Molecular Psychiatry (2015) 20, 1386–1396

www.nature.com/mp

ORIGINAL ARTICLE

Reciprocal causation models of cognitive vs volumetric cerebral intermediate phenotypes for schizophrenia in a pan-European twin cohort

T Toulopoulou[1,2,3], N van Haren[4], X Zhang[5,6], PC Sham[1,5,6], SS Cherny[1,5,6], DD Campbell[5,6], M Picchioni[7,8], R Murray[9], DI Boomsma[10], HH Pol[4], R Brouwer[4], H Schnack[4], L Fañanás[11,12], H Sauer[13], I Nenadic[13], M Weisbrod[14,15], TD Cannon[16] and RS Kahn[4]

In aetiologically complex illnesses such as schizophrenia, there is no direct link between genotype and phenotype. Intermediate phenotypes could help clarify the underlying biology and assist in the hunt for genetic vulnerability variants. We have previously shown that cognition shares substantial genetic variance with schizophrenia; however, it is unknown if this reflects pleiotropic effects, direct causality or some shared third factor that links both, for example, brain volume (BV) changes. We quantified the degree of net genetic overlap and tested the direction of causation between schizophrenia liability, brain structure and cognition in a pan-European schizophrenia twin cohort consisting of 1243 members from 626 pairs. Cognitive deficits lie upstream of the liability for schizophrenia with about a quarter of the variance in liability to schizophrenia explained by variation in cognitive function. BV changes lay downstream of schizophrenia liability, with 4% of BV variation explained directly by variation in liability. However, our power to determine the nature of the relationship between BV deviation and schizophrenia liability was more limited. Thus, while there was strong evidence that cognitive impairment is causal to schizophrenia liability, we are not in a position to make a similar statement about the relationship between liability and BV. This is the first study to demonstrate that schizophrenia liability is expressed partially through cognitive deficits. One prediction of the finding that BV changes lie downstream of the disease liability is that the risk loci that influence schizophrenia liability will thereafter influence BV and to a lesser extent. By way of contrast, cognitive function lies upstream of schizophrenia, thus the relevant loci will actually have a larger effect size on cognitive function than on schizophrenia. These are testable predictions.

Molecular Psychiatry (2015) **20,** 1386–1396; doi:10.1038/mp.2014.152; published online 2 December 2014

INTRODUCTION

Although clinically reliable, the psychiatric nosologies embodied in the Diagnostic and Statistical Manual for Mental disorders (DSM-5) and the International Classification of Diseases (ICD-10) do not represent sufficiently refined phenotypes for molecular genetic studies.[1] This is because clinical phenotypes reflect aetiologically and biologically heterogeneous samples. The argument is not new.[2,3] Identifying genetically mediated vulnerability indexes that are more proximal to the underlying pathophysiology (that is, intermediate phenotypes) of psychiatric disorders, such as, schizophrenia, could help clarify the underlying biology and assist in the search for genetic vulnerability variants for these disorders. Intermediate phenotypes are hypothetically more homogeneous and less phenotypically complex than the clinical construct of schizophrenia. They are also quantitative in nature. Because the genes responsible are likely to be expressed in a quantitative manner, rather than dichotomously, it should be easier to detect genetic variants for them.

Over the past 20 years, mainly through prospective, high-risk and family study designs, numerous candidate intermediate phenotypes of schizophrenia have been proposed and rejected. Some early candidate intermediate phenotypes now do not appear to fulfill endophenotypic criteria at all while others are still promising.[4–9] Intelligence and memory deficits are frequently cited to meet the criteria. These deficits are detected in children who go on to develop schizophrenia[10–12] in prodromal[13] and high-risk states[14] and in patients experiencing their first episode of psychosis.[15] Similar impairments are found in patients' relatives, while recent twin studies confirmed that these aspects of cognition share some of their genetic influences with schizophrenia.[5,8,16,17] Similar findings have been reported for magnetic resonance imaging (MRI) volumetric data[1,18,19] although studies

[1]The State Key Laboratory of Brain and Cognitive Sciences, The University of Hong Kong, Hong Kong, SAR, China; [2]Department of Psychology, The University of Hong Kong, Hong Kong, SAR, China; [3]Department of Basic and Clinical Neuroscience, The Institute of Psychiatry, Psychology and Neuroscience, King's College, London, UK; [4]Department of Psychiatry, Brain Centre Rudolf Magnus, University Medical Centre Utrecht, Utrecht, The Netherlands; [5]Department of Psychiatry, The University of Hong Kong, Hong Kong, SAR, China; [6]Centre for Genomic Sciences, The University of Hong Kong, Hong Kong, SAR, China; [7]St Andrew's Academic Centre, Cliftonville, Northampton, UK; [8]Department of Forensic and Neurodevelopmental Science, Institute of Psychiatry, London, UK; [9]Department of Psychosis, King's College London, London, UK; [10]Department of Biological Psychology, VU University, Amsterdam, The Netherlands; [11]Anthropology Unit, Department of Animal Biology, Biomedicine Institute of the University of Barcelona (IBUB), University of Barcelona, Barcelona, Spain; [12]Centre for Biomedical Research Network on Mental Health (CIBERSAM), Instituto de Salud Carlos III, Madrid, Spain; [13]Department of Psychiatry and Psychotherapy University Hospital of Jena, Jena, Germany; [14]Department of Psychiatry and Psychotherapy, SRH Clinic Karlsbad-Langensteinbach, Karlsbad-Langensteinbach, Germany; [15]Section for Experimental Psychopathology and Neurophysiology, Centre for Psychosocial Medicine, University of Heidelberg, Heidelberg, Germany and [16]Departments of Psychology and Psychiatry, Yale University, New Haven, CT, USA. Correspondence: Dr T Toulopoulou, Department of Psychology, The University of Hong Kong, The Jockey Club Tower, 6/F, the Centennial Campus, Pokfulam Road, Hong Kong, SAR, China.
E-mail: timothea@hku.hk and timothea.toulopoulou@kcl.ac.uk
Received 18 June 2014; revised 1 September 2014; accepted 6 October 2014; published online 2 December 2014

that used genetic modelling either produced comparatively weaker results[20,21] or did not support such a role.[22] Although attempts to identify the genetic variants that underpin candidate intermediate phenotypes have so far been mostly unsuccessful,[23] some offer potential,[24,25] especially when the focus has been on the genetic architecture of the underlying cognitive neurobiology.[26] Together these findings suggest that deviations in cognitive and some brain volumetric phenotypes are an integral part of the pathophysiology of schizophrenia, possibly reflecting biological susceptibility.

To date, very few candidate intermediate phenotypes have been subjected to experimental testing by twin modelling. This has allowed major fissures to develop in our knowledge as the twin method represents the optimal experimental designs to control for genetic effects. Without these assurances, we cannot evaluate each trait's validity as an intermediate phenotype. Furthermore, with the exception of our own earlier study that included volumetric frontal lobe and cognitive data in a combined family and twin sample,[22] no other study has to date genetically modelled multiple intermediate phenotype classes (for example, cognitive and brain structure) or addressed the genetic relationship between these.

Here we address these issues by pooling cognitive and MRI volumetric data from all available European twin samples with schizophrenia. Our first aim was to define the relative contributions of genetic and environmental factors on the associations between the two candidate intermediate phenotypes, cognition and brain volume (BV), with schizophrenia liability. Model fitting allowed us to quantify the maximum shared genetic influences between schizophrenia and the other candidate intermediate phenotypes. The second aim was to quantify the shared genetic influences between the cognitive deficits and BV. This would identify whether these candidate intermediate phenotypes share common genetic causation. The final and particularly novel aim was to go beyond merely identifying any genetic correlations and instead to resolve the direction of causality between the candidate intermediate phenotypes and the liability to schizophrenia. To address these crucial experimental and clinical questions, we designed reciprocal models to resolve the directionality of causation between the three phenotypes. The first aim was to replicate the existing small body of research in a larger sample. The second and third aims intended to fill a significant gap in the schizophrenia intermediate phenotype literature.

MATERIALS AND METHODS

Participants

Data were available for 1243 members from 626 pairs (395 monozygotic (MZ); 231 dizygotic (DZ)), including 43 MZ pairs concordant for schizophrenia, 63 MZ and 56 DZ pairs discordant for schizophrenia (Table 1). Twin samples with clinical and cognitive data collected at the University Medical Centre Utrecht, Utrecht (UMCU, the Netherlands), the Institute of Psychiatry, London (IoP, United Kingdom), the National Public Health Institute of Finland (NPHI, Finland) and the University of Barcelona, Barcelona, (UoB, Spain) were pooled and combined with the MRI European multicentre twin database[21] to produce a European multi-phenotype twin database. The MRI database also included data from Jena University Hospital, Jena (JUH, Germany). The Dutch site contributed three cohorts. The discordant twin sample (MZ and DZ) have been described previously.[27] The control twins were recruited from the twin sample of the Department of Psychiatry at the UMCU, the Netherlands and from the population-based Netherlands Twin Register.[28,29] The British schizophrenia twins also have been described previously.[5,22,30,31] They were referred to the Institute of Psychiatry, London from across the country by their consulting psychiatrists while the British control twins were recruited from the Institute of Psychiatry Volunteer Twin Register and by national media advertisements. The Finnish twins were drawn from a twin cohort that comprised of all same-sex twins born in Finland from 1940 through 1957 in which both members of each pair were alive and residing in Finland as of 1967.[32] The Spanish twins were drawn from the University of Barcelona Twin Register and by advertisements in the national media.[33] The German twins were recruited from across Germany and formed part, as indeed did all other samples described above, of the European Twin Study Network on Schizophrenia (EUTwinsS) consortium. The studies were approved by their respective ethics committees, and all participants gave written informed consent before participating.

