Published in final edited form as:

Biol Psychiatry. 2020 February 01; 87(3): 282–293. doi:10.1016/j.biopsych.2019.08.031.

Neurobiological divergence of the positive and negative schizophrenia subtypes identified upon a new factor-structure of psychopathology using non-negative factorization: An international machine-learning study

Ji Chen[1,2], **Kaustubh R. Patil**[1,2], **Susanne Weis**[1,2], **Kang Sim**[3,4], **Thomas Nickl-Jockschat**[5,6], **Juan Zhou**[7], **André Aleman**[8], **Iris E. Sommer**[8,9], **Edith J. Liemburg**[10], **Felix Hoffstaedter**[1,2], **Ute Habel**[11,12], **Birgit Derntl**[13], **Xiaojin Liu**[1,2], **Lydia Kogler**[13], **Christina Regenbogen**[11,12], **Vaibhav A. Diwadkar**[14], **Jeffrey A. Stanley**[14], **Valentin Riedl**[15], **Renaud Jardri**[16], **Oliver Gruber**[17], **Aristeidis Sotiras**[18], **Christos Davatzikos**[19,20], **Simon B. Eickhoff**[1,2,*], **Pharmacotherapy Monitoring and Outcome Survey (PHAMOUS) investigators**

[1]Institute of Neuroscience and Medicine, Brain and Behaviour (INM-7), Research Center Jülich, Jülich, Germany [2]Institute of Systems Neuroscience, Medical Faculty, Heinrich Heine University Düsseldorf, Düsseldorf, Germany [3]Department of General Psychiatry, Institute of Mental Health, Singapore [4]Research Division, Institute of Mental Health, Singapore [5]Iowa Neuroscience Institute, Carver College of Medicine, University of Iowa, Iowa City, IA, USA [6]Department of Psychiatry, Carver College of Medicine, University of Iowa, Iowa City, IA, USA [7]Center for Cognitive Neuroscience, Neuroscience and Behavioral Disorders Program, Duke-National University of Singapore Medical School, Singapore [8]Department of Neuroscience, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands [9]BCN Neuroimaging Center, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands [10]University of Groningen, University Medical Center Groningen, Rob Giel Research center, Groningen, The Netherlands [11]Department of Psychiatry, Psychotherapy and Psychosomatics, RWTH Aachen University, Germany [12]JARA-Institute Brain Structure Function Relationship, Research Center Jülich and RWTH Aachen, Jülich, Germany [13]Department of Psychiatry and Psychotherapy, University of Tübingen, Tübingen, Germany [14]Department of Psychiatry and Behavioral Neuroscience, Wayne State University, Detroit, MI, USA [15]Abteilung für diagnostische und interventionelle Neuroradiologie, Klinikum rechts der Isar, Technische Universität München, Germany [16]Univ Lille, CNRS UMR9193, SCA Lab & CHU Lille, Fontan Hospital, CURE platform,

[*]**Corresponding author: Univ.-Prof. Dr. med. Simon B. Eickhoff**, Institute of Systems Neuroscience, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany & Institute of Neuroscience and Medicine, Brain and Behaviour (INM-7), Research Center Jülich, 52428 Jülich, Germany. Tel: +49 2461 61 1791; S.Eickhoff@fz-juelich.de.

Model availability:

A Dimensions and Clustering Tool for assessing schizophrenia Symptomatology (DCTS) is available at: http://webtools.inm7.de/sczDCTS/

Disclosures

All of the authors report no biomedical financial interests or potential conflicts of interest.

Lille, France [17]Section for Experimental Psychopathology and Neuroimaging, Department of General Psychiatry, Heidelberg University, Heidelberg, Germany [18]Department of Radiology and Institute for Informatics, School of Medicine, Washington University in St. Louis, St. Louis, Missouri, USA [19]Center for Biomedical Image Computing and Analytics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, USA [20]Department of Radiology, Section of Biomedical Image Analysis, Perelman School of Medicine, University of Pennsylvania, Philadelphia, USA

Abstract

Objective: Disentangling psychopathological heterogeneity in schizophrenia is challenging and previous results remain inconclusive. We employed advanced machine-learning to identify a stable and generalizable factorization of the “Positive and Negative Syndrome Scale (PANSS)”, and used it to identify psychopathological subtypes as well as their neurobiological differentiations.

Methods: PANSS data from the Pharmacotherapy Monitoring and Outcome Survey cohort (1545 patients, 586 followed up after 1.35±0.70 years) were used for learning the factor-structure by an orthonormal projective non-negative factorization. An international sample, pooled from nine medical centers across Europe, USA, and Asia (490 patients), was used for validation. Patients were clustered into psychopathological subtypes based on the identified factor-structure, and the neurobiological divergence between the subtypes was assessed by classification analysis on functional MRI connectivity patterns.

Results: A four-factor structure representing negative, positive, affective, and cognitive symptoms was identified as the most stable and generalizable representation of psychopathology. It showed higher internal consistency than the original PANSS subscales and previously proposed factor-models. Based on this representation, the positive-negative dichotomy was confirmed as the (only) robust psychopathological subtypes, and these subtypes were longitudinally stable in about 80% of the repeatedly assessed patients. Finally, the individual subtype could be predicted with good accuracy from functional connectivity profiles of the ventro-medial frontal cortex, temporoparietal junction, and precuneus.

Conclusions: Machine-learning applied to multi-site data with cross-validation yielded a factorization generalizable across populations and medical systems. Together with subtyping and the demonstrated ability to predict subtype membership from neuroimaging data, this work further disentangles the heterogeneity in schizophrenia.

Keywords

non-negative factorization; brain imaging; subtyping; machine learning; multivariate classification; schizophrenia

INTRODUCTION:

Schizophrenia is a heterogeneous disorder with marked inter-individual variability of psychopathology, which is related to treatment response and long-term outcomes (1,2). Earlier clinical subtypes (e.g., hebephrenic or paranoid) were eliminated in recent

nosological classifications due to poor diagnostic stability, validity, and utility (3). Considerable efforts have been devoted to better understand and categorize schizophrenia phenomenology by factorizing symptoms into cardinal dimensions or clustering patients into psychopathological subtypes based on scales such as the Positive and Negative Syndrome Scale (PANSS), a well-established assessment of schizophrenia psychopathology (4).

The three PANSS subscales (negative, positive, and general psychopathology [GPS]) are generally suggested to not optimally and adequately capture the latent organization of schizophrenia symptomatology; items within a subscale show modest internal consistency (5), while those across subscales are strongly correlated (6–8). Previous factorizations of the PANSS have been inconsistent, advocating solutions between four and seven factors (6–11). Although a five-factor structure was most frequently proposed (9–11,14), it has continuously failed to be confirmed in independent samples (12–15). Interpretations of previous factor models are, furthermore, complicated by lack of sparsity (all items contribute to any factor) (16) and co-existence of positive and negative weights (17). Finally, most previous studies investigated rather small and geographically restricted samples, raising doubts over generalization to different populations and medical systems as systematic cross-validation to assess stability and generalizability have rarely been performed. Previous work on psychopathological subtyping is likewise inconclusive (18–20) with added concerns related to longitudinal stability and neurobiological differentiability. These aspects are particularly relevant in the emerging context of precision psychiatry, and raise the following questions: Do psychopathological subtypes represent stable patient characteristics, and do they relate to divergent neurobiological substrates that are identifiable from brain imaging data? Functional MRI (fMRI) parameters may serve as an endophenotype, underpinning the symptomatic heterogeneity (21), which have added ample valuable insights into the neural pathophysiology of schizophrenia and its relation to clinical presentations (22,23). However, whether, and to what extent, the brain functional connectivity (FC) could discriminate psychopathological subtypes remains unknown. A successful classification using endophenotypical characteristics would support distinctiveness of symptomatically derived schizophrenia subtypes expressed along the cardinal axes of psychopathology.

In the present study, we addressed the aforementioned questions as follows: 1) a robust, cross-validated, and interpretable factor-structure of schizophrenia psychopathology was identified based on PANSS scores of over 2000 patients using an unsupervised machine-learning approach (orthonormal projective non-negative matrix factorization [OPNMF]) (24–27); 2) core schizophrenia subtypes were derived by applying soft-clustering to the identified factor-structure, whose longitudinal stability was evaluated in repeatedly assessed patients; and 3) neurobiological differentiation of those subtypes based on resting-state FC (rsFC) patterns was investigated by cross-validated classification analysis, serving as a biological validation of a clinical (multivariate) construct.

METHODS AND MATERIALS:

Sample

We used two large datasets collectively providing individual-item PANSS scores for 2035 schizophrenia patients: 1) a subset of 1545 patients (586 followed up after 1.35±0.70 years)

with complete individual-item PANSS scores and a diagnosis of schizophrenia (DSM-IV criteria) retrieved from the Pharmacotherapy Monitoring and Outcome Survey (PHAMOUS) database (28,29). This dataset was recruited from four institutions located in the Netherlands and assessed with a uniform protocol; 2) a deliberately heterogeneous sample pooled from nine centers located in Europe, the USA, and Asia (490 patients; Table 1; Table S1). This international dataset covers a broad range of clinical states, settings, and medical systems, making it ideal to evaluate the generalization of our factor-model to new and diverse populations. Diagnoses in the international sample were established based on the DSM-IV, DSM-IV-TR, or DSM-V criteria (Supplement). At all sites, data was acquired in accordance with the declaration of Helsinki and after obtaining informed consent from the patients. Approval for the pooled re-analysis was obtained from the ethics committee of the HHU Düsseldorf.

