Does imaging genetics reveal shared mechanisms behind psychotic symptom profile?

Wenhao Jiang

Follow this and additional works at: https://scholarworks.gsu.edu/psych_diss

Recommended Citation
Jiang, Wenhao, "Does imaging genetics reveal shared mechanisms behind psychotic symptom profile?." Dissertation, Georgia State University, 2021.
doi: https://doi.org/10.57709/23343992

This Dissertation is brought to you for free and open access by the Department of Psychology at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Psychology Dissertations by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.
DOES IMAGING GENETICS REVEAL SHARED MECHANISMS BEHIND PSYCHOTIC SYMPTOM PROFILE?

by

WENHAO JIANG

Under the Direction of Jessica A. Turner, PhD

ABSTRACT

Current diagnoses of schizophrenia (SZ) and bipolar disorder (BD) are classified by phenomenological principles and clinical descriptions. The boundaries of the disorders are merging with accumulating shared genetic and brain mechanisms being uncovered. Imaging genetics is a useful tool to understand the impact of genetic variations on the brain. It also enables capturing the behavioral implication of those genes and associated brain alterations.

This study aimed to reveal the associations among sets of genetic variations, structural brain abnormalities, and clinical symptom profiles shared in schizophrenia and bipolar disorders by imaging genetics and multivariate approaches. First, we mapped the symptom profiles onto brain patterns. Distinct structural brain patterns guided with symptom profiles represented by
positive and negative syndrome scale (PANSS), through parallel independent component analysis (pICA) were extracted. Brain patterns related to positive symptoms, mood, and apathy were discovered in SZ and BD. Second, we investigated the relationships of symptoms and brain patterns regardless of diagnostic categories by projecting each disorder’s structural brain and PANSS patterns into the other disorder group (e.g., projecting patterns from schizophrenia to bipolar and vice versa) to reassess the associations. The projected brain patterns showed associations with broad symptoms rather than the original PANSS patterns. Finally, we explored the potential shared genetic mechanisms behind symptom-brain patterns by investigating the effect of polygenic risk scores (PRS) from the Psychiatric Genomics Consortium (PGC). Both SZ and BD PRS were significantly associated with the positive symptom-related brain patterns in SZ. Higher genetic risks contributed to more severe gray matter concentration (GMC) reductions in the temporal regions of SZ patients, and it may lead to worse positive symptoms. Correspondingly, in the BD, both SZ and BD PRS were significantly associated with the mood symptom-related brain patterns. Higher risks contributed to more severe gray matter concentration (GMC) reductions in the frontal-temporal-parietal circuits with worse mood symptoms. The polygenic effects behind the apathy component may be subtle. The results helped improve the understanding of categories of psychotic disorders starting from schizophrenia and bipolar disorder. It may essentially contribute to the more precise diagnosis and treatment for heterogeneous populations with psychosis.

INDEX WORDS: Schizophrenia, Bipolar disorder, Diagnostic categories, Brain Structure, Positive and Negative Symptom Scale, Imaging genetics, Multivariate Analysis
DOES IMAGING GENETICS REVEAL SHARED MECHANISMS BEHIND PSYCHOTIC SYMPTOM PROFILE?

by

WENHAO JIANG

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2021
DOES IMAGING GENETICS REVEAL SHARED MECHANISMS BEHIND PSYCHOTIC
SYMPTOM PROFILE?

by

WENHAO JIANG

Committee Chair: Jessica A. Turner
Committee: Jingyu Liu
Tricia King
Vince D. Calhoun

Electronic Version Approved:

Office of Graduate Services
College of Arts and Sciences
Georgia State University
August 2021
DEDICATION

This dissertation is dedicated to my dearest wife, parents, and friends. I can’t imagine to work through these years without your warm and strong support oversea or at my side!
ACKNOWLEDGEMENTS

I want to say thank you to all who have guided, helped, and supported me through the whole course of graduate school. Dr. Turner, I always feel that choosing to start another doctoral training in the field of psychology/neuroscience with you is the best decision I’ve made so far. Jess, you encourage me to stay on the psychotic disorders diagnostic boundaries research and provide all necessary resources to push it forward. For that, I will always be grateful. Thank you to my committee members, Dr. Liu, Dr. King, and Dr. Calhoun. Your thoughtful insight and guidance have made my dissertation much stronger. I also want to thank everyone who has contributed to the data used in this dissertation (B-SNIP, COBRE, FBIRN3, HUBIN, MCIC, and TOP). Without your efforts and kindness in advancing open science and data sharing, any further analysis and scientific significance would be impossible to reach. Thank you to the whole Turner lab. Kelly and Jesse, thank you for your help in managing the data against all odds. Thank you to our collaborators at the Center for Translational Research in Neuroimaging and Data Science for method and analytic tool supports. Finally, I want to thank Georgia State University’s Brains and Behavior programs for providing the funding to conduct my graduate research.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS .................................................................................................................. V

LIST OF TABLES ............................................................................................................................... IX

LIST OF FIGURES ............................................................................................................................... X

1 INTRODUCTION AND LITERATURE REVIEW ........................................................................... 1

1.1 Purpose of the Study ...................................................................................................................... 1

1.2 Genetic overlap between Schizophrenia and Bipolar Disorder ................................................. 3

1.3 Structural Brains alterations in Schizophrenia and Bipolar Disorder ......................................... 4

  1.3.1 Structural Brain Alteration in SZ ............................................................................................... 4

  1.3.2 Structural Brain Alteration in BD ............................................................................................. 5

  1.3.3 Potential Brain Alterations Shared by SZ and BD .................................................................. 6

1.4 Imaging Genetics of Psychotic Symptoms and Cognition ........................................................... 7

1.5 Multivariate Analysis and Data Projection ..................................................................................... 10

2 AIM 1: MAP STRUCTURAL ALTERATIONS ONTO PSYCHOTIC SYMPTOM PROFILE(S) IN SZ AND BD .................................................................................................................. 13

2.1 Aim and hypothesis ......................................................................................................................... 13

2.2 Methods ......................................................................................................................................... 13

  2.2.1 Participants ................................................................................................................................. 13

  2.2.2 Positive and negative symptom scale ....................................................................................... 16

  2.2.3 Preprocessing of imaging data ................................................................................................. 17
2.2.4 Parallel Independent Component Analysis .................................................. 17

2.2.5 Statistical analyses .................................................................................... 18

2.3 Results ........................................................................................................... 20

2.3.1 SZ group results ......................................................................................... 20

2.3.2 BD group results ......................................................................................... 26

3 AIM 2: PROJECTION OF EXTRACTED BRAIN PATTERNS INTO EITHER
DIAGNOSTIC GROUP ......................................................................................... 32

3.1 Aim and hypothesis ....................................................................................... 32

3.2 Methods ......................................................................................................... 32

3.2.1 Participants ................................................................................................ 32

3.2.2 Data Projection ........................................................................................... 32

3.2.3 Statistical Analyses ..................................................................................... 33

3.3 Results ........................................................................................................... 33

4 AIM 3: GENETIC MECHANISM BEHIND IDENTIFIED BRAIN PATTERNS .. 42

4.1 Aim and hypotheses ....................................................................................... 42

4.2 Methods ......................................................................................................... 42

4.2.1 Participants ................................................................................................ 42

4.2.2 Preprocessing of Genetic Data .................................................................... 43

4.2.3 Polygenic Risk Score Calculation ............................................................... 43

4.2.4 Statistical Analyses ..................................................................................... 44
4.3 Results ................................................................................................................................................. 45

5 DISCUSSION AND CONCLUSIONS ................................................................................................. 50

5.1 Summary ............................................................................................................................................... 50

5.2 Symptom profiles reveal distinct brain patterns in SZ and BD ............................... 52

5.3 The relationship of projected brain patterns and symptoms across disorders .... 59

5.4 Genetic effects on symptom profiles driven brain patterns across disorders...... 62

5.5 Limitations ........................................................................................................................................... 65

5.6 Conclusions and future directions ................................................................................................. 68

REFERENCES ........................................................................................................................................... 71
LIST OF TABLES

Table 1 Scanning and site information for each dataset.................................................. 14
Table 2 Demographic and scanning information of SZ participants .................................. 15
Table 3 Demographic and scanning information of BD participants.................................. 16
Table 4 Significant MANCOVA results of study site differences on PANSS items.............. 21
Table 5 PANSS items contributing most to the first PANSS component with |Z| > 1 ........... 22
Table 6 Brain regions, volumes and peak coordinates in SZ group.................................... 24
Table 7 Brain regions, volumes and peak coordinates in BD group.................................... 29
Table 8 Projected brain component loadings’ correlation with BD PANSS item scores ....... 38
Table 9 Projected brain component (mood) loadings’ correlation with SZ PANSS item scores 40
Table 10 Demographic information for polygenic risk in combined SZ sample............... 42
Table 11 Demographic information for polygenic risk in B-SNIP BD sample..................... 43
LIST OF FIGURES

Figure 1 The research strategy of imaging genetics .......................................................... 2

Figure 2 Correlation between loading coefficients of the top PANSS component and the top structural MRI component. .......................................................................................... 23

Figure 3 Parallel ICA results in SZ group......................................................................... 25

Figure 4 Parallel ICA results in BD group – ‘mood’ component ...................................... 28

Figure 5 Parallel ICA results in BD group – ‘apathy’ component.................................... 28

Figure 6 PANSS Item scores in SZ vs. BD...................................................................... 35

Figure 7 PANSS Item variations in SZ vs. BD................................................................. 36

Figure 8 Projecting brain and PANSS components from SZ to BD................................. 37

Figure 9 Projecting brain and PANSS components (mood) from BD to SZ.................... 39

Figure 10 Projecting brain and PANSS components (apathy) from BD to SZ.................. 41

Figure 11 The SZ polygenic risk and correlation results in SZ group............................... 45

Figure 12 The BD polygenic risk and correlation results in SZ group.............................. 46

Figure 13 The polygenic risk results in BD group............................................................ 47

Figure 14 The correlation results of PRS and projected mood component in BD group. .... 48

Figure 15 Comparing original apathy component and similar brain component generated from ICA in B-SNIP BD brains.................................................................................. 49

Figure 16 The correlation between PRS and apathy-like brain components...................... 49
1 INTRODUCTION AND LITERATURE REVIEW

1.1 Purpose of the Study

The current categories of psychiatric disorders are mainly based on phenomenological observation and clinical descriptions (Feighner et al., 1972; Jiang, King, & Turner, 2019; Lawrie, O'Donovan, Saks, Burns, & Lieberman, 2016b). Though these descriptions are reliable, they are not established on precise pathological bases, thus facing difficulties in clinical heterogeneities (Feighner et al., 1972; Robins & Guze, 1970). Additionally, similar psychotic symptoms aggregate in different disorders and families. Underlying the aggregation, shared biological mechanisms, including genetics and neural circuitry, are found (Bipolar Disorder Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2018; Cross-Disorder Group of the Psychiatric Genomics et al., 2013; Noordermeer et al., 2017; van Haren, Bakker, & Kahn, 2008). These findings suggest that the boundaries of psychiatric disorders are merging (Lawrie, O'Donovan, Saks, Burns, & Lieberman, 2016a; Pearlson, Clementz, Sweeney, Keshavan, & Tamminga, 2016).

In current research attempts, schizophrenia (SZ) and bipolar disorder (BD) are often studied together to elucidate this cross diagnostic issue due to their clinical similarities, high genetic correlation, and partially shared brain structural deficits (Andreassen et al., 2013; Bipolar Disorder Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2018; Stahl et al., 2019) SZ patients generally show smaller brain volumes, overall gray matter reductions in fronto-temporal, thalamo-cortical, subcortical-limbic circuits, and enlargement of ventricles (partially shared by BD) (Palaniyappan et al., 2015). These brain alterations induced in part by genetic variations (van Haren et al., 2008) are bridging the gap between gene and the intermediate phenotype (Arslan, 2015; Gottesman & Gould, 2003), and even clinical symptoms
of disorders (Nenadic, Gaser, & Sauer, 2012) (see Figure 1). With the above so-called *imaging genetics* method (Hashimoto et al., 2015; Meyer-Lindenberg & Weinberger, 2006; Thompson, Ge, Glahn, Jahanshad, & Nichols, 2013; Turner et al., 2006), some exploratory studies have been done to subdivide SZ and BD into different diagnostic groups regardless of their original diagnosis (Clementz et al., 2016; Ivleva et al., 2016; Meda et al., 2015; Tandon et al., 2016). However, further investigation to link these overlapped imaging and genetic variations with the core symptom profiles is needed before moving forward into clinical practice (Tamminga et al., 2013). To answer how exactly do potential imaging and genetic findings explain the clinical symptom profiles and ensure stability, we will review related literature and try to discuss possible methods in the following parts. The followings were partially published in author’s previous review article (Jiang et al., 2019).

![Figure 1](image.png)

**Figure 1** The research strategy of imaging genetics (Jiang et al., 2019)
1.2 Genetic overlap between Schizophrenia and Bipolar Disorder

Besides the clinical observation of similarity in both disorders' symptoms and prognosis, growing evidence from genetic and imaging studies confirms this overlap (Cross-Disorder Group of the Psychiatric Genomics, 2013). Genes contribute greatly to the etiology of SZ and BD. The heritability of SZ was around 80%, while in BD, it is estimated to be over 70% through meta-analysis twin studies (Lichtenstein et al., 2009; Sullivan, Kendler, & Neale, 2003). Additionally, an overall genetic correlation of 0.68 between SZ and BD has been demonstrated based on common variants (Cross-Disorder Group of the Psychiatric Genomics, 2013).

The Val158Met single polymorphism (SNP) of catechol-O-methyltransferase (COMT), the Val66Met SNP of brain-derived neurotrophic factor (BDNF), and the Ser704Cys SNP of disrupted-in-SZ 1 (DISC1) are the most well-known risk genes identified by candidate gene analysis (Hashimoto et al., 2015). In contrast, the first few reports of genome-wide association study (GWAS) demonstrated several loci associated with SZ, including Zinc finger protein 804A (ZNF804A), neurogranin (NRGN), and the major histocompatibility complex (MHC) region. GWAS studies with increased sample size discovered more SZ related loci, and some of these loci were shared by BD and other psychiatric disorders.(O'Donovan et al., 2008; Schwab & Wildenauer, 2013). ZNF804A is the first discovered marker that may increase risk for both SZ and BD, and meta-analysis has also supported its role (Williams et al., 2011). The combined SZ and BD GWAS study from the Psychiatric Genomics Consortium (PGC) has identified Calcium Voltage-Gated Channel Subunit Alpha1 C (CACNA1C), Ankyrin-3(ANK3), and Inter-Alpha-Trypsin Inhibitor Heavy Chain 3-4 (ITIH3-ITIH4) as a risk for both disorders (Schizophrenia Psychiatric Genome-Wide Association Study, 2011; Sullivan, Daly, & O'Donovan, 2012). Later, introducing a pleiotropy-informed conditional false discovery rate, 14 loci were associated with
both disorders, and CACNA1C and ITIH4 were identified again (Andreassen et al., 2013). The PGC's diagnostic specificity of 5 disorders analyses has also shown 5'-nucleotidase, cytosolic II (NT5C2), and coiled-coil domain containing 68 (CCDC68) are associated with both disorders (Cross-Disorder Group of the Psychiatric Genomics, 2013). The combined GWAS studies will continue to reveal more critical loci, but the functional implication and role of these distinct genes in SZ and BD will need further investigation.

