Research paper

Polygenic signal for symptom dimensions and cognitive performance in patients with chronic schizophrenia

Rose Mary Xavier[a,*], Jennifer R. Dungan[b], Richard S.E. Keefe[c], Allison Vorderstrasse[d]

[a] Neuropsychiatry Section, Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, 1034 Gates Pavilion, HUP, 3400 Spruce Street, Philadelphia, PA 19104, United States
[b] Duke University School of Nursing, United States
[c] Psychiatry and Behavioral Sciences, Duke University School of Medicine, United States
[d] NYU Rory Meyers College of Nursing, United States

A B S T R A C T

Genetic etiology of psychopathology symptoms and cognitive performance in schizophrenia is supported by candidate gene and polygenic risk score (PRS) association studies. Such associations are reported to be dependent on several factors - sample characteristics, illness phase, illness severity etc. We aimed to examine if schizophrenia PRS predicted psychopathology symptoms and cognitive performance in patients with chronic schizophrenia. We also examined if schizophrenia associated autosomal loci were associated with specific symptoms or cognitive domains.

Case-only analysis using data from the Clinical Antipsychotics Trials of Intervention Effectiveness-Schizophrenia trials ($n = 730$). PRS was constructed using Psychiatric Genomics Consortium (PGC) leave one out genome wide association analysis as the discovery data set. For candidate region analysis, we selected 105-schizophrenia associated autosomal loci from the PGC study.

We found a significant effect of PRS on positive symptoms at p-threshold (P_T) of 0.5 ($R^2 = 0.007$, $p = 0.029$, empirical $p = 0.029$) and negative symptoms at P_T of 1e-07 ($R^2 = 0.005$, $p = 0.047$, empirical $p = 0.048$). For models that additionally controlled for neurocognition, best fit PRS predicted positive (p-threshold 0.01, $R^2 = 0.007$, $p = 0.013$, empirical $p = 0.167$) and negative symptoms (p-threshold 0.1, $R^2 = 0.012$, $p = 0.004$, empirical $p = 0.329$). No associations were seen for overall neurocognitive and social cognitive performance tests. Post-hoc analyses revealed that PRS predicted working memory and vigilance performance but did not survive correction. No candidate regions that survived multiple testing corrections were associated with either symptoms or cognitive performance. Our findings point to potentially distinct pathogenic mechanisms for schizophrenia symptoms.

1. Introduction

Schizophrenia is a highly complex and disabling disorder with heritability estimates as high as 80% (Sullivan et al., 2003). The genetic risk of schizophrenia established from twin/family studies, and genome wide association studies denote a complex polygenic architecture with several common (108 candidate regions) and rare variants contributing to schizophrenia susceptibility (Rees et al., 2015). Clinically, schizophrenia is heterogeneous with wide variations in symptom presentation. Despite advances in technology that facilitate biological inquiry, pathogenic mechanisms of schizophrenia still remain largely unknown. Heterogeneity at the clinical level and complexity at the molecular level have been cited as strong reasons for this (Rasetti and Weinberger,

2011). In an attempt to address these challenges, some studies have focused on a sub-phenotype approach to examine molecular mechanisms of symptoms (Fanous et al., 2012). This is a validated approach and there is growing evidence for its utility in advancing the field (Jones et al., 2016; Sengupta et al., 2017).

Though there are multiple approaches to study biological and molecular mechanisms of symptoms, two common genetic approaches seen in the literature are: (1) association studies that examine single nucleotide polymorphism's (SNP) effect on symptoms and (2) polygenic risk score (PRS) associations which utilize an aggregate measure of genetic susceptibility to symptoms by taking into account the additive effects of all significant variants across multiple genes and regulatory areas in the entire genome. When SNP associations point to potential

functional pathways, PRS is a taken as evidence for cumulative genetic risk. Since it is an indicator of potentially true genetic risk, examining PRS correlations for symptoms can help inform which symptoms are a result of stronger genetic liability.

Studies that have examined the molecular basis of symptoms have found genetic evidence for specific symptom dimensions in schizophrenia. Evidence from candidate gene association studies identify the possibility of shared (neurotransmitter systems, neuronal development and maintenance) and unique associations to positive and negative symptom dimensions (Xavier and Vorderstrasse, 2017), though functional mechanisms are still unclear. Polygenic scores for schizophrenia have shown associations with symptom dimensions, though with inconsistencies. PRS correlations are reported in the literature for negative/disorganized dimensions not just in schizophrenia patients (Fanous et al., 2012) but also in adolescents in the general population where it predicted negative and anxiety symptoms (Jones et al., 2016). But a recent study had negative findings and the association of PRS with negative symptoms was not replicated in first episode psychosis; instead a polygenic loading was found for general psychopathology dimension and anxiety symptoms (Sengupta et al., 2017). Though in this study a PRS association with negative symptoms was reported in the Caucasian only sub-group (Sengupta et al., 2017).

There is extensive evidence for early onset and persistent cognitive impairments in schizophrenia (Green et al., 2004; Keefe et al., 2005). Such deficits are considered to be a major cause of functional impairment and a quest to understand pathogenic mechanisms of cognitive deficits remains a priority. Though the association between cognitive ability and psychosis is known (Morgan et al., 2014) with evidence supporting a shared genetic basis for cognitive ability and neuropsychiatric illnesses such as schizophrenia (Hagenaars et al., 2016; Hill et al., 2016), the exact nature of this relationship remains obscure (Johnson et al., 2016).

General cognitive ability in the general population has shown to be substantially heritable with large consortium studies identifying multiple candidate associations indicating a polygenic inheritance (Davies et al., 2015). In a recent GWAS meta-analysis of 35,298 healthy individuals, rs76114856 in the *CENPO* gene and rs6669072 on chromosome 1 were found to be associated with cognitive performance (Trampush et al., 2017). In schizophrenia and psychotic illnesses, several candidate genes (ex: *COMT*, *DTNBP1*, *NRG1*, *DISC1*, *ERBB4*) implicating neurotransmitter pathways involving dopaminergic (Green et al., 2014) and glutamatergic signaling (Greenwood et al., 2011, 2013, 2016) have found to be associated with cognitive performance and/or cognitive deficits in patients, though such findings have not produced knowledge that is translatable to clinical practice (Ehrenreich and Nave, 2014).

The genetic etiology of cognition is also supported by PRS studies which associate a higher polygenic burden with greater cognitive decline between the ages of 11 and 70 and a general lower cognitive ability at the age of 70 (McIntosh et al., 2013). In the Philadelphia neurodevelopmental cohort, investigators also found schizophrenia PRS associations for speed of emotion identification, an aspect of social cognition (Germine et al., 2016). Earlier studies utilizing genetic risk scores did not find associations with cognitive domains (Yeo et al., 2014) or IQ in schizophrenia (Van Scheltinga et al., 2013), though the lack of association could be explained by smaller sample sizes.