Clinical assessment

All participants underwent full psychiatric evaluation using available hospital records and structured clinical interviews. Exclusion criteria for all subjects were the presence of significant medical or neurological illnesses, alcohol or other drug dependence and significant past head trauma. Zygosity was based on DNA polymorphisms.

Cognitive assessment

Cognitive data were available for 1049 participants (Table 1). All sites, but Jena, contributed intelligent quotient (IQ) estimates from country-specific standardized versions of the Wechsler Adult Intelligence Scale-Revised[34] (WAIS-R: Utrecht, Helsinki) or of its successor, Wechsler Adult Intelligence Scale—Third Edition[35] (WAIS-III: London; Barcelona). All sites but London used an abbreviated battery to reliably estimate IQ. Specifically, all sites used the Vocabulary and the Block Design subtests of the WAIS. In addition, each site contributed data on additional WAIS subtests as follows: Information (Utrecht; London and Barcelona), Digit Span (London), Arithmetic (London), Comprehension (Utrecht and London), Similarities (London and Helsinki), Letter Number Sequencing (London), Picture Completion (Utrecht and London), Picture Arrangement (Utrecht, and London), Digit Symbol Coding (London and Helsinki), and Matrix Reasoning (London and Barcelona). The Utrecht site used four subtests to estimate IQ for each individual: Information and Picture Completion or Comprehension and Picture Arrangement in addition to Vocabulary and Block design. In London, IQ was calculated in the standard procedure for $>80\%$ of the sample. For the remaining sample, an abbreviated version of the WAIS-III was used to estimate IQ (Information, Block Design, Arithmetic, Digit Symbol Coding).

In addition, the London, Helsinki and Barcelona sites contributed data on aspects of verbal or visual memory function from the Wechsler Memory Scale—Revised (WMS-R)[36] or from a later version Wechsler Memory Scale —III (WMS-III)[37] WMS data were available for the Logical Memory subtest, immediate and delayed recall (London, Helsinki and Barcelona), Verbal Paired Associates (London), Visual Reproduction, immediate and delayed recall (London and Helsinki), Digit Span (London) and Visual Memory Span (London).

MRI processing

Scanner-type, MRI acquisition protocols for each site and information on the calibration algorithms developed are given in detail in Van Haren *et al.*[21] and in Schnack *et al.*[38] MRI scans from the collections described above were processed by the Department of Psychiatry at the University Medical Centre Utrecht using an established processing pipeline.[21] See Supplementary Information on MRI acquisition and software. High-quality MRI scans were available for 700 individuals (Table 1). All sites apart from Barcelona contributed MRI data. Data were available for the following BVs: intracranial, total brain, cerebral, cerebellar, grey matter, white matter volumes, lateral and third ventricular volumes and extracerebral cerebrospinal fluid.

Statistical analyses

Standardization. Cognitive data were collected at different sites, in some cases using different version of the same test. Compatibility was ensured through a standardization procedure regressing out the effect of age and gender, at each site separately, using the mean and s.d. of the respective control group. All standardized variables in the control group at each site had a mean of 0 and variance of 1. The analysis was conducted in SPSS 20 (IBM SPSS Statistics for Windows, Armonk, NY, USA).[39]

Group mean comparison. After pooling data across sites, generalized estimating equations (GEE) were applied in SPSS[39] to estimate the linear model parameters to identify differences in cognition and BV between patients with schizophrenia, their unaffected co-twins and healthy controls. Years in full-time education and intracranial volume were used as covariates for the cognitive data and the BV data, respectively. Robust

Table 1. Research participants

	MRI		*Cognition*		*Total*
	N	*Age, mean (s.d.)*	N	*Age, mean (s.d.)*	N
Site (UMCU, IoP, NPHI, UoB, JUH)					
N.(person)	328, 144, 190, 0, 38		235, 388, 189, 237, 0		328, 450, 190, 237, 38
N.(pair)	169, 72, 103, 0, 19		130, 201, 105, 119,0		169, 216, 103, 119, 19
MZ concordant					
N.(person)	52 (35/17)		77 (53/24)		80 (55/25)
N.(pair)	27 (18/9)	37.32 (9.23)	41 (28/13)	38.25 (10.61)	43 (29/14)
MZ discordant affected					
N.(person)	57 (30/27)		43 (24/19)		63 (36/27)
N.(pair)	57 (30/27)	37.88 (11.73)	43 (24/19)	40.88 (11.51)	63 (36/27)
MZ discordant, unaffected co-twin					
N.(person)	56 (28/28)		42 (25/17)		63 (34/29)
N.(pair)	56 (28/28)	38.11 (11.93)	42 (25/17)	40.88 (11.51)	63 (34/29)
MZ healthy control					
N.(person)	253 (137/116)		492 (191/301)		564 (226/338)
N.(pair)	128 (70/58)	37.18 (11.10)	252 (100/152)	37.60 (12.36)	284 (115/169)
DZ discordant affected					
N.(person)	43 (22/21)		53 (30/23)		56 (31/25)
N.(pair)	43 (20/19/0/4)	43.83 (9.70)	53 (23/21/4/5)	45.12 (10.91)	56 (25/23/4/4)
DZ discordant, unaffected co-twin					
N.(person)	43 (21/22)		48 (21/27)		55 (26/29)
N.(pair)	43 (20/19/0/4)	44.22 (10.02)	48 (19/20/4/5)	44.07 (11.02)	55 (24/23/4/4)
DZ healthy control					
N.(person)	196 (86/110)		294 (102/192)		333 (113/220)
N.(pair)	99 (29/41/27/2)	37.12 (11.83)	154 (36/82/32/4)	36.86 (12.83)	170 (36/91/39/4)
Total					
N.(person)	700 (359/341)		1049 (446/603)		1243 (542/701)
N.(pair)	363 (216MZ/147DZ)	38.44 (11.26)	555 (341MZ/214DZ)	38.49 (12.32)	626 (395MZ/231DZ)

Abbreviations: Concordant, concordant for schizophrenia; Discordant affected, ill co-twin from pairs discordant for schizophrenia; DZ, dizygotic; IoP, the Institute of Psychiatry, London, UK; JUH, Jena University Hospital, Germany; MZ, monozygotic; MZ/DZ discordant unaffected co-twin, unaffected co-twin from MZ or DZ pairs discordant for schizophrenia; N.(pair), number of pairs; N.(person), number of individuals included in the genetic analyses; NPHI, the National Public Health Institute of Finland, Finland; UMCU, University Medical Centre Utrecht, Utrecht, the Netherlands. The Dutch site contributed data from three cohorts, see text; UoB, University of Barcelona, Barcelona, Spain; Total, total number of participants with either MRI or cognitive data. For some pairs, only one member contributed data. The mean and s.d. of age are based on pairs.

Huber–White sandwich estimators were used to safeguard against misspecification in variance/covariance matrix and to control for any within-twin pair correlation.

Correlational analyses. Correlational analyses were performed between all cognitive and neuroimaging measures, and expressed separately for patients, co-twins and controls as a heatmap showing the pattern of correlations, and the clustering of variables using 'pheatmap' in R 2.14.2.[40,41]

Bivariate genetic model. Structural equation modelling in OpenMx[42] was applied to identify a model in which the variance and covariance of the phenotypes were separated into genetic and environmental influences. Additive genetic factors (A) relate to the total additive genetic influences that impact on a phenotype; common environment (C) represents environmental factors that are common or shared among family members while unique environment (E) signifies environmental factors that are unique to individual members of a family and contribute to making members of the same family different. The full-information maximum likelihood was used to generate the expected covariance matrix and estimate the free parameters. The likelihood ratio chi-squared statistic, minus two times log likelihood difference (−2LL) was used to compare the fit of the nested models.

