

## ARTICLE OPEN



# Transdiagnostic types of formal thought disorder and their association with gray matter brain structure: a model-based cluster analytic approach

Frederike Stein <sup>1,2</sup>✉, Anna Merle Gudjons<sup>1,2</sup>, Katharina Brosch <sup>1,2,3</sup>, Luca Mira Keunecke<sup>1</sup>, Julia-Katharina Pfarr<sup>1,2</sup>, Lea Teutenberg<sup>1,2</sup>, Florian Thomas-Odenthal <sup>1,2</sup>, Paula Usemann<sup>1,2</sup>, Hanna Wersching<sup>1</sup>, Adrian Wroblewski<sup>1,2</sup>, Kira Flinkenflügel <sup>4</sup>, Janik Goltermann <sup>4</sup>, Dominik Grotegerd <sup>4</sup>, Susanne Meinert <sup>4,5</sup>, Katharina Thiel<sup>4</sup>, Alexandra Winter <sup>4</sup>, Nina Alexander<sup>1,2</sup>, Tim Hahn <sup>4</sup>, Hamidreza Jamalabadi <sup>1,2</sup>, Andreas Jansen<sup>1,2,6</sup>, Axel Krug <sup>7</sup>, Igor Nenadić <sup>1,2</sup>, Benjamin Straube<sup>1,2</sup>, Udo Dannlowski<sup>4</sup> and Tilo Kircher <sup>1,2</sup>

© The Author(s) 2025

Formal thought disorder (FTD) is a complex syndrome affecting language and thought processes in psychotic and affective disorders. Clustering (i.e., identification of data-driven clinical subtypes) establishes latent (sub-) structures into psychopathological syndromes. A latent profile analysis (LPA) of FTD symptoms was conducted in 1 032 patients diagnosed with Schizophrenia-Spectrum-Disorders (n = 107), Major Depressive (n = 800), and Bipolar Disorder (n = 125). Clusters were compared for cognition and psychopathology. Associations with gray matter volume (GMV) and cortical surface (gyrification, cortical complexity, sulcal depth) were explored using T1-weighted MRI data, analyzed with CAT12. Robustness-analyses in an age- and sex-matched subsample (n = 321) with the same n for each diagnosis (n = 107) were applied. LPA revealed 4 transdiagnostic clusters: *minimal* FTD, *poverty*, *inhibition*, *severe* FTD that remained stable in an age- and sex-matched subsample and in each diagnosis separately. Patients exhibiting *severe* FTD compared to *minimal* FTD showed GMV reductions in the right superior and middle frontal gyri. *Inhibition* showed a GMV reduction in the right inferior and middle temporal gyri, and fusiform gyrus compared with *minimal* and *severe* FTD. Sulcal depth was reduced around the left insula, superior temporal sulcus and temporal pole in the *poverty* cluster, and in the bilateral insula in the *severe* cluster, both compared to the *inhibition* cluster. No results for cortical thickness, gyrification, and complexity were found. Results from the total sample could be replicated in the matched subsample. Our results unravel the clinical heterogeneity of FTD psychopathology across affective and psychotic disorders. Associations of FTD clusters with neuroanatomical substrates imply language-related brain structures being involved in thought and language impairment.

*Molecular Psychiatry* (2025) 30:4286–4295; <https://doi.org/10.1038/s41380-025-03009-w>

## INTRODUCTION

Impairments in speech production and aberrations in the form of thought are clinically termed formal thought disorder (FTD). FTD is a core feature of Schizophrenia (SZ) but also occurs in other disorders including Schizoaffective Disorder (SZA), Major Depressive Disorder (MDD), or Bipolar Disorder (BD) [1–7]. The presence of FTD is considered as a marker of illness severity, FTD can predict new episodes, and (re-) hospitalization is both more likely and significantly longer for patients with FTD [8]. Moreover, it is associated with poor social functioning [9] and impacts the perceived quality of life in SZ and SZA patients (henceforth referred to as schizophrenia-spectrum-disorders, SSD) [10]. Of importance is the widely unacknowledged fact, that the prevalence of FTD in MDD ranges from 36 to 53%, in SZ from 50 to 81%, and in SZA up to 60% [8]. Moreover, a meta-analysis reported no quantitative difference of FTD in acute BD and SZ [2].

FTD is a multifaceted syndrome. Studies examining the factorial structure of FTD symptomatology in SZ have identified factors ranging from two [2, 11] to six positive/disorganized factors [12], while one negative domain was consistently identified [11]. Positive FTD factors mostly include symptoms such as derailment and pressure of speech. In contrast, the negative domain can be characterized by a decrease in the amount of produced speech [3, 13]. Yet, research on both FTD symptomatology and underlying neurobiological mechanisms was almost exclusively performed in SZ. We have investigated transdiagnostic factors of FTD across affective and psychotic disorders [6, 7], lately using FTD symptoms from the Scales for the Assessment of Positive and Negative Symptoms (SAPS; SANS) [14, 15]. Hereof, three factors (i.e., disorganization, emptiness, incoherence) were delineated across MDD, BD, and SSD patients [7]. This factor solution was replicated using a different FTD rating scale [16].

<sup>1</sup>Department of Psychiatry and Psychotherapy, University of Marburg, Marburg, Germany. <sup>2</sup>Center for Mind, Brain and Behavior, University of Marburg, Marburg, Germany. <sup>3</sup>Institute of Behavioral Science, Feinstein Institutes for Medical Research, Manhasset, NY, USA. <sup>4</sup>Institute for Translational Psychiatry, University of Münster, Münster, Germany. <sup>5</sup>Institute for Translational Neuroscience, University of Münster, Münster, Germany. <sup>6</sup>Core-Facility Brainimaging, Faculty of Medicine, University of Marburg, Marburg, Germany. <sup>7</sup>Department of Psychiatry und Psychotherapy, University Hospital, Bonn, Germany. ✉email: frederike.stein@staff.uni-marburg.de

Received: 1 July 2024 Revised: 12 March 2025 Accepted: 27 March 2025  
Published online: 11 April 2025

Lately, there has been an increasing interest in data-driven and statistical tools such as clustering techniques and finite mixture modelling to assess clinical heterogeneity in psychiatric disorders with regard to symptoms, clinical course and underlying biological mechanisms [17]. In contrast to the above mentioned FTD symptom factor analyses, unsupervised clustering attempts to group patients based on indicator variables whereas factor analyses group a set of variables, in this case symptoms. While factor analyses can inform about dimensions of an underlying construct across a given population, cluster analyses help to better understand heterogeneity within a given population by identifying subgroups [18] and therefore allows a better patient typology of the clinical reality. Clustering algorithms have been employed to explore sub-structures of various psychopathological phenomena, e.g., for subtypes of depression [19] or cognitive subtypes in SZ [20] as well as transdiagnostic clustering approaches, focusing on cognition [21], psychotic symptoms [22] or other symptom domains and risk factors [23]. Although FTD is a prominent syndrome not only in SSD but also in BD and MDD, cluster analytic approaches to disentangle FTD symptom and severity subtypes across these disorders are lacking.

There are to our knowledge only two studies that have investigated FTD subtype clustering, restricted diagnostically to recent onset psychosis patients. By applying three different clustering algorithms (k-means, hierarchical, partitioning around medoids), the study by Oeztuerk et al. identified two subgroups distinguished by high vs. low FTD in recent onset psychosis and associated the former with lower neurocognitive performance, social and occupational functioning but not with global disease severity [24]. Subsequently, these two subtypes were predicted using GMV and functional MRI measures by Buciuman et al. The GMV and functional MRI measures of the salience, dorsal attention, visual, and ventral attention networks classified the two FTD subtypes with a combined multimodal balanced accuracy of 77% [25].

The vast majority of studies on the GMV correlates of FTD were done in SZ, however with heterogeneous findings. Among the most consistent results on the neuroanatomical correlates of FTD are GMV reductions in the bilateral superior temporal (STG) and inferior frontal gyri (IFG), partly corresponding to Wernicke's and Broca's language areas [3, 26–29]. This points to the assumption that structural aberrations of areas associated with language production and processing contribute to FTD. From a functional perspective, including both task-based and resting-state analyses, consistent activation changes have been observed in the left STG and in the posterior, ventral, and dorsal regions of the middle temporal gyrus (MTG) [30, 31].

From a more dimensional perspective, only few studies investigated associations of GMV and FTD across diagnoses. Hereof, a disorganization dimension was negatively correlated with the left middle occipital/angular gyrus, while emptiness (i.e., negative FTD) was negatively correlated with the left hippocampus/thalamus [7]. Moreover, we expanded this approach by investigating brain networks (i.e., GMV and white matter connections) being associated with these FTD dimensions across diagnoses [32]. Network-based analyses yielded subnetworks that predominantly comprised brain regions implicated in speech across both hemispheres [7, 32]. In contrast to these previous factor analytic studies from our group identifying dimensions (i.e., latent factors of FTD), the present study aims to elucidate subtypes of FTD by clustering individual patients based on FTD psychopathology into transdiagnostic groups.

