Genetic overlap between obsessive-compulsive disorder, related symptoms in the population and insulin signaling: etiological implications

Janita Bralten, Ph.D. [1,2,][*], Joanna Widomska, M.Sc. [2,3][*], Ward De Witte, B.Sc. [1], Dongmei Yu, MS [4,5],

Carol A. Mathews, M.D. [6], Jeremiah M. Scharf M.D., Ph.D.[4,5], Jan Buitelaar[2,7,8], Jennifer Crosbie, Ph.D. [9,10] Russell Schachar, M.D. [9,10], Paul Arnold M.D., Ph.D. [11,12,13] Mathieu Lemire, Ph.D. [9], Christie L. Burton, Ph.D. [9], Barbara Franke, Ph.D. [1,2,7,][*], Geert Poelmans, M.D., Ph.D. [1,][*,†]

[1]Department of Human Genetics, Radboud University Medical Center, Nijmegen, The Netherlands

[2]Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands

[3]Department of Cognitive Neuroscience, Radboud University Medical Center, Nijmegen, The Netherlands

[4]Psychiatric and Neurodevelopmental Genetics Unit, Center for Genomic Medicine, Department of Psychiatry, Massachusetts General Hospital, Boston, Massachusetts, USA.

[5]Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA. Department of Neurology, MGH, Boston, MA

[6]Department of Psychiatry and Genetics Institute, University of Florida, Gainesville, Florida, USA

[7]Department of Psychiatry, Radboud University Medical Center, Nijmegen, The Netherlands

[8]Karakter Child and Adolescent Psychiatry, Nijmegen, The Netherlands

[9]Neurosciences & Mental Health Program, Hospital for Sick Children, Toronto, Canada

[10]Department of Psychiatry, University of Toronto, Toronto, Canada

[11]Genetics & Genome Biology, Hospital for Sick Children, Toronto, Canada

[12]Mathison Centre for Mental Health Research and Education, University of Calgary, Calgary, Canada

[13]Departments of Psychiatry & Medical Genetics; Hotchkiss Brain Institute, Cumming School of Medicine, University of Calgary, Calgary, Canada

* equal contribution

† corresponding author:

Geert Poelmans, M.D., Ph.D.

Department of Human Genetics, Radboud University Medical Center,

P.O. Box 9101, 6500 HB Nijmegen, The Netherlands

Tel.: +31 630345956 ; E-mail: geert.poelmans@radboudumc.nl

Abstract

Objective: Obsessive-compulsive symptoms in the population have been linked to obsessive-compulsive disorder (OCD) in previous genetic and epidemiological studies. Genetic studies also show a link to insulin signalling. Here we aim to assess the presence and extent of genetic and biological overlap between OCD, OCD symptoms in the population, and insulin signalling, making use of the largest data sets currently available.

Methods: We used phenotypic and genetic data from a population based cohort (n=650) of children and adolescents to conduct genome-wide association studies (GWAS) to six factors derived from exploratory factor analyses on OCD symptom scores. We performed polygenic risk score analyses to check whether these OCD symptom traits had a shared genetic etiology with clinically diagnosed OCD (using GWAS data of the PGC, n=2688 OCD cases and 7037 controls). Subsequently we investigated potential shared biology performing gene-set analyses with an earlier defined set of 51 OCD-associated genes centered around insulin-regulated synaptic function and polygenic risk score analyses of five peripheral insulin signaling-related traits (type 2 diabetes (T2D) and the blood levels of four T2D biomarkers (n between 42,854-159,208)).

Results: We found genetic sharing between diagnosed OCD and four out of six factors based on OCD symptoms in the population: 'impairment', 'contamination/cleaning', 'guilty taboo thoughts', and 'symmetry/counting/ordering'. Gene-set analysis with insulin-related genes revealed an association with 'symmetry/counting/ordering'. We also identified genetic sharing between OCD, the total score of OCD symptoms and six of the derived symptom factors with the peripheral insulin signaling-related traits. We were able to validate part of our results on symmetry/counting/ordering and contamination/cleaning in an independent population-based cohort (n=5047).

Conclusions: . Our findings suggest that GWASs of OCD symptoms in population-based cohorts could be used to discover OCD-relevant genes. Our results also imply that altered insulin signaling, as

relevant to T2D, is also involved in different aspects of compulsive and obsessive behaviours and

clinical OCD. This may open up a new field of brain-based insulin-related disorders.

Introduction

Obsessive-compulsive disorder (OCD) is a psychiatric condition characterized by persistent, intrusive thoughts and urges (obsessions) and repetitive, intentional behaviours (compulsions). Typically, the repetitive behaviours are performed to reduce anxiety caused by obsessions (Association, 2013). OCD affects an estimated 2.5% of the world's population (Karno et al., 1988; Weissman et al., 1994). It is a heterogeneous disorder with more than 90% of patients showing one or more comorbid disorders, including Tourette syndrome (Goodman et al., 2006), attention-deficit/hyperactivity disorder (ADHD), anxiety disorders, mood disorders, substance use disorders, and eating disorders (Torresan et al., 2013). Twin studies have shown that OCD is moderately heritable, and ~40% of the phenotypic variance can be explained by genetic factors and a higher genetic load has been reported in children and adolescents (Taylor, 2011; van Grootheest et al., 2009). The genetic architecture of OCD is thought to be complex, with multiple genetic variants of small effect size contributing to its etiology; this has hampered the identification and replication of genetic susceptibility factors for the disorder. Meta-analysis of hypothesis-driven candidate gene association studies has implicated serotoninergic and catecholaminergic genes in OCD, while studies focusing on glutamatergic and neurotrophic genes have shown inconsistent results regarding OCD risk (Taylor, 2013). Thus far, two independent genome-wide association studies (GWASs) provided the first hypothesis-free genetic analyses for OCD (Mattheisen et al., 2015; Stewart et al., 2013), but neither GWAS nor a subsequent meta-analysis of 2688 OCD cases and 7037 healthy controls (International Obsessive Compulsive Disorder Foundation Genetics and Studies, 2017) yielded genome-wide significant findings, likely because of lack of power. The study was able to show that the polygenic signal present in either of the two samples was able to predict OCD status in the other sample, indicating the polygenic nature of the disorder {International Obsessive Compulsive Disorder Foundation Genetics, 2017 #54}.

The diagnostics of OCD are based solely on interviews on clinical symptoms; currently, no genetic or biological markers are available with high enough specificity and accuracy (American Psychiatric

5

Association, 2013). Symptom-based subtyping of OCD through factor and cluster analyses has consistently identified specific OCD symptom clusters or dimensions, with the most reliable dimensions including contamination/cleaning, doubt/checking, symmetry/ordering, unacceptable/taboo thoughts, and hoarding (Calamari et al., 2004; Leckman et al., 1997; Mataix-Cols et al., 1999; Summerfeldt et al., 1999). Looking at the symptoms underlying the diagnosis of OCD, it is clear that such behaviours are also present in the general population (Alvarenga et al., 2015; Fullana et al., 2009; Park et al., 2016). Indeed, population studies show that up to 25% of individuals endorse obsessions and/or compulsions as defined in the DSM-IV (Fullana et al., 2009). These OCD symptoms show a widespread distribution in the general population (Abramowitz et al., 2014; Park et al., 2016), and there is growing evidence suggesting that OCD represents the upper extreme of this distribution. Population-based twin studies indicate that OCD symptoms are heritable; their heritability ranges from 30 to 77% (den Braber et al., 2016; Mathews et al., 2014). In addition, genetic factors were found to contribute to specific (dimensions of) OCD symptoms, including contamination/cleaning (Brakoulias et al., 2016; Burton et al., 2018; Chacon et al., 2007) and checking/ordering (Burton et al., 2018; Kohlrausch et al., 2016). Genetic continuity between OCD symptoms in the population and clinical OCD is suggested by the fact that polygenic risk scores (PRS) based on clinical OCD GWAS data significantly predicted OCD symptoms in an independent population-based sample of 6931 individuals (den Braber et al., 2016). This result was confirmed in a second population-based study of 3982 individuals (Taylor et al., 2018). This genetic continuity suggests that genetic studies of OCD symptoms in the general population could aid in the identification of susceptibility loci for clinical OCD and provide insight in specific symptom domains affected by individual genetic risk factors.

