PREDICTING CONVERGENCE FROM MAJOR DEPRESSIVE DISORDER TO BIPOLAR DISORDER UTILIZING ELECTRONIC HEALTH RECORDS AND POLYGENIC SCORES

YAEL BESHAW[1,4], LOTEM EFRAT[1,5], Aditya Pimplaskar [2,3], Loes Olde Loohuis [3]

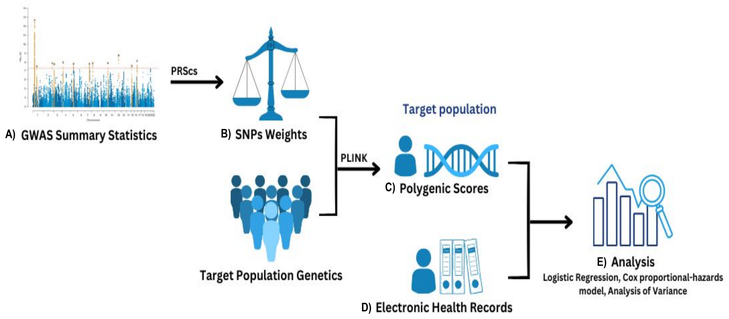
[1] Bruins In Genomics Summer Program, Institute for Quantitative and Computational Biosciences, UCLA, Los Angeles, CA, USA, [2] Bioinformatics Interdepartmental Program, UCLA, Los Angeles, CA, USA, [3] Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, David Geffen School of Medicine, UCLA, Los Angeles, CA, USA
[4] University of Maryland, College Park, College of Computer, Mathematical, and Natural Science, College Park, MD, USA, [5] University of North Carolina at Chapel Hill, Gillings School of Global Public Health, Chapel Hill, NC, USA

INTRODUCTION

- **Bipolar disorder (BD) is a severe mood disorder,** characterized by depressive and manic/hypomanic episodes at a population prevalence of about 1%, with a strong genetic component as indicated by SNP-heritability estimates of 18%[1].
- An estimated 50-75% of patients with BD are frequently **initially diagnosed with major depressive disorder (MDD)** due to an initial onset of a depressive episode [2,3].
- Delayed diagnosis of BD results in ineffective treatment, **increasing patient morbidity and risk of suicidality.**
  - MDD diagnosed BD patients may receive antidepressants instead of mood stabilizers as a first-line treatment, exacerbating symptoms of mood switches to mania [4]. Thus contributing an estimated 9.2 years reduction in expected lifespan, in which about 1 in 5 patients with BD commit suicide [5].
- On average, it takes **7 years between the onset of illness and diagnosis**. This underscores the importance of precision psychiatry approaches in enhancing diagnostic accuracy and treatment efficacy, motivating our project.
- **In this work, we utilize polygenic scores (PGS), estimates of individual's genetic liability to a trait or disease, to predict conversion from MDD to BD.**

MATERIALS AND METHODS

Figure 1: Methods Flowchart



A) **GWAS Summary Statistics:** 13 publicly available summary statistics for 7 psychiatric phenotypes. Implemented standard quality control; excluded duplicate SNPs, retained SNPs with an imputation info score > 0.8 and Minor Allele Frequency (MAF) > 0.01, and excluded complementary pairs (A/T or G/C).

B) **SNPs Weights:** Generated SNP effect sizes for each GWAS with PRS-CS.

C) **Polygenic Scores (PGS):** Computed with PLINK, utilizing SNP Weights and Target Population Genetics from the Paisa Project Cohort [2].
   a) Target Population Genetics; implemented standard quality controls; genotyping rate > 0.99, sample missingness< 0.02, Hardy-Weinberg Equilibrium P>1 x 10^{-6}, & MAF <1%.
   b) Conducted principal components analysis (PCA) and regressed PGS on the first 10 PCs, accounting for any underlying structure and variability in the data.
   c) Z-scored residuals of regression, resulting in our final, normalized PGS.

D) **Electronic Health Records (EHR):** Identified patients diagnosed with Major Depressive Disorder (MDD) as defined by ICD-10 subcodes F32 and/or F33. Refined cohort by excluding individuals < 2 years old, with missing data, with a diagnosis of Schizophrenia (ICD-10 subcode F2), with a diagnosis of Bipolar Disorder (BD) preceding MDD, and individuals without follow-up appointments after initial MDD diagnosis.