Looking at other morphometric MRI measures such as cortical thickness, few studies are available investigating FTD in SZ and none across diagnoses. Using a network-based approach, Palaniyappan et al. (2020) reported the presence of positive FTD in SZ to be correlated with reduced cortical thickness in the language network located in the superior temporal cortex, and

also related positive FTD severity with cortical thickness in the fronto-parietal network [33]. A comprehensive understanding of FTD requires the integration of both volume- and surface-based morphometric analyses. GMV reflects macroscopic structural changes associated with neurodevelopmental or adulthood processes, while surface-based metrics such as cortical thickness, sulcal depth, and gyrification capture finer details of cortical morphology and folding [17, 18, 34–36], usually linked to early (genetic and fetal) neurodevelopment. By combining these complementary methods, our study offers novel insights into its neuroanatomical basis.

Despite the presence of FTD in MDD, BD and SZ [6], there are – to the best of our knowledge – no studies available on transdiagnostic FTD clustering across different illness stages and none on their neuro-structural correlates, particularly not in a large patient sample. The present study's aims were therefore (i) to identify FTD clusters (i.e., subtypes) in a large transdiagnostic sample (MDD, BD, SSD) using a model-based clustering algorithm; (ii) to characterize the FTD clusters with regard to psychopathology (other than FTD) and neurocognitive performance; and (iii) to explore associations between FTD clusters and GMV as well as cortical surface measures using MRI morphometric methods. For (i) it was expected that the sample would sub-divide into at least two clusters based on FTD symptomatology ranging from low to severe FTD. These clusters were (ii) hypothesized to differ in neuropsychological tests that have previously been shown to correlate with the presence of FTD in SZ, i.e., attention, executive function, working memory and verbal fluency. Finally, (iii) structural correlations in brain regions previously associated with negative and positive FTD in SZ patients, in particular in the extended language network, were expected across MDD, BD, and SSD diagnoses.

## MATERIAL AND METHODS

### Participants

For the present analyses, participants from MACS cohort of the FOR2107 consortium, a bi-central study on the neurobiology of affective and psychotic disorders [37] were included. The sample comprised a broad spectrum of acutely ill to remitted patients from the Departments of Psychiatry and Psychotherapy, University Hospital of Marburg, and the Institute for Translational Psychiatry in Münster, both Germany as well as local participating hospitals and outpatient clinics (for details see [37]). All procedures were approved by the local Ethics Committees according to the Declaration of Helsinki and written informed consent was obtained from each subject. Patients with verbal IQ < 80, history of head trauma or unconsciousness, current substance dependence, current intake of Benzodiazepines, and neurological illness were excluded. Subsequent to a quality check of the T1-weighted images and the remove of incomplete data, 1 032 patients (aged 18–65) who met DSM-IV-TR criteria (SCID-I diagnosed) for MDD ( $n = 800$ ,  $f = 529/m = 271$ ), BD ( $n = 125$ ,  $f = 66/m = 59$ ), SSD ( $n = 107$ ), SZ ( $n = 67$ ,  $f = 30/m = 37$ ) and SZA ( $n = 40$ ,  $f = 22/m = 18$ ) (see Table 1) were analyzed.

### Clinical and neuropsychological assessment

Psychopathology was assessed during a semi-structured clinical interview, incorporating SCID-I and various psychopathology scales. All interviewers were trained with the evaluation of the scales used. Positive and negative symptoms were assessed with the Scale for the Assessment of Positive Symptoms (SAPS) [15] and the Scale for the Assessment of Negative Symptoms (SANS) [14]. For the evaluation of manic and depressive symptoms, the Young Mania Rating Scale (YRMS) [38] and the Hamilton Rating Scale for Depression (HAM-D) [39] were employed, respectively. For the present analysis, the single items of the SANS alogia subscale and SAPS positive formal thought disorder subscale were used, as well as two items of the YRMS and one item of the HAM-D describing an impairment in the form of thought and language.

In addition, performance in several neurocognitive domains was assessed using a broad neuropsychological test battery. Verbal episodic memory was assessed with the German version of the California verbal learn- and memory retention test [40] including the sum of correct

**Table 1.** Sample characteristics (N = 1 032).

	MDD (n = 800)	BD (n = 125)	SSD (SZ/SZA) (n = 107)	Group comparison (F-value in brackets)
Age	36.55 (13.15)	41.04 (11.77)	38.07 (11.7)	$p < 0.001^a$ (8.23)
Sex	271m / 529f	59m / 66f	55m / 52f	$p < 0.001$
Years of education	13.22 (2.71)	14.08 (2.77)	12.46 (2.71)	$p < 0.001^b$ (9.91)
TIV	1560.46 (152.98)	1580.42 (146.19)	1580.23 (188.16)	$p = 0.255$ (1.37)
HAM-D sum	8.35 (6.38)	6.96 (5.9)	6.69 (5.79)	$p < 0.01^c$ (4.56)
YRMS sum	1.43 (2.1)	3.9 (5.99)	2.41 (4.73)	$p < 0.001^d$ (18.21)
SAPS sum	0.62 (2.01)	2.39 (4.36)	8.93 (11.48)	$p < 0.001^e$ (156.35)
SAPS positive formal thought disorder	0.32 (1.3)	1.74 (3.25)	2.79 (3.99)	$p < 0.001^e$ (77.28)
SAPS hallucinations	0.08 (0.5)	0.17 (0.79)	1.88 (4.13)	$p < 0.001^f$ (75.99)
SAPS delusions	0.16 (0.87)	0.3 (0.85)	3.94 (5.99)	$p < 0.001^f$ (154.86)
SANS sum	7.42 (8.67)	5.56 (7.13)	12.81 (11.54)	$p < 0.001^f$ (14.69)
SANS alogia (negative formal thought disorder)	0.49 (1.3)	0.61 (1.32)	1.75 (2.6)	$p < 0.001^f$ (33.1)
SANS anhedonia	2.9 (3.5)	1.94 (3.05)	2.78 (3.47)	$p < 0.001^g$ (4.24)

Mean (standard deviation), SANS (Scale for the Assessment of Negative Symptoms), SAPS (Scale for the Assessment of Positive Symptoms), YMRS (Young Mania rating scale, HAM-D (Hamilton rating scale for Depression), TIV (Total intracranial volume).

<sup>a</sup>MDD < BD.

<sup>b</sup>MDD, SSD < BD.

<sup>c</sup>BD, SSD < MDD.

<sup>d</sup>MDD < BD, SSD; SSD < BD.

<sup>e</sup>MDD < BD, SSD; BD < SSD.

<sup>f</sup>SSD > BD, MDD.

<sup>g</sup>MDD > BD.

remembered words. Verbal working memory was tested using forward and reverse letter-number-spans [41] and visuospatial memory was assessed with a block span task [42]. Executive functions were evaluated using the Trail-Making-Test (TMT version A and B) [43]. During the D2-test for attention [44], participants were asked to cross out a target letter (“d”) among similar distractors, providing a measure of sustained and selective attention as well as accuracy. For the assessment of verbal fluency, the Regensburger verbal fluency test (RWT) [45] was employed, providing a measure of lexical (“name words starting with p”) and semantic (“name animals”) verbal fluency.

### Identification of formal thought disorder clusters (subtypes) using latent profile analysis

Latent profile analysis is a statistical technique for classification of individuals of a given population into homogenous latent clusters or subtypes based on a certain set of observed categorical (LCA) or continuous (LPA) variables, also referred to as indicators. As a model-based cluster analytic approach, it is a case of finite mixture modelling and provides more flexibility than traditional cluster algorithms, since it is based on an explicit model of the data [46]. It follows the assumption that the data is generated from a mixture of underlying probability distributions, where each component can be interpreted as a cluster that is more homogenous in the distribution of variable means, variances and covariances [47]. Subjects are assigned to these unobserved (latent) classes based on their probability of belonging to each class given the pattern of scores across indicator variables [48].

Here, latent profile analyses were performed using the R package mclust (version 5.4.7) [49], designed for model-based clustering, classification and density estimation based on finite Gaussian mixture modelling using the Expectation Maximization algorithm, as well as the R package tidyLPA (version 1.0.8) [50] which provides an interface to mclust functions for performing latent profile analysis. As indicators we used z-standardized items of the SANS, SAPS, YMRS, and HAM-D psychopathology scales assessing impairment of language and thought. Supplement 1 lists all latent profile indicators.

