Integrative phenomic analysis of psychosocial factors and their genetic associations in major depressive disorder using a latent topic model approach

Michelle Wang[1], Xiangfei Meng[1,2], Kieran J. O'Donnell[1,2,3], Jean Caron[1,2], Yue Li[1]

[1] McGill University, Montreal, QC, Canada; [2] Douglas Research Center, Montreal, QC, Canada;
[3] Yale Child Study Center & Department of Obstetrics Gynecology & Reproductive Sciences, Yale School of Medicine, Yale University, New Haven, CT, USA

McGill initiative in Computational Medicine

CIHR IRSC Canadian Institutes of Health Research / Instituts de recherche en santé du Canada

Introduction

- Major depressive disorder (MDD) is a serious public health issue
- Both genetic and environmental factors contribute to the clinical heterogeneity of MDD (Maj et al., 2020)

Aims:

- To identify meaningful MDD-related latent topics from a set of psychosocial factors
- To explore the etiopathogenesis of MDD by considering both genetic and psychosocial prepositions.

Methods

Data

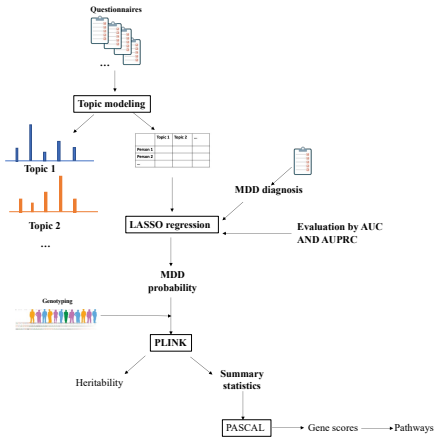
- Zone d'Épidémiologie Psychiatrique du Sud-Ouest de Montréal (ZEPSOM) cohort (Caron et al., 2012)
- 1351 participants
- ~8 million genetic variants (SNPs)
- ~500 psychosocial factors (e.g., clinical MDD diagnosis, psychological distress, drug/alcohol abuse)

Latent topic model

- Collapsed variational Bayesian inference of the LDA model: MixEHR (Li et al., 2020)

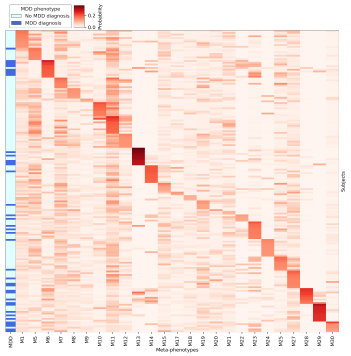
Results

1. General analysis pipeline and latent topic model



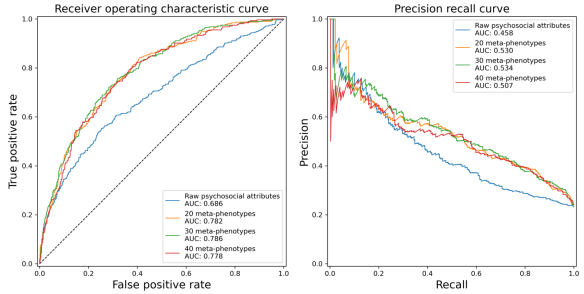
2. Inferring meaningful MDD topics ("meta-phenotypes") from psychosocial questionnaire data

6 meta-phenotypes with strong positive (M6, M22, M13) or negative (M10, M11, M5) correlation with MDD diagnosis

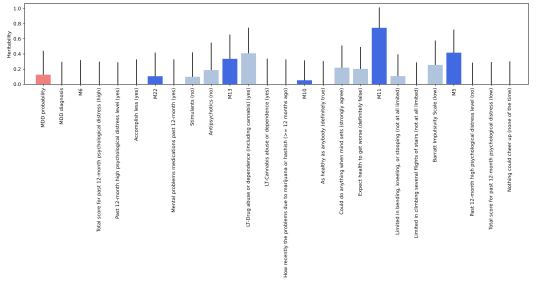


3. Classification of MDD using inferred meta-phenotypes

- Higher accuracy for meta-phenotypes compared to raw psychosocial attributes
- LASSO regression → "MDD probability"



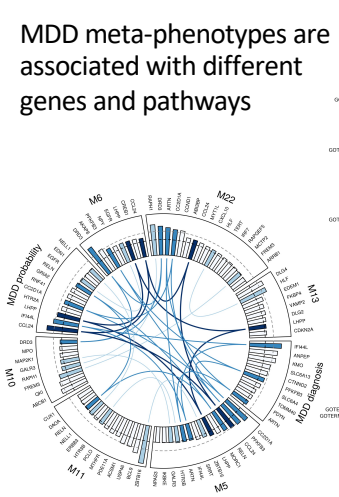
4. SNP-heritability

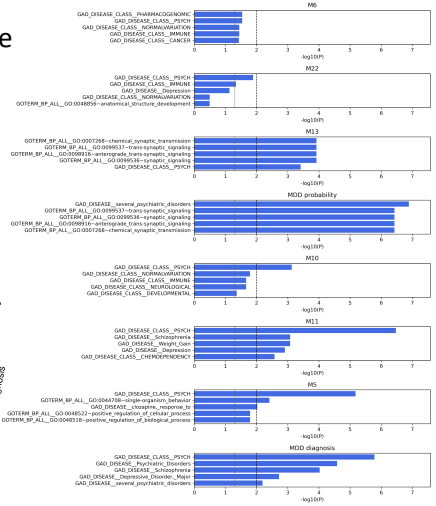


- MDD probability (0.126) > diagnosis (1×10^{-6})
- Some meta-phenotypes have higher heritability estimates than their top-3 psychosocial attributes

5. Gene score and pathway enrichment analysis

MDD meta-phenotypes are associated with different genes and pathways





Conclusion

- Topic modeling approach confers not only accurate MDD risk prediction, but also meaningful associations between genetic and psychosocial prepositions
- Pathway analysis using MDD meta-phenotypes reveals different underlying pathways which can help interpret MDD clinical heterogeneity

References

Caron, J. et al. (2012) Prevalence of psychological distress and mental disorders, and use of mental health services in the epidemiological catchment area of Montreal South-West. *BMC Psychiatry*, 12(1), 183.

Li, Y. et al. (2020) Inferring multimodal latent topics from electronic health records. *Nature Communications*, 11(1), 2536.

Maj, M. et al. (2020) The clinical characterization of the adult patient with depression aimed at personalization of management. *World Psychiatry*, 19, 269-293.