Advancing Precision Psychiatry: Brain Based Predictive Machines for Population-Scale Mental Health Mapping

Principal Investigator: Thomas Wolfers
Co-Principal Investigator: Hao-Chun Yang
HomePage: `https://mhm-lab.github.io`
Email: dr.thomas.wolfers@gmail.com
Phone: +4917684649464
Affiliations: 1) Department of Psychiatry and Psychotherapy 2) German Center for Mental Health partner Tübingen 3) Cluster of Excellence Machine Learning 4) IMPRS-IS, Tübingen, Germany

1 Abstract

Mental health diagnostics often rely on symptom-based assessments, limiting the discovery of predictive markers. This proposal utilizes normative modeling and cutting-edge machine learning techniques to decode individual heterogeneity, predicting brain architecture and function across regions. The goal is to revolutionize mental health diagnostics by offering personalized insights for improved, reliable, and fair patient care. Leveraging extensive reference samples, this research aims to impact population health by unraveling nuanced mental health patterns. The innovative algorithms developed here will address the complexity of mental health data, potentially uncovering novel markers and paving the way for tailored interventions and treatments.

2 Research Statement

Mental disorders, such as schizophrenia, pose complex challenges in diagnosis and treatment due to their multifaceted nature and reliance on symptom assessments [1–4]. This reliance often leads to misdiagnosis, especially among children and minority groups, highlighting the critical need for objective biomarkers in mental health diagnostics [5;6]. Current biomarker research faces limitations stemming from conventional case-control paradigms and supervised modeling techniques. The dominant case-control approach, limited by biased data selection and challenges in establishing causation, restricts the generalizability of findings. Similarly, supervised modeling techniques encounter scarcity or bias in labeled data, potentially leading to overfitting, ethical concerns, and oversimplified representations of mental health complexities. Addressing these limitations requires a comprehensive approach integrating diverse methodologies and high-quality, diverse datasets to ensure accuracy, generalizability, and ethical considerations in predictive models for mental health research. This endeavor not only aims to develop objective biomarkers but also seeks to impact broader mental health populations, emphasizing the collaborative effort needed to advance the field.

3 Objective and Expected Deliverables

3.1 Multi-modal Data Platform for Predictive Modelling

Objective: Addressing the challenges posed by the exponential growth of brain imaging data, our primary objective is to co-develop a robust data resource for biological and brain imaging information. This resource will curate diverse cohorts crucial for advancing modeling technologies, aiming to expand our existing brain imaging repository of over 50,000 scans through collaboration with national and international partners. We have established collaborations with scientist at the Radboud university in Nijmegen, the Netherlands (Prof. Marquand, Prof. Franke, Prof Beckmann,

Prof. Buitelaar, Prof. Greven), the university of Oslo, Norway (Prof. Westlye, Prof. Andreassen), the national institute of mental health and neuroscience in Bangalore, India, and the University of Oxford, UK and many other places. Leveraging expertise and resources from multidisciplinary collaborators, we will efficiently process and annotate new open-source samples, fostering a comprehensive multi-modal platform for predictive modeling in mental health research.

Deliverables: 1) Curated multi-modal biological and brain imaging dataset exceeding 100,000 scans. 2) Streamlined data processing pipeline for efficient and standardized data handling.

Challenges and Mitigation Strategies: 1) Data Heterogeneity: Handling diverse data sources and modalities poses a significant challenge. To address this, our collaborative efforts will focus on harmonizing data formats, establishing robust preprocessing pipelines, and developing fusion techniques to integrate information across modalities seamlessly.

3.2 Development of Representation Learning Methodologies across Modalities

Objective: Expanding representation learning techniques across diverse imaging modalities is paramount to derive nuanced insights into mental health conditions. Our aim is to employ sophisticated unsupervised and self-supervised learning methods, such as Contrastive Predictive Coding (CPC) and Variational Autoencoders (VAEs), to extract meaningful representations from heterogeneous data without reliance on labeled information. Furthermore, integrating cutting-edge anomaly detection techniques into our representation learning frameworks will enable a more comprehensive understanding of disease comorbidity. By identifying and characterizing anomalies within single-modal and multimodal data spaces, we aim to delineate distinct disease subtypes and variations, thereby refining our ability to categorize and predict individual-level responses to interventions.

Deliverables: 1) Enhanced Disease Subtype Characterization: By uncovering subtle anomalies and patterns within multimodal data, our approach will provide deeper insights into the heterogeneity of mental health disorders. This will aid in refining disease subtype categorization, potentially leading to more tailored and effective treatment strategies. 2) Individual-Level Predictive Models: The development of individual-level prediction models based on the comprehensive predictome will empower clinicians with personalized tools for prognosis and treatment planning. These models will consider the complex interplay of various modalities, offering a more holistic view of an individual's mental health status and response to interventions.

Challenges and Mitigation Strategies: 1) Model Interpretability: Interpretability of complex representation learning models is crucial for clinical adoption. We will employ techniques such as attention mechanisms and B-cos[7] frameworks to ensure transparency and enhance the clinical utility of our predictive models.