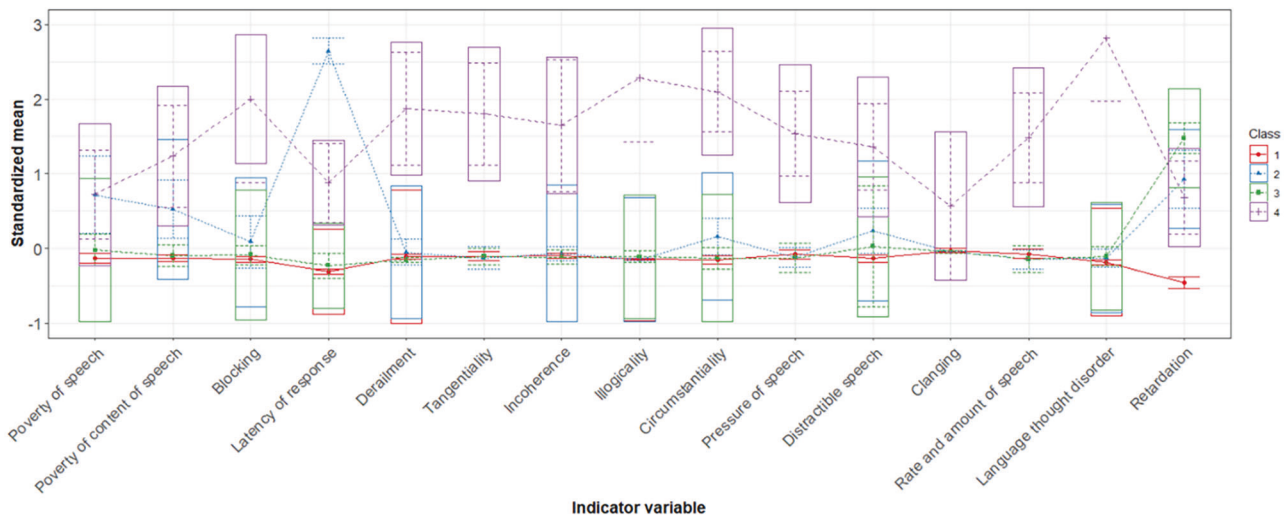
The decision upon the number of latent classes in a model and the evaluation of criteria in this process remains a matter of debate. Therefore,

we used several criteria. These included the Bayesian Information Criterion (BIC) [51], the bootstrapped likelihood ratio test (BLRT, compares the improvement in model fit between a k–1 class model and k class model) [46], entropy (measure of classification uncertainty that should be close to 1) [52], the minimum of the average latent class probabilities for most likely class membership, also reflecting classification certainty [53], and the minimum class size (classes should not contain less than 5% of the sample, although it might be preferable to include smaller classes if they are clearly interpretable). All in all, while a number of fit indices and diagnostic criteria should be considered in order to make an objective decision on the number of classes, theoretical assumptions and interpretability of the model should also be taken into account during the process of model selection [54]. Therefore, we tested models including 2 to 5 classes and compared these models with regard to the above criteria. After final model selection, subjects were assigned to classes based on estimated posterior class membership probabilities and in a next step the resulting classes were entered into subsequent analyses with the aim to validate and characterize FTD clusters regarding clinical features and neurocognitive performance. Therefore, we used analyses of variance (ANOVA) and Tukey post-hoc tests in R to investigate differences between the identified clusters with regard to FTD symptoms, general psychopathology, and neuropsychological performance. In case assumptions for parametric testing were not met, a non-parametric approach was used. Correction for multiple testing was performed using the Benjamini-Hochberg approach [55].

As the DSM-IV-TR diagnostic groups were unequally distributed, we wanted to rule out potential confounding effects of formal diagnosis. Therefore, we re-ran clustering analysis in an age- and sex-matched sample with an equal diagnosis distribution (each  $n = 107$  of MDD, BD, SSD, total  $N = 321$ ) (see supplemental eTable 6). Matching was performed using the “MatchIt” package [56] in R [57]. In addition, FTD cluster analysis was applied in each diagnostic group separately.

### MRI acquisition and preprocessing

Details of MRI data acquisition and all preprocessing steps can be found in supplement 3 and 4. For the present study, we investigated the following surface measures as derived by CAT12: cortical thickness, sulcal depth, cortical complexity, and gyrification.



**Fig. 1 Four-cluster item profiles for FTD symptoms.** Note: Estimated standardized sample means of SANS, SAPS, YMRS and HAM-D items used as indicator variables in the LPA across clusters. Boxes encompass  $\pm 1$ SD. Cluster 1 (*minimal FTD*), cluster 2 (*poverty*), cluster 3 (*inhibition*), cluster 4 (*severe FTD*).

### Voxel and surface based morphometry statistical analyses

Morphometry analyses (GMV, cortical thickness, sulcus depth, gyrification, cortical complexity) were performed using the CAT12 toolbox as implemented in SPM12. Smoothed GM images were entered into a full factorial model (one-way ANOVA) with the groups obtained via the LPA included as one factor with four levels. As covariates of no interest, age, sex, total intracranial brain volume (TIV), and two dummy coded variables accounting for site and a body-coil exchange, as recommended by Vogelbacher et al. [58], were included in the analyses. To account for medication, we added the Medication Load Index (MedIndex) as additional covariate accounting type and amount of current medication adherence [59]. As surface values usually do not depend on TIV, it was not included as a covariate in SBM analyses (see CAT12 manual).

For second level analyses, an absolute threshold masking with a threshold value of 0.1 was applied. In order to detect differences in GMV and surface measures between latent clusters, voxel-wise parametric t-tests were calculated between each group obtained via the LPA. An initial cluster-forming threshold of  $p < 0.0001$  was applied and results were corrected for multiple comparisons controlling family-wise error (FWE) at  $p < 0.05$  cluster-level.

To investigate whether transdiagnostic brain correlates of FTD were driven by DSM-IV-TR diagnostic categories, we performed post-hoc interaction analyses using an ANCOVA design in R with the clinical DSM diagnosis and cluster membership as factors and significant clusters from VBM and SBM as dependent variables. Therefore, eigenvariate values approximating mean GMV and vertex values for surface measures of significant clusters were extracted. Due to the unbalanced distribution of DSM-IV-TR categories, we used whole brain clusters of the total sample and tested them as regions of interest (ROIs) in an equally distributed sample matched for age and sex. Using the batch mode, the search space was restricted for each significant cluster from the total sample beforehand. Same covariates as for whole brain analyses in the total sample were accounted for, using a threshold of  $p < 0.05$ , uncorrected with  $k \geq 5$ .

## RESULTS

### Model comparison and selection

For the LPA, two different statistical models with different covariance structures were estimated. One restricted the variance to be equal across clusters and covariance to be zero, i.e., conditional independence of the indicator variables. For the other model type, covariance was additionally estimated and constrained to be equal across clusters. The latter showed an overall better fit than model 1 (covariance fixed to 0) for  $< 5$  clusters. For  $> 5$  clusters, model 1 showed a notable decrease in BIC and boundary estimates equaling 0 and 1, which is an indicator for convergence at a local instead of a global maximum [60].

The 4-cluster solution had a lower BIC, AIC, and a significant BLRT ( $p < 0.01$ ), relative to the 3-class solution. With including additional clusters (5-cluster solution), AIC and BIC increased and entropy and minimum latent class posterior probability were lower than in the 4-cluster model (see eTable 1a for detailed fit indices). Therefore, the 4-cluster solution is presented here (Fig. 1). Cluster membership was extracted based on the estimated posterior latent class probability indicating most likely class membership for each subject (see eTable 1b and c).

In the 4-cluster solution, cluster 1 (*minimal FTD*) was the largest comprising 729 subjects (70.6%), followed by cluster 3 (*inhibition*) including 164 subjects (15.9%), and cluster 2 (*poverty*) constituted of 80 subjects (7.8%) of the sample. The smallest group was cluster 4 (*severe FTD*) with 59 subjects (5.7%). Supplemental eTable 1 summarizes the cluster sizes and estimated posterior latent class probability for each cluster. Figure 2a, b show the distribution of DSM-IV-TR diagnostic categories and remission status across clusters.

We re-ran our analyses in an age- and sex-matched subsample with equal representation across diagnostic groups and conducted separate analyses within each diagnostic group (MDD, BD, SSD). In all cases, the 4-cluster solution consistently emerged as the best fit, demonstrating the robustness of our transdiagnostic clustering model (see supplemental eTables 6–13 and eFigs. 3–6). Remarkably, 90% of participants in the matched subsample were assigned to the same cluster as in the full sample, underscoring the stability of the clustering solution across different sample compositions. In the matched subsample, cluster distribution closely mirrored that of the full sample: *minimal FTD* remained the largest, including 202 patients (62.9%), followed by *severe FTD* with 48 participants (15.0%), *poverty* with 41 participants (12.8%), and *inhibition* with 30 patients (9.3%). Consistent with our main findings, DSM-IV-TR diagnostic categories were represented across all four clusters in the matched sample.

