Identifying Neurobiological Heterogeneity in Clinical High-Risk Psychosis: A Data-Driven Biotyping Approach Using Resting-State Functional Connectivity

Xiaochen Tang[1,2#], Yanyan Wei[1#], Jiaoyan Pang[3], Lihua Xu[1], Huiru Cui[1], Xu Liu[1], Yegang Hu[1], Mingliang Ju[1], Yingying Tang[1], Bin Long[1], Wei Liu[2], Tianhong Zhang[1*], Jijun Wang[1*]

[1] Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai, China;
[2] School of Psychology, Shanghai Normal University, Shanghai, China;
[3] School of Government, Shanghai University of Political Science and Law, Shanghai, China;
*Correspondence to: jijunwang27@163.com.

Background:

Psychiatric disorders, particularly schizophrenia, have significant heterogeneity across dimensions such as symptomatology, illness duration, and treatment response. This presents a challenge in understanding and treating these disorders effectively, as current pharmacological interventions only work in about half of patients.

To address this, researchers have focused on early identification and study of Clinical High-Risk (CHR) psychosis, which exhibits considerable variability in clinical outcomes. A data-driven approach using machine learning and high-dimensional biomarkers from resting-state functional networks is proposed to dissect the neurobiological heterogeneity of the psychosis spectrum. We hypothesize that this methodology will identify and validate distinct neurobiological biotypes of CHR, suggesting distinct neurobiological patterns..

Methods:

This study recruited 239 participants from the SHARP program, comprising 151 CHR subjects and 88 age, sex, and education-matched healthy controls (HCs). Initial screening involved the self-report Prodromal Questionnaire-Brief version (PQ-B). Neurocognitive assessments were conducted using the Chinese version of MCCB.

Neuroimaging data were acquired using a 3 T Siemens MR B17 (Verio) system equipped with a 32-channel head coil. Functional connectome reconstruction was performed using Conn and SPM12 software.

The single-cell interpretation through multi-kernel learning (SIMLR) algorithm was used to identify a homogeneous subgroup of CHR subjects. The results are logged for validation, and a multiclass support vector machine (SVM) model is trained to assess reproducibility across iterations.

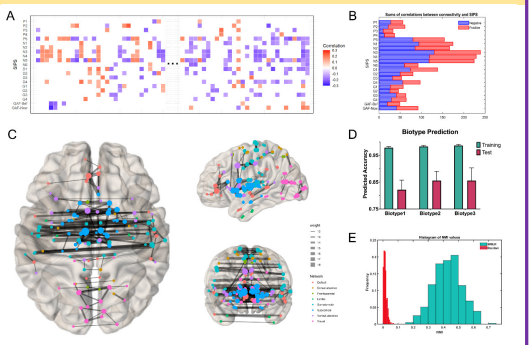

Figure 1. ***Construction of Biotype through Functional Connectivity (FC) Features. A)*** *Selection process of FC features. Correlations were calculated between 19 SIPS items, two GAF scores, and FC measures. B) Bar graph summarizing the sum of significant correlations between connectivity and SIPS/GAF items. C) Visualization of the significantly correlated functional connections, displaying only those with correlations in more than 12 items. D) Biotype prediction accuracy. E) Histogram of Normalized Mutual Information (NMI) distribution.*

Results:

The study found that Biotype 2 was significantly worse in spatial working memory performance and Biotype 3 had greater symptom severity than Biotype 1, and the CHR group performed poorly on MCCB tests, with the exception of the CPT-IP.

Biotype 1 showed enhanced connections within bilateral somatomotor networks, while Biotype 2 showed increased connectivity among bilateral subcortical networks. Biotype 3 showed increased connectivity between limbic and subcortical networks.

The biotype-specific abnormalities followed a progressive pattern, where Biotype 1 exhibited the mildest abnormalities and Biotype 3 the most severe. Furthermore, contrasting patterns of connectivity were observed between the bilateral somatomotor networks (Biotype 1 vs. Biotype 2, 3) and within subcortical areas, including the thalamus and striatum (Biotype 2 vs. Biotype 3).

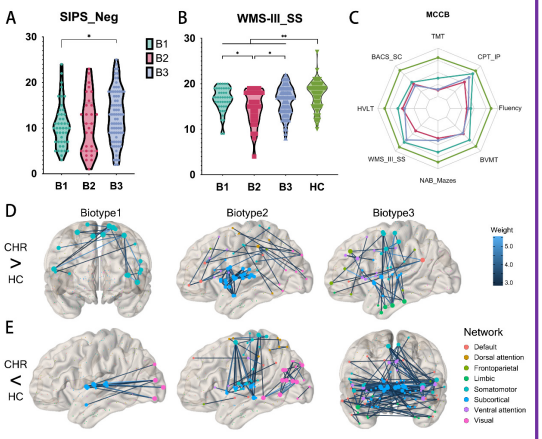

Figure 2. ***Clinical, Neuropsychological and Functional Connectivity Profiles of CHR Biotypes. A, B)*** *Post-hoc comparison of SIPS negative symptom and WMS-III_SS (Working Memory) scores across biotypes;* ***C)*** *Performance in cognitive function tests of the MCCB;* ***D, E)*** *Enhanced or reduced FC in each biotype compared to HC, with line color corresponding to the magnitude of statistical differences.*

Conclusion:

The distinct cognitive, symptomatic, and FC profiles of the biotypes underscore the complexity of psychosis risk states and suggest that different pathophysiological mechanisms may contribute to the risk of psychosis. The identification of these biotypes represents a significant step towards the development of personalized prevention and intervention strategies.

Acknowledgments:

We are grateful to the patients, their families and the volunteers who gave their time to participate in this study. Written informed consent was obtained from all participants or their legal guardians (Ethical No.2020-100).This work was supported by the National Natural Science Foundation of China (82171544, 82101582); Science and Technology Commission of Shanghai Municipality Grant (202240268); Clinical Research Plan of SHDC (SHDC2022CRD026, SHDC2020CR4066).