**Supplementary Materials**

***Supplementary Materials and Methods***

*DNA isolation and genotyping*

DNA was isolated and genotyped for 887 participants admitted to the department of psychiatry at ZNA (Ziekenhuis Netwerk Antwerpen, Belgium) who provided informed consent for DNA extraction and analyses. One 10 ml ethylenediaminetetraacetic acid (EDTA) tube was filled during standard blood draws at the ward. DNA was extracted in the clinical molecular genetics laboratory of the University Medical Center Utrecht (UMCU). Samples were brought to a DNA concentration of 50 ng/μl with a total concentration of 200 ng DNA per participant. Subsequently, samples were sent in two batches to the Human Genotyping Facility of Erasmus Medical Center (Erasmus MC) Rotterdam for Global Screening Array v.1 (GSA) by Illumina, Santa Clara (California), USA, that has excellent validity and reliability *(85)*.

*Genetic quality control*

Quality control (QC) and genetic-ancestry principal component analysis (PCA) were done with PLINK 1.9 *(75)* and performed on the two batches separately (Table S1). Pre-imputation involved the creation of a superset with the highest quality SNPs for subsequent sample QC. The superset of SNPs was created by excluding those with genotype call rates <0.01, minor allele frequencies (MAF) <0.1, Hardy-Weinberg equilibrium (HWE) <10-4, and linkage disequilibrium (LD) *r*2 >0.2, with a window size of 50 and window shifting of a step size of 5. Using the superset, subjects were removed who: 1) had a mismatch in their sex between reported and genotyped; 2) were too extremely hetero- or homozygous (their *F*-values differed ≥3 SDs from the mean in the whole cohort); 3) were related (their pi-hat was above 0.1: one of each pair was randomly excluded); and 4) were cohort outliers (had values for the first two ancestry principal components (PCs) that deviated ≥3 SDs from the mean of the whole cohort).

This was followed by a regular SNP-level QC for exclusion of ill performing SNPs: variants with genotyping rate <0.01, MAF < 0.01, HWE p-value <10-5 were thus excluded. European ancestry was checked by comparing with the HapMap population: individuals were removed who deviated ≥3 SDs from first and second genetic ancestry PCs from the Northern and Western European (CEU) population.

Lastly, before imputation, genotypic data was compared with the Haplotype Reference Consortium panel. Strands, alleles, positions and frequency differences were checked. Chromosomes were pre-phased and imputed using the Michigan Imputation Server *(86)*. Post-imputation QC was performed to include reliable SNPs: variants that had a MAF >0.05 and LD *r*2 ≥0.8 were included, resulting in 5,211,700 SNPs available to generate a polygenic risk scores (PRS) in 762 individuals remaining after QC.

Table S1. QC steps of genotype data

|  |  |  |
| --- | --- | --- |
|  *Quality control steps*a | *Subjects* | *SNPs* |
| *Batch 1* | *Batch 2* | *Batch 1* | *Batch 2* |
| Data available for QC genotyping | **537** | **350** | **686,082** | **692,319** |
| *Pre-imputation steps (separate batches)* |
| Individuals >0.05 missing genotypes | -4 | -0 |  |
| Creating SNP superset | 533 | 350 | 686,082 | 692,319 |
| Genotype rate <0.01, MAF <0.1, HWE <1×10-4, LD pruning (50 5 0.2) |  | -545,698 | -530,759 |
| Perform subject-level QC with SNP superset | 533 | 350 | 140,384 | 161,560 |
| Sex check, heterozygosity (≥3 SD), relatedness, pi-hat >0.1, genetic outliers (≥3 SD) | -20 | -62 |  |
| Normal SNP QC | 513 | 288 | 686,082 | 692,319 |
| Genotype rate <0.01, MAF <0.01, HWE <1×10-5 |  | -227,781 | -197,418 |
| Compare with HapMap | 513 | 288 | 458,301 | 494,901 |
| Removal of genetic outliers (≥3 SD) from HapMap-CEU | -1 | -18 |  |
| Retained after pre-imputation QC | 512 | 270 | 458,301 | 494,901 |
| *Post-imputation steps (merged batches)* |
| Imputed total |  | 16,271,699 |
| QC (MAF <0.05, LD R2 ≥0.8) | -11,059,999 |
| FINAL post-imputation TOTAL | **762b** | **5,211,700** |

Abbreviations: QC=quality control; SD=standard deviation; SNP=single nucleotide polymorphism; MAF=minor allele frequency; HWE=Hardy-Weinberg equilibrium; LD=linkage disequilibrium; PC=principal component; HapMap=haplotype map; GSA = global screening array.

a ‘-’ is referring to excluded in this QC step.

b 762 individuals retained after post-imputation QC, of those 40 were excluded after EEG preprocessing.

*Polygenic risk score calculation*

The summary statistics of cross-disorder (cross-disorder group; CDG) *(87)*, MDD *(88)*, schizophrenia (SCZ) *(89)*, and alcohol dependence (AD) *(90)* were used to generate PRSs *(84)*. Polygenic scores for antidepressant response were not included, since they demonstrated to be weak predictors *(6)*. If only odds ratios (ORs) were reported in the summary statistics, ORs were log-converted to beta values as effect sizes. To that end, the beta values, effective allele, and p-values were extracted from all summary statistics.

