Associations between psychosis endophenotypes across brain functional, structural, and cognitive domains

R. Blakey¹,†, S. Ranlund¹,²,†, E. Zartaloudi¹,†, W. Cahn³, S. Calafato¹, M. Colizzi², B. Crespo-Facorro⁴,⁵, C. Daniel³, Á. Diez-Revuelta¹,⁶, M. Di Forti², GROUP‡, C. Iyegbe², A. Jablensky⁷, R. Jones¹, M.-H. Hall⁸, R. Kahn³, L. Kalaydjieva⁹, E. Kravariti², K. Lin²,¹⁰, C. McDonald¹¹, A. M. McIntosh¹²,¹³, PEIC‡, M. Picchioni², J. Powell², A. Presman¹, D. Rujescu¹⁴,¹⁵, K. Schulze², M. Shaikh¹⁶,¹², J. H. Thygesen¹, T. Toulopoulou²,¹⁷,¹⁸,¹⁹, N. Van Haren³, J. Van Os²⁰,², M. Walshe¹,², WTCCC²§, R. M. Murray², and E. Bramon¹,²,²¹

¹Division of Psychiatry, University College London, London, UK; ²Institute of Psychiatry Psychology and Neuroscience at King’s College London and South London and Maudsley NHS Foundation Trust, London, UK; ³Department of Psychiatry, Brain Centre Rudolf Magnus, University Medical Center Utrecht, Utrecht, The Netherlands; ⁴CIBERSAM, Centro Investigación Biomédica en Red Salud Mental, Madrid, Spain; ⁵Department of Psychiatry, University Hospital Marqués de Valdecilla, School of Medicine, University of Cantabria–IDIVAL, Santander, Spain; ⁶Laboratory of Cognitive and Computational Neuroscience – Centre for Biomedical Technology (CTB), Complutense University and Technical University of Madrid, Madrid, Spain; ⁷Centre for Clinical Research in Neuropsychiatry, The University of Western Australia, Perth, Western Australia, Australia; ⁸Psychology Research Laboratory, Harvard Medical School, McLean Hospital, Belmont, MA, USA; ⁹Harry Perkins Institute of Medical Research and Centre for Medical Research, The University of Western Australia, Perth, Australia; ¹⁰Nuffield Department of Population Health, University of Oxford, Oxford, UK; ¹¹Department of Psychiatry, Clinical Science Institute, National University of Ireland Galway, Ireland; ¹²Division of Psychiatry, University of Edinburgh, Royal Edinburgh Hospital, Edinburgh, UK; ¹³Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, UK; ¹⁴Department of Psychiatry, Ludwig-Maximilians University of Munich, Munich, Germany; ¹⁵Department of Psychiatry, Psychotherapy and Psychosomatics, University of Halle Wittenberg, Halle, Germany; ¹⁶North East London Foundation Trust, London, UK; ¹⁷Department of Psychology, Bilkent University, Main Campus, Bilkent, Ankara, Turkey; ¹⁸Department of Psychology, the University of Hong Kong, Pokfulam Rd, Hong Kong SAR, China; ¹⁹The State Key Laboratory of Brain and Cognitive Sciences, The University of Hong Kong, The Hong Kong Jockey Club Building for Interdisciplinary Research, Hong Kong SAR, China; ²⁰Department of Psychiatry and Psychology, Maastricht University

Author for correspondence: Dr S. Ranlund, Ph.D., siri.ranlund@kcl.ac.uk and e.bramon@ucl.ac.uk.
† Contributed equally as joint first authors.
‡ Co-authors who are members of the Psychosis Endophenotypes International Consortium (PEIC) and the Genetic Risk and Outcome of Psychosis (GROUP) consortium are listed at the end of this paper.
§ Members of the Wellcome Trust Case Control Consortium 2 (WTCCC2) are listed in the Supplement.
Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0033291717002860.

All authors declare that they have no financial interests or potential conflicts of interest.
Abstract

Background.—A range of endophenotypes characterise psychosis, however there has been limited work understanding if and how they are inter-related.

Methods.—This multi-centre study includes 8754 participants: 2212 people with a psychotic disorder, 1487 unaffected relatives of probands, and 5055 healthy controls. We investigated cognition [digit span (N = 3127), block design (N = 5491), and the Rey Auditory Verbal Learning Test (N = 3543)], electrophysiology [P300 amplitude and latency (N = 1102)], and neuroanatomy [lateral ventricular volume (N=1721)]. We used linear regression to assess the interrelationships between endophenotypes.

Results.—The P300 amplitude and latency were not associated (regression coef. −0.06, 95% CI −0.12 to 0.01, p = 0.060), and P300 amplitude was positively associated with block design (coef. 0.19, 95% CI 0.10–0.28, p < 0.001). There was no evidence of associations between lateral ventricular volume and the other measures (all p > 0.38). All the cognitive endophenotypes were associated with each other in the expected directions (all p < 0.001). Lastly, the relationships between pairs of endophenotypes were consistent in all three participant groups, differing for some of the cognitive pairings only in the strengths of the relationships.

Conclusions.—The P300 amplitude and latency are independent endophenotypes; the former indexing spatial visualisation and working memory, and the latter is hypothesised to index basic processing speed. Individuals with psychotic illnesses, their unaffected relatives, and healthy controls all show similar patterns of associations between endophenotypes, endorsing the theory of a continuum of psychosis liability across the population.

Keywords
Lateral ventricular volume; P300; schizophrenia; verbal memory; unaffected relatives; working memory

Introduction

Psychotic disorders, including schizophrenia and bipolar disorder, have considerable heritability with estimates ranging between 60 and 85% (Cardno et al. 1999; Smoller & Finn, 2003; Sullivan et al. 2012), and there is evidence of significant genetic overlap between these disorders (Lee et al. 2013). Psychoses are complex genetic disorders where many common variants contribute small increments of risk, and rare variants contribute greater risks (Gratten et al. 2014; Geschwind & Flint, 2015). While many common loci and some rare variants have now been identified (Stefansson et al. 2008; Stone et al. 2008; Walsh et al. 2008; Xu et al. 2008; Purcell et al. 2009; Grozeva et al. 2011; Sklar et al. 2011; Lee et al. 2013; Ripke et al. 2013, 2014; Green et al. 2015), little is known about their functional roles and the mechanisms through which they lead to the disease (Geschwind & Flint, 2015; Harrison, 2015).
Endophenotypes could help us gain a better understanding of the underlying neurobiology (Gottesman & Gould, 2003; Cannon & Keller, 2006; Gur et al. 2007). These are biological markers which are heritable, co-segregate with a disorder within families, are observed in unaffected family members at a higher rate than in the general population, and are expressed in an individual whether or not the illness is active (Gottesman & Gould, 2003). Endophenotypes could thus be used to better understand the mechanisms underlying the associations between genetic variants and the disorder (Hall & Smoller, 2010; Braff, 2015).

