bioRxiv preprint doi: https://doi.org/10.1101/059352; this version posted June 20, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Do Regional Brain Volumes and Major Depressive Disorder Share Genetic Architecture: a study in Generation Scotland (n=19,762), UK Biobank (n=24,048) and the English Longitudinal Study of Ageing (n=5,766)

Eleanor M. Wigmore (BSc)[1*], Toni-Kim Clarke (PhD)[1], Mark J. Adams (PhD)[1], Ana M. Fernandez-Pujals (MSc)[1], Jude Gibson (BSc)[1], Gail Davies (PhD)[2], Lynsey S. Hall (BSc)[1], Yanni Zeng (MSc)[1], Pippa A. Thomson (PhD)[2,3], Caroline Hayward (PhD)[3], Blair H. Smith (MD)[4], Lynne J. Hocking (PhD)[5], Sandosh Padmanabhan (PhD)[6], Ian J. Deary (PhD)[2,7], David J. Porteous (PhD)[3], Kristin K. Nicodemus (PhD)[2,3], Andrew M. McIntosh (MD)[1,2]

1. Division of Psychiatry, University of Edinburgh, Royal Edinburgh Hospital, Edinburgh, UK, EH10 5HF
2. Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, 7 George Square, Edinburgh, UK, EH8 9JZ
3. Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Crewe Road South, Edinburgh, UK, EH4 2XU
4. Division of Population Health Sciences, University of Dundee, Dundee, UK
5. Division of Applied Medicine, University of Aberdeen, Aberdeen, UK
6. Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK
7. Department of Psychology, University of Edinburgh, UK

*Corresponding author

Eleanor M. Wigmore
Division of Psychiatry
University of Edinburgh
Royal Edinburgh Hospital
Edinburgh EH10 5HD
+44 (0)131 537 6687
e.m.wigmore@sms.ed.ac.uk

Number of words (main text)	3,292
Number of references	48
Number of Tables	2
Number of Figures	2

bioRxiv preprint doi: https://doi.org/10.1101/059352; this version posted June 20, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Major depressive disorder (MDD) is a highly debilitating and heritable disorder. It is commonly associated with subcortical volumetric abnormalities, the most replicated of these being reduced hippocampal volume. Using the most recent published data from ENIGMA consortium's genome-wide association study (GWAS) of regional brain volume, we sought to test whether there is shared genetic architecture between subcortical brain volumes and MDD. Using LD score regression utilising summary statistics from ENIGMA and the Psychiatric Genomics Consortium, we demonstrated that hippocampal volume was genetically correlated with MDD (r_G=0.46, *P*=0.02), although this did not survive multiple comparison testing. None of other six regions were genetically correlated and amygdala volume heritability was too low for analysis. We also generated polygenic risk scores (PRS) to assess potential pleiotropy on regional brain volumes and MDD in three cohorts (Generation Scotland; Scottish Family Health Study (n=19,762), UK Biobank (n=24,048) and the English Longitudinal Study of Ageing (n=5,766)). We used logistic regression to examine volumetric PRS and MDD and performed a meta-analysis across the three cohorts. No regional volumetric PRS demonstrated any significant association with lifetime MDD or recurrent MDD. In this study we provide evidence that hippocampal volume and MDD have a shared genetic architecture providing further evidence of the potential mechanistic importance of the hippocampus in depression. We found no evidence to support a shared genetic architecture for MDD with any other subcortical region.

Keywords

Subcortical volumes, major depressive disorder, polygenic risk scores, genetic correlation, LD score regression, hippocampal volume, shared genetic architecture

bioRxiv preprint doi: https://doi.org/10.1101/059352; this version posted June 20, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

Major depressive disorder (MDD) is a debilitating disorder that accounts for a large proportion of disease burden world-wide (1). It is a complex disorder that is influenced by both genetic and environmental factors with a heritability of approximately 37% estimated from twin studies (2). A recent genome-wide association study (GWAS) identified two loci of genome-wide significance in MDD (3). Nevertheless, the aetiology of the disorder has largely remained elusive and its heritability unaccounted for by currently identified variants.

Reports of lower brain volumes in cross-sectional studies are common in MDD but small sample sizes have led to poorly replicated results. Enhancing Neuro-Imaging Genetics through Meta-Analysis (ENIGMA) completed a large MDD case-control meta-analysis of subcortical volumes (n=8,927) demonstrating a significant association with reduced hippocampal volume (4). Numerous other studies have also demonstrated a link between hippocampal reduction and MDD and it is one of the most robustly associated regions (5). Other brain regions have shown limited and sometimes contradictory evidence for association with MDD. Smaller amygdala volume has been associated with depressive symptoms (6, 7) and MDD status (8), however larger amygdala volume has also been associated with the disorder (9). A 2013 meta-analysis concluded that, as well as hippocampus, smaller putamen and thalamus volumes were associated with late life MDD, although fewer studies have examined these regions (10). Another meta-analysis additionally found an association with smaller caudate nucleus volumes (11). The nucleus accumbens has not been widely associated with MDD status but a smaller volume has been implicated in the lethality of suicidal acts within mood disorder sufferers (12). Pallidum volume and intracranial volume (ICV) have not been associated with MDD in any meta-analysis to date.

bioRxiv preprint doi: https://doi.org/10.1101/059352; this version posted June 20, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Subcortical structures' volumes are known to be influenced by both genetic and environmental factors and have been demonstrated to be highly heritable ranging from 0.44 to 0.88 (13). The previously reported lower brain volumes in MDD and the high heritability of these structures means they could be of interest as an intermediate phenotype (13). A GWAS on regional brain volumes has recently been completed by the ENIGMA Consortium (14), providing an important opportunity to examine the genetic overlap between subcortical brain volumes and MDD. Overlap between genes involved in MDD and subcortical regions have been explored previously. The majority of studies have focused on candidate genes, such as the serotonin transporter (5-HTTLPR) and findings are often contradictory (15). No studies, to our knowledge, have examined the genetic overlap between common risk variants for MDD and subcortical volumes on a genome-wide basis using molecular data.