The bivariate genetic model partitions the variance of each phenotype into A, C and E components, which were represented by path loadings labeled as a_n^2, c_n^2 and e_n^2 (n = phenotype 1 or 2). Further, the model specifies the correlations between the respective components of schizophrenia and cognition or BV as ra, rc and re, to reflect the degrees to which the A, C or E factors that influenced schizophrenia liability overlap with those impacting on the second phenotype (for example, cognition). The contributions of A, C and E to phenotypic correlation was calculated by combining ra, rc and re with a^2, c^2 and e^2, as follows:

$$\text{Rph} - \text{a} = \sqrt{(\text{a}_1{}^2)} * \text{ra} * \sqrt{(\text{a}_2{}^2)}$$

$$\text{Rph} - \text{c} = \sqrt{(\text{c}_1{}^2)} * \text{rc} * \sqrt{(\text{c}_2{}^2)}$$

$$\text{Rph} - \text{e} = \sqrt{(\text{e}_1{}^2)} * \text{re} * \sqrt{(\text{e}_2{}^2)}$$

Further, we calculated for each parameter the 95% confidence interval (95% CI). As with our previous work, schizophrenia liability was related to affection status via a liability threshold model. The parameters of this liability threshold model were fixed to the point estimates derived from meta-analysis[43] as follows: $a^2 = 0.81$, $c^2 = 0.11$, and $e^2 = 0.08$, and lifetime

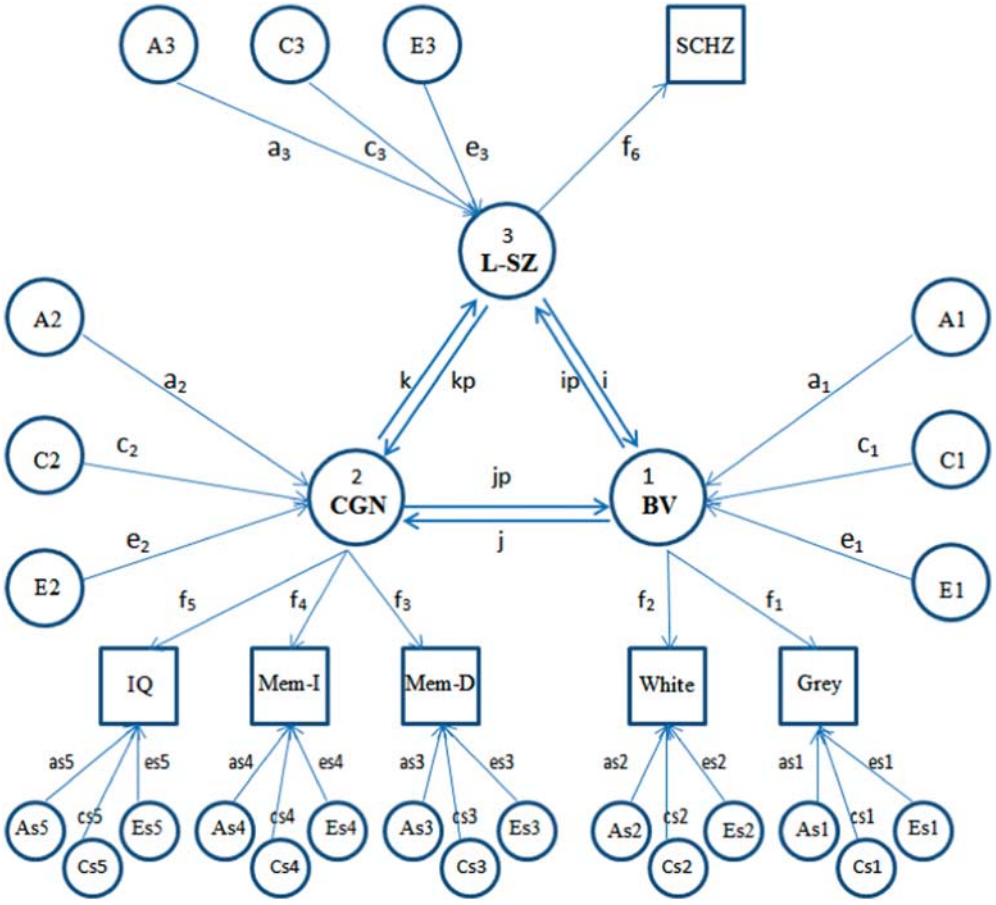

Figure 1. Latent phenotypes in reciprocal model. L-SZ = liability of schizophrenia; BV = brain volume; CGN = cognition; SCHZ = schizophrenia; white = cerebral white matter volume; grey = cerebral grey matter volume; Mem-I = Memory, immediate recall; Mem-D = Memory, delayed recall; A = genetic influence; C = common environmental influence; E = personal environmental influence; i, j, k, ip, kp and jp = single headed arrows representing causation paths. As = specific genetic influence and measurement error; Cs = specific common environmental influence and measurement error; Es = specific personal environmental influence and measurement error.

population prevalence = 1%.[43] In addition, we re-ran the analyses with several different point estimates of A, C and E for schizophrenia, to check whether the results are highly sensitive to the values of the fixed parameters. Only the measures that turned out to be statistically significantly different in the group mean comparison between co-twins and controls were included here. For cerebral volumes, we modelled grey and white matter as the primary tissue types.

Common pathways model. Multivariate model fitting was used to explore the relationship among schizophrenia, cognition and BVs (Supplementary Figure S1). The model employed three latent phenotypes to capture the common variance of inter-related observed data. In this way, information from multiple measures was utilized to maximize statistical power. As with the bivariate genetic models, common pathways models were fitted only on those measures that the previous level of analyses indicated as potentially important indicators of risk.

The first latent phenotype, schizophrenia liability (L-SZ), related to affection status and represented continues disease liability through liability threshold model. The second latent phenotype represented cognition (CGN) and was related to IQ, immediate memory and delayed memory by free parameter on path. The final latent phenotype BV extracted the common variance from cerebral grey and white matter volumes. For each latent phenotype, one of the path loadings from latent phenotype to an observed variable was fixed to 1.

A trivariate genetic common pathways common factor model[44] was used to estimate the genetic and environmental correlations among schizophrenia liability, cognition and BVs, simultaneously. The model also allowed estimation of A, C and E factors specific to the individual measures (except for affection status). All path coefficients were standardized.

Reciprocal model. The common pathways model allows for genetic and environmental correlations among phenotypes without making any assumptions about the direction of the relationship. The model implicitly assumes a common set of genetic and/or environmental influences to any pair of phenotypes. However, another possibility is that there is a causal relationship among phenotypes, such as the presence of brain abnormalities in causing schizophrenia, as opposed to perhaps a common set of genes that predispose to both a certain physical change in the brain and also to schizophrenia. Reciprocal models designed for family data were employed to address this issue and to resolve the direction of causation between the latent factors given the presence of a phenotypic causal relationship.[45] This model is depicted in Figure 1.

The cross-correlation between one phenotype in the first twin and the second phenotype in the second twin may provide important information about the direction of causality between these two phenotypes.[45] In the reciprocal model, the direction of causation among variables is embodied by two opposite single-headed arrows between each two latent phenotype variables, instead of double-headed arrows on A, C and E parts as in the correlated model. Thus, in reciprocal models the variance of the phenotype is presumed to result from the other phenotypes' variance, besides their own specific genetic and environmental influence. Like with the common pathways model, the reciprocal model has three latent phenotypes: L-SZ, CGN, and BV. Each latent phenotype was modelled to be influenced by the other two latent factors through the causal paths in Figure 1, including from clockwise paths: i, j, and k and from counter

clockwise paths: ip, jp, and kp. Further, the model included specific genetic and environmental effects, which also incorporated measurement error, for each latent phenotype, path loadings as, cs and es as well as ACE paths for the observed cognitive and BV phenotypes. As with the previous model, the liability threshold model was used to relate schizophrenia liability to affection status. All parameters were standardized.

RESULTS

Group mean comparison

Co-twins from the discordant pairs showed deficits on all cognitive measures compared with controls. Further, they showed deviations in two (cerebral volume and cerebral white matter) of the nine volumes tested compared with controls. The third ventricle showed a non-significant trend (Supplementary Table S1).

Correlational analyses

The first cluster includes measurements of immediate recall and delayed recall; the second, cerebral volumes, total BV, grey matter volume, intracranial volume, cerebellum and white matter volume. IQ segregated with both clusters 1 and 2. The remaining variables, that is, the lateral ventricle volumes, the third ventricle and extracellular cerebrospinal fluid do not cluster together. Overall, correlations between the cognitive and the BV measurements were modest ranging from respectable (for example, IQ and grey matter: 0.58) to small or non-existent. Supplementary Figure S2 shows a heatmap representing the correlations for the control group only and the spontaneous clustering of variables.

Bivariate genetic model

Genetic modelling was constrained to those variables that differed significantly between co-twins and controls in the group mean comparison analyses and, for the cerebral volumes, to grey and white matter as the primary tissue types.

The parameter ra reflects the degree to which the same genetic factors impact on two traits (for example, liability to schizophrenia and BVs) and was significant for all phenotypes. Genetic influences on schizophrenia liability and IQ were correlated by −0.62. These were slightly higher for immediate and delayed recall. For grey matter volume and schizophrenia liability, the genetic influences were correlated by −0.36. The correlation was lower for the white matter. The portion of the phenotypic correlations (Rph-t) due to shared genetic and environmental factors, both common and unique, are expressed as Rph-a, Rph-c and Rph-e, respectively. Thus, for delayed memory the portion of the phenotypic correlation (Rph-t = −0.76) due to shared genetic effects (Rph-a) was −0.64, suggesting that 84% of the phenotypic correlation was due to genetic factors (Rph-a −0.64/Rph-t $-0.76 \times 100 = 84\%$). Supplementary Table S2 shows the parameter estimates of the bivariate ACE model fitting.

Common pathways model

To explore the genetic and environmental relationship between traits, a model was built to incorporate the six observed phenotypes. First, we constructed a Cholesky model and used Cholesky decomposition of the variance of the observed phenotypes. Then we employed the common pathways model that included three latent phenotypes, L-SZ, CGN and BV, and compared its fit with the Cholesky model. The common pathways model fit was not significantly worse than the Cholesky one ($\Delta - 2LL = 29.68$, $\Delta df = 27$, $P = 0.33$) and was more parsimonious.

Figure 2 shows the standardized results of the common pathways model, with *a*, *b* and *c* representing the A, C and E components, respectively. For each observed phenotype, the variance is divided into two parts. The latent phenotype captures the shared variances of the variables, and the residual the specific variance for each variable, including its measurement error, also divided into genetic and environmental factors.