Factorization of PANSS using OPNMF

OPNMF (25,27) decomposes given data (PANSS) into two non-negative matrices; a basis matrix (dictionary) with factors as columns which can be readily generalized to new data due to the projective constraint, and a factor-loading matrix representing symptomatology of individual patients along these factors. The orthonormality constraint promotes a sparse, and hence, interpretable representation. For choosing the number of factors, a set of sophisticated evaluation strategies were implemented (Supplement, Figures S1–S2).

We first applied OPNMF to the PANSS scores from the 1545 PHAMOUS patients with the number of factors ranging from two to eleven. The optimal number of factors was identified by using cross-validation in 10,000 split-half analyses. The PHAMOUS sample was split into two halves, and on each split-sample, OPNMF was performed to derive the dictionary. The congruency between item-to-factor assignments, based on its largest coefficient, was assessed using the adjusted Rand-index (30) and variation of information (31), along with the concordance index (32) between the dictionaries. We also quantified out-of-sample reconstruction error by projecting one split-sample data on the dictionary from the other split-sample. A lower increase in out-of-sample error compared to within-sample reconstruction error indicates better generalizability. This split-sample analysis was repeated on the international dataset. Additional bootstrap and 10-fold cross-validation analyses were conducted on each of the two samples independently.

Most critically, we assessed stability and generalizability between the factorizations of the PHAMOUS sample (good for learning a structure due to size) and the international sample (good for validation due to heterogeneity). We performed OPNMF independently on the bootstrapped samples from each dataset. The resulting factorizations were then compared using the approaches described above. That is, for each number of factors, we assessed stability by comparing the dictionaries obtained from factorization of bootstrapped samples (PHAMOUS vs. international). Most importantly, we also evaluated generalization to new data by measuring increase in reconstruction error for the international data following projection on the PHAMOUS dictionary. This cross-sample evaluation was repeated after accounting for between-dataset differences in sample size, age, and illness duration. Leave-one-site-out validation, on both the PHAMOUS and international samples, was performed to

check for site bias. We also repeated all analyses after removing outliers or including the repeated PANSS measurements (Supplement). Factorizations of the pooled (PHAMOUS +international) sample, as well as the stability and accuracy of PHAMOUS generated dictionaries in estimating out-of-sample loadings/item-scores were additionally assessed (Supplement).

After identifying the optimal PANSS factorization, the international sample was projected onto this PHAMOUS-derived dictionary to obtain factor-loadings for subsequent analyses (except for longitudinal analysis as reassessments were only available in the PHAMOUS sample), in order to avoid “double dipping”/“leakage” that would occur if scores were analyzed in the same dataset used to derive the dictionary.

Internal consistency and relationship among variables

Internal consistency of the optimal OPNMF model, as well as the PANSS subscales (as reference), was assessed using Cronbach’s alpha, where higher values indicate more closely related items within a set. Relationships between the OPNMF factor-loadings were assessed using linear and partial correlations (controlling for symptom severity, i.e., total PANSS score), including bootstrap stability analyses. The OPNMF factor-loadings were correlated with the three PANSS subscales. Correlations between individual items were also computed. For comparison, we performed an exploratory factor analysis (EFA) on the PHAMOUS sample and a confirmatory factor analysis (CFA) on the international sample, as well as a principal component analysis (PCA) on both samples. Effects of gender, age, illness duration, and symptom severity on the OPNMF factor-loadings were analyzed in the international sample (N=393, with complete information). Following a MANOVA to assess effects on all the loadings, individual 4-way ANOVAs were performed on each loading to identify their association with the demographic and clinical features (corroborated by bootstrap and leave-one-site-out analyses). Current drug dosages of anti-psychotic medication, available for 149 patients, were olanzapine-equivalent transformed (33), and included in the 4-way ANOVA models for a supplementary analysis.

Psychopathological subtypes

After adjusting for age, gender, illness duration, and symptom severity, factor-loadings were used as features for clustering patients into psychopathological subtypes. After confirming the data clusterability (34), we applied fuzzy c-means clustering (35) which provided cluster membership likelihoods for each patient. The optimal cluster number was determined based on the fuzzy silhouette index (36), the Xie and Beni index (37), and partition entropy (35). Stability was tested by leave-one-site-out replication, subsampling, and bootstrap resampling (Supplement). Given the heterogeneous nature of schizophrenia and the observation of multiple patients with ambiguous memberships, a cutoff over the membership likelihoods was adopted to remove cluster ambiguous patients. For this, additional evidence from Gaussian mixture modeling (GMM) was considered. Specifically, patients were clustered again using GMM, and the optimal cluster number was determined by Bayesian information criterion (38). After assigning patients to the clusters, we took the intersection of the c-means and GMM results. A cutoff was chosen, based on the c-means membership likelihoods that well discriminated the patients inside the intersection from those outside,

while also retaining a decent sample size for classification. This filtering step is critical, as ambiguous patients might obscure otherwise classifiable rsFC patterns for the identified subtypes. Afterwards, differences between subtypes regarding factor-loadings and demographic and clinical features were ascertained by permutation tests (39). To assess longitudinal stability, the same c-means clustering was applied to the repeated assessments of the PHAMOUS sample. The optimal dictionary, identified on the PHAMOUS 1545 patients without repeatedly assessed PANSS scores, was used for projection to yield the factor-loadings. After excluding ambiguous cases, patients assigned to the same clusters in both initial and follow-up stages were regarded as longitudinally stable (Supplement). For comparison, the same clustering was additionally performed on the factor-loadings without any covariates or symptom severity adjustment, as well as on the PANSS subscales/items both with and without covariates adjustment (Supplement).

Classifying psychopathological subtypes from rsFC

Multivariate classification analysis was conducted on patients from the international sample for whom imaging data was available after excluding those with ambiguous subtype assignment, low image quality, and excessive head motion (N=84; Figure S24). After standard preprocessing (Supplement), regional time-series were extracted based on a parcellation scheme with 600 cortical (40) and 36 subcortical parcels (41), adjusted for confounders (42), and used to compute the functional connectome. We tested each parcel for whether its pattern of rsFC to all other parcels allowed to classify subtype membership in novel subjects. Resulting parcel-wise accuracies yielded a whole-brain map indicating the classifiable power of each parcel's connectivity profile. The radial basis function kernel support vector machine classifier, which can deal with the potentially nonlinear relationship between the psychopathological and the neural spaces, was employed. A stratified 10-fold cross-validation was implemented to assess the out-of-sample classification performance (Figure S25). Effects of age, gender, site, illness duration, symptom severity, and head-motion parameters were adjusted using a linear-regression model fitted only in the training sample (43). Significance of the parcel-wise accuracy was estimated by permutation tests, followed by false discovery rate (FDR) correction for multiple comparisons (Supplement). Parcels surviving FDR were functionally characterized (http://brainmap.org/)(44) (Supplement).

RESULTS:

Dimensions of psychopathology

The most robust and generalizable model consisted of three factors for the PHAMOUS sample (Figure 1A), which effectively combined the positive and affective symptoms compared to the optimal four-factor model for the international dataset (Figure 1B). The additional factor in the international dataset may relate to the higher prevalence of psychotic symptoms, particularly auditory hallucinations, compared to the PHAMOUS sample, which contains more chronic patients (Table 1). Consequently, a four-factor model was identified as the most stable and, importantly, generalizable model of psychopathology in the cross-sample evaluation (Figure 1C). The first dimension mainly represents *negative* symptoms, such as blunted affect and apathy (Figure 1D). The second represents *positive* symptoms,

e.g., delusions and hallucinations. The third factor comprises symptoms, such as depression, anxiety, and tension, reflecting an *affective* dimension, while the last factor represents *cognitive* impairments. Notably, only a few items contributed to multiple dimensions, e.g., active social avoidance contributed to both negative and affective factors.

All findings were fully confirmed by (i) bootstrap and 10-fold cross-validation, (ii) removing outliers (18 patients), (iii) adding PANSS data from follow-up examination in the PHAMOUS sample, (iv) leave-one-site-out validation, (v) accounting for between-dataset differences in sample size, age, and illness duration, (vi) pooling the two datasets with cross-validation and out-of-sample generalization assessments, and (vii) loading/item-score predictions across factor-solutions, bootstrapped samples, and sites (Figures S3–S11).

Internal consistency and relationship among variables

Items within a factor showed higher and more homogeneous positive correlations (while fewer anti-correlations) for the OPNMF factors than the PANSS subscales (Figure 2A–2B, Figure S12). Internal consistency of our OPNMF four-factor structure (positive: Cronbach's alpha=0.75; negative: 0.92; affective: 0.85; cognitive: 0.83) was on average higher than that for the PANSS subscales (positive: 0.72; negative: 0.87; GPS: 0.87), previously reported factor models (ranging from 0.6 [excited] to 0.9 [negative])(7–9,45), and the EFA models derived from the current sample (0.49-0.91, Table S2). All PHAMOUS-derived 4-7 factor EFA-models could not be confirmed in the international sample, i.e., inadequate fit (Table S2). Compared to PCA, OPNMF showed better generalizability (Figure S13). The positive and negative factors were highly correlated with the positive and negative PANSS subscales, respectively, both before (r=.92 and r=.97) and after (r=.85 and r=.89) controlling for symptom severity (Figure S14). Interestingly, after adjusting for symptom severity, the cognitive factor did not correlate with GPS (r=.02).