Albeit lacking significant findings in individual GWASs of OCD, we earlier performed an integration of the top ranked-results from the existing studies; our 'molecular landscape' of OCD suggested involvement of genes regulating postsynaptic dendritic spine formation and function through central nervous system (CNS) insulin-dependent signaling (van de Vondervoort et al., 2016). Links of OCD and OCD symptoms with dysregulated *peripheral* insulin signalling, especially by diabetes mellitus,

have been previously reported. Increased obsessive symptom levels have been found in men with type 1 diabetes (T1D) (Winocour et al., 1990). Moreover, OCD symptoms were found to be positively correlated with blood levels of glycosylated hemoglobin (HbA1c), a diagnostic measure of type 2 diabetes (T2D) (Kontoangelos et al., 2013). OCD patients also were found to show markedly higher levels of fasting glucose (a characteristic of T2D)(Albert et al., 2013) and have a higher risk of developing T2D (Isomura et al., 2017).

Based on such evidence, we here used a genetic approach to assess the presence and extent of biological overlap between OCD, OCD symptoms in the population, and insulin signalling, making use of the largest data sets currently available.. We do this in four steps, first we factor out phenotypic heterogeneity using an exploratory factor analyses on obsessive/compulsive symptoms measured in a childhood population based cohort. Secondly we investigate the presence of shared genetic etiologies between OCD and the total and factorized obsessive compulsive symptoms. Third by focusing on insulin-related traits we study whether there is a shared biology between OCD and obsessive compulsive traits. As a fourth and final step we try to validate and extend our findings from the first three steps in an independent population sample.

Methods

Sample, phenotypic and genetic data

We studied OCD symptoms in the Philadelphia Neurodevelopmental Cohort (PNC) (Calkins et al., 2015; Gur et al., 2014; Satterthwaite et al., 2016; Satterthwaite et al., 2014). This cohort includes 8719 children and adolescents aged 8-21 years with neurobehavioral phenotypes and genome-wide genotyping data. Participants in the PNC provided written consent for genomic studies when they presented for paediatric services to the Children's Hospital of Philadelphia health care network. Notably, participants were not recruited from psychiatric clinics, and the sample is not enriched for individuals who seek psychiatric help. OCD symptoms were assessed with a computerized version of

7

the Kiddie-Schedule for Affective Disorders and Schizophrenia (K-SADs)) (Kaufman et al., 1997) called GO-ASSESS.

For the current study, we selected 22 GO-ASSESS questions that corresponded to the diagnostic criteria for OCD (question numbers OCD001-OCD017, OCD024-OCD025, OCD032-OCD034; see **Supplementary Table 1**). The answers to questions OCD032 and OCD033 were re-categorized from a 10-point scale into binary responses (i.e., with scores from 0 to 4 converted into a score of 0 for "no" and scores from 5 to 10 converted into a score of 1 for "yes") for compatibility with the other 20 questions.

Data was selected if questions OCD001-OCD008 and OCD011-OCD17 were completed, which related to the presence of obsessions and/or compulsions (**Supplementary Table 1**). If those questions were all answered "no", we allowed the questions on the consequences of obsessions and compulsions (OCD009, OCD010, OCD024, OCD025, OCD032, OCD033, OCD034) to be left blank, as no consequences are expected if no symptoms are present. The scores for each of the questions were summed to create a total OCD symptom score (range 0-22).

Genome-wide genotyping in the PNC cohort has been performed in waves using six different genotyping platforms. Genotypes are available through the NIMH Database of Genotypes and Phenotypes (dbGaP), study accession ID phs000607.v1.p1. (https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000607.v1.p1&phv= 194292&phd=&pha=&pht=3445&phvf=28&phdf=&phaf=&phtf=&dssp=1&consent=&temp=1). We assessed the phenotypic distributions per platform, and for this study used the two genetic platforms with a minimum of 1000 individuals with all phenotypic data available, HumanOmniExpress-12v1.0(OMNI) and Human610_Quadv1_B(Quad) (n=1989 and n=1257, respectively). Hence, genome-wide genotyping data was available for 3246 individuals who also completed the 22 OCD questions. Ancestry was addressed by only selecting individuals of self-reported European descent and by using multidimensional scaling (MDS) analysis. Ancestry outliers (i.e., of non-European descent) were

excluded based on visual inspection of the first two principal components, leaving 3041 individuals for analyses (n=1860 genotyped on Quad and n=1181 genotyped on OMNI). As one of the main aims of our study was to assess the genetic overlap between OCD and OCD symptoms in the population, we only used phenotypic and genetic data from those PNC participants who answered positively on at least one of the questions related to the presence of obsessions and/or compulsions(and hence scored at least 1 on the total OCD symptom score). This resulted in a final sample of 650 individuals for the subsequent factor and genome-wide association analyses (n=418 genotyped on Quad and n=232 genotyped on OMNI).

Factor analysis

First, using SPSS 20.0 (SPSS Technologies, Armonk, NY, USA), we determined the internal consistency (Cronbach's α) of the answers to the 22 questions that together constitute the total OCD symptom scores in the 650 PNC participants. We then conducted a factor analysis of the scores on the 22 OCD questions using the Promax method to determine which combination of factors explains the largest portion of the observed variance in the total score of OCD symptoms.

Genome-wide association analyses

Quality control filtering was applied to genetic data of the final sample of participants to remove single nucleotide polymorphisms (SNPs) with low minor allele frequency (<0.01), poor genotype call rate (<95%), and deviations from Hardy-Weinberg equilibrium ($P<1 \times 10^{-6}$). The imputation protocol used MaCH (Li et al., 2010) for haplotype phasing, and minimac (Howie et al., 2012) for imputation. Quality control filtering was applied to remove imputed SNPs with low imputation quality score (info<0.6) and low minor allele frequency (<0.05).

We assessed whether the distribution of the OCD symptom scores and the scores for each of the OCD symptom factors showed enough variation to perform genome-wide analyses. When the scores fell within the limits of a normal distribution (i.e., a skewness and kurtosis of the distribution between -1 and 1), we used a continuous trait design for the genome-wide association analysis.

Otherwise, we used a pseudocase-control design, in which all individuals with a score of 0 for a factor were defined as 'controls' and compared against the 'pseudocases', i.e. all individuals with a score of 1 or more for that factor.

Genome-wide association studies (GWASs) were carried out with mach2qtl (Li et al., 2010) using the total OCD symptom score and the scores for those factors that showed enough variation in their score as phenotypes, with age and gender included as covariates; GWAS was performed separately for each genotyping platform. Genome-wide results were then combined in an inverse-variance-weighted meta-analysis using METAL (Willer et al., 2010), accounting for genomic inflation.

Shared genetic etiology analyses

OCD with OCD symptoms

We determined the level of shared genetic etiology between diagnosed OCD and OCD symptoms in the population using the summary statistics from the meta-analysis of the two published GWASs of OCD (International Obsessive Compulsive Disorder Foundation Genetics and Studies, 2017) (data provided through the Psychiatric Genomics Consortium (PGC) for 2688 OCD cases of European ancestry and 7037 genomically matched controls) as the 'base' sample for polygenic risk score (PRS)-based analyses in PRSice (Euesden et al., 2015). The summary statistics from the GWASs of the different OCD symptoms - i.e., the total OCD symptom scores and the scores for the OCD symptom factors in the PNC were used as the 'target' samples for the PRS-based analyses. Clumping and the protocol used in PRSice are described in the **Supplementary Methods**.

The calculated P-values for genetic sharing were aggregated and corrected for multiple comparisons using the false discovery rate (FDR) method (Benjamini and Hochberg, 1995; Glickman et al., 2014). .