The path loadings of white matter and grey matter volumes on the BV latent phenotype are 0.82 and 0.69, respectively, and the square of these path loadings, 0.68 and 0.48, express the proportion of the variance that BV explains for each variable. Similarly, the path loadings of the cognitive measures to the latent phenotype CGN are 0.47 for IQ, 0.88 for immediate recall and 0.85 for delayed recall, with the proportion of the variance explained by CGN calculated as 22% for IQ, 78% for the immediate recall and 73% for the delayed recall.

The first latent phenotype represents the liability to schizophrenia, where the genetic and environmental contributions are already fixed. The latent phenotypes BV and CGN both have high heritability, nearly 1 (BV = A1 square of BV+A2 square of BV = $-0.212^2+0.959^2 = 0.045+0.919 = 0.964$) and 0.61(CGN = 0.158 +0.001+0.448 = 0.61), respectively, with environmental factors C and E contributing 0.19 and 0.21 to CGN, and close to 0 to BV.

The genetic correlation between the 'cognition' and the 'liability to schizophrenia' latent factors was −0.51 (95%CI −0.27, −0.85), indicating that 51% of the genetic factors that influence schizophrenia liability also impact on cognition. This correlation is higher than the correlation between the 'liability to schizophrenia' and the 'brain volume' latent factor at −0.22 (95%CI −0.02, −0.43). The tendency for higher correlations involving the 'cognition' latent factor is also seen in the phenotypic correlations and in the phenotypic correlation explained by overlapping genes. Supplementary Table S3 shows the genetic and environmental correlations between each of the two latent phenotypes derived from the common pathways model.

Reciprocal model

The full reciprocal model (model 1 in Table 2) included six causal paths between three latent factors (Figure 1). To test the direction of causation between these latent factors, we dropped the causal paths sequentially until we identified the first least explanatory causal path, then the second and third until the model fit was significantly worse than the full model. Specifically, we started by dropping sequentially all possible causal paths between the latent factors (models 2–7).

When we dropped the causal path which indicated that liability to schizophrenia is accounted for by variation in cognition (model 6), the model fit deteriorated significantly compared with the full reciprocal model. This suggested that that causal path is important and could not be dropped. In contrast, when we dropped the causal path which indicated that liability to schizophrenia is explained by variation in BV (model 3), the model was largely unaffected ($\Delta - 2LL = 0.01$, $\Delta df = 1$, $P = 0.92$) and had the smallest Akaike information criterion (AIC) value (AIC = 1496.74), suggesting that this causal path did not significantly contribute to the full model and could be dropped in subsequent steps.

Then we dropped two causal paths simultaneously. The path already identified (causal path from BV to L-SZ in Figure 1) and sequentially one other (models 8–12 in Table 2). When the causal paths from schizophrenia liability (model 8) or cognition (model 9) to BV deviation were dropped at the same time as the causal path from BV deviation to schizophrenia liability, the models were significantly different from model 3 (model 8: $P = 0.02$; model 9: $P < 0.001$). Similarly, model 11 without the causal paths from BV deviation to schizophrenia and from BV deviation to cognitive deficit was marginally significantly worse than model 3 ($P = 0.07$). Model 12 without the causal path from cognition to schizophrenia was also significantly worse than model 3 ($P < 0.001$). When we dropped the path indicating direction of causation from BV deviation to schizophrenia and from schizophrenia liability to cognitive deficit (model 10), the model did not significantly differ from the full reciprocal model (and the AIC value was the

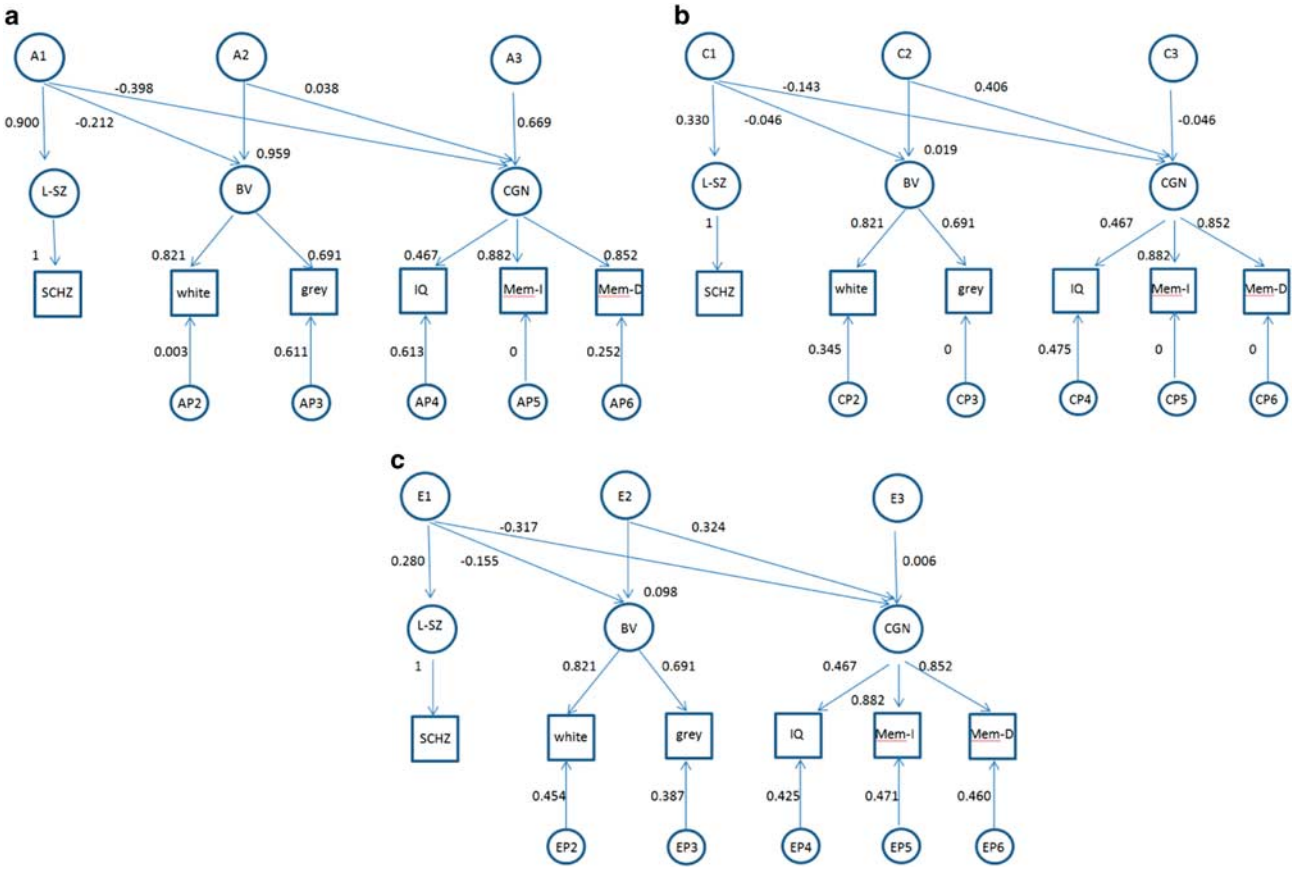

Figure 2. Estimated parameters of common pathways model. (**a**) Genetic component of the model. (**b**) Shared environmental component of the model. (**c**) Personal environmental component of the model. The figures show standardized results of the common pathways, and panels **a**, **b** and **c** correspondingly present the A, C and E part. L-SZ = liability of schizophrenia; BV = brain volume; CGN = cognition; SCHZ = schizophrenia; white = cerebral white matter volume; grey = cerebral grey matter volume; Mem-I = immediate recall; Mem-D = delayed recall; A = genetic influence; C = common environmental influence; E = personal environmental influence; AP = unique genetic influence; CP = unique common environmental influence; EP = unique personal environmental influence.

smallest), suggesting that these causal paths are not important and could be dropped in the subsequent steps below. We repeated the same process, dropping two causal paths simultaneously, including the second least important path from models 2 to 7, then the third and so on (models 13–21 in Table 2).

Based on model 10, we went on to drop three causal paths simultaneously: the two least important causal paths from the previous steps and then one by one all remaining causal paths (nested models of model 10: models 22–25). Model 23 was not significantly worse than model 10, so in addition to the paths suggested by model 10 we can drop path from BV to cognition without significant loss of fit. We then went on to drop four causal paths based on model 23 (nested models of model 23: models 29–31). These models differed significantly from model 23, suggesting that additional paths beyond those indicated in model 23 could not be dropped. We depict the model with the smallest AIC, model 10, in Figure 3.

Model 16 with a path reversed fitted the data almost equally well (based on the AIC) with model 10. We therefore repeated the process described above for model 10 (nested models of model 16: models 26–28; nested models of model 26: models 32–33). To summarize the results in Table 2: (1) the causal path from cognition to schizophrenia liability was present in all models; (2) the path from schizophrenia liability to cognition was consistently dropped, and (3) there is little robust evidence to determine which path direction between schizophrenia liability to BV is dropped. About 23% of the variance in liability to schizophrenia is explained from variation in cognitive function. For BV, this was approximately $\leqslant 5\%$ regardless of the path direction (data available on request).