In addition to candidate gene analysis, polygenetic methods combine a large number of loci, allowing the overlap between disorders to be broadly assessed (Cross-Disorder Group of the Psychiatric Genomics, 2013) As mentioned above, a genetic correlation around 0.68 blurs the distinction across categories again and indicates a broad genetic mechanism for SZ and BD.

1.3 Structural Brains alterations in Schizophrenia and Bipolar Disorder

1.3.1 Structural Brain Alteration in SZ

Anatomical changes in fronto-temporal, thalamo-cortical, subcortical-limbic circuits, enlargement of ventricles, and widespread white matter fibers abnormalities are found in structural brain studies in SZ; still, heterogeneity in brain abnormality patterns is shown across different studies (Dazzan et al., 2015). With the growing sample size and collaboration through different sites, many large meta-analyses have provided new information. The ENIGMA SZ working group’s meta-analysis of subcortical regions across several thousand subjects reported the consistent findings of the smaller hippocampus, amygdala, thalamus, accumbens, and intracranial volumes, but larger pallidum and lateral ventricles (van Erp et al., 2016). The putamen and caudate volume effects were not reliable across different populations and studies even with this sample size, indicating the possibility of clinical heterogeneity affecting those regions (van Erp et al., 2016). In addition, widespread cortical thinning and smaller surface area
with the strongest effect in frontal and temporal regions were identified in SZ again in an ENIGMA SZ study (van Erp et al., 2018).

Functional imaging studies yield comparable results with structural studies. The reduced activation in the prefrontal cortex, superior temporal gyrus, thalamus, frontal lobe, and parietal lobe has been reported in either resting state or task fMRI (Cao, Dixson, Meyer-Lindenberg, & Tost, 2016). Besides all these regions of hypoactivity, hyperactivity has also been reported in the hippocampus and putamen (Cao et al., 2016). Besides, functional connectivity, network analysis, and graph theory application to functional imaging data form an even larger neuroanatomical map and will only be discussed in the following sections when studied using an imaging genetic method and with relatively consistent functional implications.

1.3.2 Structural Brain Alteration in BD

The anatomical changes in BD are milder and less consistent than SZ (Hibar et al., 2018; Hibar et al., 2016; van Erp et al., 2016; van Erp et al., 2018). Various meta-analyses of structural MRI studies have been carried out in BD and healthy controls. The only consistent findings based on previous literature were the enlargement of the right lateral ventricular (Kempton, Geddes, Ettinger, Williams, & Grasby, 2008; McDonald et al., 2004). Other studies may indicate dorsolateral prefrontal cortex (Dickstein et al., 2005), ventrolateral prefrontal cortex (Blumberg et al., 2006), superior temporal gyrus (Elvsashagen et al., 2013; Hegarty et al., 2012; Lyoo et al., 2006), the anterior cingulate (Fornito et al., 2008; Fornito et al., 2009) and subcortical region including amygdala with the thalamus (McDonald et al., 2004) are potentially affected in BD and controls. The ENIGMA BD working group identified consistent and robust grey matter deficits in subcortical regions, including the hippocampus, thalamus, and enlarged lateral ventricles (Hibar et al., 2016). In cortical regions, cortical thinning was found in frontal, temporal, and
parietal regions e.g. pars opercularis, left fusiform gyrus and left rostral middle frontal cortex (Hibar et al., 2018). The functional findings of BD can be less consistent than the structural findings, though they reflect similar regions. The most commonly detected hypo/hyper-activation in prefrontal-frontal-limbic circuits indicated disrupted emotional regulation in BD patients (Chen, Suckling, Lennox, Ooi, & Bullmore, 2011; Strakowski, Delbello, & Adler, 2005; Townsend & Altshuler, 2012).

1.3.3 Potential Brain Alterations Shared by SZ and BD

Many of the above regions have been identified as structural or functional commonalities among multiple diagnostic categories, including SZ and BD (Isobe et al., 2016; Schwarz, Tost, & Meyer-Lindenberg, 2016). SZ and BD both have deficits in the frontal-temporal-parietal circuit (Hibar et al., 2018; van Erp et al., 2018), prefrontal-striatal-thalamic-cerebellar circuit (Soh et al., 2015), default mode network (Soh et al., 2015; Wang et al., 2015), insula (Maggioni, Bellani, Altamura, & Brambilla, 2016; Selvaraj et al., 2012), anterior cingulate cortex (Arnone et al., 2009; Ellison-Wright & Bullmore, 2010), and other numerous brain regions (Yu et al., 2010).

In addition, starting from the same point as genetics, imaging research efforts are trying to redraw the boundaries between psychotic disorders with shared imaging features. From Bipolar-SZ Network on Intermediate Phenotypes (B-SNIP) Consortium, Clementz et al. applied neurobiological measures among SZ, BD, and SAD their first-degree relatives, and unrelated controls, and tried to regroup them into “biotypes” rather than DSM catalogs (Clementz et al., 2016), using a wide selection of psychotic biomarkers and functional brain activity measures. Clementz et al. then identified three “biotypes” which were also found to be more heritable than their original DSM diagnoses. Sensorimotor reactivity and cognitive control performance distinguish three biotypes: Biotype 1 patients showed serious impairment across sensorimotor
reactivity and cognitive control. Biotype 2 patients show only cognitive control deficits, and biotype 3 patients seem to be the mildest in cognitive symptoms. The B-SNIP group has also been trying to find other factors contributing to its biotyping; one attempt is to use the low-frequency fluctuations (ALFF/fALFF) across the SZ, BD, and SAD from a large family study (Meda et al., 2015). More recently, gray matter density was checked in these three biotypes, and the density loss followed the same order as cognitive decline: biotype 1 showed whole-brain gray matter density loss, while type 2 showed largely overlapping results with type 1, and the largest effects were found in fronto-temporal circuits, parietal cortex, and cerebellum. The findings were much more localized and of less magnitude for type 1. Type 3 only showed small reductions in frontal, cingulate, and temporal regions despite their similar DSM diagnoses (Ivleva et al., 2016).

In a larger perspective, both highlighted by ENIGMA and B-SNIP studies, the most robust and stable GM alterations shared between SZ and BD may spread throughout the whole brain with a strong emphasis on the frontal-temporal-parietal circuits. In addition, this circuit plays an important role in shared anatomical bases for potential biotypes between SZ and BD. These findings help mark the targets for the phenotypic and genetic implications behind common brain alterations in SZ and BD.

1.4 Imaging Genetics of Psychotic Symptoms and Cognition

The imaging genetic studies incorporating a complete pathway from genes and structural brain alterations to clinical symptoms are very limited (Hashimoto et al., 2015; Sagar & Pattanayak, 2017; Shahab et al., 2018; Strik, Stegmayer, Walther, & Dierks, 2017). Researchers have been trying to map psychotic symptoms of disorders like SZ and BD onto brain structure and function for decades (Strik et al., 2017). It is generally understood that symptoms including thought disorders (Cavelti, Kircher, Nagels, Strik, & Homan, 2018; Kaplan et al., 1993; Liddle et
al., 1992; Sabri et al., 1997) and auditory verbal hallucinations (Hare et al., 2017; Hare et al., 2018; Kompus, Westerhausen, & Hugdahl, 2011; Stegmayer et al., 2017) can be mapped to semantic regions, affective and paranoid syndromes to limbic regions (Garety & Freeman, 2013; Kapur, 2003; Lebow & Chen, 2016; Ramirez, Moscarello, LeDoux, & Sears, 2015) and motor symptoms to the executive domain (Kendler, 2016; Peralta & Cuesta, 2017; van Harten, Walther, Kent, Sponheim, & Mittal, 2017). However, these results can be quite broad, and lack replications (Insel et al., 2010; Strik et al., 2017).

Some imaging genetic studies in cognition may add to our knowledge based on the aggregate pathway of single cognitive domains and shared risk genes between SZ and BD. Working memory deficit is a fundamental and critical characteristic in SZ and BD. The most well studied intermediate imaging phenotype was the connection abnormalities between dorsolateral prefrontal cortex and hippocampus. Among the shared risk genes, ZNF804A (Esslinger et al., 2009; Paulus et al., 2013; Rasetti et al., 2011) and CACNA1C (Paulus et al., 2014) have been associated with the prefrontal cortex and hippocampus connection alteration. In healthy subjects, the risk allele of ZNF804A was associated with increased connectivity. COMT (Tan et al., 2012), Regulator of G protein signaling 4 (RGS4) (Buckholtz et al., 2007), and COMT X glutamate metabotropic receptor 3 (GRM3) epistasis are connected to prefrontal cortex-parietal coupling in healthy subjects. For episodic and long-term memory, which is also often disturbed in SZ and BD, possible intermediate imaging phenotypes included decreased coupling of the hippocampus-parietal cortex, hippocampus, ventrolateral prefrontal cortex, and bilateral hippocampus (Bearden, Woogen, & Glahn, 2010; Guo, Ragland, & Carter, 2019; Shinozaki & Potash, 2014). However, the genetic association within this thread is elusive (Cao et
al., 2016). CACNA1C (Erk et al., 2010) and NRGN (Krug et al., 2013) might be involved in these memory processes as well.

There are several other cognitive functions often impaired both in SZ and BD. Essential risk genes including NOS1, CACNA1C, and ZNF804A affected attentions deficits through alterations in the prefrontal cortex and inferior frontal gyrus (Zhang et al., 2015), inferior parietal lobule and medial frontal gyrus (Thimm et al., 2011), anterior cingulate cortex (Thurin et al., 2013), and bilateral hippocampus (Bigos et al., 2010). In emotion-processing, ZNF804A (Esslinger et al., 2009), DRD2 (Blasi et al., 2009), COMT (Drabant et al., 2006), and MIR137 (Mothersill et al., 2014) have been proved for their roles in affecting the amygdala, anterior cingulated, and prefrontal cortex which was directly connected the various stages of emotion processing. The reduced theory of mind capabilities tends to hinder SZ and BD patients from understanding themselves and others’ mental states. ZNF804A risk alleles along with alterations in the prefrontal cortex, parietal cortex, and temporal-parietal junction were found critical in such social information processes (Mohnke et al., 2014; Walter et al., 2011).

In summary, current imaging genetic studies with behavioral implications among SZ and BD are mostly limited to functional brain measures (Jiang et al., 2019). In addition, most of these trials were not directly carried out in patients with SZ and BD (Cao et al., 2016). Inevitably, besides cognitive impairment, clinical symptoms are even more critical for the diagnostic issue and future work of disorder boundaries. Finally, candidate gene analyses reveal relatively small effect sizes and explain limited brain alterations (Hibar et al., 2015). In contrast, this study focuses on extracting structural brain deficits associated with clinical symptom profiles to overcome such limitations. The common allele suggested by SZ and BD GWAS study will be collected to weight the polygenic risk on the extracted brain deficits in either diagnostic group.
Thus, insight can be provided for possible shared imaging genetic mechanisms behind SZ/BD symptoms.

### 1.5 Multivariate Analysis and Data Projection

As discussed above, though candidate gene analyses reveal some genetic effects on brain alterations, the effect size and consistency across different studies are limited. Combining genetic risk as a polygenic risk score may help resolve this issue (Bogdan et al., 2017; Euesden, Lewis, & O'Reilly, 2015; Hashimoto et al., 2015). However, imaging phenotypes can be another essential source resulting in such inconsistency (Winkler et al., 2010). The genetic loci or SNPs may have potential effects on multiple brain regions rather than a single region. Previously identified anatomical atlas and the region of interest studies based on a detailed atlas may fail to capture the complete genetic influence (Desikan et al., 2006). In addition, brain regions themselves are not independent, and they are likely to demonstrate some covariation (Jiang et al., 2019; Mennigen et al., 2019).

Multivariate analysis such as independent component analysis (ICA) or source-based morphometry (SBM) is useful in constructing such brain phenotypes considering their covariance (Xu, Groth, Pearlson, Schretlen, & Calhoun, 2009). The multivariate analysis produces distinct spatial brain structural circuits that covary in the data. All brain voxels and the relationship among them are taken into consideration simultaneously. Multivariate analysis has proved its usefulness and reproducibility in the decomposing structural brain (Gupta et al., 2015) and its association with heritability (Sprooten et al., 2015; Turner et al., 2012) in psychiatric disorders.

Similarly, the mapping of psychiatric symptomatology onto specific brain structural alterations has yielded inconsistent findings. Most previous studies have calculated associations
between symptoms and structural brain measures using predefined brain regions based on atlas parcellations or voxels, and total or factor scores from assessment scales on the other hand. Even though these approaches are informative, they might lack sensitivity to detect alterations associated with specific symptoms or clusters of symptoms. In one step further than SBM/ICA, one approach to overcome these shortcomings is to apply parallel independent component analysis (pICA), which identifies patterns of alterations within two or more modalities simultaneously (Calhoun & Sui, 2016; Liu et al., 2009; Pearlson, Liu, & Calhoun, 2015). The extension to ICA/SBM algorithm, pICA, performs individual ICA on multiple data modalities, ‘in parallel’, maximizing independence within each modality while optimizing the correlation of independent components between the data modalities. As a result of this type of analysis, correlation coefficients between the two modalities' independent components are estimated. Individuals' loading coefficients for each independent component can then be used to interrogate group differences further. A recent study employing pICA on functional connectivity measures of the default mode network (DMN) and genetic data found that genes regulating neurodevelopmental processes are associated with functional hypoconnectivity of the DMN in patients with schizophrenia (Meda et al., 2015). Applying pICA will provide meaningful insights into the etiology of psychiatric disorders, encouraging further multimodal investigations.

In addition to multivariate analysis in a single data, projection of multivariate findings can again enhance the stability and contribute to the cross disorder issue. In new data (e.g. another disorder), projection helps achieve the loading coefficients corresponding to the original components such as brain patterns, gene/SNP sets, and clinical profiles. It makes it possible to reassess the association between projected patterns in the new data (Chen et al., 2019). In this way, the mapping of symptomatology onto the brain in SZ can be projected into BD patients, and
it will help to reveal whether association remains valid between psychiatric symptom profile and brain alterations despite the diagnostic categories.