Although there is an extensive literature identifying and validating endophenotypes for psychosis, fewer studies have examined the relationships between different endophenotypes. Studies conducted so far have mainly analysed the associations between different cognitive measures (Toomey et al. 1998; Dickinson et al. 2002, 2006; Sullivan et al. 2003; Gladsjo et al. 2004; Sheffield et al. 2014; Seidman et al. 2015), but there is a lack of literature examining brain structural-cognitive and electrophysiological–cognitive pairings. Moreover, the inclusion of unaffected relatives in these studies has been rare, yet examining relatives – who carry increased genetic risk but have no illness or treatment confounding factors – is crucial for establishing the utility of these markers for genetic research.

This study seeks to investigate the relationships between the following electrophysiological, neurocognitive, and neuroanatomical endophenotypes for psychosis:

- **P300 event-related potential**: Reduced amplitude and prolonged latency of the P300 wave have consistently been found in patients with psychotic illnesses as well as in unaffected relatives, compared with controls (Blackwood et al. 1991; Weisbrod et al. 1999; Pierson et al. 2000; Winterer et al. 2003; Bramon et al. 2005; Price et al. 2006; Schulze et al. 2008; Bestelmeyer et al. 2009; Diez et al. 2013; Light et al. 2015; Turetsky et al. 2015). The P300 amplitude is thought to be a correlate of attention and working memory (Naatanen, 1990; Ford, 2014). Although the latency has been less precisely characterized, it is thought to index classification speed (Polich, 2007, 2011).

- **Cognitive performance**: Deficits on cognitive tests such as digit span (measuring working memory), block design (measuring working memory and spatial visualisation), and the Rey Auditory Verbal Learning Task (RA VLT) immediate and delayed recall (measuring short- and long-term verbal memory, respectively) are common and persistent across psychotic disorders (Heinrichs & Zakzanis, 1998; Gur et al. 2007; Bora et al. 2009; Stone et al. 2011; Bora & Pantelis, 2015; Kim et al. 2015b; Lee et al. 2015). Abnormalities are often observed before the onset of the illness as well as in unaffected relatives (Glahn et al. 2006; Saperstein et al. 2006; Snitz et al. 2006; Birkett et al. 2008; Horan et al. 2008; Forbes et al. 2009; Reichenberg et al. 2010; Ivleva et al. 2012; Park & Gooding, 2014; Gur et al. 2015).

- **Lateral ventricular volume**: Increased ventricular volume is a highly replicated finding in patients with psychosis compared with controls (Sharma et al. 1998; Fannon et al. 2000; Wright et al. 2000; Shenton et al. 2001; McDonald et al. 2002, 2006; Strasser et al. 2005; Boos et al. 2007; Crespo-Facorro et al. 2009;
This multi-centre study, seeking to investigate the relationships between multi-modal endophenotypes, includes the largest sample yet of individuals with psychosis, their unaffected first-degree relatives, and controls. The main objective is to facilitate the use of endophenotypes for genetic research into psychosis, which requires well defined and characterised measures. The aim of this study was therefore to examine the relationships between different endophenotype pairs, and in particular, to characterise the P300 event related potential in the context of well-defined cognitive markers.

Methods and materials

Sample and clinical assessments

The total sample included 8754 participants: 2212 individuals with a diagnosis of a psychotic disorder (see Table 1 for a breakdown of diagnoses), 1487 of their unaffected first-degree relatives (with no personal history of psychosis), and 5055 healthy controls (with no personal or family history of psychosis). Relatives and controls were not excluded if they had a personal history of non-psychotic disorders (such as depression or anxiety), provided they were well and off psychotropic medication at the time of testing and for the preceding 12 months.

To confirm or rule out a DSM-IV (APA, 1994) diagnosis, all participants underwent a structured clinical interview with either the Comprehensive Assessment of Symptoms and History (Andreasen et al. 1992), the Structured Clinical Interview for DSM Disorders (Spitzer et al. 1992), the Schedule for Affective Disorders and Schizophrenia (Endicott & Spitzer, 1978) or the Schedule for Clinical Assessment in Neuropsychiatry, Version 2.0 (Wing et al. 1990). Participants were excluded if they had a history of neurologic disease or a loss of consciousness due to a head injury.

Recruitment took place across 11 locations in Australia and Europe (Germany, Holland, Spain, and the UK) (see online Supplementary Table S1 in the supplement). Participants provided written informed consent, and the study was approved by the respective ethical committees at each of the 11 participating centres.

The main focus of this paper is an analysis of the associations between different endophenotype domains, which represents new and unpublished data. Some centres have previously published comparisons in endophenotype performance between groups (patients, relatives, and controls) (Weisbrod et al. 1999; Hulshoff Pol et al. 2002; McDonald et al. 2002; Steel et al. 2002; Bramon et al. 2005; Johnstone et al. 2005; Hall et al. 2006b; Price et al. 2006; Schulze et al. 2006; González-Blanch et al. 2007; Crespo-Facorro et al. 2009; Waters et al. 2009; Wobrock et al. 2009; Toulopoulou et al. 2010; Collip et al. 2013). Here we also present results of a mega-analysis of the combined multi-centre sample.
Neuropsychological assessments

The Wechsler Adult Intelligence Scale, revised version (Wechsler, 1981) or third edition (Wechsler, 1997), were administered to participants. Performance on two subtests was used for analyses: the combined forward and backward digit span (measuring attention and working memory) and block design (measuring spatial visualisation). The Rey Auditory Verbal Learning Test (Rey, 1964), including both immediate and delayed recall (assessing short- and long-term verbal memory, respectively), was also administered. Higher scores on the cognitive tasks indicate better performance. Full methodology for each contributing site is reported elsewhere (Johnstone et al. 2005; Crespo-Facorro et al. 2007; González-Blanch et al. 2007; Waters et al. 2009; Touloupoulou et al. 2010; Walters et al. 2010; Korver et al. 2012).

EEG data collection and processing

Electrophysiological data were obtained from three sites (online Supplementary Table S1). EEG data acquisition and processing methods varied slightly between sites as summarised below. The full methods for each site are reported elsewhere (Weisbrod et al. 1999; Bramon et al. 2005; Hall et al. 2006b; Price et al. 2006; Waters et al. 2009).