We performed SECA for the significant findings from the PRS-based analyses (http://neurogenetics.qimrberghofer.edu.au/SECA; Nyholt, 2014). In SECA, association results rather than individual genotyped data are analyzed to test for genetic pleiotropy (the same SNPs affecting both traits) and concordance (the agreement in SNP effect directions across both traits) between two

genetically determined traits. Concordance can be negative (the effect direction in test dataset is negative when the effect direction is positive in the reference dataset, and vice versa), positive (the effect direction is positive or negative in both datasets) or null when there is no evidence for concordance. We used SECA to calculate empirical P-values for pleiotropy and concordance between all traits that emerged from the PRS-based analyses as having a significant shared genetic etiology. SECA P-values were considered significant if they exceeded a Bonferroni-corrected threshold accounting for the number of tests we performed.

Shared biology analyses

Gene-set analyses

For gene-set analysis, we first used a set of OCD-associated genes that had been included in a molecular landscape of OCD, which implicated postsynaptic dendritic spine formation and function through insulin-dependent signaling (van de Vondervoort et al., 2016). This resulted in a set of 52 unique genes, of which the 51 autosomal genes were included insubsequent analyses. Gene set analysis was performed using the Multimarker Analysis of GenoMic Annotation (MAGMA) software (de Leeuw et al., 2015), see **Supplementary Methods**. P-values for gene-set association were considered significant if they exceeded a Bonferroni-corrected threshold accounting for the number of phenotypes tested. For significant gene-set associations, we also looked at the gene-wide P-values for each of the individual genes in the set.

OCD and OCD symptoms with peripheral insulin signaling-related traits

To determine the level of genetic sharing between five peripheral insulin signaling-related traits and OCD as well as OCD symptoms, we also conducted PRS-based analyses in PRSice (Euesden et al., 2015). As base samples, we used summary statistics data from GWASs of the following peripheral insulin signaling-related traits: Type 2 diabetes (T2D) (GWAS of 26,676 cases and 132,532 controls) (Scott et al., 2017), blood levels of HbA1c (GWAS of 123,665 general population subjects; increased HbA1c levels are a diagnostic measure of T2D) (Wheeler et al., 2017), fasting insulin (GWAS of

11

108,557 general population subjects; increasing fasting insulin levels lead to insulin resistance, the pathological hallmark of T2D) (Scott et al., 2012), fasting glucose (GWAS of 133,010 general population subjects; increased fasting glucose levels (i.e. hyperglycemia) are a characteristic of T2D)(Scott et al., 2012), and the blood levels of glucose 2 hours after an oral glucose challenge (2-hour glucose or 2hGlu, which is a clinical measure of glucose tolerance used in the diagnosis of T2D) (GWAS of 42,854 general population subjects) (Scott et al., 2012). As target samples of the PRS-based analyses, we used the summary statistics from the OCD GWAS meta-analysis as well as the summary statistics from the GWASs of the total OCD symptom scores and the scores for the OCD symptom factors in the PNC. Using the same approach as described above, we then calculated P-values of shared genetic etiology between the five peripheral insulin-signaling related traits and OCD/OCD symptoms, and we also corrected these P-values using the FDR method. We performed SECA for the significant findings from the PRS-based analyses to calculate empirical P-values for pleiotropy and concordance between all traits that emerged from the PRS-based analyses as having a significant shared genetic etiology. SECA P-values were considered significant if they exceeded a Bonferroni-corrected threshold accounting for the number of tests we performed.

Validation analyses in an independent population sample

In order to validate and possibly expand our findings, we performed PRS-based, SECA, and gene-set analyses using data from GWASs of OCD symptoms in an independent population sample. The 'Spit for Science' project recruited 16,718 children and adolescents aged 6-17 years at a local science museum (Crosbie et al., 2013). OCD symptoms were measured using the Toronto Obsessive-Compulsive Scale (TOCS), a novel, validated 21-item parent- or self-report questionnaire that covers a wide variation in OCD symptoms (Park et al., 2016), see **Supplementary Methods**. Each item was scored from -3 (far less often than others of the same age) to +3 (far more often than others of the same age). The total score shows the wide distribution of OCD symptoms in the community with some skew to the right (Park et al., 2016).

12

For the validation analyses, we first assessed which TOCS questions could be grouped into OCD symptom factors similar to those calculated based on the data from the PNC. We identified two OCD symptom factors (symmetry/counting/ordering and contamination/cleaning) that were similar, see Supplementary Table 4.. We conducted 'continuous trait design' genome-wide association analyses for each factor. In total, genome-wide genotyping data for 5047 individuals of Caucasian descent entered the analysis. A description of genotyping, quality control and imputation can be found elsewhere (Burton et al.,biorxiv 248484). A linear regression model was used, adjusting for age, sex, respondent (questionnaire filled by parent or child), genotyping platform (Illumina HumanCoreExome-12 (v1-0) bead-chip or the HumanOmni1-Quad (v1.0) bead-chip), and the first 6 principal components from MDS. The significance between the SNP imputed allele dosage and the response was calculated from a Wald test. Only SNPs with minor allele frequency > 0.01 and imputation quality (allelic R2) > 0.60 were tested. Using the PRS-based methods and criteria described above, we used the summary statistics of the GWASs of the two TOCS-based OCD symptom factors to identify the level of shared genetic etiology between 1) the two TOCS-based OCD symptom factors and the corresponding factors based on data from the smaller PNC, 2) OCD (as derived from the GWAS meta-analysis data of PGC) and the two TOCS-based factors, and 3) the five peripheral insulin signaling-related traits and the two TOCS-based factors. For significant findings from the PRS-based analyses, we also conducted SECA analyses. Gene-set analysis between the set of 51 insulin signaling genes and the two TOCS-based OCD symptom factors was also performed.

Results

Factor analysis

The internal consistency of the answers to the 22 questions that constitute the total score of OCD symptoms was satisfactory (Cronbach's α=0.69). **Supplementary Figure 1A** shows the total score distribution with a mean of 6.4 (s.d.=3.35). Factor analysis of the 22 OCD symptom questions revealed an eight factors solution as the best-fitting model, explaining 58.6% of the variance in the

total score. We named these eight OCD symptom factors 'impairment', 'symmetry/counting/ordering', 'contamination/cleaning', 'aggressive taboo thoughts', 'repetition', 'guilty taboo thoughts', 'distress' and 'religious taboo thoughts' (**Table 1**). In **Supplementary Figure 1B**, the distributions of the scores on the eight factors are shown.

Genome-wide association analyses

We conducted GWASs of the total OCD symptom score and the scores for six OCD symptom factors. . Based on the distributions of the scores, we used a continuous trait design for the GWASs of the total OCD symptom score and the factors 'impairment', 'symmetry/counting/ordering', and 'guilty taboo thoughts'; a pseudocase-control design was used for analysis of the factors 'contamination/cleaning', 'aggressive taboo thoughts', and 'distress'. The distribution of the scores on the OCD symptom factors 'repetition' and 'religious taboo thoughts' showed too little variation to be taken forward.

Shared genetic etiology analyses

OCD with OCD symptoms

With the polygenic risk score based method we found statistically significant evidence for overlap between genetic variants increasing the risk for diagnosed OCD and genetic variants associated with three OCD symptom factors: 'impairment', 'contamination/cleaning', and 'guilty taboo thoughts' (see Supplementary **Figure 2** and **Table 2**). SECA analysis yielded significant evidence - after Bonferroni correction - of genetic pleiotropy (i.e., the same SNPs affecting two traits) between OCD and 'contamination/cleaning' as well as 'guilty taboo thoughts'. We also found significant evidence of negative genetic concordance between OCD and 'impairment' as well as 'guilty taboo thoughts' while there was a significant, positive concordance between OCD and contamination/cleaning (see **Supplementary Table 2**).

