ARTICLE

Open Access

Shared genetic etiology between obsessive-compulsive disorder, obsessive-compulsive symptoms in the population, and insulin signaling

Janita Bralten[1,2], Joanna Widomska[2,3], Ward De Witte[1], Dongmei Yu[4,5], Carol A. Mathews[6], Jeremiah M. Scharf[4,5,7], Jan Buitelaar [2,3,8], Jennifer Crosbie[9,10], Russell Schachar[9,10], Paul Arnold [11,12,13], Mathieu Lemire[9], Christie L. Burton [9], Barbara Franke [1,2,14] and Geert Poelmans[1]

Abstract

Obsessive-compulsive symptoms (OCS) in the population have been linked to obsessive-compulsive disorder (OCD) in genetic and epidemiological studies. Insulin signaling has been implicated in OCD. We extend previous work by assessing genetic overlap between OCD, population-based OCS, and central nervous system (CNS) and peripheral insulin signaling. We conducted genome-wide association studies (GWASs) in the population-based Philadelphia Neurodevelopmental Cohort (PNC, 650 children and adolescents) of the total OCS score and six OCS factors from an exploratory factor analysis of 22 questions. Subsequently, we performed polygenic risk score (PRS)-based analysis to assess shared genetic etiologies between clinical OCD (using GWAS data from the Psychiatric Genomics Consortium), the total OCS score and OCS factors. We then performed gene-set analyses with a set of OCD-linked genes centered around CNS insulin-regulated synaptic function and PRS-based analyses for five peripheral insulin signaling-related traits. For validation purposes, we explored data from the independent Spit for Science population cohort (5,047 children and adolescents). In the PNC, we found a significant shared genetic etiology between OCD and 'guilty taboo thoughts'. In the Spit for Science cohort, we additionally observed genetic sharing between 'symmetry/counting/ordering' and 'contamination/cleaning'. The CNS insulin-linked gene-set also associated with 'symmetry/counting/ordering' in the PNC. Further, we identified genetic sharing between peripheral insulin signaling-related traits: type 2 diabetes with 'aggressive taboo thoughts', and levels of fasting insulin and 2 h glucose with OCD. In conclusion, OCD, OCS in the population and insulin-related traits share genetic risk factors, indicating a common etiological mechanism underlying somatic and psychiatric disorders.

Introduction

Obsessive-compulsive disorder (OCD) is a heterogeneous psychiatric condition characterized by persistent, intrusive thoughts and urges (obsessions) and repetitive, intentional behaviours (compulsions)[1]. OCD affects 2–3%

of the world's population[2,3]. OCD is moderately heritable, with approximately 40% of the phenotypic variance explained by genetic factors, and a higher genetic heritability has been reported in childhood onset OCD[4–6]. The genetic architecture of OCD is complex, with multiple genetic variants of small effect size contributing to its etiology. This has hampered the identification and replication of genetic susceptibility factors. A 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[7]. Neither

Correspondence: Geert Poelmans (geert.poelmans@radboudumc.nl)
[1]Department of Human Genetics, Radboud University Medical Center, Nijmegen, The Netherlands
[2]Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands
Full list of author information is available at the end of the article.
These authors contributed equally: Janita Bralten, Joanna Widomska, Barbara Franke, Geert Poelmans

of the two independent genome-wide association studies (GWASs) of OCD[8,9] nor a subsequent meta-analysis (2,688 cases and 7,037 controls)[10]—including children, adolescents and adults—yielded genome-wide significant findings, likely due to lack of power. However, the meta-analysis demonstrated that the polygenic signal from either sample predicted OCD status in the other sample, indicating the polygenic nature of the disorder[10].

The diagnosis of OCD is based solely on clinical symptoms, and no genetic or biological markers are available with sufficient specificity and accuracy to be clinically actionable[11]. However, factor analyses of OCD symptoms have consistently identified specific OCD symptom clusters or dimensions, with the most reliable including contamination/cleaning, doubt/checking, symmetry/ordering, and unacceptable/taboo thoughts[11–14].

Obsessive-compulsive symptoms (OCS) are also present in the general population[15–17]. Indeed, 21 to 38 % of individuals in the population endorse obsessions and/or compulsions, although only a small minority (2–3 %) meet the DSM-5 criteria for a clinical OCD diagnosis[18,19]. OCS are also heritable, with their heritability ranging from 30 to 77%[20,21]. In addition to contributing to overall OCS, genetic factors contribute to specific OCS dimensions, including contamination/cleaning[22–24] and checking/ordering[24,25]. Genetic overlap between clinical OCD and OCS in the population is suggested by the fact that polygenic risk scores (PRS) based on OCD GWAS data significantly predict OCS in two population-based samples of 6,931 and 3,982 individuals, respectively[20,26]. Moreover, a very recent analysis found that compulsive symptoms in the general population overlap with the genetic liability for clinical OCD[27], and incorporating compulsive symptom GWAS data in a meta-analysis with OCD GWAS data yielded new findings in gene-based and gene enrichment analyses. Therefore, genetic studies of OCS 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.

