predictive algorithm uses Ridge Logistic Regression (with penalty hyper-parameter lambda = 10.0). The overall process applies the Bootstrap Bias Corrected Cross Validation, a protocol for algorithm hyper-parameter tuning during performance estimation and multiple tie adjustment [31] , [35] . The technical analysis report is in SI-Appendix-1.

Results

1. Classification between BD patients and Unaffected Controls

The AutoML classification analysis produced a Ridge Logistic Regression model with high AUC for the positive class bipolar (93%), based on 25 characteristic biomarkers. AUC, and Average precision values and confidence intervals(CIs), ROC curve and main optimized classification threshold dependent metrics for Accuracy / Balanced Accuracy are shown in Image 1.The BD and HC groups are completely separated and coherent in the UMAP plot (Image 2).

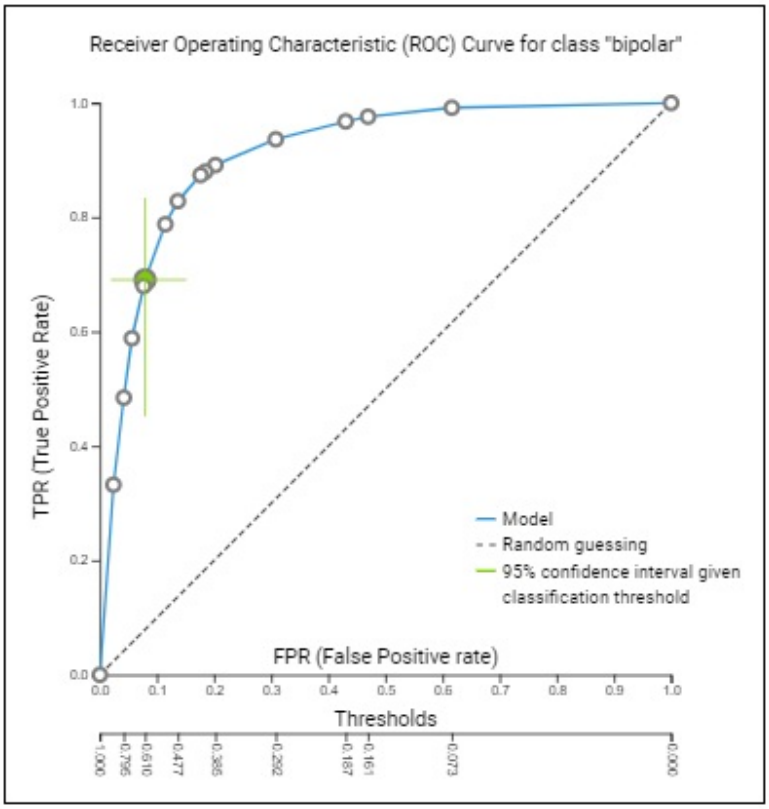

Image 1. Using the best performing model option of in the platform, the AUC for the positive class bipolar is 0.929 (~93%) , with a 95% confidence interval (CI) between 0.868 - 0.977 and average precision 0.955, with a 95% CI 0.914 - 0.986. Accuracy has been calculated at 0.843, Precision at 0.906 and Specificity at 0.921 (full data in SI, Image 1). The classification threshold (0,61) has been optimized and determined for Accuracy / balanced accuracy. Classification as positive is performed when out-of-sample predicted probability is above this given threshold (0,61).