Over individual patients, the loadings on our four factors were significantly inter-correlated, with negative and positive factors showing lowest (r=.32, averaged over 10,000 bootstraps), negative and affective factors highest (r=.70) correlations (Figure 2C). After controlling for symptom severity, positive and negative factors became anti-correlated (r=−.59, Figure S15).

MANOVA revealed a significant influence of symptom severity on the joint factor-loadings (p<.001). Follow-up 4-way ANOVAs showed that symptom severity had a significant effect on each factor (all p<.001, all 6β0.07, Figure 2D). The cognitive factor showed a trend towards a positive relationship with illness duration (p=.081, β=0.014) and a significant negative relationship with age (p=.033, β=−0.015, Figure 2D), though both covariates were collinear (r=.65). In contrast, loadings on the negative factor were higher for older individuals (p=.11, β=0.016) and lower for those with longer illness duration (p=.18, β=−0.015). Gender-differences were not observed in any factor. Bootstrap (Figure 2E) and leave-one-site-out analyses corroborated the aforementioned ANOVA findings (Supplement). Adding olanzapine-equivalent dosage to the 4-way ANOVA did not reveal any significant association with medication.

Psychopathological subtypes

Fuzzy c-means clustering on the adjusted loadings revealed an optimal two-cluster solution (Figure 3A–3B, Figures S16–S17). Although GMM demonstrated an optimal three-cluster solution, one of the clusters was diffusely distributed in space containing patients from both the c-means clusters. This GMM cluster was excluded as it would not represent any specific subtype (Figure 3C). Patients inside the c-means GMM intersection had higher c-means membership likelihoods (roughly>0.7) to belong to their own cluster than those outside the intersection ($p<.01$, *Wilcoxon rank-sum test*; Figure 3D). We chose the "cluster cores" using the likelihoods of c-means with a cutoff of 0.7. Resultantly, two core subtypes were defined after filtering out 50 ambiguous patients from each of the two c-means clusters (Figure 3E). The first subtype showed a psychopathological profile dominated by negative and affective symptomatology (subtype A). The other featured prominent positive symptoms (subtype B; all $p<.001$ in permutation tests) (Figure 3F, Figure S18). Importantly, subtypes did not differ in gender distribution, age, or illness duration (all $p>.05$), but subtype B showed higher symptom severity ($p=.008$). The same two-cluster solution was replicated on the PHAMOUS patients with complete demographic and clinical information ($N=1326$; 56% of the 603 ambiguous patients in subtype B when hard-clustered), and those with repeatedly assessed PANSS scores ($N=527$; 45% ambiguous). Almost 80% of the reassessed patients retained their subtype, with subtype A being more stable (85%) (Supplement). The additional clustering analyses supported our four-factor model with covariates adjustment for a clinically meaningful subtyping (Figures S19–S23).

Classifying psychopathological subtypes from rsFC

The rsFC profile of the parcel located in right ventro-medial prefrontal cortex (vmPFC) yielded the highest out-of-sample classification accuracy (70% of patients not used for training were assigned to the correct psychopathological subtypes), followed by parcels in the right temporoparietal junction (TPJ), the bilateral precuneus, and the posterior cingulate cortex (PCC). Permutation tests showed that the top 104 classifiable parcels were significant ($p<.05$) against chance (i.e., randomized subtype labels), and 53 parcels survived FDR correction ($q<.05$). Of note, parcels are labeled by their microanatomical location, with their functional implications (Table S4). Classification with additional global mean signal removal or with a RS-subcortical parcellation replacement (7 parcels [46]; as a control analysis instead of the finer Brainnetome subcortical parcellation), replicated these results (Supplement; Figure S26).

DISCUSSION

By factorizing the PANSS scores from a large sample using OPNMF and cross-validating the results in a heterogeneous multi-site dataset, we revealed a robust, replicable, and generalizable four-factor structure comprising negative, positive, affective, and cognitive dimensions across populations, settings, and medical systems. Based on this four-factor structure, two core psychopathological subtypes were obtained which showed good longitudinal stability and could be discriminated by regional rsFC patterns, with right vmPFC showing the highest (70%) classification accuracy.

Relationship to previous factor models of the PANSS

The three PANSS subscales do not reflect the latent structure of this inventory well (5–7). In turn, our model represents a stable, generalizable and well interpretable description of schizophrenia psychopathology suited for representing the full range of acute and chronic symptoms. Resonating with this view, a pyramidical model proposed by the PANSS developers comprised four components (6). Three of these (negative, positive, and affective dimensions) showed good agreement with our model. The fourth component, however, isolated only excitement, while cognitive disturbances were distributed across all dimensions or discarded. Such a representation is obviously at odds with the importance of cognitive dysfunction, and has prompted the proposal of more complex models (7–11), e.g., a recent five-factor model reflecting negative, positive, depressed, excited, and cognitive dimensions (45). However, replicability, external validity, and generalization remain a concern for these models (47). As a striking example, White and colleagues (12) found that, none of 20 tested models fit their data adequately, and put forward a new pentagonal model, which later (together with 24 others) also could not be confirmed (13). In the same study (14), the authors developed an improved five-factor model using 10-fold cross-validation. However, it still failed to be confirmed, along with 31 other five-factor models, in a later study involving two large Chinese samples (15). In our sample, inadequate fit for EFA models with 4-7 factors was also manifested, and the fifth OPNMF factor, compared to the four factors, showed the poorest out-of-sample loading predictions (Figure S10B). These facts, as a whole, point to a fundamental instability of five-factor models (9–15,44). Addressing these concerns, the current work was not only based on a large sample for model identification, but, importantly, focused strongly on cross-validated stability and out-of-sample generalization. Critically, the external validation was based on a heterogeneous, international sample, and the optimal model suggested a single factor to combine both the cognitive and excited symptoms. This view is corroborated by observations that cognitive and excited symptomatology is highly correlated (45) and share similar neurobiological substrates (48,49).

Internal consistency and relationship among variables

Although we identified the optimal representation by its robustness and ability to generalize to new populations, the positive and negative dimensions of our model also showed better internal consistency than the PANSS subscales while differentiating the broad “general psychopathology”. Moreover, our affective and cognitive factors showed higher internal consistency than those reported in previous factor-models (7,8,10,45). Finally, correlations between individual items within OPNMF factors were higher and more homogeneous compared to the PANSS subscales.

Matching previous reports (4,8,45), negative and positive factors from our model were least related before, and showed a strong anti-correlation after, controlling for symptom severity. The inversed age vs. illness duration effect on negative symptoms implies that this effect may be more related to age than illness duration. This is intriguing from the perspective of early aging/degeneration, but needs to be taken with caution, as age and illness duration are highly correlated. Co-linear variables in a single linear regression model make it difficult to

disentangle their respective effects on the negative factor, as well as on the cognitive factor (50).

Psychopathological subtypes, longitudinal stability, and neurobiological differentiability

Our results revealed two distinct schizophrenia subtypes featuring predominantly positive and negative symptoms, respectively. The subtypes were longitudinally stable and could be classified from neuroimaging data. Such a positive-negative dichotomy has been widely supported (51,52). Finer distinctions have been proposed, but show poor replicability (18–20). The inconsistency of a finer subtyping may relate to idiosyncrasies in small samples from a single geographical region and to the lack of explicit analyses of stability and replicability. Moreover, longitudinal stability of our new subtypes was higher than that reported for traditional clinical subtypes or for a positive/negative/mixed topology (53–55). Interestingly, we found subtype A to be particularly stable. Previous studies indicated that both mixed and negative symptom states increase over time, whereas psychotic expressions usually diminish outside acute episodes and over time (54,55). Future studies with a longer follow-up duration are desired; the mean of 1.35 years' follow-up assessed in the present study is not a long period in schizophrenia. Additionally, the employed soft-clustering method better accommodates ambiguous patients compared to previous hard-clustering methods, and patients with ambiguous memberships can be furthermore filtered out with appropriate cutoffs to improve the ability of detecting neurobiological distinctions between subtypes. Nonetheless, the cutoff value chosen in the present study should be noted as heuristic. The cluster ambiguous patients might represent a transient group lying in-between the two more differentiated subtypes.

The current top classifiable brain regions are all implicated in schizophrenia pathophysiology and processes relevant to the psychopathological distinction (22,56–60). Previous findings in the literature relating fMRI parameters to differential symptoms, so far, exclusively relied on group-level analyses, while our approach bridged an important gap between neurobiological divergence and distinct symptomatic patterns at the individual level. To our knowledge, it is the first study to successfully classify psychopathological subtypes in schizophrenia. Of note, the current classification accuracy was similar to those previously reported for classifications of schizophrenia patients vs. healthy participants (61–63). The demonstrated neurobiological differentiability corroborates the currently identified schizophrenia subtypes expressed along the four OPNMF dimensions.