Data driven approaches have validated the veracity of the polygenic risk score approach (Chen et al., 2017), but PRS results vary across studies and are reported to be dependent on factors such as the sample, stage and/or severity of the disease (Cooke Bailey and Igo, 2016). One study reported PRS to be associated with treatment resistance (Frank et al., 2015), though another study found the evidence to be inconclusive (Martin and Mowry, 2016). PRS has also been reported to be associated with frequent hospitalizations in patients, with a suggested possibility that the association likely comes from the common variants involved in deterioration during the course of the illness (Meier et al.,

2016). But PRS predictions of symptoms and cognitive dimensions have not been specifically examined in patients with chronic schizophrenia.

Our study aimed to examine whether schizophrenia PRS is associated with psychopathology symptoms and cognitive dimensions (neurocognition and social cognition) in a sample of patients with chronic schizophrenia. We also examined schizophrenia associated candidate regions (Ripke et al., 2014) for region specific associations with symptoms and cognitive dimensions in this sample.

2. Method

2.1. Subjects

This study is a secondary analysis of data from 741 subjects with genetic data in the Clinical Antipsychotics Trials of Intervention Effectiveness (CATIE)-schizophrenia trial. CATIE (2001–2004) was a United States based multisite trial designed to assess differences in antipsychotic efficacy. We refer you to Sullivan et al. (2007) for details on the study design, patient selection criteria, sample details and genotyping. For the CATIE study, a total of 1493 patients with chronic schizophrenia were recruited from several sites to participate in the trial and of these patients, 741 provided a DNA sample. The study excluded first episode patients, treatment resistant patients, patients with mood disorders or features and patients whose symptoms were directly a result of substance use or abuse. Study diagnoses for patients were determined by CATIE investigators using the Structured Clinical Interview for DSM-IV (First, 1995). In addition to the consent for the main trial, patients also provided a separate consent for genetic studies (Sullivan et al., 2007).

2.2. Genetic data

Genotyping for CATIE genetic samples were done by Perlegen Sciences (Mountain View, CA, USA) using Affymetrix 500K 'A' and custom 164K chipsets. We obtained the genetic data set from the National Institutes of Mental Health (NIMH) Repository and Genomics Resource (https://www.nimhgenetics.org) after IRB and relevant data use approvals. Data received from the NIMH repository included information on 741 CATIE subjects and 751 controls. Information was available for a total of 495,172 single nucleotide polymorphisms (SNPs) in coordinates mapped in human genome build 17 (hg17, University of California, Santa Cruz). The data we received were pre-processed with some quality control filters applied by both Perlegen and CATIE investigators. SNPs were removed for low quality, minor allele frequency (maf) < 0.01 and hardy-weinberg equilibrium (hwe) < 1e-05. Subjects with > 5% missing data were also removed. We did additional data processing by removing 21 subjects from the total sample (18 controls and 3 cases) with sex discrepancies in phenotypic and genotypic data and 12 subjects (8 cases and 4 controls) for relatedness. After extracting only cases from the sample and performing LiftOver (https://genome.ucsc.edu/cgi-bin/hgLiftOver) to update to a recent human genome build (hg19) we were left with 730 cases and 486,895 SNPs.

2.3. Measures

2.3.1. Psychopathology symptom variables

Psychopathology symptoms were measured by the Positive and Negative Syndrome Scale (PANSS) which is an extensively validated 30 item scale in which each Likert type item is rated from 1 to 7 with higher scores indicating more severe symptoms (Kay et al., 1987). We used the NIMH consensus model to construct our psychopathology symptom variables as a previous study done on CATIE data has shown that this model fit the data the best (Stefanovics et al., 2014). The NIMH consensus model uses 20 out of the 30 items in PANSS and constructs five symptom dimensions – positive, negative, disorganized, excited and depressed. Please see Table 1 for details on these variables.

Table 1
Variables analyzed.

Variable		Description	Measures/tests
Psycho-pathology symptoms	Positive	Sum of PANSS items P1, P3, P5 and G9	PANSS
	Negative	Sum of PANSS items N1, N2, N3, N4, N6, G7	
	Disorganized	Sum of PANSS items P2, N5 and G11	
	Excited	Sum of PANSS items P4, P7, G8, G14	
	Depressed	Sum of PANSS items G2, G3 AND G6	
Cognitive domains	Neurocognition	Composite score of standardized domain scores of:	
		1. Working memory	Computerized test of visuo-spatial working memory
		2. Verbal memory	Hopkins Verbal Learning Test
		3. Vigilance	Continuous Performance Test
		4. Processing speed	−Controlled oral word association test
			−Category Instance
			−Grooved peg board
			Wechsler adult intelligence scale
		5. Reasoning	−Wisconsin card sorting test
			−Wechsler intelligence scale for children
	Social cognition	Total score is number of correct responses on facial emotion recognition task. Rescaled from 0-30, to 15-30 with all scores below 15 assigned a value of 15	Facial Emotion Recognition Task

PANSS - Positive and Negative Syndrome Scale; P_N - Nth item on positive subscale of PANSS; N_N - Nth item on negative subscale of PANSS; G_N - Nth item on general psychopathology subscale of PANSS.

2.3.2. Cognitive variables

We tested two cognitive variables – neurocognition and social cognition (Table 1). The neurocognition variables were the composite scores computed from predefined domains-working memory, verbal memory, processing speed, reasoning and vigilance. Details on the composite score, domain scores, neurocognitive tests used to compute domain scores, and measurement characteristics are provided in Keefe et al. (2006). Social cognition was measured by the facial emotion recognition task (Kerr and Neale, 1993). Since the task measures facial emotion recognition, is only one component of social cognition from here on we refer to it as emotion recognition. Briefly, in this task subjects are asked to discriminate if two faces presented have the same emotional expression. The total score is the number of correctly discriminated faces and ranges between 0 and 30. The data distribution for this variable was skewed with 16 subjects scoring < 15. Subsequently, we rescaled the variable to a lowest score of 15 and all subjects who scored 15 or below were assigned a value of 15.

2.3.3. Genetic variables-PRS

To examine if schizophrenia PRS was predictive of symptom and cognitive variables in chronic schizophrenia we first performed imputation on the genotyped data. Autosomes were imputed using 1000 Genomes Project Phase 3 data as the reference set. Imputation was done in following steps: (1) Applied strict quality filters using thresholds for hwe 1e-06 and removed individuals who had > 2% missing data after which 721 subjects and 421,802 SNPs remained. (2) Pre-phased the filtered data in SHAPEIT software package by first checking for strand alignment issues, removing strand inconsistencies and then estimating haplotypes (Howie et al., 2012). (3) Conducted imputation in chunks of 5 million bases (Mb) using IMPUTE2 software package (Marchini and Howie, 2010). (4) Applied post imputation quality control filters (the same filters used in the Psychiatric Genomics Consortium analysis: maf 0.1 and info score cutoff of 0.9) to select the list of markers (Ripke et al., 2014).

