Integrating polygenic allele burden information and phenomic data to characterize complex disease trajectories in severe mental illness

Monika Budde[1], Ivan Kondofersky[2,3], Kristina Adorjan[1], Fanny Aldinger[1], Heike Anderson-Schmidt[1,4], Till F. M. Andlauer[5], Laura Flatau[1], Katrin Gade[1,4], Urs Heilbronner[1], János Kálmán[1], Sergi Papiol[1], Peter Falkai[6], Fabian J. Theis[2,3], Nikola S. Mueller[2] & Thomas G. Schulze[1]

[1]Institute of Psychiatric Phenomics and Genomics (IPPG), Medical Center of the University of Munich, Germany
[2]Institute of Computational Biology, Helmholtz Zentrum Munich, Germany
[3]Center for Mathematics, Chair of Mathematical Modeling of Biological Systems, Technical University of Munich, Germany
[4]Department of Psychiatry and Psychotherapy, University Medical Center, Georg-August-University, Göttingen, Germany
[5]Department of Translational Research in Psychiatry, Max Planck Institute of Psychiatry Munich, Germany
[6]Department of Psychiatry and Psychotherapy, Medical Center of the University of 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 (QoL).

Aims

I) Identification of cross diagnostic longitudinal clusters of patients as a basis for biological studies

II) Testing the association of these clusters with schizophrenia-polygenic risk scores (SZ-PRS)

I) Identification of longitudinal clusters

Project: KFO241/PsyCourse (www.kfo241.de; www.psycourse.de)

- Multi-site project in Germany & Austria (Fig.1)
- DSM-IV diagnoses: SZ, SZA, BD
- Longitudinal, naturalistic study design (Fig.2)
- Extensive phenotyping; blood samples
- Recruitment ongoing

Figure 2: Longitudinal study design

Sample

- N = 198 (completed all 4 study visits)
- Sex: 46 % female
- Mean (SD) age at baseline: 46.92 (12.43) years
- Diagnoses: SZ: 49 %; BD: 38%; SZA: 12 %
- Mean (SD) age at onset: 29.67 (11.09) years
- Mean (SD) duration of illness: 17.25 (11.64) years
- Treatment at baseline: 78% outpatients; 22 % hospitalized
- Data from 16 centers

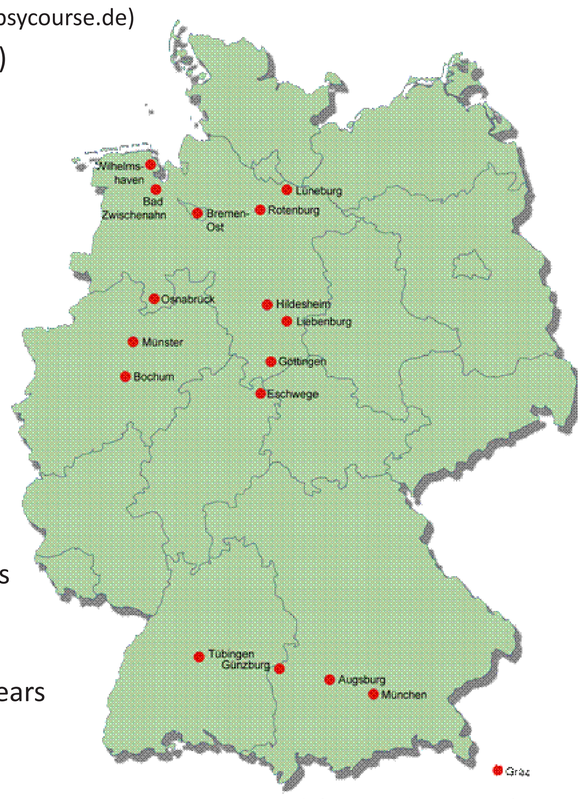

Figure 1: Recruitment network (KFO241/PsyCourse)

Methods

Step 1: Dimension reduction (Fig. 3)

- Apply factor analysis for mixed data (FAMD) to a set of 106 variables each measured at 4 time points → identify main latent dimension behind these data
- result: individual trajectories across time on dimension 1

Step 2: Longitudinal clustering (Fig. 4)

- k-means clustering for longitudinal data on individual trajectories on dimension 1

