Identifying PTSD Biotypes Using Explainable AI (XAI) for Personalized Treatment Strategies

(Authors: Zhu, Luyu, Bradley, Fatima, Jie, Orren Ravid, Yoojean Kim, Seonjoo Lee, Yuval Neria, Raj Morey)

Post-traumatic stress disorder (PTSD) is a complex mental health condition characterized by a wide range of cognitive, emotional, and behavioral symptoms. These symptoms often co-occur with other disorders such as major depressive disorder (MDD) and anxiety, making both diagnosis and treatment particularly challenging. With high rates of comorbidity and significant heterogeneity in symptom profiles, there is an urgent need for more personalized and precise treatment approaches. One promising avenue for achieving this goal is the use of Explainable AI (XAI) techniques, which can help uncover distinct neurobiological biotypes of PTSD and enhance our understanding of the disorder's underlying mechanisms.

This study leverages deep variational autoencoders (DVAEs), a cutting-edge XAI method, to identify neurobiologically distinct PTSD biotypes for the first time. By analyzing data from 3084 adult participants (1298 PTSD patients and 1786 controls) from the ENIGMA-PTSD working group, the study aims to uncover latent neurobiological features that differentiate subtypes of PTSD based on brain structure. The DVAE model is trained on controls' data to establish normative baselines for brain structural measures such as volume, cortical thickness, and surface area. These latent features are then used in clustering analyses to detect and characterize PTSD biotypes. XAI techniques are applied to interpret the contributions of specific brain regions, helping to pinpoint neuroanatomical differences that contribute to the emergence of these biotypes.

Aims of the Study:

1. **Identify distinct PTSD biotypes** based on neurobiological differences in brain structure, using deep variational autoencoders (DVAEs) and explainable AI (XAI) techniques.
2. **Evaluate the clinical relevance** of the identified PTSD biotypes by comparing PTSD severity, depression comorbidity, and classification performance across biotypes.
3. **Assess the predictive utility** of PTSD biotypes in treatment outcomes, specifically responses to prolonged exposure (PE) therapy.

The clinical utility of these identified biotypes is evaluated through a two-step process. First, we compare clinical measures (e.g., PTSD severity, comorbid depression) across PTSD biotypes to understand their distinct profiles. Second, we assess classification performance between each PTSD biotype and controls using machine learning algorithms such as random forests (RF) and support vector machines (SVM), validated by a K-fold cross-validation strategy. The classification performance is quantified by area under the curve (AUC) and accuracy (ACC) metrics, providing insight into the diagnostic potential of these biotypes. Finally, we examine the utility of these biotypes in predicting treatment outcomes, specifically responses to prolonged exposure therapy (PE).

Two distinct biotypes of PTSD were identified, with significant differences in PTSD severity and depression comorbidity. Notably, classification models that focused on specific biotypes outperformed models using the entire PTSD cohort, demonstrating the potential for more precise diagnosis. XAI analysis highlighted the pericalcarine region as a key brain structure distinguishing the two biotypes. Furthermore, the clinical relevance of these biotypes was confirmed through treatment outcome data: patients in biotype 2 showed a 75% reduction in PTSD symptoms following PE therapy, compared to a 57% reduction in biotype 1.

This study underscores the power of combining XAI with advanced neuroimaging models to reveal meaningful subtypes of PTSD. By incorporating these techniques into clinical decision-making, we can better understand the neurobiological underpinnings of PTSD, leading to more personalized and effective treatment strategies. These findings pave the way for a new era of precision mental health care, where treatments are tailored to the specific needs of each individual based on their unique neurobiological profile.