In summary, multivariate analysis helps to construct the symptom profiles, structural brain patterns, and genetic risk covarying through the data. Thus, it will reduce the bias of selecting behavioral and brain phenotypes and maximize the modalities' potential. The projection of these multivariate findings helps validate discovered associations across different disorders with the potential shared mechanism. Also, polygenic methods rather than single variation enhance the capturing of genetic effects in a more realistic background. Therefore, to answer how exactly do potential imaging and genetic findings explain the clinical symptom profiles and ensure stability, we extracted the brain structural effects behind clinical profiles from both SZ and BD through pICA. Additionally, we examined whether these brain alterations were shared by SZ and BD or specific to the original diagnostic groups. Finally, we explored the polygenic factors that modify brain patterns. Thus, we presented potential pathways from genes to brain to clinical symptom profiles behind SZ and BD.
2 AIM 1: MAP STRUCTURAL ALTERATIONS ONTO PSYCHOTIC SYMPTOM PROFILE(S) IN SZ AND BD

2.1 Aim and hypothesis

In this study, the first aim was to map structural alterations of brains onto the psychotic symptom profile(s) in SZ and BD. The symptom profiles were represented and described by the Positive and Negative Symptom Scale (PANSS). The correlations between loading coefficients of brain and PANSS components extracted by pICA were tested, and the patterns of brain and profiles of symptoms were examined. We hypothesized that, given the different symptom patterns in the different disorders, the symptoms profiles represented by the PANSS will extract distinct brain structural components in both SZ and BD. These analyses have been published in part in author’s previous publications (Jiang et al., 2020; Mennigen et al., 2019).

2.2 Methods

2.2.1 Participants

For the whole project, data were pooled from seven major studies: TOP (Thematic Organized Psychosis research; Oslo, Norway) (Athanasiu et al., 2010), FBIRN 3 (Functional imaging Biomedical Information Research Network, multiple sites in the USA) (Potkin et al., 2009; Wible et al., 2009), COBRE (Center of Biomedical Research Excellence, Albuquerque, NM, USA) (Aine et al., 2017; Gupta et al., 2015), B-SNIP (Bipolar and Schizophrenia Network for Intermediate Phenotypes, multiple sites in the USA) (Meda et al., 2014), MCIC (MIND Clinical Imaging Consortium, Albuquerque, NM, USA) (Gollub et al., 2013), and HUBIN
(Human Brain Informatics, Stockholm, Sweden.) (Nesvag et al., 2008). Data were collected under the approval of local institutional review boards, and all participants provided informed consent. All studies provided structural MRI data as well as PANSS item scores. In total, data were obtained from 1065 patients with a diagnosis of SZ, 268 patients with a diagnosis of BD, and 1011 NC, as confirmed by the Structured Clinical Interview for Diagnosis (SCID) for DSM-IV or DSM-IVTR conducted at each study site (First, 2004). All studies provided information on age at the time of the scan, sex, duration of illness, and medication status (in varying details) for each participant. Scanning information is included in table 1.

**Table 1** Scanning and site information for each dataset

<table>
<thead>
<tr>
<th>Study</th>
<th>Size</th>
<th>Sites</th>
<th>Scanner (T)</th>
<th>Sequence</th>
<th>Voxel size (mm)</th>
<th>Orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>FBIRN3</td>
<td>356</td>
<td>8</td>
<td>Siemens Tim Trio (3)</td>
<td>MPRAGE</td>
<td>1.1 x 0.9 x 1.2</td>
<td>Sagittal</td>
</tr>
<tr>
<td>TOP</td>
<td>414</td>
<td>1</td>
<td>Siemens (1.5)</td>
<td>MPRAGE</td>
<td>1.33 x 0.94 x 1</td>
<td>Sagittal</td>
</tr>
<tr>
<td>COBRE</td>
<td>156</td>
<td>1</td>
<td>Siemens Tim Trio (3)</td>
<td>MPRAGE</td>
<td>1 x 1 x 1</td>
<td>Sagittal</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Philips Achieva (3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Siemens Allegra (3)</td>
<td>MPRAGE</td>
<td>1 x 1 x 1</td>
<td>Sagittal</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Siemens Trio (3)</td>
<td>IR-SPGR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GE Signa HDxt (3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Siemens Trio (3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B-SNIP</td>
<td>395</td>
<td>5</td>
<td>Siemens Allegra (3)</td>
<td>MPRAGE</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Siemens Trio (3)</td>
<td>IR-SPGR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GE Signa HDxt (3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Siemens Trio (3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Siemens (1.5)</td>
<td>Grad Echo</td>
<td>0.625 x 0.625 x 1.5</td>
<td>Coronal</td>
</tr>
<tr>
<td>MCIC</td>
<td>233</td>
<td>4</td>
<td>GE Signa (1.5)</td>
<td>Grad Echo</td>
<td>0.664 x 0.664 x 1.6</td>
<td>Coronal</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Siemens Trio (3)</td>
<td>MPRAGE</td>
<td>0.625 x 0.625 x 1.5</td>
<td>Coronal</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Siemens (1.5)</td>
<td>Grad Echo</td>
<td>0.625 x 0.625 x 1.5</td>
<td>Coronal</td>
</tr>
<tr>
<td>HUBIN</td>
<td>195</td>
<td>1</td>
<td>GE Signa (1.5)</td>
<td>SPGR</td>
<td>1 x 1 x 1</td>
<td>Coronal</td>
</tr>
</tbody>
</table>

In aim 1, data of the SZ group were pooled from three major studies: TOP, FBIRN 3, and COBRE. In total, data were obtained from 342 patients with a diagnosis of schizophrenia,
schizoaffective, or schizotypal disorders as confirmed by the Structured Clinical Interview for Diagnosis (SCID) for DSM-IV or DSM-IV-TR conducted at each study site (Mennigen et al., 2019). Demographic and scanning information is presented in Table 2.

**Table 2** Demographic and scanning information of SZ participants

<table>
<thead>
<tr>
<th></th>
<th>FBIRN 3 (156)</th>
<th>TOP (110)</th>
<th>COBRE (71)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>39.4 ± 11.7a</td>
<td>30.9 ± 8.2b</td>
<td>36.7 ± 12.3a</td>
</tr>
<tr>
<td>Duration of illness (years)</td>
<td>18.0 ± 11.8a</td>
<td>7.1 ± 6.3b</td>
<td>16.0 ± 12.0a</td>
</tr>
<tr>
<td>AP medication (in%)</td>
<td>95.6</td>
<td>88.2</td>
<td>91.5</td>
</tr>
<tr>
<td>CPZ equivalent (in mg/d)</td>
<td>542.3 ± 1271.4</td>
<td>n/a</td>
<td>516.3 ± 1095.4</td>
</tr>
<tr>
<td>PANSS total score</td>
<td>59.2 ± 14.6</td>
<td>61.8 ± 17</td>
<td>59.9 ± 14.3</td>
</tr>
<tr>
<td>Field strength (T)</td>
<td>3</td>
<td>1.5</td>
<td>3</td>
</tr>
<tr>
<td>Sequence</td>
<td>MPRAGE</td>
<td>MPRAGE</td>
<td>MPRAGE</td>
</tr>
<tr>
<td>Voxel size (mm)</td>
<td>1.1 x 0.9 x 1.2</td>
<td>1.3 x 0.9 x 1</td>
<td>1 x 1 x 1</td>
</tr>
</tbody>
</table>

AP – antipsychotic medication, CPZ – chlorpromazine equivalent. The Superscript a and b indicated the result of ANOVA with post-hoc pairwise comparisons (with Bonferroni correction).

The BD study used data from 110 patients with bipolar I (n=69) or II (n=41) diagnoses from the TOP study (Jiang et al., 2020; Rimol et al., 2012). Patients met the following criteria to be included in the study: 1) aged from 18 to 65 years; 2) understood and spoke a Scandinavian language; 3) had no history of severe head trauma; 4) had IQ over 70. The diagnosis of bipolar I (DSM-IV 296.0-7) and bipolar 2 (DSM-IV 296.89) was established by the Structured Clinical Interview for DSM-IV Axis I Disorders. At the stage of enrollment, the mood status of the patients was identified as depressive (n=60), manic (n=5), euthymic (n=40), mixed (n=2), and unspecified (n=3). In addition, 54 patients (approximately 50%) experienced at least one lifetime
psychotic episode. PANSS scores and structural MRI data were available for all the participants, and detailed demographic and scanning information is presented in Table 3.

### Table 3 Demographic and scanning information of BD participants

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>34.87</td>
<td>11.65</td>
</tr>
<tr>
<td>Females (N/%)</td>
<td>72 (65%)</td>
<td></td>
</tr>
<tr>
<td>Duration of illness (years)</td>
<td>12.13</td>
<td>10.09</td>
</tr>
<tr>
<td>PANSS positive score</td>
<td>9.81</td>
<td>3.28</td>
</tr>
<tr>
<td>PANSS negative score</td>
<td>9.75</td>
<td>3.54</td>
</tr>
<tr>
<td>PANSS general score</td>
<td>25.42</td>
<td>5.82</td>
</tr>
<tr>
<td>PANSS total score</td>
<td>44.96</td>
<td>10.33</td>
</tr>
<tr>
<td>Field strength (T)</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>Sequence</td>
<td>MPRAGE</td>
<td></td>
</tr>
<tr>
<td>Voxel size (mm)</td>
<td>1.33 x 0.94 x 1</td>
<td></td>
</tr>
<tr>
<td>Scanning orientation</td>
<td>Sagittal</td>
<td></td>
</tr>
</tbody>
</table>

SD = standard deviation

#### 2.2.2 Positive and negative symptom scale

The PANSS is one of the most commonly used questionnaires to assess symptomatology in patients with SZ (Kay, Fiszbein, & Opler, 1987). It allows for the assessment of dimension-specific abnormalities across positive, negative, and general symptoms. In PANSS, the positive symptoms score includes delusions, conceptual disorganization, hallucinations, excitement, grandiosity, suspiciousness, and hostility. The negative scores includes blunted affect, emotional withdrawal, poor rapport, passive/apathetic social withdrawal, difficulty in abstract thinking, lack of spontaneity and flow of conversation, and stereotyped thinking. Importantly, PANSS has already proved its usefulness in discovering the potential relationship between clinical dimensions and brain alterations (Koutsouleris et al., 2008).
2.2.3 Preprocessing of imaging data

T1-weighted structural MRI data were normalized to the standard Montreal Neurological Institute (MNI) template using a 12-parameter affine model, resliced to a voxel size of 2 x 2 x 2mm, and segmented into GM, white matter, and cerebrospinal fluid using Statistical Parametric Mapping 12 (SPM12, http://www.fil.ion.ucl.ac.uk/spm/software/spm12/), ending with normalized segmented images, prior to smoothing. To identify possible outliers, individual scans were correlated with the group-generated GM template, and outliers were visually checked. Scans showing low correlations with the group-generated GM template ($\rho < 0.98$) were discarded. In addition, since the pICA would be carried out independently on SZ and BD, the regression of age, gender, head motions (subjects with available functional images) and/or site was performed on each diagnostic group separately. Voxel-wise regression was applied to remove the effects of these three variables on patients with SZ. However, only age and sex were included in the regression model for patients with BD. Finally, structural MRI data were smoothed with a 10mm Gaussian kernel (Silver, Montana, Nichols, & Alzheimer's Disease Neuroimaging, 2011).

2.2.4 Parallel Independent Component Analysis

In aim 1, the fusion ICA Toolbox (FIT, http://mialab.mrn.org/software/fit/) was used to perform pICA on the smoothed, normalized structural MRI data and raw PANSS scores (30 items per participant) on SZ and BD separately. The minimum description length (MDL) algorithm (Rissanen, 1978) was used to estimate the optimal component numbers of the structural MRI data. Each PANSS item was designed to represent distinct psychopathological symptoms. Considering the inner structure of the PANSS (positive, negative, and general dimensions), the model order was set to three independent components. Importantly, a PANSS
component derived from the ICA reflected any combination of PANSS items based on the data itself, and the three PANSS components did not necessarily reflect the preset inner structure. Within the pICA framework, a GM mask was applied to structural MRI data before ICA was performed. To ensure stability of component estimation, infomax ICA ran 20 times, and the central point of 20 runs was selected as the final component using ICASSO (Himberg, Hyvarinen, & Esposito, 2004).

The numbers of pairs of GM components and PANSS components varying from 1 to 3 (default setting) were included and tested according to the actual constrain situation. Components that showed the highest (from 1\textsuperscript{st} to 3\textsuperscript{rd}) correlation was constrained in the analysis in order to yield more stable results. More specifically, the correlation between loading coefficients of this structural MRI-PANSS pair was enhanced through pICA, while independence within each modality was further maximized.

2.2.5 \textit{Statistical analyses}

Statistical analysis was performed in MATLAB 2017b. (MATLAB and Statistics Toolbox Release 2017b, the MathWorks, Inc., Natick, MA) and SPSS 17.0 (SPSS, Inc., Chicago, IL).

2.2.5.1 \textit{Correlation between loading coefficients of sMRI components and PANSS components}

To determine whether the loading coefficients of sMRI components correlated with the loading coefficients of the PANSS components, the correlation analysis was performed within the extracted pairs of sMRI components and PANSS components during the pICA process by fusion toolbox. Note that age, gender, and site effects were previously removed from the structural brains. The conservative Bonferroni correction was employed for multiple testing. This
threshold was decided by all the combinations of all extracted components from both modalities. To indicate significance, $P$ values had to exceed $0.05/(\text{number of sMRI components} \times \text{number of PANSS components})$.

In addition, all pICA analyses were validated by 10-fold validation (Salman et al., 2019; van der Gaag, Hoffman, et al., 2006). Each validation included 90% of the individuals of the sample and was balanced across different study sites if applicable. Parameter settings were the same as in the original analysis. The results from each iteration were examined for overlap, i.e., correlation, with the original pICA output.

2.2.5.2 Potential confounders in SZ

Since the data of SZ patients were selected from multi-center studies, the potential effects of study site and illness duration on PANSS scores were explored in a multivariate analysis of co-variance (MANCOVA) framework: the 30 PANSS items (multivariate dependent variable) were tested for associations with the study site (independent variable) while controlling for illness duration (covariate). In addition, the PANSS total scores (dependent variable) were tested for an association with study site (independent variable) controlling for illness duration (covariate) employing analysis of co-variance (ANCOVA).

Duration of illness (DOI) and medication status were typical confounders in studies involving patients with SZ. In a regression framework, the possible effects of DOI (independent variable) on the loading coefficients of the sMRI component and the PANSS component (dependent variable) were tested. We tested for the effect of medication on both modality loading coefficients in participants for whom chlorpromazine (CPZ) equivalents were available. Spearman’s rho was tested because of a skewed distribution.
To further explore the relationship between the traditional PANSS dimension scores and the data-driven dimensions of the PANSS and sMRI components derived from the pICA, correlation analyses were performed on PANSS positive, negative, and general scores and the loading coefficients of the components. An alpha level of .05 was applied as our significance threshold for all the above analyses.