In summary, EEG was collected from 17 to 20 electrodes placed according to the International 10/20 system (Jasper, 1958). The P300 event related potential was obtained using a standard two-tone frequency deviant auditory oddball paradigm, with standard (‘non target’) tones of 1000 Hz and rare (‘target’) tones of 1500 Hz. The number of tones presented varied from 150 to 800, the tones were 80 dB or 97 dB, lasted for 20–50 ms, and the inter-stimulus interval was between 1 and 2 s. The majority of participants (93.4%) were asked to press a button in response to ‘target’ stimuli, but a subset were asked to close their eyes and count ‘target’ stimuli in their head.

The data were continuously recorded in one of three ways: 500 Hz sampling rate and 0.03–120 Hz band pass filter; 200 Hz sampling rate and 0.05–30 Hz band pass filter; or 400 Hz sampling rate and 70 Hz low-pass filter. Linked earlobes or mastoids were used as reference and vertical, and in most cases also horizontal, electro-oculographs were recorded at each site and used to correct for eye-blink artefacts using regression based weighting coefficients (Semlitsch et al. 1986). After additional manual checks, artefact-free epochs were included and baseline corrected before averaging. The averaged waveforms to correctly detected targets were then filtered using 0.03 or 0.05 Hz high-pass and 30 or 45 Hz low-pass filters. The peak amplitude and latency of the P300 were measured at electrode location PZ (parietal midline), within the range of 250–550 ms post-stimulus.

MRI data collection and processing

MRI data acquisition and image processing varied between sites; see previous publications and the supplementary materials for an outline of the methods used for each centre (Barta et al. 1997; Frangou et al. 1997; Hulshoff Pol et al. 2002; McDonald et al. 2002, 2006; McIntosh et al. 2004, 2005a, b; Schulze et al. 2006; Crespo-Facorro et al. 2009; Dutt et al. 2009; Mata et al. 2009; Wobrock et al. 2009; Habets et al. 2011; Collip et al. 2013). Field strengths included 1, 1.5 or 3 Tesla. Lateral ventricular volumes were measured using
automatic or semi-automatic region of interest analyses, and included the body, frontal, occipital, and temporal horns.

**Statistical methods**

**Mega-analysis of group comparisons**—Endophenotype measures were first standardised for each site separately using the mean and standard deviation within each site. Linear regression analyses for each measure were used to establish whether endophenotype performance differed according to group (patients, relatives, and controls). The outcome in each regression model was the endophenotype measure and the main predictor was group. These analyses were adjusted for age, gender, clinical group, study site and, where significant, group x site.

**Associations between endophenotypes**—Linear regression models were used to investigate associations between each pair of endophenotypes. Potential effect modification by group membership was assessed by specifying in the statistical model a term for the interaction between the predictor of the endophenotype pair and group (patient, relative, control). Where we found evidence that the relationship between a pair of endophenotypes differed according to group, associations are reported separately for patients, relatives, and controls. Where there was no evidence of effect modification, the interaction term was dropped from the model, and associations are reported for the whole sample adjusted for group. These analyses were adjusted for age, gender, clinical group, and study site.

In all analyses, we accounted for correlations between individuals within families using robust standard errors. 63% of the participants had no other family member taking part, but the study also included 1056 families of 2–11 members each (85% of the families had only two members included in the sample). This family clustering violates the independence of observations assumption in linear regression. To account for this clustered structure in the dataset we created a new variable ‘family ID’ that was shared by all individuals in each family. Then we used the variance estimator with the robust cluster option in all the linear regression models. This allowed us to account for the within-family correlations and maintain correct type-1 error rates. This is a standard approach in family studies (Shaikh et al. 2013; Bramon et al. 2014; Ranlund et al. 2014).

We examined the distribution of residuals and plots of residuals v. fitted values for all models and were able to rule out departures from normality and heteroscedasticity. Lateral ventricular volume showed a positively skewed distribution and to account for this we used bootstrap methods for analyses where this is the outcome variable. Heteroscedasticity was not found to be a concern for ventricular volumes. $p$ values are not presented for the models which used bootstrapping; instead, we examined the 95% bias-corrected confidence intervals to check for statistical significance at the 5% level ($p = 0.05$).

Although we tested seven endophenotypes, we expect measurements within domains to be correlated and thus a correction of $p$ values by seven tests through Bonferroni or other methods was deemed too stringent for a hypothesis-driven study such as this (Rothman, 1990; Savitz & Olshan, 1995; Perneger, 1998). We therefore corrected for associations between three domains (EEG, MRI, cognition), with a corrected significance threshold of
0.05/3 = 0.0167, that we rounded to the slightly more stringent cut-off of $p < 0.01$. Statistical analyses were conducted using STATA version 13.

**Results**

**Sample characteristics**

The sample characteristics are summarised in Table 1. Patients were on average 12.4 years younger than relatives (95% CI: 11.4–13.4; $p < 0.001$) and 11.9 years younger than controls (95% CI: 11.1–12.7; $p < 0.001$). There was no evidence of any age difference between relatives and controls. There was a lower proportion of females than males among patients than among relatives and controls (32.1%, 58.0%, and 51.5% respectively; global $p < 0.001$).

**Group comparisons on endophenotype performance**

As shown in Fig. 1 and Table 2, differences between the three participant groups on the endophenotypes followed the expected pattern with performance improving from patients through to relatives and controls. We found evidence that patients’ scores differed significantly from those of controls with smaller P300 amplitudes, delayed P300 latency, larger lateral ventricular volumes and deficits in digit span, block design and RVLT immediate recall. When compared with controls, the unaffected relatives showed reduced P300 amplitude, delayed P300 latency and poorer performance in digit span and block design.

**Associations between endophenotype pairs**

**Associations which do not differ according to clinical group**—Associations between endophenotype pairs where there was no evidence of effect modification by group are reported in Table 3. There was no evidence of an association between the P300 amplitude and latency at the 1% level of statistical significance (coef. $-0.06$, 95% CI $-0.12$ to $0.01$, $p = 0.06$). The P300 amplitude was positively associated with digit span (coef. $0.15$, 95% CI $0.04$–$0.26$, $p = 0.009$) and block design (coef. $0.19$, 95% CI $0.10$–$0.28$, $p < 0.001$) performances, but not with either of the RAVLT measures. The P300 latency showed weak evidence of a negative association with digit span (coef. $-0.15$, 95% CI $-0.28$ to $-0.03$, $p = 0.017$). Lateral ventricular volume showed no evidence of an association with any of the other measures. All cognitive pairings were significantly positively associated (all $p < 0.001$).