In a study aiming to identify molecular mechanisms underlying OCD, we earlier performed integration of the top-ranked results from the existing GWASs with genes implicated in OCD through other evidence. This resulted in a 'molecular landscape' that suggested the involvement of genes regulating postsynaptic dendritic spine formation and function through central nervous system (CNS) insulin-dependent signalling[28]. Support for a role of dysregulated insulin signaling in OCD and OCS comes from studies showing increased OCS in men with type 1 diabetes[29] and from a study indicating that OCD patients have a higher risk of developing type 2 diabetes (T2D)[30]. Furthermore, OCS were found to be positively correlated with blood levels of glycosylated hemoglobin (HbA1c), a diagnostic measure of T2D[31], and OCD patients had markedly higher levels of fasting glucose (a characteristic of T2D)[32].

In this paper, we aimed to assess the presence and extent of genetic overlap between OCD, OCS in the population, and insulin signaling, using the largest available data sets. Specifically, we parsed phenotypic heterogeneity using an exploratory factor analysis of OCS measured in a population cohort of children and adolescents. Subsequently, using PRS-based analyses, we investigated the presence of shared genetic etiologies between OCD and the total and factorized OCS. We then assessed genetic sharing between OCD, OCS, and insulin-related traits. In addition to assessing shared genetic etiologies, we tested for potential overlapping biology using gene-set analyses. Lastly, we extended our findings in an independent population cohort of children and adolescents.

Methods
Sample, phenotypic, and genetic data
We studied OCS in the Philadelphia Neurodevelopmental Cohort (PNC)[33–37], which includes 8,719 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 to the Children's Hospital of Philadelphia health care network. OCS were assessed with GO-ASSESS, a computerized version of the Kiddie-Schedule for Affective Disorders and Schizophrenia (K-SADs)[38]. For the current study, we selected 22 GO-ASSESS questions that corresponded to the diagnostic criteria for OCD (Supplementary Table 1). Participants were included if they answered the questions related to obsessions and/or compulsions. If those questions were all answered "no", we allowed the questions on the consequences of obsessions and compulsions to be left blank, as no consequences are expected if no symptoms are present. The scores for each of the questions (0 for "no" and 1 for "yes") were then summed to create a total OCS score (range 0–22). Genome-wide genotyping in the PNC cohort was performed in waves using six different genotyping platforms (details in Supplementary Methods). As a primary aim of our study was to assess the genetic overlap between OCD and OCS 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. This resulted in a final sample of 650 individuals for the subsequent factor and genome-wide association analyses.

Factor analysis
First, using SPSS 23 (SPSS Technologies, Armonk, NY, USA), we determined the internal consistency (Cronbach's α) of the 22 questions that constitute the total OCS scores in the 650 PNC participants. We then conducted a

factor analysis of the scores on the 22 questions using Promax rotation to determine the number of factors that, when combined, explains the largest portion of the observed variance in the total OCS score. Specifically, we considered scree plots, eigenvalues >1 and the cumulative variance explained when selecting the number of factors and assigned questions to factors based on the highest absolute loading value.

Genome-wide association analyses

Quality control filtering was applied to the genetic data to remove non-European individuals, single nucleotide polymorphisms (SNPs) with low minor allele frequency (MAF) (<0.05), poor genotype call rate (<95%), and deviations from Hardy-Weinberg equilibrium ($P < 1 \times 10^{-6}$) (details in the Supplementary Methods). The imputation protocol used MaCH[39] for haplotype phasing and minimac[40] for imputation. Imputed SNPs with low imputation quality score (r^2<0.8) and low MAF (<0.05) were removed. If the total OCS score or the scores for the OCS factors fell within the limits of a normal distribution (i.e., a skewness and kurtosis between -1 and 1), we used a continuous trait design for the genome-wide association analysis. Otherwise, we used a pseudo case-control design, in which all individuals with a score of 0 for a factor were defined as 'controls' and compared against the 'pseudo cases', i.e., all individuals with a score of 1 or more for that factor. GWASs were carried out with mach2qtl[39] using the total OCS score and the scores for those factors that showed sufficient variation as phenotypes, with age, gender, and the four principal components from MDS included as covariates. GWASs were performed separately for each genotyping platform and combined in an inverse-variance-weighted meta-analysis using METAL[41], accounting for genomic inflation.

Shared genetic etiology analyses
OCD and OCS

First, we determined the level of shared genetic etiology between diagnosed OCD and OCS in the population. For this, we used the summary statistics from the meta-analysis of the two published GWASs of OCD[10] (data provided through the Psychiatric Genomics Consortium (PGC) for 2,688 OCD cases and 7,037 controls) as the 'base' sample for polygenic risk score (PRS)-based analyses in PRSice[42]. The summary statistics from the GWASs of the different OCS in the PNC were used as the 'target' samples for the PRS-based analyses. For details see Supplementary Methods.