In this current study, we sought to test the hypothesis that the genetic architecture of MDD is shared with that of multiple subcortical brain regions. We employed two techniques; the first, LD score regression (16, 17), estimates the genetic correlation between these traits using summary statistics from the ENIGMA and PGC consortia. The second method, polygenic risk scoring (18), utilises ENIGMA summary statistics to generate individual level polygenic profile scores of each brain region's volume. We then calculated the association of polygenic scores with MDD status in three cohorts and combined them in a meta-analysis.

Methods

Cohort Descriptions and genotyping

Generation Scotland: Scottish Family Health Study (GS:SFHS)

bioRxiv preprint doi: https://doi.org/10.1101/059352; this version posted June 20, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

GS:SFHS is a family based cohort with phenotypic data for 24,080 (mean age=47.6, s.d.=15.4) and genome-wide data available for 20,032 participants. Recruitment for this cohort has been described previously (19). Diagnosis of MDD was made using the structured clinical interview for DSM-IV disorders (SCID) after screening positive during interview questions (n=19,762, cases=2,643) (20). Bipolar disorder patients (n=76) were excluded from this study.

Details of DNA extraction have been previously described (21). Genotyping was completed at the Wellcome Trust Clinical Research Facility Genetics Core, Edinburgh (www.wtcrf.ed.ac.uk) using the Illumina HumanOmniExpressExome -8v1.0 Beadchip and Infinium chemistry (22) and processed using GenomeStudio Analysis Software v2011.1. Quality Control (QC) utilised the inclusion threshold as follows; missingness per individual <1%, missingness per single nucleotide polymorphism (SNP) <1%, Hardy-Weinberg Equilibrium (HWE) p-value $>1x10^{-6}$, minor allele frequency (MAF) >1%. 556,705 SNPs passed QC criteria.

UK Biobank

UK Biobank is an open resource cohort with phenotypic data for 502,664 (mean age=56.5, s.d.= 8.1) between the ages of 40-69 recruited within the United Kingdom between 2006-2010 and genome-wide data available for 152,734 participants. Our study was conducted under UK Biobank application 4844. Study design and recruitment has been described previously (23) but, in brief, participants were asked to complete a touchscreen questionnaire and additional data was collected by nurse interview. GS:SFHS and related individuals (n=35,752) were excluded from this sample. MDD status was based upon putative MDD phenotype defined by Smith et al., (2013) (24) (n=24,048). Participants with mild depressive

bioRxiv preprint doi: https://doi.org/10.1101/059352; this version posted June 20, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

symptoms were removed based on this definition and self-reported bipolar disorder participants (n=1,211) have been excluded. Subcortical volumes for nucleus accumbens, amygdala, caudate nucleus, hippocampus, pallidum, putamen and thalamus were measured by T1-weighted structural imaging. The UK Biobank imaging protocol has been described elsewhere (http://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=1977). The mean of the sum of left and right volume was taken for each subcortical region. ICV was generated by the sum of white matter, grey matter and ventricular cerebrospinal fluid volumes. Imaging data for the eight structures was available for 4,446 participants of which 968 had genetic data available.

Genotyping was completed utilising two Affymetrix arrays; BiLEVE (n=49,979) and the UK Biobank Axiom (n=102,750). Details have been described previously (25). Initial genotyping QC was performed by UK Biobank (26). Additional filtering was then applied to participants with poor heterozygosity or missingness, QC failure, non-British White ancestry, gender mismatch, genotype missing > 2%, and relatedness within UK Biobank and to the GS:SFHS sample ($r > 0.0442$). SNPs were filtered by HWE $p < 10e-6$ and MAF < 0.01. 731,536 SNPs passed QC criteria.

English Longitudinal Study of Ageing (ELSA)

ELSA is a prospective cohort study of health and ageing collected in 2002 with six follow-up waves taken at two-year intervals. At wave 1 (baseline) phenotypic data were available for 12,003 (mean age=63.9, s.d.=10.7) and genotypic data available for 7,452 participants. Details of this cohort have been described previously (27). MDD status in this study was defined using a shortened form of the Centre of Epidemiological Studies – Depression scale (CES-D scale) (completed by 5,752 participants with genomic data). This consisted of 8 questions, rather than the original 20, with a “no”/”yes” response which was converted to a

bioRxiv preprint doi: https://doi.org/10.1101/059352; this version posted June 20, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

binary 0/1, respectively, although positive questions i.e. "During the past week, were you happy?", were scored in reverse; 0 being "yes" and 1 being "no". After summing the scores, a dummy variable of MDD status was classified as those with a score of 4 or above, as in previous studies (28). Self-reported "manic depressive" (n=41) individuals were excluded.

Genotyping was completed in 2013/14 on 7,452 participants on the Illumina Omni 2.5-8 chip. QC was completed at the University College London Genetics Institute and further using the same inclusion thresholds as used for GS:SFHS with >1.3 million SNPs passing QC criteria.

Linkage disequilibrium (LD) score regression

Genetic correlation of subcortical structures and MDD were measured using the LD score regression technique (16, 17). In brief, this technique utilises GWAS summary statistics to examine heritability of a trait and genetic correlation between traits, in this study we used summary data from ENIGMA and PGC. The sample was pruned on missingness and MAF thresholds.