**Associations which differ according to clinical group**—For three pairs of cognitive endophenotypes, we found evidence of an interaction with group. This indicates that the association between these endophenotype pairs differs between patients, relatives, and controls, as reported in Fig. 2 (and online Supplementary Table S3 in the Supplement). In all three cases, the relationship between endophenotype pairs was in the same direction for the three groups, differing only in magnitude.

There was strong evidence that digit span and RAVLT immediate and delayed recall were positively associated with scores on the block design task in all three groups (patients, relatives, and controls). The magnitude of each association was greater among patients than
controls (all $p < 0.01$), but there was no evidence that the strength of the relationship among relatives was different from that among controls (all $p > 0.03$). See online supplementary Table S3 for full results.

**Discussion**

This study examined the relationships between different multimodal psychosis endophenotypes in a large multi-centre sample of patients, their unaffected first-degree relatives, and controls.

Our mega-analysis confirms that both patients and relatives showed reduced amplitudes and prolonged latencies of the P300, compared with controls, replicating past findings and providing further evidence that these are endophenotypes for psychosis (Turetsky et al. 2000; Bramon et al. 2005; Price et al. 2006; Schulze et al. 2008; Thaker, 2008; Bestelmeyer et al. 2009; Díez et al. 2013). We found no evidence of association between the P300 amplitude and latency, indicating that these are independent measures. To examine whether variability on P300 amplitude and latency could potentially affect the correlations between these, we tested for heteroscedasticity between clinical groups. The standard deviations between the patient, relative, and control groups did not vary significantly and are thus unlikely to explain the lack of correlation between P300 amplitude and latency performance. In contrast to our results, Hall et al. (Hall et al. 2006a) and Polich et al. (Polich, 1992; Polich et al. 1997) found a negative correlation between the amplitude and latency. Notably however, these past studies included only small samples (up to 128 participants) compared with our study ($N = 1083$), and they did not take into account covariates such as age and gender that are known to influence both P300 parameters (Goodin et al. 1978; Polich et al. 1985; Conroy & Polich, 2007; Chen et al. 2013). Furthermore, in the studies by Polich et al. (Polich, 1992; Polich et al. 1997) the amplitude – latency correlation was strongest over frontal electrodes, and not parietal as investigated in our current study. More recently, Hall et al. (2014) found a negative correlation between the amplitude and latency in a sample of 274 patients with psychosis and controls after controlling for age and gender effects. Further research is thus needed to clarify the relationship between the P300 amplitude and latency. However, our findings in this large sample suggest that the measures are independent, indexing separate brain functions.

We found associations between the P300 amplitude and both digit span and block design, as in previous smaller studies (Souza et al. 1995; Polich et al. 1997; Fjell & Walhovd, 2001; Hermens et al. 2010; Kaur et al. 2011; Dong et al. 2015b). According to the context-updating theory (Heslenfeld, 2003; Kujala & Naatanen, 2003), the P300 amplitude is an attention-driven, context-updating mechanism, which subsequently feeds into memory stores (Polich, 2007, 2011). Hence, one would expect the amplitude to be associated with cognitive tasks that require attention and working memory, such as digit span and block design (Näätänen, 1990; Baddeley, 1992; Ford, 2014). The context-updating theory provides a possible explanation for the association between P300 amplitude and block design, since this task requires a constant update of the mental representation of the blocks, in order to complete the target pattern (Polich, 2007, 2011). The lack of evidence for associations between P300 amplitude and the RAVLT tests support the idea that the neurobiology of...
verbal memory is distinct from the attentional and working memory processes linked to the P300 amplitude (Polich, 2011).

The P300 latency showed evidence of a trend-level association with digit span, and no evidence of an association with the other measures. Previous studies have provided conflicting results, with some reporting associations with attention and working memory (Polich et al. 1983), while others have not (Fjell & Walhovd, 2001; Walhovd & Fjell, 2003; Dong et al. 2015b). The P300 latency has been conceptualised as a measure of classification speed (Polich, 2011; van Dinteren et al. 2014). Investigating the relationship between behavioural reaction times (i.e. the speed of button press in the task) and the P300 latency, some have found associations (Bashore et al. 2014) while others have not (Ramchurn et al. 2014). Furthermore, there is a substantial body of research showing that the P300 latency as well as reaction times increase (that is they slow down) with ageing in healthy participants (Polich, 1996; Chen et al. 2013). Based on our findings we hypothesise that the P300 latency is a specific measure of processing speed at a basic neuronal level. In contrast, block design and the RAVLT task – while influenced by processing speed – reflect wider cognition including spatial abilities and verbal memory. The more complex elements to these tasks may therefore obscure effects of a simple processing speed, and hence explain the lack of association with P300 latency. The trend-level association with digit span performance – a task dependent on attention and short-term working memory – is in line with this interpretation too.

In terms of lateral ventricular volume, there was no evidence of a relationship with any other endophenotype investigated. Enlargement of cerebral ventricles remains the best replicated biological marker in schizophrenia and bipolar disorder, according to several meta-analyses (Kempton et al. 2010; Olabi et al. 2011; De Peri et al. 2012; Fusz-Poli et al. 2013; Fraguas et al. 2016; van Erp et al. 2016; Huhtaniska et al. 2017; Moberget et al. 2017). Our hypothesis that ventricular volumes would correlate with other endophenotypes of a functional nature was not confirmed by our data. Of course for such analyses our sample size was modest ranging 428–1001 and lack of statistical power could be a potential reason. Keilp et al. (Keilp et al. 1988) found an association with verbal memory and others have found enlarged lateral ventricles to be associated with poorer motor speed (Antonova et al. 2004; Hartberg et al. 2011; Dong et al. 2015a). A limitation of our study is the heterogeneity of the MRI methodology between study sites, which might have obscured any true associations. We conclude that ventricular volumes do not seem to exert a detectable influence on brain function in terms of cognition or cortical neurophysiology, however association studies of structural-functional biomarkers in larger samples are needed.