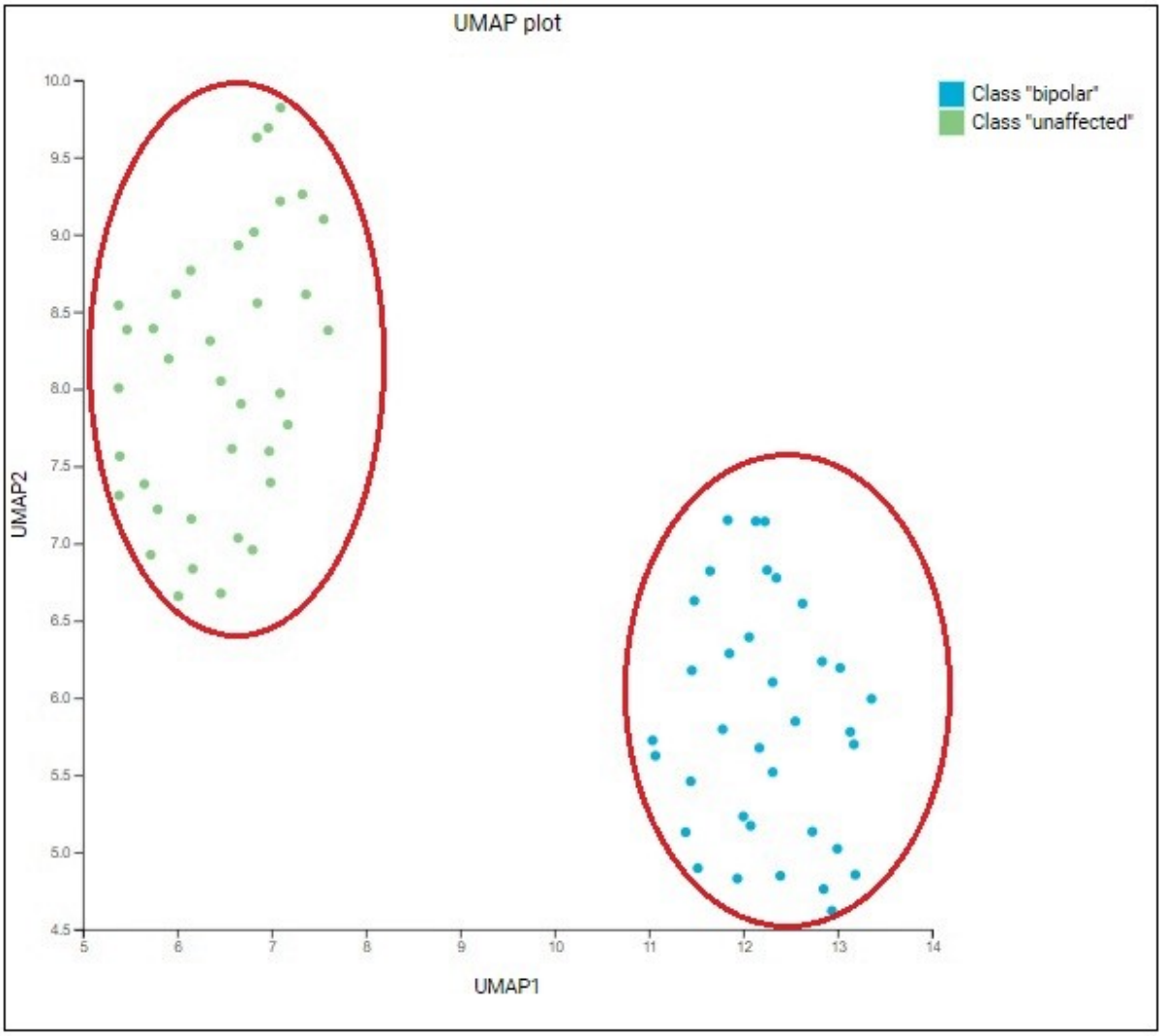

Image 2. Complete separation of BD patients from unaffected controls, in UMAP plots based on all the 25 selected biomarkers. In the Box Plot contrasting the cross-validated predicted probability of belonging to a specific class against the actual class of the samples, the medians are ~0,72 for the class “bipolar” and ~0,18 for the class “unaffected” (SI, Image 2).

2. Biomarker and Biosignature identification

The algorithm selected the most important 25, out of 28869 features (trancriptomes from the genes studied) in the original dataset for the reference signature, used for prediction of BD based on these samples. Inclusion of the 6 most important features (gene transcriptomes from **RNU6-576P, MIR194-2, GDPD5, CARD16, RABGGTA, KREMEN2**), achieves predictive performance (PP) 99,603%. Inclusion of the most important feature RNU6-576P, leads to 76,6% PP, inclusion of the first and second (MIR194-2) most important feature achieves 85,8% PP and

additional inclusion of the third most important feature (GDPD5) achieves 92,7% PP. The progressive feature inclusion plot for 6 most important of 25 identified features is presented in Image 3.