Summary statistics for the regional brain volume GWAS completed by ENIGMA were downloaded from http://enigma.ini.usc.edu/enigma-vis/. The GWAS was completed on 11,840 participants on eight MRI volumetric measures; nucleus accumbens, amygdala, caudate nucleus, hippocampus, pallidum, putamen, thalamus and ICV (14).

Summary statistics for the MDD GWAS completed by the MDD Working Group of the Psychiatric Genomics Consortium were downloaded from http://www.med.unc.edu/pgc/downloads. The study examined 9,238 MDD cases and 8,039 controls (29).

bioRxiv preprint doi: https://doi.org/10.1101/059352; this version posted June 20, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Polygenic Risk Scoring (PRS)

Construction of PRS was completed in PLINK software (30), which has been previously described (18). Summary statistics were taken from the ENIGMA GWAS (14) (details above) to construct PRS using five *P* value thresholds: 0.01, 0.05, 0.1, 0.5 and 1. All five thresholds are reported in models of subcortical volume PRS predicting their respective volume in UK Biobank and the best predictive threshold was carried forward into models associating MDD status in all three cohorts. The *P* value thresholds carried forward were; nucleus accumbens: 0.01, amygdala: 0.1, caudate nucleus: 0.5, hippocampus: 0.01, ICV: 0.5, pallidum: 0.5, putamen: 0.1 and thalamus: 0.05. Scores for GS:SFHS and UK Biobank were computed on the raw genotypes but in ELSA, PRS were computed on imputed data.

Statistical Analysis

Mixed linear model analyses were completed in ASReml-R (http://www.vsni.co.uk/software/asreml/) for GS:SFHS with MDD status as the dependent variable and volume PRS fitted as the independent variable. The model was adjusted for age, sex and the first four principal components (PCs), to control for population stratification. An additive matrix (expected relatedness derived from pedigree information) was fitted as a random effect to account for the family structure in GS:SFHS. Wald's conditional F-test was used to calculate P values for all fixed effects and the variance explained was calculated by division of the difference in the sum of residual variance and additive genetic effect in the null model (without PRS) with the full model (with PRS). To adjust for the use of linear mixed regression models being applied to a binary dependent variable in a structured dataset, the fixed effects and standard errors from the linear model were transformed utilising a Taylor series approximation from the linear scale to the liability scale (further details on this method

bioRxiv preprint doi: https://doi.org/10.1101/059352; this version posted June 20, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

are given elsewhere) (31). Since hippocampal volumetric differences have been more closely associated with recurrent MDD and early illness onset (4, 32), hippocampus PRS regression analyses were also run with recurrent MDD, number of episodes, MDD duration and age of onset as dependent variables (for further details see Supplementary Materials).

Logistic regression utilising generalised linear models in R version 3.2.3 (www.r-project.org) was used for both UK Biobank and ELSA cohorts. Related individuals were removed in this analysis as one of the assumptions of the generalised linear model is independent observations. Models were adjusted for age, sex and the first 15 PCs and models predicting subcortical structures were also adjusted for ICV. For ELSA, related individuals were excluded and models were adjusted for age, sex and the first four PCs. Hippocampus volume PRS was also examined with recurrent MDD, number of episodes, MDD duration and age of diagnosis for UK Biobank however this data was not available for ELSA (see Supplementary Materials).

In order to increase power, fixed effect meta-analysis, weighted by standard error, of the beta values relating PRS scores to MDD was carried out using the 'meta' package (version 4.3-2) (33) in R.

Results

Genetic correlation

Using LD score regression, we calculated SNP-based heritability estimates for the eight subcortical regions and MDD, utilising summary data from GWAS completed by ENIGMA (14) and PGC (29) respectively. Estimates for subcortical volumes ranged from the nucleus

bioRxiv preprint doi: https://doi.org/10.1101/059352; this version posted June 20, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

accumbens (h^2=0.0855, s.e.=0.0438) to the putamen (h^2=0.297, s.e.=0.051), however amygdala heritability was non-significant, and MDD SNP heritability was calculated at 0.204 (s.e.=0.0386) (**Table 1**). Genetic correlation between each subcortical region and MDD was then calculated, however due to low estimated heritability the amygdala could not be carried forward into this analysis. Hippocampal volume demonstrated significant genetic correlation with MDD (r_G=0.460, s.e.=0.200, *P*=0.0213) (**Table 1**), although this did not survive multiple testing correction (**Supplementary Table S2**). No other subcortical volume was genetically correlated with MDD (**Table 1**).

Polygenic risk score (PRS)

Subcortical PRS were calculated in UK Biobank to examine the variance explained with respect to their own volume. PRS were positively associated with their respective volume in four of the eight structures across the 5 *P* value thresholds; caudate nucleus, ICV, putamen and thalamus; hippocampus was significantly associated at a *P* value threshold of 0.01 only. These results retained significance after multiple test correction across the 5 thresholds, however only raw *P* values have been reported. Nucleus accumbens, amygdala and pallidum PRS did not demonstrate any association with their respective volume. The variance explained by PRS was small for all volumes, with the largest reported in the caudate nucleus (R^2=0.0102, beta=0.117, *P*=0.000119) (**Figure 1** and **Supplementary Table S1**).