With regard to group comparisons, although patients showed enlarged lateral ventricles compared with controls, a very well supported finding (Wright et al. 2000; Steen et al. 2006; Cahn et al. 2009; Kempton et al. 2010), having adjusted by age and sex we observed no volume differences between relatives and controls. This is consistent with the latest meta-analysis of brain structure in relatives of patients with schizophrenia (Boos et al. 2007), and suggests that enlarged ventricles in patients are less heritable than previously thought. Instead, they might be related to illness progression, or to environmental effects or antipsychotic medication, as seen in both animal models of antipsychotic exposure (Dorph-

For all cognitive measures, patients performed less well than controls, consistent with extensive literature (Ayres et al. 2007; Horan et al. 2008; Bora et al. 2010, 2014; Fusar-Poli et al. 2012; Bora & Murray, 2014; Fatouros-Bergman et al. 2014; Stone et al. 2015). For the digit span and block design, there were also statistically significant differences between relatives and controls, suggesting a possible effect of increased genetic risk for psychosis. However, this was not seen for the immediate or delayed recall of the RAVLT task, where controls and relatives had similar performance. While some have reported verbal memory impairments in relatives of patients (Sitskoorn et al. 2004; Wittorf et al. 2004; Massuda et al. 2013), other studies have not (Üçok et al. 2013; Kim et al. 2015a). These findings suggest that working memory and spatial visualisation might represent more promising endophenotypes for genetic research into psychosis than verbal memory.

The associations between pairs of cognitive measures were strong and in the expected directions, as per previous findings (Dickinson et al. 2002; Sullivan et al. 2003; Gladsjo et al. 2004; Sheffield et al. 2014; Seidman et al. 2015). It is interesting to note that for some cognitive measures, the relationships interacted with group; however, the direction of the effect remained the same across patients, relatives, and controls. The interaction effects with group were found exclusively amongst the cognitive measures, and not in any of the other domains. This is possibly due to the larger sample sizes for the cognitive measures, yielding greater statistical power and enabling the detection of subtle interaction effects.

Both the lack of interaction effects for most associations investigated, and the gradient effects identified (where there was an interaction), are consistent with the notion that endophenotype impairments characterising psychosis represent a continuum that includes both relatives and the general population. Ultimately this continuum reflects the underlying variation in genetic liability of developing the disease (Johns & van Os, 2001; Wiles et al. 2006; Allardyce et al. 2007; Esterberg & Compton, 2009; Ian et al. 2010; DeRosse & Karlsgodt, 2015).

This study has several limitations: Firstly, association analyses could only be done for those participants with data available for pairs of endophenotypes and this led to relatively smaller samples for some of the associations. Secondly, there was a mismatch in age and gender between patients and relatives. The group of relatives has older individuals and more females compared with the group of patients who are younger and include more males. This is a common occurrence in psychosis family studies because the onset of psychosis in typically in youth. Most of the families who participated in the study include unaffected parents (with greater participation of mothers) and their affected and unaffected offspring. Family studies in psychosis are less likely to recruit affected parents. Because of this, we recruited a control group with a wider age range than either the other groups and with a balanced gender distribution so as to improve the age and sex matching across the two key comparisons (controls v. patients, controls v. relatives). Furthermore, since age and sex remains a potential confounder, we included age and sex as co-variates in the models throughout the study. As shown in online Supplementary Table S4 in the supplement, there
was no evidence of model instability based on the estimates and confidence interval width between the models with and without age and sex.

Another limitation of this study is that we were unable to account for potential moderators such as tobacco, other drug use and medication. Also, information about participants’ socioeconomic status was not available. These clinical and demographic variables could have a potentially important influence on how the three clinical groups perform on endophenotypes. However, the main analyses, which was to investigate associations between endophenotypes are all done within-individuals and are thus less likely to be influenced by exposure to drugs and medication. As for clinical variables such as depression, the sample included 5.5% of individuals with a history of depression. Depression did not constitute an exclusion criterion for our study because it is such a prevalent disorder that if excluded it would probably make our findings hard to generalize. We have re-analysed the group comparisons excluding all participants with a history of depression and the overall findings are unchanged.

A further potential limitation was the heterogeneity of methods between study sites; differences in cognitive test versions and variation on the EEG and MRI protocols all introduced greater variability into the data. All measures were standardised within centres to minimise this variability. Despite this challenge, it is precisely through this multi-centre effort that we were able to achieve a very large sample, the key strength of this study. As the Psychiatric Genomics Consortium’s work shows, large international collaborations are essential in genetic studies of common diseases and traits (Sklar et al. 2011; Lee et al. 2013; Smoller et al. 2013; Ripke et al. 2014). A further strength of this study is the use of regression models as opposed to the correlation approach frequently seen in the literature (Brewer et al. 1970; Polich et al. 1983, 1997; Breteler et al. 1994; Brillinger, 2001; Kim et al. 2003), which allowed us to account for some important confounding factors, such as ageing effects. Not only did this approach reduce vulnerability to spurious correlations, but it allowed the examination of interesting interaction effects across groups.

In summary, this study has investigated the relationships between endophenotypes for psychosis, including measures of cognition, electrophysiology, and brain structure. We have shown that cognitive measures are associated with each other as expected, and we have provided support for the notion that the amplitude and latency of the P300 are independent endophenotypes. The P300 amplitude is an index of spatial visualisation and working memory, while the latency is hypothesised to be a correlate of basic speed of processing. Individuals with psychotic illnesses, their unaffected relatives, and healthy controls all have similar patterns of associations between all pairs of endophenotypes, endorsing the theory of a continuum of liability of developing psychosis across the population.

Co-authors who are members of the Psychosis Endophenotypes International Consortium (PEIC):

Maria J. Arranz1,2, Steven Bakker3, Stephan Bender4,5, Elvira Bramon6,2, David Collier7,2, Benedicto Crespo-Facorro8,9, Marta Di Forti2, Jeremy Hall10, Mei-Hua Hall11, Conrad Iyegbe2, Assen Jablensky12, René S. Kahn3, Luba Kalaydjieva13, Eugenia Kravariti2, Stephen M Lawrie10, Cathryn M. Lewis2, Kuang Lin2,14, Don H. Linszen15, Ignacio
Mata\textsuperscript{16,9}, Colm McDonald\textsuperscript{17}, Andrew M McIntosh\textsuperscript{10,18}, Robin M. Murray\textsuperscript{2}, Roel A. Ophoff\textsuperscript{19}, Marco Picchioni\textsuperscript{2}, John Powell\textsuperscript{2}, Dan Rujescu\textsuperscript{20,21}, Timothea Touloumpoulou\textsuperscript{2,22,23}, Jim Van Os\textsuperscript{24,2}, Muriel Walshe\textsuperscript{6,2}, Matthias Weisbrod\textsuperscript{25,5}, and Durk Wiersma\textsuperscript{26}.