Finally, we re-ran the analyses with several different point estimates of A, C and E for schizophrenia and confirmed that the results were not highly sensitive to the values of the fixed parameters (data available on request).

DISCUSSION

We addressed the nature of the covariance between schizophrenia liability and multiple volumetric and cognitive candidate intermediate phenotypes as well as the direction of any cause and effect relationships between them in this unique international cohort of twins with schizophrenia. We confirmed our previous findings in the smaller twin samples that cognition and BV share genetic influences with schizophrenia liability. We also found, for the first time, evidence indicating that schizophrenia liability is in part caused by cognitive deficit.

Our first aim was to explore whether selected candidate intermediate phenotypes for schizophrenia pertaining to cognition and BV share their genetic influences with the disorder. We found that both broad types of candidate intermediate phenotypes were caused partially by common genetic influences shared with schizophrenia. This pattern was confirmed both when

Table 2. Reciprocal model fitting

No.	*Name*	*ep*	*−2LL*	*df*	*AIC*	*diffLL*	*diffdf*	*P-value*	*comp.*	*Note*
0	Common pathways	53	11470.45	4983	1504.45					
1	Reciprocal	53	11470.73	4986	1498.73				0	
2	Drop i	52	11471.14	4987	1497.14	0.41	1	0.52	1	Drop one path
3	Drop ip	52	11470.74	4987	1496.74	0.01	1	0.92	1	
4	drop j	52	11473.43	4987	1499.43	2.70	1	0.10	1	
5	Drop jp	52	11473.40	4987	1499.4	2.67	1	0.10	1	
6	Drop k	52	11478.95	4987	1504.95	8.22	1	0	1	
7	Drop kp	52	11471.42	4987	1497.42	0.69	1	0.41	1	
8	Drop ip and i	51	11476.22	4988	1500.22	5.49	2	0.06	1	Drop two paths, including 'ip'
						5.48	1	0.02	3	
9	Drop ip and jp	51	11482.41	4988	1506.41	11.68	2	0	1	
						11.67	1	0	3	
10	**Drop ip and kp**	**51**	**11471.4**	**4988**	**1495.42**	**0.69**	**2**	**0.71**	1	
						0.68	**1**	**0.41**	3	
11	Drop ip and j	51	11474.00	4988	1498	3.26	2	0.20	1	
						3.25	1	0.07	3	
12	Drop ip and k	51	11478.95	4988	1502.95	8.22	2	0.02	1	
						8.21	1	0	3	
13	Drop i and j	51	11474.81	4988	1498.81	4.07	2	0.13	1	Drop two paths, including 'i'
						3.67	1	0.06	2	
14	Drop i and jp	51	11491.25	4988	1515.25	20.51	2	0	1	
						20.11	1	0	2	
15	Drop i and k	51	11479.00	4988	1503	8.27	2	0.02	1	
						7.86	1	0.01	2	
16	Drop i and kp	51	11471.64	4988	1495.64	0.91	2	0.64	1	
						0.50	1	0.48	2	
17	Drop j and jp	51	11474.26	4988	1498.26	3.53	2	0.17	1	Drop two paths, including 'j'
						0.83	1	0.36	4	
18	Drop j and k	51	11480.17	4988	1504.17	9.44	2	0.01	1	
						6.74	1	0.01	4	
19	Drop j and kp	51	11474.05	4988	1498.05	3.32	2	0.19	1	
						0.63	1	0.43	4	
20	Drop jp and k	51	11483.66	4988	1507.66	12.93	2	0	1	Drop two paths, including 'jp'
						10.27	1	0	5	
21	Drop jp and kp	51	11473.42	4988	1497.42	2.69	2	0.26	1	
						0.02	1	0.88	5	
22	Drop ip, kp and jp	50	11482.52	4989	1504.52	11.79	3	0.01	1	Drop three paths, including 'ip and kp' (nested models of model 10)
						11.11	1	0	10	
23	Drop ip, kp and j	50	11474.30	4989	1496.30	3.57	3	0.31	1	
						2.88	1	0.09	10	
24	Drop ip, kp and i	50	11478.77	4989	1500.77	8.04	3	0.05	1	
						7.35	1	0.01	10	
25	Drop ip, kp and k	50	11480.44	4989	1502.44	9.71	3	0.02	1	
						9.02	1	0	10	
26	Drop i, kp and j	50	11474.88	4989	1496.88	4.15	3	0.25	1	Drop three paths, including 'i and kp' (nested models of model 16)
						3.24	1	0.07	16	
27	Drop i, kp and jp	50	11491.51	4989	1513.51	20.78	3	0	1	
						19.87	1	0	16	
28	Drop i, kp and k	50	11575.53	4989	1597.53	104.80	3	0	1	
						103.89	1	0	16	
29	Drop ip, j, kp and jp	49	11482.57	4990	1502.57	11.84	4	0.02	1	Drop four paths, including 'ip and kp and j' (nested models of model 23)
						8.27	1	0	23	
30	Drop ip, j, kp and i	49	11480.03	4990	1500.03	9.30	4	0.05	1	
						5.73	1	0.02	23	
31	Drop ip, j, kp and k	49	11628.34	4990	1648.34	157.60	4	0	1	
						154.04	1	0	23	
32	Drop i, j, kp and jp	49	11503.03	4990	1523.03	32.29	4	0	1	Drop four paths, including 'i and kp and j' (nested models of model 26)
						28.15	1	0	26	
33	drop i, j, kp and k	49	11576.14	4990	1596.14	105.40	4	0	1	
						101.26	1	0	26	

Abbreviations: AIC, Akaike information criterion; comp., the model which is chosen for comparison; df, degree of freedom; diffLL, the difference of minus 2 log likelihood between two models; diffdf, the difference of the degree of freedom between two models; ep, estimate parameter; Model, reciprocal model; No., Model number; −2LL, minus 2 log likelihood. 0 in *P*-value column indicates $P < 0.01$. i, j, k, ip, jp and kp represent causal paths; causal path i, L-SZ to BV; causal path j, BV to CGN; causal path k, CGN to L-SZ; causal path ip, BV to L-SZ; causal path jp, CGN to BV; causal path kp, L-SZ to CGN (see Figure 1). The process involved dropping the causal paths one by one until the first least important causal path is identified (models 2–7). Then dropping the first (models 8–12), second (13–16), third (17–19) and fourth (20–21) least important causal paths with one other path at a time. Next, dropping three paths based on the best model (for example, nested models of best model (based on AIC), model 10; models 22–25) and on model 16 that had AIC close to model 10 (models 26–28) and then four paths (nested models of model 23: models 29–31 and nested models of model 26: models 32–33) until the model was significantly worse than the full model. Causal path k from cognition to schizophrenia is always significant while path kp from schizophrenia liability to cognition can be dropped. Bold values: Model 10 (Drop ip and kp): Model 10 did not significantly differ from the full reciprocal model (and the AIC value was the smallest), suggesting that these paths are not important and could be dropped in subsequent steps.