Structural PRS were selected at the threshold which best predicted its own volume (nucleus accumbens=0.01, amygdala=0.1, caudate nucleus=0.5, hippocampus=0.01, ICV=0.5, pallidum=0.5, putamen=0.1, thalamus=0.05) and tested for prediction of MDD status. No PRS for any volume was significantly associated with MDD status in any of the cohorts (**Table 2**). In order to increase power we completed a meta-analysis of the three cohorts. No

bioRxiv preprint doi: https://doi.org/10.1101/059352; this version posted June 20, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

heterogeneity was identified in any of the meta-analyses. We found no association between any structural PRS and MDD (**Figure 2a** and **Supplementary Fig. S1**). We examined recurrent MDD in association with hippocampal volume PRS in GS:SFHS and UK Biobank only as we could not obtain episode data in the ELSA cohort; the association was non-significant (OR=0.98, *P*=0.0850) (**Figure 2b**). Further, hippocampal volume PRS was not significantly associated with number of episodes (beta=-0.00390, *P*=0.425), MDD duration (beta=-0.00110, *P*=0.414) or age of onset (beta=0.0142, *P*=0.291) (**Supplementary Fig. S2**).

Discussion

In this study we investigated whether there was evidence of shared genetic architecture between subcortical volumes and MDD. We conducted this analysis in order to test if the reported phenotypic associations between brain volumes and MDD in previous studies were due to shared genetic factors. Genetic correlations between regional brain volume and MDD show that hippocampal volume and MDD are partially influenced by common genetic variants (r_G=0.46, s.e.=0.200, *P*=0.0213), although did not survive correction for multiple testing. No other brain volume showed evidence of shared genetically aetiology with MDD. A meta-analysis of data from three studies, totalling 49,576 individuals including 11,552 cases, found no evidence of association between any regional brain volume PRS and MDD, including the hippocampus. Since previous neuroimaging evidence suggests that decreased hippocampal volumes could occur as a consequence of recurrent depressive episodes and early illness onset (4, 32), we examined hippocampal volume PRS in association with recurrent MDD, number of episodes, MDD duration, and age of onset but, similarly, we did not observe a significant association.

bioRxiv preprint doi: https://doi.org/10.1101/059352; this version posted June 20, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The genetic correlation reported in hippocampal volume is novel, so far as any of the authors are aware, but this finding was not significant after correcting for multiple testing and was not validated by PRS analysis. This apparent discrepancy could be due to several reasons. Firstly, linkage disequilibrium (LD) information is utilised in LD score regression to measure genetic correlation, however in the process of PRS calculation, all SNPs are pruned to a much smaller independent set. Previous simulation studies have demonstrated that predictive capabilities of PRS are greatly enhanced when utilising LD information (34). This implies that LD pruning may be removing causal SNPs and those more closely tagging causal variants, resulting in a loss of information and predictive accuracy. Secondly, each dataset had a different MDD definition; GS:SFHS utilised the SCID, ELSA MDD was defined utilising the CES-D and UK Biobank MDD was generated using self-reported information. Whist the PGC MDD definition most closely matches that of GS:SFHS MDD, the GS:SFHS sample was population based rather than identified from a clinically ascertained samples. The apparent lack of replication using alternative methodologies may therefore be due to a number of factors related to ascertainment differences. Thirdly, PRS have not explained a large amount of variance within their own trait in almost any analysis conducted to date, with the possible exception of schizophrenia. The PRS that explained most variance in hippocampal volume only explained 0.621%. PRS do not conventionally include non-additive and epistatic effects that may be able to explain a greater proportion of total phenotypic variance (35). It may be important to examine the non-additive genetic effects contributing to hippocampal volume in future studies.

We conclude that that the widely replicated hippocampal atrophy demonstrated in MDD is caused partly by shared genetic factors, although did not withstand multiple test correction. Animal models have previously demonstrated that increased stress can drive decreased

bioRxiv preprint doi: https://doi.org/10.1101/059352; this version posted June 20, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

hippocampal neurogenesis (and therefore increased atrophy) (36) and this reduced neurogenesis can lead to depressive-like symptoms (37). Stress is a well-established environmental risk factor associated with MDD (38) and inhibition of glucocorticoid receptors has been shown to normalise hippocampal neurogenesis (39) and relieve symptoms in psychotic major depression (40). Furthermore, increased duration of depression has also been related to more pronounced hippocampal reductions (41). Our results suggest that genes determining hippocampal volume may also be risk factors for MDD. Given the previous literature linking both hippocampal structure and MDD to stress, it is possible that gene-environment interactions (GxE) could further explain their correlation.

Hippocampal volume reductions are also widely associated with other psychiatric disorders such as schizophrenia. A similar analysis examined the genetic correlation between subcortical volumes and schizophrenia finding no significant correlations (42). This is suggestive that the genetic correlation observed could be specific to hippocampal atrophy in MDD. However, these results are only indicative of a genetic correlation between the two traits and further research would be necessary to provide confirmative evidence and the directionality of any causal relationships.

Subcortical volume PRS were not associated with their own volume in three out of the eight structures and was only associated with hippocampal volume at one threshold. The sample was a cohort of 968 participants and is perhaps underpowered to detect an association. Power of the PRS is also limited by the size of the initial ENIGMA GWAS (n=11,840), larger discovery sample sizes greatly improve the accuracy of PRS (43, 44). Of the PRS that were associated with their phenotype, the largest amount of variance explained was 1% with the majority predicting ~0.6%. The amount of variance explained is therefore very low although

bioRxiv preprint doi: https://doi.org/10.1101/059352; this version posted June 20, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

this is fairly common in PRS studies (45) with one of the largest explained variance by PRS reported in schizophrenia (~7% on the liability scale) (44). It is therefore perhaps unsurprising that there was no significant association with MDD in the PRS which were not associated with their own phenotype. Results from the other subcortical PRS should be treated with caution as these explained little variance with respect to their own phenotype.