OCD, OCS and peripheral insulin signaling-related traits

To determine the level of genetic sharing between five peripheral insulin signaling-related traits and OCD as well as OCS, we conducted PRS-based analyses in PRSice[42], as described above. As base samples, we used summary statistics data from GWASs of the following peripheral insulin signaling-related traits: type 2 diabetes (T2D) and the blood levels of four T2D markers: HbA1c, fasting insulin, fasting glucose and glucose 2 h after an oral glucose challenge (2hGlu) (details in Supplementary Methods). As target samples of the PRS-based analyses, we used the summary statistics from the OCD GWAS meta-analysis and the GWASs of the total OCS score and the scores for the OCS factors in the PNC. Multiple comparisons correction for all tests performed (i.e., for the tests assessing genetic sharing between OCD and OCS and the tests assessing genetic sharing between the five insulin-related traits and OCD as well as OCS) was done using the Benjamini-Hochberg false discovery rate (FDR) method. With this method, we aggregated the calculated p-values of the shared genetic etiology analyses[43,44], which is similar to the approaches used in earlier studies working with multiple phenotypes and PRSice[45,46].

Gene-set analyses

We first compiled a set of genes encoding proteins from our molecular landscape of OCD[28] (see above). In this paper by van de Vondervoort et al., the OCD landscape was built based on proteins that have been implicated in the disease through different types of genetic evidence. Firstly, proteins were included if their corresponding genes have been implicated in OCD through SNPs from the published GWASs that are associated at $P < 1.00E-04$ and are located within the gene and/or 100 kb of flanking downstream and upstream sequences. In addition, genes/proteins were included that have been implicated in other ways in OCD etiology. After critical evaluation of the literature, only genes/proteins that have received support through findings from (genetic) animal studies, gene mutations, and/or two or more independent candidate gene association studies (or at least nominal significance in meta-analysis) and/or mRNA/protein expression studies, were included. Our selection resulted in a set of 51 autosomal genes for subsequent analyses. Using the GWAS results of the total OCS score and the scores for the OCS factors, competitive gene-set analyses were then performed using the Multimarker Analysis of GenoMic Annotation (MAGMA) software[47], see Supplementary Methods. *P*-values were considered significant if they exceeded a Bonferroni-corrected threshold accounting for the number of phenotypes tested ($P < 0.05/7$ tests (total OCS score and six OCS factors) = 7.14E-03). For significant gene-set associations, we looked at the individual gene-wide *P*-values and applied Bonferroni correction ($P < 0.05/51$ genes in the gene-set = 9.80E-04).

Validation analyses in an independent population sample

In order to validate and possibly expand our findings, we performed PRS-based and gene-set analyses using data

from GWASs of OCS in an independent population sample: the 'Spit for Science' project, which includes 16,718 children and adolescents aged 6–17 years recruited from a local science museum[48]. OCS were measured using the Toronto Obsessive-Compulsive Scale (TOCS), a validated 21-item parent-or self-report questionnaire[15]. TOCS items are scored from −3 (far less often than others of the same age) to +3 (far more often than others of the same age). We first assessed which TOCS questions could be grouped into OCS factors similar to those calculated based on the PNC data. Only two OCS factors (being 'symmetry/counting/ordering' and 'contamination/cleaning') were similar and therefore could be used for validation purposes, see Supplementary Table 2. Genome-wide genotyping data for 5,047 individuals of Caucasian descent entered the 'continuous trait' GWAS analysis for each factor. A description of genotyping, quality control and imputation can be found elsewhere[49] and GWAS details in the Supplementary Methods. Using summary statistics of the GWASs of the two TOCS OCS factors, we examined the shared genetic etiology between OCD and the TOCS OCS factors, and between the five peripheral insulin signaling-related traits and the TOCS OCS factors. As described above for the PNC data, multiple comparisons correction was done using the FDR method for all tests performed in the Spit for Science cohort. Gene-set analyses between the set of 51 genes from the OCD landscape and the two TOCS OCS factors were also performed.

Results
Factor analysis
The internal consistency between the scores on the 22 OCS questions from the PNC was satisfactory (Cronbach's $\alpha = 0.69$). Supplementary Fig. 1A shows the total score distribution (mean = 6.4, s.d. = 3.35). Factor analysis revealed an eight factors solution as the best-fitting model, explaining 58.6% of the variance in the total score. We named these eight OCS factors 'impairment', 'symmetry/counting/ordering', 'contamination/cleaning', 'aggressive taboo thoughts', 'repetition', 'guilty taboo thoughts', 'distress', and 'religious taboo thoughts' (Table 1; factor score distributions in Supplementary Fig. 1B).