We constructed PRS using the methods suggested by the International Schizophrenia Consortium (Purcell et al., 2009) using the 'fastscore' option in PRSice software package (Euesden et al., 2015). To curate a most informative SNP list for PRS calculations the SNPs are first filtered to address linkage disequilibrium (LD). We used the clump option to address LD by applying thresholds of $r^2 \geq 0.1$ and distance of 500 kb. Since SNPs in the major histocompatibility complex region of the genome have long range LD, we also removed SNPs in the MHC region (26–33 Mb) prior to constructing PRS. PRS was calculated on an a priori set of significance thresholds (P_T = 5e-08, 1e-07, 1e-06, 1e-05, 1e-04, 1e-03, 0.01, 0.05, 0.1,0.2,0.3,0.4 and 0.5) to identify the best fit PRS that was most predictive of an association. Multiple comparisons were addressed by applying 10^6 permutations.

2.3.4. SNPs for set based analysis

Results of GWAS for complex traits suggest that the disease or trait associations are enriched in genomic regions with several risk variants in a single locus and as such, a set based association test assessing the effect of an aggregated set of SNPs has been suggested to be more powerful than single SNP analyses (Bakshi et al., 2016). Hence, to address our second aim we chose a set based analysis. We first extracted SNP sets by using the genomic locations reported for the validated 108 schizophrenia associated loci (Ripke et al., 2014). Excluding three regions in the X-chromosome we had a total of 105 sets.

2.4. Statistical analyses

2.4.1. Polygenic risk scores

Analyses were run in PRSice using PRS as a continuous measure. We tested associations between PRS, symptoms and cognitive variables using linear regression models. We tested two models for symptom variables: (1) the first model controlled for the effects of base covariates (age, sex and ancestry using four multidimensional scaling components). (2) Due to statistically significant correlations observed between neurocognitive variables and symptom dimensions, we also tested a second model adjusting for neurocognition in addition to base covariates. For cognitive variables, we tested one model controlling for the effects of base covariates. Results are presented as change in variance (R^2 of model minus R^2 of the reduced covariates only model) and overall significance of the model. We report both unadjusted p-values as well as permutation based p-values addressing multiple comparisons issue to identify best fit PRS.

2.4.2. Set based analysis

This analysis was implemented in Plink2 (https://www.cog-genomics.org/plink2/general_usage) (Chang et al., 2015) using the set option. A set based analysis follows a series of steps: (1) Single SNP analysis is run for all SNPs within the sets. (2) Within each set, a specified number of SNP(s) that are below a specified significance threshold are selected based on a set LD criterion. (3) The set is then

permuted using the phenotype status to obtain an empirical *p*-value adjusted to address multiple comparisons within each set (Purcell et al., 2007). We set the SNP selection significance threshold at 0.05, LD criterion at $r^2 > 0.8$, used a top SNP approach selecting only the top hit SNP within each set and permuted the data set 10,000 times. Since we had 105 sets in the analysis, we applied false discovery rate (FDR, Benjamini-Hochberg procedure with an alpha threshold of 0.05) correction. We tested two models for association analyses with symptom variables, the first model adjusting for base covariates and a second model adjusting for neurocognition in addition to base covariates. We tested cognitive variables adjusting for just the base covariates.

3. Results

3.1. PRS associations

For model 1 which controlled for age, sex and ancestry, we found a significant effect of PRS on positive symptoms at P_T of 0.5 ($R^2 = 0.007$, $p = 0.029$, empirical $p = 0.029$) and a significant effect on negative symptoms at P_T of 1e-07 ($R^2 = 0.005$, $p = 0.047$, empirical $p = 0.048$), though the effect on negative symptoms was not in the predicted direction ($\beta = -180.5$). Please refer to S1 Table 1 for full results. Fig. 1 shows model 1 bar plots for PRS predictions of positive symptoms and negative symptoms.

Model 2 which additionally controlled for neurocognition showed findings in a different direction. We found a significant effect of PRS on positive and negative symptoms at higher SNP selection thresholds. Best fit PRS at P_T of 0.01 predicted positive symptoms ($R^2 = 0.007$, $p = 0.013$, empirical $p = 0.167$) and a significant effect on negative symptoms at P_T of 0.1 ($R^2 = 0.012$, $p = 0.004$, empirical $p = 0.329$) and did not survive permutation correction. Please refer to S2 Table 2 for full results. Fig. 2 shows model 2 bar plots for PRS predictions of positive symptoms and negative symptoms.

There were no significant effects of PRS on disorganized, depressed or excited symptom dimensions. We did not find evidence for PRS predictions of the neurocognitive composite measure or emotion recognition. Post hoc analyses of neurocognitive domains revealed a significant association of PRS with working memory ($P_T = 1e-05$, $R^2 = 0.006$, $p = 0.035$, empirical $p = 0.15$) and vigilance ($P_T = 1e-04$, $R^2 = 0.008$, $p = 0.020$, empirical $p = 0.09$) (Fig. 3), but they did not survive permutation correction.

3.2. Set based associations

Results from our set based analysis are presented in Tables 2 and 3. Though we had a few sets with variants that were significantly associated with symptom dimensions, no loci survived correction for multiple set comparisons. Similarly, no variants associated with neurocognition or emotion cognition survived correction. Results from our post hoc analyses of set based associations with neurocognitive subdomains are presented in supplementary Table 3. We found that after multiple set comparison correction was applied, variant rs12706998, an intergenic variant, was significantly associated with working memory (FDR $p = 0.031$) and variant rs17269617, also an intergenic variant, (FDR $p = 0.020$) was associated with verbal memory.