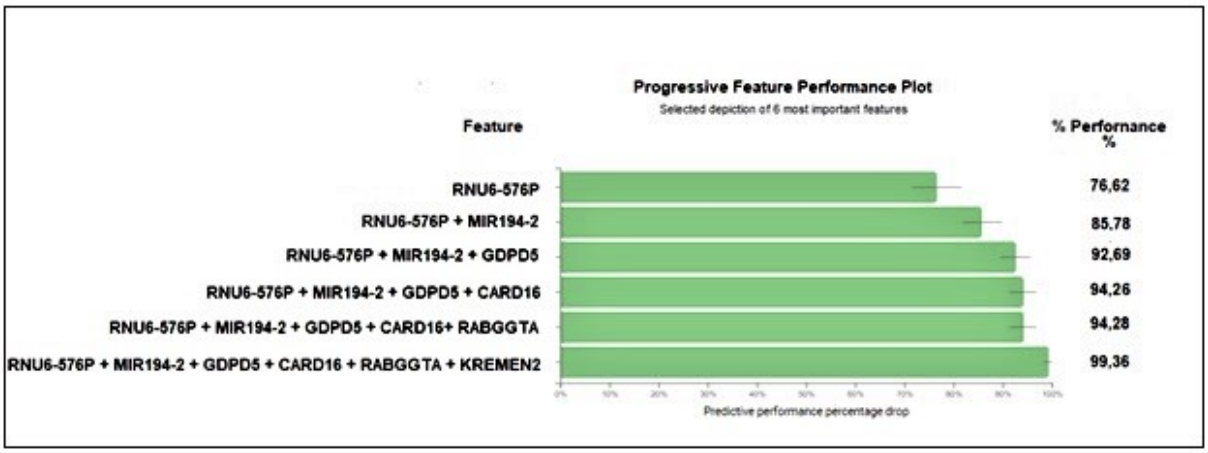

Image 3. Progressive feature inclusion plot. This plot reports the predictive performance (in percentage) that can be achieved by using only part of the features. The features are added one at the time, starting from the most important and ending with the complete signature. Grey lines indicate 95% confidence intervals. In this image the predictive performance of the 6 most important features is presented.

The 21 known or probable functional roles of RNU6-576P, MIR194-2, GDPD5, in the nervous system and CNS diseases, were found after a thorough review of the relevant literature (using and the gene aliases from Gene Cards), which produced 36 articles, and are presented in SI-Appendix 2. The most consistent and important finding is that RNU6-576P, MIR194-2 and GDPD5, have been associated with epilepsy (but not until now with bipolar disorder). Both epilepsy and bipolar disorder are characterized by episodic functional deregulation in the CNS [40], co-occur [41] , share common symptoms and precipitating factors [42] , [43] , their treatment with antiepileptics / mood stabilizers is partially overlapping [44], and potential pathophysiological links have been proposed recently [45] , regarding aberrant neuronal excitation-inhibition related to ANK-3 gene expression. Finally, epileptiform EEG discharges are connected to progress and worse course of disease in BDII patients [46] and manic symptoms are more common in patients with temporal lobe epilepsy [42]. Significantly, alterations in RNU6-576P and MIR194-2 expression are connected to temporal lobe epilepsy [47], [48] , [49] , [50] , [51] , which shares the most common symptoms and pathways with Bipolar Disorders I and II {41] - [46] . Expression of gene RNU6-576P is the most overexpressed small non-coding mRNA in the hippocampus of patients with mesial temporal lobe epilepsy [47] and the most important identifying biomarker in the cerebellum of BD patients in this study.

