

## ORIGINAL ARTICLE

## Dissection of major depressive disorder using polygenic risk scores for schizophrenia in two independent cohorts

HC Whalley<sup>1</sup>, MJ Adams<sup>1</sup>, LS Hall<sup>1</sup>, T-K Clarke<sup>1</sup>, AM Fernandez-Pujals<sup>1</sup>, J Gibson<sup>1</sup>, E Wigmore<sup>1</sup>, J Hafferty<sup>1</sup>, SP Hagenaars<sup>1</sup>, G Davies<sup>2,3</sup>, A Campbell<sup>4</sup>, C Hayward<sup>5</sup>, SM Lawrie<sup>1</sup>, DJ Porteous<sup>4</sup>, IJ Deary<sup>2,3</sup> and AM McIntosh<sup>1,2</sup>

Major depressive disorder (MDD) is known for its substantial clinical and suspected causal heterogeneity. It is characterized by low mood, psychomotor slowing and increased levels of the personality trait neuroticism; factors also associated with schizophrenia (SCZ). It is possible that some cases of MDD may have a substantial genetic loading for SCZ. The presence of SCZ-like MDD subgroups would be indicated by an interaction between MDD status and polygenic risk of SCZ on cognitive, personality and mood measures. Here, we hypothesized that higher SCZ polygenic risk would define larger MDD case-control differences in cognitive ability, and smaller differences in distress and neuroticism. Polygenic risk scores (PRSs) for SCZ and their association with cognitive variables, neuroticism, mood and psychological distress were estimated in a large population-based cohort (Generation Scotland: Scottish Family Health Study, GS:SFHS). The individuals were divided into those with, and without, depression ( $n = 2587$  and  $n = 16\,764$ , respectively) to test for the interactions between MDD status and schizophrenia risk. Replication was sought in UK Biobank (UKB;  $n = 6049$  and  $n = 27\,476$  cases and controls, respectively). In both the cohorts, we found significant interactions between SCZ-PRS and MDD status for measures of psychological distress ( $\beta_{GS} = -0.04$ ,  $P_{GS} = 0.014$  and  $\beta_{UKB} = -0.09$ ,  $P_{UKB} \leq 0.001$  for GS:SFHS and UKB, respectively) and neuroticism ( $\beta_{GS} = -0.04$ ,  $P_{GS} = 0.002$  and  $\beta_{UKB} = -0.06$ ,  $P_{UKB} = 0.023$ ). In both the cohorts, there was a reduction of case-control differences on a background of higher genetic risk of SCZ. These findings suggest that depression on a background of high genetic risk for SCZ may show attenuated associations with distress and neuroticism. This may represent a causally distinct form of MDD more closely related to SCZ.

*Translational Psychiatry* (2016) 6, e938; doi:10.1038/tp.2016.207; published online 1 November 2016

## INTRODUCTION

Major depressive disorder (MDD) has a lifetime prevalence of approximately 16% and is a heritable and severely disabling psychiatric condition.<sup>1</sup> It is highly clinically heterogeneous, relying on diagnostic criteria that are fulfilled through the presence of an arbitrary threshold of symptoms that lack empirical validation. Further, individuals passing this threshold can have a very different combination of symptoms. As a likely consequence of causal heterogeneity, the progress in our understanding of depression's underlying risk factors and biological mechanisms has been limited.

MDD is characterized by sustained negative affect, psychological distress and by differences in personality and behavioural traits, such as increased levels of the personality trait neuroticism.<sup>2</sup> People with MDD also manifest, on average, reduced executive function, attention, processing speed and working memory,<sup>3</sup> but relatively preserved general cognitive ability.<sup>4</sup> These behavioural and personality features are partly heritable, enduring features of illness that are often genetically correlated with liability to MDD.<sup>5–8</sup> Notably however, these clinical and personality associations are not specific to MDD and also occur in other psychiatric disorders including schizophrenia (SCZ).<sup>9–11</sup>

Although MDD is a notoriously heterogeneous disorder, evidence is beginning to emerge to suggest potentially distinguishable aetiological sub-categories.<sup>12</sup> For example, heritability

estimates are considerably higher for early-onset cases ( $\leq 25$  years) in comparison with later-onset MDD.<sup>13</sup> Early-onset MDD has also been associated with a more malignant form of the disorder, with increased numbers of episodes, greater symptom severity, increased functional and behavioural impairments, and personality traits such as increased neuroticism.<sup>14,15</sup> Genetic studies also suggest there may be differential pathways underlying these sub-categories grouped according to the age of onset<sup>16,17</sup> or symptom profiles.<sup>12</sup>

Genetic factors account for ~37% of liability to MDD and depressive symptoms.<sup>18,19</sup> Genome-wide association studies have indicated that ~30% of the heritability for MDD is accounted for by an additive polygenic genetic architecture, where risk is conferred by an accumulation of many alleles of small effect.<sup>20</sup> Molecular genetic approaches also demonstrate shared genetic architectures across the disorders.<sup>21</sup> For example, two papers have shown genetic correlation between MDD and SCZ using both genome-wide association analysis identifying shared risk loci,<sup>21</sup> or an approach estimating genetic correlation from sampled single-nucleotide polymorphisms (SNPs).<sup>22</sup> These findings have been interpreted as showing that there are genetic variants that raise liability to both MDD and SCZ. Recently, however, a second possibility has been proposed—that the genetic correlation between SCZ and MDD is caused not by pleiotropy, but by the misclassification of SCZ cases as MDD. This possibility has recently

<sup>1</sup>Division of Psychiatry, Royal Edinburgh Hospital, University of Edinburgh, Edinburgh, UK; <sup>2</sup>Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK; <sup>3</sup>Department of Psychology, University of Edinburgh, Edinburgh, UK; <sup>4</sup>Centre for Genetics and Molecular Medicine, Institute of Genetics and Molecular Medicine, Western General Hospital, University of Edinburgh, Edinburgh, UK and <sup>5</sup>MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK. Correspondence: Dr HC Whalley, Division of Psychiatry, Royal Edinburgh Hospital, University of Edinburgh, Morningside Park, Edinburgh EH10 5HF, UK. E-mail: heather.whalley@ed.ac.uk

