***Comparing genetic risk scores and family history as predictors of schizophrenia using Nordic registers***

Supplemental Tables & Figures

***Table of Contents***

Figure S1. Flow chart of sample processing 2

Figure S2. Nagelkerke R2 explained by individual GRS 3

Figure S3. Boxplots for GRS mean differences in samples with positive and negative family history 4

Table S1. Intercepts from LD-score regression 5

Table S2. Number of SNPs used for GRS calculation. 6

Table S3. Genetic correlations of SCZ, BIP, MDD, ASD, BMI, & EDU in discovery sets 7

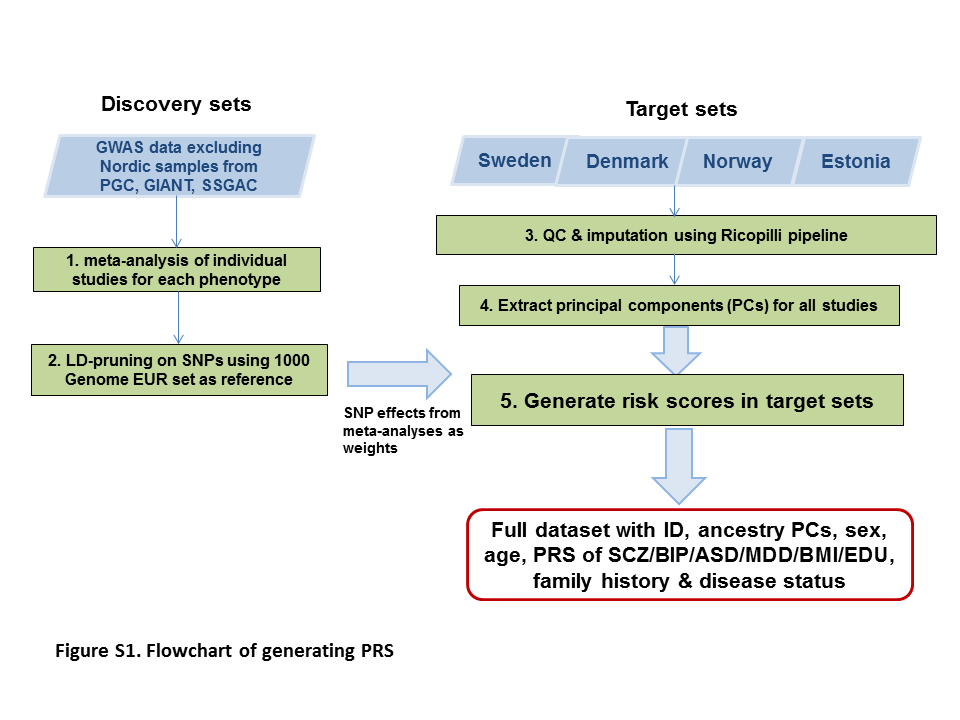
Table S4. Mean case-control differences in GRS and the variance explained. 8

Table S5. Correlations of GRS for SCZ, BIP, MDD, ASD, BMI, & EDU in target sets 9

Table S6. Separate & joint estimation of family history & GRS (excluding Norway) 10

References 11

## Figure S1. Flow chart of sample processing



## Figure S2. Nagelkerke R2 explained by individual GRS

The variance in schizophrenia (Nagelkerke R2) explained by GRS of four psychiatric disorders, schizophrenia (SCZ), bipolar disorder (BIP), major depressive disorder (MDD), and autism spectrum disorder (ASD), in the target samples from Denmark (DENM), Estonia (ESTO), Norway (NORW) and Sweden (SWE). Eight spikes per disorder per sample set corresponds to eight *P*-value cutoffs (*PT*): ≤5x10-8, ≤1x10-5, ≤1x10-3, ≤0.01, ≤0.05, ≤0.1, ≤0.5, and ≤1, given their associations with these disorders.

## Figure S3. Boxplots for GRS mean differences in samples with positive and negative family history



Among both schizophrenia cases and controls, a positive family history was associated with increased GRS of schizophrenia. Among cases, the mean GRS for those with a positive family history was 0.47, compared to 0.32 for those with a negative family history (P=7x10-5); and among controls, the mean GRS were 0.06 versus -0.27 (P=0.0001), respectively.