This study has other notable limitations; for instance this study only explored the effects of common genetic variants and it would therefore be important to examine rare variants in MDD and subcortical volumes to generate a more complete picture of their genetic architecture. PGC GWAS, despite being one of largest GWAS for MDD (~17,000), is potentially still too small to have power to detect common variants (29), likewise the ENIGMA study too could have insufficient power. The lower heritability, higher prevalence and likely heterogeneity of MDD results in less precise estimates of marker weights from GWAS (46), decreasing the power to detect genetic correlations with other phenotypes. This may explain why the genetic correlation between hippocampal volume and MDD did not withstand multiple comparison testing. Larger genome-wide analysis would be necessary to generate confirmatory conclusions. The estimates for SNP heritability, calculated using LD score regression, were lower than have been previously described (47). For instance, in MDD the SNP heritability has been reported to be at 0.32 (48) whereas we observed a value of 0.20. This same lower SNP heritability has been reported previously using LD score regression on subcortical volumes (42). We can draw no conclusions about shared genetic architecture between amygdala and MDD due to the PRS not being associated with its phenotype and being unable to calculate the genetic correlation due to low heritability estimates.

Despite these limitations, we provide evidence of a genetic correlation between hippocampal

bioRxiv preprint doi: https://doi.org/10.1101/059352; this version posted June 20, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

volume and MDD, however, we could not demonstrate an association utilising PRS techniques. Low explanation of variance and loss of LD information were notable limitations in our PRS analysis. We therefore conclude; that the well-established relationship between hippocampal volume and MDD could be in part driven by genetic factors.

Acknowledgements

This investigation was supported by the Wellcome Trust 104036/Z/14/Z (STRADL, Stratifying Resilience and Depression Longitudinally). Generation Scotland received core funding from the Chief Scientist Office of the Scottish Government Health Directorate CZD/16/6 and the Scottish Funding Council HR03006. We thank all families, practitioners and the Scottish School of Primary Care involved in the recruitment process as well as the entirety of Generation Scotland team; interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, healthcare assistants and nurses. We are grateful towards the Dr Mortimer and Theresa Sackler foundation for the financial support for this work. This research has been conducted using the UK Biobank resource and we would therefore like to thank all participants and coordinators in this cohort.'Samples from the English Longitudinal Study of Ageing DNA Repository (EDNAR), which receives support from the National Institute on Aging (NIA) and the Economic and Social Research Council (ESRC), were used in this study. We thank contributors and the ELSA participants. IJD is supported by MRC and BBSRC funding to the University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology (MR/K026992/1).

Conflict of Interest

AMM has received financial support from Pfizer (formerly Wyeth), Janssen and Lilly. IJD and DJP were participants in UK Biobank. The remaining authors declare no conflict of

bioRxiv preprint doi: https://doi.org/10.1101/059352; this version posted June 20, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

interest.

References

1. Ustün T, Ayuso-Mateos J, Chatterji S, Mathers C, Murray C. Global burden of depressive disorders in the year 2000. Br J Psychiatry. 2004;184:386-92.
2. Sullivan P, Neale M, Kendler K. Genetic epidemiology of major depression: review and meta-analysis. Am J Psychiatry. 2000;157(10):1552-62.
3. CONVERGE consortium. Sparse whole-genome sequencing identifies two loci for major depressive disorder. Nature. 2015;523(7562):588-91.
4. Schmaal L, DJ V, van Erp T, Sämann P, Frodl T, Jahanshad N, et al. Subcortical brain alterations in major depressive disorder: findings from the ENIGMA Major Depressive Disorder working group. Mol Psychiatry. 2015;doi: 10.1038/mp.2015.69.
5. Arnone D, McIntosh A, Ebmeier K, Munafò M, Anderson I. Magnetic resonance imaging studies in unipolar depression: systematic review and meta-regression analyses. Eur Neuropsychopharmacol. 2012;22(1):1-16.
6. Depping M, Wolf N, Vasic N, Sambataro F, Thomann P, Christian Wolf R. Specificity of abnormal brain volume in major depressive disorder: a comparison with borderline personality disorder. J Affect Disord. 2015;174:650-7.
7. van Mierlo T, Chung C, Foncke E, Berendse H, van den Heuvel O. Depressive symptoms in Parkinson's disease are related to decreased hippocampus and amygdala volume. Mov Disord. 2015;30(2):245-52.
8. Kronenberg G, Tebartz van Elst L, Regen F, Deuschle M, Heuser I, Colla M. Reduced amygdala volume in newly admitted psychiatric in-patients with unipolar major depression. J Psychiatr Res. 2009;43(13):1112-7.
9. Hamilton J, Siemer M, Gotlib I. Amygdala volume in major depressive disorder: a meta-analysis of magnetic resonance imaging studies. Mol Psychiatry. 2008;13(11):993-1000.
10. Sexton C, Mackay C, Ebmeier K. A systematic review and meta-analysis of magnetic resonance imaging studies in late-life depression. Am J Geriatr Psychiatry. 2013;21(2):184-95.
11. Koolschijn P, van Haren N, Lensvelt-Mulders G, Hulshoff Pol H, Kahn R. Brain volume abnormalities in major depressive disorder: a meta-analysis of magnetic resonance imaging studies. Hum Brain Mapp. 2009;30(11):3719-35.
12. Gifuni A, Ding Y, Olié E, Lawrence N, Cyprien F, Le Bars E, et al. Subcortical nuclei volumes in suicidal behavior: nucleus accumbens may modulate the lethality of acts. Brain Imaging Behav. 2016;10(1):96-104.
13. den Braber A, Bohlken M, Brouwer R, van 't Ent D, Kanai R, Kahn R, et al. Heritability of subcortical brain measures: a perspective for future genome-wide association studies. Neuroimage. 2013;83:98-102.
14. Hibar DP, Stein JL, Renteria ME, Arias-Vasquez A, Desrivieres S, Jahanshad N, et al. Common genetic variants influence human subcortical brain structures. Nature. 2015;520(7546):224-9.
15. Won E, Ham B. Imaging genetics studies on monoaminergic genes in major depressive disorder. Prog Neuropsychopharmacol Biol Psychiatry. 2016;64:311-9.
16. Bulik-Sullivan B, Loh P, Finucane H, Ripke S, Yang J, Schizophrenia Working Group