Limitations and Considerations

We assessed factor-structure, subtypes, and their neurobiological differentiations with a particular emphasis on robustness and generalization. This conservative approach seems necessary given current concerns of non-replicability in biomedical research, but may have contributed to the fact that we corroborated the clinically well-established positive-negative distinction rather than identifying more differentiated subtypes. We note, though, that a recent imaging-based clustering also provided evidence for two subtypes (64), and stress that the current analysis of a large, heterogeneous sample did not reveal any evidence for a more fine-grained differentiation among the patients. It, thus, remains to be seen whether an additional differentiation between these two core subtypes may be robustly revealed by

analyzing substantially larger samples, or whether previously proposed additional subtypes represent distinctions that could be found in a particular dataset but are not universally present. We also acknowledge that patients were on their regular medication as prescribed by the attending psychiatrists, and the current results might thus be confounded by direct and indirect mediation effects thereof. However, it stands to reason that a multi-site study, pooling patients from different psychiatrists with differential medication strategies, will render medication largely as a source of random variation in our data. Such noise would effectively make it harder to identify generalizable psychopathological factors, robust subtypes, and, in particular, train models that work well for out-of-sample classification of subtype membership. We would thus argue that the current results should not be driven by medication effects, but rather represent general structures of psychopathology and schizophrenia subtypes. In addition, rs-fMRI has its own limitations, such as variability across scanning sessions and the issue of confounding factors (65–67). We focused on rs-fMRI because it could temporally better map the likewise state-dependent psychopathology compared to structural MRI.

Using advanced machine-learning with cross-sample validation, the present study suggested a stable and generalizable four-factor model of PANSS. This representation allowed for the definition of a reliable positive-negative subtype differentiation that showed good longitudinal stability and a neurobiological divergence in rsFC. Overall, the present work further disentangled the heterogeneity of schizophrenia, possibly allowing for the design of more specifically targeted treatments.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

PHAMOUS-investigators:

Agna A. Bartels-Velthuis (University of Groningen, University Medical Center Groningen, University Center for Psychiatry, Rob Giel Research Center; Lentis Psychiatric Institute, Groningen, The Netherlands); Richard Bruggeman (University of Groningen, University Medical Center Groningen, University Center for Psychiatry, Rob Giel Research Center and Faculty of Behavioural and Social Sciences, Department of Clinical Psychology & Developmental Neuropsychology, Groningen, The Netherlands); Stynke Castelein (Lentis Psychiatric Institute; University of Groningen, Faculty of Behavioural and Social Sciences, Department of Clinical Psychology & Experimental Psychopathology, Groningen, The Netherlands); Frederike Jörg (GGZ Friesland Mental Health Care Organization, Leeuwarden, The Netherlands); Gerdina H.M. Pijnenborg (University of Groningen, Faculty of Behavioural and Social Sciences, Department of Clinical Psychology & Experimental Psychopathology; GGZ Drenthe Mental Health Care Organization, Dennenweg; University of Groningen, University Medical Center Groningen, University Center for Psychiatry, Psychosis Department, Groningen, The Netherlands), Henderikus Knegtering (University of Groningen, University Medical Center Groningen, University Center for Psychiatry, Rob Giel Research Center; Lentis Psychiatric Institute, Groningen, The Netherlands), Ellen Visser (University of Groningen, University Medical Center Groningen, University Center for Psychiatry, Rob Giel Research Center, Groningen, The Netherlands).

This study was supported by the Deutsche Forschungsgemeinschaft (DFG, EI 816/4-1), the National Institute of Mental Health (R01-MH074457), the Helmholtz Portfolio Theme “Supercomputing and Modeling for the Human Brain”, and the European Union’s Horizon 2020 Research and Innovation Programme under Grant Agreement No. 720270 (HBP SGA1) and 785907 (HBP SGA2). Ji Chen has received a Ph.D fellowship from the Chinese Scholarship Council. Also, acknowledgment goes to Asadur Chowdury, PhD (Brain Imaging Research Division, Wayne State University School of Medicine, Detroit, Michigan), who contributed to the early arrangement and communication of the Wayne-State dataset.

REFERENCES:

1. Lally J, MacCabe JH (2015): Antipsychotic medication in schizophrenia: a review. Br Med Bull 114:169–179. [PubMed: 25957394]
2. Lang F, Kösters M, Lang S, Becker T, Jäger M (2013): Psychopathological long-term outcome of schizophrenia-a review. Acta Psychiatr Scand 127: 173–182. [PubMed: 23136879]
3. Braff DL, Ryan J, Rissling AJ, Carpenter WT (2013): Lack of use in the literature from the last 20 years supports dropping traditional schizophrenia subtypes from DSM-5 and ICD-11. Schizophr Bull 39: 751–753. [PubMed: 23674819]
4. Kay SR, Fiszbein A, Opler LA (1987): The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr Bull 13: 261–276. [PubMed: 3616518]
5. Peralta V, Cuesta MJ (1994): Psychometric properties of the positive and negative syndrome scale (PANSS) in schizophrenia. Psychiatry Res 53: 31–40. [PubMed: 7991730]
6. Kay SR, Sevy S (1990): Pyramidical model of schizophrenia. Schizophr Bull 16: 537–545. [PubMed: 2287938]
7. Emsley R, Rabinowitz J, Torreman M, Group R-I-EPGW (2003): The factor structure for the Positive and Negative Syndrome Scale (PANSS) in recent-onset psychosis. Schizophr Res 61: 47–57. [PubMed: 12648735]
8. Van den Oord EJ, Rujescu D, Robles JR, Giegling I, Birrell C, József Bukszár, et al. (2006): Factor structure and external validity of the PANSS revisited. Schizophr Res 82: 213–223. [PubMed: 16229988]
9. Kim JH, Kim SY, Lee J, Oh KJ, Kim YB, Cho ZH (2012): Evaluation of the factor structure of symptoms in patients with schizophrenia. Psychiatry Res 197: 285–289. [PubMed: 22364933]
10. Levine SZ, Rabinowitz J (2007): Revisiting the 5 dimensions of the Positive and Negative Syndrome Scale. J Clinical Psychopharmacology 27: 431–436.
11. Wallwork R, Fortgang R, Hashimoto R, Weinberger D, Dickinson D (2012): Searching for a consensus five-factor model of the Positive and Negative Syndrome Scale for schizophrenia. Schizophr Res 137: 246–250. [PubMed: 22356801]
12. White L, Harvey PD, Opler L, Lindenmayer J (1997): Empirical assessment of the factorial structure of clinical symptoms in schizophrenia. Psychopathology 30: 263–274. [PubMed: 9353855]
13. van der Gaag M, Cuijpers A, Hoffman T, Remijsen M, Hijman R, de Haan L, et al. (2006a): The five-factor model of the Positive and Negative Syndrome Scale I: confirmatory factor analysis fails to confirm 25 published five-factor solutions. Schizophr Res 85: 273–279. [PubMed: 16730430]
14. van der Gaag M, Hoffman T, Remijsen M, Hijman R, de Haan L, van Meijel B, et al. (2006b): The five-factor model of the Positive and Negative Syndrome Scale II: a ten-fold cross-validation of a revised model. Schizophr Res 85(1–3): 280–287. [PubMed: 16730429]
15. Jiang J, Sim K, Lee J (2013): Validated five-factor model of positive and negative syndrome scale for schizophrenia in Chinese population. Schizophr Res 143(1): 38–43. [PubMed: 23148897]
16. Trninić V, Jelaska I, Štalec J (2013): Appropriateness and limitations of factor analysis methods utilized in psychology and kinesiology: Part II. Fizička kultura 67: 1–17.
17. Devarajan K (2008): Nonnegative matrix factorization: an analytical and interpretive tool in computational biology. PLoS Comput Biol 4: e1000029. [PubMed: 18654623]
18. Dickinson D, Pratt DN, Giangrande EJ, Grunnagle M, Orel J, Weinberger D, et al. Attacking heterogeneity in schizophrenia by deriving clinical subgroups from widely available symptom data. Schizophr Bull 44: 101–113. [PubMed: 28369611]
19. Dollfus S, Everitt B, Ribeyre JM, Assouly-Besse F, Sharp C, Petit M (1996): Identifying subtypes of schizophrenia by cluster analyses. Schizophr Bull 22: 545–555. [PubMed: 8873304]
20. Helmes E, and Landmark J (2003): Subtypes of schizophrenia: a cluster analytic approach. Can J Psychiatry 48: 702–708. [PubMed: 14674054]
21. Gottesman II, Gould TD (2003): The endophenotype concept in psychiatry: etymology and strategic intentions. Am J Psychiatry 160(4): 636–645. [PubMed: 12668349]