### Formal thought disorder psychopathological characteristics

Clusters were compared with respect to average ratings across all indicator variables using ANOVA and Tukey post-hoc tests. Clusters differed significantly across all indicator variables used in the LPA (Fig. 1 for a visualization of the estimated latent FTD profiles, eFig. 1 and eTable 2 for unstandardized means and results of post-hoc tests). To summarize, *minimal FTD* showed no or only subclinical FTD symptomatology, whereas *poverty* was characterized by moderate FTD symptoms manifesting as a quantitative

deficit of produced speech with especially increased latency of response but also signs of circumstantiality. Subjects of *inhibition* seemed to exhibit inhibited speech and thought process but at most a mild deficit in the amount of speech produced. *Severe* FTD reflected the most severe symptoms with a tendency towards but not exclusively positive FTD symptoms, manifesting as an overall increase of speech production with disorganized patterns of speech marked by circumstantial and disjointed utterances that convey little substantial information (Fig. 1). FTD psychopathological characteristics of the 4-cluster solution in the age- and sex-matched subsample with same n per diagnosis can be found in supplemental eTable 5.

### Clinical and demographic characteristics

Supplemental eTable 3 shows demographic and clinical characteristics of the identified clusters. Identified clusters differed significantly across all psychopathology rating scales and subscales included. In summary, *minimal* FTD showed less overall negative and positive symptoms as well as depressive symptoms compared to all other clusters. *Poverty* was found to exhibit prominent negative symptoms compared to *minimal* FTD and *inhibition* and additionally exhibited a significantly higher average in hallucinations. Individuals of *inhibition* were found to have highest ratings in anhedonia but showed significantly lower ratings than *poverty* and *severe* FTD in all remaining subscales referring to positive and negative symptoms. Clinical and demographic characteristics of the 4-cluster solution in the age- and sex-matched subsample with same n per diagnosis can be found in supplemental eTable 11.

### Neurocognitive characteristics

Clusters were further compared with respect to performance in neuropsychological domains that have previously been correlated with FTD severity in SZ. Subjects in *minimal* FTD performed significantly better than all other cluster in the domains of phonemic verbal fluency and attention. Clusters 2 and 4 showed significantly lower scores than clusters 1 and 3 in terms of executive function, working memory, semantic verbal fluency, and episodic verbal memory (supplemental eFig. 2). Overall, those clusters that were characterized by *poverty* and *severe* FTD could also be distinguished by poor neurocognitive functioning across all inspected domains (eTable 4). Neurocognitive characteristics of the 4-cluster solution in the age- and sex-matched subsample with same n per diagnosis can be found in supplemental eTable 12.

### Association of formal thought disorder clusters with GMV and sulcal depth

Morphometry analyses revealed significant differences between the identified FTD clusters in GMV and sulcal depth (Figs. 3 and 4). We identified one FTD cluster main effect cluster for GMV ( $k = 844$ ,  $x/y/z = 45/-6/-38$ ,  $p = 0.012$  FWE cluster-level corrected) comprising the right inferior and middle temporal gyri and the right fusiform gyrus. For sulcal depth, we identified a significant cluster comprising the left insula ( $k = 110$ ,  $x/y/z = -37/1/-21$ ,  $p = 0.045$  FWE cluster-level corrected). Post-hoc parametric t-tests revealed GMV and sulcal depth clusters that differed significantly between the *poverty* and *inhibition* FTD cluster and clusters with *minimal* and *severe* FTD, respectively. Investigations of cortical thickness, gyrification, and complexity did not result in significant effects. Table 2 lists all significant effects.

To further confirm the stability of our neuroanatomical findings, we analyzed brain structural differences across FTD clusters within the age- and sex-matched subsample, using the significant clusters identified in the total sample as ROIs. All significant neuroanatomical clusters from the total sample were successfully replicated in the matched subsample (see supplemental eTable 13).

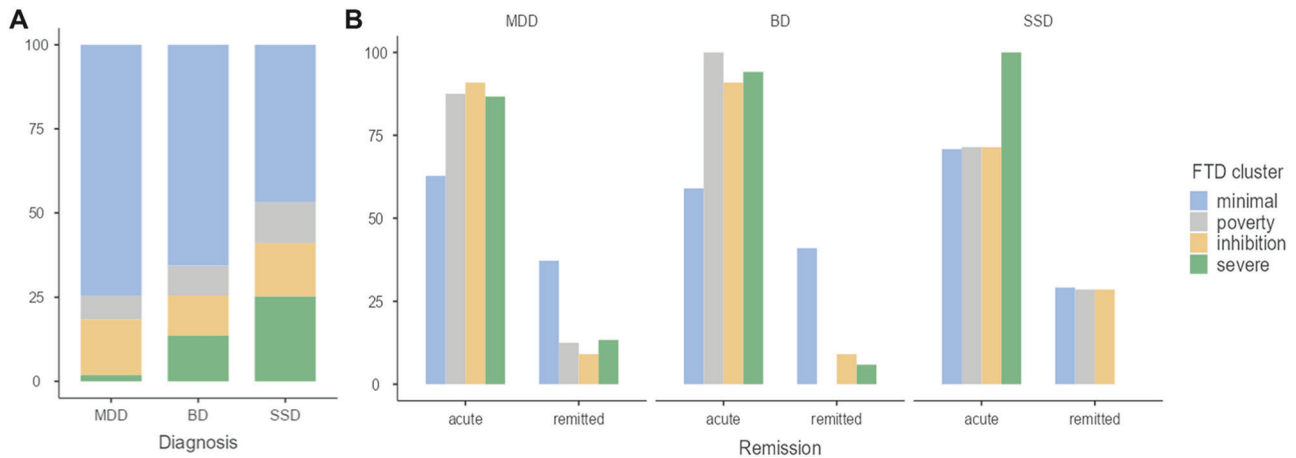
### DISCUSSION

The aim of the present study was to identify latent patient clusters (i.e., psychopathological subtypes) based on FTD psychopathology in a large transdiagnostic sample of patients with MDD, BD, and SSD, and to investigate gray matter as well as surface brain structure associated with the identified FTD clusters. There were four FTD clusters, i.e., *minimal* and *severe* FTD, while the other two manifested as a low amount of produced speech (*poverty*) and inhibited speech and thought process (*inhibition*). Identified FTD clusters differed in neurocognitive domain test performance as well as GMV and sulcal depth patterns. sMRI analyses revealed differences between the *inhibition* cluster compared to both *minimal* and *severe* FTD in the right inferior temporal lobe. *Minimal* and *severe* FTD differed in the GMV of the right superior frontal lobe. Sulcal depth was reduced around the left insula (*inhibition* vs. *poverty*) as well as bilaterally (*inhibition* vs. *severe* FTD). Psychopathological clusters and their neural correlates were present across diagnosis and stable when correcting for current psychotropic medication. Additionally, robustness analyses confirmed that the four-cluster solution remained valid in an age- and sex-matched subsample and within each diagnosis independently. Furthermore, the neuroanatomical findings were replicated in the matched subsample, demonstrating that the observed structural differences across FTD clusters were consistent and not influenced by sample imbalances.