Shared biology analyses

Gene-set analyses

MAGMA-based gene-set analysis for the 51 OCD genes containing 33,329 SNPs (effective number of SNPs after adjusting for LD structure=2,189) revealed significant association for 'symmetry/counting/ordering' (P=0.0038) after Bonferroni correction (P=0.05/7 tests). Within the gene-set significantly associated with 'symmetry/counting/ordering', none of the individual genes showed gene-wide association after Bonferroni correction (**Supplementary Table 3**). No significant associations were found for total OCD symptoms or the five other OCD symptom factors.

OCD and OCD symptoms with peripheral insulin signaling-related traits

T2D

In the PRS-based and SECA analyses with the five peripheral insulin signaling-related traits as base samples we found statistically significant evidence for overlap and pleiotropy between genetic variants increasing the risk for T2D and genetic variants associated with OCD, the total score of OCD symptoms, and 'aggressive taboo thoughts', we also found a significant shared genetic etiology between T2D and 'symmetry/counting/ordering' and both positive and negative concordances were observed (see **Table 4**).

Insulin signalling-related traits

For a summary of the findings, please see Figure #, in short blood HbA1c levels showed a shared genetic etiology with total OCD symptoms, 'impairment', 'contamination/cleaning', 'aggressive taboo thoughts' and 'distress' and we found evidence for pleiotropy for all of these except for 'distress'. Blood levels of fasting insulin showed genetic sharing and significant pleiotropy with OCD risk and 'distress'. For blood levels of fasting glucose, we found a shared genetic etiology with 'symmetry/counting/ordering', 'contamination/cleaning' and 'aggressive taboo thoughts' and pleiotropy with 'contamination/cleaning'. Lastly, we observed a shared genetic etiology between 2hGlu blood levels and OCD, 'contamination/cleaning' and 'guilty taboo thoughts', and pleiotropy

with OCD as well as 'contamination/cleaning'. Both positive and negative concordances were observed, see **Table 4**.

Validation analyses in an independent population sample

Two OCD symptom factors with similar questions as in the 'symmetry/counting/ordering' and 'contamination/cleaning' factors based on the PNC data were identified (**Supplementary Table 4** and **Supplementary Figure 4**). A continuous trait design GWAS of the two TOCS-based OCD symptom factors and subsequent PRS-based and SECA analyses (**Supplementary Table 5** and **Supplementary Figure 5A-D** ; **Supplementary Table 6**) showed that on a genetic level these factors were indeed significantly overlapping with the PNC data (Symmetry/counting/ordering TOCS and PNC R^2 = 0.89 % ; FDR-corrected P = 8.09E-03 and Contamination/cleaning TOCS and PNC R^2 = 0.42 % ; FDR-corrected P = 4.87E-02) justifying our use of these factors for subsequent validation analyses Diagnosed OCD shows genetic overlap with 'symmetry/counting/ordering $_{TOCS}$' and 'contamination/cleaning $_{TOCS}$', with the later providing a validation of our finding of genetic overlap between OCD and the 'contamination/cleaning' factor based on the PNC data. SECA analyses revealed a negative concordance between OCD and 'symmetry/counting/ordering $_{TOCS}$' as well as 'contamination/cleaning $_{TOCS}$', with the latter being the opposite of what we found (i.e., a positive concordance) for OCD and the 'contamination/cleaning' factor based on the PNC data. Furthermore, we found a shared genetic etiology and positive concordance between T2D and 'contamination/cleaning $_{TOCS}$' and between blood levels of HbA1c and 'contamination/cleaning $_{TOCS}$'. The latter findings of genetic sharing and a positive concordance between HbA1c levels and 'contamination/cleaning $_{TOCS}$' validate our results for HbA1c levels and the 'contamination/cleaning' factor based on the PNC data. Lastly, gene-set analysis for the OCD landscape genes in the two OCD symptom factors revealed no associations surviving Bonferroni correction. Summary of the PRS-based and SECA analyses are presented in Table 4.

Discussion

In this study, we explored the genetic overlap between OCD, OCD symptoms in the population, and insulin signaling. Surprisingly, using the results from the published GWASMA of OCD and phenotypic and genetic data of 650 children and adolescents from the general population, we did not find a shared genetic etiology between OCD and total OCD symptoms scores, but we did find significant evidence for a shared genetic etiology between OCD and three OCD factors, 'impairment', 'contamination/cleaning', and 'guilty taboo thoughts'. To investigate potential shared biology we started from the insulin link found in the clinical OCD GWAS top hits and tested an insulin gene-set and insulin-based traits for their link with OCD and OCD traits. We found a genetic link between the insulin gene-set and symmetry/counting/ordering as well as multiple insulin based traits with OCD and OCD traits. Validation was observed in the larger Spit for Science cohort for the genetic overlap between OCD and the 'contamination/cleaning' factor. In this cohort, we were also able to expand our results by finding genetic sharing between OCD and the 'symmetry/counting/ordering' factor.

Our findings of genetic sharing between diagnosed OCD and related symptoms in the population are in keeping with the literature about genetic factors contributing to specific (dimensions of) OCD symptoms and, importantly, the notion of (at least partial) genetic continuity between OCD symptoms and diagnosed OCD. OCD and three out of six analysed OCD symptom factors, 'impairment', 'contamination/cleaning', and 'guilty taboo thoughts', showed genetic overlap in PRS-based analyses. Using SECA as an alternative method, we also observed pleiotropy for OCD and two of those, i.e. 'contamination/cleaning' and 'guilty taboo thoughts'. SECA-based analysis of concordance - the agreement in SNP effect direction across two traits - revealed a positive concordance for OCD and 'contamination/cleaning'; however, unexpectedly, negative concordance was observed for 'impairment' and 'guilty taboo thoughts', suggesting opposite directions for the contributions of alleles to disorder and trait. Concordance analysis in the Spit for Science cohort also yielded a negative concordance between OCD and both the 'symmetry/counting order' and 'contamination/cleaning' factors. For the latter factor, this was the opposite of our analysis in the

17

PNC that revealed a positive concordance. This discrepancy may be due to the fact that, although they share variants conveying genetic susceptibility, the two 'contamination/cleaning' factors do not consist of exactly the same questions and may therefore cover slightly different 'contamination/cleaning' symptoms.

OCD and OCD symptoms have been phenotypically linked to an altered (peripheral) insulin signalling. To find out whether these links are based on a shared genetic etiology, we used two strategies. Firstly, we performed a gene-set analysis including data on 51 OCD-associated genes centered around insulin-regulated synaptic function, showing significant association with 'symmetry/counting/ordering. Finding association with OCD symptoms in PNC prompted us to perform more comprehensive analyses using the results of published GWASs of T2D and the blood levels of HbA1c, fasting insulin, fasting glucose and 2hGlu. We indeed found a shared genetic etiology between T2D and OCD, the total OCD symptoms score, and 'aggressive taboo thoughts' additional overlap was seen for 'symmetry/counting/ordering' and 'contamination/cleaning' (validated in an independent cohort). We also found shared genetic etiologies between multiple insulin based traits with OCD and OCD traits, . Taken together, our findings provide support for 'dysregulated' peripheral insulin signaling as seen in T2D as a biological process contributing to both OCD and OCD symptoms. Further evidence for a role of (altered) peripheral insulin signaling in OCD etiology is suggested by the fact that selective serotonin reuptake inhibitors (SSRIs) - the first-line pharmacological treatment for OCD - positively affect diabetic parameters when used to treat depressive symptoms in T2D (i.e. decreasing HbA1c levels and insulin requirement, and increasing insulin sensitivity) (Janardhan Reddy et al., 2017). In addition, a recent study demonstrated that bilateral deep brain stimulation (DBS), a safe and effective treatment option for pharmaco-resistant OCD, of the ventral striatum not only resulted in a reduction of OCD symptoms but also a decrease of fasting insulin levels in the blood of both an OCD patient with T2D and a cohort of non-diabetic OCD patients (Ter Horst et al., 2018). Insulin in the central nervous system (CNS)- either entering from the periphery by crossing the blood brain barrier (Margolis and Altszuler, 1967) or synthesized in the CNS (Clarke et al., 1986) - has

18

important non-metabolic functions in addition to neuronal glucose uptake, including modulating synaptic and dendritic plasticity (Chiu et al., 2008) , learning and memory (Dou et al., 2005; Zhao et al., 1999).