22. Walton E, Hibar DP, van Erp TG, et al. (2018): Prefrontal cortical thinning links to negative symptoms in schizophrenia via the ENIGMA consortium. Psychol Med 48(1): 82–94. [PubMed: 28545597]
23. Su TW, Hsu TW, Lin YC, Lin CP (2015): Schizophrenia symptoms and brain network efficiency: A resting-state fMRI study. Psychiatry Res: Neuroimaging 34(2): 208–218.
24. Varikuti DP, Genon S, Sotiras A, et al. (2018): Evaluation of non-negative matrix factorization of grey matter in age prediction. NeuroImage 173: 394–410. [PubMed: 29518572]
25. Sotiras A, Resnick SM, Davatzikos C (2015): Finding imaging patterns of structural covariance via non-negative matrix factorization. NeuroImage 108: 1–16. [PubMed: 25497684]
26. Sotiras A, Toledo JB, Gur RE, Gur RC, Satterthwaite TD, Davatzikos C (2017): Patterns of coordinatedcortical remodeling during adolescence and their associations with functional specialization and evolutionary expansion. Proc Natl Acad Sci U S A 114(13): 3527–3532. [PubMed: 28289224]
27. Yang Z, Oja E (2010): Linear and nonlinear projective nonnegative matrix factorization. IEEE Trans Neural Netw 21: 734–749. [PubMed: 20350841]
28. Bartels-Velthuis AA, Visser E, Arends J, Pijnenborg GHM, Wunderink L, Jörg F, et al. (2018): Towards a comprehensive routine outcome monitoring program for people with psychotic disorders: The Pharmacotherapy Monitoring and Outcome Survey (PHAMOUS). Schizophr Res 197: 281–287. [PubMed: 29395613]
29. Liemburg EJ, Nolte IM, Klein HC, Knegtering H (2018): Relation of inflammatory markers with symptoms of psychotic disorders: a large cohort study. Prog Neuropsychopharmacol Biol Psychiatry 86 :89–94. [PubMed: 29778547]
30. Hubert L, Arabie P (1985): Comparing partitions. J Classif 12: 193–218.
31. Meilă M (2007): Comparing clusterings-an information based distance. J Multivar Anal 98: 873–895.
32. Raguideau S, Plancade S, Pons N, Leclerc M, Laroche B (2016): Inferring Aggregated Functional Traits from Metagenomic Data Using Constrained Non-negative Matrix Factorization: Application to Fiber Degradation in the Human Gut Microbiota. PLoS Comput Biol 12: e1005252. [PubMed: 27984592]
33. Gardner DM, Murphy AL, O'Donnell H, Centorrino F, Baldessarini RJ (2010): International consensus study of antipsychotic dosing. Am J Psychiatry 167: 686–693. [PubMed: 20360319]
34. Lawson RG, Jurs PC (1990): New index for clustering tendency and its application to chemical problems. J Chem Inf Comput Sci 30: 36–41.
35. Bezdek JC (1981): Objective function clustering In: Pattern recognition with fuzzy objective function algorithms. Springer; pp 43–93.
36. Campello RJ, Hruschka ER (2006): A fuzzy extension of the silhouette width criterion for cluster analysis. Fuzzy Sets Syst 157: 2858–2875.
37. Xie XL, Beni G (1991): A validity measure for fuzzy clustering. IEEE Trans Pattern Anal Mach Intell 13: 841–847.
38. Fraley C, Raftery AE (2002): Model-based clustering, discriminant analysis, and density estimation. J Am Stat Assoc 97: 611–631.
39. Kaiser J (2007): An exact and a Monte Carlo proposal to the Fisher-Pitman permutation tests for paired replicates and for independent samples. Stata J 7: 402–412.
40. Schaefer A, Kong R, Gordon EM, Laumann T, Zuo XN, Holmes A, et al. (2017): Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. Cereb Cortex 28: 3095–3114.
41. Fan L, Li H, Zhuo J, Zhang Y, Wang J, Chen L, et al. (2016): The human brainnetome atlas: a new brain atlas based on connectional architecture. Cereb Cortex 26: 3508–3526. [PubMed: 27230218]
42. Varikuti DP, Hoffstaedter F, Genon S, Schwender H, Reid AT, Eickhoff SB (2017): Resting-state test–retest reliability of a priori defined canonical networks over different preprocessing steps. Brain Struct Funct 222: 1447–1468. [PubMed: 27550015]
43. Snoek L, Miletic S, Scholte HS (2019): How to control for confounds in decoding analyses of neuroimaging data. Neuroimage 184: 741–760. [PubMed: 30268846]

44. Genon S, Reid A, Langner R, Amunts K, and Eickhoff SB. (2018): How to characterize the function of a brain region. Trends Cogn Sci 22(4): 350–364. [PubMed: 29501326]
45. Rodriguez-Jimenez R, Bagney A, Mezquita L, Martinez-Gras I, Sanchez-Morla E, Mesa N, et al. (2013): Cognition and the five-factor model of the positive and negative syndrome scale in schizophrenia. Schizophr Res 143: 77–83. [PubMed: 23201306]
46. Choi EY, Yeo BTT, Buckner RL (2012): The organization of the human striatum estimated by intrinsic functional connectivity. J Neurophysiol 108(8): 2242–2263. [PubMed: 22832566]
47. Lehoux C, Gobeil M-H, Lefèbvre A-A, Maziade M, Roy M-A (2009): The five-factor structure of the PANSS: a critical review of its consistency across studies. Clin Schizophr Relat Psychoses 3: 103–110.
48. Nishimura Y, Takizawa R, Muroi M, Marumo K, Kinou M, Kasai K (2011): Prefrontal cortex activity during response inhibition associated with excitement symptoms in schizophrenia. Brain Res 1370: 194–203. [PubMed: 21059348]
49. Oh J, Chun J-W, Jo HJ, Kim E, Park H, Lee B, et al. (2015): The neural basis of a deficit in abstract thinking in patients with schizophrenia. Psychiatry Res Neuroimaging 234: 66–73.
50. Wurm LH, Fisicaro SA (2014): What residualizing predictors in regression analyses does (and what it does not do). J Mem Lang 72: 37–48.
51. Kay SR, Singh MM (1989): The positive-negative distinction in drug-free schizophrenic patients: Stability, response to neuroleptics, and prognostic significance. Arch Gen Psychiatry 46: 711–718. [PubMed: 2568824]
52. Andreasen NC, Olsen S (1982): Negative v positive schizophrenia. Definition and validation. Arch Gen Psychiatry 39: 789–794. [PubMed: 7165478]
53. Kendler KS, Gruenberg AM, Tsuang MT (1985): Subtype stability in schizophrenia. Am J Psychiatry 142: 827–832. [PubMed: 4014504]
54. Deister A, Marneros A (1993): Long-term stability of subtypes in schizophrenic disorders: a comparison of four diagnostic systems. Eur Arch Psychiatry Clin Neurosci 242: 184–190. [PubMed: 8461344]
55. Kulhara P, Chandiramani K (1990): Positive and negative subtypes of schizophrenia: a follow-up study from India. Schizophr Res 3(2): 107–116. [PubMed: 2278975]
56. Schilbach L, Derntl B, Aleman A, Caspers S, Clos M, Diederen KMJ, et al. (2016): Differential patterns of dysconnectivity in mirror neuron and mentalizing networks in schizophrenia. Schizophr Bull 42: 1135–1148. [PubMed: 26940699]
57. Vercammen A, Knegtering H, den Boer JA, Liemburg EJ, Aleman A (2010): Auditory hallucinations in schizophrenia are associated with reduced functional connectivity of the temporo-parietal area. Biol Psychiatry 67: 912–918. [PubMed: 20060103]
58. Derntl B, Finkelmeyer A, Eickhoff S, Kellermann T, Falkenberg DI, Schneider F, et al. (2010): Multidimensional assessment of empathic abilities: neural correlates and gender differences. Psychoneuroendocrinology 35: 67–82. [PubMed: 19914001]
59. Nenadic I, Yotter RA, Sauer H, Gaser C (2015): Patterns of cortical thinning in different subgroups of schizophrenia. Br J Psychiatr 206: 479–483.
60. Shaffer JJ, Peterson MJ, McMahon MA, Bizzell J, Calhoun V, van Erp TGM, et al. (2015): Neural correlates of schizophrenia negative symptoms: distinct subtypes impact dissociable brain circuits. Mol neuropsychiatry 1: 191–200. [PubMed: 27606313]
61. Orban P, Dansereau C, Desbois L, Mongeau-Pérusse V, Giguère CE, Nguyen H, et al. (2018): Multisite generalizability of schizophrenia diagnosis classification based on functional brain connectivity. Schizophr Res 192: 167–171. [PubMed: 28601499]
62. Mikolas P, Melicher T, Skoch A, Slovakova A, Matejka M, Rydlo J, et al. (2016): Diagnostic classification of patients with first-episode schizophrenia spectrum disorders from resting-fMRI. Eur Neuropsychopharmacology 26: S490.
63. Rozycki M, Satterthwaite TD, Koutsouleris N, Erus G, Doshi J, Wolf DH, et al. (2017): Multisite machine learning analysis provides a robust structural imaging signature of schizophrenia detectable across diverse patient populations and within individuals. Schizophr Bull 44: 1035–1044.