Discussion:

Main Findings

The classification between the Bipolar and unaffected control groups was completed in <1 hour, with accuracy ~93% and without overlaps between the produced sets of individuals. Welsh t-test for the 6 most important genes, established that the differences in expression between patients with Bipolar Disorder and Unaffected Controls, are statistically meaningful (SI, Image 3). Classification using the JADBIO platform can be considered a reliable means and produces robust results, with potential research interest and physical meaning. The single most important identifier was by far the RNU6-576P small non-coding RNA, accounting for ~77% of total feature importance and also highly deregulated in temporal lobe epilepsy. The role of small non-coding RNAs and pseudogenes is a new area of intense research regarding their role in the onset of psychotic disorders, depression and bipolar disorder [48] , [49]. The role of MIR-194-2 expression in epilepsy has been better studied and a constant pattern of down-regulation has been documented in various epilepsy studies [50] - [53] .

Limitations of the study

The present study was based on a relatively small sample of patients with Bipolar Disorder Types I and II, with an increased analogy of deaths from suicide and was based on post-mortem tissue sampling. Genetic differences between patients with BD I and BD II have been suggested [54] , [55] , [56] , [57] , using family databases, but neuroimaging differences have not been confirmed [58]. The bipolar spectrum is highly heterogeneous, with many different biotypes and their probable neurobiological and functional variance [4] , [59] , [60] , [61] , [62] ; different biotypes can be fully represented only in large samples [4] . The bipolar spectrum includes 7 of the 37 patients with bipolar disorder had committed suicide, a number close to known prevalence of death by suicide in BD. Suicide mainly occurs during the depressive state of the disease and – occasionally- during a manic episode and could be connected to certain patterns of gene expression [61] , biomarker differences [63] and vary during an acute or prolonged depressive or manic or mixed episode) [64] , including the cerebellum. Finally, the genetic characteristics of post- mortem brain tissue sampling could be divergent from the same characteristics of the living brain, in health and disease; still they remain one of the cornerstones of research on the neurobiology of the CNS and its disorders [65] , [66] , [67] [68] .

References

1. *Diagnostic and Statistical Manual of Mental Disorders.*. 5th ed. Washington D.C.: American Psychiatric Association; 2013.

2. Duffy A, Vandeleur C, Heffer N, Preisig M. The clinical trajectory of emerging bipolar disorder among the high-risk offspring of bipolar parents: current understanding and future considerations. *Int J BipolarDisord.*. November 2017;1:37

3. Fountoulakis K, Grunze H, Vieta E, al. e. The International College of Neuro-Psychopharmacology (CINP) Treatment Guidelines for Bipolar Disorder in Adults (CINP-BD-2017), Part 3: The Clinical Guidelines. *Int J Neuropsychopharmacol.*. 2017;2:180-195.

4. Ching C, Hibar D, Gurholt T, al e. What we learn about bipolar disorder from large-scale neuroimaging: Findings and future directions from the ENIGMA Bipolar Disorder Working Group. *HumBrainMapp* :1-27.

5. Charney A, Mullins N, Park Y, Xu J. On the diagnostic and neurobiological origins of bipolar disorder. *TranslPsychiatry*. Apr 2020;1:118.

6. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017. *lancet*. Nov 2018;10159:1789-1858.

7. Wang Y, Wang J, Jia Y, et al. Shared and Specific Intrinsic Functional Connectivity Patterns in Unmedicated Bipolar Disorder and Major Depressive Disorder. *Sci Rep*. 2017;7(1):3570. Published 2017 Jun 15. doi:10.1038/s41598-017-03777-8

8. Adamaszek M, D'Agata F, Ferrucci R, et al. Consensus Paper: Cerebellum and Emotion. *Cerebellum*. 2017;16(2):552-576. doi:10.1007/s12311-016-0815-8

9. Romer AL, Knodt AR, Houts R, et al. Structural alterations within cerebellar circuitry are associated with general liability for common mental disorders. *Mol Psychiatry*. 2018;23(4):1084-1090. doi:10.1038/mp.2017.57

10. Moorhead TW, McKirdy J, Sussmann JE, et al. Progressive gray matter loss in patients with bipolar disorder. *Biol Psychiatry*. 2007;62(8):894-900. doi:10.1016/j.biopsych.2007.03.005