Received 4 August 2016; accepted 30 August 2016

**Table 1.** Demographics, clinical, behavioural and temperament measures for GS:SFHS individuals in the current study

	Controls (n = 16 764) <sup>a</sup>		MDD cases (n = 2587) <sup>a</sup>		Significance T/Ch <sup>2</sup> (P-value)
	Mean	s.d.	Mean	s.d.	
<i>Demographics</i>					
Mean age (years) <sup>a</sup>	47.61	15.27	46.36	12.88	15.73 (7.31 × 10 <sup>-5</sup> )
Gender (M:F) <sup>a</sup>	7218:9546	—	746:1841	—	186.54 (2.20 × 10 <sup>-16</sup> )
<i>Cognitive</i>					
Logical memory (n = 16 764;n = 2587)	30.70	8.48	31.16	7.89	9.16 (2.48 × 10 <sup>-3</sup> )
Digit symbol (n = 16 566;n = 2566)	72.32	17.32	71.59	16.21	1.47 (2.25 × 10 <sup>-1</sup> )
Mill Hill vocabulary (n = 16 456;n = 2563)	30.11	4.72	31.21	4.79 <sup>3</sup>	1.70 (1.92 × 10 <sup>-1</sup> )
Verbal fluency (n = 16 553;n = 2575)	39.62	11.67	40.40	11.91	13.53 (2.35 × 10 <sup>-4</sup> )
<i>Clinical and trait-related features</i>					
Psycholog distress (GHQ)	14.93	7.55	22.70 <sup>a</sup>	12.62	1797.00 (< 1.00 × 10 <sup>-100</sup> )
Neuroticism	3.46	2.94	6.46 <sup>a</sup>	3.32	2155 (< 1.00 × 10 <sup>-100</sup> )
<i>PRS measure (measures scaled)</i>					
SCZ	-0.02	1.00	0.10	1.02	24.00 (9.69 × 10 <sup>-7</sup> )

Abbreviations: GHQ, General Health Questionnaire; GS:SFHS, Generation Scotland: Scottish Family Health Study; MDD, major depressive disorder; PRS, polygenic risk score; SCZ, schizophrenia. <sup>a</sup>Sample size as indicated unless otherwise specified in descriptive variable column, numbers for controls and MDD, respectively. Controlled for age, sex, four-multidimensional scaling ancestry components and relatedness.

been tested using a novel method aimed at using genotype data to distinguish pleiotropy from heterogeneity, where early evidence suggested that some cases of SCZ may have been misclassified as MDD, though this did not reach significance in this proof of technique study.<sup>23</sup> If confirmed, this would imply that some cases of MDD would be a hidden ‘forme fruste’ of SCZ, or that some individuals with MDD may be on a trajectory towards SCZ, but have yet to manifest the clinical phenotype.

The effects of polygenic risk for SCZ have also been demonstrated on a number of related behavioural and cognitive phenotypes. It has been consistently demonstrated, for example, that polygenic risk for schizophrenia is associated with decreased cognitive ability in healthy control samples.<sup>24–27</sup> However, as most studies have examined these relationships in unaffected controls, it is unclear whether the associations between SCZ risk and behaviour/cognitive ability are ameliorated or magnified by the presence of depression. Clarifying this relationship is important because, if MDD consists of a subtype that is closely related to schizophrenia, then that subtype would be expected to show associations with behaviour and cognitive ability that resembled those of SCZ itself.

In the current paper, we sought to stratify depression using a continuous scale of SCZ polygenic risk score (PRS), hypothesizing that case–control differences would depend on the genetic risk for schizophrenia; in other words, that there would be a significant interaction between SCZ-PRS and MDD status. We made two predictions: (1) based on the premise that deficits in cognition are greater in SCZ than in MDD, we first predicted that the high SCZ-PRS and MDD case–control differences for cognitive variables would be greater compared with those with lower SCZ PRS scores; (2) as neuroticism and psychological distress differ to a greater extent in MDD than SCZ,<sup>28</sup> we also predicted that SCZ-PRS and MDD status interactions would be found for these variables, but that they would be the reverse direction to cognition (attenuated in those with high SCZ PRS). We sought to test these predictions in a population-based cohort study, the Generation Scotland cohort: The Scottish Family Health Study sample (GS:SFHS).<sup>29,30</sup> Replication of any significant findings was sought in UK Biobank (UKB).

## MATERIALS AND METHODS

### Main study participants

*Generation Scotland: the Scottish Family Health Study.* GS:SFHS is a family- and population-based cohort study. The individuals were recruited at random through general medical practices across Scotland. Initial data collection for GS:SFHS took place between February 2006 and March 2011. The complete study protocol has been described in detail elsewhere.<sup>29–31</sup> Details of the mental health and related assessments are summarized below. Ethical approval was provided by NHS Tayside Research Ethics Committee (reference 05/S1401/89). Written consent for the use of data was obtained from all the participants.

The full cohort consists of 23 690 individuals who were over 18 years at the time of recruitment and 21 516 of these attended the research clinic. The present study includes 19 351 individuals on whom genome-wide genotype data were available (demographic details presented in Table 1). Pedigree information was available for all the participants, and this has subsequently been validated against estimates of relatedness estimated using genome-wide SNP data on 19 995 individuals.

*Assessment of MDD in GS:SFHS.* MDD was diagnosed using the structured clinical interview for the Diagnostic and Statistical Manual of Mental Disorders (SCID).<sup>32</sup> A brief questionnaire was administered to the participants to screen for MDD. The participants were asked ‘Have you ever seen anybody for emotional or psychiatric problems?’ and ‘Was there ever a time when you, or someone else, thought you should see someone because of the way you were feeling or acting?’ If they answered yes to either of these questions (21.7% screened positive), they were asked to complete the SCID.<sup>32</sup> If they answered no to both of these questions, they were assigned control status. Those who completed the SCID but did not meet the criteria for MDD or other major psychiatric disorder were also defined as controls. The individuals who declined to complete the screening questionnaire or the SCID had their MDD status set to missing. The individuals with a diagnosis of bipolar disorder were removed for the purposes of the current investigation.

*Cognitive assessment in GS:SFHS.* The cognitive abilities were assessed using four tests. Immediate and delayed scores from the recall section of one story of the Wechsler Logical Memory III UK test were summed to provide a measure of verbal declarative memory.<sup>33</sup> The Wechsler Digit Symbol Coding test was used to measure processing speed.<sup>33</sup> The verbal ability was assessed using the Mill Hill Vocabulary Scale, junior and senior synonyms.<sup>34</sup> The executive function was measured using the letter-based phonemic verbal fluency test (letters C, F and L, for 1 min each).<sup>35</sup> All the test scores were scaled to a mean of zero and standard deviation of one.