4. Discussion

The current study aimed to examine if a schizophrenia PRS was associated with psychopathology symptom and cognitive dimensions in addition to potential candidate associations in a sample of patients with chronic schizophrenia. The biological basis of symptom dimensions and performance in cognitive domains (Glahn et al., 2014) are deemed to be less complex than the schizophrenia phenotype itself and as such we evaluated for overlap with schizophrenia genetic liability. In our chronic sample, we found a significant PRS association with negative
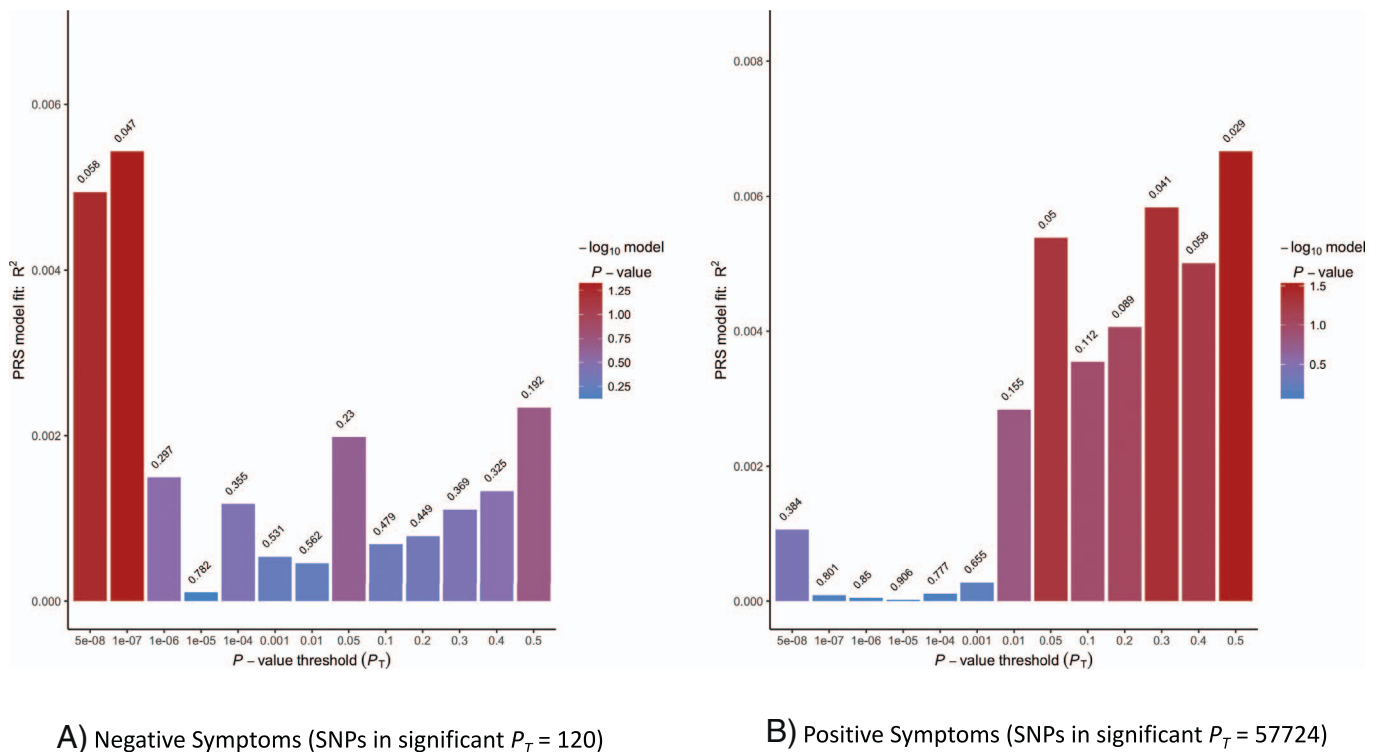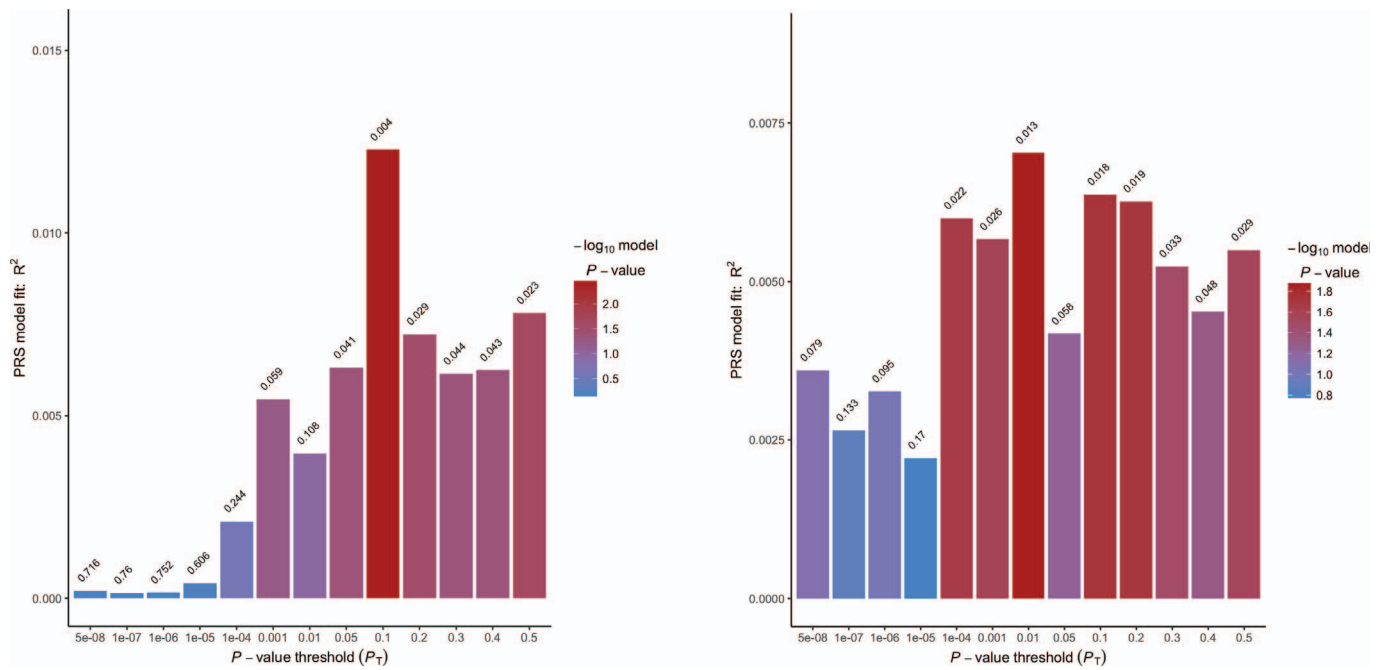

A) Negative Symptoms (SNPs in significant P_T = 120)

B) Positive Symptoms (SNPs in significant P_T = 57724)

Fig. 1. Model 1 fit for polygenic risk score predictions on positive and negative symptom dimensions. The plots show model 1 results of 13 analyses based on SNP set without linkage disequilibrium. For each P_T, SNPs are selected if significant at that threshold and coefficients and effect sizes estimated. Values above each bar are unadjusted *p*-values of phenotype from regression analyses. For 1A, the best fit PRS for negative symptoms is at the P_T of 1e-07 and explains roughly 0.5% of the variance. For 1B, the best fit PRS for positive symptoms is at P_T of 0.5 and roughly explains 0.7% of the variance.

A) Negative Symptoms (SNPs in significant P_T = 23929)

B) Positive Symptoms (SNPs in significant P_T = 6834)

Fig. 2. Model 2 fit for polygenic risk score predictions on positive and negative symptom dimensions. The plots show model 2 results of 13 analyses based on SNP set without linkage disequilibrium. For each P_T, SNPs are selected if significant at that threshold and coefficients and effect sizes estimated. Values above each bar are unadjusted *p*-values of phenotype from regression analyses. For 2A, the best fit PRS for negative symptoms is at P_T of 0.1 and explains roughly 1.2% of the variance. For 2B, the best fit PRS for positive symptoms is at P_T of 0.01 and roughly explains 0.7% of the variance.