3.3 Clinical Validation of Longitudinal Inference

Objective: Longitudinal studies play a pivotal role in understanding the trajectory of mental health conditions, especially in assessing short-term changes due to interventions like pharmacological treatments or behavioral therapies. However, existing models often lack the granularity required to capture subtle yet crucial changes occurring within shorter intervals. Our primary objective is to pioneer advanced time series modeling methods, such as graphical models, recurrent neural networks (RNNs), and attention-based models enabling precise and granular longitudinal inference with high temporal resolution. By initially applying these models to the ABCD[8] dataset, a large-scale population-based dataset of children and adolescents followed longitudinally, we will validate and refine our models, paving the way for clinical application and translation.

Deliverables: 1) Improved Disease Progression Estimation: Our refined longitudinal models will

contribute to a deeper understanding of disease progression trajectories, aiding clinicians in better comprehending the course and evolution of various mental health conditions. 2) Individual-Level Longitudinal Changes Post-Intervention: By applying our refined models to assess individual-level longitudinal changes post-intervention, we anticipate providing unparalleled insights into treatment efficacy and individual response patterns. This will significantly enhance pharmaceutical imaging studies and offer personalized interventions, thereby advancing precision medicine in mental health care.

Challenges and Mitigation Strategies: 1) Clinical Translation and Utility: Ensuring clinical utility and ease of integration into clinical practice is critical. We will collaborate closely with healthcare professionals to tailor our longitudinal inference models to their needs, emphasizing interpretability and usability in real-world settings. 2) Ethical Considerations in Longitudinal Data Analysis: Upholding ethical standards in handling longitudinal data requires robust privacy safeguards and adherence to ethical guidelines. We will implement strict protocols to protect patient privacy and ensure responsible data usage throughout the research process.

4   Previous Works

Our prior research endeavors have laid a solid foundation for the proposed study, yielding significant advancements in understanding mental health dynamics[9–11]. Our exploration of age-based normative models unveiled substantial heterogeneity within clinical groups, providing crucial insights into the complexity of mental health conditions. Moreover, recent strides have been made in integrating cognition into our modeling frameworks and linking these cognitive aspects to polygenic risk scores[12]. Scaling our methodologies to encompass populations exceeding 50,000 individuals has further solidified our ability to derive comprehensive insights at a larger scale[13]. These achievements not only signify our expertise in the field but also serve as a springboard for the proposed research. Our collaborations with esteemed partners and the exceptional research environment at the University of Tübingen have been instrumental in these achievements. These partnerships have fostered an environment conducive to innovative research, providing the necessary resources and expertise to ensure the feasibility and excellence of the proposed project.

References

[1] S. Kapur, A. G. Phillips, and T. R. Insel. Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Molecular psychiatry*, 17(12):1174–1179, 2012.

[2] E. Scarr et al. Biomarkers for psychiatry: the journey from fantasy to fact, a report of the 2013 cinp think tank. *International Journal of Neuropsychopharmacology*, 18(10):pyv042, 2015.

[3] K. E. Stephan, S. Iglesias, J. Heinzle, and A. O. Diaconescu. Translational perspectives for computational neuroimaging. *Neuron*, 87(4):716–732, 2015.

[4] P. F. Hitchcock, E. I. Fried, and M. J. Frank. Computational psychiatry needs time and context. *Annual review of psychology*, 73:243–270, 2022.

[5] E. C. Merten, J. C. Cwik, J. Margraf, and S. Schneider. Overdiagnosis of mental disorders in children and adolescents (in developed countries). *Child and adolescent psychiatry and mental health*, 11:1–11, 2017.

[6] M. A. Gara et al. A naturalistic study of racial disparities in diagnoses at an outpatient behavioral health clinic. *Psychiatric Services*, 70(2):130–134, 2019.

[7] M. Böhle, M. Fritz, and B. Schiele. Holistically explainable vision transformers. *arXiv preprint arXiv:2301.08669*, 2023.

[8] B. J. Casey et al. The adolescent brain cognitive development (abcd) study: imaging acquisition across 21 sites. *Developmental cognitive neuroscience*, 32:43–54, 2018.

[9] R. Dinga et al. Evaluating the evidence for biotypes of depression: Methodological replication and extension of. *NeuroImage: Clinical*, 22:101796, 2019.

[10] T. Wolfers et al. Replicating extensive brain structural heterogeneity in individuals with schizophrenia and bipolar disorder. *Human brain mapping*, 42(8):2546–2555, 2021.

[11] S. Rutherford et al. The normative modeling framework for computational psychiatry. *Nature protocols*, 17(7):1711–1734, 2022.

[12] R. Kjelkenes et al. Mapping normative trajectories of cognitive function and its relation to psychopathology symptoms and genetic risk in youth. *Biological Psychiatry Global Open Science*, 3(2):255–263, 2023.

[13] S. Rutherford et al. Charting brain growth and aging at high spatial precision. *elife*, 11:e72904, 2022.