**Table 2.** PRSs SCZ×MDD status interactions in both samples

	GS:SFHS (Controls n = 16 764, MDD n = 2587) <sup>a</sup>		UK Biobank (Controls n = 27 476, MDD n = 6049) <sup>a</sup>		
	Statistics (β)	P-value	Statistics (β)	P-value	
<i>Cognitive</i>					
Composite 'g' factor	-0.0220	2.67 × 10 <sup>-1</sup>	Composite 'g' factor	-0.0434	1.75 × 10 <sup>-1</sup>
Logical memory (n = 16 764; n = 2587)	-0.0028	8.65 × 10 <sup>-1</sup>	Memory	-0.0290	3.54 × 10 <sup>-1</sup>
Digit symbol (n = 16 566; n = 2566)	<b>-0.0313</b>	<b>4.31 × 10<sup>-2</sup></b>	Symb digit substitution (n = 7799; n = 1767)	0.0197	6.52 × 10 <sup>-1</sup>
Mill Hill vocabulary (n = 16 456; n = 2563)	0.0009	9.49 × 10 <sup>-1</sup>	Reaction time (n = 27 299; n = 6008)	0.0119	6.45 × 10 <sup>-1</sup>
Verbal fluency (n = 16 553; n = 2575)	-0.0010	7.42 × 10 <sup>-1</sup>	Verb-num reasoning (n = 26 847; n = 5929)	-0.0282	2.92 × 10 <sup>-1</sup>
<i>Clinical and trait-related features</i>					
Psycholog distress (GHQ)	<b>-0.0402</b>	<b>1.44 × 10<sup>-2</sup></b>	Psycholog distress (PHQ) (n = 26 581; n = 5969)	<b>-0.0903</b>	<b>8.03 × 10<sup>-4</sup></b>
Neuroticism	<b>-0.0483</b>	<b>1.96 × 10<sup>-3</sup></b>	Neuroticism	<b>-0.0602</b>	<b>2.28 × 10<sup>-2</sup></b>

Abbreviations: GHQ, General Health Questionnaire; GS:SFHS, Generation Scotland: Scottish Family Health Study; MDD, major depressive disorder; PHQ, Patient Health Questionnaire; PRS, polygenic risk score; SCZ, schizophrenia; symb, symbol; verb-num, verbal-numerical. <sup>a</sup>Sample size as indicated unless otherwise specified in descriptive variable column, numbers for controls and MDD, respectively. Controlled for age, sex, population stratification and relatedness. Significant interactions are highlighted in bold.

General cognitive ability (g) scores were obtained by conducting principal component analysis of the tests, and saving the scores on the first unrotated principal component, on which all the tests had substantial loadings.<sup>36</sup>

*Measures of neuroticism and psychological distress GS:SFHS.* Measures of psychological distress were assessed using the 28-item General Health Questionnaire (GHQ-28),<sup>37</sup> which consists of four subscales designed to assess the following: (1) somatic symptoms, (2) anxiety and insomnia, (3) social dysfunction and (d) 'severe depression'. A single measure of global affective symptoms was derived from item responses, and the standard Likert scoring totals were used in the current study. The Eysenck Personality Questionnaire-Revised Short Form provided a measure of neuroticism.<sup>38</sup> For the numbers of individuals completing each measure, see Table 2.

*Genotyping and polygenic profiling in GS:SFHS.* The blood samples were collected using standard operating procedures and stored at the Wellcome Trust Clinical Research Facility Genetics Core, Edinburgh (www.wtcrf.ed.ac.uk) where they were genotyped using the IlluminaHumanOmniExpressExome-8v1.0 BeadChip and Infinium chemistry.<sup>39</sup> The genotypes were then processed using the IlluminaGenomeStudio Analysis software v2011.1. The details of blood collection and DNA extraction are provided elsewhere.<sup>40</sup>

The PRSs were calculated using the method described previously,<sup>41</sup> using summary data from the most recent genome-wide association studies from the Psychiatric Genomics Consortium (2014). Briefly, the SNPs were excluded if they had a minor allele frequency <1%, deviated significantly from Hardy-Weinberg equilibrium ( $P < 1 \times 10^{-6}$ ) in the total sample of founder individuals, or had a call rate <99%. Clump-based linkage disequilibrium pruning ( $r^2 = 0.2$ , 300 kb window) was performed to create a SNP set in linkage equilibrium. Before creating PRSs, all strand-ambiguous SNPs were removed from the GS:SFHS data set and the PRS for each individual was calculated using the cumulative sum of the number of each SNP allele multiplied by the log of the odds ratio for SCZ across their whole genome. Five SNP set scores were generated using the  $P$ -value threshold cut-offs of 1, 0.5, 0.1, 0.05 and 0.01. Our primary analyses concerned those SNPs from the Psychiatric Genomics Consortium data that met a significance level of  $P = 0.5$  consistent with the previous studies.<sup>41</sup> The findings for other thresholds are contained within the Supplementary Material.

#### Replication sample: UKB

This study includes replication data from the UKB Study (<http://www.ukbiobank.ac.uk>). UKB received ethical approval from the Research Ethics Committee (REC, reference: 11/NW/0382). UKB is a health research

resource that aims to improve the prevention, diagnosis and treatment of a range of illnesses. Through UKB, 502 655 community-dwelling participants were recruited between 2006 and 2010 in the United Kingdom.<sup>42</sup> They underwent cognitive, neuroticism and physical assessments, provided blood, urine and saliva samples for future analysis, gave detailed information about their backgrounds and lifestyles, and agreed to have their health followed longitudinally. For the present study, genome-wide genotyping data were available on 33 525 individuals. This data set comprised individuals who survived the quality control process, were independent of individuals within GS:SFHS, who were unrelated and who did not meet criteria for another major psychiatric disorder. The individuals were assigned to a narrow definition of MDD including cases of moderate or recurrent MDD, ( $n = 6049$  and  $n = 27 476$  cases and controls, respectively), see Supplementary Material for further information.

#### Cognitive assessment in UKB

Three cognitive tests were used in the present study to generate measures of general intelligence 'g'. These tests, which cover the cognitive domains of reaction time, memory and verbal-numerical reasoning were collected at the baseline assessment. In addition, the data from a second assessment, which provided measures of processing speed (similar to the Digit Symbol Coding test, the Symbol Digit Substitution test) were also assessed (further details in Supplementary Material). All the test scores were scaled to a mean of zero and standard deviation of 1. General cognitive ability (g) was estimated using the principal component analysis approach described above.<sup>36</sup>

#### Measures of neuroticism and psychological distress in UKB

In UKB, the participants completed the Neuroticism scale of the Eysenck Personality Questionnaire-Revised Short Form (EPQ-R Short Form).<sup>43</sup> As a measure of psychological distress, UKB participants undertook the Patient Health Questionnaire, which is a multipurpose instrument for screening, diagnosing, monitoring and measuring the severity of depression. As in GS:SFHS, the scores were scaled to a mean of zero and standard deviation of 1.

#### Genotyping and polygenic profiling in UKB

The details of the array design, genotyping, quality control and imputation have been described previously.<sup>44</sup> Quality control included removal of participants based on missingness, relatedness, gender mismatch and non-British ancestry. Polygenic profiles were created for SCZ in all the genotyped participants using PRSice.<sup>45</sup> PRSice calculates the sum of alleles associated with the phenotype of interest across many genetic loci,

weighted by their effect sizes estimated from a genome-wide association study of that phenotype in an independent sample. Before creating the scores, the SNPs with a minor allele frequency < 1% were removed and clumping was used to obtain SNPs in linkage equilibrium with an  $r^2 < 0.25$  within a 200 bp window. Multiple scores were then created for each phenotype containing SNPs selected according to the significance of their association with the phenotype. The genome-wide association studies summary data were used to create polygenic profiles for SCZ in the UKB participants, at thresholds of  $P < 0.01$ ,  $P < 0.05$ ,  $P < 0.1$ ,  $P < 0.5$  and all SNPs. As in the main GS:SFHS analysis, current results focus on  $P < 0.5$ .