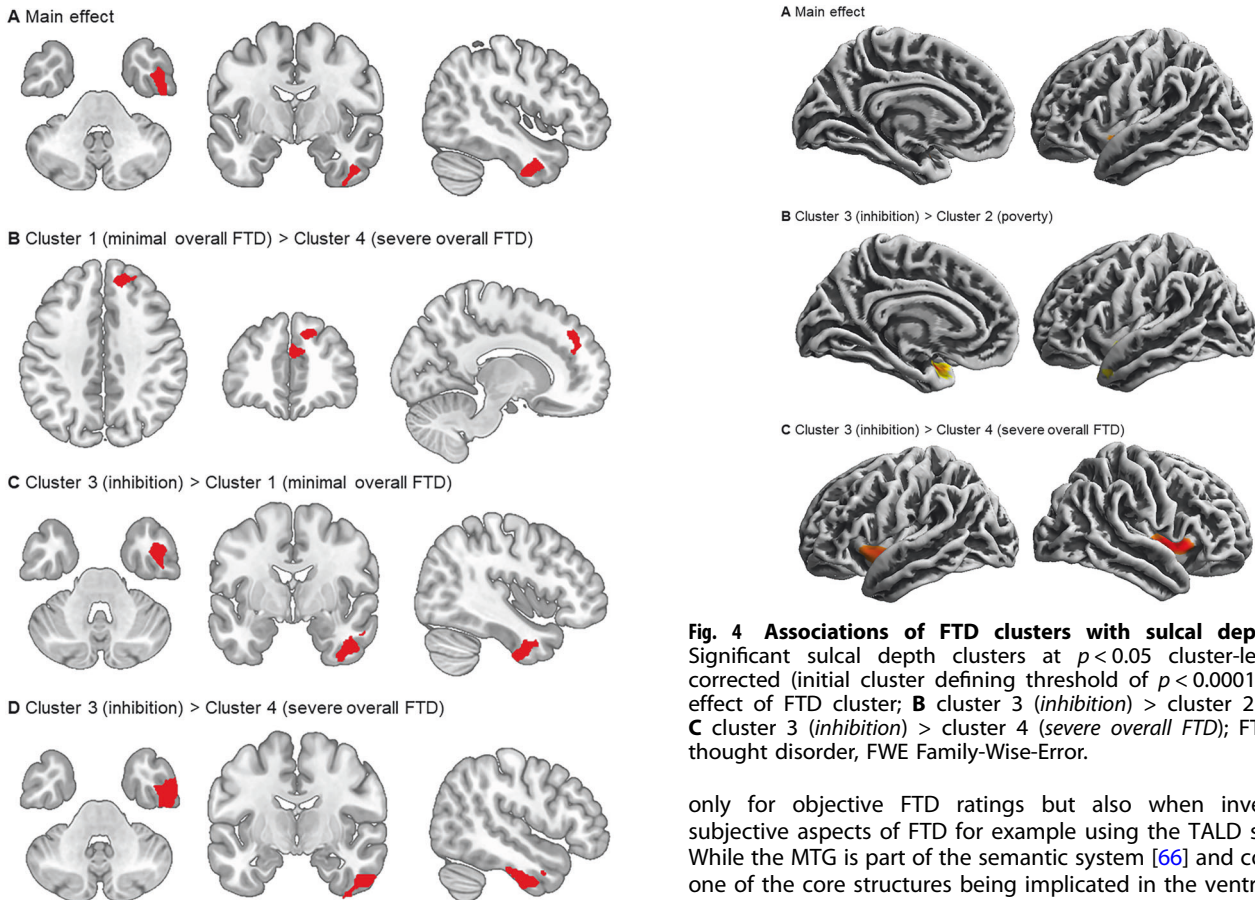
Based on our findings, several new insights emerge. First, identified FTD clusters were not distinguished with regard to positive and negative FTD, but rather differed in their severity of all FTD symptoms while showing a tendency towards one symptom (e.g., poverty of speech). In contrast to factor analytic studies which aggregate symptoms, we could demonstrate that when clustering individual patients, pFTD and nFTD occur concomitantly in the same patient. Interestingly, Oeztuerk et al. [24] also identified low vs. high FTD clusters in recent-onset psychosis. In their study, high FTD was characterized by conceptual disorganization, poverty of speech and poverty of content of speech, and increased response latency, similarly to our results of the *severe* cluster [24]. We further identified one distinctly negative FTD patient cluster who exhibited mainly *poverty*, in line with factor analytic studies in SZ [11] as well as one *inhibition* patient cluster. Importantly, diagnoses were distributed across the identified latent clusters such that every DSM-IV-TR diagnosis was present in each FTD cluster, indicating that FTD is not an exclusive feature of SSD or bipolar mania [2, 7]. *Severe* FTD, collapsing acute and remitted patients, comprised a higher percentage of BD and SSD vs. MDD subjects suggesting that patterns of severe disorganized speech are more prevalent in these disorders [8] which is unsurprising because this is part of the DSM-IV/V diagnostic classification. Yet strikingly, both, the *severe* and the *minimal* clusters are distributed almost evenly across the three diagnostic categories in the acute patients (Fig. 2b). Additionally, results showed that each FTD cluster was present even in remission across diagnoses, underscoring the significance of FTD across various illness stages.

Second, regarding associations of patient clusters with neuropsychological test domain performance, *poverty* and *severe* FTD were characterized by poor performance in lexical and semantic verbal fluency, attention, executive functioning, working memory and verbal episodic working memory. This supports the notion of FTD being related to or being in part a result of impaired neurocognitive functioning (e.g., dysexecutive functioning, impaired semantic access) [3, 61–63], not only in SSD patients but also in affective disorders (i.e., transdiagnostic samples) [64, 65].

Third, as hypothesized, the most prominent gray matter structural aberrations were found in the *severe* FTD symptom cluster: In comparison to the *minimal* FTD patients, they showed reduced GMV in the right superior/medial frontal and to the *inhibition* cluster in the right inferior/middle temporal gyri. Volume reductions in both the temporal and frontal lobes have been highly consistent findings in association with FTD in SZ [3, 26]. Not



**Fig. 2** Distribution of diagnosis and clinical status across identified FTD clusters. **Note:** **A** Relative distribution of diagnoses. Diagnoses were distributed across all 4 clusters with MDD patients constituting the largest proportion of every cluster except for cluster 4 (*severe* FTD), where SSD patients outnumbered all other patient groups, while MDD patients composed the smallest percentage. **B** Acute/remitted patients within latent clusters.



**Fig. 3** Associations of FTD clusters with gray matter volume. **Note:** Significant GMV clusters at  $p < 0.05$  cluster-level FWE-corrected (initial cluster defining threshold of  $p < 0.0001$ ). **A** Main effect of FTD cluster; **B** cluster 1 (*minimal overall FTD*) > cluster 4 (*severe overall FTD*); **C** cluster 3 (*inhibition*) > cluster 1 (*minimal overall FTD*); **D** cluster 3 (*inhibition*) > cluster 4 (*severe overall FTD*); FTD formal thought disorder, FWE Family-Wise-Error.

**Fig. 4** Associations of FTD clusters with sulcal depth. **Note:** Significant sulcal depth clusters at  $p < 0.05$  cluster-level FWE-corrected (initial cluster defining threshold of  $p < 0.0001$ ). **A** Main effect of FTD cluster; **B** cluster 3 (*inhibition*) > cluster 2 (*poverty*); **C** cluster 3 (*inhibition*) > cluster 4 (*severe overall FTD*); FTD formal thought disorder, FWE Family-Wise-Error.

only for objective FTD ratings but also when investigating subjective aspects of FTD for example using the TALD scale [29]. While the MTG is part of the semantic system [66] and constitutes one of the core structures being implicated in the ventral stream of the language network [32] mapping phonological representations onto lexical conceptual representations [67], the right inferior temporal gyrus (ITG) has been correlated with the articulation rate in SZ during speech production tasks [68]. While the present study focused on structural MRI correlates of FTD, functional neuroimaging studies have demonstrated altered activity patterns in almost similar brain regions as reported here [3, 26, 28, 30, 31, 61, 69, 70]. For example, an activation likelihood estimation meta-analysis on FTD highlighted the role of the dorsal

**Table 2.** Differences between identified FTD clusters in GMV and sulcal depth.

	H	x	y	z	t	k	p FWE cluster-level	p inter-action analyses	Partial. $\eta^2$
<b>GMV</b>									
cluster 1 ( <i>minimal overall FTD</i> ) > cluster 4 ( <i>severe overall FTD</i> )									
55.5% superior frontal gyrus 44.5% medial frontal gyrus	R	12	42	38	3.99	800	0.048	0.404	0.022
cluster 3 ( <i>inhibition</i> ) > cluster 1 ( <i>minimal overall FTD</i> )									
50.8% inferior temporal gyrus 19.5% fusiform gyrus 18.2% middle temporal gyrus	R	42	-3	-34	4.93	926	0.029	0.060	0.027
cluster 3 ( <i>inhibition</i> ) > cluster 4 ( <i>severe overall FTD</i> )									
85.6% inferior temporal gyrus 11.5% middle temporal gyrus 2.9% fusiform gyrus	R	48	-9	-36	4.45	2404	<b>&lt;0.0001</b>	0.063	0.020
<b>Sulcal depth</b>									
cluster 3 ( <i>inhibition</i> ) > cluster 2 ( <i>poverty</i> )									
45% insula 37% superior temporal sulcus 16% temporal pole	L	-36	1	-22	3.97	238	<b>0.004</b>	0.145	0.019
cluster 3 ( <i>inhibition</i> ) > cluster 4 ( <i>severe overall FTD</i> )									
100% insula	R	40	5	-2	3.57	287	<b>0.001</b>	0.778	0.009
100% insula	L	-39	2	-7	3.31	188	<b>0.001</b>	0.667	0.005

Note: Interaction analyses refer to the post-hoc investigation of clinical DSM diagnoses on brain structural differences between FTD clusters. Therefore, an ANCOVA design in R with the clinical DSM diagnosis and cluster membership as factors and significant clusters from VBM and SBM as dependent variables were used. R right, L left, H hemisphere; FWE Family-Wise-Error; k cluster extend; bold letters indicate significance after correction for multiple testing (Benjamini Hochberg).

and ventral MTG in semantic processing of language related information [31].