64. Sun H, Lui S, Yao L, Deng W, Xiao Y, Zhang W, et al. (2015): Two patterns of white matter abnormalities in medication-naive patients with first-episode schizophrenia revealed by diffusion tensor imaging and cluster analysis. JAMA Psychiatry 72: 678–686. [PubMed: 25993492]

65. Zuo XN, Kelly C, Adelstein JS, Klein DF, Castellanos FX, Milham MP (2010): Reliable intrinsic connectivity networks: test–retest evaluation using ICA and dual regression approach. Neuroimage 49: 2163–2177. [PubMed: 19896537]

66. Varikuti DP, Hoffstaedter F, Genon S, Schwender H, Reid AT, Eickhoff SB (2017): Resting-state test–retest reliability of a priori defined canonical networks over different preprocessing steps. Brain Struct Funct 222(3): 1447–1468. [PubMed: 27550015]

67. Bright MG, Tench CR, Murphy K (2017): Potential pitfalls when denoising resting state fMRI data using nuisance regression. Neuroimage 154: 159–168. [PubMed: 28025128]

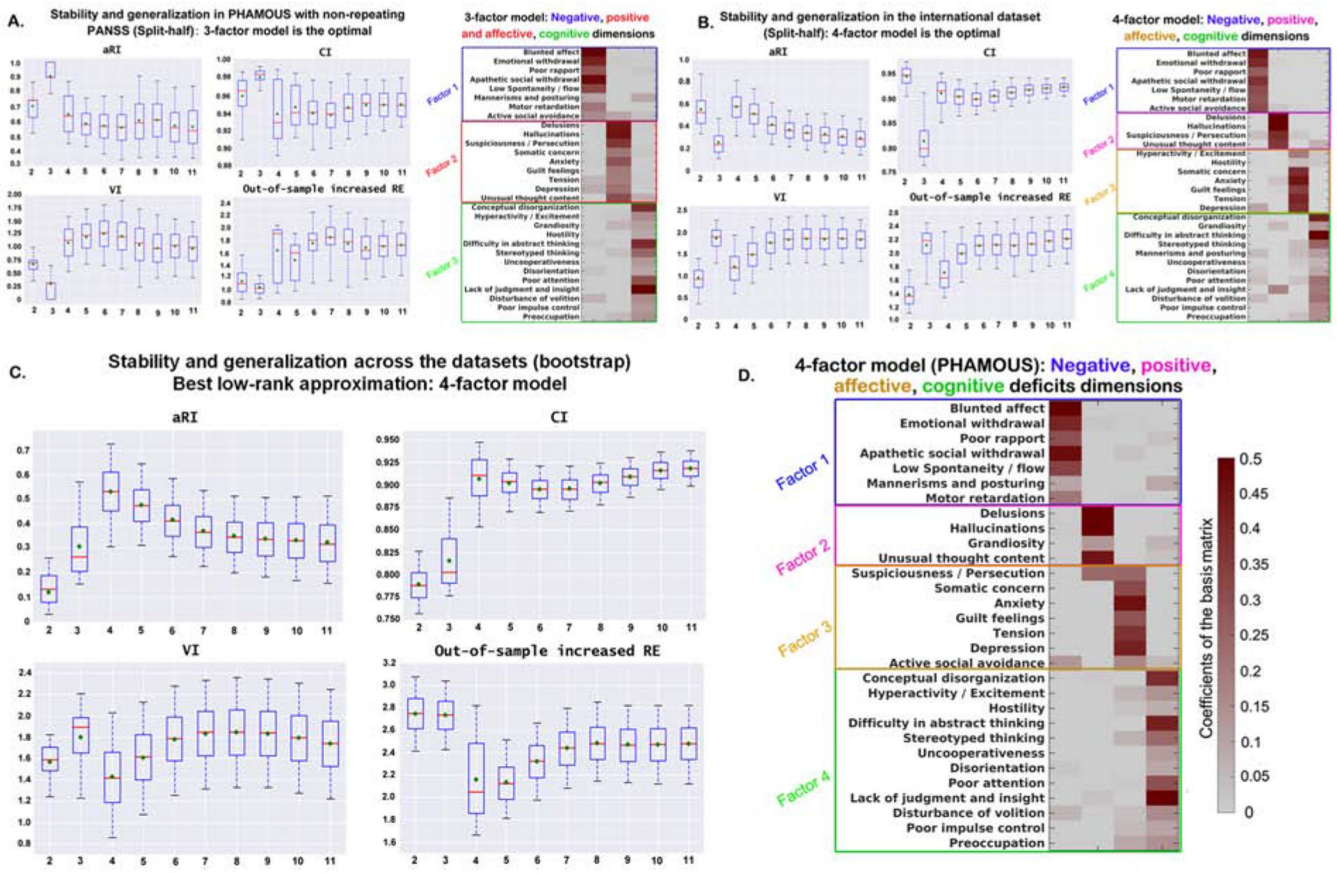

Figure 1. Split-half cross-validation (10,000 repetitions) of stability and generalizability of the factor-solutions derived by OPNMF[a]

[a]The three indices; aRI, VI and CI demonstrate the factor stability, while out-of-sample increased RE reflects the performance of generalizability. Box-plots show stability and generalizability results of the factor solutions. Higher values for aRI and CI (upper row) indicate higher stability. Lower values for VI and out-of-sample increase in RE (bottom row) indicate better stability and generalizability, respectively. For the box-plots, the red line depicts the median, the green diamond depicts the mean, and the whiskers represent the 5th and 95th percentiles. For the factor-models, the weight of an item in assigning to a specific psychopathological factor (columns of the matrix) is color coded according to the coefficients by a heat color map, from grey (minimum) to dark red (maximum). aRI, adjusted Rand index; VI, variation of information; CI, concordance index; RE, reconstruction error. **Panel A) shows the best factor solution derived from the PHAMOUS data (1545 patients)**. According to the four aforementioned evaluation indices, a three-factor model was indicated as the best since both of the mean and median values for VI and out-of-sample increase in RE achieve the lowest, and the aRI and CI reach the highest at that point. **Panel B) shows the best factor solution derived from the international sample (490 patients)**. As shown, four factors is the optimal solution, as mean and median values of VI and out-of-sample increase in RE achieve the local minimum, while the aRI reaches maxima and the CI reaches a local maximum. **Panel C) shows the best factor solution identified by the bootstrap comparison of the two datasets**

(PHAMOUS vs. international). As shown, a four-factor solution is optimal, as the mean and median values of adjusted RI and CI reach the maximum, while the mean and median values of VI and median value of out-of-sample increase in RE achieve the minimum. **Panel D) shows the most stable and generalizable four-factor structure derived from the PHAMOUS sample, serving as the best basis for future studies**. This four-factor model consists of a negative (factor 1), a positive (factor 2), an affective (factor 3), and a cognitive (factor 4) factor which were named based on the items they contained.

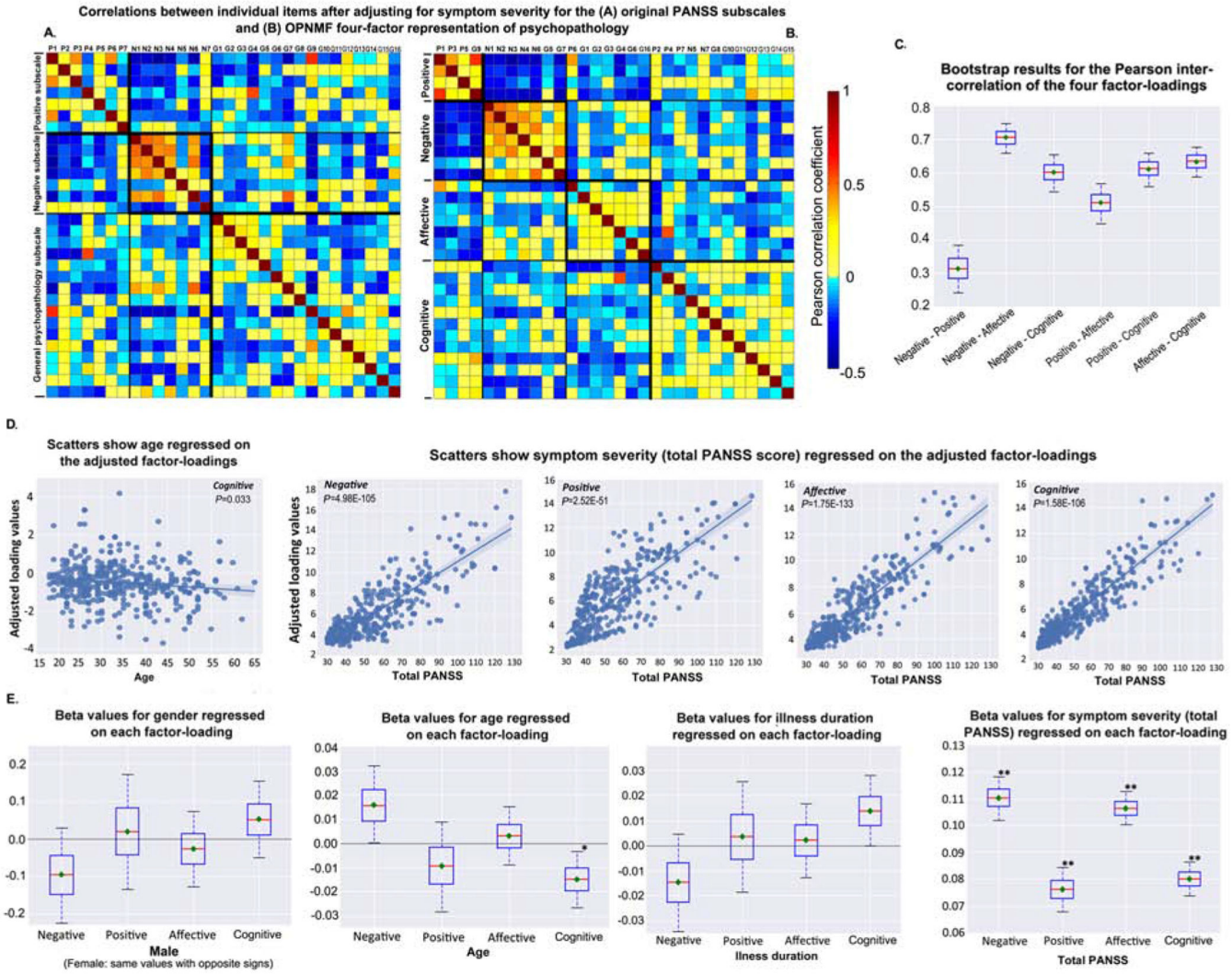

Figure 2. Inter-item correlations, relationship between factors, socio-demographic and clinical information[a]