### Statistical analysis

All the analyses were conducted in ASReml-R ([www.vsnr.co.uk/software/asrem](http://www.vsnr.co.uk/software/asrem), version 3.0). The associations were examined between PRSs for SCZ and cognition (cognitive factor 'g', and contributing tests), psychological distress (GHQ total) and neuroticism (from the Eysenck personality questionnaire) using mixed linear model association analysis in GS:SFHS. Age, age<sup>2</sup>, sex, four-multidimensional scaling ancestry components and polygenic profile scores were entered as fixed effects. As GS:SFHS is a family-based study, to control for relatedness family structure was fitted as a random effect by creating an inverse relationship matrix using pedigree kinship information. SCZ-PRS and MDD status were entered as main effects in the model and PRS  $\times$  MDD status effects were examined by including the interaction term. Where significant interactions were found, further tests were conducted in controls and MDD groups separately. Wald's conditional F-test was used to calculate the significance of fixed effects. The  $P$ -values presented are raw  $P$ -values uncorrected for multiple testing, results were considered significant if they were reported in both the cohorts ( $P < 0.05$ ). The reported  $\beta$ -values are standardized values. The proportion of phenotypic variance explained by PRS was calculated by multiplying the profile score by its corresponding regression coefficient and estimating its variance. This value was then divided by the variance of the observed phenotype to yield a coefficient of determination between 0 and 1 and converted into a percentage.<sup>46</sup>

Significant association and interaction effects were also explored within the independent sample UKB where equivalent measures were available. In UKB, the polygenic risk profile scores were examined for their association with observed phenotypes and for group interaction effects in ASReml-R using the same methods as above (details in Supplementary Material), but without the inclusion of a genetic relationship matrix owing to the large data set and unrelated nature of the filtered UKB study population used in the current investigation.

## RESULTS

### Demographic, cognitive, trait-related features of MDD and PRS scores in GS:SFHS

The demographic details are presented in Table 1 for GS:SFHS. The groups defined according to depression status were significantly different in terms of age, sex, for the cognitive measures of logical memory and verbal fluency, with the greatest differences between groups for measures of psychological distress (GHQ) and for neuroticism (Table 1), higher in MDD cases. The MDD cases and controls also differed significantly on the SCZ PRS score, with the MDD individuals scoring higher.

The demographic details of the individuals from UKB included in the current study are presented in Supplementary Table 1. There were significant differences between the groups in terms of age and gender. The groups also differed according to psychological distress and neuroticism, with the MDD cases scoring higher than controls. Differences between the groups did not reach significance for the individual cognitive tests, and the MDD cases and controls differed significantly on the SCZ PRS score, with the MDD individuals scoring higher.

### Interaction between MDD status and polygenic risk for schizophrenia, in GS:SFHS and UKB

The results for tests of interaction are presented in Table 2 and Figures 1a and b. Significant PRS SCZ  $\times$  group interactions effects were seen across both cohorts for measures of psychological

distress ( $\beta_{GS} = -0.0402$ ,  $P_{GS} = 1.44 \times 10^{-2}$ ;  $\beta_{UKB} = -0.0903$ ,  $P_{UKB} = 8.03 \times 10^{-4}$ , GS:SFHS and UKB, respectively, Figure 1) and for neuroticism ( $\beta_{GS} = -0.0483$ ,  $P_{GS} = 1.96 \times 10^{-3}$ ;  $\beta_{UKB} = -0.0602$ ,  $P_{UKB} = 2.28 \times 10^{-2}$ ). A significant interaction was also reported for measures of processing speed in GS:SFHS ( $\beta_{GS} = -0.0313$ ,  $P_{GS} = 4.31 \times 10^{-2}$ ), but this was not replicated in the UKB sample.

In terms of the direction of associations within the diagnostic groups, in both GS:SFHS and UKB samples, there was a significant positive association in the control group for measures of psychological distress ( $\beta_{GS} = 0.0400$ ,  $P_{GS} = 1.46 \times 10^{-8}$ ;  $\beta_{UKB} = 0.0827$ ,  $P_{UKB} = 3.25 \times 10^{-14}$ ) and for neuroticism ( $\beta_{GS} = 0.0369$ ,  $P_{GS} = 9.15 \times 10^{-6}$ ;  $\beta_{UKB} = 0.0539$ ,  $P_{UKB} = 4.79 \times 10^{-19}$ , GS:SFHS and UKB; see Table 3). There were no significant associations within the MDD groups for either measure, in either cohort (see Table 3, Figures 1 a and b).

In terms of psychological distress, the proportion of variance explained was 0.15 and 0.22% for control individuals, for GS:SFHS and UKB, respectively, and 0.02%, 0.01% for MDD cases. For measures of neuroticism, values were 0.17 and 0.30% in the controls and < 0.01% in the MDD cases (see Table 3 and Figure 2).

For both distress and neuroticism, the SCZ-PRS  $\times$  MDD status interactions remained significant after modelling if the individuals were experiencing a current depressive episode at the time of assessment according to the SCID, and remained significant when modelling if individuals were taking antidepressant medication at the time of assessment according to self-report; full details in Supplementary Material.