11. Eker C, Simsek F, Yılmazer EE, et al. Brain regions associated with risk and resistance for bipolar I disorder: a voxel-based MRI study of patients with bipolar disorder and their healthy siblings. *Bipolar Disord*. 2014;16(3):249-261. doi:10.1111/bdi.12181

12. Mahon K, Wu J, Malhotra AK, et al. A voxel-based diffusion tensor imaging study of white matter in bipolar disorder. *Neuropsychopharmacology*. 2009;34(6):1590-1600. doi:10.1038/npp.2008.216

13. Redlich R, Almeida JJ, Grotegerd D, et al. Brain morphometric biomarkers distinguishing unipolar and bipolar depression. A voxel-based morphometry-pattern classification approach [published correction appears in JAMA Psychiatry. 2014 Dec 1;71(12):1408]. *JAMA Psychiatry*. 2014;71(11):1222-1230. doi:10.1001/jamapsychiatry.2014.1100

14. Phillips JR, Hewedi DH, Eissa AM, Moustafa AA. The cerebellum and psychiatric disorders. *Front Public Health*. 2015;3:66. Published 2015 May 5. doi:10.3389/fpubh.2015.00066

15. Shinn AK, Roh YS, Ravichandran CT, Baker JT, Öngür D, Cohen BM. Aberrantcerebellarconnectivity in bipolardisorderwithpsychosis. *Biol Psychiatry Cogn Neurosci Neuroimaging*. 2017;2(5):438-448. doi:10.1016/j.bpsc.2016.07.002

16. Wang Y, Zhong S, Jia Y, et al. Interhemispheric resting state functional connectivity abnormalities in unipolar depression and bipolar depression. *Bipolar Disord*. 2015;17(5):486-495. doi:10.1111/bdi.12315

17. Liang MJ, Zhou Q, Yang KR, et al. Identify changes of brain regional homogeneity in bipolar disorder and unipolar depression using resting-state FMRI. *PLoS One*. 2013;8(12):e79999. Published 2013 Dec 4. doi:10.1371/journal.pone.0079999

18. Liu CH, Ma X, Wu X, et al. Regional homogeneity of resting-state brain abnormalities in bipolar and unipolar depression. *Prog Neuropsychopharmacol Biol Psychiatry*. 2013;41:52-59. doi:10.1016/j.pnpbp.2012.11.010

19. Maloku E, Covelo IR, Hanbauer I, et al. Lower number of cerebellar Purkinje neurons in psychosis is associated with reduced reelin expression. *Proc Natl Acad Sci U S A*. 2010;107(9):4407-4411. doi:10.1073/pnas.0914483107

20. Fatemi SH, Stary JM, Earle JA, Araghi-Niknam M, Eagan E. GABAergic dysfunction in schizophrenia and mood disorders as reflected by decreased levels of glutamic acid decarboxylase 65 and 67 kDa and Reelin proteins in cerebellum [published correction appears in Schizophr Res. 2005 May 1;74(2-3):287. Hossein Fatemi, S [corrected to Fatemi, S Hossein]]. *Schizophr Res*. 2005;72(2-3):109-122. doi:10.1016/j.schres.2004.02.017

21. Cecil KM, DelBello MP, Sellars MC, Strakowski SM. Proton magnetic resonance spectroscopy of the frontal lobe and cerebellar vermis in children with a mood disorder and a familial risk for bipolar disorders. *J Child Adolesc Psychopharmacol*. 2003;13(4):545-555. doi:10.1089/104454603322724931

22. Su L, Cai Y, Xu Y, Dutt A, Shi S, Bramon E. Cerebral metabolism in major depressive disorder: a voxel-based meta-analysis of positron emission tomography studies. *BMC Psychiatry*. 2014;14:321. Published 2014 Nov 19. doi:10.1186/s12888-014-0321-9

23. Altamura AC, Bertoldo A, Marotta G, et al. White matter metabolism differentiates schizophrenia and bipolar disorder: a preliminary PET study. *Psychiatry Res*. 2013;214(3):410-414. doi:10.1016/j.pscychresns.2013.08.011