**PEIC affiliations:**

\textsuperscript{1}Fundació de Docència i Recerca Mútua de Terrassa, Universitat de Barcelona, Catalonia, Spain.

\textsuperscript{2}Institute of Psychiatry, Psychology and Neuroscience, King’s College London, De Crespigny Park, London SE5 8AF, UK.

\textsuperscript{3}University Medical Center Utrecht, Department of Psychiatry, Rudolf Magnus Institute of Neuroscience, The Netherlands.

\textsuperscript{4}Child and Adolescent Psychiatry, University of Technology Dresden, Fetscherstrasse 74, 01307 Dresden, Germany.

\textsuperscript{5}General Psychiatry, Vossstraße 4, 69115 Heidelberg, Germany.

\textsuperscript{6}Division of Psychiatry & Institute of Cognitive Neuroscience, University College London, UK.

\textsuperscript{7}Discovery Neuroscience Research, Lilly, UK.

\textsuperscript{8}University Hospital Marqués de Valdecilla, IDIVAL, Department of Psychiatry, School of Medicine, University of Cantabria, Santander, Spain.

\textsuperscript{9}CIBERSAM, Centro Investigación Biomédica en Red Salud Mental, Madrid, Spain.

\textsuperscript{10}College of Biomedical and Life Sciences, Cardiff University, CF24 4HQ Cardiff, UK.

\textsuperscript{11}Mclean Hospital, Harvard Medical School, Belmont MA, USA.

\textsuperscript{12}Centre for Clinical Research in Neuropsychiatry, The University of Western Australia, Perth, Australia.

\textsuperscript{13}Western Australian Institute for Medical Research and Centre for Medical Research, The University of Western Australia, Perth, Australia.

\textsuperscript{14}Nuffield Department of Population Health, University of Oxford, Oxford, UK.

\textsuperscript{15}Academic Medical Centre University of Amsterdam, Department of Psychiatry, Amsterdam The Netherlands.

\textsuperscript{16}Fundacion Argibide, Pamplona, Spain.

\textsuperscript{17}The Centre for Neuroimaging & Cognitive Genomics (NICOG) and NCBES Galway Neuroscience Centre, National University of Ireland Galway, Galway Ireland.

\textsuperscript{18}Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, UK.
Co-authors who are members of the Genetic Risk and Outcome of Psychosis (GROUP) consortium:

Richard Bruggeman, MD, PhD, Department of Psychiatry, University Medical Center Groningen, University of Groningen; Wiepke Cahn, MD, PhD, Department of Psychiatry, Rudolf Magnus Institute of Neuroscience, University Medical Center Utrecht; Lieuwe de Haan, MD, PhD, Department of Psychiatry, Academic Medical Center, University of Amsterdam; René S. Kahn, MD, PhD, Department of Psychiatry, Rudolf Magnus Institute of Neuroscience, University Medical Center Utrecht, Utrecht, the Netherlands; Carin Meijer, PhD, Department of Psychiatry, Academic Medical Center, University of Amsterdam; Inez Myin-Germeys, PhD, South Limburg Mental Health Research and Teaching Network, EURON, Maastricht University Medical Center; Jim van Os, MD, PhD, South Limburg Mental Health Research and Teaching Network, EURON, Maastricht University Medical Center, Maastricht, the Netherlands, and King’s College London, King’s Health Partners, Department of Psychosis Studies, Institute of Psychiatry, London, England; and Agna Bartels, PhD, Department of Psychiatry, University Medical Center Groningen, University.

**Supplementary Material**

Refer to Web version on PubMed Central for supplementary material.

**Acknowledgements.**

We would like to thank all the patients, relatives, and controls who took part in this research, as well as the clinical staff who facilitated their involvement. This work was supported by the Medical Research Council (G0901310) and the Wellcome Trust (grants 085475/B/08/Z, 085475/Z/08/Z). We thank the UCL Computer Science Cluster team for their excellent support. This study was supported by the NIHR Biomedical Research Centre at University College London (mental health theme) and by the NIHR Biomedical Research Centre for Mental Health at the South London and Maudsley NHS Foundation Trust and Institute of Psychiatry Kings College London.
E. Bramon thanks the following funders: BMA Margaret Temple grants 2016 and 2006, MRC- Korean Health Industry Development Institute Partnering Award (MC_PC_16014), MRC New Investigator Award and a MRC Centenary Award (G0901310), National Institute of Health Research UK post-doctoral fellowship, the Psychiatry Research Trust, the Schizophrenia Research Fund, the Brain and Behaviour Research foundation’s NARSAD Young Investigator Awards 2005, 2008, Wellcome Trust Research Training Fellowship.

Further support: The Brain and Behaviour Research foundation’s (NARSAD’s) Young Investigator Award (Grant 22604, awarded to C. Iyegbe). The BMA Margaret Temple grant 2016 to Johan Thygesen. European Research Council Marie Curie award to A Díez-Revuelta.

The infrastructure for the GROUP consortium is funded through the Geestkracht programme of the Dutch Health Research Council (ZON-MW, grant number 10-000-1001), and matching funds from participating pharmaceutical companies (Lundbeck, AstraZeneca, Eli Lilly, Janssen Cilag) and universities and mental health care organizations (Amsterdam: Academic Psychiatric Centre of the Academic Medical Center and the mental health institutions: GGZ Ingeest, Arkin, Dijk en Duin, GGZ Rivierduinen, Erasmus Medical Centre, GGZ Noord Holland Noord. Maastricht: Maastricht University Medical Centre and the mental health institutions: GGZ Eindhoven en de kempen, GGZ Breburg, GGZ Oost-Brabant, Vincent van Gogh voor Geestelijke Gezondheid, Mondriaan Zorggroep, Prins Clauscentrum Sittard, RIAGG Roermond, Universitair Centrum Sint-Jozef Kortenberg, CAPRI University of Antwerp, PC Ziekeren Sint-Truiden, PZ Sancta Maria Sint-Truiden, GGZ Overpelt, OPZ Rekem, Groningen: University Medical Center Groningen and the mental health institutions: Lentis, GGZ Friesland, GGZ Drenthe, Dimence, Mediant, GGNet Warnsveld, Yuliis Dordrecht and Pararnsis psycho-medical center (The Hague). Utrecht: University Medical Center Utrecht and the mental health institutions Altrecht, GGZ Centraal, Riagg Amersfoort and Delta).