Regarding cortical surface, patients' brains in the *poverty* and *severe* FTD clusters showed sulcal flattening around the left insula, STS, temporal pole, and bilateral insula when compared to *inhibition*. This is in line with previous studies finding the bilateral GMV of the insula [29, 71] and cerebral blood flow in the right insula (along with MTG) [29, 72] to correlate with negative FTD. Sulcal changes have been associated with early neurodevelopmental impacts [73, 74], and in particular insular alterations with negative symptoms and poor psycho-social functioning in SZ patients [75, 76]. We could now extend and specify these findings, as our analyses revealed sulcal flattening in these clusters to be associated with negative FTD symptomatology across MDD, BD, and SSD. More generally our structural findings indicate that particular FTD syndromes may arise from distinct brain structural changes. This aligns well with our previous findings on structural network connectivity being associated with transdiagnostic FTD dimensions [32]. However, while our findings indicate specific differences in brain structure across FTD clusters, the cause remains speculative. Future longitudinal studies could clarify these relationships and help determine directionality.

Finally, differences in brain structure between FTD clusters were present across clinical DSM-IV-TR diagnoses and were independent of current medication. Brain regions that have been found to display reduced volume in SZ, were associated with FTD clusters across diagnoses, indicating that FTD related structural aberrations do not confine to psychotic disorders. This finding aligns with studies indicating a large (biological) overlap across affective and psychotic disorders including shared genetics [77], environmental risk factors [78], brain structural changes [7, 79–84], and psychopathology [23, 85].

#### Future directions

Future research should investigate the stability of FTD clusters over time with the examination of sociodemographic factors as potential modulators. Longitudinal studies within and across

specific diagnostic groups could provide valuable insights into how FTD clusters evolve with illness progression or in response to treatment. While FTD has been closely linked to illness progression in SZ, we suggest that its significance may (similarly) extend to other conditions, such as MDD.

#### Limitations

The present study has some limitations. First, FTD was not assessed with a specific FTD scale but with 15 items devoted to language and thought-related symptoms from general psychopathology scales. Including items from specific FTD scales (such as the TALD scale [6]) in cluster analyses might lead to the identification of somewhat different FTD clusters, although the SAPS/SANS used in our study correlate highly with TLC and TALD. However, we have used widely recognized scales enhancing the generalizability and practical relevance of our findings. Second, since there is a lack of consensus about the criteria by which the correct model should be selected in LPA, the decision on the number of extracted clusters might be to a small extent subjective. While several criteria were evaluated in the current study, model selection was also guided by interpretability of FTD clusters. Third, aberrations in gray matter brain structure between FTD clusters resulted from group comparisons without healthy participants, because of the small variance in FTD present in the latter group [3, 6]. Nevertheless, this approach is most suitable for the exploration of differences between FTD clusters across mental disorders. Fourth, the cross-sectional design limits the ability to draw conclusions about causality or directionality between neuroanatomical differences and FTD symptomatology. Fifth, the majority of patients included to the present study were diagnosed with MDD. Nevertheless, several robustness checks confirmed the transdiagnostic validity of FTD clusters.

#### CONCLUSION

This study offers several strengths and new insights. First, data-driven approaches as employed in the present study provide valuable

insight into the clinical co-occurrence of psychopathologies highlighting their nature as multidimensional constructs. By employing a transdiagnostic perspective, this work demonstrates how advanced clustering methods can reveal the heterogeneity within and across psychiatric disorders and therefore assist the establishment of symptom-specific typologies (i.e., syndromes and sub-syndromes), their neural correlates, and eventually treatments (which are mostly syndrome based in psychiatry). Second, combining data-driven approaches on psychopathology with the examination its of neuroanatomical correlates, our study contributes to understanding the neural basis of complex syndromes like FTD that transcends traditional diagnostic boundaries. Rather than aligning with any single diagnosis, our findings indicate that a network of structural brain changes is more closely associated with specific symptoms or syndromes, reinforcing the importance of a neurobiological transdiagnostic approach in future etiological research. Finally, we extended existing knowledge on the neural correlates of FTD by demonstrating that cortical surface alterations, in addition to volumetric changes, are implicated in thought and language impairment.

#### DATA AVAILABILITY

The raw data collected in this study is not openly accessible to protect participant consent and confidentiality. Nevertheless, the FOR2107 consortium is a highly valuable resource for the global research community and is, in principle, available to qualified scientific researchers affiliated with non-commercial research organizations worldwide. Researchers interested in accessing the data must submit a formal research proposal that clearly defines the specific research questions, methodology, and intended statistical analyses. Applications are reviewed by the study's principal investigators, please contact Professors Tilo Kircher and Udo Dannlowski for further information.

#### CODE AVAILABILITY

See supplementary materials for R code used in the LPA. Other codes and data used for the present analyses as well as generated outputs are available. Please contact the corresponding author (FS) for access.

#### REFERENCES

- Palaniyappan L. Dissecting the neurobiology of linguistic disorganisation and impoverishment in schizophrenia. *Semin Cell Dev Biol.* 2022;129:47–60.
- Yalincetin B, Bora E, Binbay T, Ulas H, Akdede BB, Alptekin K. Formal thought disorder in schizophrenia and bipolar disorder: a systematic review and meta-analysis. *Schizophr Res.* 2017;185:2–8.
- Kircher T, Bröhl H, Meier F, Engelen J. Formal thought disorders: from phenomenology to neurobiology. *Lancet Psychiatry.* 2018;5:515–26.
- Hinzen W, Palaniyappan L. The 'L-factor': language as a transdiagnostic dimension in psychopathology. *Prog Neuropsychopharmacol Biol Psychiatry.* 2024;131:110952.
- Cuesta MJ, Peralta V. Testing the hypothesis that formal thought disorders are severe mood disorders. *Schizophr Bull.* 2011;37:1136–46.
- Kircher T, Stratmann M, Ghazi S, Schales C, Frauenheim M, et al. A rating scale for the assessment of objective and subjective formal thought and language disorder (TALD). *Schizophr Res.* 2014;160:216–21.
- Stein F, Buckenmayer E, Brosch K, Meller T, Schmitt S, Ringwald KG, et al. Dimensions of formal thought disorder and their relation to gray- and white matter brain structure in affective and psychotic disorders. *Schizophr Bull.* 2022;48:902–11.
- Roche E, Creed L, Macmahon D, Brennan D, Clarke M. The epidemiology and associated phenomenology of formal thought disorder: a systematic review. *Schizophr Bull.* 2015;41:951–62.
- Yalınçetin B, Ulaş H, Var L, Binbay T, Akdede BB, Alptekin K. Relation of formal thought disorder to symptomatic remission and social functioning in schizophrenia. *Compr Psychiatry.* 2016;70:98–104.
- Tan EJ, Thomas N, Rossell SL. Speech disturbances and quality of life in schizophrenia: differential impacts on functioning and life satisfaction. *Compr Psychiatry.* 2014;55:693–8.
- Roche E, Lyne JP, O'Donoghue B, Segurado R, Kinsella A, Hannigan A, et al. The factor structure and clinical utility of formal thought disorder in first episode psychosis. *Schizophr Res.* 2015;168:92–98.
- Peralta V, Cuesta MJ. Dimensional structure of psychotic symptoms: an item-level analysis of SAPS and SANS symptoms in psychotic disorders. *Schizophr Res.* 1999;38:13–26.