A) Working Memory (Significant SNPs in P_T = 423)

B) Vigilance (Significant SNPs in P_T = 2611)
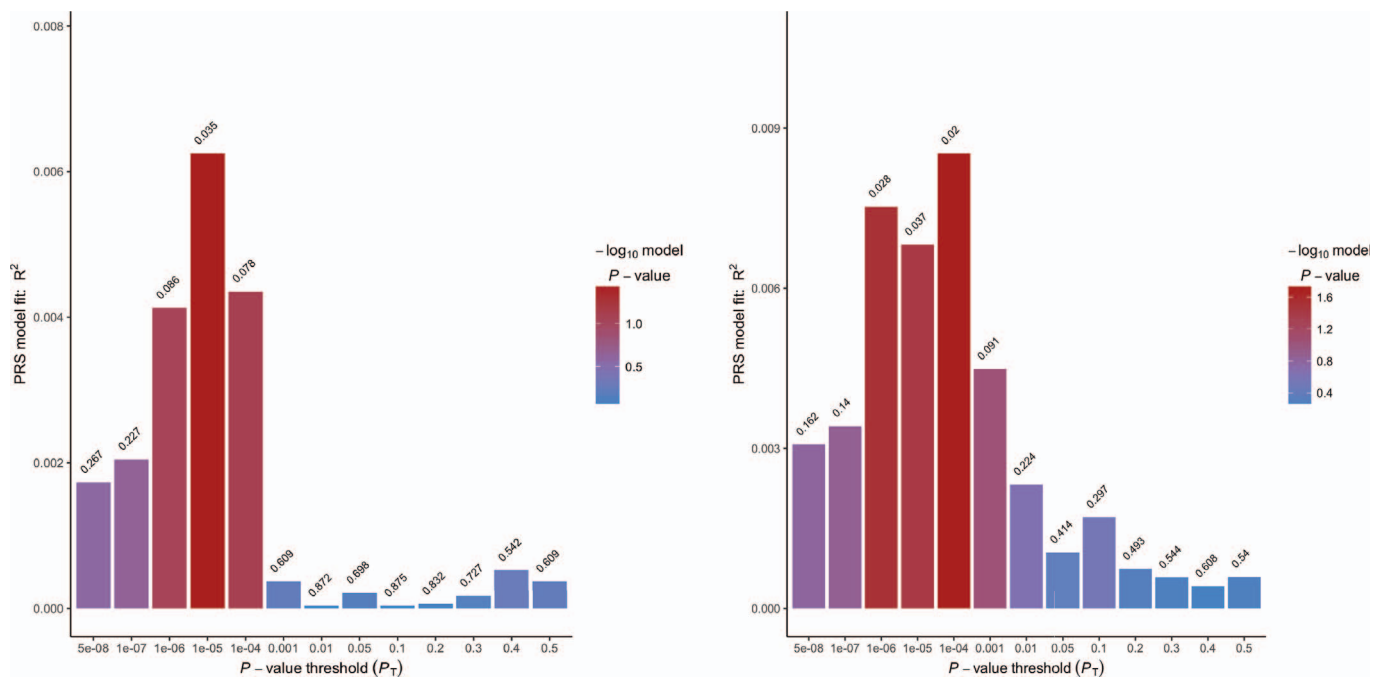
Fig. 3. Model fit for polygenic risk score predictions of neurocognitive domains (working memory and vigilance) from post hoc analyses. The plots show the results of 13 analyses per variable based on SNP set without linkage disequilibrium. For each P_T, SNPs are selected if significant at that threshold and coefficients and effect sizes estimated. Values above each bar are unadjusted *p*-values of phenotype from regression analyses. For 3A, the best fit PRS for working memory is at P_T of 1e-05 and explains roughly 0.6% of the variance. For 3B, the best fit PRS for vigilance is at P_T of 1e-04 and roughly explains 0.8% of the variance.

Table 2

Associations (p < 0.05 with 10,000 permutations within each set) of quantitative psychopathology symptom phenotypes.

Symptom phenotypes	Loci	Top variant in set	p-Val in set	Beta	Gene	Relevant biological function[a]
Positive	chr7:137039644–137085244	rs16874961	0.0028	−0.94	DGKI (Diacylglycerol kinase iota)	Intracellular signal transduction
	chr1:30412551–30437271	rs1009080[b]	0.0158	0.76	[c]	
	chr1:177247821–177300821	rs3176443	0.0241	0.71	FAM5B (BMP/retinoic acid inducible neural specific protein 2)	Positive regulation of neuronal differentiation, protein forms the cell component of neuronal cell body
Negative	chr11:133808069–133852969	rs3758927	0.003	1.75	IGSF9B (protein turtle homolog B)	G-protein coupled receptor, cell component of postsynaptic membrane, dendrite, neuronal cell body.
	chr11:113317794–113423994	rs7115090	0.0111	−1.05	[c]	
	chr18:53453389–53804154	rs1960054	0.0228	1.54	[c]	
	chr7:110034393–110106693	rs211814	0.0243	4.26	[c]	
	chr2:185601420–185785420	rs16826183	0.0274	−2.44	ZNF804A (zinc finger protein 804A)	Implicated in several neurocognitive processes
	chr1:149998890–150242490	rs11205334	0.0281	−3.42	PLEKHO1 (Plecksrin homolog domain containing family O member 1)	NA
Dis-organized	chr1:97792625–98559084	rs10783063	0.0284	1.44	DPYD (Dihydropyrimidine dehydrogenase)	NA
	chr15:84661161–85153461	rs171158441	0.0312	−1.41	ADAMTSL3 (ADAMTS like protein 3)	NA
	chr2:146416922–146441832	rs13006403[b]	0.0413	0.87		
	chr7:104598064–105063064	rs10249746	0.01	0.67	[c]	
	chr7:86403226–86459326	rs2214011[b]	0.024	−1.05	GRM3(Metabotropic glutamate receptor 3)	G-protein coupled receptor signaling pathway, ionotropic glutamate pathway, neuro system process, neuron-neuron synaptic transmission
Excited	chr6:73132701–73171901	rs2815712[b]	0.034	−0.51	[c]	
	chr3:52541105–52903405	rs1573815[b]	0.0495	−0.56	TMEM110(Transmembrane protein 110 -MUSTN1 related)	Calcium channel regulator activity
	chr6:84279922–84407274	rs16871626	0.0092	1.54	SNAP91 (Clathrin coat assembly protein AP180)	Intracellular protein transport, receptor mediated endocytosis, chemical synaptic transmission
Depressed	chr17:2095899–2220799	rs749240[b]	0.0473	−0.42	SMG6 (Telomerase binding protein EST1A)	Telomere maintenance, RNA stability
	chr11:123394636–123395986	rs49366823	0.0477	0.038	GRAMD1B (Gram domain containing protein 1)	NA
	chr19:50067499–50135399	rs2288920	0.0052	0.61	PRR12 (proline rich protein 12)	NA
	chr2:193848340–194028340	rs17662626	0.0212	−1.08	[c]	
	chr16:58669293–58682833	rs9922575[b]	0.0459	−0.41	[c]	

NA-information not available.