24. McCarthy MJ, Liang S, Spadoni AD, Kelsoe JR, Simmons AN. Whole brain expression of bipolar disorder associated genes: structural and genetic analyses. *PLoS One*. 2014;9(6):e100204. Published 2014 Jun 18. doi:10.1371/journal.pone.0100204

25. Chen C, Cheng L, Grennan K, et al. Two gene co-expression modules differentiate psychotics and controls. *Mol Psychiatry*. 2013;18(12):1308-1314. doi:10.1038/mp.2012.146

26. Chen C, Meng Q, Xia Y, et al. The transcription factor POU3F2 regulates a gene coexpression network in brain tissue from patients with psychiatric disorders. *Sci Transl Med*. 2018;10(472):eaat8178. doi:10.1126/scitranslmed.aat8178

27. Bracher-Smith M, Crawford K, Escott-Price V. Machine learning for genetic prediction of psychiatric disorders: a systematic review. *Mol Psychiatry*. 2021;26(1):70-79. doi:10.1038/s41380-020-0825-2

28. Karthik, S., Sudha, M. Predicting bipolar disorder and schizophrenia based on non-overlapping genetic phenotypes using deep neural network. *Evol. Intel.***14,** 619–634 (2021). https://doi.org/10.1007/s12065-019-00346-y

29. Wang D, Liu S, Warrell J, et al. Comprehensive functional genomic resource and integrative model for the human brain. *Science*. 2018;362(6420):eaat8464. doi:10.1126/science.aat8464

30. Hernandez LM, Kim M, Hoftman GD, et al. Transcriptomic Insight Into the Polygenic Mechanisms Underlying Psychiatric Disorders. *Biol Psychiatry*. 2021;89(1):54-64. doi:10.1016/j.biopsych.2020.06.005

31. Tsamardinos I, Charonyktakis P, Lakiotaki K, et al. Just Add Data: Automated Predictive Modeling and BioSignature Discovery. bioRxiv; 2020. DOI: 10.1101/2020.05.04.075747.

32. Marios Adamou, Grigoris Antoniou, Elissavet Greasidou, Vincenzo Lagani, Paulos Charonyktakis, and Ioannis Tsamardinos. 2018. Mining Free-Text Medical Notes for Suicide Risk Assessment. In Proceedings of the 10th Hellenic Conference on Artificial Intelligence (SETN '18). Association for Computing Machinery, New York, NY, USA, Article 47, 1–8. DOI:https://doi.org/10.1145/3200947.3201020

33. Adamou M, Antoniou G, Greasidou E, et al. Toward Automatic Risk Assessment to Support Suicide Prevention. *Crisis*. 2019;40(4):249-256. doi:10.1027/0227-5910/a000561

34. Karstoft KI, Tsamardinos I, Eskelund K, Andersen SB, Nissen LR. Applicability of an Automated Model and Parameter Selection in the Prediction of Screening-Level PTSD in Danish Soldiers Following Deployment: Development Study of Transferable Predictive Models Using Automated Machine Learning. *JMIR Med Inform*. 2020;8(7):e17119. Published 2020 Jul 22. doi:10.2196/17119

35. Karaglani M, Gourlia K, Tsamardinos I, Chatzaki E. Accurate Blood-Based Diagnostic Biosignatures for Alzheimer's Disease via Automated Machine Learning. *J Clin Med*. 2020;9(9):3016. Published 2020 Sep 18. doi:10.3390/jcm9093016

36. Nissen LR, Tsamardinos I, Eskelund K, Gradus JL, Andersen SB, Karstoft KI. Forecasting military mental health in a complete sample of Danish military personnel deployed between 1992-2013. *J Affect Disord*. 2021;288:167-174. doi:10.1016/j.jad.2021.04.010

37. Lakiotaki K, Vorniotakis N, Tsagris M, Georgakopoulos G, Tsamardinos I. BioDataome: a collection of uniformly preprocessed and automatically annotated datasets for data-driven biology. *Database (Oxford)*. 2018;2018:bay011. doi:10.1093/database/bay011

38. “Download, preprocess, annotate and analyze omics data sets”. Kleanthi Lakiotaki. Computer Science Department, Univeristy of Crete .2017-12-01