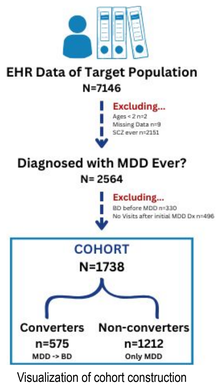
E) **Analysis Plan:** Focusing on PGS, conducted logistic regressions to identify significant predictors of diagnosis, Cox regressions to evaluate predictors of diagnosis conversion and the time to conversion and ANOVA to assess BD PGS differences across subgroups.

Table 1: Summary of Selected GWAS

Trait	Study	Population Ancestry
Suicidality - Psychiatric controls (Sdiag_m21) - General population controls (Sgenpop_m21))	Mullins, 2021	EUR, EAS, AA
Suicide Attempt (SA_d23)	Docherty, 2023	EUR, EAS, AFR, LAT
Schizophrenia (SCZ_t22)	Trubetskoy, 2022	EUR, EAS, AFR, LAT
Schizophrenia (SCZ_l19)	Lam, 2019	EUR, EAS
Major Depression (MDD_a23)	Als, 2023	EUR
Major Depression (MDD_m24)	Meng, 2024	AFR, EAS, SA, LAT
Bipolar Disorder (BD_m21)	Mullins, 2021	EUR
ADHD (ADHD_d23)	Demontis, 2023	EUR
Postpartum Depression - European ancestry (PPD_g23eur) - Trans ancestry meta analysis (PPD_g23tam)	Guintivano, 2023	EUR, EAS, AFR, LAT
Substance Use Disorder - African ancestry (SUD_h23_afr) - European ancestry (SUD_h23_eur)	Hatoum, 2023	EAS, AFR

Table of selected GWAS for PGS generation as discussed in Figure 1a with psychiatric trait of GWAS, Study, and Population Ancestry information included for reference.

Figure 2: Cohort Construction



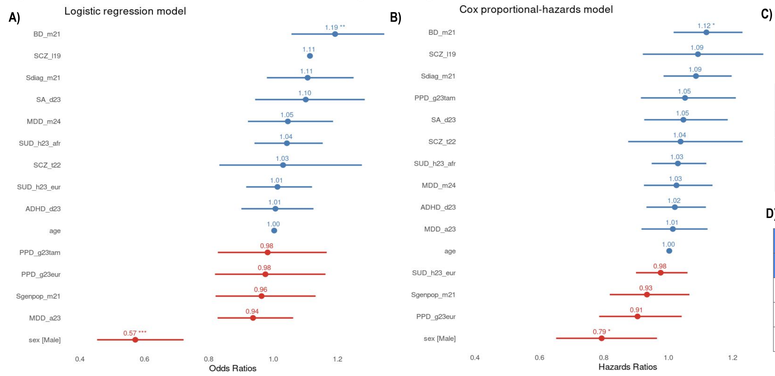
Visualization of cohort construction methodology as outlined in Figure 1d.

RESEARCH QUESTION

Are psychiatric polygenic scores able to accurately predict the conversion to Bipolar Disorder within Major Depressive Disorder patients in the Paisa region of Colombia?

RESULTS

Figure 3: PGS Predictors of Conversion in Logistic Regression and Cox Models



C) PGS evaluated individually: Logistic Regression

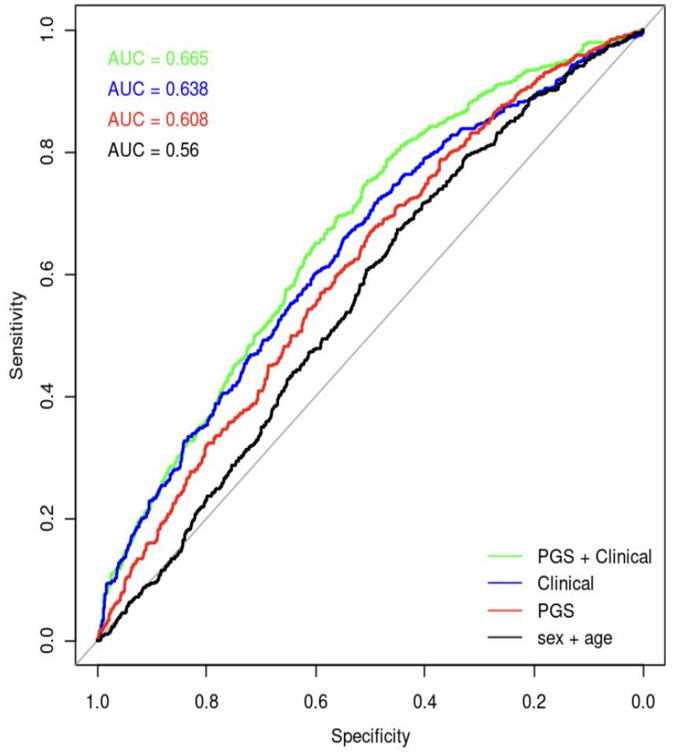
Conversion ~ age + sex + PGS		
PGS	**Odds Ratio**	**p-value**
BD_m21	1.27	9.10e-06
SA_d23	1.18	1.65e-03
SCZ_l19	1.24	7.25e-05
SCZ_t22	1.22	1.57e-04