The sample from Spain was collected at the Hospital Universitario Marqués de Valdecilla, University of Cantabria, Santander, Spain, under the following grant support: Carlos III Health Institute PI020499, PI050427, PI060507, Plan Nacional de Drugs Research Grant 2005-Orden sco/3246/2004, SENY Fundació Research Grant CI 2005–0380007 and Fundación Marqués de Valdecilla API07/011. We wish to acknowledge Biobanco HUMV-IDIVAL for hosting and managing blood samples and IDIVAL. Neuroimaging Unit for imaging acquirement and analysis.

References


Psychol Med. Author manuscript; available in PMC 2019 May 14.


Psychosis Endophenotypes International Consortium PEI, Wellcome Trust Case-Control Consortium 2


Psychol Med. Author manuscript; available in PMC 2019 May 14.


Fig. 1.
Estimated marginal means (adjusted for average age, gender, and study site) of standardised endophenotype scores by group (patients, relatives, and controls). Error bars represent standard errors of the means. RAVLT, Rey auditory verbal learning task.
Fig. 2.
Interactions between group (patient, relative, and control) and endophenotype pairs (standardised scores). Graphs are adjusted for covariates (age, gender, and study site), and include 95% confidence intervals. RAVLT, Rey auditory verbal learning task.
Table 1.

Sample characteristics (N = 8754)

<table>
<thead>
<tr>
<th></th>
<th>Patients with psychosis</th>
<th>Unaffected relatives</th>
<th>Controls</th>
<th>Total sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size, N (%)</td>
<td>2212 (25.3%)</td>
<td>1487 (17.0%)</td>
<td>5055 (57.7%)</td>
<td>8754</td>
</tr>
<tr>
<td>Age, mean years (S.D.)</td>
<td>31.6 (10.6)</td>
<td>46.0 (15.8)</td>
<td>45.5 (16.2)</td>
<td>42.6 (15.8)</td>
</tr>
<tr>
<td>Age range (years)</td>
<td>16–79</td>
<td>16–85</td>
<td>16–89</td>
<td>16–89</td>
</tr>
<tr>
<td>Gender (% female)</td>
<td>32.1%</td>
<td>58.0%</td>
<td>51.5%</td>
<td>47.7%</td>
</tr>
<tr>
<td>Diagnoses; N(%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>1396 (63.1%)</td>
<td>-</td>
<td>-</td>
<td>1396 (15.9%)</td>
</tr>
<tr>
<td>Bipolar I disorder</td>
<td>135 (6.1%)</td>
<td>-</td>
<td>-</td>
<td>135 (1.5%)</td>
</tr>
<tr>
<td>Psychosis NOS</td>
<td>168 (7.6%)</td>
<td>-</td>
<td>-</td>
<td>168 (1.9%)</td>
</tr>
<tr>
<td>Schizopreniform disorder</td>
<td>158 (7.1%)</td>
<td>-</td>
<td>-</td>
<td>158 (1.8%)</td>
</tr>
<tr>
<td>Schizoaffective disorder</td>
<td>124 (5.6%)</td>
<td>-</td>
<td>-</td>
<td>124 (1.4%)</td>
</tr>
<tr>
<td>Brief psychotic disorder</td>
<td>56 (2.5%)</td>
<td>-</td>
<td>-</td>
<td>56 (0.6%)</td>
</tr>
<tr>
<td>Other psychotic illness</td>
<td>175 (7.9%)</td>
<td>-</td>
<td>-</td>
<td>175 (2.0%)</td>
</tr>
<tr>
<td>Depression</td>
<td>246 (16.5%)</td>
<td>232 (4.6%)</td>
<td>478 (5.5%)</td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td>47 (3.2%)</td>
<td>24 (0.5%)</td>
<td>71 (0.8%)</td>
<td></td>
</tr>
<tr>
<td>Other non-psychotic illness</td>
<td>62 (4.2%)</td>
<td>106 (2.1%)</td>
<td>168 (1.9%)</td>
<td></td>
</tr>
<tr>
<td>No psychiatric illness</td>
<td>1132 (76.1%)</td>
<td>4693 (92.8%)</td>
<td>5825 (66.5%)</td>
<td></td>
</tr>
</tbody>
</table>

Endophenotypes N=sample size, Mean (SD) of raw scores unadjusted for covariates

<table>
<thead>
<tr>
<th></th>
<th>N = 397</th>
<th>N = 379</th>
<th>N= 313</th>
<th>N=1089</th>
</tr>
</thead>
<tbody>
<tr>
<td>P300 amplitude (μV)</td>
<td>10.5 (6.1)</td>
<td>11.0 (6.7)</td>
<td>13.7 (7.0)</td>
<td>11.6 (6.7)</td>
</tr>
<tr>
<td>P300 latency (ms)</td>
<td>382.6 (55.3)</td>
<td>390.8 (56.1)</td>
<td>356.9 (39.1)</td>
<td>378.2 (53.3)</td>
</tr>
<tr>
<td>Lateral ventricular volume (cm³)</td>
<td>700</td>
<td>337</td>
<td>684</td>
<td>1721</td>
</tr>
<tr>
<td>Block Design (% of max. score)</td>
<td>49.9 (27.9)</td>
<td>47.4 (25.6)</td>
<td>60.4 (21.2)</td>
<td>56.6 (23.8)</td>
</tr>
<tr>
<td>Digit Span (% of max. score)</td>
<td>460</td>
<td>136</td>
<td>2531</td>
<td>3127</td>
</tr>
<tr>
<td></td>
<td>Patients with psychosis</td>
<td>Unaffected relatives</td>
<td>Controls</td>
<td>Total sample</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-------------------------</td>
<td>----------------------</td>
<td>----------------</td>
<td>--------------</td>
</tr>
<tr>
<td></td>
<td>47.4 (15.9)</td>
<td>40.0 (4.5)</td>
<td>51.5 (14.5)</td>
<td>50.4 (14.9)</td>
</tr>
<tr>
<td>RAVLT immediate recall (no. of words recalled)</td>
<td>N = 1232</td>
<td>N = 934</td>
<td>N = 1377</td>
<td>N = 3543</td>
</tr>
<tr>
<td></td>
<td>7.6 (2.2)</td>
<td>8.4 (2.1)</td>
<td>8.7 (2.0)</td>
<td>8.2 (2.2)</td>
</tr>
<tr>
<td>RAVLT delayed recall (no. of words recalled)</td>
<td>N = 1224</td>
<td>N = 927</td>
<td>N = 1358</td>
<td>N = 3509</td>
</tr>
<tr>
<td></td>
<td>2.1 (1.0)</td>
<td>2.9 (1.0)</td>
<td>2.9 (0.9)</td>
<td>2.6 (1.0)</td>
</tr>
</tbody>
</table>

S.D., standard deviation; NOS, not otherwise specified; RAVLT, Rey auditory verbal learning task.