The current results should be viewed in light of some strengths and limitations. A clear strength is the use of quantitative symptom scores collected through questionnaires in the general population; this has the advantage that a large number of individuals can be assessed in a relatively cheap way. Moreover, using samples selected from the population may reduce selection bias, which can occur when patient samples are analysed (e.g. individuals suffering from several comorbid disorders are more likely to present for clinical care) (Gibbs N., 1996]. The main limitation of the current study is that the sample size of the GWASs based on the data from the PNC is too small to discover new single genetic variant associations. However, this sample size was large enough to provide proof of concept - through shared genetic etiology and gene-set analyses - for genetic sharing between OCD, related symptoms in the population, and insulin signaling measures. The fact that we were able to validate part of our findings in the independent Spit for Science cohort adds credibility to our work in PNC. Another possible limitation of this study may be the fact that the proportions of the variance in the target phenotypes being explained by the base phenotypes emerging from the PRS-based analyses are quite small. However, these 'variances explained' are in fact (much) higher than those found in similar analyses, e.g. our study of the genetic overlap between autism and autistic traits (Bralten et al., 2018), with the highest variance explained being 0.54 %, and the polygenic risk derived from a GWAS of OCD explaining (only) 0.20% of the variance in OCD symptoms in a population sample (den Braber et al., 2016) . Moreover, as the variance explained is dependent on the size of the 'base sample' for the generation of the polygenetic risk scores (Dudbridge, 2013), the observed variances explained with the summary statistics from the still relatively small GWASMA of OCD as base sample may be underestimates.

In conclusion, we identified genetic overlap between OCD, OCD symptoms in the population and both peripheral and CNS insulin signaling. These findings have important implications, as GWASs of OCD symptoms in (existing) population-based cohorts - that are relatively easy and cheap to collect compared to case-control samples - could be used to identify novel OCD-relevant genes, which will in turn further our understanding of the biology underlying OCD. Our results imply that altered insulin signaling is not only relevant for somatic disorders, but also is involved in the etiology of brain disorders and traits, in this case OCD (symptoms). Further studies are needed to disentangle the contributions of peripheral and central (CNS) insulin production and signaling to these disorders and traits.

Acknowledgements

References

Abramowitz, J.S., Fabricant, L.E., Taylor, S., Deacon, B.J., McKay, D., and Storch, E.A. (2014). The relevance of analogue studies for understanding obsessions and compulsions. Clin Psychol Rev 34, 206-217.

Adam, Y., Meinlschmidt, G., Gloster, A.T., and Lieb, R. (2012). Obsessive-compulsive disorder in the community: 12-month prevalence, comorbidity and impairment. Soc Psychiatry Psychiatr Epidemiol 47, 339-349.

Albert, U., Aguglia, A., Chiarle, A., Bogetto, F., and Maina, G. (2013). Metabolic syndrome and obsessive-compulsive disorder: a naturalistic Italian study. Gen Hosp Psychiatry 35, 154-159.

Alvarenga, P.G., Cesar, R.C., Leckman, J.F., Moriyama, T.S., Torres, A.R., Bloch, M.H., Coughlin, C.G., Hoexter, M.Q., Manfro, G.G., Polanczyk, G.V., et al. (2015). Obsessive-compulsive symptom dimensions in a population-based, cross-sectional sample of school-aged children. J Psychiatr Res 62, 108-114.

American Psychiatric Association (2013). Diagnostic and Statistical Manual of Mental Disorders, Vol 5th edn. (Arlington, VA).

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B (Methodological) 57, 289-300.

Bloch, M.H., Landeros-Weisenberger, A., Rosario, M.C., Pittenger, C., and Leckman, J.F. (2008). Meta-analysis of the symptom structure of obsessive-compulsive disorder. Am J Psychiatry 165, 1532-1542.

Brakoulias, V., Starcevic, V., Martin, A., Berle, D., Milicevic, D., and Viswasam, K. (2016). The familiality of specific symptoms of obsessive-compulsive disorder. Psychiatry Res 239, 315-319.

Bralten, J., van Hulzen, K.J., Martens, M.B., Galesloot, T.E., Arias Vasquez, A., Kiemeney, L.A., Buitelaar, J.K., Muntjewerff, J.W., Franke, B., and Poelmans, G. (2018). Autism spectrum disorders and autistic traits share genetics and biology. Mol Psychiatry 23, 1205-1212.

Burton, C.L., Park, L.S., Corfield, E.C., Forget-Dubois, N., Dupuis, A., Sinopoli, V.M., Shan, J., Goodale, T., Shaheen, S.M., Crosbie, J., et al. (2018). Heritability of obsessive-compulsive trait dimensions in youth from the general population. Transl Psychiatry 8, 191.

Christie L Burton, C.L., Wright, L., Shan, J., Xiao, B., Dupuis, A., Goodale, T, Shaheen, S.M., Corfield, E.C., Arnold, P.D., Schachar, R.J., Crosbie, J. Utility of the SWAN Scale for ADHD Trait-Based Genetic Research: A Validity and Polygenic Risk Study bioRxiv 248484

Calamari, J.E., Wiegartz, P.S., Riemann, B.C., Cohen, R.J., Greer, A., Jacobi, D.M., Jahn, S.C., and Carmin, C. (2004). Obsessive-compulsive disorder subtypes: an attempted replication and extension of a symptom-based taxonomy. Behav Res Ther 42, 647-670.

Calkins, M.E., Merikangas, K.R., Moore, T.M., Burstein, M., Behr, M.A., Satterthwaite, T.D., Ruparel, K., Wolf, D.H., Roalf, D.R., Mentch, F.D., et al. (2015). The Philadelphia Neurodevelopmental Cohort: constructing a deep phenotyping collaborative. J Child Psychol Psychiatry 56, 1356-1369.

Chacon, P., Rosario-Campos, M.C., Pauls, D.L., Hounie, A.G., Curi, M., Akkerman, F., Shimabokuro, F.H., de Mathis, M.A., Lopes, A.C., Hasler, G., *et al.* (2007). Obsessive-compulsive symptoms in sibling pairs concordant for obsessive-compulsive disorder. Am J Med Genet B Neuropsychiatr Genet *144B*, 551-555.
Chiu, S.L., Chen, C.M., and Cline, H.T. (2008). Insulin receptor signaling regulates synapse number, dendritic plasticity, and circuit function in vivo. Neuron *58*, 708-719.

Clarke, D.W., Mudd, L., Boyd, F.T., Jr., Fields, M., and Raizada, M.K. (1986). Insulin is released from rat brain neuronal cells in culture. J Neurochem *47*, 831-836.

Crosbie, J., Arnold, P., Paterson, A., Swanson, J., Dupuis, A., Li, X., Shan, J., Goodale, T., Tam, C., Strug, L.J., *et al.* (2013). Response inhibition and ADHD traits: correlates and heritability in a community sample. J Abnorm Child Psychol *41*, 497-507.

de Leeuw, C.A., Mooij, J.M., Heskes, T., and Posthuma, D. (2015). MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput Biol *11*, e1004219.

den Braber, A., Zilhao, N.R., Fedko, I.O., Hottenga, J.J., Pool, R., Smit, D.J., Cath, D.C., and Boomsma, D.I. (2016). Obsessive-compulsive symptoms in a large population-based twin-family sample are predicted by clinically based polygenic scores and by genome-wide SNPs. Transl Psychiatry *6*, e731.

Dou, J.T., Chen, M., Dufour, F., Alkon, D.L., and Zhao, W.Q. (2005). Insulin receptor signaling in long-term memory consolidation following spatial learning. Learn Mem *12*, 646-655.

du Mortier, J.A.M., Visser, H.A.D., van Balkom, A., van Megen, H., Hoogendoorn, A.W., Glas, G., and van Oppen, P. (2019). Examining the factor structure of the self-report Yale-Brown Obsessive Compulsive Scale Symptom Checklist. Psychiatry Res *271*, 299-305.