D) PGS evaluated individually: Cox proportional-hazards

(Conversion, time) ~ age + sex + PGS		
PGS	**Hazard Ratio**	**p-value**
BD_m21	1.17	1.80e-04
SCZ_l19	1.18	1.17e-04
SCZ_t22	1.18	1.91e-04

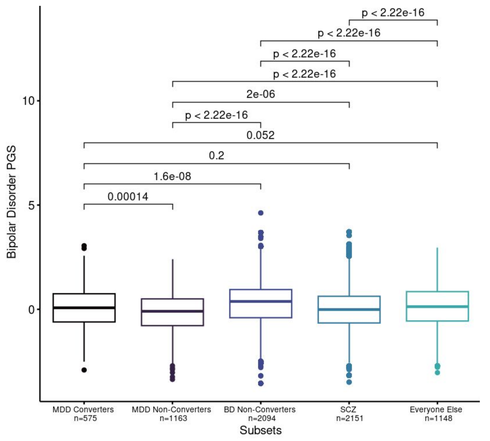
A, B) In both the logistic and Cox regressions, BD PGS and Female sex are found to be significant predictors of conversion and hazard rate of conversion. At a threshold of Bonferroni-corrected p<0.05, no other PGS are significant predictors. **C)** When evaluated individually in a logistic regression model, BD, SCZ and SA PGS are significant predictors of conversion. **D)** BD and SCZ remain significantly associated with conversion in the time-to-event model.

Figure 4: Logistic Regression Models Comparisons



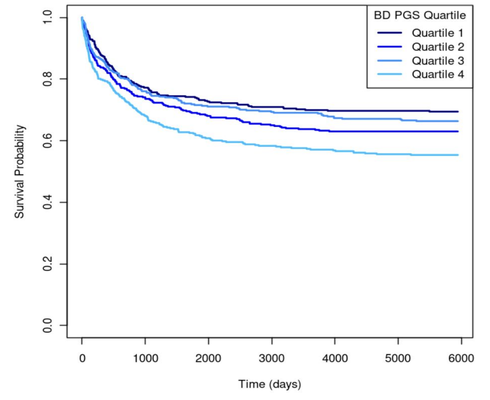
Comparison of logistic regression models by Area under the ROC curve (AUC). In this comparison, integrating both the PGS and Clinical provides the best performing model. We observe that between stand alone Clinical and PGS models, the Clinical model performs slightly better.

Figure 5: Comparison of BD PGS across Subsets



Boxplots depicting distribution of BD PGS scores across subsets of the Paisa Project cohort; MDD Converters and MDD Non-Converters (included in final cohort (Figure 2)), BD Non-Converters, SCZ, and Everyone Else (not included in final cohort). ANOVA testing for differences of means verify that all pairwise differences are significant with the exception of MDD Non-Converters and Everyone Else & MDD Converters and SCZ.

Figure 6: Conversion Survival Curve by BD PGS



Predicted survivor function for a Cox proportional hazards model stratified by BD PGS Quartiles. Median conversion survival times are as follows: Q4 - 1761 days, Q3 - 2599 days, Q2 - 2445 days, and Q1 - 2682 days.

SUMMARY AND DISCUSSION

- **We find that genetics can be leveraged to predict conversion when added to clinical predictors. The model incorporating both clinical data and PGS demonstrated superior performance compared to models with clinical data or PGS alone (Figure 4).**
- Within our cohort, 575 out of 1738 MDD patients converted to BD, resulting in a conversion rate of 33% with mean conversion time of 858 days, or about 2.35 years (Figure 2, Figure 6).
- In addition to BD PGS and female sex, key clinical predictors for BD diagnosis conversion included MDD hospitalization, and specific ICD-10 MDD subcodes: F333 (recurrent, severe with psychotic symptoms), F323 (single episode, severe with psychotic features), and F338 (other recurrent depressive disorders) (Figure 4).
  - These subcodes highlight the importance of MDD severity and psychotic features in relation to conversion.
  - Logistic regression analyses also identified BD, SA and SCZ PGS as significant predictors for the conversion from MDD to BD when evaluated individually (Figure 3C).
- Cox regression analysis corroborate the logistic regression findings and provided additional insights into the hazard rate of conversion. (Figure 3B)
  - Stratified survival analysis revealed that the group with the highest BD genetic liability had a higher conversion rate compared to the group with the lowest BD genetic liability (Figure 6).
- In an ANOVA comparing BD PGS across subsets of the original EHR cohort, differences in means corroborate findings in our logistic regression. Participants with an initial BD diagnosis (BD Non-Converters) have a significantly greater BD PGS than converters (MDD Converters) (Figure 5).
- These results provide insights into genetic risk factors for MDD-BD conversion, with implications for psychiatric precision medicine improvements in patient care.