39. Torrey EF, Webster M, Knable M, Johnston N, Yolken RH. The stanley foundation brain collection and neuropathology consortium. *Schizophr Res*. 2000;44(2):151-155. doi:10.1016/S0920-9964(99)00192-9

40. Mula M, Marotta AE, Monaco F. Epilepsy and bipolar disorders. *Expert Rev Neurother*. 2010;10(1):13-23. doi:10.1586/ern.09.139

41. Bakken IJ, Revdal E, Nesvåg R, et al. Substance use disorders and psychotic disorders in epilepsy: a population-based registry study. *Epilepsy Res*. 2014;108(8):1435-1443. doi:10.1016/j.eplepsyres.2014.06.021

42. Lyketsos CG, Stoline AM, Longstreet P, et al. Mania in Temporal Lobe Epilepsy, Neuropsychiatry, Neuropsychology & Behavioral Neurology: 1993; 6(1):19-25

43. Bostock ECS, Kirkby KC, Garry MI, Taylor BVM. Systematic Review of Cognitive Function in Euthymic Bipolar Disorder and Pre-Surgical Temporal Lobe Epilepsy. *Front Psychiatry*. 2017;8:133. Published 2017 Aug 9. doi:10.3389/fpsyt.2017.00133

44. Haggarty SJ, Karmacharya R, Perlis RH. Advances toward precision medicine for bipolar disorder: mechanisms & molecules. *Mol Psychiatry*. 2021;26(1):168-185. doi:10.1038/s41380-020-0831-4

45. Lopez AY, Wang X, Xu M, et al. Ankyrin-G isoform imbalance and interneuronopathy link epilepsy and bipolar disorder. *Mol Psychiatry*. 2017;22(10):1464-1472. doi:10.1038/mp.2016.233

46. Drange OK, Sæther SG, Finseth PI, et al. Differences in course of illness between patients with bipolar II disorder with and without epileptiform discharges or other sharp activity on electroencephalograms: a cross-sectional study. *BMC Psychiatry*. 2020;20(1):582. Published 2020 Dec 7. doi:10.1186/s12888-020-02968-4

47. Niu X, Zhu HL, Liu Q, Yan JF, Li ML. MiR-194-5p serves as a potential biomarker and regulates the proliferation and apoptosis of hippocampus neuron in children with temporal lobe epilepsy. *J Chin Med Assoc*. 2021;84(5):510-516.

48. Yoshino Y, Dwivedi Y. Non-Coding RNAs in Psychiatric Disorders and Suicidal Behavior. *Front Psychiatry*. 2020;11:543893. Published 2020 Sep 15. doi:10.3389/fpsyt.2020.543893

49. Barbash S, Simchovitz A, Buchman AS, Bennett DA, Shifman S, Soreq H. Neuronal-expressed microRNA-targeted pseudogenes compete with coding genes in the human brain. *Transl Psychiatry*. 2017;7(8):e1199. Published 2017 Aug 8. doi:10.1038/tp.2017.163

50. Cava C, Manna I, Gambardella A, Bertoli G, Castiglioni I. Potential Role of miRNAs as Theranostic Biomarkers of Epilepsy. *Mol Ther Nucleic Acids*. 2018;13:275-290. doi:10.1016/j.omtn.2018.09.008

51. An N, Zhao W, Liu Y, Yang X, Chen P. Elevated serum miR-106b and miR-146a in patients with focal and generalized epilepsy. *Epilepsy Res*. 2016;127:311-316. doi:10.1016/j.eplepsyres.2016.09.019

52. Li MM, Jiang T, Sun Z, et al. Genome-wide microRNA expression profiles in hippocampus of rats with chronic temporal lobe epilepsy. *Sci Rep*. 2014;4:4734. Published 2014 Apr 22. doi:10.1038/srep04734

53. Mills JD, van Vliet EA, Chen BJ, et al. Coding and non-coding transcriptome of mesial temporal lobe epilepsy: Critical role of small non-coding RNAs. *Neurobiol Dis*. 2020;134:104612. doi:10.1016/j.nbd.2019.104612