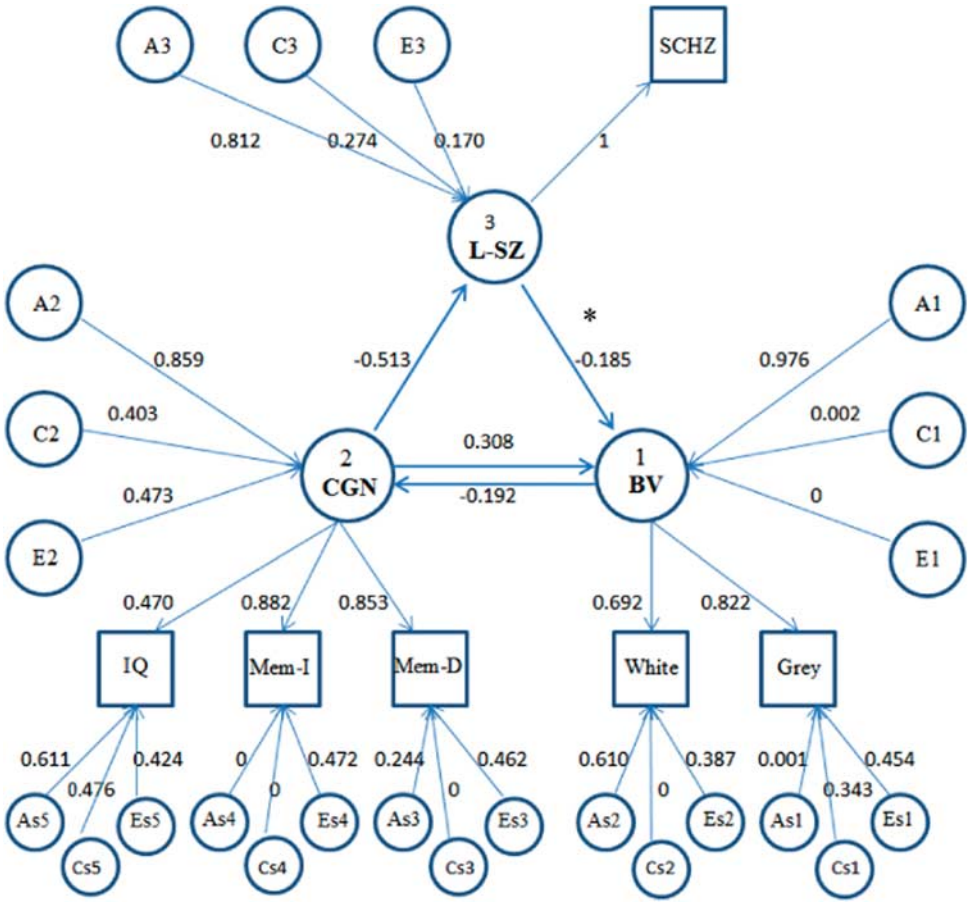

Figure 3. Estimated parameters of the reciprocal model. Estimated parameters based on model 10 in Table 2. L-SZ = liability of schizophrenia; BV = brain volume; CGN = cognition; white = cerebral white matter volume; grey = cerebral grey matter volume; IQ = estimated intelligence quotient; Mem-I = immediate recall; Mem-D = delayed recall; A = genetic influence; C = common environmental influence; E = personal environmental influence. As, Cs, Es = Specific genetic and environmental effects and measurement error incorporated into the model as the residual of the measured phenotypes. The latent phenotype CGN explained 22% (that is, the square of path loading $0.470^2 = 0.22 \times 100 = 22\%$), 78% and 73%, respectively, of the variance in IQ, immediate recall and delayed recall. White matter and grey matter volumes shared 48% and 68%, respectively, of their variance across the BV latent variable. The three latent phenotypes in this Figure interacted through four causal paths based on model 10 (that is, six causal paths of the full reciprocal model minus the two dropped): (1) CGN to L-SZ; (2) L-SZ to BV; (3) CGN to BV, and (4) BV to CGN. The causal paths suggest that cognitive deficit is upstream of the liability of schizophrenia and that about a quarter of the variance in liability to schizophrenia is explained from variation in cognitive function ($-0.513^2 = 0.26 \times 100 = 26\%$). The residual would be explained by other ACE factors not considered in the model. Brain volume is downstream of schizophrenia liability with approximately 4% (Direct: $-0.185^2 = 0.034 \times 100 = 3.4\%$) in brain volume variation explained from direct variation in schizophrenia liability and 0.35% through the indirect component of CGN (Indirect: $-0.513^2 \times (0.308^2 - 0.192^2) = -0.0595^2$). The two opposite direction causal paths between cognition and brain volume suggest a reciprocal relationship. Ten percent in brain volume is accounted for by variation in cognitive function while about 3.7% of variation in cognitive function is explained by variation in brain volume. *Model 16 with a path reversed suggests about 5% of the variance in liability to schizophrenia is explained from variation in brain volume, fitted the data almost equally well (based on the AIC) with model 10. The causal paths from cognition to schizophrenia liability and from cognition to brain volume were significant in all models that fitted the data well. The percentage of the variance in liability to schizophrenia explained from variation in cognitive function is comparable, if slightly higher, in all other models, including model 16, with that of model 10 (data available on request).

cognition and BV were modelled as observed traits (bivariate model) and as latent variables (common pathways model). Much higher estimates for covariance between cognition and schizophrenia compared with the BV and schizophrenia were detected in line with our previous work.[5,8,46,47]

Though critical, the question of the direction of causation has never been directly addressed before in the schizophrenia literature. Unlike other areas of science where direct manipulation of causal factors is possible, research into schizophrenia, and indeed any other psychiatric disorder, is often not easily amenable to such processes.[48,49] We cannot, for instance, experimentally increase cognitive deficit and reduce brain size in children, so that we could examine the consequences these might have on the subsequent adult rates of schizophrenia. When the question of cause and effect cannot be explored in any other way, reciprocal causation models in data collected from monozygotic and dizygotic twins may offer a reasonable alternative to explore causative relationships.[45,48] We have extended the reciprocal causation model from two to three latent phenotypes. We have used this to explore the causal relationship between schizophrenia, BV and cognitive deficit.

The models that identified cognitive deficit(s) as the determinant of schizophrenia liability fitted the data better than models that specified volumetric deviation as the cause of the disease.

This suggests that the genetic influences responsible for cognitive function cause variation in schizophrenia liability while the 'disease process' underlying schizophrenia causes variation in BV. It is important to distinguish the disease and the liability to disease. In the modelling, it is the liability to disease that is the latent factor, not the disease itself. We would interpret the liability to disease as a process, which is a quantitative trait present in everyone. The disease process is already elevated in patients at sub-threshold level before the onset of illness and in some clinically normal relatives of patients, and this can explain the presence of BV changes in these groups. One prediction of the interpretation that the schizophrenia disease process causes BV change is that the risk loci which influence schizophrenia risk will also influence BV, but with a smaller effect size than for schizophrenia. For cognitive function, one would predict that a proportion of the loci for schizophrenia would have a larger effect size for cognitive function than for schizophrenia.

The idea that cognitive deficits lie at an intermediate position on the pathophysiological pathway that links genes to the clinical phenotype is not new. For example, numerous retrospective and prospective studies of patients with schizophrenia have found evidence of cognitive deficit predating the onset of the illness.[50–52] These impairments manifest in early childhood and remain stable until early adolescence.[52,53] Twin and family studies indicate that deficits are transmitted within families[5,8,54,55] and that they are influenced by genetic variants implicated in schizophrenia.[56] Finally, the timing of the onset of schizophrenia coincides with the normal maturation of higher-order cognition, and many patients fail to achieve full maturation during that developmental period.[57,58] Together, these findings suggest that cognitive deficit is important in understanding schizophrenia but do not provide evidence that cognitive deficit is in the causative pathway. Results from our study suggest that this is the case.

The reciprocal models suggested that cognitive deficits are a phenotype that lies closer to the true underlying causes.[48,59] In this respect, understanding the molecular and cellular basis of cognition could provide critical new insights in the aetiology and pathophysiology of schizophrenia. Understanding how molecules influence brain structure and function to then influence cognitive processes such as memory[60,61] or how structural architecture influences dynamic functional networks[62] are increasingly the subject of research in this field, yet the specific mechanisms through which brain structure and function manifests itself as cognition remains very poorly understood.[63,64]

We found that cognitive deficit and volumetric brain abnormality share influences and also influence each other. The degree of shared influences was small with about half due to common genetic causation. Previous findings on structure–function relationships in healthy people[65–68] and in chronic,[69–71] first episode[72] and ultra-high-risk[73] samples are mixed. Results varied depending on the measures included and were at best modest; thus our results are in keeping with these previous findings. Although we do not fully understand how neurodevelopmental disease, IQ and chronicity influence these relationships across time, it is sometimes assumed that the direction of causation is from neuroanatomy to cognition.[68] We found that cognition shared a small reciprocal relationship with BV, but with the former having more influence on the latter. Specifically, about 10% in BV was accounted for by direct variation in cognitive function and 3.7% of variation in cognitive function was explained by direct variation in BV. We do not know what drives this reciprocal relationship but understanding the cellular and molecular mechanisms underlying neurogenesis and synaptic plasticity and how these relate to environmental factors is likely to be very important.

Inferential errors can occur when the assumptions that underpin genetic modelling are violated.[45,48] We adopted several approaches to reduce the likelihood of these errors. First, we selected and measured multiple indices for each of the two latent phenotypes: BV and cognition. Second, we incorporated measurement error into the models to parse out their statistical impact. Finally, we tested multiple causal pathways to comprehensively explore the direction of causation. The mode of transmission in our data approximated the optimal mode of inheritance under which these models work best providing reassurance in the results. Despite these efforts, these same questions should be tested by complimentary research methods, for example, by longitudinal cross-lagged correlations.

Another issue is that our results may be dependent on the measures included in this study, and as such, generalizability of the findings is subject to future validation. In our model, BV is a latent variable indexed by observed grey and white matter volumes. In order for other variables to be added, they should be highly correlated with grey and white matter (otherwise they are unlikely to be measuring the same latent variable) but should not be mathematically derived from grey and white matter as is the case with total cerebral volume (that would lead to collinearity in the model). For these reasons, we could not include third ventricle volume. Nonetheless, fitting reciprocal causation models to other measures of cortical integrity or cognition such as cortical thickness, connectivity or decision making will be of interest. One of the models with a path reversed fitted the data almost equally well with the best model (based on the AIC), indicating that we had insufficient power to determine the direction of causation between liability to schizophrenia and BV. The causal path from cognition to schizophrenia liability, on the other hand, was present in all models. The relationship between the liability threshold model used in the analysis and psychotic symptoms is intimate. If an individual's liability exceeds the threshold at any point in their life, the disorder will develop and as a consequence symptoms will be manifest, if not, there will be a near absence of symptoms. Here we were interested in exploring the aetiology of DSM-IV schizophrenia as a liability construct rather than transient psychotic symptoms around the time of scanning. Nevertheless, how the causative model relates to continuous state transient variables (psychotic symptoms) as opposed to persistent trait categorical one (diagnosis—as we used originally) is interesting.

Another limitation is that all the models are likely to represent simplifications of reality. The reciprocal model, for example, ignores possible correlations between the A, C and E of the observed traits. Another way of putting this is that we lack a general model that incorporates both the reciprocal and the common pathways. Such a general model has not been developed.

Finally, the nature of causality is that it influences outcome. We used a mathematical model to explain causality. Of course, the true test will be an interventional study to determine whether by improving cognition the liability to develop schizophrenia will be reduced. This paper helps to justify such a study. Although we have interpreted the finding as evidence of a causative relationship between cognition and schizophrenia, we cannot exclude the possibility that cognitive dysfunction may not 'cause' schizophrenia but that individuals with lower cognitive ability might be more likely to be diagnosed with the disorder.