- Nikzad AH, Cong Y, Berretta S, Hänsel K, Cho S, Pradhan S, et al. Who does what to whom? graph representations of action-predication in speech relate to psychopathological dimensions of psychosis. *Schizophrenia.* 2022;8:1–10.
- Andreasen NC. The Scale for the Assessment of Negative Symptoms (SANS): conceptual and theoretical foundations. *Br J Psychiatry Suppl.* 1989;49–58.
- Andreasen NC. Scale for the assessment of positive symptoms. Iowa City: University of Iowa; 1984.
- Tang SX, Hänsel K, Cong Y, Nikzad AH, Mehta A, Cho S, et al. Latent factors of language disturbance and relationships to quantitative speech features. *Schizophr Bull.* 2023;49:S93–S103.
- Marquand AF, Wolfers T, Mennes M, Buitelaar J, Beckmann CF. Beyond lumping and splitting: a review of computational approaches for stratifying psychiatric disorders. *Biol Psychiatry Cogn Neurosci Neuroimaging.* 2016;1:433–47.
- Iniesta R, Stahl D, McGuffin P. Machine learning, statistical learning and the future of biological research in psychiatry. *Psychol Med.* 2016;46:2455–65.
- Ulbricht CM, Chrysanthopoulou SA, Levin L, Lapane KL. The use of latent class analysis for identifying subtypes of depression: a systematic review. *Psychiatry Res.* 2018;266:228–46.
- Geisler D, Walton E, Naylor M, Roessner V, Lim KO, Charles Schulz S, et al. Brain structure and function correlates of cognitive subtypes in schizophrenia. *Psychiatry Res.* 2015;234:74–83.
- Lee J, Rizzo S, Altschuler L, Glahn DC, Miklowitz DJ, Sugar CA, et al. Deconstructing bipolar disorder and schizophrenia: a cross-diagnostic cluster analysis of cognitive phenotypes. *J Affect Disord.* 2017;209:71–79.
- Dwyer DB, Kalman JL, Budde M, Kambeitz J, Ruef A, Antonucci LA, et al. An investigation of psychosis subgroups with prognostic validation and exploration of genetic underpinnings: the PsyCourse Study. *JAMA Psychiatry.* 2020;77:523–33.
- Pelin H, Ising M, Stein F, Meinert S, Meller T, Brosch K, et al. Identification of transdiagnostic psychiatric disorder subtypes using unsupervised learning. *Neuropsychopharmacology.* 2021;46:1895–905.
- Oeztuerk OF, Pignoni A, Wenzel J, Haas SS, Popovic D, Ruef A, et al. The clinical relevance of formal thought disorder in the early stages of psychosis: results from the PRONIA study. *Eur Arch Psychiatry Clin Neurosci.* 2022;272:403–13.
- Buciuman MO, Oeztuerk OF, Popovic D, Enrico P, Ruef A, Bieler N, et al. Structural and functional brain patterns predict formal thought Disorder's severity and its persistence in recent-onset psychosis: results from the PRONIA Study. *Biol Psychiatry Cogn Neurosci Neuroimaging.* 2023;8:1207–17.
- Cavelti M, Kircher T, Nagels A, Strik W, Homan P. Is formal thought disorder in schizophrenia related to structural and functional aberrations in the language network? A systematic review of neuroimaging findings. *Schizophr Res.* 2018;199:2–16.
- Sumner PJ, Bell IH, Rossell SL. A systematic review of the structural neuroimaging correlates of thought disorder. *Neurosci Biobehav Rev.* 2018;84:299–315.
- Palaniyappan L, Homan P, Alonso-Sanchez MF. Language network dysfunction and formal thought disorder in schizophrenia. *Schizophr Bull.* 2023;49:486–97.
- Maderthaner L, Pavlidou A, Lefebvre S, Nadesalingam N, Chapellier V, von Känel S, et al. Neural correlates of formal thought disorder dimensions in psychosis. *Schizophr Bull.* 2023;49:S104–S114.
- Chen J, Wensing T, Hoffstaedter F, Cieslik EC, Müller VI, Patil KR, et al. Neurobiological substrates of the positive formal thought disorder in schizophrenia revealed by seed connectome-based predictive modeling. *Neuroimage Clin.* 2021;30:102666.
- Wensing T, Cieslik EC, Müller VI, Hoffstaedter F, Eickhoff SB, Nickl-Jockschat T. Neural correlates of formal thought disorder: an activation likelihood estimation meta-analysis. *Hum Brain Mapp.* 2017;38:4946.
- Stein F, Gruber M, Mauritz M, Brosch K, Pfarr J-K, Ringwald KG, et al. Brain structural network connectivity of formal thought disorder dimensions in affective and psychotic disorders. *Biol Psychiatry.* 2024;95:629–38.
- Palaniyappan L, Al-Radaideh A, Gowland PA, Liddle PF. Cortical thickness and formal thought disorder in schizophrenia: an ultra high-field network-based morphometry study. *Prog Neuropsychopharmacol Biol Psychiatry.* 2020;101:109911.
- Dahnke R, Yotter RA, Gaser C. Cortical thickness and central surface estimation. *Neuroimage.* 2013;65:336–48.
- Yotter RA, Nenadic I, Ziegler G, Thompson PM, Gaser C. Local cortical surface complexity maps from spherical harmonic reconstructions. *Neuroimage.* 2011;56:961–73.
- Desai R, Liebenthal E, Possing ET, Waldron E, Binder JR. Volumetric vs. surface-based alignment for localization of auditory cortex activation. *Neuroimage.* 2005;26:1019–29.
- Kircher T, Wöhr M, Nenadic I, Schwarting R, Schrott G, Alferink J, et al. Neurobiology of the major psychoses: a translational perspective on brain structure and function — the FOR2107 consortium. *Eur Arch Psychiatry Clin Neurosci.* 2019;269:949–62.

38. Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: reliability, validity and sensitivity. *Br J Psychiatry*. 1978;133:429–35.
39. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry*. 1960;23:56–62.
40. Helmstaedter C, Lendt M, Lux S. Verbaler lern- und Merkfähigkeitstest. Göttingen: Hogrefe; 2001.
41. Gold JM, Carpenter C, Randolph C, Goldberg TE, Weinberger DR. Auditory working memory and wisconsin card sorting test performance in schizophrenia. *Arch Gen Psychiatry*. 1997;54:159–65.
42. Wechsler D. Wechsler memory scale-revised: manual. San Antonio: Harcourt Brace Jovanovich; 1987.
43. Reitan RM. Trail making test. Manual for administration and scoring. Tuscon, USA: Reitan Neuropsychology Laboratory; 1979.
44. Brickenkamp R. Test D2 Aufmerksamkeitsbelastungstest. Göttingen: Hogrefe; 2002.
45. Aschenbrenner A, Tucha O, Lange K. RWT regensburger wortflüssigkeitstest. Göttingen: Hogrefe; 2000.
46. McLachlan G, Peel D. Finite mixture models. Hoboken, NJ, USA: John Wiley & Sons, Inc; 2000.
47. Magidson J, Vermunt J. Latent class models for clustering: A comparison with K-means. *Canadian J Mark Res. Professional Marketing Research Society*. 2002;20:36–43.
48. Muthén BO, Muthén LK. Integrating person-centered and variable-centered analyses: growth mixture modeling with latent trajectory classes. *Alcohol Clin Exp Res*. 2000;24:882–91.
49. Scrucca L, Fop M, Murphy T, Raftery A. mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *The R Journal*. 2016;8/1, pp. 289–317.
50. Rosenberg JM, Beymer PN, Anderson DJ, van Lissa CJ, Schmidt JA. tidyLPA: an R package to easily carry out latent profile analysis (LPA) using open-source or commercial software. *J Open Source Softw*. 2019;3:978.
51. Schwarz G. Estimating the dimension of a model. *AnnStatist*. 1978;6:461–4.
52. Celeux G, Soromenho G. An entropy criterion for assessing the number of clusters in a mixture model. *Journal of Classification*. 1996;13:195–212.
53. Jung T, Wickrama KAS. An introduction to latent class growth analysis and growth mixture modeling. *Soc Personal Psychol Compass*. 2008;10:302–17.
54. Nyland KL, Asparouhov T, Muthén BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo Simulation Study. *Struct Equ Modeling*. 2007;14:535–69.
55. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol*. 1995;57:289–300.
56. Ho D, Kosuke I, King G, Stuart E. Matchit: nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, 42;1–28. <https://doi.org/10.18637/jss.v042.i08>.
57. R Development Core Team. A language and Environment for Statistical Computing. R Foundation for Statistical Computing, 2008, Vienna.
58. Vogelbacher C, Möbius TWD, Sommer J, Schuster V, Dannowski U, Kircher T, et al. The Marburg-Münster Affective Disorders Cohort Study (MACS): A quality assurance protocol for MR neuroimaging data. *Neuroimage*. 2018;172:450–60.
59. Redlich R, Almeida JR, Grotegerd D, Opel N, Kugel H, Heindel W, et al. Brain morphometric biomarkers distinguishing unipolar and bipolar depression: a voxel-based morphometry-pattern classification approach. *JAMA Psychiatry*. 2014;71:1222–30.
60. Geiser C. *Datenanalyse mit mplus: eine anwendungsorientierte Einführung*. Wiesbaden: VS Verlag für Sozialwissenschaften; 2010. vol. 1
61. Fuentes-Claramonte P, López-Araquistain L, Sarró S, Sans-Sansa B, Ortiz-Gil J, Maristany T, et al. Brain functional correlates of formal thought disorder in schizophrenia: examining the frontal/dysexecutive hypothesis. *Psychol Med*. 2021;51:2446–53.
62. Stirling J, Hellewell J, Blakey A, Deakin W. Thought disorder in schizophrenia is associated with both executive dysfunction and circumscribed impairments in semantic function. *Psychol Med*. 2006;36:475–84.
63. Bora E, Yalincetin B, Akdede BB, Alptekin K. Neurocognitive and linguistic correlates of positive and negative formal thought disorder: a meta-analysis. *Schizophr Res*. 2019;209:2–11.
64. Nagels A, Fahrman P, Stratmann M, Ghazi S, Schales C, Frauenheim M, et al. Distinct neuropsychological correlates in positive and negative formal thought disorder syndromes: the thought and language disorder scale in endogenous psychoses. *Neuropsychobiology*. 2016;73:139–47.
65. Stein F, Schmitt S, Brosch K, Meller T, Pfarr J-K, Ringwald K, et al. State of illness-dependent associations of neuro-cognition and psychopathological syndromes in a large transdiagnostic cohort. *J Affect Disord*. 2023;324:589–99.
66. Binder JR, Desai RH, Graves WW, Conant LL. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb Cortex*. 2009;19:2767–96.
67. Davey J, Thompson HE, Hallam G, Karapanagiotidis T, Murphy C, De Caso I, et al. Exploring the role of the posterior middle temporal gyrus in semantic cognition: Integration of anterior temporal lobe with executive processes. *Neuroimage*. 2016;137:165–77.
68. Kircher TJ, Liddle PF, Brammer MJ, Williams SCR, Murray RM, McGuire PK. Reversed lateralization of temporal activation during speech production in thought disordered patients with schizophrenia. *Psychol Med*. 2002;32:439–49.
69. Sumner PJ, Bell IH, Rossell SL. A systematic review of task-based functional neuroimaging studies investigating language, semantic and executive processes in thought disorder. *Neurosci Biobehav Rev*. 2018;94:59–75.
70. Alonso-Sánchez MF, Hinzen W, He R, Gati J, Palaniyappan L. Perplexity of utterances in untreated first-episode psychosis: an ultra-high field MRI dynamic causal modelling study of the semantic network. *J Psychiatry Neurosci*. 2024;49:E252.
71. Palaniyappan L, Mahmood J, Balain V, Mougín O, Gowland PA, Liddle PF. Structural correlates of formal thought disorder in schizophrenia: an ultra-high field multivariate morphometry study. *Schizophr Res*. 2015;168:305–12.
72. Winkelbeiner S, Cavelti M, Federspiel A, Kunzelmann K, Dierks T, Strik W, et al. Decreased blood flow in the right insula and middle temporal gyrus predicts negative formal thought disorder in schizophrenia. *Schizophr Res*. 2018;201:432–4.
73. Yan J, Cui Y, Li Q, Tian L, Liu B, Jiang T, et al. Cortical thinning and flattening in schizophrenia and their unaffected parents. *Neuropsychiatr Dis Treat*. 2019;15:935.
74. Csernansky JG, Gillespie SK, Dierker DL, Anticevic A, Wang L, Barch DM, et al. Symmetric abnormalities in sulcal patterning in schizophrenia. *Neuroimage*. 2008;43:440–6.
75. Takahashi T, Kido M, Sasabayashi D, Nakamura M, Furuichi A, Takayanagi Y, et al. Gray matter changes in the insular cortex during the course of the schizophrenia spectrum. *Front Psychiatry*. 2020;11:659.
76. Uwatoko T, Yoshizumi M, Miyata J, Ubukata S, Fujiwara H, Kawada R, et al. Insular gray matter volume and objective quality of life in schizophrenia. *PLoS ONE*. 2015;10:e0142018.
77. Anttila V, Bulik-Sullivan B, Finucane HK, Walters RK, Bras J, Duncan L, et al. Analysis of shared heritability in common disorders of the brain. *Science*. 2018;360:eaap8757.
78. Uher R, Zwickler A. Etiology in psychiatry: embracing the reality of poly-gene-environmental causation of mental illness. *World Psychiatry*. 2017;16:121–9.
79. Repple J, Gruber M, Mauritz M, de Lange SC, Winter NR, Opel N, et al. Shared and specific patterns of structural brain connectivity across affective and psychotic disorders. *Biol Psychiatry*. 2023;93:178–86.
80. Stein F, Meller T, Brosch K, Schmitt S, Ringwald K, Pfarr JK, et al. Psychopathological syndromes across affective and psychotic disorders correlate with gray matter volumes. *Schizophr Bull*. 2021;47:1740–50.
81. Goodkind M, Eickhoff SB, Oathes DJ, Jiang Y, Chang A, Jones-Hagata LB, et al. Identification of a common neurobiological substrate for mental illness. *JAMA Psychiatry*. 2015;72:305–15.
82. Winter NR, Leenings R, Ernsting J, Sarink K, Fisch L, Emden D, et al. Quantifying deviations of brain structure and function in major depressive disorder across neuroimaging modalities. *JAMA Psychiatry*. 2022;79:879–88.
83. Brosch K, Stein F, Schmitt S, Pfarr JK, Ringwald KG, Thomas-Odenthal F, et al. Reduced hippocampal gray matter volume is a common feature of patients with major depression, bipolar disorder, and schizophrenia spectrum disorders. *Mol Psychiatry*. 2022;27:4234–4243.
84. McCutcheon RA, Pillinger T, Guo X, Rogdaki M, Welby G, Vano L, et al. Shared and separate patterns in brain morphometry across transdiagnostic dimensions. *Nat Ment Health*. 2023;1:55–65.
85. Stein F, Lemmer G, Schmitt S, Brosch K, Meller T, Fischer E, et al. Factor analyses of multidimensional symptoms in a large group of patients with major depressive disorder, bipolar disorder, schizoaffective disorder and schizophrenia. *Schizophr Res*. 2020;218:38–47.

## ACKNOWLEDGEMENTS

The FOR2107 cohort project was approved by the Ethics Committees of the Medical Faculties, University of Marburg (AZ:07/14) and University of Münster (AZ:2014-422-b-5). We are deeply indebted to all study participants and staff. A list of acknowledgements can be found here: [www.for2107.de/acknowledgements](http://www.for2107.de/acknowledgements).

## AUTHOR CONTRIBUTIONS

FS: Funding Acquisition, Conceptualization, Data Curation, Formal Analysis, Methodology, Project Administration, Visualization, Writing – Original Draft Preparation. AMG: Data Curation, Formal Analysis, Visualization, Writing – Review & Editing. TK: Funding Acquisition, Conceptualization, Investigation, Methodology, Project Administration, Resources, Supervision, Writing – Review & Editing. SM, NA, TH, HJ, AJ, AK,

IN, BS, UD: Funding Acquisition, Project Administration, Investigation, Writing – Review & Editing. KB, LMK, JKP, LT, FTO, US, HW, AW, KF, JG, DG, KT, AW: Data Curation, Investigation, Writing – Review & Editing. All authors contributed to the development of the article and approved the final submitted version.

## FUNDING

This work is part of the German multicentre consortium “Neurobiology of Affective Disorders. A translational perspective on brain structure and function”, funded by the German Research Foundation (DFG, Research Unit FOR2107). Principal investigators are Tilo Kircher (KI 588/14-1, KI 588/14-2, KI 588/22-1), Udo Dannlowski (DA 1151/5-1, DA 1151/5-2, DA 1151/11-1), Igor Nenadić (NE 2254/1-2, NE 2254/2-1, NE 2254/3-1, NE 2254/4-1), Axel Krug (KR 3822/5-1, KR 3822/7-2), Tim Hahn (HA 7070/2-2), Andreas Jansen (JA 1890/7-1, JA 1890/7-2), and Benjamin Straube (STR 1146/18-1). The study was supported by the German Research Foundation (DFG) to Frederike Stein (STE3301/1-1) project number 527712970. This work was in part supported by the DFG SFB/TRR 393 consortium, project number 521379614 and by the DYNAMIC initiative, which is funded by the LOEWE program of the Hessian Ministry of Science and Arts (Grant Number: LOEWE1/16/519/03/09.001(0009)/98). The study was in part supported by grants from UKGM and Forschungscampus Mittelhessen to Igor Nenadić. Open Access funding enabled and organized by Projekt DEAL.

## COMPETING INTERESTS

Biomedical financial interests or potential conflicts of interest: Tilo Kircher received unrestricted educational grants from Servier, Janssen, Recordati, Aristo, Otsuka, and neuraxpharm. All other authors declare no conflict of interest.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41380-025-03009-w>.

**Correspondence** and requests for materials should be addressed to Frederike Stein.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025