### 2.2.5.3 Potential confounders in BD

BD patients were collected from one site. However, DOI and medication status were also assessed as in the SZ. DOI (predictor) on the loading coefficients of the sMRI component and the PANSS component (dependent variable) was tested in a regression framework. We also tested the medication effect on the loading coefficients of both modalities with Spearman’s rho. Different from CPZ in SZ samples, the current medication status was included as a binary variable according to whether or not the patients were currently on antipsychotic, antiepileptic, or lithium treatment.

In addition, the relations of original PANSS subscale scores (positive, negative, and general symptoms), PANSS loading coefficients and sMRI loading coefficients from pICA were assessed following the same step in SZ patients for correlations. An alpha level of .05 was applied as our significance threshold for all the above analyses.

### 2.3 Results

#### 2.3.1 SZ group results

##### 2.3.1.1 The PANSS scores

The PANSS total score did not differ significantly between sites of the three different studies after controlling for illness duration ($F = 1.57$, df = 7, $p = 0.14$). However, the MANCOVA revealed significant differences in several PANSS items (see Table 4).
Table 4 Significant MANCOVA results of study site differences on PANSS items corrected for illness duration at p < 0.05, uncorrected

<table>
<thead>
<tr>
<th>PANSS item</th>
<th>F statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>P2 Conceptual disorganization</td>
<td>2.18</td>
<td>0.028</td>
</tr>
<tr>
<td>P7 Hostility</td>
<td>2.83</td>
<td>0.005</td>
</tr>
<tr>
<td>N2 Emotional withdrawal</td>
<td>2.01</td>
<td>0.045</td>
</tr>
<tr>
<td>N3 Poor rapport</td>
<td>4.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>N5 Difficulty abstract thinking</td>
<td>12.39</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>N6 Lack of spontaneity</td>
<td>2.15</td>
<td>0.031</td>
</tr>
<tr>
<td>N7 Stereotyped thinking</td>
<td>2.29</td>
<td>0.021</td>
</tr>
<tr>
<td>G4 Tension</td>
<td>2.13</td>
<td>0.033</td>
</tr>
<tr>
<td>G5 Mannerism</td>
<td>2.86</td>
<td>0.004</td>
</tr>
<tr>
<td>G6 Depression</td>
<td>2.78</td>
<td>0.005</td>
</tr>
<tr>
<td>G9 Unusual thought content</td>
<td>2.18</td>
<td>0.029</td>
</tr>
<tr>
<td>G10 Disorientation</td>
<td>3.83</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>G11 Poor attention</td>
<td>2.46</td>
<td>0.013</td>
</tr>
<tr>
<td>G12 Lack of judgement</td>
<td>4.12</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>G13 Disturbance of volition</td>
<td>3.51</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>G16 Active social avoidance</td>
<td>3.1</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Importantly, items that contributed most to the PANSS component, including delusional symptoms (P1), suspiciousness/persecution (P6), and anxiety (G2), were not significantly associated with study site (see Table 5).
Table 5 PANSS items contributing most to the first PANSS component with $|Z| > 1$

<table>
<thead>
<tr>
<th>PANSS item</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1 Delusion</td>
<td>2.87</td>
</tr>
<tr>
<td>P6 Suspiciousness</td>
<td>2.52</td>
</tr>
<tr>
<td>G2 Anxiety</td>
<td>1.24</td>
</tr>
<tr>
<td>P3 Hallucination</td>
<td>1.14</td>
</tr>
<tr>
<td>N3 Poor rapport</td>
<td>-1.06</td>
</tr>
<tr>
<td>G13 Disturbance of volition</td>
<td>-1.09</td>
</tr>
<tr>
<td>G8 Uncooperativeness</td>
<td>-1.22</td>
</tr>
<tr>
<td>G10 Disorientation</td>
<td>-1.23</td>
</tr>
<tr>
<td>G5 Mannerisms and posturing</td>
<td>-1.31</td>
</tr>
</tbody>
</table>

2.3.1.2 Parallel Independent Component Analysis

The correlation between the structural MRI and PANSS components was Pearson’s $r = 0.25$ ($p = 2.62e^{-06}$, Bonferroni-corrected threshold $= 6.67^{-04}$) shown in Figure 2. The 10-fold validation showed that results in 8 out of the 10 validation runs were consistent with findings of the original analysis. Given that the correlation was positive, individuals with greater structural loading coefficients for this component also showed greater symptom profile loadings.
Figure 2 Correlation between loading coefficients of the top PANSS component and the top structural MRI component. Pearson’s $r = 0.25$, $p = 2.62 \times 10^{-06}$ (Bonferroni-corrected threshold = $6.67 \times 10^{-04}$) (Mennigen et al., 2019).

The structural MRI component consisted of lower GM concentration in bilateral inferior temporal gyrus merging with fusiform gyrus and the gyrus rectus, and higher GM concentration in bilateral pre- and postcentral gyrus (see Figure 3a and Table 6).

The PANSS component was most strongly positively weighted on delusional symptoms (P1), hallucinatory behavior (P3), suspiciousness/persecution (P6), and anxiety (G2). Poor rapport (N3), mannerism and posturing (G5), uncooperativeness (G8), disorientation (G10), and disturbance of volition (G13) were negatively associated with this component (Figure 3b).

In summary, patients with higher reductions in inferior temporal gyrus and fusiform gyrus and increased GM concentration in pre- and postcentral gyrus, also exhibit a more severe symptom profile consisting of higher delusional symptoms, suspiciousness, and anxiety and lower symptoms of mannerism, disorientation, and uncooperativeness.
The 10-fold validation showed high correlation with the original pICA run with a mean correlation of Pearson’s r = 0.81; in 2 out of 10 validation runs the directionality of the structural MRI component and therefore its association with the PANSS component was inverted.

**Table 6** Brain regions, volumes and peak coordinates in SZ group

<table>
<thead>
<tr>
<th>Structural MRI clusters</th>
<th>Brain regions</th>
<th>Volume (cc)^3</th>
<th>Brodmann area</th>
<th>Z-score and peak coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 1</td>
<td>R Fusiform Gyrus</td>
<td>0.59</td>
<td>20, 36</td>
<td>-2.8 (22,2,-40)</td>
</tr>
<tr>
<td></td>
<td>R Inferior Temporal Gyrus</td>
<td>0.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cluster 2</td>
<td>R Fusiform Gyrus</td>
<td>0.83</td>
<td>20, 36, 37</td>
<td>-3.6 (46,-38,-28)</td>
</tr>
<tr>
<td></td>
<td>R Inferior Temporal Gyrus</td>
<td>1.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cluster 3</td>
<td>L Fusiform Gyrus</td>
<td>0.85</td>
<td>20, 36, 37</td>
<td>-3.3 (-42,-34,-22)</td>
</tr>
<tr>
<td></td>
<td>L Inferior Temporal Gyrus</td>
<td>0.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cluster 4</td>
<td>Rectal Gyrus</td>
<td>0.78</td>
<td>11, 47</td>
<td>-3.0 (10,26,-26)</td>
</tr>
<tr>
<td>Cluster 5</td>
<td>R Precentral Gyrus</td>
<td>3.2</td>
<td>3, 4, 6</td>
<td>3.8 (18,-28,70)</td>
</tr>
<tr>
<td></td>
<td>R Postcentral Gyrus</td>
<td>2.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cluster 6</td>
<td>L Precentral Gyrus</td>
<td>2.5</td>
<td>3, 4, 6</td>
<td>4.0 (-12,-22,-72)</td>
</tr>
<tr>
<td></td>
<td>L Postcentral Gyrus</td>
<td>0.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^Z-threshold > 2, and clusters size > 150 voxel. L - left, R - right.
a) Structural MRI component

Figure 3 Parallel ICA results in SZ group - the top a) structural MRI and b) PANSS components. PANSS items with $|Z| > 1$: P1 – delusion symptoms; P3 – hallucinatory behavior; P6 – suspiciousness/ persecution; N3 – poor rapport; G2 – anxiety; G5 – mannerisms and posturing;
G8 – uncooperativeness; G10 – disorientation; G13 – disturbance of volition (Mennigen et al., 2019).

2.3.1.3 Additional analyses

The regression models including loading coefficients of the structural MRI component or PANSS component (dependent variable) and illness duration (predictor) showed no significant association ($\beta_{\text{sMRI}} = 0.25$, $t_{\text{sMRI}} = 0.45$, $p_{\text{sMRI}} = 0.65$; $\beta_{\text{PANSS}} = 0.03$, $t_{\text{PANSS}} = 0.62$, $p_{\text{PANSS}} = 0.54$).

Within the subsample with available CPZ equivalents there were no significant correlations between CPZ equivalents and the structural MRI loading coefficients (Spearman’s $\rho = 0.08$, $p = 0.25$) or loading coefficients for the PANSS component (Spearman’s $\rho = 0.12$, $p = 0.1$). CPZ equivalents did not differ significantly between the two studies providing these data ($t = -0.15$, $df = 197$, $p = 0.89$).

The correlation between the traditional PANSS dimension scores and the experimentally derived PANSS loading coefficients revealed high correlations with PANSS positive (Pearson’s $r = 0.83$, $p < 0.001$), negative (Pearson’s $r = 0.51$, $p < 0.001$), and general (Pearson’s $r = 0.85$, $p < 0.001$) scores.

Likewise, loading coefficients of the structural MRI component were significantly correlated with PANSS positive (Pearson’s $r = 0.23$, $p < 0.001$), negative (Pearson’s $r = 0.14$, $p = 0.008$), and general (Pearson’s $r = 0.22$, $p < 0.001$) scores.

2.3.2 BD group results

2.3.2.1 Parallel Independent Component Analysis

Two pairs of sMRI and PANSS components showed significant correlations, passing the Bonferroni correction threshold of $6.67 \times 10^{-44}$ (determined by 25 sMRI components and 3
PANSS components). In the first pair of components, the correlation between the structural MRI and PANSS components was Pearson’s $r = -0.41$ ($p = 7.67 \times 10^{-06}$, see Figure 4). In the second pair, the correlation was Pearson’s $r = -0.35$ ($p = 1.80 \times 10^{-04}$, see Figure 5). The negative correlation indicated that greater loading coefficients of sMRI correlated with lower loadings of symptom profile represented by PANSS.

In the first pair, the sMRI component showed positively contributing GM concentration mainly in right middle/superior temporal gyrus in the more medial aspect (extended to inferior frontal gyrus, see Figure 4 and Table 7). The corresponding PANSS component was positively weighted strongly on anxiety (G2), guilty feelings (G3) and depression (G6); the lack of spontaneity and flow of conversation (N6) was negatively associated with this component (see Figure 4). Participants with higher preserved GM concentration in right superior/middle temporal gyrus exhibited reduced severity in these symptoms consisting of anxiety, guilty feelings and depression.
Figure 4 Parallel ICA results in BD group – ‘mood’ component a) structural MRI and b) PANSS component. PANSS items with |Z| >1: N6 – Lack of spontaneity and flow of conversation; G2 – Anxiety; G3 – Guilty feelings; G6 – Depression (Jiang et al., 2020).

In the second pair, positively contributing GM concentration was mainly located in bilateral middle frontal gyrus, bilateral parietal lobule including postcentral gyrus with temporal regions on the left side (left inferior, middle and superior temporal gyrus), and negatively contributing GM concentration was located in left supramarginal gyrus for the sMRI component (see Figure 5 and Table 7). The PANSS component was positively weighted strongly on blunted affect (N1), emotional withdrawal (N2), passive/apathetic social withdrawal (N4), depression (G6) and active social avoidance (G16), and excitement (P4) and guilty feeling (G3) was negatively associated with this component (Figure 5). In summary, participants with higher preserved GMC in bilateral frontal, parietal and left temporal regions show milder severity for these symptoms.

Figure 5 Parallel ICA results in BD group – ‘apathy’ component a) structural MRI and b) PANSS components. PANSS items with |Z| >1: P4 – Excitement; N1 – Blunted affect; N2 –
Emotional withdrawal; N4 – Passive/apathetic social withdrawal; G3 – Guilty feelings; G6 – Depression; G16 – Active social avoidance (Jiang et al., 2020).

Table 7 Brain regions, volumes and peak coordinates in BD group

<table>
<thead>
<tr>
<th>Component 1</th>
<th>Brain regions</th>
<th>Brodmann area</th>
<th>Volumes (cc&lt;sup&gt;3&lt;/sup&gt;)</th>
<th>Peak coordinates (Z and coordinates)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster 1</td>
<td>R superior temporal gyrus (extended to inferior frontal gyrus &amp; insula)</td>
<td>48</td>
<td>3.6</td>
<td>6.0 (48, 14, -6)</td>
</tr>
<tr>
<td>Cluster 2</td>
<td>R middle temporal gyrus</td>
<td>39</td>
<td>2.7</td>
<td>7.8 (38, -62, 22)</td>
</tr>
</tbody>
</table>

Component 2

| Cluster 1/2 | L/R parietal lobe (postcentral gyrus) | 3,3       | 5.7/5.5                  | 3.3(-40,-26, 40)/5.5(46,-26, 36) |
| Cluster 3/4 | L/R frontal lobe (middle frontal gyrus) | 9,8       | 5.4/3.0                  | 5.4(-26,32,32)/6.4(22,12, 46) |
| Cluster 5   | L inferior middle temporal gyrus | 37         | 3.4                      | 7.6(-46,-58,-6)                     |
| Cluster 6   | L inferior parietal lobule | 41         | 2.9                      | 7.6(-46,-46,22)                     |
| Cluster 7   | L parietal lobe (L angular gyrus) | 7          | 2.7                      | 7.3(-30,-60,38)                     |
| Cluster 8   | L supramarginal gyrus | 40         | 4.6                      | -7.3(-42,-46,38)                     |

For both components, Z score was set > 3, and cluster volumes were set > 2 cc<sup>3</sup> to retain their most significant results. Brodmann areas are listed for the peak coordinates. The directions of
peak Z scores indicate whether the brain region contributed positively or negatively to the component as red for positive and blue for negative in brain maps.

### 2.3.2.2 Additional analyses

The 10-fold validation showed high correlation with the original pICA results in both of the pairs of components. For the first pair, 7 out of 10 validation runs showed similar results to the original output, with one of run showing an inverted direction of sMRI and PANSS correlation to the other runs. The first pair presented a mean correlation of Pearson’s $r = 0.75$ of above replicated runs. The second pair also presented similar stability, with 7 out of 10 validation runs showing similar results to the original output and two of the runs showing inverted direction of sMRI and PANSS correlation to the other runs. The second pair presented a mean correlation of Pearson’s $r = 0.67$ of its replicated runs.