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

Euesden, J., Lewis, C.M., and O'Reilly, P.F. (2015). PRSice: Polygenic Risk Score software. Bioinformatics *31*, 1466-1468.

Fullana, M.A., Mataix-Cols, D., Caspi, A., Harrington, H., Grisham, J.R., Moffitt, T.E., and Poulton, R. (2009). Obsessions and compulsions in the community: prevalence, interference, help-seeking, developmental stability, and co-occurring psychiatric conditions. Am J Psychiatry *166*, 329-336.

Gibbs N. Nonclinical populations in research on Obsessive-Compulsive Disorder: A critical review. Clinical Psychology Review 1996

Glickman, M.E., Rao, S.R., and Schultz, M.R. (2014). False discovery rate control is a recommended alternative to Bonferroni-type adjustments in health studies. J Clin Epidemiol *67*, 850-857.

Goodman, W.K., Storch, E.A., Geffken, G.R., and Murphy, T.K. (2006). Obsessive-compulsive disorder in Tourette syndrome. J Child Neurol *21*, 704-714.

Gur, R.C., Calkins, M.E., Satterthwaite, T.D., Ruparel, K., Bilker, W.B., Moore, T.M., Savitt, A.P., Hakonarson, H., and Gur, R.E. (2014). Neurocognitive growth charting in psychosis spectrum youths. JAMA Psychiatry *71*, 366-374.

Hagenaars, S.P., Harris, S.E., Davies, G., Hill, W.D., Liewald, D.C., Ritchie, S.J., Marioni, R.E., Fawns-Ritchie, C., Cullen, B., Malik, R., *et al.* (2016). Shared genetic aetiology between cognitive functions and physical and mental health in UK Biobank (N=112 151) and 24 GWAS consortia. Mol Psychiatry *21*, 1624-1632.

Harris, S.E., Hagenaars, S.P., Davies, G., David Hill, W., Liewald, D.C.M., Ritchie, S.J., Marioni, R.E., Metastroke Consortium, I.C.f.B.P.G.-W.A.S., International Consortium for Blood Pressure Genome-Wide Association, S., Aging, C.C., *et al.* (2017). Molecular genetic contributions to self-rated health. Int J Epidemiol *46*, 994-1009.

Iervolino, A.C., Rijsdijk, F.V., Cherkas, L., Fullana, M.A., and Mataix-Cols, D. (2011). A multivariate twin study of obsessive-compulsive symptom dimensions. Arch Gen Psychiatry *68*, 637-644.
International Obsessive Compulsive Disorder Foundation Genetics, C., and Studies, O.C.D.C.G.A. (2017). Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis. Mol Psychiatry.

Isomura, K., Brander, G., Chang, Z., Kuja-Halkola, R., Ruck, C., Hellner, C., Lichtenstein, P., Larsson, H., Mataix-Cols, D., and Fernandez de la Cruz, L. (2017). Metabolic and Cardiovascular Complications in Obsessive-Compulsive Disorder: A Total Population, Sibling Comparison Study With Long-Term Follow-up. Biol Psychiatry.

Janardhan Reddy, Y.C., Sundar, A.S., Narayanaswamy, J.C., and Math, S.B. (2017). Clinical practice guidelines for Obsessive-Compulsive Disorder. Indian J Psychiatry *59*, S74-S90.

Karno, M., Golding, J.M., Sorenson, S.B., and Burnam, M.A. (1988). The epidemiology of obsessive-compulsive disorder in five US communities. Arch Gen Psychiatry *45*, 1094-1099.

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., Williamson, D., and Ryan, N. (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. Journal of the American Academy of Child and Adolescent Psychiatry *36*, 980-988.

Kohlrausch, F.B., Giori, I.G., Melo-Felippe, F.B., Vieira-Fonseca, T., Velarde, L.G., de Salles Andrade, J.B., and Fontenelle, L.F. (2016). Association of GRIN2B gene polymorphism and Obsessive Compulsive disorder and symptom dimensions: A pilot study. Psychiatry Res *243*, 152-155.

Kontoangelos, K., Raptis, A.E., Papageorgiou, C.C., Papadimitriou, G.N., Rabavilas, A.D., Dimitriadis, G., and Raptis, S.A. (2013). The association of the metabolic profile in diabetes mellitus type 2 patients with obsessive-compulsive symptomatology and depressive symptomatology: new insights. Int J Psychiatry Clin Pract *17*, 48-55.

Leckman, J.F., Grice, D.E., Boardman, J., Zhang, H., Vitale, A., Bondi, C., Alsobrook, J., Peterson, B.S., Cohen, D.J., Rasmussen, S.A., *et al.* (1997). Symptoms of obsessive-compulsive disorder. Am J Psychiatry *154*, 911-917.

Li, Y., Willer, C.J., Ding, J., Scheet, P., and Abecasis, G.R. (2010). MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. Genet Epidemiol *34*, 816-834.

Mancini, F., and Gangemi, A. (2004). Fear of guilt from behaving irresponsibly in obsessive-compulsive disorder. J Behav Ther Exp Psychiatry *35*, 109-120.

Margolis, R.U., and Altszuler, N. (1967). Insulin in the cerebrospinal fluid. Nature *215*, 1375-1376.

Mataix-Cols, D., Rauch, S.L., Manzo, P.A., Jenike, M.A., and Baer, L. (1999). Use of factor-analyzed symptom dimensions to predict outcome with serotonin reuptake inhibitors and placebo in the treatment of obsessive-compulsive disorder. Am J Psychiatry *156*, 1409-1416.

Mathews, C.A., Delucchi, K., Cath, D.C., Willemsen, G., and Boomsma, D.I. (2014). Partitioning the etiology of hoarding and obsessive-compulsive symptoms. Psychol Med *44*, 2867-2876.

Mattheisen, M., Samuels, J.F., Wang, Y., Greenberg, B.D., Fyer, A.J., McCracken, J.T., Geller, D.A., Murphy, D.L., Knowles, J.A., Grados, M.A.*, et al.* (2015). Genome-wide association study in obsessive-compulsive disorder: results from the OCGAS. Mol Psychiatry *20*, 337-344.

McKay, D., Abramowitz, J.S., Calamari, J.E., Kyrios, M., Radomsky, A., Sookman, D., Taylor, S., and Wilhelm, S. (2004). A critical evaluation of obsessive-compulsive disorder subtypes: symptoms versus mechanisms. Clin Psychol Rev *24*, 283-313.

Nyholt, D.R. (2014). SECA: SNP effect concordance analysis using genome-wide association summary results. Bioinformatics *30*, 2086-2088.

Park, L.S., Burton, C.L., Dupuis, A., Shan, J., Storch, E.A., Crosbie, J., Schachar, R.J., and Arnold, P.D. (2016). The Toronto Obsessive-Compulsive Scale: Psychometrics of a Dimensional Measure of Obsessive-Compulsive Traits. Journal of the American Academy of Child and Adolescent Psychiatry *55*, 310-318.e314.

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

Rachman, S. (1993). Obsessions, responsibility and guilt. Behav Res Ther *31*, 149-154.

Satterthwaite, T.D., Connolly, J.J., Ruparel, K., Calkins, M.E., Jackson, C., Elliott, M.A., Roalf, D.R., Ryan Hopsona, K.P., Behr, M., Qiu, H.*, et al.* (2016). The Philadelphia Neurodevelopmental Cohort: A publicly available resource for the study of normal and abnormal brain development in youth. Neuroimage *124*, 1115-1119.

Satterthwaite, T.D., Elliott, M.A., Ruparel, K., Loughead, J., Prabhakaran, K., Calkins, M.E., Hopson, R., Jackson, C., Keefe, J., Riley, M.*, et al.* (2014). Neuroimaging of the Philadelphia neurodevelopmental cohort. Neuroimage *86*, 544-553.