54. Gershon ES, Hamovit J, Guroff JJ, et al. A family study of schizoaffective, bipolar I, bipolar II, unipolar, and normal control probands. *Arch Gen Psychiatry*. 1982;39(10):1157-1167. doi:10.1001/archpsyc.1982.04290100031006

55. Heun R, Maier W. The distinction of bipolar II disorder from bipolar I and recurrent unipolar depression: results of a controlled family study. *Acta Psychiatr Scand*. 1993;87(4):279-284. doi:10.1111/j.1600-0447.1993.tb03372.x

56. Sadovnick AD, Remick RA, Lam R, et al. Mood Disorder Service Genetic Database: morbidity risks for mood disorders in 3,942 first-degree relatives of 671 index cases with single depression, recurrent depression, bipolar I, or bipolar II. *Am J Med Genet*. 1994;54(2):132-140. doi:10.1002/ajmg.1320540208

57. McGrath BM, Wessels PH, Bell EC, Ulrich M, Silverstone PH. Neurobiological findings in bipolar II disorder compared with findings in bipolar I disorder. *Can J Psychiatry*. 2004;49(12):794-801. doi:10.1177/070674370404901202

58. Hajek T, Cullis J, Novak T, et al. Brain structural signature of familial predisposition for bipolar disorder: replicable evidence for involvement of the right inferior frontal gyrus. *Biol Psychiatry*. 2013;73(2):144-152. doi:10.1016/j.biopsych.2012.06.015

59. Duffy A, Vandeleur C, Heffer N, Preisig M. The clinical trajectory of emerging bipolar disorder among the high-risk offspring of bipolar parents: current understanding and future considerations. *Int J Bipolar Disord*. 2017;5(1):37. Published 2017 Nov 22. doi:10.1186/s40345-017-0106-4

60. Diaz AP, Bauer IE, Sanches M, Soares JC. Neuroanatomic and Functional Neuroimaging Findings. *Curr Top Behav Neurosci*. 2021;48:173-196. doi:10.1007/7854_2020_174

61. Guzman-Parra J, Streit F, Forstner AJ, et al. Clinical and genetic differences between bipolar disorder type 1 and 2 in multiplex families. *Transl Psychiatry*. 2021;11(1):31. Published 2021 Jan 11. doi:10.1038/s41398-020-01146-0

62. Kung CH, Lee SY, Chang YH, et al. Poorer sustained attention in bipolar I than bipolar II disorder. *Ann Gen Psychiatry*. 2010;9(1):8. Published 2010 Feb 15. doi:10.1186/1744-859X-9-8

63. Sher L, Sublette ME, Grunebaum MF, Mann JJ, Oquendo MA. Plasma testosterone levels and subsequent suicide attempts in males with bipolar disorder [published online ahead of print, 2021 Nov 4]. *Acta Psychiatr Scand*. 2021;10.1111/acps.13381. doi:10.1111/acps.13381

64. Valtonen HM, Suominen K, Haukka J, et al. Differences in incidence of suicide attempts during phases of bipolar I and II disorders. *Bipolar Disord*. 2008;10(5):588-596. doi:10.1111/j.1399-5618.2007.00553.x

65. Lewis DA. The human brain revisited: opportunities and challenges in postmortem studies of psychiatric disorders. *Neuropsychopharmacology*. 2002;26(2):143-154. doi:10.1016/S0893-133X(01)00393-1

66. McCullumsmith RE, Meador-Woodruff JH. Novel approaches to the study of postmortem brain in psychiatric illness: old limitations and new challenges. *Biol Psychiatry*. 2011;69(2):127-133. doi:10.1016/j.biopsych.2010.09.035

67. Harrison PJ. Using our brains: the findings, flaws, and future of postmortem studies of psychiatric disorders. *Biol Psychiatry*. 2011;69(2):102-103. doi:10.1016/j.biopsych.2010.09.008

68. Sullivan K, Pantazopoulos H, Liebson E, et al. What can we learn about brain donors? Use of clinical information in human postmortem brain research. *Handb Clin Neurol*. 2018;150:181-196. doi:10.1016/B978-0-444-63639-3.00014-1