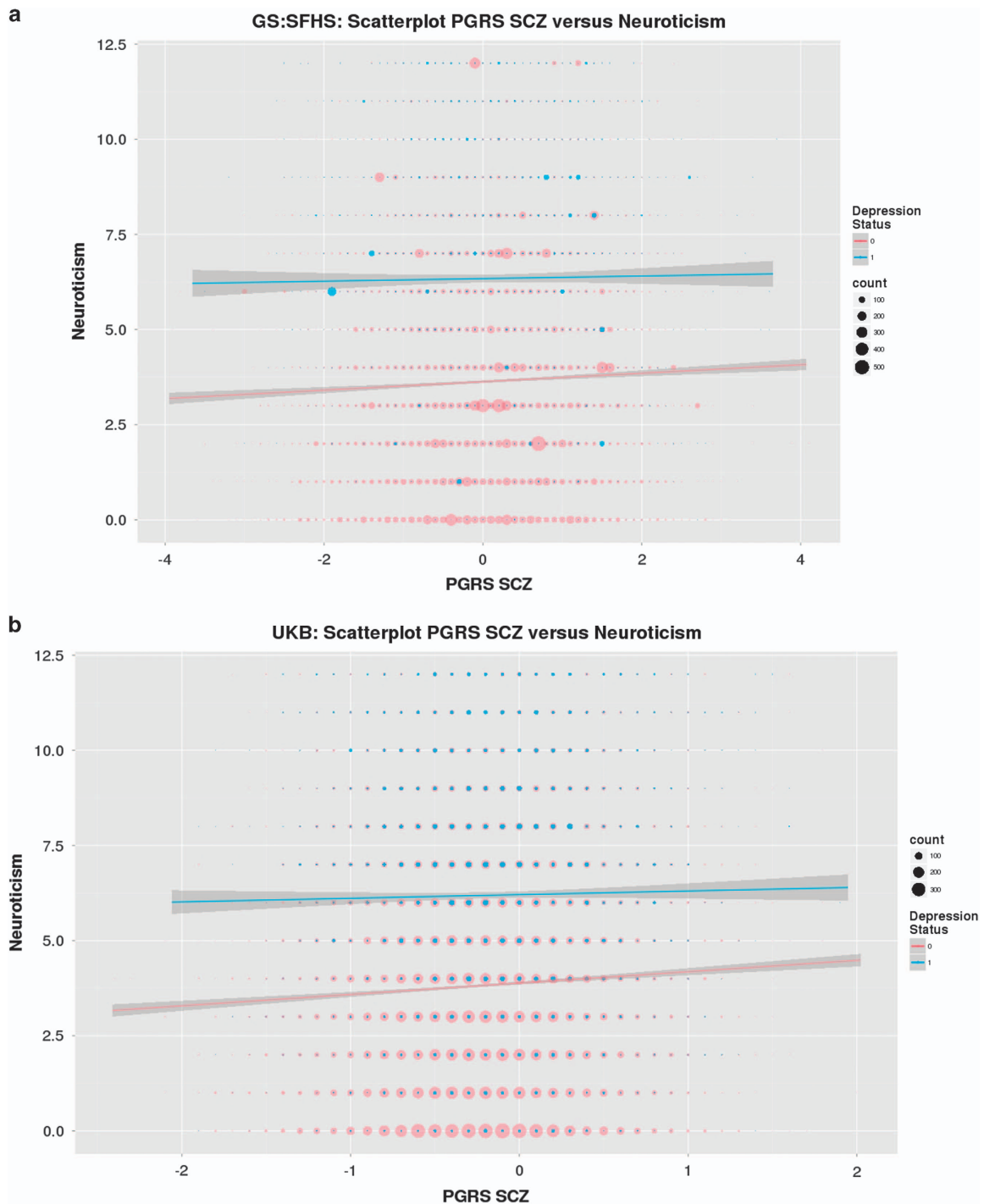
## DISCUSSION

Consistent with our predictions, we report MDD case-control differences in neuroticism and distress that were attenuated by higher SCZ PRS scores, found in both the cohorts. We also demonstrate that the greatest variance explained for neuroticism and distress was higher in controls than in the MDD cases, supporting the suggestion that the MDD samples contain a more heterogeneous group comprising differing aetiological subtypes. Although we also report larger MDD case-control differences with respect to cognition in the context of higher SCZ PRS scores in GS:SFHS, this was not replicated in the independent UKB cohort.

### PRS SCZ associations and interactions with clinical and trait features

Neuroticism is considered to be a stable heritable trait<sup>47</sup> characterized by high tension, irritability, dissatisfaction, shyness, low mood and reduced self-confidence.<sup>38</sup> Psychological distress is a more generalized measure of psychiatric well-being or psychological health than a specific psychopathological categorization. These two traits have been reported to share strong genetic links,<sup>7</sup> and there has been considerable support for overlapping genetic risk factors affecting neuroticism and MDD.<sup>48</sup> Direct links between genetic risk for schizophrenia and these measures, however, are less commonly reported than associations with cognitive deficits, as described above.

These measures demonstrated significant interaction effects between SCZ-PRS and MDD case-control status, seen in both GS:SFHS and UKB cohorts. From Figure 1, it can be seen that for controls, neuroticism scores were positively correlated with SCZ PRSs, which has indeed been previously reported.<sup>49</sup> However, for MDD cases, the regression line indicated that neuroticism scores did not depend on PRS SCZ. Similar patterns were observed for psychological distress. We are not aware of this relationship having been reported previously. In addition, the variance explained for these measures was greater in control individuals (~0.20%) than in the MDD cases (~0.01%). We suggest that this may be attributable to the aetiologically diverse subgroups within the MDD cases. The origins of these interactions were an



**Figure 1.** (a and b) Scatterplot of SCZ PRS and neuroticism according to MDD status in GS:SFHS, and UK Biobank (UKB), respectively. The size of circles represents the degree of overlapping data points. For clarity, histograms of the distribution of neuroticism scores in cases and controls for both cohorts are presented in the Supplementary Figures 1a and b. GS:SFHS, Generation Scotland: Scottish Family Health Study; MDD, major depressive disorder; PRS, polygenic risk score (PGRS in the figure); SCZ, schizophrenia.

attenuation of case–control differences on a background of higher genetic risk of SCZ, seen in both cohorts. Neuroticism and psychological distress may therefore be less closely related to the aetiology of depression in the context of high genetic risk for schizophrenia (see Figure 1). It is, however, also possible that the absence of a similar gradient of effect in the MDD cases could be owing to the substantial effects of depression on distress and neuroticism, leading to a ceiling effect whereby the additional

effect of SCZ PRS on these phenotypes is undetectable. However, we consider that this is unlikely as the variance of these measures are not in fact smaller in the MDD groups in comparison with the control group (which would be the case if there were ceiling effects, see Table 1), and so truncation of these scores at the extreme end seems unlikely. Together, these findings suggest that depression on a background of high genetic risk for schizophrenia may represent a somewhat causally distinct form of the condition.

**Table 3.** Associations between PRSs for SCZ with cognitive, clinical and trait-related features of MDD in cases/controls separately