Genome-wide association analyses
Based on the distributions of the scores, we used a continuous trait design for the GWASs of the total OCS score and the factors 'impairment', 'symmetry/counting/ordering', and 'distress'. A pseudo case-control design was used for the factors 'contamination/cleaning' and 'aggressive taboo thoughts', and 'guilty taboo thoughts'. The distribution of the scores on the OCS factors 'repetition' and 'religious taboo thoughts' showed too little

variation to be taken forward (Supplementary Fig. 1B). Because of the lack of power—with only 650 individuals per GWAS—we do not report individual GWAS results, but we have used the GWAS results for PRS-based analyses.

Shared genetic etiology analyses
OCD and OCS
We found statistically significant evidence for a shared genetic etiology between diagnosed OCD and the population-based OCS factor 'guilty taboo thoughts' ($R^2 = 2.28\%$; $P = 2.52E-03$) (Supplementary Fig. 2 and Table 2).

OCD, OCS, and peripheral insulin signaling-related traits
We found statistically significant evidence for a shared genetic etiology between T2D and 'aggressive taboo thoughts' ($R^2 = 1.86\%$; $P = 5.95E-03$) (Supplementary Fig. 3A and Table 3). Fasting insulin levels showed genetic sharing with OCD ($R^2 = 0.26\%$; $P = 7.67E-05$) and for HbA1c and fasting glucose levels, we did not find evidence of genetic sharing. Lastly, we observed genetic sharing between 2 h glucose levels and OCD ($R^2 = 0.14\%$; $P = 4.75E-03$) (Supplementary Fig. 3B–E and Table 3).

Gene-set analyses
MAGMA-based gene-set analysis for the CNS insulin signalling genes extracted from our earlier-defined OCD landscape containing 33,329 SNPs (effective number of SNPs after adjusting for LD structure = 2,189) revealed a significant association with 'symmetry/counting/ordering' ($P = 4.08E-03$) (Supplementary Table 3). Within the significant gene-set, none of the individual genes showed gene-wide association (Supplementary Table 4). No significant associations were found with the total OCS score or the five other OCS factors.

Validation analyses in an independent population sample
Two OCS factors were similar between the PNC and Spit for Science cohort, i.e., 'symmetry/counting/ordering' and 'contamination/cleaning' (Supplementary Table 2 and Supplementary Fig. 4A, B). Using summary statistics of the GWASs of these two factors, we found that diagnosed OCD shows genetic sharing with 'symmetry/counting/ordering$_{TOCS}$' ($R^2 = 0.49\%$; FDR-adjusted $P = 2.42E-05$) and 'contamination/cleaning$_{TOCS}$' ($R^2 = 0.23\%$; FDR-adjusted $P = 4.07E-03$).

We also observed a shared genetic etiology between T2D and 'contamination/cleaning$_{TOCS}$' ($R^2 = 0.28\%$; FDR-adjusted $P = 1.59E-03$) (Supplementary Table 5 and Supplementary Fig. 5A–C). Gene-set analysis for the OCD landscape genes in the two OCS$_{TOCS}$ factors revealed no significant associations.

All results from the PRS-based analyses are summarized in Table 4.

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 obsessive-compulsive 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
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)

Table 1 continued

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 OCS score as well as the scores for six OCS factors.

	P_T	P-value	R^2	nSNPs
Total OCS score	0.05	4.72E-01	0.04%	26,653
Impairment	0.2	2.61E-01	0.51%	81,294
Symmetry/counting/ordering	0.001	3.53E-01	0.23%	1,007
Contamination/cleaning	0.2	1.12E-01	0.85%	81,294
Aggressive taboo thoughts	0.4	2.28E-01	0.46%	135,230
Guilty taboo thoughts	**0.3**	**2.52E-03**	**2.28%**	**110,197**
Distress	0.1	4.05E-01	0.13%	47,020

Shown in this table are the best SNP P-value thresholds (P_T) for the PRSice analyses between OCD ('base' sample) and the total OCS score and six OCS factors ('target' samples), their Benjamini-Hochberg adjusted P-values (P-value) 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.

Discussion

In this study, we extended previous work by assessing genetic etiologies between OCD, OCS in the population, and CNS and peripheral insulin signaling. While previous studies[20,26] have yielded a shared genetic etiology between OCD and the total population-based OCS score, our aalyses using phenotypic and genetic data of 650 children and adolescents from the population (PNC cohort) found genetic sharing between OCD and the OCS factor 'guilty taboo thoughts'. In the larger Spit for Science cohort ($n = 5,047$), we expanded our results by showing genetic sharing between OCD and 'symmetry/counting/ordering' as well as 'contamination/cleaning'. Our findings are in keeping with the literature suggesting (at least partial) genetic overlap between OCD and population-based OCS[20,22–24,27]. Since OCD is genetically correlated with other psychiatric disorders (e.g., Anorexia Nervosa, Major Depressive Disorder and Tourette Syndrome[50]), future studies investigating OCS as (a) shared trait(s) between

Table 3 PRS-based results for shared genetic etiology between five peripheral insulin-signaling-related traits and OCD and OCS.