bioRxiv preprint doi: https://doi.org/10.1101/059352; this version posted June 20, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

of the Psychiatric Genomics Consortium, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet. 2015a;47(3):291-5.
17. Bulik-Sullivan B, Finucane H, Anttila V, Gusev A, Day F, Loh P, et al. An atlas of genetic correlations across human diseases and traits. Nat Genet. 2015b;47(11):1236-41.
18. International Schizophrenia Consortium, Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature. 2009;460(7256):748-52.
19. Smith B, Campbell H, Blackwood D, Connell J, Connor M, Deary I, et al. Generation Scotland: the Scottish Family Health Study; a new resource for researching genes and heritability. BMC Med Genet. 2006;7:74.
20. Fernandez-Pujals A, Adams M, Thomson P, McKechanie A, Blackwood D, Smith B, et al. Epidemiology and Heritability of Major Depressive Disorder, Stratified by Age of Onset, Sex, and Illness Course in Generation Scotland: Scottish Family Health Study (GS:SFHS). PLoS One. 2015;10(11):e0142197.
21. Kerr S, Campbell A, Murphy L, Hayward C, Jackson C, Wain L, et al. Pedigree and genotyping quality analyses of over 10,000 DNA samples from the Generation Scotland: Scottish Family Health Study. BMC Med Genet. 2013;14(38).
22. Gunderson K. Whole-genome genotyping on bead arrays. Methods Mol Biol. 2009;529:197-213.
23. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 2015;12(3):e1001779.
24. Smith D, Nicholl B, Cullen B, Martin D, Ul-Haq Z, Evans J, et al. Prevalence and characteristics of probable major depression and bipolar disorder within UK biobank: cross-sectional study of 172,751 participants. PLoS One. 2013;8(11):e75362.
25. Wain L, Shrine N, Miller S, Jackson V, Ntalla I, Soler Artigas M, et al. Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a genetic association study in UK Biobank. Lancet Respir Med. 2015;3(10):769-81.
26. UK Biobank. Genotyping and quality control of UK Biobank, a large-scale, extensively phenotyped prospective resource. http://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=155580. Accessed 22 June, 2015.
27. Steptoe A, Breeze E, Banks J, Nazroo J. Cohort profile: the English longitudinal study of ageing. Int J Epidemiol. 2013;42(6):1640-8.
28. Marshall A, Jivraj S, Nazroo J, Tampubolon G, Vanhoutte B. Does the level of wealth inequality within an area influence the prevalence of depression amongst older people? Health Place. 2014;27:194-204.
29. Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium, Ripke S, Wray N, Lewis C, Hamilton S, Weissman M, et al. A mega-analysis of genome-wide association studies for major depressive disorder. Mol Psychiatry. 2013;18(4):497-511.
30. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira M, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559-75.
31. International Genetics of Ankylosing Spondylitis Consrtium (IGAS), Cortes A, Hadler J, Pointon J, Robinson P, Karaderi T, et al. Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. Nat Genet. 2013;45(7):730-8.
32. Videbech P, Ravnkilde B. Hippocampal volume and depression: a meta-analysis of MRI studies. Am J Psychiatry. 2004;161(11):1957-66.

bioRxiv preprint doi: https://doi.org/10.1101/059352; this version posted June 20, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

33. Schwarzer G. meta: General Package for Meta-Analysis. R package version 4.3-2. Available at http://CRAN.R-project.org/package=meta. Accessed April 29, 2016.
34. Chatterjee N, Wheeler B, Sampson J, Hartge P, Chanock S, Park J. Projecting the performance of risk prediction based on polygenic analyses of genome-wide association studies. Nat Genet. 2013;45(4):400-5.
35. Nicodemus K, Hargreaves A, Morris D, Anney R, Gill M, Corvin A, et al. Variability in Working Memory Performance Explained by Epistasis versus Polygenic Scores in the ZNF804A Pathway. JAMA Psychiatry. 2014;71(7):778-85.
36. Warner-Schmidt J, Duman R. Hippocampal neurogenesis: opposing effects of stress and antidepressant treatment. Hippocampus. 2006;16(3):239-49.
37. Snyder J, Soumier A, Brewer M, Pickel J, Cameron H. Adult hippocampal neurogenesis buffers stress responses and depressive behaviour. Nature. 2011;476(7361):458-61.
38. Heim C, Binder E. Current research trends in early life stress and depression: review of human studies on sensitive periods, gene-environment interactions, and epigenetics. Exp Neurol. 2012;233(1):102-11.
39. Mayer J, Klumpers L, Maslam S, de Kloet E, Joëls M, Lucassen P. Brief treatment with the glucocorticoid receptor antagonist mifepristone normalises the corticosterone-induced reduction of adult hippocampal neurogenesis. J Neuroendocrinol. 2006;18(8):629-31.
40. Flores B, Kenna H, Keller J, Solvason H, Schatzberg A. Clinical and biological effects of mifepristone treatment for psychotic depression. Neuropsychopharmacology. 2006;31(3):628-36.
41. MacQueen G, Campbell S, McEwen B, Macdonald K, Amano S, Joffe R, et al. Course of illness, hippocampal function, and hippocampal volume in major depression. Proc Natl Acad Sci U S A. 2003;100(3):1387-92.
42. Franke B, Stein J, Ripke S, Anttila V, Hibar D, Van Hulzen K, et al. Genetic influences on schizophrenia and subcortical brain volumes: large-scale proof of concept. Nat Neurosci. 2016;19(3):420-31.
43. Dudbridge F. Power and Predictive Accuracy of Polygenic Risk Scores. PLoS Genet. 2013;9(3):e1003348.
44. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. Nature. 2014;511(7510):421-7.
45. Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. Lancet. 2013;381(9875):1371-9.
46. Levinson D, Mostafavi S, Milaneschi Y, Rivera M, Ripke S, Wray N, et al. Genetic studies of major depressive disorder: why are there no genome-wide association study findings and what can we do about it? Biol Psychiatry. 2014;76(7):510-2.
47. Ge T, Nichols T, Lee P, Holmes A, Roffman J, Buckner R, et al. Massively expedited genome-wide heritability analysis (MEGHA). Proc Natl Acad Sci U S A. 2015;112(8):2479-84.
48. Milaneschi Y, Lamers F, Peyrot W, Abdellaoui A, Willemsen G, Hottenga J, et al. Polygenic dissection of major depression clinical heterogeneity. Mol Psychiatry. 2016;21(4):516-22.