Generation Scotland: SFHS						
	Controls (n = 16 764)			MDD cases (n = 2587)		
	β	P-value	R <sup>2</sup>	β	P-value	R <sup>2</sup>
<i>Cognitive</i>						
Composite 'g' factor	-0.0773	2.61 × 10 <sup>-17</sup>	0.3470	-0.0906	6.50 × 10 <sup>-5</sup>	0.6671
Logical memory	-0.0600	2.33 × 10 <sup>-13</sup>	0.3119	-0.0764	3.00 × 10 <sup>-5</sup>	0.8353
Digit symbol	<b>-0.0654</b>	<b>3.27 × 10<sup>-20</sup></b>	<b>0.3258</b>	<b>-0.0906</b>	<b>7.35 × 10<sup>-8</sup></b>	<b>1.0900</b>
Mill Hill vocabulary	-0.0251	1.36 × 10 <sup>-3</sup>	0.0379	-0.0131	5.01 × 10 <sup>-1</sup>	0.0173
Verbal fluency	-0.0152	6.70 × 10 <sup>-2</sup>	0.0190	-0.0064	7.54 × 10 <sup>-1</sup>	0.0110
<i>Clinical and trait-related features</i>						
Psycholog distress (GHQ)	<b>0.0400</b>	<b>1.46 × 10<sup>-8</sup></b>	<b>0.1487</b>	<b>0.0111</b>	<b>6.99 × 10<sup>-1</sup></b>	<b>0.0149</b>
Neuroticism	<b>0.0369</b>	<b>91.15 × 10<sup>-6</sup></b>	<b>0.1661</b>	<b>-0.0002</b>	<b>9.91 × 10<sup>-1</sup></b>	<b>0.0009</b>
UK Biobank						
	Controls (n = 27 476)			MDD cases (n = 6049)		
	β; z ratio	P-value	R <sup>2</sup>	β; z ratio	P-value	R <sup>2</sup>
<i>Cognitive</i>						
Composite 'g' factor	-0.1203	6.99 × 10 <sup>-13</sup>	0.3156	-0.1696	4.12 × 10 <sup>-9</sup>	0.6162
Memory	0.0657	1.66 × 10 <sup>-10</sup>	0.1536	0.0490	2.56 × 10 <sup>-2</sup>	0.0863
Symbol digit substitution	-0.1006	2.79 × 10 <sup>-7</sup>	0.2816	-0.0785	5.00 × 10 <sup>-2</sup>	0.1802
Reaction time	0.0390	4.79 × 10 <sup>-4</sup>	0.0422	0.0509	3.68 × 10 <sup>-2</sup>	0.0692
Verbal-numerical reasoning	-0.1101	2.16 × 10 <sup>-21</sup>	0.3467	-0.1390	2.88 × 10 <sup>-8</sup>	0.5495
<i>Clinical and trait-related features</i>						
PHQ	<b>0.0827</b>	<b>3.25 × 10<sup>-14</sup></b>	<b>0.2218</b>	<b>-0.0259</b>	<b>4.20 × 10<sup>-1</sup></b>	<b>0.0111</b>
Neuroticism	<b>0.0539</b>	<b>4.79 × 10<sup>-19</sup></b>	<b>0.3026</b>	<b>0.0055</b>	<b>6.82 × 10<sup>-1</sup></b>	<b>0.0029</b>

Abbreviations: GHQ, General Health Questionnaire; MDD, major depressive disorder; PHQ, Patient Health Questionnaire; PRS, polygenic risk score; SCZ, schizophrenia; SFHS, Scottish Family Health Study. R<sup>2</sup> represents estimate of variance in trait explained by polygene score in percentage. Significant interactions are highlighted in bold.



**Figure 2.** Bar chart of percentage variance explained for distress and neuroticism in GS:SFHS and UK Biobank (UKB). GS:SFHS, Generation Scotland: Scottish Family Health Study; MDD, major depressive disorder; PGRS, polygenic risk score; SCZ, schizophrenia.

PRS SCZ associations and interactions with cognition

Previous studies have consistently reported a negative relationship between cognitive function and SCZ PRS, indicating shared genetic risk between SCZ and deficits in cognition.<sup>24–26,50</sup> With the exception of one previous study where genetic loading for SCZ

(based on proximal family history) was reported to have a substantial negative impact on neurocognition in mood disorder patients,<sup>51</sup> we are not aware of any other studies examining this relationship in the context of MDD. In the current study, the data from the four individual subtests that contributed to the general

factor indicated a significant interaction between SCZ-PRS and MDD case–control status for the digit symbol coding test. This test is a measure of processing speed, which is known to be significantly impaired in SCZ patients, even in relation to other neuropsychological measures.<sup>52,53</sup> Within GS:SFHS, significant negative associations between SCZ-PRS and processing speed were reported in both controls and the MDD cases. In other words, the individuals in GS:SFHS with MDD demonstrated greater deficits in tests of processing speed in those who had higher genetic loading for SCZ versus those at low risk. However, the lack of formal replication means these results should be viewed with caution.

We are aware of one other recent large study using PRSs, including those for SCZ, to determine potential subtypes within MDD.<sup>12</sup> This study took the approach of determining subgroup classification based on symptom profiles (typical/atypical) and then compared groups using various PRS measures. They reported that typical MDD had a greater genetic overlap with PRS SCZ than atypical MDD, which was more closely related to cardiometabolic PRS such as body mass index and triglycerides.<sup>12</sup> Although similar, it is however difficult to make direct comparisons between this and the current study, since we used the opposite approach of categorizing according to PRS SCZ and then examined associations with features such as cognition, personality features and psychological distress, which were not described in the Milanese *et al.*<sup>12</sup> study.

#### Limitations

Although this is a large and arguably well-powered study with independent replication, there are important limitations that should be considered. Of note is that the data used to derive polygenic scores do not presently fully account for all heritability attributable to common variation, nor do these measures reflect the contribution of rare variation and copy-number variants. In addition, the degree of variance explained by the SCZ PRS score for these phenotypes is relatively modest (0.5–1%). It should be considered, however, that they are consistent with other previously reported figures of 0–2% for similar phenotypes and in similarly unselected population-based cohorts, and that the higher figures reported in the literature (7–18%) generally relate to the phenotype of SCZ case–control status (see ref. 26).

Another limitation relates to consistency of testing between these large data sets. Particularly relevant are the differences in administration of the tests of processing speed. Within GS:SFHS, a pen and paper version of the Digit Symbol Coding task from the Wechsler Adult Intelligence Scale III was used, where participants had a 2 min time limit to complete the task. In UKB, the format of the task was a computerized version where the participant had to click on the number that matched the symbol shown and the participant had 1 min to complete the test. Importantly, there were also fewer individuals included in the DSC assessments in UKB ( $n=7799$  controls and  $n=1767$  MDD individuals) than for other cognitive measures, and in relation to GS:SFHS ( $n=16\,566$  controls and  $n=2566$  MDD individuals) as this test was conducted at a follow-up assessment rather than at baseline.

The approach we have currently applied consisted of presenting two separate analyses for GS:SFHS and UKB, declaring significance when both the studies reported findings with  $P < 0.05$ . However, we further note that a more stringent approach would be to apply multiple comparison correction within each cohort separately. Applying such a correction for the summary cognitive measure and measures of clinical trait features, the main interactions for distress and neuroticism remained significant in both the cohorts (GS:SFHS,  $P_{\text{FDR corrected}} = 0.0216$ ,  $P_{\text{FDR corrected}} = 0.0059$ ; in UKB  $P_{\text{FDR corrected}} = 0.0024$ ,  $P_{\text{FDR corrected}} = 0.0342$ , for distress and neuroticism, respectively (FDR, false discovery rate)).