In summary, one issue that affects reliability, replicability and effect size in genetic association studies is the characterization of the clinical phenotype, still usually based on psychiatric descriptive categories. There is a fundamental mismatch between twin and adoption studies, which indicate large genetic contributions to schizophrenia, and the results from large genome-wide association studies. Current molecular genetic studies mainly focus on the genetic variants associated with the clinical diagnosis, which is likely to be heterogeneous, harbouring aetiologically and biologically distinct populations, rather than with the vulnerability indexes that are more proximal to the underlying pathophysiology (that is, intermediate phenotypes). Little is known about the relative contributions of genetic and environmental influences on

the association of candidate intermediate phenotypes and schizophrenia liability while the direction of causation between candidate intermediate phenotypes and schizophrenia liability has never been explored before. Using all known available European MRI volumetric and cognitive data from twins with schizophrenia, we have shown for the first time that cognitive deficit lies in the intermediate pathway linking genes to schizophrenia, with about a quarter of the variance in liability to schizophrenia being explained from variation in cognitive function. One testable prediction of this finding is that the proportion of the loci for schizophrenia would have a larger effect size for cognitive function than for schizophrenia.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

We acknowledge support from the Research Grant Council (Hong Kong) through a General Research Fund grant award (Toulopoulou, PI; Sham, co-PI), NARSAD (through a Young Investigator Award to Toulopoulou) and the European Community's Sixth Framework Programme through a Marie Curie Training Network (MRTN-CT-2006-035987) called the European Twin Study Network on Schizophrenia (EUTwinsS).

REFERENCES

1 Waters-Metenier S, Toulopoulou T. Putative structural neuroimaging endophenotypes in schizophrenia: a comprehensive review of the current evidence. *Future Neurol* 2011; **6**: 679–715.
2 Jablensky A. The nature of psychiatric classification: issues beyond ICD-10 and DSM-IV. *Aust NZ J Psychiatry* 1999; **33**: 137–144.
3 Wessel J, Schork AJ, Tiwari HK, Schork NJ. Powerful designs for genetic association studies that consider twins and sibling pairs with discordant genotypes. *Genet Epidemiol* 2007; **31**: 789–796.
4 Greenwood TA, Braff DL, Light GA, Cadenhead KS, Calkins ME, Dobie DJ *et al.* Initial heritability analyses of endophenotypic measures for schizophrenia—the consortium on the genetics of schizophrenia. *Arch Gen Psychiatry* 2007; **64**: 1242–1250.
5 Toulopoulou T, Picchioni M, Rijsdijk F, Hua-Hall M, Ettinger U, Sham P *et al.* Substantial genetic overlap between neurocognition and schizophrenia—genetic modeling in twin samples. *Arch Gen Psychiatry* 2007; **64**: 1348–1355.
6 Aukes MF, Alizadeh BZ, Sitskoorn MM, Selten JP, Sinke RJ, Kemner C *et al.* Finding suitable phenotypes for genetic studies of schizophrenia: heritability and segregation analysis. *Biol Psychiatry* 2008; **64**: 128–136.
7 Waters-Metenier S, Toulopoulou T. Qualifying brain functional MRI parameters as endophenotypes in schizophrenia. *Future Neurol* 2010; **5**: 817–838.
8 Toulopoulou T, Goldberg TE, Mesa IR, Picchioni M, Rijsdijk F, Stahl D *et al.* Impaired intellect and memory: a missing link between genetic risk and schizophrenia? *Arch Gen Psychiatry* 2010; **67**: 905–913.
9 Gur RC, Calkins ME, Satterthwaite TD, Ruparel K, Bilker WB, Moore TM *et al.* NEurocognitive growth charting in psychosis spectrum youths. *JAMA Psychiatry* 2014; **71**: 366–374.
10 Seidman LJ, Cherkerzian S, Goldstein JM, Agnew-Blais J, Tsuang MT, Buka SL. Neuropsychological performance and family history in children at age 7 who develop adult schizophrenia or bipolar psychosis in the New England Family Studies. *Psychol Med* 2013; **43**: 119–131.
11 Cannon TD, Bearden CE, Hollister JM, Rosso IM, Sanchez LE, Hadley T. Childhood cognitive functioning in schizophrenia patients and their unaffected siblings: a prospective cohort study. *Schizophr Bull* 2000; **26**: 379–393.
12 Cannon M, Caspi A, Moffitt TE, Harrington H, Taylor A, Murray RM *et al.* Evidence for early-childhood, pan-developmental impairment specific to schizophreniform disorder—results from a longitudinal birth cohort. *Arch Gen Psychiatry* 2002; **59**: 449–456.
13 Seidman LJ, Giuliano AJ, Meyer EC, Addington J, Cadenhead KS, Cannon TD *et al.* Neuropsychology of the prodrome to psychosis in the NAPLS consortium: relationship to family history and conversion to psychosis. *Arch Gen Psychiatry* 2010; **67**: 578–588.
14 Keshavan MS, DeLisi LE, Seidman LJ. Early and broadly defined psychosis risk mental states. *Schizophr Res* 2011; **126**: 1–10.
15 Joyce E, Hutton S, Mutsatsa S, Gibbins H, Webb E, Paul S *et al.* Executive dysfunction in first-episode schizophrenia and relationship to duration of untreated psychosis: the West London Study. *Br J Psychiatry Suppl* 2002; **181**: S38–S44.
16 Owens SF, Picchioni MM, Rijsdijk FV, Stahl D, Vassos E, Rodger AK *et al.* Genetic overlap between episodic memory deficits and schizophrenia: results from the Maudsley Twin Study. *Psychol Med* 2011; **41**: 521–532.
17 Owens SF, Rijsdijk F, Picchioni MM, Stahl D, Nenadic I, Murray RM *et al.* Genetic overlap between schizophrenia and selective components of executive function. *Schizophr Res* 2011; **127**: 181–187.
18 Wood SJ, Pantelis C, Velakoulis D, Yucel M, Fornito A, McGorry PD. Progressive changes in the development toward schizophrenia: studies in subjects at increased symptomatic risk. *Schizophr Bull* 2008; **34**: 322–329.
19 Pantelis C, Velakoulis D, McGorry PD, Wood SJ, Suckling J, Phillips LJ *et al.* Neuroanatomical abnormalities before and after onset of psychosis: a cross-sectional and longitudinal MRI comparison. *Lancet* 2003; **361**: 281–288.
20 Rijsdijk FV, Van Haren NEM, Picchioni MM, McDonald C, Toulopoulou T, Pol HEH *et al.* Brain MRI abnormalities in schizophrenia: same genes or same environment? *Psychol Med* 2005; **35**: 1399–1409.
21 van Haren NEM, Rijsdijk F, Schnack HG, Picchioni MM, Toulopoulou T, Weisbrod M *et al.* The genetic and environmental determinants of the association between brain abnormalities and schizophrenia: the schizophrenia twins and relatives consortium. *Biol Psychiatry* 2012; **71**: 915–921.
22 Owens SF, Picchioni MM, Ettinger U, McDonald C, Walshe M, Schmechtig A *et al.* Prefrontal deviations in function but not volume are putative endophenotypes for schizophrenia. *Brain* 2012; **135**: 2231–2244.
23 Need AC, Attix DK, McEvoy JM, Cirulli ET, Linney KL, Hunt P *et al.* A genome-wide study of common SNPs and CNVs in cognitive performance in the CANTAB. *Hum Mol Genet* 2009; **18**: 4650–4661.
24 Lencz T, Knowles E, Davies G, Guha S, Liewald D, Starr J *et al.* Molecular genetic evidence for overlap between general cognitive ability and risk for schizophrenia: a report from the Cognitive Genomics consorTium (COGENT). *Mol Psychiatry* 2013; **19**: 168–174.
25 Lichtenstein P, Yip BH, Bjork C, Pawitan Y, Cannon TD, Sullivan PF *et al.* Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet* 2009; **373**: 234–239.
26 Tan HY, Chen AG, Kolachana B, Apud JA, Mattay VS, Callicott JH *et al.* Effective connectivity of AKT1-mediated dopaminergic working memory networks and pharmacogenetics of anti-dopaminergic treatment. *Brain* 2012; **135**: 1436–1445.
27 Hulshoff Pol HE, Brans RG, van Haren NE, Schnack HG, Langen M, Baaré WF *et al.* Gray and white matter volume abnormalities in monozygotic and same-gender dizygotic twins discordant for schizophrenia. *Biol Psychiatry* 2004; **55**: 126–130.
28 Willemsen G, De Geus EJ, Bartels M, Van Beijsterveldt C, Brooks AI, Estourgie-van Burk GF *et al.* The Netherlands Twin Register biobank: a resource for genetic epidemiological studies. *Twin Res Hum Genet* 2010; **13**: 231–245.
29 Baare WFC, van Oel CJ, Pol HEH, Schnack HG, Durston S, Sitskoorn MM *et al.* Volumes of brain structures in twins discordant for schizophrenia. *Arch Gen Psychiatry* 2001; **58**: 33–40.
30 Picchioni M, Toulopoulou T, Pauli A, Valli I, Ettinger U, Fu C *et al.* Exploring genetic and environmental influences on brain function in schizophrenia. *Schizophr Bull* 2011; **37**: 150.
31 Picchioni MM, Walshe M, Toulopoulou T, McDonald C, Taylor M, Waters-Metenier S *et al.* Genetic modelling of childhood social development and personality in twins and siblings with schizophrenia. *Psychol Med* 2010; **40**: 1305–1316.
32 Cannon TD, Kaprio J, Lonnqvist J, Huttunen M, Koskenvuo M. The genetic epidemiology of schizophrenia in a Finnish twin cohort—a population-based modeling study. *Arch Gen Psychiatry* 1998; **55**: 67–74.
33 Goldberg X, Alemany S, Rosa A, Picchioni M, Nenadic I, Owens SF *et al.* Substantial genetic link between iq and working memory: Implications for molecular genetic studies on schizophrenia. the European twin study of Schizophrenia (EUTwinsS). *Am J Med Genet B* 2013; **162B**: 413–418.
34 Wechsler D. *Wechsler Adult Intelligence Scale–Revised Manual*. The Psychological Corporation: New York, NY, USA, 1981.
35 Wechsler D. *Wechsler Adult Intelligence Scale–Third Edition: Administration and Scoring Manual*. The Psychological Corporation: London, UK, 1997.
36 Wechsler D. *Manual for the Wechsler Memory Scale–Revised*. The Psychological Corporation: San Antonio, TX, USA, 1987.
37 Wechsler D. *Wechsler Memory Scale—Third edition. Administration and Scoring Manual*. The Psychological Corporation: USA, 1997.
38 Schnack HG, van Haren NE, Hulshoff Pol HE, Picchioni M, Weisbrod M, Sauer H *et al.* Reliability of brain volumes from multicenter MRI acquisition: a calibration study. *Hum Brain Mapp* 2004; **22**: 312–320.
39 Corp. I. IBM. *SPSS Statistics for Windows Version 20.0*. IBM Corp: Armonk, NY, USA, 2011.
40 Kolde R. pheatmap: Pretty Heatmaps. R Package Version 0.7.4., 2014; http://CRAN.R-project.org/package=pheatmap.