In the regression model including loadings coefficients of sMRI components as dependent variable and DOI values as predictors, there was no significant association between the DOI and component 1 ($\beta = -0.10, t = -1.06, p = 0.29$) or component 2 ($\beta = -0.004, t = -0.043, p = 0.97$). The regression model including loadings coefficients of PANSS components as dependent variable and DOIs as predictors also showed no significant association in component 1 ($\beta = 0.01, t = 0.08, p = 0.94$) or component 2 ($\beta = -0.05, t = -0.49, p = 0.63$).

There was no significant correlation between medication status and the sMRI loading coefficients in component 1 (Spearman’s $\rho = 0.10, p = 0.33$) or component 2 (Spearman’s $\rho = -0.06, p = 0.56$). In addition, there was no significant correlation between medication status and the PANSS loading coefficients in component 1 (Spearman’s $\rho = -0.06, p = 0.57$) or component 2 (Spearman’s $\rho = -0.14, p = 0.88$).
In the first pair, the correlation between the original PANSS subscale scores and the pICA derived PANSS loading coefficients revealed no association with PANSS positive (Pearson’s $r = 0.017$, $p = 0.86$), negative (Pearson’s $r = 0.15$, $p = 0.88$), or general (Pearson’s $r = 0.062$, $p = 0.53$) scores. In addition, sMRI component loading coefficients were not significantly correlated with PANSS positive (Pearson’s $r = 0.023$, $p = 0.82$), negative (Pearson’s $r = -0.039$, $p = 0.69$), and general (Pearson’s $r = -0.031$, $p = 0.75$) scores.

In the second pair, there was also no significant correlation found between the pICA PANSS component loading and PANSS positive (Pearson’s $r = 0.074$, $p = 0.45$), negative (Pearson’s $r = 0.11$, $p = 0.27$), or general (Pearson’s $r = 0.11$, $p = 0.26$) scores. In the sMRI, the loading coefficients were not significantly correlated with PANSS positive (Pearson’s $r = 0.072$, $p = 0.45$), negative (Pearson’s $r = -0.031$, $p = 0.75$) and general (Pearson’s $r = -0.061$, $p = 0.53$) scores.

All aims will discussed together in the final section.
3  AIM 2: PROJECTION OF EXTRACTED BRAIN PATTERNS INTO EITHER DIAGNOSTIC GROUP

3.1  Aim and hypothesis

In this study, the second aim was to determine if SZ brain components accounted for corresponding symptom profiles when projected to the BD samples and vice versa. The correlations between the projected loading coefficients of brain and PANSS components were tested, and additional correlation analyses between projected loadings of brain patterns and item scores of PANSS were tested. We hypothesized that the brain component from SZ or BD will be associated with corresponding PANSS component when they are projected to the other diagnostic group.

3.2  Methods

3.2.1  Participants

The participants included in aim 2 stayed the same with final subjects included in aim 1. Totally 337 SZ and 109 BD participants were included in this analysis.

3.2.2  Data Projection

The brain and PANSS components of each diagnostic group were projected to each other’s preprocessed imaging data and PANSS data to extract the projected loading coefficients. Here we used brain components projection to explain the whole process: 1) We used the pinv function in Matlab to extract all the coordinates of the brain components (positive symptom profile related brain component, for example) in the original SZ data (only including voxels whose GMC value is larger than 0); 2) We extracted the brain component according to above coordinates in the whole output matrix of the pICA in SZ data, and we named it as matrix 1; 3)
We used the intersect function in Matlab and generated the shared mask from the mask files of SZ and BD data; 4) We read in the preprocessed BD subjects’ images and reshaped it as matrix 2; 5) Through matrix 1 multiplied by matrix 2 within the shared mask generated in 3), we got the projected SZ brain pattern in BD brain images; 6) The projected loading coefficients were extracted for further analysis. The projection of BD brain components and PANSS components was made through a similar process as above.

The projection analysis ensured the same brain and PANSS components would be kept the same as in the original group, so it was possible to re-assess their correlations in the other group. It had to be pointed out here, by making a projection, we assumed the projected pattern existed in both of the datasets (Chen et al., 2019; Luo et al., 2018).

3.2.3 Statistical Analyses

Statistical analysis was performed in MATLAB 2017b.( MATLAB and Statistics Toolbox Release 2017b, the MathWorks, Inc., Natick, MA ) and SPSS 17.0 (SPSS, Inc., Chicago, IL).
3.2.3.1 Correlation between projected loading coefficients of sMRI components and PANSS components

We tested whether the projected sMRI loading coefficient correlated with the projected PANSS loading coefficient and was performed within the sMRI components and the corresponding PANSS component in both SZ and BD patients. The conservative Bonferroni correction was employed for multiple testing, and a corrected p < .05 was applied as our significance threshold for all the analyses.

3.2.3.2 Correlation between projected loading coefficients of sMRI components and PANSS item scores

Aware of the possibility of the non-existence of PANSS components across the diagnostic group, we also explored the potential relationship between projected sMRI components and PANSS items. Correlation analyses were performed between sMRI loadings and 30 PANSS item scores, positive, negative, and total scores. An alpha level of .05 will be applied as our significance threshold for all the above analyses.

3.3 Results

3.3.1 PANSS item scores and variations in SZ and BD

We graphed PANSS items scores and variation levels in SZ and BD for a general map to show how these SZ and BD populations might differ from each other clinically. SZ participants generally showed a higher level of item scores in most PANSS items. However, BD participants showed higher scores in anxiety, guilty feelings, and depression (see Figure 6).
Figure 6 PANSS Item scores in SZ vs. BD. In the figure a), blue showed the scores of SZ, and yellow showed scores of BD. PANSS items highlighted in previous analysis were marked with text. In the figure b) (the radar map), blue showed the scores of SZ, and red showed the scores of BD.
We also graphed PANSS items score variation levels in SZ and BD. SZ participants generally showed higher variation levels in most PANSS items. However, BD participants showed higher or similar variation levels in anxiety, guilty feelings, tension, and depression (see Figure 7).

Figure 7 PANSS Item variations in SZ vs. BD. In the figure a), blue showed the variations of SZ, and red showed scores of BD. PANSS items showing similar variation levels were marked with text. In the figure b) (the radar map), blue showed the scores of SZ, and red showed the scores of BD.
3.3.2 Projection results from SZ to BD

After the projection of both positive symptom profiles and related brain patterns from SZ to BD, the correlation between brain component and PANSS component showed insignificant correlation (Pearson’s $r = 0.02$, $p = 0.84$). However, it was weakly in the same direction as the original findings (Figure 8).

![Figure 8](image)

Figure 8 Projecting brain and PANSS components from SZ to BD. The upper part of the figure showed original results in SZ (the positive symptom profile and corresponding brain component). The lower part of the figure showed after projecting both of brain and PANSS, they showed insignificant correlation while stayed same direction (Pearson’s $r = 0.02$, $p = 0.84$).
In addition, the projected brain components showed significant correlation with several PANSS items in BD participants, as shown in Table 8.

**Table 8** Projected brain component loadings’ correlation with BD PANSS item scores

<table>
<thead>
<tr>
<th>PANSS order</th>
<th>PANSS items</th>
<th>Pearsons’ r</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive_05</td>
<td>Grandiosity</td>
<td>0.21</td>
<td>0.026*</td>
</tr>
<tr>
<td>Positive_07</td>
<td>Hostility</td>
<td>0.22</td>
<td>0.022*</td>
</tr>
<tr>
<td>General_14</td>
<td>Poor impulse control</td>
<td>0.25</td>
<td>0.010*</td>
</tr>
<tr>
<td>General_16</td>
<td>Active social avoidance</td>
<td>0.19</td>
<td>0.043*</td>
</tr>
<tr>
<td>Total positive symptom score</td>
<td>-</td>
<td>0.18</td>
<td>0.068</td>
</tr>
<tr>
<td>Total negative symptom score</td>
<td>-</td>
<td>0.026</td>
<td>0.79</td>
</tr>
<tr>
<td>Total general symptom score</td>
<td>-</td>
<td>0.18</td>
<td>0.065</td>
</tr>
<tr>
<td>Total PANSS score</td>
<td>-</td>
<td>0.16</td>
<td>0.087</td>
</tr>
</tbody>
</table>

The significant correlations (P value < .05) were marked with asterisk.

### 3.3.3 Projection results from BD to SZ

#### 3.3.3.1 Projection from BD mood symptom profiles and related brain component

After the projection of both mood symptom profiles and related brain patterns from BD to SZ, the correlation between brain component and PANSS component showed an insignificant correlation (Pearson’s r = -0.06, p = 0.27) while it stayed the same direction as the original findings (Figure 9).
Figure 9 Projecting brain and PANSS components (mood) from BD to SZ. The upper part of the figure showed original results in BD (the mood symptom profile and corresponding brain component). The lower part of the figure showed after projecting both of brain and PANSS, they showed insignificant correlation while stayed same direction (Pearson’s r = -0.06, p = 0.27).

However, the projected brain components showed significant correlations with several PANSS items in SZ participants. For detailed results, see table 9.
Table 9: Projected brain component (mood) loadings’ correlation with SZ PANSS item scores

<table>
<thead>
<tr>
<th>PANSS order</th>
<th>PANSS items</th>
<th>Pearsons’ r</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive_02</td>
<td>Conceptual disorganization</td>
<td>-0.12</td>
<td>0.026*</td>
</tr>
<tr>
<td>Positive_04</td>
<td>Excitement</td>
<td>-0.11</td>
<td>0.047*</td>
</tr>
<tr>
<td>Positive_05</td>
<td>Grandiosity</td>
<td>-0.15</td>
<td>0.0050*</td>
</tr>
<tr>
<td>Negative_05</td>
<td>Difficulty in abstract thinking</td>
<td>-0.12</td>
<td>0.033*</td>
</tr>
<tr>
<td>General_08</td>
<td>Uncooperativeness</td>
<td>-0.12</td>
<td>0.024*</td>
</tr>
<tr>
<td>General_10</td>
<td>Disorientation</td>
<td>-0.17</td>
<td>0.0020*</td>
</tr>
<tr>
<td>General_15</td>
<td>Preoccupation</td>
<td>-0.17</td>
<td>0.0020*</td>
</tr>
<tr>
<td>Total positive symptom score</td>
<td>-</td>
<td>-0.13</td>
<td>0.020*</td>
</tr>
<tr>
<td>Total negative symptom score</td>
<td>-</td>
<td>-0.083</td>
<td>0.13*</td>
</tr>
<tr>
<td>Total general symptom score</td>
<td>-</td>
<td>-0.12</td>
<td>0.030*</td>
</tr>
<tr>
<td>Total PANSS score</td>
<td>-</td>
<td>-0.14</td>
<td>0.012*</td>
</tr>
</tbody>
</table>

The significant correlations (P value < .05) were marked with asterisk.

3.3.3.2 Projection form BD apathy symptom profiles and related brain component

After the projection of both apathy symptom profiles and related brain patterns from BD to SZ, the correlation between brain component and PANSS component showed insignificant correlation (Pearson’s r = -0.018, p = 0.75) while it stayed the same direction as the original findings (Figure 10).

However, the projected brain components showed significant correlation with PANSS item of grandiosity (Pearson’s r = -0.12, p = 0.023).
Figure 10 Projecting brain and PANSS components (apathy) from BD to SZ. The upper part of the figure showed original results in BD (the apathy symptom profile and corresponding brain component). The lower part of the figure showed after projecting both of brain and PANSS, they showed insignificant correlation while stayed same direction (Pearson’s $r = -0.018$, $p = 0.75$).
4 AIM 3: GENETIC MECHANISM BEHIND IDENTIFIED BRAIN PATTERNS

4.1 Aim and hypotheses

The final aim was to determine if the polygenic risks from SZ and BD GWAS results (using diagnostic groups as phenotypes) accounted for the identified brain components extracted by psychotic symptom profiles in previous aims. We calculated both the SZ and BD risk scores for all individuals. We hypothesized that both polygenic risks from SZ and BD will correlate with the brain structural components identified in either diagnostic group. The polygenic risk based on disorder risk of SZ will account for the brain structural components extracted by symptom profiles in the BD group and vice versa.

4.2 Methods

4.2.1 Participants

This aim included two sets of data. To calculate the polygenic risk for SZ, a combined dataset of 526 SZ participants and 611 controls were used. For the calculation of polygenic risk for BD, we were not able to acquire genetic data from TOP, so the 156 BD participants from B-SNIP and the same 611 control participants were used. Table 10 and Table 11 provide detailed demographic and site information.

Table 10 Demographic information for polygenic risk in combined SZ sample

<table>
<thead>
<tr>
<th>Study Name</th>
<th>SZ Number</th>
<th>Number of control</th>
<th>Age</th>
<th>Gender (F:M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FBIRN 3</td>
<td>107</td>
<td>120</td>
<td>38.4±11.2</td>
<td>50:177</td>
</tr>
<tr>
<td>TOP</td>
<td>90</td>
<td>163</td>
<td>33.7±9.1</td>
<td>110:143</td>
</tr>
<tr>
<td>COBRE</td>
<td>23</td>
<td>21</td>
<td>38.0±13.1</td>
<td>11:34</td>
</tr>
<tr>
<td>B-SNIP</td>
<td>132</td>
<td>103</td>
<td>35.5±12.4</td>
<td>89:146</td>
</tr>
<tr>
<td>MCIC</td>
<td>99</td>
<td>121</td>
<td>33.7±11.1</td>
<td>73:147</td>
</tr>
<tr>
<td>HUBIN</td>
<td>4</td>
<td>17</td>
<td>37.4±11.0</td>
<td>7:14</td>
</tr>
<tr>
<td>NW</td>
<td>71</td>
<td>66</td>
<td>33.0±13.3</td>
<td>50:87</td>
</tr>
</tbody>
</table>

Ages were displayed as mean ± standard deviation
Table 11 Demographic information for polygenic risk in B-SNIP BD sample

<table>
<thead>
<tr>
<th>Site</th>
<th>GP</th>
<th>GT</th>
<th>JS</th>
<th>CT</th>
<th>MB</th>
<th>MK</th>
</tr>
</thead>
<tbody>
<tr>
<td>BD</td>
<td>30</td>
<td>28</td>
<td>50</td>
<td>21</td>
<td>7</td>
<td>20</td>
</tr>
<tr>
<td>Number</td>
<td>34.8±12.8</td>
<td>35.8±12.1</td>
<td>34.2±14.4</td>
<td>40.3±10.5</td>
<td>43.1±15.3</td>
<td>35.1±12.1</td>
</tr>
<tr>
<td>Age</td>
<td>16:14</td>
<td>19:9</td>
<td>36:14</td>
<td>16:5</td>
<td>7</td>
<td>13:7</td>
</tr>
</tbody>
</table>

This sample shared the control groups in the polygenic risk calculation. Ages were displayed as mean ± standard deviation

4.2.2 Preprocessing of Genetic Data

Site-wise genetic data were merged based on the common SNPs among different studies (N = 24.2 millions in SZ data and N = 10.9 millions in BD data). Subjects with >10% SNP values missing and SNPs with > 5% values missing were removed from the sample. The minor allele frequency was set to 1%, while 10e-6 was set as the threshold for Hardy-Weinberg equilibrium. Relatives or family members were checked, and subjects and subjects with PI_HAT > 0.18 were removed accordingly. All genetic data used was the hg19 build. Finally, 3.7 millions SNPs were included in SZ group, and 3.5 millions SNPs were included in the BD group.

