

Brain Based Predictive Machines for 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 diagnosis and treatment primarily rely on symptom-based assessments, hindering the discovery of predictive markers. Leveraging normative modeling and machine learning, this proposal aims to chart the individual heterogeneity, predicting brain architecture/function across regions to enhance precision in psychiatry. Using extensive reference samples and novel modeling techniques, this research seeks to revolutionize mental health diagnostics, offering personalized insights for improved patient care.

2 Research Goals

Mental disorders (e.g. schizophrenia) are complex phenotypes¹⁻⁴, yet their diagnosis and treatment are solely based on the assessment of symptoms²⁻⁵. The dynamic nature of symptom development, coupled with the lack of objective diagnostic tests, can lead to misdiagnosis, particularly in children and minority groups^{6,7}. Consequentially, growing interest in developing reliable quantitative biomarkers for detecting related symptoms has recently emerged. Biomarker studies rely on the dominant case-control paradigm, for which a group of healthy individuals is compared to a group of patients, which generally occludes individual-level differences contributing the development of a complex brain disorders. These disorders do not follow a clear-cut distinction between cases-controls but are more complex in terms of biology and aetiology. Therefore, we develop a set of machine learning methods, inspired by approaches used in paediatrics for which a child's height is charted against age. In our team we have earlier developed methods under the banner of normative modelling, for which we mapped individual differences in reference to a population for more than a hundred thousand different brain imaging-based biomarkers 6-9. We could show that only about 2-10% of the individuals with schizophrenia show extreme deviations in the same brain regions (REF). This finding generalizes to other complex brain disorders and diseases and has been replicated extensively 6,9-11. We recently showed that quantifying deviations in brain structure at the level of the brain network led to an increase of overlap between patients 12, highlighting the possibility that mental disorders are network disorders 13. Here we propose innovation based on ...

Challenges in current researches Data-wise:

1. case-control paradigm: The dominant case-control paradigm, for which a group of healthy individuals is compared to a group of patients, occludes individual-level differences contributing to disorder or disease. Method-wise: supervised machine learning

1. Requires large matched patient and control datasets, which can be challenging and costly to obtain
2. does not account Heterogeneity / comorbidity across different mental disorders

Proposed: Requires exploring a robust and reliable computational method to model individual Heterogeneity across biological factors -i- aiding in the identification of novel illness subtypes and enhancing generalizability to new data and populations without relying on labeled supervision.

3 Research Plan and Expected Outcomes

3.1 Platform and resource for predictive modelling

3.2 Development of normative / unsupervised predictome technology

3.3 Clinical validation, enhanced diagnostic precision and the identification of novel illness subtypes

3.4 Longitudinal Inference

3.5 Fairness: children, under representative

4 Previous Works

During the Emmy Noether Programme I will chart the normative predictome by estimating the architecture or function of one brain region from the collection of other brain regions across modalities for one of the largest reference cohorts assembled to date incorporating relevant covariates and confounds into different machine learning models. This approach extends the use of our earlier developed normative modelling framework and has shown superior performance in preliminary comparisons (Figure 2 A). As a proof of principle we estimated cortical thickness from thalamic volume in a small sample of 500 individuals and detected a plausible predicted pattern (Figure 2 B) that mimics the established relationship between the cortex and different thalamic nuclei 14. To develop this approach further, I will pursue the following strategy. First, I will assemble a large-scale database for model development in close collaboration with my local, national, and international partners. Second, I will implement machine learning methods charting the normative predictome. More concretely, I will estimate the cortical predictome for which I will predict the structure of the cortex from different subcortical brain regions. In addition, I will develop a method to estimate the lateral predictome, for which I will predict the right side of the brain from a multivariate estimate of the left and vice versa improving current univariate approaches of laterality assessment. Furthermore, I will estimate the multimodal predictome for which I will predict one imaging modality from another. Third, I will validate the predictome based normative models against clinical phenotypes. The availability of a large database allows for extensive validations against different disorders and diseases. Concretely, the estimated predictomes will be used to map inter-individual differences of people with the same illness. Fourth, I will translate our models to estimate longitudinal change. Clinical trials are generally designed to detect effects of treatments that occur in terms of months, weeks, hours, minutes or even seconds. The approach proposed here will be able to capture predictive patterns between brain regions at high temporal precision. The deeply phenotyped ABCD dataset 15, which is a large scale population based dataset of children and adolescents which are consulted about every two years, is an excellent resource for initial longitudinal validations. Further, data from clinical trials gathered by partnering consortia will be included which contain clinical information (e.g., TOP, EU-AIMS, NeuroImage).

My previous work provides an outstanding basis for the here proposed research 9,16,17. Using age-based normative models I have shown a high degree of heterogeneity in clinical groups (Figure 1). Recently, my colleagues and I modelled cognition within the normative modelling framework and linked this to polygenic risk scores 18, and worked on scaling age-based normative models to populations of more than 50k individuals 19, developed deep learning technology for normative modelling in the context of aging 20 and integrated it into the PCNtoolkit, a python based statistics and machine learning toolkit for normative modelling 16. Currently, I focus my work on deep covariate embedding, for which we use the abstraction layers of a neural network as embedding space for a normative model. This architecture allows us to identify individual level profiles of deviation from minimally preprocessed brain imaging data (Figure 3). Combined this work forms the basis for my research in Tübingen on the normative predictome. My background and expertise in relevant research domains in combination with excellent partners and an embedding in an outstanding research environment at the University of Tübingen as well as the availability of large scale datasets contributes to the feasibility and excellence of the proposed work.

References

- [1] Thomas R Insel et al. Disruptive insights in psychiatry: transforming a clinical discipline. *The Journal of clinical investigation*, 119(4):700–705, 2009.
- [2] Shitij Kapur, Anthony G Phillips, and Thomas 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.

- [3] Klaas E Stephan, Sandra Iglesias, Jakob Heinzle, and Andreea O Diaconescu. Translational perspectives for computational neuroimaging. *Neuron*, 87(4):716–732, 2015.
- [4] Peter F Hitchcock, Eiko I Fried, and Michael J Frank. Computational psychiatry needs time and context. *Annual review of psychology*, 73:243–270, 2022.
- [5] Elizabeth Scarr, Mark J Millan, Sabine Bahn, Alessandro Bertolino, Christoph W Turck, Shitij Kapur, Hans-Jürgen Möller, and Brian Dean. 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.
- [6] Eva Charlotte Merten, Jan Christopher Cwik, Jürgen Margraf, and Silvia Schneider. Overdiagnosis of mental disorders in children and adolescents (in developed countries). *Child and adolescent psychiatry and mental health*, 11:1–11, 2017.
- [7] Michael A Gara, Shula Minsky, Steven M Silverstein, Theresa Miskimen, and Stephen M Strakowski. A naturalistic study of racial disparities in diagnoses at an outpatient behavioral health clinic. *Psychiatric Services*, 70(2):130–134, 2019.