41 Wilkinson L, Friendly M. The history of the cluster heat map. *Am Stat* 2009; **63**: 179–184.
42 Boker S, Neale M, Maes H, Wilde M, Spiegel M, Brick T *et al.* OpenMx: an open source extended structural equation modeling framework. *Psychometrika* 2011; **76**: 306–317.
43 Sullivan PF, Kendler KS, Neale MC. Schizophrenia as a complex trait—evidence from a meta-analysis of twin studies. *Arch Gen Psychiatry* 2003; **60**: 1187–1192.
44 Neale M, Cardon L. *Methodology for Genetic Studies of Twins and Families*. Springer, 1992.
45 Heath AC, Kessler RC, Neale MC, Hewitt JK, Eaves LJ, Kendler KS. Testing hypotheses about direction of causation using cross-sectional family data. *Behav Genet* 1993; **23**: 29–50.
46 Owens SF, Rijsdijk FV, Picchioni MM, Murray RM, Toulopoulou T. Genetic overlap between executive function and schizophrenia-the Maudsley Twin Study. *Schizophr Bull* 2009; **35**: 109.
47 Rodger AK, Owens SF, Rijsdijk FV, Picchioni MM, Waters-Metenier S, Murray RM *et al.* Genetic overlap between memory and schizophrenia. The Maudsley Twin Study. *Schizophr Bull* 2009; **35**: 118.
48 Duffy DL, Martin NG. Inferring the direction of causation in cross-sectional twin data—theoretical and empirical considerations. *Genet Epidemiol* 1994; **11**: 483–502.
49 Neale MC, Walters E, Heath AC, Kessler RC, Perusse D, Eaves LJ *et al.* Depression and parental bonding—cause, consequence, or genetic covariance. *Genet Epidemiol* 1994; **11**: 503–522.
50 Meier MH, Caspi A, Reichenberg A, Keefe RS, Fisher HL, Harrington H *et al.* Neuropsychological decline in schizophrenia from the premorbid to the postonset period: evidence from a population-representative longitudinal study. *Am J Psychiatry* 2013; **171**: 91–101.
51 Jones P, Rodgers B, Murray R, Marmot M. Child developmental risk-factors for adult schizophrenia in the British 1946 birth cohort. *Lancet* 1994; **344**: 1398–1402.
52 MacCabe JH, Wicks S, Lofving S, David AS, Berndtsson A, Gustafsson JE *et al.* Decline in cognitive performance between ages 13 and 18 years and the risk for psychosis in adulthood. A Swedish longitudinal cohort study in males. *JAMA Psychiatry* 2013; **70**: 261–270.
53 Reichenberg A, Caspi A, Harrington H, Houts R, Keefe RS, Murray RM *et al.* Static and dynamic cognitive deficits in childhood preceding adult schizophrenia: a 30-year study. *Am J Psychiatry* 2010; **167**: 160–169.
54 Goldberg TE, Ragland JD, Torrey EF, Gold JM, Bigelow LB, Weinberger DR. Neuropsychological assessment of monozygotic twins discordant for schizophrenia. *Arch Gen Psychiatry* 1990; **47**: 1066–1072.
55 Toulopoulou T, Mapua-Filbey F, Quraishi S, Kravariti E, Morris RG, McDonald C *et al.* Cognitive performance in presumed obligate carriers for psychosis. *Br J Psychiatry* 2005; **187**: 284–285.
56 Prata DP, Mechelli A, Fu CHY, Picchioni M, Toulopoulou T, Bramon E *et al.* Epistasis between the DAT 3 ' UTR VNTR and the COMT Val158Met SNP on cortical function in healthy subjects and patients with schizophrenia. *Proc Natl Acad Sci USA* 2009; **106**: 13600–13605.
57 Catts VS, Fung SJ, Long LE, Joshi D, Vercammen A, Allen KM *et al.* Rethinking schizophrenia in the context of normal neurodevelopment. *Front Cell Neurosci* 2013; **7**: 1–27.
58 Insel TR. Rethinking schizophrenia. *Nature* 2010; **468**: 187–193.
59 Rothman K. Modern Epidemiology, vol. 31. *Little, Brown and Co.: Boston, MA, USA* 1986.
60 Buxbaum AR, Wu B, Singer RH. Single β-actin mRNA detection in neurons reveals a mechanism for regulating its translatability. *Science* 2014; **343**: 419–422.
61 Park HY, Lim H, Yoon YJ, Follenzi A, Nwokafor C, Lopez-Jones M *et al.* Visualization of dynamics of single endogenous mRNA labeled in live mouse. *Science* 2014; **343**: 422–424.
62 Park HJ, Friston K. Structural and functional brain networks: from connections to cognition. *Science* 2013; **342**: 579.
63 Bilder RM, Howe AG, Sabb FW. Multilevel models from biology to psychology: mission impossible? *J Abnorm Psychol* 2013; **122**: 917–927.
64 Mercado E. Neural and cognitive plasticity: from maps to minds. *Psychol Bull* 2008; **134**: 109–137.
65 Martin-Loeches M, Bruner E, de la Cuetara JM, Colom R. Correlation between corpus callosum shape and cognitive performance in healthy young adults. *Brain Struct Funct* 2013; **218**: 721–731.
66 Karama S, Colom R, Johnson W, Deary IJ, Haier R, Waber DP *et al.* Cortical thickness correlates of specific cognitive performance accounted for by the general factor of intelligence in healthy children aged 6 to 18. *Neuroimage* 2011; **55**: 1443–1453.
67 Bruner E, Martin-Loeches M, Burgaleta M, Colom R. Midsagittal brain shape correlation with intelligence and cognitive performance. *Intelligence* 2011; **39**: 141–147.
68 Johnson W, Jung RE, Colom R, Haier RJ. Cognitive abilities independent of IQ correlate with regional brain structure. *Intelligence* 2008; **36**: 18–28.
69 Voineskos AN, Foussias G, Lerch J, Felsky D, Remington G, Rajji TK *et al.* Neuroimaging evidence for the deficit subtype of schizophrenia. *JAMA Psychiatry* 2013; **70**: 472–480.
70 McIntosh AM, Moorhead TWJ, McKirdy J, Hall J, Sussmann JED, Stanfield AC *et al.* Prefrontal gyral folding and its cognitive correlates in bipolar disorder and schizophrenia. *Acta Psychiatry Scand* 2009; **119**: 192–198.
71 Toulopoulou T, Grech A, Morris RG, Schulze K, McDonald C, Chapple B *et al.* The relationship between volumetric brain changes and cognitive function: a family study on schizophrenia. *Biol Psychiatry* 2004; **56**: 447–453.
72 Quan MN, Lee SH, Kubicki M, Kikinis Z, Rathi Y, Seidman LJ *et al.* White matter tract abnormalities between rostral middle frontal gyrus, inferior frontal gyrus and striatum in first-episode schizophrenia. *Schizophr Res* 2013; **145**: 1–10.
73 Meijer JH, Schmitz N, Nieman DH, Becker HE, van Amelsvoort TAMJ, Dingemans PM *et al.* Semantic fluency deficits and reduced grey matter before transition to psychosis: a voxelwise correlational analysis. *Psychiatry Res* 2011; **194**: 1–6.

Supplementary Information accompanies the paper on the Molecular Psychiatry website (http://www.nature.com/mp)