## Table S1. Intercepts from LD-score regression

|  |  |  |
| --- | --- | --- |
|  | **h2 intercept (s.e.)** | **gcov intercept (s.e.)** |
| **Target set:** |  |  |
| **Nordic schizophrenia studies** | 1.023 (0.009) |  |
| **Discovery sets (excluding Nordic studies):** | | |
| **SCZ (29 studies)** | 1.039 (0.012) | 0.020 (0.008) |
| **BIP (25 studies)** | 1.045 (0.009) | 0.009 (0.006) |
| **MDD (27 studies)** | 1.006 (0.009) | 0.009 (0.005) |
| **ASD** | 0.982 (0.008) | -0.003 (0.006) |
| **BMI** | 0.648 (0.012) | 0.007 (0.005) |
| **EDU** | 0.949 (0.012) | -0.005 (0.007) |

The intercepts from the univariate and bivariate models of LD-score regressions. The univariate models can be used to estimate SNP heritability; the intercept (h2 intercept) should be close to 1, unless the GWAS fails to control for population substructure (population stratification or cryptic relatedness; h2 intercept > 1), or the GWAS is over-conservative and corrected for the genomic control parameter (GC-correction; h2 intercept <1). The majority of h2 intercepts were close to 1, except the discovery set of BMI which was significantly lower than one (due to their use of lambda-GC correction). The bivariate models can be used to estimate genetic correlation; the intercept (gcov intercept) should be close to zero unless there are overlapping or related samples among the two sample sets. We examined the gcov intercepts between the target set and each of the discovery sets, and none were significantly different from zero, suggesting limited evidence for sample overlap.

## Table S2. Number of SNPs used for GRS calculation.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Study** | **N SNPs in target** |  | **SCZ**4 | **BIP**5 | **MDD**6 | **ASD**7 | **BMI**8 | **EDU**9 |
| **Swe1** | 83784951 | *Overlapped* | 4721784 | 4787309 | 4766627 | 4889639 | 1811308 | 4868763 |
|  |  | *clumped* | 151218 | 160096 | 158647 | 174399 | 103846 | 168963 |
| **S234** | 82822341 | *Overlapped* | 4721787 | 4787389 | 4766698 | 4889878 | 1811369 | 4869387 |
|  |  | *clumped* | 151218 | 160101 | 158651 | 174406 | 103865 | 168980 |
| **Swe5** | 82622871 | *Overlapped* | 4721797 | 4787384 | 4766695 | 4889732 | 1811348 | 4869076 |
|  |  | *clumped* | 151218 | 160099 | 158648 | 174401 | 103863 | 168973 |
| **Swe6** | 82979991 | *Overlapped* | 4721802 | 4787389 | 4766700 | 4889740 | 1811350 | 4869073 |
|  |  | *clumped* | 151218 | 160099 | 158649 | 174403 | 103860 | 168971 |
| **Denmark** | 45323472 | *Overlapped* | 4377590 | 4411095 | 4405224 | 4401885 | 1705345 | 4433921 |
|  |  | *clumped* | 105353 | 108276 | 107906 | 107895 | 85505 | 108759 |
| **Norway** | 29718162 | *Overlapped* | 2879870 | 2894357 | 2888494 | 2896592 | 1298131 | 2913028 |
|  |  | *clumped* | 79430 | 79923 | 79675 | 80614 | 73018 | 80954 |
| **Estonia** | 33476313 | *Overlapped* | 3194275 | 3215888 | 3208959 | 3236384 | 1469068 | 3231717 |
|  |  | *clumped* | 88153 | 90301 | 89800 | 92172 | 84135 | 91806 |

Notes:

1. No post-imputation QC was applied.
2. Post-imputation selected SNPs (MAF>0.05) with high imputation quality (INFO>0.8).
3. Post-imputation selected SNPs (MAF>0.05) with high imputation quality (probability of best-guess genotype > 0.8 and per-SNP missing rate < 0.01).
4. Based on the results from a meta-analysis of PGC2-SCZ studies[1](#_ENREF_1) excluding Nordic samples, 4762682 SNPs passed all filters (*Methods*).
5. Based on PGC2-BIP results excluding Nordic samples, 4839971 SNPs passed all filters.
6. Based on PGC2-MDD results excluding Nordic samples, 4818724 SNPs passed all filters.
7. Based on PGC2-ASD results, 5046391 SNPs pass all filters.
8. Based on published GWA summary results for BMI[2](#_ENREF_2), 1835322 SNPs pass all filters.
9. Based on published GWA summary results for educational attainment excluding Swedish samples[3](#_ENREF_3), 4926845 SNPs pass all filters.