LIMITATIONS

- Samples in GWAS Summary Statistics and Linkage Disequilibrium References were majority European (Table 1), differing from our South American admixed target data. Future studies aim to investigate results utilizing diversified biobanks, GWAS, and LD references.
- Informative presence bias- presence of a person's information in an EHR is affected by the person's health status- impacted our cohort criteria (excluding those with < 1 visit after initial MDD Dx). This may potentially skew our data towards those with higher healthcare utilization, impacting predictive power and generalizability in our models.

FUTURE DIRECTIONS

- Analyze the genetic and environmental implications of the hospital site for data collection (CSJDM vs HOMO).
- Replicate study with PGS of mania onset BD vs. depression onset BD
- Consider other psychiatric disorders that may be predictive of conversion.
- Stratify population by characteristics (i.e., high vs low PGS) and analyze clinical predictor model performance accordingly.
- New release of genetic data for 50k patients will facilitate well-powered discovery of genetic variants associated with conversion using GWAS

ACKNOWLEDGEMENTS

We acknowledge the UCLA Bruins-in-Genomics Summer Program and it's sponsors; University of California, Los Angeles (UCLA), UCLA Department of Quantitative and Computational Biology, National Science Foundation and National Institutes of Health. We would also like to thank and recognize the participants of and the team behind the Paisa Project for their time and commitment. Lastly, we thank and appreciate our mentor Aditya Pimplaskar, Principal Investigator Dr. Loes Olde Loohuis, and all who are apart of the Loes Lab for their mentorship, guidance, support, and efforts in making this project possible.

REFERENCES

[1] Mullins N, Forstner AJ, O'Connell KS, Coombes B, Coleman JRI, Qiao Z, Als TD, Bigdeli TB, Børte S, Bryois J, Charney AW, Drange OK, Gandal MJ, Hagenaars SP, Ikeda M, Kamitaki N, Kim M, Krebs K, Panagiotaropoulou G, et al. Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. Nat Genet. 2021 Jun;53(6):817-829. doi: 10.1038/s41588-021-00857-4. Epub 2021 May 17. PMID: 34002096; PMCID: PMC8192451.
[2] Service SK, De La Hoz J, Diaz-Zuluaga AM, Arias A, Pimplaskar A, Luu C, Mena L., Valencia J, Ramirez MC, Bearden CE, Sabbab C, Reus VI, López-Jaramillo C, Freimer NB, Lochuis LMO. Predicting diagnostic conversion from major depressive disorder to bipolar disorder: an EHR based study from Colombia. medRxiv [Preprint]. 2023 Oct 2:2023.09.28.23296092. doi: 10.1101/2023.09.28.23296092. PMID: 37873340; PMCID: PMC10593019.
[3] Yang R, Zhao Y, Tan Z, Lai J, Chen J, Zhang X, Sun J, Chen L, Lu K, Cao L, Liu X. Differentiation between bipolar disorder and major depressive disorder in adolescents: from clinical to biological biomarkers. Front Hum Neurosci. 2023 Sep 15;17:1192544. doi: 10.3389/fnhum.2023.1192544. PMID: 37780961; PMCID: PMC10540438.
[4] Nestsiarovich A, Reps JM, Matheny ME, DuVall SL, Lynch KE, Beaton M, Jiang X, Spotnitz M, Pfohl SR, Shah NH, Torre CO, Reich CG, Lee DY, Son SJ, You SC, Park RW, Ryan PB, Lambert CG. Predictors of diagnostic transition from major depressive disorder to bipolar disorder: a retrospective observational network study. Transl Psychiatry. 2021 Dec 20;11(1):642. doi: 10.1038/s41398-021-01760-6. PMID: 34930903; PMCID: PMC8688463.
[5] Depression and Bipolar Support Alliance. Bipolar Depression Statistics. Available from https://www.dbsalliance.org/education/bipolar-disorder/bipolar-disorder-statistics/