R01 MH133840 (PI: Zhu) 04/01/2025-3/31/2030 (funded)

Abstract: Current psychiatric diagnoses rely on categorical classifications that may not fully capture the shared and distinct neurobiological mechanisms across PTSD, MDD, GAD, and SAD. This approach does not account for biological heterogeneity, limits treatment personalization, and treats comorbid conditions as confounds rather than potential sources of meaningful insights. Neuroimaging research has provided valuable findings, but challenges such as small sample sizes, methodological variability, and limited generalizability have constrained its clinical utility. Additionally, existing studies have yet to establish reliable biomarkers that can predict treatment response. To address these limitations, this R01-funded study integrates large-scale multimodal neuroimaging data and advanced computational techniques to refine psychiatric classification, improve treatment prediction, and contribute to precision psychiatry.

This R01 grant aims to improve the classification and treatment of PTSD, MDD, GAD, and SAD by identifying transdiagnostic biomarkers and biotypes through large-scale multimodal neuroimaging and advanced computational modeling. By leveraging neuroimaging data from the ENIGMA consortia (N=10,875) and an independent replication dataset from the UK Biobank, the grant will first use a top-down approach based on the Research Domain Criteria (RDoC) framework to identify transdiagnostic biomarkers associated with symptom domains. Multivariate regression models will be applied to multimodal neuroimaging data, including structural and resting-state functional MRI, to examine neural correlates of shared and disorder-specific symptoms such as disrupted sleep, irritability, and trauma-related distress. Integrating comorbid features into the analysis will provide a more comprehensive understanding of how these neural circuits contribute to symptomatology across disorders. To ensure robustness, penalized regression methods such as elastic nets and LASSO regression will be used for feature selection, with findings validated in the UK Biobank dataset.

In addition to a hypothesis-driven approach, the grant will incorporate data-driven methods to identify transdiagnostic biotypes, addressing the biological heterogeneity of these disorders. A stacked Denoising Variational Autoencoder (DVAE) framework will extract latent variables from high-dimensional multimodal neuroimaging data to identify biotypes that transcend traditional diagnostic categories. These biotypes will be validated through a DVAE visualization tool, supervised machine learning models, and comparisons of demographic, clinical, and neurobiological profiles. To assess external validity, transfer learning techniques will be used to replicate the findings in the UK Biobank dataset.

Finally, the grant aims to bridge neuroimaging research with clinical practice by predicting treatment response using the biomarkers and biotypes identified. Machine learning

algorithms, including support vector regression and random forest regression, will be employed to predict symptom reduction in individuals undergoing Prolonged Exposure therapy (ENIGMA-PTSD) or SSRI treatment (ENIGMA-MDD). The predictive performance of biomarkers, biotypes, and traditional clinical features will be compared to evaluate which models provide the most clinically relevant insights. The study will also test hypotheses regarding treatment-specific outcomes, such as whether biotypes dominated by negative affect network dysfunction are more responsive to cognitive-behavioral therapy, while those associated with the salience and pain-affective network respond better to pharmacological interventions.

As the Principal Investigator (PI) of this R01 grant, Dr. Xi Zhu is responsible for the overall direction, execution, and successful completion of the research. She provides the strategic framework for employing a top-down Research Domain Criteria (RDoC) approach to identify transdiagnostic biomarkers while spearheading the data-driven discovery of biotypes using advanced machine learning techniques. Her expertise in neuroimaging analysis, multivariate modeling, and deep learning methodologies is integral to the study's aims.

Dr. Zhu leads the identification of transdiagnostic biomarkers of symptom domains by guiding the application of multivariate regression models to large-scale multimodal neuroimaging data from ENIGMA (N=10,875). Building on her previous work comparing resting-state functional connectivity (rs-FC) patterns in PTSD and PTSD+MDD, she ensures the integration of comorbid features rather than treating them as confounds. She oversees the selection of penalized regression methods such as elastic nets and LASSO regression to optimize feature selection, ensuring the robustness of identified biomarkers. To validate these findings, she directs the replication efforts in the UK Biobank dataset, reinforcing the study's generalizability.

Beyond hypothesis-driven approaches, Dr. Zhu is the developer of the stacked Denoising Variational Autoencoder (DVAE) framework, the core method for identifying transdiagnostic biotypes across PTSD, MDD, GAD, and SAD. She leads the implementation of DVAE models on multimodal neuroimaging data, extracting latent variables that capture shared and distinct neurobiological features. Her expertise is pivotal in training and optimizing DVAE models, setting hyperparameters, applying denoising techniques, and evaluating performance through loss functions. Following latent feature extraction, she oversees the clustering analysis using ensemble-based k-means approaches, ensuring the biotypes are biologically and clinically meaningful.

Dr. Zhu also developed an interpretable deep learning visualization tool that plays a central role in biotype validation. She leads efforts to extend this tool to multimodal neuroimaging inputs, allowing the team to quantify the contribution of each imaging feature to the latent biotype structure. Under her guidance, supervised machine learning algorithms, such as support vector machines (SVM) and random forest (RF) classifiers, will be used to test the discriminability of biotypes from one another and from healthy controls. She oversees comparisons of biotypes across demographic, clinical, and neurobiological characteristics, using statistical analyses such as t-tests and ANOVA to assess their relevance.

A crucial aspect of Dr. Zhu's role is ensuring the generalizability of biotypes through external validation in the UK Biobank dataset. She directs the use of transfer learning techniques, leveraging pre-trained DVAE models from ENIGMA to encode and reconstruct UK

Biobank data, followed by biotype identification using the same procedures. Her expertise in handling large-scale datasets and model robustness ensures that the findings are replicable and broadly applicable.

Finally, Dr. Zhu's prior research on brain-based biotypes predicting treatment response in PTSD directly informs the study's treatment prediction component. She leads the quantification of biomarkers and biotypes to predict response to SSRIs for MDD and Prolonged Exposure (PE) therapy for PTSD. She oversees the application of machine learning models, including support vector regression and random forest regression, to compare the predictive power of neurobiological markers against standard clinical measures. By guiding the evaluation of whether specific neural network-dominated biotypes predict different treatment responses, she ensures the study's clinical relevance.

Through her leadership in neuroimaging, machine learning, and data-driven modeling, Dr. Zhu plays a pivotal role in achieving the grant's overarching goal: developing biologically informed transdiagnostic biomarkers and biotypes that can improve classification, enhance treatment prediction, and advance precision psychiatry.