In summary, we consider the main novel finding is not just the significant relationship between PRS SCZ and neuroticism in controls, but the lack of that clear relationship in MDD, inferring heterogeneity. These findings are consistent with a model in which genetic risk for schizophrenia predicts depressive traits in the general population, but that neuroticism and psychological distress may be less closely related to the aetiology of depression on a background of high genetic risk for schizophrenia. This may represent a somewhat causally distinct form of MDD more closely related to SCZ. The study of the genetic basis of variation in such measures is likely to further the understanding of mechanisms by which SCZ genes affect neural function in the context of health and depressive illness.

#### CONFLICT OF INTEREST

AMM has previously received grant support from Pfizer, Lilly and Janssen. These studies are not connected to the current investigation. SML has previously received research grant support from Pfizer, Roche, Abbvie and Sunovion, as well as personal fees from Roche, Sunovion and Janssen, outside the present work. The remaining authors declare no conflict of interest.

#### ACKNOWLEDGMENTS

This study is supported by a Wellcome Trust Strategic Award 'Stratifying Resilience and Depression Longitudinally' (STRADL; Reference 104036/Z/14/Z). Generation Scotland (GS:SFHS) received core support from the Chief Scientist Office (CSO) of the Scottish Government Health Directorates (grant number CZD/16/6) and the Scottish Funding Council (HR03006). Genotyping of the GS:SFHS samples was carried out by the Genetics Core Laboratory at the Wellcome Trust Clinical Research Facility, Edinburgh, Scotland and was funded by the Medical Research Council of UK and the Wellcome Trust. Ethics approval for the study was given by the NHS Tayside committee on research ethics (reference 05/S1401/89). We are grateful to all the families who took part, the general practitioners and the Scottish School of Primary Care for their help in recruiting them, and the whole Generation Scotland team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, healthcare assistants and nurses. This research has also been conducted using the UKB Resource. HCW is supported by a JMAs SIM fellowship from the Royal College of Physicians of Edinburgh and by an ESAT College Fellowship from the University of Edinburgh. Part of the work was undertaken in The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology (CCACE), part of the cross council Lifelong Health and Wellbeing Initiative (MR/K026992/1); funding from the Biotechnology and Biological Sciences Research Council (BBSRC) and Medical Research Council (MRC) is gratefully acknowledged. Age UK (The Disconnected Mind project) also provided support for the work undertaken at CCACE.

#### REFERENCES

- 1 Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR *et al*. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA* 2003; **289**: 3095–3105.
- 2 Kennedy SH. Core symptoms of major depressive disorder: relevance to diagnosis and treatment. *Dialogues Clin Neurosci* 2008; **10**: 271–277.
- 3 Austin MP, Mitchell P, Goodwin GM. Cognitive deficits in depression: possible implications for functional neuropathology. *Br J Psychiatry* 2001; **178**: 200–206.
- 4 Navrady LB, Ritchie SJ, Chan SWY, Kerr DM, Adams MJ, Hawkins E *et al*. Intelligence and neuroticism in relation to depression and psychological distress: evidence of interaction using data from Generation Scotland: Scottish Family Health Study and UK Biobank. *bioRxiv* 2016; <http://biorxiv.org/content/early/2016/02/12/039545>.
- 5 Glahn DC, Curran JE, Winkler AM, Carless MA, Kent JW Jr, Charlesworth JC *et al*. High dimensional endophenotype ranking in the search for major depression risk genes. *Biol Psychiatry* 2012; **71**: 6–14.
- 6 Kendler KS, Gatz M, Gardner CO, Pedersen NL. Personality and major depression: a Swedish longitudinal, population-based twin study. *Arch Gen Psychiatry* 2006; **63**: 1113–1120.
- 7 Luciano M, Huffman JE, Arias-Vasquez A, Vinkhuyzen AA, Middeldorp CM, Giegling I *et al*. Genome-wide association uncovers shared genetic effects among personality traits and mood states. *Am J Med Genet B Neuropsychiatr Genet* 2013; **159B**: 684–695.
- 8 Middeldorp CM, de Moor MH, McGrath LM, Gordon SD, Blackwood DH, Costa PT *et al*. The genetic association between personality and major depression or