4.2.3 Polygenic Risk Score Calculation

We employed the GWAS results from Psychiatric Genomics Consortium (PGC) to calculate the polygenic risk score (PRS) in both SZ and BD. PRSice-2 (https://www.prsice.info/) (Choi & O'Reilly, 2019) were used. In the SZ polygenic risk calculation, the PGC schizophrenia study (Schizophrenia Working Group of the Psychiatric Genomics, 2014) (https://www.med.unc.edu/pgc/results-and-downloads/scz/) was used as the base file while case/control situation in the target data was used as a phenotype to decide the threshold applied in PRS modeling. Similarly, the base data for the BD analyses was from the PGC bipolar GWAS study, with more than 13 million SNPs (Stahl et al., 2019)
The full set of GWAS results with more than 9 million SNPs was used for PRS analysis (Schizophrenia Working Group of the Psychiatric Genomics, 2014).

The diagnostic phenotypes were specified as a binary trait. We ran the typical PRS calculations, where the association between the PRS and diagnostic phenotypes in the target data was tested through logistic regression. The best fitting model of diagnostic phenotypes was defined as the set of SNPs for which the PRS explained the greatest variance. The algorithm then produced the best fitting p-value threshold, and it gave only one risk score per subject according to this threshold (Choi & O'Reilly, 2019). We then checked the correlation between the risk score and the loading coefficients of the brain components identified in both SZ and BD in aim 1. A similar process was done in the BD risk calculation.

### 4.2.4 Statistical Analyses

Statistical analysis was performed in MATLAB 2017b. (MATLAB and Statistics Toolbox Release 2017b, the MathWorks, Inc., Natick, MA) and SPSS 17.0 (SPSS, Inc., Chicago, IL). We analyzed whether the polygenic risk score was significantly correlated to the loading coefficients of the sMRI components extracted above. Both polygenic risk scores calculated from SZ and BD risk were used on all sMRI component loadings. To detail, in the SZ group, SZ polygenic risk scores calculated based on SZ case/control difference were tested for correlation with the loading coefficients of all sMRI components highlighted in aim 1. We also tested the correlation between BD polygenic risk scores calculated based on SZ case/control difference and the same brain loadings. The same correlation analyses were run in the BD group. Analyses in the BD group only included BD participants from the B-SNIP dataset. Note that the age, gender, and site effects were previously removed from the brains, and no DOI or medication effect was
detected on previous brain components. Since only one brain feature and two sets of PRS were used in SZ group, and two brain features and two sets of PRS were used in BD, a threshold of \( p < .05 \) was applied as our significance for all the analyses.

### 4.3 Results

#### 4.3.1 Polygenic risk and its correlation with brain component in SZ group

In the SZ group, the most sensitive \( p \)-value threshold to compute PRS based SZ risk was 0.0013, with 13% of the variance in the case vs. control status explained (\( p = 1.30 \times 10^{-21} \), Figure 11a). Under this threshold, a total of 2.9 millions SNPs were included in the PRS model. 186 SZ participants were found with both available polygenic risk scores and brain loadings. The polygenic risk scores were highly correlated with positive symptom profile related brain component loadings (Pearson’s \( r = 0.19 \), \( p = 8.42 \times 10^{-3} \), Figure 11b).

![Figure 11](image)

**Figure 11** The SZ polygenic risk and correlation results in SZ group. A) SZ polygenic risk on case vs. control phenotypes in SZ group, by threshold \( p \)-value on the x axis and variance explained by PRS on the y axis. B) The correlation between polygenic risk score and positive symptom profile-related brain component loadings (Pearson’s \( r = 0.19 \), \( p = 8.42 \times 10^{-3} \)).
The most sensitive $p$-value threshold to compute PRS based BD risk was 0.0044, with 9.5% of the variance in the case vs. control status explained ($p = 3.50E-16$, Figure 12a). Under this threshold, a total of 3.1 millions SNPs were included in the PRS model. The polygenic risk scores were highly correlated with positive symptom profile related brain component loadings (Pearson’s $r = 0.21$, $p = 5.11E-03$, Figure 12b).

**Figure 12** The BD polygenic risk and correlation results in SZ group, similar to Figure 11. The correlation between polygenic risk score and positive symptom profile-related brain component loadings was Pearson’s $r = 0.21$, $p = 5.11E-03$.

### 4.3.2 Polygenic risk and its correlation with brain component in BD group

The most sensitive $p$-value threshold to compute PRS based BD risk was 1.00E-04, with 8.8% of the variance in the case vs. control status explained ($p = 5.60E-11$, Figure 13a). Under this threshold, a total of 2.9 millions SNPs were included in the PRS model. The most sensitive $p$-value threshold to compute PRS based SZ risk was 0.012, with 6.1% of the variance
in the case vs. control status explained ($p = 8.00E-09$, Figure 13b). Under this threshold, a total of 2.8 millions SNPs were included in the PRS model.

### 4.3.2.1 Correlation analysis between PRS and mood component

Among the 156 BD participants, we were able to access the usable images for 140 out of the total B-SNIP sample. The projected loadings in these brains from previously identified brain component of mood and apathy components were used in following analyses. The BD risk scores were correlated with mood symptom profile related brain component loadings (Pearson’s $r = -0.17$, $p = 0.040$, Figure 14a).

**Figure 13** The polygenic risk results in BD group. In the figure, a) the BD polygenic risk based on case vs. control phenotypes in BD group, by threshold $p$ value on the x axis and variance explained by PRS on the y axis. The figure b) showed SZ polygenic risk on case vs. control phenotypes in BD group similar to figure a).

The SZ risk scores were also correlated with mood symptom profile related brain component loadings (Pearson’s $r = -0.33$, $p = 6.21E-05$, Figure 14b).
Figure 14 The correlation results of PRS and projected mood component in BD group. In figure a), the BD polygenic risk scores were negatively correlated with the projected mood profile-related brain component loadings (Pearson’s r = -0.17, p = 0.040). In figure b) the SZ polygenic risk scores were negatively correlated with the same projected mood profile related brain component loadings (Pearson’s r = -0.33, p = 6.21E-05).

4.3.2.2 Correlation analysis between PRS and apathy component

No direct relationship between the PRS and projected apathy brain component loadings was found. However, considering the possibilities of missing information during the data projections, we did an another ICA analysis on the B-SNIP BD participants’ brain independently (25 brain components). We extracted the brain components most similar to apathy component, both by visual checking, and with a maximal GMC correlation (Pearson’s r = 0.20, p < 10E-06). See figure 15.
**Figure 15** Comparing original apathy component and similar brain component generated from ICA in B-SNIP BD brains. The upper brains were the original apathy brain component, and the lower part showed the most similar brain component generated independently in B-SNIP. The GMC values of both brain components were correlated at $r = 0.20, p < 10^{-06}$. The SZ risk scores were found correlated with brain component loadings (Pearson’s $r = -0.31, p = 2.30E-04$). See figure 16.

**Figure 16** The correlation between PRS and apathy-like brain components. PRS was negative correlated with this brain component (Pearson’s $r = -0.31, p = 2.30E-04$).
5 DISCUSSION AND CONCLUSIONS

5.1 Summary

This study first investigated the brain patterns guided by symptom profiles within the SZ and BD framework, and potentially shared mechanism were observed at genetic, brain, and clinical levels. We identified distinct brain patterns correlated with positive symptom, mood, and apathy symptom profiles across SZ and BD. In addition, after projecting across diagnostic groups, these brain patterns were correlated with grandiosity, as well as several remaining positive, and general psychotic symptoms. It indicated these brain pattern were connected to more complicated and broad symptom profiles. Most importantly, both polygenic risk scores based on SZ and BD risk were highly correlated with the brain patterns (loading coefficients in the GMC brain components) identified here. It suggests that the highly shared genomic risks exhibit strong effects on brain circuits alterations behind different psychiatric disorders as SZ and BD.

Through data-driven multivariate analysis such as pICA, we extracted these brain patterns behind typical symptom profiles of psychotic and mood disorders, respectively. Specifically, these highlighted brain regions were spreading across the frontal (sensorimotor regions and inferior/middle frontal gyrus), temporal (inferior/middle/superior temporal gyrus and fusiform gyrus), and parietal (sensorimotor regions and inferior parietal lobule) lobes, which were also the cortical deficits shared by SZ and BD (Hibar et al., 2018; van Erp et al., 2018). Lower GMC in these regions consistently correlated with worse clinical symptoms detected by PANSS (higher PANSS scores) across SZ and BD groups.

To further investigate the relationships behind symptoms profiles and identified brain patterns across disorders, we did the projection analysis to determine the disorder specificity or
generalizability of these findings. We found that the brain components previously identified in each disorder separately showed a more complicated and partially shared clinical significance across psychotic and mood symptoms in the other disorder. The brain pattern identified in SZ showed various connections with grandiosity, hostility, poor impulse control, and active social avoidance in BD. Correspondingly, the brain patterns discovered by mood and apathy symptom profiles in BD, showed associations with grandiosity, excitement, conceptual disorganization, disorientation, and preoccupation when projected into SZ participants. A single symptom was not anchored to a single brain region, and brain regions could play multiple roles in other symptoms (Hare et al., 2018; Walton et al., 2018). Our findings enhanced this understanding across diagnostically.

We tried to explore whether there was a shared genetic mechanism or risks behind the brain patterns detected in the final aim. Based on the most extensive genetic studies from the PGC cohort, we produced the PRS from the alleles carrying SZ risk that best accounted for diagnostic effects in SZ and BD groups. We also calculated the PRS from the alleles carrying BD risk in the same way for both groups. Notably, both sets of SZ and BD risk scores were correlated with brain patterns identified regardless of the original diagnostic groups. Specifically, in the SZ participants, higher genetic risk of SZ and BD accounted for lower GMC (more severe GMC reductions) in the sensorimotor and temporal regions as identified by positive symptom profiles. As in the BD group, same correlations were detected between both disorders’ PRS and mood symptom related brain component within the frontal-temporal-parietal circuits. However, the relationship between PRS and apathy component was subtle. Our hypotheses of a shared genetic mechanism were partially supported.
5.2 Symptom profiles reveal distinct brain patterns in SZ and BD

5.2.1 Positive symptom profile revealed temporal and sensorimotor brain patterns in SZ

The brain structures underlying psychiatric symptoms are not mapped to a single symptom. Networks of regions are related to groups of symptoms (Hare et al., 2017; Modinos et al., 2013; Shinn et al., 2012; Walton et al., 2018). The positive symptoms main PANSS component correlated with GMC in a network mainly located in the inferior temporal and pre- and postcentral regions (sensorimotor areas). The PANSS component differed from the traditional dimensions of the PANSS. It could be broadly described as symptomatology characteristic of a chronic form previously described as paranoid schizophrenia (Tandon et al., 2013). In addition, this PANSS component's loading coefficients correlated strongly with the PANSS general and positive scores, indicating that participants with more severe overall symptoms also showed higher loading coefficients for this component.

As for the structural MRI data, lower GMC often observed in schizophrenia is generally understood as reductions in neuron density (Harrison, 1999; Todtenkopf et al., 2005) and changes in the glia compartment (Fornito et al., 2009). Reduced GMC concentration and cortical thickness often coexist (Narr et al., 2005), though they might reflect different pathophysiological processes (Fornito et al., 2009). Also, these reductions could be caused by antipsychotic treatment (Huhtaniska et al., 2017; Vita et al., 2015). However, our post-hoc analyses did not reveal any significant association between the loading coefficients of structural brain component and CPZ equivalents.

The brain pattern identified here replicated some previous multivariate findings regarding cortical thickness and GM volume in schizophrenia. Padmanabhan et al. (2015) applied factor analyses to structural brains (GM volume, cortical thickness, and surface area) showing
associations with the PANSS positive scale. A temporal factor that included the cortical thickness of the bilateral inferior temporal gyrus and bilateral fusiform gyrus, was identified, and it was a significant predictor of PANSS positive scores (Padmanabhan et al., 2015). We showed similar GMC alterations in inferior temporal gyrus and fusiform gyrus with a different multivariate approach with expanded associations with psychotic symptom profiles. The GMC alterations in these regions have also been associated with a lack of insight (Bergé et al., 2011; Cropley et al., 2016; Ha et al., 2004). Lack of insight was also thought to be correlated with positive symptoms, which might explain these overlapping brain structures (Pousa et al., 2017). Besides, cortical thickness of inferior temporal and fusiform gyrus exhibited the most robust reductions in schizophrenia patients, suggesting that these regions may indeed be strongly associated with the psychotic pathophysiology (van Erp et al., 2018).

Higher GMC in sensorimotor regions also contributed to the identified structural MRI component. Previous studies have revealed inconsistent findings regarding alterations of GMC in schizophrenia. Whereas smaller samples (Ha et al., 2004; Narr et al., 2005) showed higher GMC, studies with larger sample showed lower GMC (Ivleva et al., 2017). Interestingly, none of these studies observed any association with clinical symptoms as we identified, so higher GMC in sensorimotor regions was associated with a specific pattern of the PANSS, which might explain the divergent findings (Ha et al., 2004; Ivleva et al., 2017; Narr et al., 2005).

5.2.2 Mood and apathy symptom profiles revealed frontal, temporal, and parietal brain patterns

In the BD analyses, we found two GMC networks that were related to two different patterns of symptom severity profiles. Importantly, the PANSS profiles highlighted here were not constructed according to the typical preset of PANSS subscales, but the underlying
relationships with GMC patterns drove them. The mood symptom profile was correlated with GMC reductions in the right temporal regions, while in contrast, the apathy/asocial symptom profile was correlated with more widespread network deficits, including frontal, temporal, and parietal regions. It implicated the GMC deficits in the temporal lobe and frontal-temporal-parietal circuits related to clinical profiles as mood and apathy.