Scott, R.A., Lagou, V., Welch, R.P., Wheeler, E., Montasser, M.E., Luan, J., Magi, R., Strawbridge, R.J., Rehnberg, E., Gustafsson, S.*, et al.* (2012). Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. Nat Genet *44*, 991-1005.

Scott, R.A., Scott, L.J., Magi, R., Marullo, L., Gaulton, K.J., Kaakinen, M., Pervjakova, N., Pers, T.H., Johnson, A.D., Eicher, J.D.*, et al.* (2017). An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. Diabetes *66*, 2888-2902.

Shafran, R., and Watkins, E. (1996). Guilt in obsessive-compulsive disorder. Journal of Anxiety Disorders *10*, 509-516.

Shapiro, L.J., and Stewart, E.S. (2011). Pathological guilt: a persistent yet overlooked treatment factor in obsessive-compulsive disorder. Ann Clin Psychiatry *23*, 63-70.

Stewart, S.E., Yu, D., Scharf, J.M., Neale, B.M., Fagerness, J.A., Mathews, C.A., Arnold, P.D., Evans, P.D., Gamazon, E.R., Davis, L.K., *et al.* (2013). Genome-wide association study of obsessive-compulsive disorder. Mol Psychiatry *18*, 788-798.

Summerfeldt, L.J., Richter, M.A., Antony, M.M., and Swinson, R.P. (1999). Symptom structure in obsessive-compulsive disorder: a confirmatory factor-analytic study. Behav Res Ther *37*, 297-311.

Tallis, F. (1994). Obsessions, responsibility and guilt: two case reports suggesting a common and specific aetiology. Behav Res Ther *32*, 143-145.

Taylor, M.J., Martin, J., Lu, Y., Brikell, I., Lundstrom, S., Larsson, H., and Lichtenstein, P. (2018). Association of Genetic Risk Factors for Psychiatric Disorders and Traits of These Disorders in a Swedish Population Twin Sample. JAMA Psychiatry.

Taylor, S. (2011). Etiology of obsessions and compulsions: a meta-analysis and narrative review of twin studies. Clin Psychol Rev *31*, 1361-1372.

Taylor, S. (2013). Molecular genetics of obsessive-compulsive disorder: a comprehensive meta-analysis of genetic association studies. Mol Psychiatry *18*, 799-805.

Ter Horst, K.W., Lammers, N.M., Trinko, R., Opland, D.M., Figee, M., Ackermans, M.T., Booij, J., van den Munckhof, P., Schuurman, P.R., Fliers, E., *et al.* (2018). Striatal dopamine regulates systemic glucose metabolism in humans and mice. Sci Transl Med *10*.

Torresan, R.C., Ramos-Cerqueira, A.T., Shavitt, R.G., do Rosario, M.C., de Mathis, M.A., Miguel, E.C., and Torres, A.R. (2013). Symptom dimensions, clinical course and comorbidity in men and women with obsessive-compulsive disorder. Psychiatry Res *209*, 186-195.

van de Vondervoort, I., Poelmans, G., Aschrafi, A., Pauls, D.L., Buitelaar, J.K., Glennon, J.C., and Franke, B. (2016). An integrated molecular landscape implicates the regulation of dendritic spine formation through insulin-related signalling in obsessive-compulsive disorder. J Psychiatry Neurosci *41*, 280-285.

van Grootheest, D.S., Cath, D., Hottenga, J.J., Beekman, A.T., and Boomsma, D.I. (2009). Genetic factors underlie stability of obsessive-compulsive symptoms. Twin Res Hum Genet *12*, 411-419.

Weissman, M.M., Bland, R.C., Canino, G.J., Greenwald, S., Hwu, H.G., Lee, C.K., Newman, S.C., Oakley-Browne, M.A., Rubio-Stipec, M., Wickramaratne, P.J., *et al.* (1994). The cross national epidemiology of obsessive compulsive disorder. The Cross National Collaborative Group. J Clin Psychiatry *55 Suppl*, 5-10.

Wheeler, E., Leong, A., Liu, C.T., Hivert, M.F., Strawbridge, R.J., Podmore, C., Li, M., Yao, J., Sim, X., Hong, J., *et al.* (2017). Impact of common genetic determinants of Hemoglobin A1c on type 2 diabetes risk and diagnosis in ancestrally diverse populations: A transethnic genome-wide meta-analysis. PLoS Med *14*, e1002383.

Willer, C.J., Li, Y., and Abecasis, G.R. (2010). METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics *26*, 2190-2191.

Winocour, P.H., Main, C.J., Medlicott, G., and Anderson, D.C. (1990). A psychometric evaluation of adult patients with type 1 (insulin-dependent) diabetes mellitus: prevalence of psychological

dysfunction and relationship to demographic variables, metabolic control and complications. Diabetes Res *14*, 171-176.

Zhao, W., Chen, H., Xu, H., Moore, E., Meiri, N., Quon, M.J., and Alkon, D.L. (1999). Brain insulin receptors and spatial memory. Correlated changes in gene expression, tyrosine phosphorylation, and signaling molecules in the hippocampus of water maze trained rats. J Biol Chem *274*, 34893-34902.

Table 1. Item content of and loadings on the eight factors that constitute the best fitting model to explain the variance in the total score of the 22 items from the questionnaire of OCD symptoms that was completed by 650 participants from the PNC cohort.

Factor 1	*Impairment*	(14.63% of the variance in the total score explained)
Items		Factor loadings
OCD024	Did these thoughts and behaviors prevent you from doing things you normally would do?	0.537
OCD025	Did having these thoughts or behaviors bother you a lot?	0.717
OCD032	You told me (insert endorsed thoughts/behaviors). How much did having these thoughts/behaviors upset or bother you? How much did you ever feel upset or disappointed with yourself because of your thoughts/behaviors?	0.741
OCD033	How much did the thoughts/behaviors you have told me about cause problems for you at home, at school/work, or with your family or friends?	0.716
OCD034	Did you stay home from school/work because of your behaviors/thoughts?	0.339

Factor 2	*Symmetry/counting/ordering*	(10.58% of the variance in the total score explained)
Items		Factor loadings
OCD007	Have you ever been bothered by thoughts that don't make sense to you, that come over and over again and won't go away, such as need for symmetry/exactness?	0.682
OCD012	Have you ever had to do something over and over again - that would have made you feel really nervous if you couldn't do it, like: counting?	0.588
OCD013	Have you ever had to do something over and over again - that would have made you feel really nervous if you couldn't do it, like: checking (for example, doors, locks, ovens)?	0.545
OCD016	Have you ever had to do something over and over again - that would have made you feel really nervous if you couldn't do it, like: ordering or arranging things?	0.776
OCD017	Have you ever had to do something over and over again - that would have made you feel really nervous if you couldn't do it, like: doing things over and over again at bedtime, like arranging the pillows, sheets, or other things?	0.513

Factor 3	*Contamination/cleaning*	(7.54% of the variance in the total score explained)
Items		Factor loadings
OCD003	Have you ever been bothered by thoughts that don't make sense to you, that come over and over again and won't go away, such as thoughts about contamination/germs/illness?	0.871
OCD011	Have you ever had to do something over and over again - that would have made you feel really nervous if you couldn't do it, like: cleaning or washing (for example, your hands, house)?	0.757