- bipolar disorder. A polygenic score analysis using genome-wide association data. *Transl Psychiatry* 2011; **1**: e50.
- 9 Barch DM, Sheffield JM. Cognitive impairments in psychotic disorders: common mechanisms and measurement. *World Psychiatry* 2014; **13**: 224–232.
  - 10 Bagby RM, Bindseil KD, Schuller DR, Rector NA, Young LT, Cooke RG et al. Relationship between the five-factor model of personality and unipolar, bipolar and schizophrenic patients. *Psychiatry Res* 1997; **70**: 83–94.
  - 11 Van Os J, Jones PB. Neuroticism as a risk factor for schizophrenia. *Psychol Med* 2001; **31**: 1129–1134.
  - 12 Milaneschi Y, Lamers F, Peyrot WJ, Abdellaoui A, Willemsen G, Hottenga JJ et al. Polygenic dissection of major depression clinical heterogeneity. *Mol Psychiatry* 2016; **21**: 516–522.
  - 13 Lyons MJ, Eisen SA, Goldberg J, True W, Lin N, Meyer JM et al. A registry-based twin study of depression in men. *Arch Gen Psychiatry* 1998; **55**: 468–472.
  - 14 Zisook S, Lesser I, Stewart JW, Wisniewski SR, Balasubramani GK, Fava M et al. Effect of age at onset on the course of major depressive disorder. *Am J Psychiatry* 2007; **164**: 1539–1546.
  - 15 Zisook S, Rush AJ, Lesser I, Wisniewski SR, Trivedi M, Husain MM et al. Preadult onset vs. adult onset of major depressive disorder: a replication study. *Acta Psychiatr Scand* 2007; **115**: 196–205.
  - 16 Power RA, Tansey KE, Buttenschön HN, Cohen-Woods S, Bigdeli T, Hall LS et al. Genome-wide association for major depression through age at onset stratification. *Biol Psychiatry* 2016 (in press).
  - 17 Power RA, Keers R, Ng MY, Butler AW, Uher R, Cohen-Woods S et al. Dissecting the genetic heterogeneity of depression through age at onset. *Am J Med Genet B Neuropsychiatr Genet* 2016; **159B**: 859–868.
  - 18 McGue M, Christensen K. Genetic and environmental contributions to depression symptomatology: evidence from Danish twins 75 years of age and older. *J Abnorm Psychol* 1997; **106**: 439–448.
  - 19 Sullivan PF, Neale MC, Kendler KS. Genetic epidemiology of major depression: review and meta-analysis. *Am J Psychiatry* 2000; **157**: 1552–1562.
  - 20 Lubke GH, Hottenga JJ, Walters R, Laurin C, de Geus EJ, Willemsen G et al. Estimating the genetic variance of major depressive disorder due to all single nucleotide polymorphisms. *Biol Psychiatry* 2012; **72**: 707–709.
  - 21 Smoller JW, Craddock N, Kendler K, Lee PH, Neale BM, Nurnberger JI et al. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* 2013; **381**: 1371–1379.
  - 22 Hong Lee S, Ripke S, Consortium CDGotPG. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet* 2013; **45**: 984–994.
  - 23 Han B, Pouget JG, Slowikowski K, Stahl E, Lee HL, Diogo D et al. Using genotype data to distinguish pleiotropy from heterogeneity: deciphering coheritability in autoimmune and neuropsychiatric diseases. *bioRxiv* 2016.
  - 24 Lencz T, Knowles E, Davies G, Guha S, Liewald DC, Starr JM 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* 2014; **19**: 168–174.
  - 25 McIntosh AM, Gow A, Luciano M, Davies G, Liewald DC, Harris SE et al. Polygenic risk for schizophrenia is associated with cognitive change between childhood and old age. *Biol Psychiatry* 2013; **73**: 938–943.
  - 26 Liebers DT, Pirooznia M, Seiffudin F, Musliner KL, Zandi PP, Goes FS. Polygenic risk of schizophrenia and cognition in a population-based survey of older adults. *Schizophr Bull* 2016; **42**: 984–991.
  - 27 Jones HJ, Stergiakouli E, Tansey KE, Hubbard L, Heron J, Cannon M et al. Phenotypic manifestation of genetic risk for schizophrenia during adolescence in the general population. *JAMA Psychiatry* 2016; **73**: 221–228.
  - 28 Fanous AH, Kendler KS. The genetic relationship of personality to major depression and schizophrenia. *Neurotox Res* 2004; **6**: 43–50.
  - 29 Smith BH, Campbell A, Linksted P, Fitzpatrick B, Jackson C, Kerr SM et al. Cohort profile: Generation Scotland: Scottish Family Health Study (GS:SFHS). The study, its participants and their potential for genetic research on health and illness. *Int J Epidemiol* 2012; **42**: 689–700.
  - 30 Smith BH, Campbell H, Blackwood D, Connell J, Connor M, Deary IJ et al. Generation Scotland: the Scottish Family Health Study; a new resource for researching genes and heritability. *BMC Med Genet* 2006; **7**: 74.
  - 31 Kerr SM, Campbell A, Murphy L, Hayward C, Jackson C, Wain LV et al. Pedigree and genotyping quality analyses of over 10000 DNA samples from the Generation Scotland: Scottish Family Health Study. *BMC Med Genet* 2013; **14**: 38.
  - 32 First MB, Spitzer RL, Miriam G, JBW Williams. *Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition (SCID-I/P)*. Biometrics Research, New York State Psychiatric Institute: New York, NY, USA, 2002.
  - 33 Weschler D. *WAIS-III Administration and Scoring Manual*. Psychological Corporation: London, UK, 1998.
  - 34 Raven J. *The Mill Hill vocabulary scale*. H.K.Lewis: London, UK, 1965.
  - 35 Lezak MD. *Neuropsychological Assessment*. Oxford University Press: New York, NY, USA, 1995.
  - 36 Johnson W, Bouchard TJJ, Krueger RF, McGue M, Gottesman II. Just one g: consistent results from three test batteries. *Intelligence* 2004; **32**: 95–107.
  - 37 Goldberg DP, Hillier VF. A scaled version of the General Health Questionnaire. *Psychol Med* 1979; **9**: 139–145.
  - 38 Eysenck HJ, Eysenck SBG. *Manual for the Eysenck personality inventory*. University of London Press: London, UK, 1964.
  - 39 Gunderson KL. Whole-genome genotyping on bead arrays. *Methods Mol Biol* 2009; **529**: 197–213.
  - 40 Clarke TK, Hall LS, Fernandez-Pujals AM, MacIntyre DJ, Thomson P, Hayward C et al. Major depressive disorder and current psychological distress moderate the effect of polygenic risk for obesity on body mass index. *Transl Psychiatry* 2015; **5**: e592.
  - 41 Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**: 748–752.
  - 42 Allen NE, Sudlow C, Peakman T, Collins R. UK biobank data: come and get it. *Sci Transl Med* 2014; **6**: 224ed4.
  - 43 Deary IJ, Bedford A. Some origins and evolution of the EPQ-R (short form) neuroticism and extraversion items. *Pers Individ Dif* 2011; **50**: 1213–1217.
  - 44 Hagenars SP, Harris SE, Davies G, Hill WD, Liewald DC, Ritchie SJ et al. Shared genetic aetiology between cognitive functions and physical and mental health in UK Biobank (N = 112 151) and 24 GWAS consortia. *Mol Psychiatry* 2016.
  - 45 Euesden J, Lewis CM, O'Reilly PF. PRSice: Polygenic Risk Score software. *Bioinformatics* 2015; **31**: 1466–1468.
  - 46 Nakagawa S, Schielzeth H. A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods Ecol Evol* 2013; **4**: 133–142.
  - 47 Bouchard TJ Jr. Genes, environment, and personality. *Science* 1994; **264**: 1700–1701.
  - 48 de Moor MH, van den Berg SM, Verweij KJ, Krueger RF, Luciano M, Arias Vasquez A et al. Meta-analysis of genome-wide association studies for neuroticism, and the polygenic association with major depressive disorder. *JAMA Psychiatry* 2015; **72**: 642–650.
  - 49 Gale CR, Hagenars SP, Davies G, Hill WD, Liewald DC, Cullen B et al. Pleiotropy between neuroticism and physical and mental health: findings from 108 038 men and women in UK Biobank. *Transl Psychiatry* 2016; **6**: e791.
  - 50 Hatzimanolis A, Bhatnagar P, Moes A, Wang R, Roussos P, Bitsios P et al. Common genetic variation and schizophrenia polygenic risk influence neurocognitive performance in young adulthood. *Am J Med Genet B Neuropsychiatr Genet* 2015; **168**: 392–401.
  - 51 Hori H, Matsuo J, Teraishi T, Sasayama D, Kawamoto Y, Kinoshita Y et al. Schizotypy and genetic loading for schizophrenia impact upon neuropsychological status in bipolar II and unipolar major depressive disorders. *J Affect Disord* 2012; **142**: 225–232.
  - 52 Zaninotto L, Guglielmo R, Calati R, Ioime L, Camardese G, Janiri L et al. Cognitive markers of psychotic unipolar depression: a meta-analytic study. *J Affect Disord* 2015; **174**: 580–588.
  - 53 Dickinson D, Ramsey ME, Gold JM. Overlooking the obvious: a meta-analytic comparison of digit symbol coding tasks and other cognitive measures in schizophrenia. *Arch Gen Psychiatry* 2007; **64**: 532–542.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016

Supplementary Information accompanies the paper on the Translational Psychiatry website (<http://www.nature.com/tp>)