[a]The four-factor structure, derived from the PHAMOUS sample with initial measure of PANSS scores, was adopted as the reference on which the international sample was projected to derive the factor-loadings. **Heat maps A-B)** show inter-item correlations for (A) the original PANSS subscales, and (B) the current OPNMF four-factor representation of psychopathology after controlling for symptom severity (total PANSS score). Correlation strength is color-coded (light yellow to red: positive correlations; cyan to blue: negative correlations). **Box-plot C)** shows the bootstrap results (repeated 10,000 times) for the Pearson correlations between the four factor-loadings. Bootstrap samples were drawn with replacement from the original international sample, and then the correlation analysis was done on them. The red line depicts the median, the green diamond depicts the mean, and the whiskers represent the 5th and 95th percentiles. **Graphs D-E)** show effects of socio-demographic and clinical features on the four factor-loadings. **D).** Scatter plots show 4-way ANOVA results of the significant negative association between age (adjusted for gender, illness duration, and total PANSS score) and the cognitive loading ($p = .033$), as well as the

significant positive associations between the symptom severity (total PANSS score) and the four factor-loadings (negative: $p = 4.98E-105$; positive: $p = 2.52E-51$; affective: $p = 1.75E-133$; cognitive: $p = 1.58E-106$) after adjusting for age, gender, and illness duration. Regression lines are depicted with a 95% confidence interval on the fitted values. **E)** shows bootstrap results for the 4-way ANOVA analysis. Bootstrap samples were drawn with replacement from the original international sample and then the ANOVAs were done on them. Boxes refer to the beta values. The red line depicts the median, the green diamond depicts the mean, and the whiskers represent the 5th and 95th percentiles.* (*Mdn*, $p < .05$), ** (*M* and *Mdn*, $p < .05$).

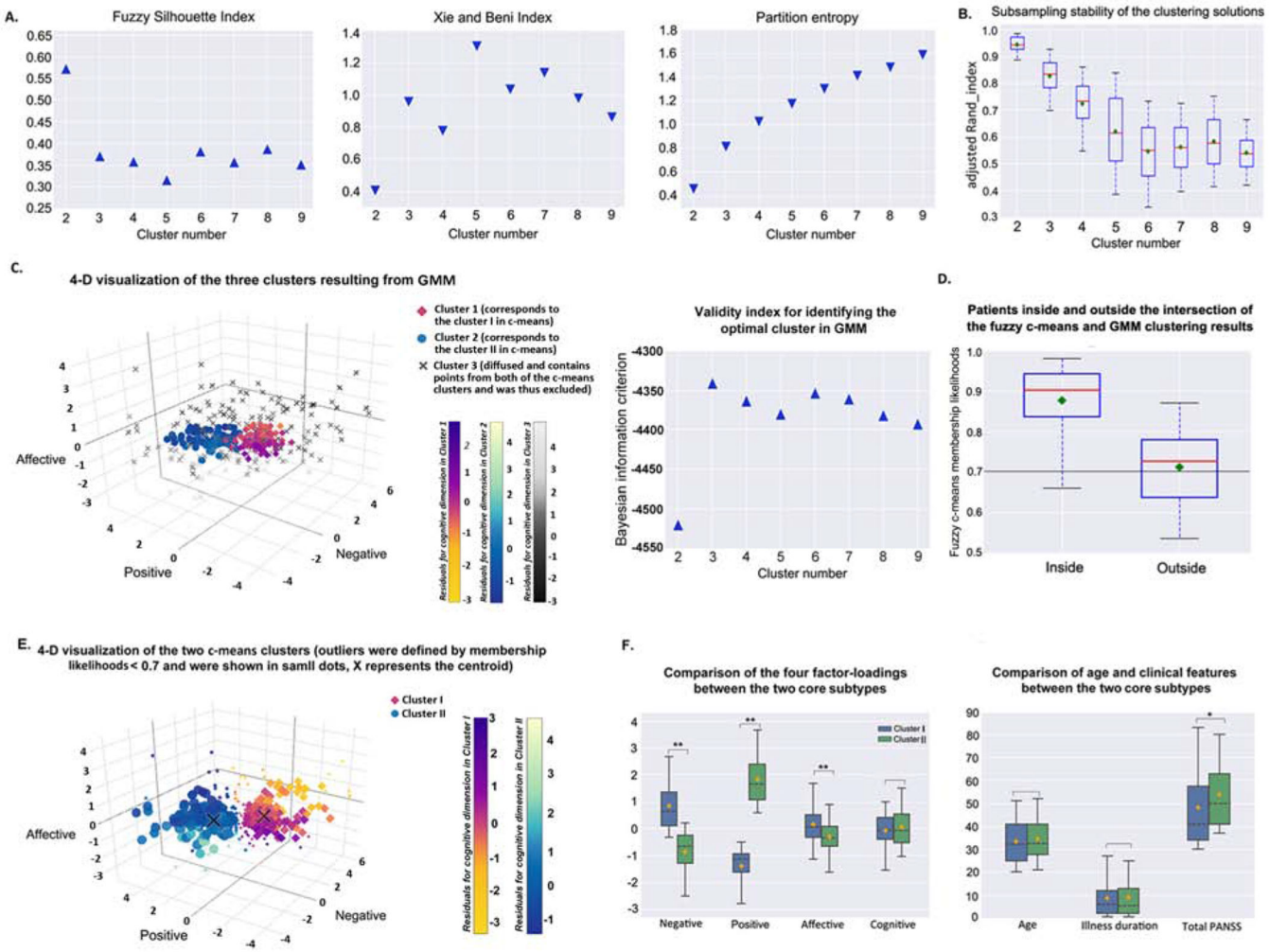

Figure 3. Fuzzy c-means clustering results of patient subgroups based on the loadings of the generalizable four-factor structure[a]

[a]**Panel A)** shows the internal validity indices used for determining the optimal cluster number. Higher values of FSI (in triangle) and lower values of XB and PE (in inverted triangle) indicate a better clustering quality. The maximum for FSI and the minimums for XB and PE all suggested a two-cluster solution. FSI and XB reflect the compactness and separation of the generated clusters, while PE reflects the fuzziness of the cluster partition, i.e., the uncertainty of the patients to be assigned to a certain cluster. **Box-plot B)** shows results of the assessment of clustering stability based on the subsampling technique. The cluster number two reaches the highest aRI. aRI reflects the convergent assignment of the patient-pairs to the clusters between the sub-samples and the original sample. **C)** Four-dimensional visualization of the optimal three GMM clusters determined by the Bayesian information criterion (a higher value indicates a better clustering solution). Magnitude of the cognitive loading was color-coded differently for the three clusters (cluster 1, corresponds to the cluster I (i.e., subtype A) in fuzzy c-means, yellow to Modena; cluster 2, corresponds to the cluster II (i.e., subtype B) in fuzzy c-means, blue to shallow flaxe; cluster 3, i.e., the excluded diffused cluster which would not present any specific subtype, black to light grey).

Boxplot D) shows the fuzzy c-means membership likelihoods of the patients inside and outside the intersection of the c-means and GMM clustering results. The black line indicates a heuristic cutoff of 0.7. **Panel E)** shows a four-dimensional visualization of the optimal fuzzy c-means two-cluster solution. Ambiguous assignments were defined by membership likelihoods < 0.7, which was selected by interacting with GMM. Those subtype ambiguous patients are shown in small dots, X represents the centroid. Magnitude of the cognitive loading is color-coded differently for the two clusters (cluster I, yellow to Modena; cluster II, blue to shallow flaxe). Grouped box-plots **F)** show the between-subtype (without subtype ambiguous patients) comparison results of the four factor-loadings, age, illness duration, and total PANSS score. Cluster I is dominated by negative and affective symptoms (i.e., subtype A), cluster II is significantly prominent in positive symptom expressions (i.e., subtype B). The black dashed line depicts the median, the yellow diamond depicts the mean, and the whiskers represent the 5th and 95th percentiles. *$p < .01$; **$p < .001$. FSI, fuzzy silhouette index; XB, Xie and Beni index; PE, partition entropy; GMM, Gaussian mixture modeling.

