M56. INTEGRATING POLYGENIC ALLELE BURDEN INFORMATION AND PHENOMIC DATA TO CHARACTERIZE COMPLEX DISEASE TRAJECTORIES IN SEVERE MENTAL ILLNESS
Monika Budde1, Ivan Kondofersky2, Kristina Adorjan1, Fanny Aldinger1, Heike Anderson-Schmidt1, Till F. M. Andlauer4, Laura Flatau1, Katrin Gade3, Urs Heilbronner4, Janos Kalman1, Sergi Papiol1, Fabian J. Theis2, Peter Falkai5, Nikola S. Mueller6, Thomas G. Schulze1

1Institute of Psychiatric Phenomics and Genomics, Medical Center of the University of Munich, 2Institute of Computational Biology, Helmholtz Zentrum Munich, Germany, 3University Medical Centre Goettingen, Germany, 4Department of Translational Research in Psychiatry, Max Planck Institute of Psychiatry Munich, Germany, 5Department of Psychiatry, Ludwig Maximillian University, Munich, Germany, 6Regulatory Networks, Institute of Computational Biology, Helmholtz Zentrum Munich, Germany

Background Bipolar disorder (BD), schizoaffective disorder (SZA) and schizophrenia (SZ) are severe mental illnesses that share - at least in parts - psychopathological features and an underlying polygenic nature. One characteristic of all three diagnoses is the highly variable disease course and outcome. This heterogeneity is one of the biggest challenges in studying the underlying biological mechanisms. Therefore, defining more homogeneous subgroups across diagnoses is a promising approach. However, there are no clear criteria as how to define a “good” or “poor” course of illness as different domains can be considered such as psychopathology, cognitive performance, psychosocial functioning, or quality of life. We aim to integrate these domains and define longitudinal clusters of patients across diagnoses. Furthermore, we explore the characteristics of these clusters and the association of cluster membership with the individual load on schizophrenia polygenic risk scores (SZ-PRS).

Methods Participants were selected from an ongoing longitudinal project carried out at several centers in Germany and Austria (www.kfo241.de; www.PsyCourse.de). We characterize patients at four time-points over an 18-month period with a comprehensive phenotyping battery. The selected sample comprised a total of 198 participants (age(SD)=46.93(12.43); 46% females) with a DSM-IV diagnosis of SZ, SZA or BD, who completed the entire study period. DNA samples were genotyped using the Illumina PsychChip and imputed using the 1000 Genomes Phase 3 reference panel. SZ-PRS were calculated for all individuals based on the PGC2 SZ summary results.

Factor analysis for mixed data (FAMD) was applied to compute abstract data dimensions in a set of 117 longitudinally measured variables, i.a. on psychopathology, cognitive performance, functioning and quality of life. Longitudinal trajectories of patients on the first dimension were used as inputs for k-mean clustering for longitudinal data. This, in turn, resulted in the identification of three distinct clusters of patients, which we used as predictive variables for SZ-PRS at 11 p-value thresholds in a linear regression model.

Results Strongest loadings on the first dimension computed by FAMD were observed for quality of life items, a global depression rating and level of functioning. Three clusters of longitudinal trajectories were identified on this dimension: A) patients who scored highly on the dimension across all time points (58.1%); B) patients with consistently low scores (26.3%); C) patients who improved from baseline to the last follow up (15.7%). There were no significant between-group differences regarding
sex, age, diagnoses, center, age at onset, and duration of illness. Cluster membership was significantly associated with the SZ-PRS with highest polygenic burden in cluster B.

**Discussion** Although the reported results are preliminary and therefore have to be interpreted with caution, the approach of longitudinal clustering in order to identify cross-diagnostic, homogeneous subgroups of patients for genetic studies is promising. The next steps will be refinement of clusters by taking more than one dimension from the FAMD into account, verification of cluster solutions in an external dataset, and exploration of associations with other biological markers.

**Disclosure:** Nothing to Disclose.