[a] Source: PANTHER database (http://pantherdb.org).

[b] Variant non-significant when controlled for neurocognition.

[c] Intergenic variants.

Table 3
Associations (p < 0.05 with 10,000 permutations within each set) of cognitive phenotypes.

Cognitive phenotypes	Loci	Top variant in set	p-Val in set	Beta	Gene	Relevant biological function[a]
Neuro-cognition	chr8:60475469–60954469	rs172279957	0.0059	−0.47	PLCB2 (1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase beta)[b]	G-protein coupled receptor binding signaling pathway, intracellular signal transduction;
	chr15:405667599–40602237	rs3784397	0.021	0.15		Thyrotropin, oxytocin, WNT signaling, alpha adrenergic, 5HT2 receptor signaling etc.
	chr2:146416922–146441832	rs10496985	0.0258	−0.26		
	chr1:149998890–150242490	rs11205334	0.0309	−0.55	PLEKHO1[b]	NA
Social cognition	chr8:27412627–27453627	rs17466684	0.0065	−0.82	CLU (Clusterin)	CCKR signaling map
	chr12:2321860–2523731	rs11062160	0.0081	−1.22	CACNA1C (Voltage dependent L type ca channel subunit alpha 1C)	5HT2, Alzheimer's disease amyloid secretase, Nicotinic acetyl choline receptor, oxytocin receptor signaling pathways
	chr19:50067499–50135399	rs2288920	0.0319	0.48	PRRG2 (Transmembrane gamma-carboxyglutamic acid protein 2)[b]	Immune system process, proteolysis, response to stimulus
	chr1:149998890–150242490	rs12073359	0.0498	0.66		

NA-information not available.
[a] Source: PANTHER database (http://pantherdb.org).
[b] Intergenic variants.

symptoms for both models (with and without neurocognition). Model that controlled for the effects of neurocognition explained a significant amount of variance (1.2%) in negative symptoms, though this did not survive permutation testing and could likely be explained by our smaller sample size. PRS associations with negative symptoms have been somewhat inconsistent in the literature with both positive (Jones et al., 2016; Fanous et al., 2012) and negative findings (Sengupta et al., 2017). Such inconsistencies could be a reflection of the heterogeneity of the sample or the sample size itself. PRS prediction accuracies are found to be strongly linked to sample size (Dudbridge, 2013).

As with negative symptoms, we also found evidence for PRS predicting positive symptoms in the two models tested. Though the amount of variance in positive symptoms (0.7%) explained by both models were the same (potentially suggesting a negligible effect of neurocognition on positive symptoms), model 2 controlling for neurocognition did not survive permutation correction and again could likely be explained by our smaller sample size. PRS association of positive symptoms is contradictory to previous studies that did not show a polygenic loading on positive symptoms (Fanous et al., 2012). This could likely be explained by the differences in sample characteristics. We had a less heterogeneous group (only chronic patients) in comparison to previous studies. It's been suggested that positive symptoms that emerge in adolescence may be strongly influenced by environmental factors such as childhood trauma and/or cannabis use, hence a lack of association in adolescence or first episode patients; PRS associations with positive symptoms may therefore emerge in patients with a later onset of such symptoms (Jones et al., 2016). Our findings could be suggestive of this difference in the pathogenicity and require further study.

We did not find an association with disorganized symptom dimension. The study that previously associated PRS with disorganized symptoms had a different factor structure which clumped negative and disorganized symptoms into a single factor (Fanous et al., 2012). We had separate factor structures for these symptom dimensions which is likely the reason for the absence of association.

Despite the polygenic signal evident in our study we did not find any SNP associations with symptom dimensions that withstood correction which could be due to our sample size. Large sample sizes are required to identify effects of common variants (Ripke et al., 2011). It is very likely that we were underpowered to detect an effect.

We did not find evidence for a strong polygenic association for neurocognition as a single composite factor. This could be due to different genetic susceptibility for different sub-domains. We confirm this in our post hoc analyses where we found that two subdomains (working memory and vigilance) of neurocognition were predicted by PRS, though they did not survive correction for multiple tests. Earlier studies have found PRS associations with working memory in the healthy general population (Hatzimanolis et al., 2015). In addition, PRS has been associated not just with behavioral measures of cognition but also brain based endophenotypes such as prefrontal inefficiency (Walton et al., 2014). Working memory has been strongly associated with prefrontal function (Perlstein et al., 2001) and this offers potential causal mechanisms for further investigation. In our study we also found PRS associations with vigilance measured by the continuous performance test. Vigilance measured by the same test has been associated with schizophrenia PRS in a healthy population (Hatzimanolis et al., 2015).

Literature also supports a polygenic basis for social cognition. Social cognition measured by emotion recognition was found to be associated with PRS in subjects between the ages of 8–21 with findings replicated in a sample of healthy adults (Germine et al., 2016). Of importance here, is that PRS was specifically associated with the speed of emotion recognition and not the ability to recognize emotion per se. Our study did not support a polygenic association and is likely due to measurement differences. Our social cognitive variable assessed for the ability to discriminate emotion and did not measure the speed, which could explain the lack of association.

Taken together, our results are suggestive of a polygenic association

for specific symptom and cognitive dimensions in chronic schizophrenia. This study has several limitations; the obvious one is the sample size which is small in comparison to the recommended *n* of 2000 (Wray et al., 2014). General limitations of the PRS analysis apply to our study-PRS identifies a polygenic signal but cannot provide specific associations that can help elucidate functional mechanisms. PRS is computed from common variants with small effects, so the contributions of rare and structural variants are not known. Additionally, in our study we only used autosomes for PRS calculations, so effects of common variants in sex chromosomes contributing to polygenic burden are not known. For our set based analysis, we utilized a top SNP approach for the loci analyzed which is a limitation as this could potentially exclude several SNPs that could be causal. Some of these loci are known to have several genes (Ripke et al., 2014).

In conclusion, in a sample of patients with chronic schizophrenia we examined if genetic susceptibility to schizophrenia could predict symptom dimensions and cognitive performance. We did not find specific variants associated with symptom or cognitive dimensions. However, we found that PRS predicted a small amount of variance in positive and negative symptom dimensions and performance on working memory and vigilance tests. Some of our findings are consistent with the literature and extend these associations in the population with chronic schizophrenia. Given that these symptoms and deficits in cognitive domains contribute heavily to morbidity in schizophrenia, these findings hold promise to inform further research and potentially guide an objective approach to diagnosis and development of precise therapeutic interventions.

Conflicts of interest

RX, JD and AV have no potential or relevant conflicts of interest to report.

RK is the founder and owner of NeuroCog Trials Inc. No other potential or relevant conflicts of interest to disclose.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scog.2018.01.001.