## Table S3. Genetic correlations of SCZ, BIP, MDD, ASD, BMI, & EDU in discovery sets

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | SCZ | BIP | MDD | ASD | BMI | EDU |
| SCZ |  |  |  |  |  |  |
| BIP | **0.743 (0.026)** |  |  |  |  |  |
| MDD | **0.448 (0.053)** | **0.435 (0.069)** |  |  |  |  |
| ASD | **0.158 (0.064)** | 0.078 (0.073) | 0.203 (0.115) |  |  |  |
| BMI | **-0.082 (0.025)** | -0.045 (0.030) | 0.031 (0.047) | -0.005 (0.038) |  |  |
| EDU | **0.093 (0.026)** | **0.214 (0.030)** | **-0.183 (0.054)** | **0.280 (0.040)** | **-0.266 (0.02)** |  |

Genetic correlations between schizophrenia and related phenotypes estimated in the discovery samples (all studies excluding Nordic sets) using bivariate LD-score regression, and rg estimates (s.e.) are presented. Significant estimates (P<0.05) are bolded.

## Table S4. Mean case-control differences in GRS and the variance explained.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | mean case/con diff in GRS (s.e.) | | | | Nagelkerke R2 | | |
|  | SCZ | BIP | MDD | ASD | all four GRS | FH | GRS+FH |
| Denmark | 0.66 (0.05) | 0.36 (0.06) | 0.17 (0.06) | 0.07 (0.06) | 0.14 | 0.06 | 0.19 |
| Estonia | 0.62 (0.07) | 0.54 (0.07) | 0.19 (0.07) | 0.07 (0.07) | 0.08 | . | . |
| Norway | 0.58 (0.07) | 0.31 (0.07) | 0.21 (0.07) | 0.21 (0.07) | 0.13 | 0.13 | 0.22 |
| Sweden | 0.61 (0.02) | 0.36 (0.02) | 0.10 (0.02) | 0.01 (0.02) | 0.12 | 0.08 | 0.17 |

No sufficient information on family history (FH) with psychiatric disorders in the Estonian cohort, hence missing R2 for the model of family history alone and the joint model of family history and GRS together.

## Table S5. Correlations of GRS for SCZ, BIP, MDD, ASD, BMI, & EDU in target sets

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | SCZ | BIP | MDD | ASD | BMI | EDU |
| SCZ | . | **0.23** | **0.10** | 0.01 | **0.06** | **-0.04** |
| BIP | **0.24** | . | **0.10** | 0.02 | 0.01 | **0.03** |
| MDD | **0.13** | **0.12** | . | **0.03** | 0.00 | **-0.06** |
| ASD | **-0.03** | 0.01 | 0.00 | . | -0.01 | **0.03** |
| BMI | **0.09** | -0.02 | 0.02 | -0.02 | . | **-0.14** |
| EDU | **-0.05** | **0.04** | **-0.03** | 0.02 | **-0.22** | . |

Correlation matrices among GRS derived in the combined sets of Nordic samples.

GRS with Pt<0.05; Lower triangle: correlation among cases, upper triangle: correlations among controls; significant correlations (P<0.05) are in bold.

## Table S6. Separate & joint estimation of family history & GRS (excluding Norway)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Separate estimation | |  | Joint estimation | |
|  | OR (95% CI) | P-value |  | OR (95% CI) | P-value |
| **Family history** | 9.19 (7.56-11.17) | 1E-109 |  | 7.86 (6.42-9.61) | 4E-89 |
| **GRS (PT ≤ 0.05)** |  |  |  |  |  |
| SCZ | 1.82 (1.74-1.90) | 1E-160 |  | 1.81 (1.72-1.89) | 3E-142 |
| BIP | 1.25 (1.20-1.30) | 1E-28 |  | 1.25 (1.20-1.31) | 3E-26 |
| MDD | 1.02 (0.99-1.06) | 0.21 |  | 1.02 (0.98-1.06) | 0.32 |
| ASD | 1.02 (0.98-1.06) | 0.29 |  | 1.03 (0.99-1.07) | 0.17 |
| **GRS of BMI & EDU**  **(PT ≤ 0.05)** |  |  |  |  |  |
| BMI | 0.96 (0.92-1.00) | 0.06 |  | 0.96 (0.92-1.00) | 0.05 |
| EDU | 1.07 (1.03-1.11) | 5E-04 |  | 1.06 (1.02-1.10) | 0.007 |

## References

1. Schizophrenia Working Group of the Psychiatric Genomics, C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421-7 (2014).

2. Locke, A.E. *et al.* Genetic studies of body mass index yield new insights for obesity biology. *Nature* **518**, 197-206 (2015).

3. Okbay, A. *et al.* Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* **533**, 539-42 (2016).