'base' sample	'target' sample	P_T	*P*-value	R^2	nSNPs
Type 2 diabetes	OCD	0.1	2.80E-01	0.03%	138,256
	Total OCS score	0.2	3.36E-01	0.28%	83,578
	Impairment	0.3	4.10E-01	0.11%	114,988
	Symmetry/counting/ordering	0.3	2.56E-01	0.53%	114,988
	Contamination/cleaning	0.5	3.53E-01	0.24%	158,182
	Aggressive taboo thoughts	**0.2**	**5.95E-03**	**1.86%**	**83,578**
	Guilty taboo thoughts	0.5	1.89E-01	0.68%	158,182
	Distress	0.3	4.72E-01	0.02%	114,988
HbA1c	OCD	0.001	4.15E-01	0.01%	1,152
	Total OCS score	0.2	7.64E-02	1.00%	64,945
	Impairment	0.2	2.28E-01	0.58%	64,945
	Symmetry/counting/ordering	0.05	4.10E-01	0.11%	21,040
	Contamination/cleaning	0.001	3.21E-01	0.33%	1,030
	Aggressive taboo thoughts	0.05	7.99E-02	0.97%	21,040
	Guilty taboo thoughts	0.1	4.03E-01	0.15%	37,363
	Distress	0.05	4.07E-01	0.13%	21,040
Fasting Insulin	**OCD**	**0.2**	**7.67E-05**	**0.26%**	**12,557**
	Total OCS score	0.1	3.74E-01	0.19%	6,564
	Impairment	0.001	4.10E-01	0.11%	328
	Symmetry/counting/ordering	0.4	4.72E-01	0.04%	18,227
	Contamination/cleaning	0.1	4.47E-01	0.07%	6,564
	Aggressive taboo thoughts	0.4	4.66E-01	0.06%	18,227
	Guilty taboo thoughts	0.001	2.61E-01	0.52%	328
	Distress	0.1	2.80E-01	0.43%	6,564
Fasting Glucose	OCD	0.4	3.53E-01	0.02%	21,586
	Total OCS score	0.3	4.16E-01	0.10%	14,814
	Impairment	0.1	4.10E-01	0.12%	6,861
	Symmetry/counting/ordering	0.001	2.80E-01	0.39%	481
	Contamination/cleaning	0.001	2.28E-01	0.58%	481
	Aggressive taboo thoughts	0.5	2.71E-01	0.49%	21,798
	Guilty taboo thoughts	0.1	3.21E-01	0.33%	6,861
	Distress	0.05	3.53E-01	0.23%	4,271
2h Glucose	**OCD**	**0.5**	**4.75E-03**	**0.14%**	**24,148**
	Total OCS score	0.1	4.72E-01	0.02%	5,202
	Impairment	0.2	3.53E-01	0.22%	9,362
	Symmetry/counting/ordering	0.4	4.10E-01	0.11%	16,970
	Contamination/cleaning	0.05	2.71E-01	0.49%	2,962
	Aggressive taboo thoughts	0.001	2.28E-01	0.59%	186
	Guilty taboo thoughts	0.001	9.74E-02	0.90%	186
	Distress	0.2	3.53E-01	0.25%	9,362

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 OCS factors ('target' samples), their Benjamini-Hochberg adjusted *P*-values (*P*-value), the variance explained (R^2) in the target sample phenotypes, and the number of SNPs (nSNPs). Significant findings are indicated in bold.

disorders could help address underlying biological mechanisms of comorbidity.

OCD and OCS have been linked to altered CNS and peripheral insulin signaling. When testing for potential overlapping biology, we found significant association between a set of 51 autosomal OCD genes centered around CNS insulin-regulated synaptic function and 'symmetry/counting/ordering'. As for peripheral insulin signaling, we found genetic sharing between T2D and—

based on the PNC data—'aggressive taboo thoughts', and —in the Spit for Science cohort—between T2D and 'contamination/cleaning'. For two out of the four T2D blood markers (blood levels of fasting insulin and 2hGlu), we also identified a shared genetic etiology with OCD. These findings provide support for 'dysregulated' peripheral insulin signaling as a biological process contributing to both OCD and population-based OCS. Further evidence for a role of (altered) peripheral insulin

Table 4 Summary of results from PRS-based analyses.