References

Bakshi, A., Zhu, Z., Vinkhuyzen, A.A., Hill, W.D., McRae, A.F., Visscher, P.M., Yang, J., 2016. Fast set-based association analysis using summary data from GWAS identifies novel gene loci for human complex traits. Sci. Rep. 6, 32894.

Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., Lee, J.J., 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 4, 7.

Chen, J., Calhoun, V.D., Pearlson, G.D., Perrone-Bizzozero, N.I., Turner, J.A., Ehrlich, S., Ho, B.C., Liu, J., 2017. Independent component analysis of SNPs reflects polygenic risk scores for schizophrenia. Schizophr. Res. 181, 83–85.

Cooke Bailey, J.N., Igo, R.P., 2016. Genetic risk scores. Curr. Protoc. Hum. Genet. 1.29 (1-1.29), 9.

Davies, G., Armstrong, N., Bis, J.C., Bressler, J., Chouraki, V., Giddaluru, S., Hofer, E., Ibrahim-Verbaas, C.A., Kirin, M., Lahti, J., Van Der Lee, S.J., 2015. Genetic contributions to variation in general cognitive function: a meta-analysis of genome-wide association studies in the CHARGE consortium (N = 53,949). Mol. Psychiatry 20 (2), 183–192.

Dudbridge, F., 2013. Power and predictive accuracy of polygenic risk scores. PLoS Genet. 9, e1003348.

Ehrenreich, H., Nave, K.A., 2014. Phenotype-based genetic association studies (PGAS) - towards understanding the contribution of common genetic variants to schizophrenia subphenotypes. Genes 5, 97–105.

Euesden, J., Lewis, C.M., O'reilly, P.F., 2015. PRSice: Polygenic Risk Score Software v1. pp. 21.

Fanous, A.H., Zhou, B., Aggen, S.H., Bergen, S.E., Amdur, R.L., Duan, J., Sanders, A.R., Shi, J., Mowry, B.J., Olincy, A., 2012. Genome-wide association study of clinical dimensions of schizophrenia: polygenic effect on disorganized symptoms. Am. J. Psychiatr. 169, 1309–1317.

First, M.B., 1995. Structured Clinical Interview for the DSM (SCID). Wiley Online Library.

Frank, J., Lang, M., Witt, S., Strohmaier, J., Rujescu, D., Cichon, S., Degenhardt, F., Nöthen, M., Collier, D., Ripke, S., 2015. Identification of increased genetic risk scores for schizophrenia in treatment-resistant patients. Mol. Psychiatry 20, 150.

Germine, L., Robinson, E., Smoller, J., Calkins, M., Moore, T., Hakonarson, H., Daly, M., Lee, P., Holmes, A., Buckner, R., 2016. Association between polygenic risk for schizophrenia, neurocognition and social cognition across development. Transl. Psychiatry 6, e924.

Glahn, D.C., Knowles, E.E., Mckay, D.R., Sprooten, E., Raventós, H., Blangero, J., Gottesman, I.I., Almasy, L., 2014. Arguments for the sake of endophenotypes: examining common misconceptions about the use of endophenotypes in psychiatric genetics. Am. J. Med. Genet. B Neuropsychiatr. Genet. 165, 122–130.

Green, M.F., Kern, R.S., Heaton, R.K., 2004. Longitudinal studies of cognition and functional outcome in schizophrenia: implications for MATRICS. Schizophr. Res. 72, 41–51.

Green, M.J., Chia, T.Y., Cairns, M.J., Wu, J., Tooney, P.A., Scott, R.J., Carr, V.J., 2014. Catechol-*O*-methyltransferase (COMT) genotype moderates the effects of childhood trauma on cognition and symptoms in schizophrenia. J. Psychiatr. Res. 49, 43–50.

Greenwood, T.A., Lazzeroni, L.C., Calkins, M.E., Freedman, R., Green, M.F., Gur, R.E., Gur, R.C., Light, G.A., Nuechterlein, K.H., Olincy, A., Radant, A.D., 2016. Genetic assessment of additional endophenotypes from the Consortium on the Genetics of Schizophrenia Family Study. Schizophr. Res. 170 (1), 30–40.

Greenwood, T.A., Lazzeroni, L.C., Murray, S.S., Cadenhead, K.S., Calkins, M.E., Dobie, D.J., Green, M.F., Gur, R.E., Gur, R.C., Hardiman, G., Kelsoe, J.R., 2011. Analysis of 94 candidate genes and 12 endophenotypes for schizophrenia from the Consortium on the Genetics of Schizophrenia. Am. J. Psychiatr. 168 (9), 930–946.

Greenwood, T.A., Swerdlow, N.R., Gur, R.E., Cadenhead, K.S., Calkins, M.E., Dobie, D.J., Freedman, R., Green, M.F., Gur, R.C., Lazzeroni, L.C., Nuechterlein, K.H., 2013. Genome-wide linkage analyses of 12 endophenotypes for schizophrenia from the Consortium on the Genetics of Schizophrenia. Am. J. Psychiatr. 170 (5), 521–532.

Hagenaars, S.P., Harris, S.E., Davies, G., Hill, W.D., Liewald, D.C., Ritchie, S.J., Marioni, R.E., Fawns-Ritchie, C., Cullen, B., Malik, R., Worrall, B.B., 2016. Shared genetic aetiology between cognitive functions and physical and mental health in UK Biobank (N = 112,151) and 24 GWAS consortia. Mol. Psychiatry 21 (11), 1624.

Hatzimanolis, A., Bhatnagar, P., Moes, A., Wang, R., Roussos, P., Bitsios, P., Stefanis, C.N., Pulver, A.E., Arking, D.E., Smyrnis, N., Stefanis, N.C., Avramopoulos, D., 2015. Common genetic variation and schizophrenia polygenic risk influence neurocognitive performance in young adulthood. Am. J. Med. Genet. B Neuropsychiatr. Genet. 168, 392–401.

Hill, W.D., Davies, G., Harris, S.E., Hagenaars, S.P., Davies, G., Deary, I.J., Debette, S., Verbaas, C.I., Bressler, J., Schuur, M., Smith, A.V., 2016. Molecular genetic aetiology of general cognitive function is enriched in evolutionarily conserved regions. Transl. Psychiatry 6 (12), e980.

Howie, B., Fuchsberger, C., Stephens, M., Marchini, J., Abecasis, G.R., 2012. Fast and accurate genotype imputation in genome-wide association studies through prephasing. Nat. Genet. 44, 955–959.

Johnson, M.R., Shkura, K., Langley, S.R., Delahaye-Duriez, A., Srivastava, P., Hill, W.D., Rackham, O.J., Davies, G., Harris, S.E., Moreno-Moral, A., Rotival, M., 2016. Systems genetics identifies a convergent gene network for cognition and neurodevelopmental disease. Nat. Neurosci. 19 (2), 223–232.