Figure Legends

Figure 1.
Significant *P* values (<0.05) are indicated with ***.** Nucleus accumbens, amygdala and pallidum PRS were not significantly associated with their respective volume at any threshold.

Figure 2.
Both plots demonstrate a negative correlation with MDD and recurrent MDD with no heterogeneity between cohorts but neither plot reaches statistical significance. TE; treatment effect (regression beta's); seTE, standard errors; OR, odds ratio; CI, confidence intervals; W(fixed), weight of individual studies in fixed effect meta-analysis.

Tables

Table 1. Genetic correlation of subcortical brain regions and MDD.

	SNP heritability		*Genetic correlation*			
Brain region	*SNP h^2*	*s.e.*	r_G	*s.e.*	*Z*	*P*
Nucleus accumbens	0.0855	0.0438	0.0458	0.210	0.218	0.828
Amygdala	-0.0277	0.0354	NA	NA	NA	NA
Caudate nucleus	0.253	0.0432	0.0752	0.130	0.580	0.562
Hippocampus	**0.137**	**0.0481**	**0.460**	**0.200**	**2.30**	**0.0213**
ICV	0.167	0.0462	0.123	0.166	0.739	0.460
Pallidum	0.171	0.049	-0.0077	0.158	-0.0491	0.961
Putamen	0.297	0.051	0.0986	0.118	0.834	0.404
Thalamus	0.125	0.0401	-0.0808	0.177	-0.457	0.648

***** The heritability of amygdala was too low to be carried forward. The *P* values shown are uncorrected for multiple testing.

Table 2. Mixed model analysis of Subcortical volumetric PRS and MDD.

PRS	Study	P value	beta	s.e.
Nucleus accumbens	**GS**	0.485	-0.0181	0.0218
	UKB	0.284	-0.0151	0.0141
	ELSA	0.659	0.0175	0.0395
Amygdala	**GS**	0.327	-0.0285	0.0217
	UKB	0.994	0.000102	0.0146
	ELSA	0.208	-0.0500	0.0398
Caudate nucleus	**GS**	0.246	0.0197	0.0224
	UKB	0.426	-0.0127	0.0159
	ELSA	0.114	-0.0630	0.0398
Hippocampus	**GS**	0.782	-0.00299	0.0211
	UKB	0.602	-0.0073	0.0140
	ELSA	0.571	-0.0222	0.0392
ICV	**GS**	0.395	0.0179	0.0225
	UKB	0.283	-0.0158	0.0147
	ELSA	0.0995	0.0652	0.0396
Pallidum	**GS**	0.752	-0.00808	0.0221
	UKB	0.234	0.0185	0.0155
	ELSA	0.658	0.0176	0.0398
Putamen	**GS**	0.303	-0.0246	0.0218
	UKB	0.220	0.0228	0.0186
	ELSA	0.695	-0.0161	0.0410
Thalamus	**GS**	0.451	-0.0192	0.0219
	UKB	0.792	0.00374	0.0142
	ELSA	0.222	-0.0482	0.0395

* Best P value threshold for PRS were carried forward in this analysis; Nucleus accumbens=0.01, Amygdala= 0.1, Caudate nucleus=0.5, Hippocampus=0.01, ICV=0.5, Pallidum=0.5, Putamen=0.1, Thalamus=0.05. No PRS demonstrated a significant association with MDD in any cohort.

Figure 1. Variance in subcortical volumes explained by their respective PRS.