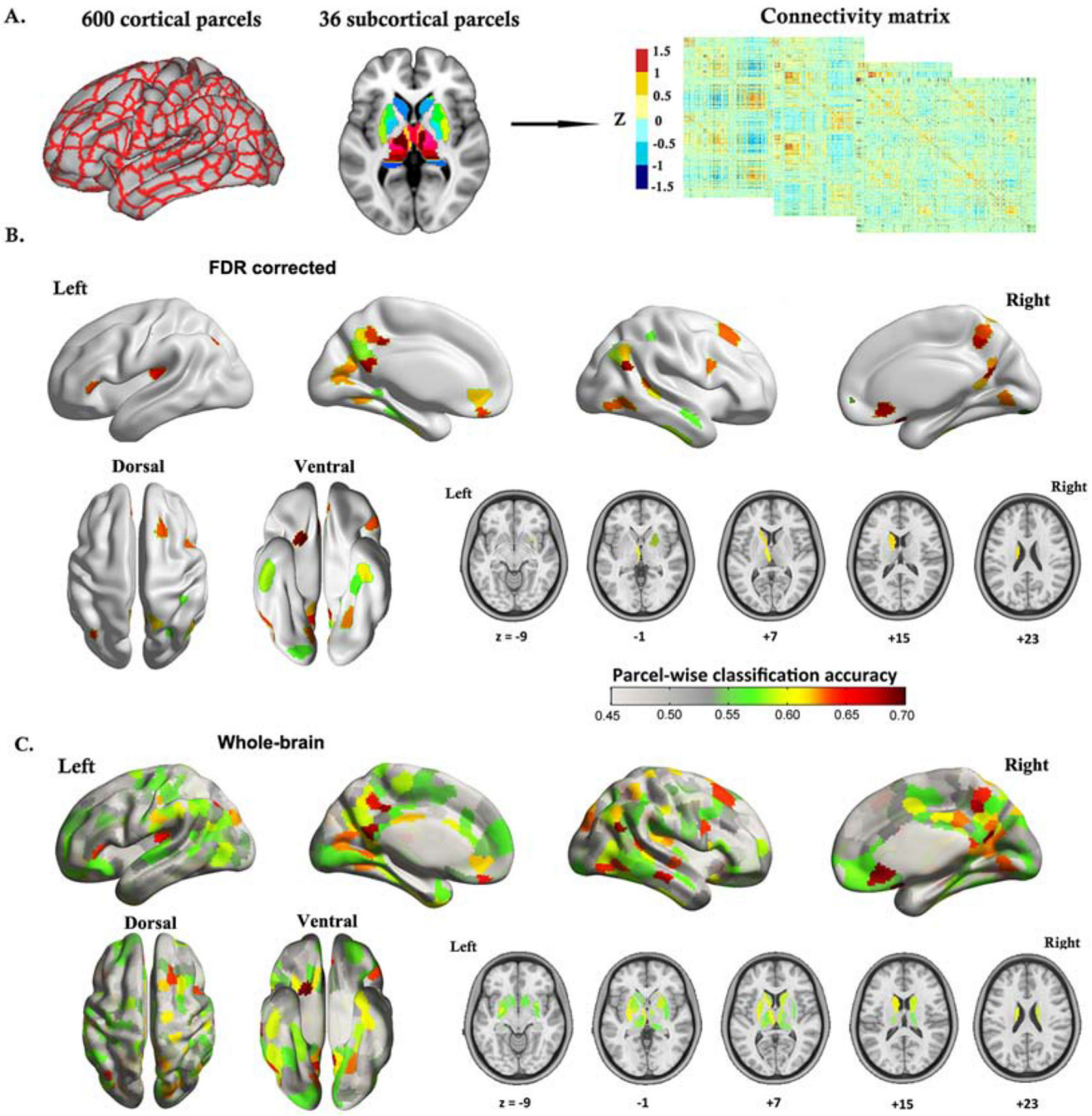

Figure 4. Classifying psychopathological subtypes from resting-state functional connectivity[3]

[a]**Panel A)** illustrates the brain parcellation scheme (600 cortical parcels plus 36 subcortical parcels), and the resting-state functional connectivity matrix that was constructed based on this parcellation system. In parcel-wise classification analysis, one column of the connectivity matrix was taken to represent the functional connectivity pattern for a single parcel. **B)** Cortical surface rendering and subcortical axial slices show parcel-wise classification results for those parcels which survived FDR correction ($q < .05$), demonstrating a neurobiological divergence between the two identified psychopathological

subtypes of schizophrenia. **C)** Cortical surface rendering and subcortical axial slices show parcel-wise classification results for the whole brain. The balanced classification accuracy is color-coded from light grey to dark red.

Table1.

Demographic and clinical characteristics of the patients with schizophrenia

Characteristics	PHAMOUS sample (N=1545)	International dataset from 9 centers (N=490)	International dataset with imaging (N=147)	Statistics	p-value
Demographic					
Age (years)[a]	44.15 (11.42)	33.82 (10.28)	34.89 (11.67)	183.51	**<.001**
Gender (male/female)	1108/437	333/157	102/45	2.45	.292
Illness during (years)[b]	18.22 (10.54)	9.13 (8.98)	11.37 (10.36)	134.71	**<.001**
PANSS					
Positive[c]	12.48 (4.91)	14.24 (5.76)	15.36 (5.50)	37	**<.001**
Negative	14.60 (6.20)	14.67 (7.21)	15.07 (6.06)	0.375	.687
General[d]	26.70 (8.16)	29.10 (11.34)	30.93 (10.97)	23.67	**<.001**
Illness severity (Total PANSS)[e]	53.78 (16.35)	58.01 (21.87)	61.36 (19.57)	19.48	**<.001**
P3 item (hallucinations)[f]	2.30 (1.47)	2.66 (1.83)	3.22 (1.91)	28.18	**<.001**
Medication[g]					
Atypical antipsychotics	**NA**	167 (34.1%)	110 (74.8%)		
Typical antipsychotics	**NA**	26 (5.3%)	8 (5.4%)		
Both A & T	**NA**	16 (3.3%)	9 (6.1%)		
None or unknown	**NA**	281 (57.3%)	20 (25.9%)		
Current antipsychotic medication[h]	**NA**	19.64 (14.15)	19.30 (12.57)		

Data are mean (SD), or n (%). *p*-values in bold indicate a significance of $p < .05$. Except for gender, which was based on χ^2 test, other statistics were all based on one-way ANOVA analyses. Of note, since the detailed medication information was missing for several patients in different proportions for those with or without imaging data in the international dataset, statistical comparisons were not conducted.

Post-hoc analysis after one-way ANOVAs showing significant pair-wise differences among the three datasets:

[a]PHAMOUS > International sample = International sample with imaging at $p < .05$, Bonferroni corrected.

[b]PHAMOUS > International = International with imaging at $p < .05$, Bonferroni corrected.

Information of illness duration was available for 1326 patients in the PHAMOUS sample, 393 patients in the international sample

[c]PHAMOUS < International sample = International sample with imaging at p< .05, Bonferroni corrected.

[d]PHAMOUS < International sample = Internationalsample with imaging at p< .05, Bonferroni corrected.

[e]PHAMOUS < International sample = Internationalsample with imaging at p< .05, Bonferroni corrected.

[f]PHAMOUS < International sample < Internationalsample with imaging at p< .05, Bonferroni corrected.

[g]211 patients with medication information in the whole international sample. 149 patients with also illness duration information were included in the ANOVA analysis.

[h]demonstrated in olanzapine-equivalent dosage (mg/day).

Abbreviations: PANSS, the Positive and Negative Syndrome Scale; A, atypical antipsychotics; T, typical antipsychotics; NA, not available.

KEY RESOURCES TABLE

Resource Type	Specific Reagent or Resource	Source or Reference	Identifiers	Additional Information
Add additional rows as needed for each resource type	Include species and sex when applicable.	Include name of manufacturer, company, repository, individual, or research lab. Include PMID or DOI for references; use “this paper” if new.	Include catalog numbers, stock numbers, database IDs or accession numbers, and/or RRIDs. RRIDs are highly encouraged; search for RRIDs at https://scicrunch.org/resources.	Include any additional information or notes if necessary.
Deposited Data; Public Database	COBRE	http://fcon_1000.projects.nitrc.org/indi/retro/cobre.html	RRID:SCR_010482	
Software; Algorithm	a Dimensions and Clutsering Tool for assessing schizophrenia Symptomatology (DCTS)	this paper; http://webtools.inm7.de/sczDCTS	N/A	
Software; Algorithm	Matlab R2016a	Mathworks	RRID:SCR_001622	
Software; Algorithm	R version 3.4.2	http://www.r-project.org/	RRID:SCR_001905	
Software; Algorithm	mclust package (version 5.3) in R version 3.4.2	http://CRAN.R-project.org/package=mclust	N/A	
Software; Algorithm	LIBSVM	https://www.csie.ntu.edu.tw/~cjlin/libsvm/	RRID:SCR_010243	
Software; Algorithm	SPM12	Wellcome Centre for Human Neuroimaging; https://www.fil.ion.ucl.ac.uk/spm/software/spm12/	RRID:SCR_007037	
Software; Algorithm	CAT12	Structural Brain Mapping Group at the Departments of Psychiatry and Neurology, University of Jena; http://www.neuro.uni-jena.de/cat/	N/A	
Other	Shaefer local-global cortical parcellation	https://github.com/ThomasYeoLab/CBIG/tree/master/stable_projects/brain_parcellation/Schaefer2018_LocalGlobal	N/A	
Other	Brainnetome atlas	http://atlas.brainnetome.org	RRID:SCR_014091	