Jones, H.J., Stergiakouli, E., Tansey, K.E., Hubbard, L., Heron, J., Cannon, M., Holmans, P., Lewis, G., Linden, D.E.J., Jones, P.B., Smith, G.D., O'donovan, M.C., Owen, M.J., Walters, J.T., Zammit, S., Davey Smith, G., 2016. Phenotypic manifestation of genetic risk for schizophrenia during adolescence in the general population. JAMA Psychiat. 73 (221–228), 8p.

Kay, S.R., Fiszbein, A., Opler, L.A., 1987. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr. Bull. 13, 261–276.

Keefe, R.S., Bilder, R.M., Harvey, P.D., Davis, S.M., Palmer, B.W., Gold, J.M., Meltzer, H.Y., Green, M.F., Miller, D.D., Canive, J.M., Adler, L.W., 2006. Baseline neurocognitive deficits in the CATIE schizophrenia trial. Neuropsychopharmacology 31 (9), 2033–2046.

Keefe, R.S., Eesley, C.E., Poe, M.P., 2005. Defining a cognitive function decrement in schizophrenia. Biol. Psychiatry 57, 688–691.

Kerr, S.L., Neale, J.M., 1993. Emotion perception in schizophrenia: specific deficit or further evidence of generalized poor performance? J. Abnorm. Psychol. 102, 312.

Marchini, J., Howie, B., 2010. Genotype imputation for genome-wide association studies. Nat. Rev. Genet. 11, 499–511.

Martin, A.K., Mowry, B., 2016. Increased rare duplication burden genomewide in patients with treatment-resistant schizophrenia. Psychol. Med. 46 (3), 469–476.

McIntosh, A.M., Gow, A., Luciano, M., Davies, G., Liewald, D.C., Harris, S.E., Corley, J., Hall, J., Starr, J.M., Porteous, D.J., 2013. Polygenic risk for schizophrenia is associated with cognitive change between childhood and old age. Biol. Psychiatry 73, 938–943.

Meier, S.M., Agerbo, E., Maier, R., Pedersen, C.B., Lang, M., Grove, J., Hollegaard, M.V., Demontis, D., Trabjerg, B.B., Hjorthøj, C., Ripke, S., 2016. High loading of polygenic risk in cases with chronic schizophrenia. Mol. Psychiatry 21 (7), 969–974.

Morgan, V.A., McGrath, J.J., Jablensky, A., Badcock, J.C., Waterreus, A., Bush, R., Carr, V., Castle, D., Cohen, M., Galletly, C., Harvey, C., 2014. Psychosis prevalence and physical, metabolic and cognitive co-morbidity: data from the second Australian national survey of psychosis. Psychol. Med. 44 (10), 2163–2176.

Perlstein, W.M., Carter, C.S., Noll, D.C., Cohen, J.D., 2001. Relation of prefrontal cortex dysfunction to working memory and symptoms in schizophrenia. Am. J. Psychiatr. 158, 1105–1113.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., De Bakker, P.I., Daly, M.J., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575.

Purcell, S.M., Wray, N.R., Stone, J.L., Visscher, P.M., O'donovan, M.C., Sullivan, P.F., Sklar, P., Ruderfer, D.M., Mcquillin, A., Morris, D.W., 2009. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature 460, 748–752.

Rasetti, R., Weinberger, D.R., 2011. Intermediate phenotypes in psychiatric disorders. Curr. Opin. Genet. Dev. 21, 340–348.

Rees, E., O'donovan, M.C., Owen, M.J., 2015. Genetics of schizophrenia. Current Opinion in Behavioral Sciences 2, 8–14.

Ripke, S., Neale, B.M., Corvin, A., Walters, J.T., Farh, K.H., Holmans, P.A., Lee, P., Bulik-Sullivan, B., Collier, D.A., Huang, H., Pers, T.H., 2014. Biological insights from 108 schizophrenia-associated genetic loci. Nature 511 (7510), 421.

Ripke, S., Sanders, A.R., Kendler, K.S., Levinson, D.F., Sklar, P., Holmans, P.A., Lin, D.Y., Duan, J., Ophoff, R.A., Andreassen, O.A., Scolnick, E., 2011. Genome-wide association study identifies five new schizophrenia loci. Nat. Genet. 43 (10), 969.

Sengupta, S.M., MacDonald, K., Fathalli, F., Yim, A., Lepage, M., Iyer, S., Malla, A., Joober, R., 2017. Polygenic Risk Score associated with specific symptom dimensions in first-episode psychosis. Schizophr. Res. 184, 116–121.

Stefanovics, E.A., Elkis, H., Zhening, L., Zhang, X.Y., Rosenheck, R.A., 2014. A cross-national factor analytic comparison of three models of PANSS symptoms in schizophrenia. Psychiatry Res. 219, 283–289.

Sullivan, P.F., Keefe, R.S., Lange, L.A., Lange, E.M., Stroup, T.S., Lieberman, J., Maness, P.F., 2007. NCAM1 and neurocognition in schizophrenia. Biol. Psychiatry 61, 902–910.

Sullivan, P.F., Kendler, K.S., Neale, M.C., 2003. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. Arch. Gen. Psychiatry 60, 1187–1192.

Trampush, J.W., Yang, M.L.Z., Yu, J., Knowles, E., Davies, G., Liewald, D.C., Starr, J.M., Djurovic, S., Melle, I., Sundet, K., Christoforou, A., 2017. GWAS meta-analysis reveals novel loci and genetic correlates for general cognitive function: a report from the COGENT consortium. Mol. Psychiatry 22 (3), 336–345.

Van Scheltinga, A.T., Bakker, S.C., Van Haren, N.E.M., Derks, E.M., Buizer-Voskamp, J.E., Cahn, W., Ripke, S., Ophoff, R.A., Kahn, R.S., 2013. Schizophrenia genetic variants are not associated with intelligence. Psychol. Med. 43 (12), 2563–2570.

Walton, E., Geisler, D., Lee, P.H., Hass, J., Turner, J.A., Liu, J., Sponheim, S.R., White, T., Wassink, T.H., Roessner, V., Gollub, R.L., Calhoun, V.D., Ehrlich, S., 2014. Prefrontal inefficiency is associated with polygenic risk for schizophrenia. Schizophr. Bull. 40, 1263–1271.

Wray, N.R., Lee, S.H., Mehta, D., Vinkhuyzen, A.A., Dudbridge, F., Middeldorp, C.M., 2014. Research review: polygenic methods and their application to psychiatric traits. J. Child Psychol. Psychiatry 55, 1068–1087.

Xavier, R.M., Vorderstrasse, A., 2017. Genetic basis of positive and negative symptom domains in schizophrenia. Biological Research For Nursing 0, 1099800417715907.

Yeo, R.A., Gangestad, S.W., Walton, E., Ehrlich, S., Pommy, J., Turner, J.A., Liu, J., Mayer, A.R., Schulz, S.C., Ho, B.C., Bustillo, J.R., 2014. Genetic influences on cognitive endophenotypes in schizophrenia. Schizophr. Res. 156 (1), 71–75.