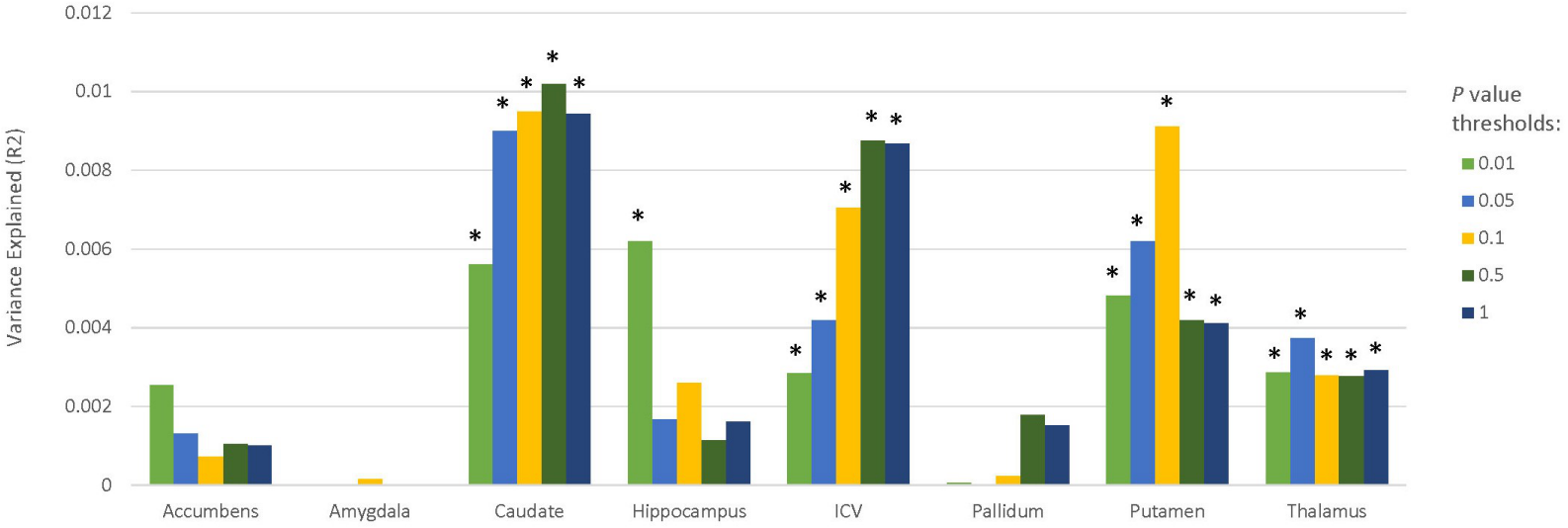
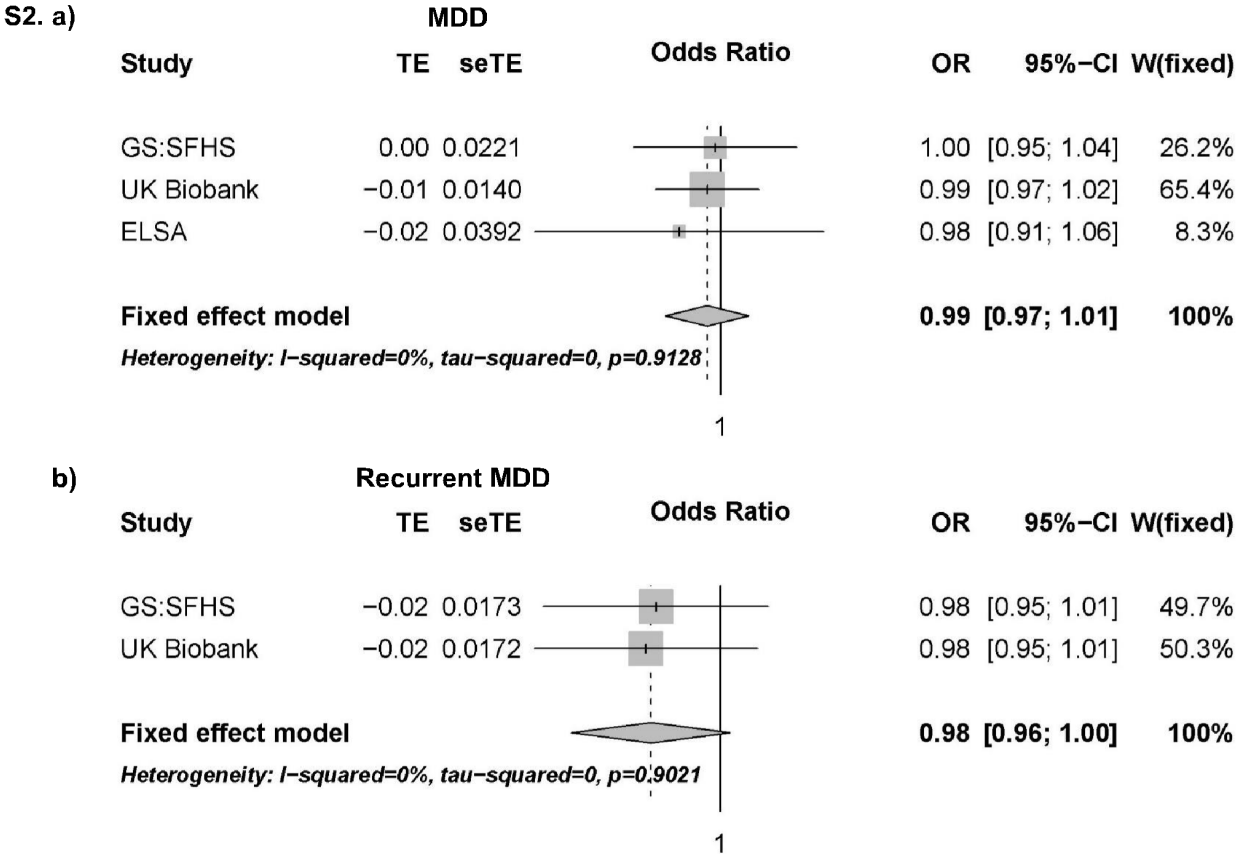
Figure 2. Forest plots of fixed effect meta-analysis of hippocampal volume PRS with MDD (a) and recurrent MDD (b) in Generation Scotland: Scottish Family Health Study (GS:SFHS), UK Biobank and English Longitudinal Study of Ageing (ELSA).