The PANSS component for mood mainly included anxiety, guilty feelings, and depression, contrasting with the symptoms of lack of spontaneity and flow of conversation but less strongly. This represented a generally negative mood component, but differing from the original design of the PANSS dimensions. This mood component showed similar affective items highlighted in studies of affective temperaments (Rihmer, Akiskal, Rihmer, & Akiskal, 2010). The temporal regions’ deficits are common in BD and other psychiatric disorder (Abe et al., 2016; Hanford, Nazarov, Hall, & Sassi, 2016; Whalley et al., 2012). However, its relationship with mood symptoms has been elusive, and few studies formed a direct association between mood symptoms and brain structure alterations (Busatto, 2013; Hozer & Houenou, 2016; Padmanabhan et al., 2015; Palaniyappan & Cousins, 2010; Selvaraj et al., 2012). Among psychiatric symptoms, temporal regions (mostly in the superior temporal gyrus) were believed to be critical in thought disorders, auditory verbal hallucinations, and overall positive symptom domains in psychosis (Cavelti et al., 2018; Morch-Johnsen et al., 2018; Morch-Johnsen et al., 2017; Padmanabhan et al., 2015; Strik et al., 2017). However, mood symptoms were usually believed to be marked across several large regions, including frontal, temporal, and limbic regions (Garety & Freeman, 2013; Lebow & Chen, 2016; Ramirez et al., 2015). The structural brain studies emphasized multiple temporal regions’ involvement, especially the superior temporal gyrus (Hanford et al., 2016). Cortical thickness in the right superior frontal and superior
temporal has previously been negatively correlated with symptom severity measured on the Hamilton Depression Rating Scale (Maller, Thaveenthiran, Thomson, McQueen, & Fitzgerald, 2014). In addition, cortical thickness reduction in the middle temporal gyrus was negatively correlated with the global functioning assessment. Despite these significant structure/mood symptom studies, imaging research in BD often reported no correlation between these structural alterations and mood symptom severity (Doan et al., 2017; Elvsashagen et al., 2013; Hanford et al., 2016; Lan et al., 2014; Ratnanather et al., 2014).

Functional studies provided some evidence to support the engagement of temporal gyrus in mood symptoms and regulations (Whalley et al., 2012). An emotional prosody task indicated patients with BD exhibited stronger activations in the bilateral superior temporal gyrus and right inferior frontal gyrus (Mitchell, Elliott, Barry, Cruttenden, & Woodruff, 2004). In addition, in an emotional memory task, the superior temporal gyrus was highlighted with stronger activations (Whalley et al., 2009). Our findings revisited the critical roles of the frontal (inferior frontal gyrus) and temporal (superior/middle temporal gyrus) regions involved in mood symptoms and processes. In summary, compared to most null or inconsistent findings in BD populations discussed above, we exhibited the association between GMC deficits in a specific frontal-temporal brain pattern and mood symptoms through multimodality multivariate analysis. Through aggregating several related items describing mood status in PANSS and localized brain spots that may work together, enhanced brain and clinical phenotypes connection was clearly revealed. It indicated the usefulness of such multimodality multivariate analysis in raising the heterogeneity of future studies. Additionally, we would try to discuss mood symptom profile related brains and its implication to our understanding of BD and disorder boundaries along with apathy in following part.
Apathy is commonly observed in psychotic disorders (Grande, Berk, Birmaher, & Vieta, 2016). The PANSS component for the apathy was strongly weighted on blunted affect, emotional withdrawal, passive/apathetic social withdrawal, depression, and active social avoidance, while negatively contributed by excitement and guilty feeling. In addition to negative symptoms, this component included all the withdrawal/avoidance items in PANSS. It reflected itself as a combination of negative symptoms and behavioral deficits involving avoidance and asociality. In schizophrenia research, a two-factor model categorized anhedonia-asociality-avolition into a general motivation and pleasure (MAP) factor and alogia/blunted affect into an expressive (EXP) factor. However, asociality was not included (Jang et al., 2016). In addition, apathy was another common term used to describe motivational deficits. Moreover, previous studies in the TOP project investigated the relationship between the Apathy Evaluation Scale (AES) and PANSS apathy scores (Faerden et al., 2009; Faerden et al., 2008; Morch-Johnsen et al., 2015). In these studies, apathy scores were formed from emotional withdrawal (N2), passive/apathetic social withdrawal (N4), and lack of spontaneity and flow of conversation (N6) were chosen to compare its capability with AES, while N2 and N4 were chosen according to the definition of apathy by Marin (Marin, Biedrzycki, & Firinciogullari, 1991). It was found that the apathy items chosen from PANSS were significantly correlated with AES.

Through the pICA approach, we found emotional withdrawal, passive/apathetic social withdrawal, blunted affect, and depression included in the apathy symptom profile. Social avoidance can be more complicated. Active social avoidance often refers to the social avoidance caused by unwarranted fear, hostility, or distrust (Hansen, Torgalsboen, Melle, & Bell, 2009). Some studies argued it was weighted on positive symptoms (van der Gaag, Cuijpers, et al., 2006), while others believed it was related to depression or anxiety (Bell, Lysaker, Beam-Goulet,
Milstein, & Lindenmayer, 1994). It was also suggested that social avoidance might not be related to the preset factors of PANSS items (White, Harvey, Opler, & Lindenmayer, 1997). We suggest active social avoidance, regardless of the cause, should be seen as part of asociality grouped into this apathy component. Besides, similar to the mood component, this apathy component was also constructed out of the subscales of PANSS. Through the automatic data reduction process, active social avoidance survived and contributed strongly to the profile. This apathy component described another unique clinical observation of BD patients in the TOP research.

The frontal-parietal-temporal circuit pattern related to the apathy component reflects significant brain deficits previously found in patients with BD (Dusi et al., 2019; Han, De Berardis, Fornaro, & Kim, 2019; Shahab et al., 2018). Connectome research suggested the frontal-parietal circuit relates to the positive and negative symptoms assessed by PANSS, though the structural connectivity could not predict symptom dimensions (Wang et al., 2018). Another informative comparison among SZ and BD was suggested by the B-SNIP (Clementz et al., 2016; Ivleva et al., 2016; Meda et al., 2015). As noted previously from that analysis, more SZ participants were in type 1 while more BD participants were in the other biotypes. The GMC reductions of different types were as following: type 1 (higher proportion of SZ) showed GMC loss across the whole brain, while type 2 (similar proportion of SZ, SAD, and BD) largely overlapped with type 1, and the largest effects were in fronto-temporal circuits, parietal lobes, and cerebellum. Type 3 (higher proportion of BD participants) showed localized GMC reductions in frontal, cingulate, and temporal regions despite the DSM diagnoses (Ivleva et al., 2016). Our findings in BD showed similar but more complicated maps detailed by specific symptom profiles. The brain patterns driven by apathy symptoms showed relative spread deficits across frontal, temporal, and parietal regions as type2 (balanced SZ, SAD, and BD grouped
together). Notably, apathy has been often differentiated from simply depressive mood for its potential psychotic features (Morch-Johnsen et al., 2015). This connection added to the evidence for placing BD between SZ and affective disorder by DSM-5 (American Psychiatric Association, 2013) while emphasizing its characteristics as psychotic disorder. However, the brain patterns driven by mood symptoms showed similar localized brain deficits in frontal and temporal regions as type3 (BD main biotype). Thus, it reflected the evidence of BD being categorized as an affective disorder. Our results stayed in line with the idea BD is between psychotic and affective disorders, and we were able to separate coexisting psychotic and affective brain patterns in another BD population.

It must be pointed out that manic symptoms were the typical positive symptoms expected in BD patients (Dell'osso et al., 2009). These positive symptoms did not emerge through PANSS in this analysis. However, this may be interpreted that these symptoms do not show the strongest relationships with brain structure; or it may be that the range of severity on those symptoms was not large in this sample. Among the three PANSS components in BD group, there was a component highlighted by excitement (Z-score = 1.4) and grandiosity (Z-score = 2.2). It indicated enough variations in positive symptom items for pICA to capture, though its correlation with the brain patterns failed to reach significance ($P = 0.073$). This brain pattern included mainly the inferior/middle frontal gyrus and cerebellum. The frontal gyrus reductions in GM, especially in inferior frontal and prefrontal regions, were reported connected with mania episodic times in follow-up studies (Abe et al., 2016; Abe et al., 2020). However, other studies found heterogeneous cortical deficits or no significant results regarding the above regions and their relationship with manic symptom (Cotovio et al., 2020). Notably, there was a divergence between BD-I and BD-II; in the BD-II participants, the cortical thinning in the prefrontal, frontal,
and temporal regions were found to connect to the depressive episodic times, not hypomania (Elvsashagen et al., 2013; Zak et al., 2019). In summary, our analysis was not able to pick up the manic symptoms, but the heterogeneity of data (especially including BD-I and BD-II at the same time) and complexity between manic symptoms and mapping to structural brains might lead to the insignificance of the findings.

Similarly, we did not get medication or DOI effects on PANSS or brain structures in the highlighted components. Previous studies found medication exposure positively affecting brain structures through various pathways (Dell'Osso, Del Grande, Gesi, Carmassi, & Musetti, 2016; Di Sero et al., 2019; Hartberg et al., 2015; Jorgensen et al., 2016), and large scale studies found medication effects in cortical and subcortical brain regions (Hibar et al., 2018; Hibar et al., 2016). The DOI showed similar findings through the strongest effect found in cortical thickness (Hibar et al., 2018). The pICA approach prioritized the correlation between the PANSS scores and structural brains during the data reduction process, so these components did not necessarily reflect brain regions most significantly affected by medication or DOI.

5.3 The relationship of projected brain patterns and symptoms across disorders

Data projection allows the re-assessments of correlations between previously associated pairs of patterns to be possible across different datasets. As mentioned, SZ and BD have different clinical symptomatology, as reflected by PANSS scores and variations. In addition, being aware that during an independent ICA process, the projected PANSS profiles patterns might not reveal themselves across diagnostically, we further checked the correlations between projected brain loadings and single PANSS item, positive, negative, general, and total scores. Thus, we avoided missing the potential association between projected brain patterns and symptoms profiles.
The projections results from SZ to BD included several exciting findings. Rather than staying connected with the same positive symptom profile across disorders, the brain pattern identified in SZ positive symptom profile showed various connections with grandiosity, hostility, poor impulse control, and active social avoidance after being projected into the BD group. Correspondingly, after being projected, the brain patterns extracted by mood and apathy symptom profiles in the BD, showed associations with grandiosity, excitement, conceptual disorganization, disorientation, and preoccupation. They also showed insignificant correlations (though the same direction) with projected PANSS mood and apathy profiles, which were previously associated.

These results were not totally out of expectations. Three brain patterns extracted by PANSS in SZ and BD fell in the shared brain regions reported previously. As mentioned before, with a large cohort meta-analysis of sMRI studies, the ENIGMA bipolar working group identified consistent and robust thinner frontal, temporal, and parietal cortex (Hibar et al., 2018) in BD patients compared to controls. These differences show overlap with brain structural differences also found in SZ (Rimol et al., 2010; Rimol et al., 2012; van Erp et al., 2018). In a broader view of various brain diseases (including SZ, BD, depression, anxiety, obsessive-compulsive disorder, and addiction), GM losses often converged and became less diagnostic-specific through a longer life span (Goodkind et al., 2015; Kaufmann et al., 2019). Importantly, the identified and projected brain patterns in our work overlapped greatly with these regions including frontal-temporal-parietal circuits and key spots as inferior/middle/superior temporal gyrus.

In the clinical observation, there were same trends of shared cognitive declines behind these common brain alterations in SZ, BD, and other psychiatric disorders (Chang et al., 2018;
Psychomotor speed, executive function, and learning abilities along with intelligence are relatively impaired simultaneously in SZ and BD, and it is also challenging to figure the brain pattern abnormalities that distinguish between SZ and BD (Chang et al., 2018; Knochel et al., 2016; Molina et al., 2011). However, the shared symptomatology with these brain circuits between SZ and BD can be elusive (Nery, Monkul, & Lafer, 2013), and inconsistent findings have been often reported even in BD patients with psychotic symptoms (Ivleva et al., 2012; Yuksel et al., 2012). Importantly, a recent multivariate analysis revealed some potentially supporting findings. Through group pICA, negative symptoms were found to be shared between SZ and depression participants by GMV circuits including caudate, thalamus, middle temporal gyrus, and inferior temporal gyrus (Qi et al., 2020).

In this work, the PANSS negative item and Hamilton Rating Scale for Depression (HRSD) were working cross-diagnostically to reveal the shared brain patterns behind the same symptom feature. Brain components in the inferior/middle temporal gyrus were highlighted again among the different diagnostic groups. In our studies, within the respective disorder group, the psychotic symptom profiles (positive symptom and apathy) related brain patterns (broader coverage of regions spanning sensorimotor cortex and frontal-temporal-parietal circuit) emphasized the psychotic characteristics of SZ and BD. The affective symptom profiles (mood symptoms consist of anxiety, depression, and guilty feelings) related brain regions (relatively localized deficits in frontal-temporal regions) reflected the affective characteristics of BD. Our projection findings, on the other hand, exhibited potentially shared psychotic features behind the frontal-temporal-parietal circuit and frontal-temporal regions among SZ and BD. The projected brains related symptoms, including positive symptoms such as grandiosity, excitement, and hostility, along with others, were strongly involved in common psychotic symptoms in SZ and BD. This
situation indicated the specific brain patterns accounting for disorder-specific symptoms profiles such as delusional positive symptoms, apathy, and mood could distinguish psychotic and affective disorders to some extent. Simultaneously, they may also potentially contributed to broadly shared psychotic features between SZ and BD while the significance was limited to the levels of clinical observations circumstantially.

The reasons why the brain patterns failed to show correlations with original the PANSS patterns after the cross-diagnostic projection should be considered. These include potentially the conservative settings applied to the pICA algorithm, and the logic of the pICA process. The pICA was designed to prioritize the correlation between the modalities and extract the most significant results according to such correlations (Pearlson et al., 2015). To improve the constrain stability, we were quite conservative in the estimated component numbers in the imaging and symptom data (twenty-five brain components and three PANSS components respectively). Also, one pair of components from both modalities in SZ, and three pairs (two passed significance) in BD were prioritized through the iterative pICA process. With above setting, only positive symptoms, mood, and apathy symptom profiles survived the pICA approach. It didn’t cover all the potentially correlated symptom profiles and brain patterns. This choice conservatively reduced the cross-diagnostic representation for the original PANSS patterns from brain components. However, despite this conservative application, we could still detect broad associations between brain patterns and symptom profiles after projection. These findings again indicated the shared brain alterations and clinical relevance behind SZ and BD.

5.4 Genetic effects on symptom profiles driven brain patterns across disorders

In the search for a polygenic mechanism, we were able to find common diagnostic gene risk behind brain patterns extracted by PANSS profiles regardless of the original diagnostic
groups. As presented above, both polygenic risks of SZ and BD were accounting for the brain alterations behind disorders with the comparable significance of the association. Our findings were indicating the highly shared genome-wide risk across psychiatric disorders.