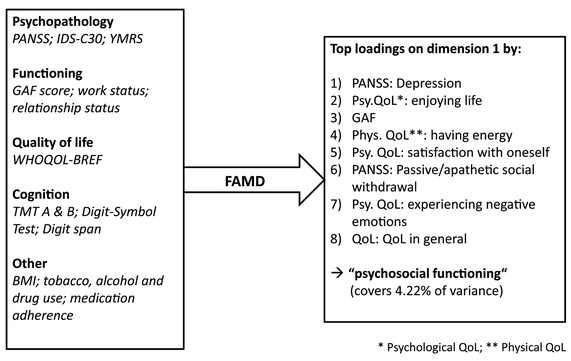

Figure 3: Dimension reduction of phenotype data

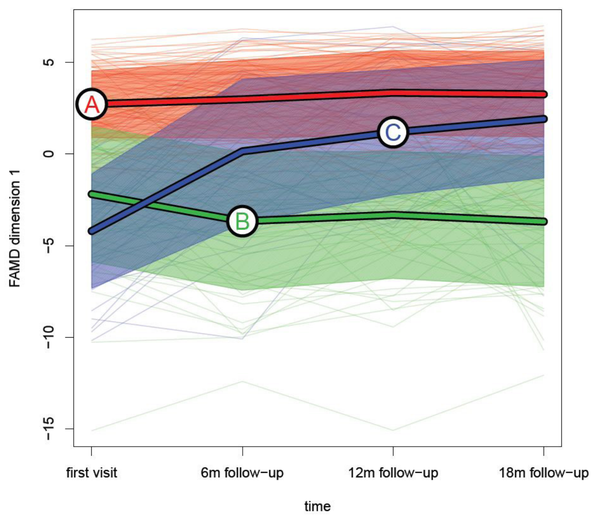

Figure 4: Longitudinal Clustering
A = “good“; B = “poor“; C = “improve”
FAMD dimension 1 = “psychosocial functioning”

Results

Based on the Calinski-Harabasz criterion, three clusters of longitudinal trajectories were identified on the dimension “psychosocial functioning”. **“Good” (A):** patients who scored highly across all time points (57%); **“poor” (B):** patients with consistently low scores (26%); **“improve” (C):** patients who improved from baseline to the last follow up (17%).

There were no significant between-group differences regarding sex, age, diagnoses, center, age at onset, and duration of illness. The mean difference in GAF score between clusters “good” and “poor” was 10-15 points. Significantly fewer patients in the “poor” group were fully employed compared to the other groups.

II) Association of longitudinal clusters and SZ-PRS

Background

SZ-PRS are associated with chronicity in SZ patients *(Meier et al., 2016, Mol Psychiatry)*

➔ How much variability of cluster membership in the dimension “psychosocial functioning” can be explained by SZ-PRS?

Methods

Genotyping and imputation

- DNA samples genotyped using the Illumina PsychChip
- imputed using the 1000 Genomes Phase 3 reference panel

Calculation of SZ-PRS

- calculated for all individuals with PLINK 1.07
- allelic effect sizes and P-values were obtained from the PGC2 SZ summary results *(SZ Working Group of the PGC 2014, Nature)*

Statistical analyses

- multinomial regression of cluster membership on SZ-PRS (11 P-value thresholds)
- covariates: age, sex, 5 principle ancestry components

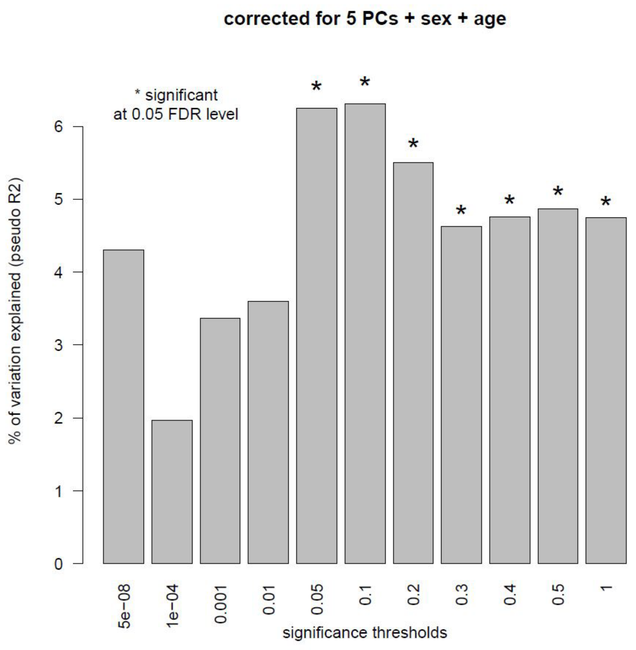

Figure 5: Multinomial repression of cluster membership on SZ-PRS

Results

From a P-value threshold of 0.05 on, up to 6% of the variability of cluster membership could be explained by SZ-PRS (significant after FDR correction for multiple testing; Fig. 5). Polygenic loading for SZ was highest in the “poor” cluster and lowest in the “improve” cluster. However, these differences between single clusters did not turn out to be significant.

Summary

We were able to identify three cross-diagnostic clusters of trajectories on a “psychosocial functioning” dimension over the course of 1.5 years. While the majority of patients stayed stable on a high (“good” cluster) or low (“poor” cluster) level, about 17% of the patients improved over time. First analyses with SZ-PRS look promising with the highest polygenic loading in the “poor“ cluster.

Limitations

- Study covers only relatively short period of time: 1.5 years
- Only small group of participants change over time
- Systematic drop out of severely impaired patients in longitudinal design
- Clusters not externally validated
- Clinical relevance of clusters questionable

Outlook: Optimizing current analyses

- increase power by repeating analyses in updated database
- capture more variance in the data by using more than one dimension from FAMD for clustering
- checking robustness of clusters with bootstrapping methods

Funding

Deutsche Forschungsgemeinschaft (DFG): SCHU 1603/5-1, SCHU 1603/7-1, TH 900/9-1

Dr. Lisa Oehler-Stiftung

The authors declare no conflicts of interest.

References

Meier SM et al., 2016. High loading of polygenic risk in cases with chronic schizophrenia. *Mol Psychiatry.*

Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014. Biological insights from 108 schizophrenia-associated genetic loci. *Nature.*