Table 1 - continued.		
Factor 4	***Aggressive taboo thoughts***	(6.02% of the variance in the total score explained)
Items		Factor loadings
OCD001	Have you ever been bothered by thoughts that don't make sense to you, that come over and over again and won't go away, such as concern with harming others/self?	0.503
OCD002	Have you ever been bothered by thoughts that don't make sense to you, that come over and over again and won't go away, such as pictures of violent things?	0.845
OCD006	Have you ever been bothered by thoughts that don't make sense to you, that come over and over again and won't go away, such as forbidden/bad thoughts?	0.578
Factor 5	***Repetition***	(5.56% of the variance in the total score explained)
Items		Factor loadings
OCD014	Have you ever had to do something over and over again - that would have made you feel really nervous if you couldn't do it, like: getting dressed over and over again?	0.782
OCD015	Have you ever had to do something over and over again - that would have made you feel really nervous if you couldn't do it, like: going in and out a door over and over again?	0.662
Factor 6	***Guilty taboo thoughts***	(5.04% of the variance in the total score explained)
Items		Factor loadings
OCD004	Have you ever been bothered by thoughts that don't make sense to you, that come over and over again and won't go away, such as fear that you would do something/say something bad without intending to?	0.758
OCD005	Have you ever been bothered by thoughts that don't make sense to you, that come over and over again and won't go away, such as feelings that bad things that happened were your fault?	0.722
Factor 7	***Distress***	(4.68% of the variance in the total score explained)
Items		Factor loadings
OCD009	Did these thoughts continue to bother you no matter how hard you tried to get rid of them or ignore them?	0.770
OCD010	Did you try not to think about (thoughts), try to keep them out of your head, or try to push the thoughts away?	0.552
Factor 8	***Religious taboo thoughts***	(4.56% of the variance in the total score explained)
Items		Factor loadings
OCD008	Have you ever been bothered by thoughts that don't make sense to you, that come over and over again and won't go away, such as religious thoughts?	0.722

Taken together, the eight factors explain 58.6% of the variance in the total score of OCD symptoms in the general population from the questionnaire.

Table 2. PRS-based results for shared genetic etiology between OCD and the total OCD symptom score as well as the scores for six OCD symptom factors [#]		
Total OCD symptoms	P_T	0.001
	P-value	2.93E-01
	R^2	0.045%
	nSNPs	1560
Impairment	**P_T**	**0.2**
	P-value	**4.22E-03**
	R^2	**1.061%**
	nSNPs	**158541**
Symmetry/counting/ordering	P_T	0.001
	P-value	1.66E-01
	R^2	0.145%
	nSNPs	1560
Contamination/cleaning	**P_T**	**0.4**
	P-value	**6.47E-07**
	R^2	**3.541%**
	nSNPs	**273491**
Aggressive taboo thoughts	P_T	0.4
	P-value	1.48E-01
	R^2	0.168%
	nSNPs	273512
Guilty taboo thoughts	**P_T**	**0.2**
	P-value	**4.11E-10**
	R^2	**5.636%**
	nSNPs	**158545**
Distress	P_T	0.2
	P-value	5.23E-02
	R^2	0.404%
	nSNPs	158541

[#] Shown in this table are the best SNP P-value thresholds (P_T) for the PRSice analyses between OCD - 'base' sample - and the total OCD symptom score as well as six OCD symptom factors ('target' samples), their FDR-corrected P-values for shared genetic etiology, the variance explained in the target sample phenotypes (R^2), and the number of SNPs (nSNPs). Significant findings are indicated in bold.

Table 3. PRS-based results for shared genetic etiology between five peripheral insulin-signalling-related traits and OCD/OCD symptoms [#]

		T2D	HbA1c	Fasting Insulin	Fasting Glucose	2hGlu
OCD	P_T	**0.1**	0.001	**0.2**	0.4	**0.5**
	P-value	**4.00E-02**	2.08E-01	**4.22E-07**	1.10E-01	**1.39E-04**
	R^2	**0.032%**	0.007%	**0.261%**	0.015%	**0.136%**
	nSNPs	**138297**	1152	**12558**	21585	**24147**
Total OCD symptoms	P_T	**0.5**	**0.2**	0.1	0.3	0.2
	P-value	**2.04E-03**	**2.69E-03**	1.34E-01	2.11E-01	1.97E-01
	R^2	**1.261%**	**1.185%**	0.189%	0.100%	0.112%
	nSNPs	**322098**	**77169**	7353	17010	10703
Impairment	P_T	0.5	**0.2**	0.001	0.05	0.2
	P-value	3.59E-01	**6.43E-03**	1.87E-01	5.34E-02	2.49E-01
	R^2	0.020%	**0.948%**	0.122%	0.400%	0.071%
	nSNPs	322098	**77169**	342	4794	10703
Symmetry/counting/ordering	P_T	**0.5**	0.4	0.4	**0.4**	0.4
	P-value	**5.06E-07**	2.29E-01	3.43E-01	**1.91E-02**	9.74E-02
	R^2	**3.982%**	0.085%	0.025%	**0.659%**	0.258%
	nSNPs	**322098**	130416	20910	**21114**	19482
Contamination/cleaning	P_T	0.5	**0.001**	0.1	**0.001**	**0.05**
	P-value	8.07E-02	**2.00E-02**	2.00E-01	**1.22E-02**	**8.31E-03**
	R^2	0.301%	**0.647%**	0.109%	**0.777%**	**0.879%**
	nSNPs	322043	**1126**	7353	**518**	**3348**
Aggressive taboo thoughts	P_T	**0.4**	**0.05**	0.4	**0.05**	0.001
	P-value	**6.08E-07**	**7.57E-05**	2.04E-01	**4.77E-02**	7.82E-02
	R^2	**3.616%**	**2.184%**	0.105%	**0.427%**	0.309%
	nSNPs	**275379**	**24321**	20909	**4794**	210
Guilty taboo thoughts	P_T	0.05	0.1	0.001	0.5	**0.1**
	P-value	1.96E-01	2.27E-01	1.20E-01	8.08E-02	**3.38E-03**
	R^2	0.112%	0.086%	0.213%	0.301%	**1.122%**
	nSNPs	54161	43920	342	25109	**5932**
Distress	P_T	0.5	**0.5**	**0.1**	0.4	0.2
	P-value	1.10E-01	**1.12E-02**	**3.67E-02**	8.61E-02	7.09E-02
	R^2	0.232%	**0.800%**	**0.492%**	0.286%	0.332%
	nSNPs	322098	**152062**	**7353**	21114	10703

[#] Shown in this table are the best SNP P-value thresholds (P_T) for the PRSice analyses between five peripheral insulin signalling-related traits ('base' samples) and OCD as well as obsessive-compulsive symptom factors ('target' samples), their FDR-corrected association p-value (P-value), the variance explained (R^2) in the target sample phenotypes, and the number of SNPs (nSNPs). Significant findings are indicated in bold.

Abbreviations: T2D = type 2 diabetes ; 2hGlu = 2-hour glucose

Table 4. Summary of the significant results of PRS-based and SECA analyses. Shown in this table are the significant FDR-corrected association p-values at the best SNP P-value thresholds along with the variance explained for each of the 'base' and 'target' sample pairs. Cell color indicates the direction of concordance identified in SECA analyses: green – negative concordance, red – positive concordance. Abbreviations: PNC – Philadelphia Neurodevelopmental Cohort, PGC – Psychiatric Genomics Consortium, TOCS - Toronto Obsessive-Compulsive Scale

			'Base' sample					
			OCD	Diabetes Type 2	2 h Glucose	Fasting Glucose	Fasting Insulin	HbA1c
PNC		Total Score		2.04E-03 1.26%				2.69E-03 1.19%
		Impairment	4.22E-03 1.06%					6.43E-03 0.95%
		Symmetry/counting/ordering		5.06E-07 3.98%		1.91E-02 0.66%		
		Contamination/cleaning	6.47E-07 3.54%		8.31E-03 0.88%	1.22E-02 0.78%		2.00E-02 0.65%
		Aggressive taboo thoughts		6.08E-07 3.62%		4.77E-02 0.43%		7.57E-05 2.18%
		Guilty taboo thoughts	4.11E-10 5.64%		3.38E-03 1.12%			
		Distress					3.67E-02 0.49%	1.12E-02 0.80%
PGC		OCD			1.39E-04 0.14%		4.22E-07 0.26%	
Spit for Science		Symmetry/counting/ordering$_{TOCS}$	2.87E-07 0.49%					
		Contamination/cleaning$_{TOCS}$	3.62E-04 0.23%	2.37E-04 0.28%				4.88E-02 0.05%

'Target' sample