*Missing data for age (717 subjects) and gender (6 subjects).

The group differences in endophenotype performance adjusted by covariates are reported in Table 2.
Table 2.

Endophenotype performance comparison across clinical groups

<table>
<thead>
<tr>
<th>Endophenotype</th>
<th>Total sample</th>
<th>Patients - controls</th>
<th>Patients - relatives</th>
<th>Relatives - controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Global, p value *</td>
<td>Mean difference (95% CI)</td>
<td>Mean difference (95% CI)</td>
<td>Mean difference (95% CI)</td>
</tr>
<tr>
<td>P300 amplitude</td>
<td>&lt;0.001</td>
<td>−0.50 (−0.71 to −0.29) p &lt;0.001</td>
<td>−0.16 (−0.32 to −0.01) p = 0.061</td>
<td>−0.34 (−0.54 to −0.14) p = 0.001</td>
</tr>
<tr>
<td>P300 latency</td>
<td>&lt;0.001</td>
<td>0.47 (0.33–0.61) p &lt;0.001</td>
<td>0.03 (−0.14–0.18) p = 0.749</td>
<td>0.44 (0.29–0.60) p &lt;0.001</td>
</tr>
<tr>
<td>Lateral ventricular volume</td>
<td>0.20 (0.08–0.32)</td>
<td>0.09 (−0.06 to 0.23)</td>
<td>0.11 (−0.04 to 0.25)</td>
<td></td>
</tr>
<tr>
<td>Digit span</td>
<td>&lt;0.001</td>
<td>−0.72 (−0.88 to −0.55) p &lt; 0.001</td>
<td>−0.14 (−0.32 to 0.05) p = 0.141</td>
<td>−0.58 (−0.77 to −0.39) p &lt;0.001</td>
</tr>
<tr>
<td>Block design</td>
<td>&lt;0.001</td>
<td>−0.91 (−1.07 to −0.75) p &lt;0.001</td>
<td>−0.08 (−0.21 to 0.04) p = 0.190</td>
<td>−0.83 (−0.97 to −0.69) p &lt;0.001</td>
</tr>
<tr>
<td>RAVLT immediate recall</td>
<td>&lt;0.001</td>
<td>−1.32 (−2.29 to −0.37) p = 0.007</td>
<td>−1.24 (−2.22 to −0.27) p = 0.012</td>
<td>−0.08 (−0.24 to 0.07) p = 0.286</td>
</tr>
<tr>
<td>RAVLT delayed recall</td>
<td>=0.123</td>
<td>−0.98 (−2.21 to 0.25) p = 0.118</td>
<td>−0.94 (−2.18 to 0.30) p = 0.136</td>
<td>−0.03 (−0.20 to 0.13) p = 0.669</td>
</tr>
</tbody>
</table>

Linear regression models investigating group differences on endophenotype performance. Endophenotype data were standardised for each site using the mean and standard deviation within each site. The main predictor was clinical group (patients, relatives, and controls). All models included age, gender, study site and, where significant, group × centre interactions. We used robust standard errors to account for correlations within families in all models.

*P value for the overall test of a group effect; Note that p values were not produced for the models that include lateral ventricular volume since we used bootstrapping, which is a percentile-based method; therefore we looked at the bias-corrected confidence intervals to check for significance.

RAVLT, Rey auditory verbal learning task; CI, confidence interval.
Table 3.

Adjusted associations between endophenotypes in the whole sample

<table>
<thead>
<tr>
<th></th>
<th>P300 latency</th>
<th>Lateral ventricular volume</th>
<th>Digit span</th>
<th>Block design</th>
<th>RAVLT immediate recall</th>
<th>RAVLT delayed recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>P300 amplitude</td>
<td>N =1083</td>
<td>N = 428</td>
<td>N = 340</td>
<td>N = 426</td>
<td>N = 255</td>
<td>N = 255</td>
</tr>
<tr>
<td></td>
<td>-0.06 (−0.12 to 0.01)</td>
<td>0.05 (−0.07 to 0.15)</td>
<td>0.15 (0.04–0.26)</td>
<td>0.19 (0.10–0.28)</td>
<td>0.11 (−0.02 to 0.25)</td>
<td>0.08 (−0.06 to 0.22)</td>
</tr>
<tr>
<td></td>
<td>p = 0.060</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P300 latency</td>
<td>-</td>
<td>N = 434</td>
<td>N = 346</td>
<td>N = 437</td>
<td>N = 254</td>
<td>N = 254</td>
</tr>
<tr>
<td></td>
<td>0.02 (−0.08 to 0.15)</td>
<td>−0.15 (−0.28 to −0.03)</td>
<td>−0.04 (−0.12 to 0.04)</td>
<td>0.03 (−0.09 to 0.15)</td>
<td>0.03 (−0.07 to 0.14)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p = 0.017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral ventricular volume</td>
<td>-</td>
<td>N = 468</td>
<td>N = 1001</td>
<td>N = 498</td>
<td>N = 492</td>
<td></td>
</tr>
<tr>
<td></td>
<td>−0.01 (−0.09 to 0.09)</td>
<td>0.02 (−0.04 to 0.09)</td>
<td>−0.04 (−0.14 to 0.06)</td>
<td>−0.02 (−0.11 to 0.09)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p &lt;0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digit Span</td>
<td>-</td>
<td>N = 2754</td>
<td>N = 291</td>
<td>N = 291</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.33 (0.30–0.36)</td>
<td>0.39 (0.28–0.49)</td>
<td>0.31 (0.20–0.42)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p &lt;0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Block Design</td>
<td>-</td>
<td>N = 2169</td>
<td>N = 2137</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.26 (0.21–0.30)</td>
<td>0.24 (0.20–0.29)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p &lt;0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAVLT immediate recall</td>
<td>-</td>
<td>N = 3505</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.76 (0.74–0.78)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>p &lt; 0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

RAVLT, Rey auditory verbal learning task.

Regression models using standardised scores, adjusted for age, gender, study site and group using robust standard errors to account for correlations within families and, where significant, group × by centre interactions.

Statistics reported are sample sizes, regression coefficients (95% confidence intervals), and p values. Note that p values were not produced for the models that include lateral ventricular volume since we used bootstrapping, which is a percentile-based method; therefore we looked at the bias-corrected confidence intervals to check for significance.