'Target' sample	'Base' sample					
	OCD	Type 2 Diabetes	HbA1c	Fasting insulin	Fasting glucose	2 h Glucose
PNC						
Total OCS score	4.72E-01	3.36E-01	7.64E-02	3.74E-01	4.16E-01	4.72E-01
	0.04%	0.28%	1.00%	0.19%	0.10%	0.02%
Impairment	2.61E-01	4.10E-01	2.28E-01	4.10E-01	4.10E-01	3.53E-01
	0.51%	0.11%	0.58%	0.11%	0.12%	0.22%
Symmetry/counting/ordering	3.53E-01	2.56E-01	4.10E-01	4.72E-01	2.80E-01	4.10E-01
	0.23%	0.53%	0.11%	0.04%	0.39%	0.11%
Contamination/cleaning	1.12E-01	3.53E-01	3.21E-01	4.47E-01	2.28E-01	2.71E-02
	0.85%	0.24%	0.33%	0.07%	0.58%	0.49%
Aggressive taboo thoughts	2.28E-01	**5.95E-03**	7.99E-02	4.66E-01	2.71E-01	2.28E-01
	0.46%	**1.86%**	0.97%	0.06%	0.49%	0.59%
Guilty taboo thoughts	**2.52E-03**	1.89E-0.1	4.03E-01	2.61E-01	3.21E-01	9.74E-02
	2.28%	0.68%	0.15%	0.52%	0.33%	0.90%
Distress	4.05E-01	4.72E-01	4.07E-01	2.80E-01	3.53E-01	3.53E-01
	0.13%	0.02%	0.13%	0.43%	0.23%	0.25%
PGC						
OCD		2.80E-01	4.15E-01	**7.67E-05**	3.53E-01	**4.75E-03**
		0.03%	0.01%	**0.26%**	0.02%	**0.14%**
Spit for Science						
Symmetry/counting/ordering $_{TOCS}$	**2.42E-05**	3.39E-01	2.71E-01	3.71E-01	2.47E-01	4.45E-01
	0.49%	0.03%	0.04%	0.02%	0.04%	0.01%
Contamination/cleaning $_{TOCS}$	**4.07E-03**	**1.59E-03**	1.68E-01	4.37E-01	4.36E-01	3.76E-01
	0.23%	**0.28%**	0.05%	0.01%	0.01%	0.02%

Shown in this table are the Benjamini-Hochberg adjusted *P*-values at the best SNP *P*-value thresholds along with the variance explained for each of the 'base' and 'target' sample pairs from PRS analyses in PRSice. Significant findings are indicated in bold. *PNC* Philadelphia Neurodevelopmental Cohort, *PGC* Psychiatric Genomics Consortium, *TOCS* Toronto Obsessive-Compulsive Scale.

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)[51]. Interestingly, SSRIs are particularly effective for treating harm-related obsessions, which are a part of 'aggressive taboo thoughts'[52]. This is in line with our finding of genetic sharing between T2D and 'aggressive taboo thoughts'. In addition, a recent study demonstrated that bilateral deep brain stimulation (DBS), a safe and effective treatment option for pharmaco-resistant OCD, not only reduced OCD symptoms but also decreased fasting insulin levels in the blood of both OCD patients with T2D and non-

diabetic OCD patients[53]. Moreover, insulin in the CNS - either entering from the periphery by crossing the blood brain barrier[54] or synthesized in the CNS[55]—has important non-metabolic functions, including modulating synaptic plasticity[56] and learning and memory[57,58].

Although it is not clear yet what the relative contributions are of dysregulated peripheral and CNS insulin signaling to OCD and OCS, we recently demonstrated that compulsivity observed in Tallyho (TH) mice, a rodent model of T2D, is potentially linked to disturbances in insulin signaling. TH mice both displayed compulsive behaviour and increased glucose levels in their dorsomedial striatum, which could be due to decreased action of peripheral and/or CNS insulin, and the glucose levels correlated with compulsivity[59].

The current results should be viewed in light of some strengths and limitations. A strength is that we used quantitative symptom scores collected through questionnaires in the general population, which has enabled us to generate OCS phenotypes that we could then perform GWASs on. Using samples selected from the community may also 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)[60]. A limitation of the current study is the small sample size of the GWASs and limited power to discover new single genetic variant associations. However, this sample size was large enough to provide proof of concept for genetic sharing between OCD, OCS in the population, and insulin signaling. A second limitation we faced was that the questions in the discovery and validation cohorts were not exactly the same, which may partly explain the lack of validation. Another limitation may be that the proportions of the variance in the target phenotypes being explained by the base phenotypes are quite small. However, these 'variances explained' are in fact similar to or higher than those found in similar analyses, e.g., the PRS derived from a GWAS of OCD explained (only) 0.20% of the variance in OCS in a population sample[20]. Moreover, as the variance explained is dependent on the size of the 'base sample' for the generation of the PRS[61], the observed variances explained with the still relatively small meta-GWAS of OCD as base sample may be underestimated.

In conclusion, we identified a shared genetic etiology between OCD, OCS in the population, and both CNS and peripheral insulin signaling. Our results imply that altered insulin signaling is not only relevant for somatic disorders but is also involved in the etiology of psychiatric disorders and related symptoms in the population, especially OCD and OCS. Further studies are needed to disentangle the contributions of peripheral and CNS insulin production and signaling to these disorders and symptoms.