It was worth pointing out, the polygenic risks and brain structural alterations in psychiatric disorders often report inconsistent findings (van der Merwe et al., 2019). The earlier studies in SZ, including around 100 genetic variations, reported no significant relationship with any cortical thickness or white matter measures as fusion images (Voineskos et al., 2016). A similar situation was described in the BD population, while BD was not usually studied alone for the association between polygenic risks and brain structures. Among the limited studies, they often reported insignificant findings (Ranlund et al., 2018). As the increasing genetic variations and sample size were included in the analysis, more encouraging results have been provided. It remains unclear whether the polygenic risk accounts for white matter changes or vast subcortical volume deficits in psychiatric disorders, including SZ and BD (Grama et al., 2020; Reus et al., 2017; Simoes et al., 2020). However, cortical deficits and their association with PRS have been identified robustly. In a combined analysis including 16 cross-sectional data, heterogeneous deficits within cortical thickness, cortical area, cortical volumes, ventricle volumes, and hippocampal subregions were found in SZ vs. controls. Among these regions, higher PRS was associated with thinner frontal and temporal brain cortices along with smaller hippocampal subregions in UK Biobank data (Alnaes et al., 2019). In BD, the situation with cortical anatomical features becomes more complicated. It is not yet clear whether the cortical thickness, GMC, or anatomical measures in BD are stably explained by PRS (Ranlund et al., 2018). It has been reported there was a positive correlation between PRS calculated based on PGC BD results and the longitudinal changes of cortical thickness in medial occipital gyrus, central sulcus, and
posterior cingulate (Abe et al., 2020). However, these regions were not typically highlighted in the frontal-temporal-parietal circuits. A multivariate study showed widespread changes in whole brain GMC might help predict the PRS in five psychiatric disorders, including SZ, BD, and others, but this would need further replication (Ranlund et al., 2018). Among all the above inconsistent findings, it seems polygenic effects on cortical regions are repeatedly reported in SZ but not BD. In addition, the localization of these spots showing the strongest PRS effects are varying from study to study. On the other hand, our results exhibit both SZ and BD polygenic risks consistently affect the cortical abnormalities in frontal, temporal, and sensorimotor regions. Notably, cortical deficits in a larger network as the frontal-temporal-parietal circuit, though extracted in BD, is only associated with SZ polygenic risk. It indicates that SZ polygenic risks show stronger and broader effects across all cortical brain deficits than SZ and BD’s shared effects.

Additionally, we did not find a direct relationship between PRS and the PANSS symptom profiles. PANSS general scores have been previously correlated with PRS, though it was in first-episode SZ rather than a longer duration of illness as in our sample (Sengupta et al., 2017). Generally, the exact map of the relationship between PRS and clinical/behavioral phenotypes is challenging to draw at this moment, especially in cross diagnostic studies (Dima & Breen, 2015; Smeland et al., 2019). In current diagnostic categories, BD-I might exhibit higher SZ RPS in BD-I with manic psychosis than BD-I with depressive psychosis, BD-II, and controls (Markota et al., 2018). In a combined population of SZ, BD, and depression, the polygenic liability for SZ, BD, and depression was associated with the risk of depression, and it was to predict early onset of depression in BD (Musliner et al., 2019). Similar SZ polygenic liability predicted the chances of occurrence and the level of mood-incongruent psychotic symptoms in BD (Allardyce et al.,
Additionally, SZ and BD showed comparable polygenic effects on the memory (Dezhina, Ranlund, Kyriakopoulos, Williams, & Dima, 2019), emotional processing (Wang et al., 2018), spatial visualization (Ranlund et al., 2018), anhedonia (Ward et al., 2019), immune and metabolic serum markers (Maj et al., 2020). More complicated results were suggested, as we mentioned above, regarding the distinguishing of diagnostic features through PRS and clinical phenotypes in SZ and BD. PRS methods revealed loci distinguishing between BD and SZ through psychotic symptoms and the age of onset of the BD participants (Bipolar Disorder Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2018). However, using the PGC data, the PRS provided only modest effects in distinguishing SZ, BD, and unaffected relatives (Calafato et al., 2018). The difficulty in revealing direct relationships between polygenic risks and clinical or behavioral phenotypes might point to the original idea and usefulness of applying imaging genetics. In our analysis, we were still able to capture PANSS associated brain alterations as the intermediate phenotypes, while the direct association between genes and PANSS measures was missing. In addition, despite the PANSS components remaining distinct in SZ and BD, the PRS appeared more useful in exploring the shared features in brains, phenotypes, and clinical observations rather than the disease-specific ones (Calafato et al., 2018; Dima & Breen, 2015; Fullerton & Nurnberger, 2019). In summary, it’s still hard to reveal the direct relationships between polygenic risk scores and clinical phenotypes. We will discuss better ways to draw the map of polygenic effects on brains and behavioral phenotypes in future directions.

5.5 Limitations

There are inevitable limitations within this work. The data used in this work were pooled from multiple datasets. The availability of other data modalities (genes, brains, and behavior)
limited the usable numbers of participants within different aims and might impair comparability among different aims. To extract PANSS profiles related to brain patterns, we were able to use 337 SZ and 109 BD. Though the same number of participants were included in projection analysis, we used quite different participants in genetic analysis. More than 1000 participants were included in polygenic risk calculation in the SZ group to utilize as many participants as possible. However, among the participants with available genetic data, 186 of SZ were finally found with both available genetic and brain imaging data. Thus, a relatively small proportion of the data were included in the final aim.

There were also some limitations in the demographic background of the data. The different studies included more than 15 sites, with a variety of data collection techniques. The BD participants, collected at one site, were diagnosed with bipolar spectrum disorders, including bipolar I and II, and around half of the patients experienced at least one psychotic episode (Nesvag et al., 2017; Wolfers et al., 2018). The history of manic episodes, along with euthymia periods and their potential effects on brain structures and PANSS profiles, could not be taken into consideration. In addition, the proportion of lifetime psychotic episode might differ from the larger BD population overall. We tried to control for age, gender, and site/scanning difference with their potential effects on the brains. Also, we evaluated DOI and medication with their effects on brains and PANSS items. However, with missing data on medications in the SZ group, we could only assess medication effects on participants with available CPZ estimates. In the BD group, considering lithium, antipsychotic, and antiepileptic treatment and their potential effects on brain structures, the medication status was checked as binary variable on the brain structures and PANSS scores. However, the exact dosage and dosage-related effects could not be examined on brains and PANSS scores.
There were chances of overfitting and randomization during the pICA approach. We made an effort to enhance the stability of the results. 10-fold validations were applied to pICA to avoid results being driven by outliers, and pICA itself was running with ICASSO to ensure stability. By introducing a 10mm smoothing kernel (Gupta et al., 2015; Segall et al., 2009; Silver et al., 2011), using a relatively strict Z score, and cluster size to filter the structural brain, the main findings were restricted to the cortical regions. It must be pointed out, in our earlier work applying SBM analysis to GMC (Gupta et al., 2015), we found similar structural brain case vs. control differences as in the current work. However, we did not identify an association between GM concentration and the PANSS positive score when applying post hoc correlation analyses.

During the projection analysis process, as highlighted in the discussion part, one inevitable limitation was that the projection did not ensure the existence of the component in the target data. The other limitation might be that the projection was limited to the shared brain mask and voxels across different datasets (Chen et al., 2019). Though we were able to do it with the full frame of the PANSS items, not a complete set of structural brain patterns identified in SZ and BD could be projected into either of the groups.

There is limited power in the final polygenic risk score analysis. As described in the demographic information, the number of participants with available genetics and images was reduced compared to the full sample size in genetic risk score calculations. Especially in the BD group, with moderate effect size (0.5 for t-test), α level at 0.05, the minimal total sample size to achieve power over 95% (assuming equal sample size in each group) is 210 (105 for each group) (Faul, Erdfelder, Lang, & Buchner, 2007). Our sample size for BD was slightly above this power requirement. However, in the actual practice of imaging genetic studies, a moderate effect size is not common (Hashimoto et al., 2015; Jiang et al., 2019), though a previous power model of
imaging genetics was suggested based on a moderate effect size of 0.5 (Carter et al., 2017). In the situation with these data (setting effect size to be small, effect size = 0.2, α level at 0.05, and power = 0.95), the minimal total sample size would be 1188 (594 subjects for each group), in which case these analyses are definitely underpowered. In addition, we were not able to use the original genetic data from the TOP dataset which had identified the correlated brain patterns and symptom profiles. We had to project the brain patterns into B-SNIP samples, which provided genetic and imaging data. During the projection process, there could be a risk of losing dataset-specific variations and the nonexistence of such components in the target data (Chen et al., 2019). Though we tried an independent ICA to extract a similar apathy component, this subtle relationship added to the difficulty of revealing the genetic mechanism behind apathy.

5.6 Conclusions and future directions

This work indicated shared structural brain alterations in frontal-temporal-parietal circuits behind distinct symptom profiles (positive, mood, and apathy) in the SZ and BD populations through multivariate analysis. Additionally, we found shared polygenic risks of SZ and BD were affecting the structural brain alterations. Despite the differences in symptom profiles, extracted brain patterns fell within the frontal-temporal-parietal network while emphasizing the critical temporal regions, including inferior, middle, and superior temporal gyrus. Importantly, they were explained by compared polygenic risk regardless of the original diagnoses. These results raised essential research questions for future studies.

Firstly, at the clinical observation level, it is critical to include more broad but symptom domain balanced assessment tools. In the current analysis, we used PANSS as it is a valuable tool in detecting the psychiatric symptoms across diagnoses. However, PANSS didn’t cover all the symptoms commonly observed clinically, and the scores for each of symptom domains were
not perfectly balanced. For instance, limited information is covered the single item of depression, and hidden information of depression can also be observed in more subtle behavioral changes. One choice is to include a wide selection of current clinical assessment scales, though it was difficult to keep so many scales consistently rated across different sites. Along with the enlarged inter-rater bias, the different scoring scales, inner structures, and factors in various tools would be hard to control in future data analyses (Pearlson et al., 2016). Imaging genetic and multivariate methods compatible symptom scales are needed to provide clinical profiles covering the whole symptom domains without current diagnostic specificity. In addition, it would need to be easy to use for reducing inter-rater bias, or it can be built as a self-report tool for large data collection. The items would be easy to read without ambiguity, and the different levels of scores would easily be distinguished from each other. The Symptom Checklist-90 (SCL-90) and the revised version might be a potentially useful choice (Derogatis & Cleary, 1977; Jiang et al., 2019; Schmitz et al., 2000). Each question or item in SCL-90 is straightforward for the rater to understand, and it covers the symptom details we would like to collect from the patients.

Secondly, more brain measures than simply gray matter, and explicitly cross-diagnostic analytic tools would be helpful. Interestingly, within the framework of multivariate analysis, novel methods as parallel group ICA and multi-site canonical correlation analysis with reference + joint ICA (MCCAR + jICA) were developed to help achieve these goals. They would be helpful in incorporating multiple modalities in imaging data, e.g. gray matter and white matter together, guided by single scale score or the full input of item scores of a specific scales (Qi et al., 2020; Qi et al., 2019; Sui et al., 2018). These more recent approaches have the benefit of being tolerant to missing data, which hampered the pICA analyses. Regarding the clinical data availability across sites and the different scales of the assessment data, these methods could
avoid the current problem while projecting the data. Due to the small effect sizes of common variations and the difficulty in separating disorder-specific SNPs from the rest of the whole genome, sparse parallel ICA was designed recently to improve genomic data's results (Duan et al., 2020). This could provide another reliable and possibly more sensitive way of exploring genetic effects on the brain and behavioral data.

Another future direction would be constructing PRS based on genetic pathways, more detailed phenotype information, and genetic factors closely connected to structural brains (Toulopoulou et al., 2019). For complicated human diseases, including psychotic disorders, PRS may have independent effects on behavioral, environmental, and clinical risk factors such as lifestyle and stressful life events (Janssens, 2019; Musliner et al., 2015). In measuring the exact effects of PRS on intermediate phenotypes as brain structures, we may need to rule out the above independent effects of PRS (Konigorski, Wang, Cigsar, & Yilmaz, 2018). Constructing PRS based on variations involved in the pathways or gene sets (e.g., abnormal long-term potentiation or electrophysiology) directly affecting brain development, morphometry, and function out of the total PRS is one potential direction (Grama et al., 2020). Another choice is collecting variations involved in certain clinical domains that are directly associated with brains, such as a specific field of cognitive functions or symptom profiles (Toulopoulou et al., 2019). It helps elucidate the functional implication and restrict the genetic effect compared with total PRS effects simultaneously. In current work, we were estimating the PRS based on diagnostic phenotypes, and we were following the idea that specific psychotic disorder risks were the source of modifying the brains. Future work would need to collect brain structures related to genetic risks while stepping out of current diagnoses' categories (Carter et al., 2017). By doing so, stronger effect sizes and future evidence of research based diagnostic criteria could be formed.
REFERENCES


Calhoun, V. D., & Sui, J. (2016). Multimodal fusion of brain imaging data: A key to finding the missing link(s) in complex mental illness. *Biol Psychiatry Cogn Neurosci Neuroimaging, 1*(3), 230-244. doi: 10.1016/j.bpsc.2015.12.005


Dusi, N., De Carlo, V., Delvecchio, G., Bellani, M., Soares, J. C., & Brambilla, P. (2019). MRI features of clinical outcome in bipolar disorder: A selected review: Special Section on "Translational and Neuroscience Studies in Affective Disorders". Section Editor, Maria Nobile MD, PhD. This Section of JAD focuses on the relevance of translational and neuroscience studies in providing a better understanding of the neural basis of affective disorders. The main aim is to briefly summaries relevant research findings in clinical
neuroscience with particular regards to specific innovative topics in mood and anxiety disorders. *J Affect Disord, 243*, 559-563. doi: 10.1016/j.jad.2018.05.066


multi-site brain image data from a clinical investigation of schizophrenia.


doi: 10.1016/j.biopsych.2016.08.030


*Schizophr Res, 11*(1), 47-54.


Nery, F. G., Monkul, E. S., & Lafer, B. (2013). Gray matter abnormalities as brain structural vulnerability factors for bipolar disorder: A review of neuroimaging studies of


Correlation of positive symptoms exclusively to hyperperfusion or hypoperfusion of
cerebral cortex in never-treated schizophrenics. *Lancet, 349*(9067), 1735-1739. doi:
10.1016/S0140-6736(96)08380-8


Group ICA for identifying biomarkers in schizophrenia: 'Adaptive' networks via spatially
constrained ICA show more sensitivity to group differences than spatio-temporal

association study identifies five new schizophrenia loci. *Nat Genet, 43*(10), 969-976. doi:
10.1038/ng.940

10.1038/nature13595

185-193. doi: 10.1023/a:1008931926181

Schwab, S. G., & Wildenauer, D. B. (2013). Genetics of psychiatric disorders in the GWAS era:


