Identifying Biotypes of OCD Using Data-Driven and Theory-Based Brain Network Models

Obsessive-compulsive disorder (OCD) is a highly heterogeneous psychiatric condition with clinical variability complicating both diagnosis and treatment. While the theory-driven approach relies on predefined hypotheses about specific brain networks and regions implicated in OCD, ensuring that the derived biotypes align with established theories of brain network functions. Data-driven computational approaches hold promise for identifying unbiased, biologically informed subtypes, or biotypes, that could enhance diagnostic precision and guide personalized interventions. In this study, we compare two data-driven approaches – denoising variational autoencoders (DVAEs) – based models (normative modeling and semi-supervised learning) with one theory-based approach to identify OCD biotypes and evaluate their clinical utility, particularly for diagnosis and treatment responsiveness.

We analyzed resting-state fMRI (rs-fMRI) data from the ENIGMA-OCD working group, comprising 2458 adult participants (1277 OCD patients, 1181 healthy controls). For theory-based approach, we extracted brain circuit scores grounded in a theory-based taxonomy (Williams et al., 2022). For data-driven approaches, we first trained a DVAE normative model exclusively on healthy controls to learn normative representations, then applied the trained model on all subjects to derive deviation scores that quantify how much each subject diverged from normative patterns. We then trained a semi-supervised DVAE model on OCD subjects alone, incorporating treatment outcome measured by pre- to post-cognitive behavior therapy (CBT) Y-BOCS changes to guide the latent space toward treatment-responsive features. Feature sets from each model was evaluated for 1) HC vs. OCD classification using Support Vector Machine (SVM) and Random Forest (RF), and 2) unsupervised clustering to identify OCD biotypes. We examined clinical utility and demographic characteristics (including medication status and treatment responsiveness) across biotypes, interpreted the deep learning models using feature-space visualization techniques (Zhu et al., 2024), and investigated the underlying biological network connectivity of the identified subgroups.

The theory-based approach yielded biotypes with significantly distinct treatment response ($p<0.05$), but the classifiers using brain circuit scores failed to discriminate OCD from HC (AUC≈0.50). In contrast, the DVAE-based normative modeling approach effectively captured healthy brain patterns, and the resulting deviation scores enabled robust classification performance between OCD and HC (AUC≈0.90). Although the biotypes derived from this normative model exhibit notable differences in symptom severity ($p=0.06$), they showed no significant variations in treatment responsiveness. Lastly, the semi-supervised DVAE resulted in biotypes that differed significantly in treatment outcomes

($p<0.05$), demonstrating the effectiveness in integrating relevant clinical information into deep learning models for specific clinical goals.
Discussion: Our findings highlight the complementary strengths of theory-driven and data-driven approaches in characterizing OCD. While the theory-based method and the semi-supervised DVAE both revealed biotypes with distinct treatment responses, the theory-based approach alone could not differentiate OCD from HC. Conversely, the DVAE normative modeling effectively classified OCD vs. HC but did not capture differences in treatment responsiveness. Taken together, our results suggest that DVAE model is an effective dimensionality reduction method when dealing with high-dimensional imaging data and can be flexibly adapted to specific clinical goals. Ultimately, the resulting models and features pave the way for more precise diagnosis and individualized treatment strategies, advancing the field toward personalized mental healthcare.