Author details
[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, MA, USA. [5]Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA. [6]Department of Psychiatry and Genetics Institute, University of Florida, Gainesville, FL, USA. [7]Department of Neurology, Massachusetts General Hospital, Boston, MA, USA. [8]Karakter Child and Adolescent Psychiatry, Nijmegen, The Netherlands. [9]Neurosciences & Mental Health Program, Hospital for Sick Children, Toronto, ON, Canada. [10]Department of Psychiatry, University of Toronto, Toronto, ON, Canada. [11]Genetics & Genome Biology, Hospital for Sick Children, Toronto, ON, Canada. [12]Mathison Centre for Mental Health Research and Education, University of Calgary, Calgary, AB, Canada. [13]Departments of Psychiatry & Medical Genetics; Hotchkiss Brain Institute, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada. [14]Department of Psychiatry, Radboud University Medical Center, Nijmegen, The Netherlands

Code availability
Code used to perform the analyses is available upon request.

Conflict of interest
J.M.S. has received honoraria from Abide Therapeutics regarding treatments for OCD and related disorders. In the past 3 years, J.K.B. has been a consultant to/member of advisory board of /and/or speaker for Shire, Roche, Medice, and Servier. He is not an employee of any of these companies, and not a stock shareholder of any of these companies. He has no other financial or material support, including expert testimony, patents, royalties. B.F. has received educational speaking fees from Medice and Shire. G.P. is director of Drug Target ID (DTID) Ltd. The remaining authors declare that they have no conflict of interest.

Publisher's note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
1. American Psychiatric Association. Diagnostic and Statistican Manual of Mental Disorders. *Fifth Edition*. (American Psychiatric Publishing, Arlington, VA, 2013).
2. Weissman, M. M. et al. The cross national epidemiology of obsessive compulsive disorder. The Cross National Collaborative Group. *J. Clin. Psychiatry* **55**, 5–10 (1994).
3. Ruscio, A. M., Stein, D. J., Chiu, W. T. & Kessler, R. C. The epidemiology of obsessive-compulsive disorder in the National Comorbidity Survey Replication. *Mol. Psychiatry* **15**, 53–63 (2010).
4. Taylor, S. Etiology of obsessions and compulsions: a meta-analysis and narrative review of twin studies. *Clin. Psychol. Rev.* **31**, 1361–1372 (2011).
5. van Grootheest, D. S., Cath, D., Hottenga, J. J., Beekman, A. T. & Boomsma, D. I. Genetic factors underlie stability of obsessive-compulsive symptoms. *Twin Res. Hum. Genet.* **12**, 411–419 (2009).

6. van Grootheest, D. S., Cath, D. C., Beekman, A. T. & Boomsma, D. I. Twin studies on obsessive-compulsive disorder: a review. *Twin Res. Hum. Genet.* **8**, 450–458 (2005).

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

8. Mattheisen, M. et al. Genome-wide association study in obsessive-compulsive disorder: results from the OCGAS. *Mol. Psychiatry* **20**, 337–344 (2015).

9. Stewart, S. E. et al. Genome-wide association study of obsessive-compulsive disorder. *Mol. Psychiatry* **18**, 788–798 (2013).

10. International Obsessive Compulsive Disorder Foundation Genetics C, Studies OCDCGAS. Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis. *Mol. Psychiatry* **23**, 1181–1188 (2018).

11. Calamari, J. E. et al. Obsessive-compulsive disorder subtypes: an attempted replication and extension of a symptom-based taxonomy. *Behav. Res. Ther.* **42**, 647–670 (2004).

12. Mataix-Cols, D., Rauch, S. L., Manzo, P. A., Jenike, M. A. & Baer, L. 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 (1999).

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

14. Leckman, J. F. et al. Symptoms of obsessive-compulsive disorder. *Am. J. Psychiatry* **154**, 911–917 (1997).

15. Park, L. S. et al. The Toronto obsessive-compulsive scale: psychometrics of a dimensional measure of obsessive-compulsive traits. *J. Am. Acad. Child Adolesc. Psychiatry* **55**, 310–318 e4 (2016).

16. Alvarenga, P. G. et al. Obsessive-compulsive symptom dimensions in a population-based, cross-sectional sample of school-aged children. *J. Psychiatr. Res.* **62**, 108–114 (2015).

17. Fullana, M. A. et al. Obsessions and compulsions in the community: prevalence, interference, help-seeking, developmental stability, and co-occurring psychiatric conditions. *Am. J. Psychiatry* **166**, 329–336 (2009).

18. Fullana, M. A. et al. Obsessive-compulsive symptom dimensions in the general population: results from an epidemiological study in six European countries. *J. Affect. Disord.* **124**, 291–299 (2010).

19. Barzilay, R. et al. Obsessive-compulsive symptomatology in community youth: typical development or a red flag for psychopathology? *J. Am. Acad. Child Adolesc. Psychiatry* **58**, 277–286 e4 (2019).

20. den Braber, A. et al. 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 (2016).

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

22. Chacon, P. et al. Obsessive-compulsive symptoms in sibling pairs concordant for obsessive-compulsive disorder. American journal of medical genetics Part B. *Neuropsychiatr. Genet.* **144B**, 551–555 (2007).

23. Brakoulias, V. et al. The familiality of specific symptoms of obsessive-compulsive disorder. *Psychiatry Res.* **239**, 315–319 (2016).

24. Burton, C. L. et al. Heritability of obsessive-compulsive trait dimensions in youth from the general population. *Transl. Psychiatry* **8**, 191 (2018).

25. Kohlrausch, F. B. et al. Association of GRIN2B gene polymorphism and obsessive compulsive disorder and symptom dimensions: a pilot study. *Psychiatry Res.* **243**, 152–155 (2016).

26. Taylor, M. J. et al. Association of genetic risk factors for psychiatric disorders and traits of these disorders in a Swedish population twin sample. *JAMA psychiatry* **76**, 280–289 (2019).

27. Smit, D. J. A. et al. Genetic meta-analysis of obsessive-compulsive disorder and self-report compulsive symptoms. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* (2019). https://doi.org/10.1002/ajmg.b.32777. [Epub ahead of print].

28. van de Vondervoort, I. et al. 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 (2016).

29. Winocour, P. H., Main, C. J., Medlicott, G. & Anderson, D. C. 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 (1990).

30. Isomura, K. et al. Metabolic and cardiovascular complications in obsessive-compulsive disorder: a total population, sibling comparison study with long-term follow-up. *Biol. Psychiatry* **84**, 324–331 (2018).

31. Kontoangelos, K. et al. 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 (2013).

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

33. Satterthwaite, T. D. et al. Neuroimaging of the Philadelphia neurodevelopmental cohort. *NeuroImage* **86**, 544–553 (2014).

34. Calkins, M. E. et al. The Philadelphia Neurodevelopmental Cohort: constructing a deep phenotyping collaborative. *J. Child Psychol. Psychiatry* **56**, 1356–1369 (2015).

35. Gur, R. C. et al. Neurocognitive growth charting in psychosis spectrum youths. *JAMA Psychiatry* **71**, 366–374 (2014).

36. Calkins, M. E. et al. The psychosis spectrum in a young U.S. community sample: findings from the Philadelphia Neurodevelopmental Cohort. *World Psychiatry* **13**, 296–305 (2014).

37. Satterthwaite, T. D. et al. The Philadelphia Neurodevelopmental Cohort: A publicly available resource for the study of normal and abnormal brain development in youth. *NeuroImage* **124**, 1115–1119 (2016).

38. Kaufman, J. et al. Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): initial reliability and validity data. *J. Am. Acad. Child Adolesc. Psychiatry* **36**, 980–988 (1997).

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

40. Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. & Abecasis, G. R. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat. Genet.* **44**, 955–959 (2012).

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

42. Euesden, J., Lewis, C. M. & O'Reilly, P. F. PRSice: polygenic risk score software. *Bioinformatics* **31**, 1466–1468 (2015).

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

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

45. Harris, S. E. et al. Molecular genetic contributions to self-rated health. *Int. J. Epidemiol.* **46**, 994–1009 (2017).

46. Hagenaars, S. P. et al. Shared genetic aetiology between cognitive functions and physical and mental health in UK Biobank (N=112 151) and 24 GWAS consortia. *Mol. Psychiatry* **21**, 1624–1632 (2016).

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

48. Crosbie, J. et al. Response inhibition and ADHD traits: correlates and heritability in a community sample. *J. Abnorm. Child Psychol.* **41**, 497–507 (2013).

49. Burton, C. L. et al. SWAN scale for ADHD trait-based genetic research: a validity and polygenic risk study. *J. Child Psychol. Psychiatry.* **60**, 988–997 (2019).

50. Brainstorm, C. et al. Analysis of shared heritability in common disorders of the brain. *Science* **360**, eaap8757 (2018).

51. Janardhan Reddy, Y. C., Sundar, A. S., Narayanaswamy, J. C. & Math, S. B. Clinical practice guidelines for obsessive-compulsive disorder. *Indian J. Psychiatry* **59**, S74–S90 (2017).

52. Landeros-Weisenberger, A. et al. Dimensional predictors of response to SRI pharmacotherapy in obsessive-compulsive disorder. *J. Affect. Disord.* **121**, 175–179 (2010).

53. Ter Horst, K. W. et al. Striatal dopamine regulates systemic glucose metabolism in humans and mice. *Sci. Transl. Med* **10**, eaar3752 (2018).

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

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

56. Chiu, S. L., Chen, C. M. & Cline, H. T. Insulin receptor signaling regulates synapse number, dendritic plasticity, and circuit function in vivo. *Neuron* **58**, 708–719 (2008).

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

58. Zhao, W. et al. 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 (1999).

59. van de Vondervoort, I. et al. Converging evidence points towards a role of insulin signaling in regulating compulsive behavior. *Transl. Psychiatry* **9**, 225 (2019).

60. Gibbs, N. Nonclinical populations in research on obsessive-compulsive disorder: a critical review. *Clin. Psychol. Rev.* **16**, 729–773 (1996).

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