SNPs that overlapped between the summary statistics GWASs (training datasets), 1,000 genomes (reference), and our dataset (target) were extracted. Then, insertions or deletions, and ambiguous SNPs, were excluded. To account for complicated LD structure of SNPs in the genome, these SNPs were clumped in two rounds using PLINK 1.90b3z *(91)*, according to previously established methods *(92, 93)*; round 1 with the default parameters (physical distance threshold 250 kb and LD threshold (*r*2) 0.5); round 2 with a physical distance threshold of 5,000 kb and LD threshold (*r*2) 0.2. Additionally, we excluded all SNPs in genomic regions with strong or complex LD structures. Sample overlap between training datasets with our target samples is unlikely since all samples belong to different cohorts and no Belgians had been included in the aforementioned GWASs used to generate PRSs.

We constructed PRSs based on risk alleles weighted by their effect sizes estimate using PLINK’s score function for 13 GWAS p-value thresholds (PT) *(86, 94)*: 5×10-8, 5×10-7, 5×10-6, 5×10-5, 5×10-4, 5×10-3, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5 and 1.

*EEG recordings*

During EEG recordings, subjects were seated in a sound and light attenuated room that was controlled at an ambient temperature of around 22°C. The participants were instructed to sit still for the duration of the recording without thought instructions. The operator did not intervene when drowsiness patterns were observed in the EEG.

Resting-state eyes closed EEG recordings for dataset 1 were acquired from 65 channels of the Electrical Geodesics Incorporated (EGI; Magstim, UK) system (dataset 1) and from 26 channels (10-20 electrode international system: Fp1, Fp2, F7, F3, Fz, F4, F8, FC3, FCz, FC4, T3, C3, Cz, C4, T4, CP3, CPz, CP4, T5, P3, Pz, P4, T6, O1, Oz, and O2) of the Neuroscan NuAmps (Compumedics, Australia; dataset 2 and 3). Data were recorded for three (dataset 1) or two (dataset 2 and 3) minutes during eyes closed condition. The sampling frequency was 500 Hz for most recordings, but 1,000 Hz for 6 recordings in dataset 1 (which were down-sampled to 500 Hz prior to further analyses). Data were referenced to Cz (dataset 1) or average mastoids with a ground at AFz (dataset 2 and 3). Horizontal eye movements were recorded with electrodes 61 and 64 (dataset 1) or electrodes placed 1.5 cm lateral to the outer canthus of each eye (dataset 2 and 3). Vertical eye movements were recorded with electrodes 5 and 62 (dataset 1) or electrodes placed 3 mm above the middle of the left eyebrow and 1.5 cm below the middle of the left bottom eyelid (dataset 2 and 3). Cartesian coordinates of the EGI system electrodes (dataset 1) were converted to spherical coordinates prior to EEG preprocessing.

*LORETA-fICA model*

The typical ICA model assumes that the source signals are not observable, statistically independent and non-Gaussian, with an unknown, but linear, mixing process *(95)*, and is described by the following formula:

*x=As*

where *x*, *A* and *s* represent matrices. In our case, these three matrices consisted of the following data:

1. Matrix *x* with 1,123 rows corresponding to all subjects of dataset 1, and the data per subject consists of 37,434 (6,239x6) columns corresponding to the spectral power at 6,239 cortical voxels for the six frequency bands. This approach, using a priori determined frequency bands, is a unique feature of the method used *(96)*.
2. Matrix *s* with 58 rows corresponding to the number of statistically independent components (i.e. functional networks), and 37,434 columns. In this way, each functional network contains 6 spatial images corresponding to neural activity of each frequency band (i.e. in a cross-frequency manner).
3. Matrix *A* with 1,123 rows and 58 columns. Thus, what remains of this data reduction for every subject is the amount of each component that was used for that subject. This amount is expressed as a loading (i.e. signed weight or score) per functional network for each subject.

**Table S2.** Significant different component loadings between homo-/heterozygotes for five variants

|  |  |  |
| --- | --- | --- |
|  | *Men* | *Women* |
| ANCOVA p-value | 0 | 1 | 2 | Post-hocp-value | *d* | ANCOVA p-value | 0 | 1 | 2 | Post-hocp-value | *d* |
| 1:67083671:T:C\_T(rs6656912) | **0.006** | m=328 | m=634 | m=829 | 0 vs 1 | 0.008 | -0.276 | 0.12 | m=591 | m=475 | m=209 | 0 vs 1 | 0.50 | 0.100 |
| N=138 | N=203 | N=45 | 0 vs 2 | 0.007 | -0.435 | N=114 | N=164 | N=58 | 0 vs 2 | 0.041 | 0.310 |
| 6:66534041:G:T\_T(rs68170059) | 0.71 | m=510 | m=482 | m=592 | 0 vs 1 | 0.43 | 0.096 | **0.004** | m=297 | m=715 | m=308 | 0 vs 1 | 0.001 | -0.354 |
| N=206 | N=154 | N=26 | 0 vs 2 | 0.93 | -0.001 | N=177 | N=137 | N=22 | 0 vs 2 | 0.89 | -0.010 |
| 9:11155055:G:T\_T(rs72694269) | 0.88 | m=555 | m=500 | m=834 | 0 vs 1 | 0.75 | 0.050 | **0.017** | m=408 | m=674 | m=3215 | 0 vs 1 | 0.08 | -0.226 |
| N=307 | N=75 | N=4 | 0 vs 2 | 0.72 | -0.220 | N=269 | N=66 | N=1 | 0 vs 2 | 0.022 | - |
| 9:11423966:T:C\_C(rs56072039) | **0.045** | m=443 | m=742 | m=412 | 0 vs 1 | 0.017 | -0.270 | 0.21 | m=410 | m=535 | m=891 | 0 vs 1 | 0.27 | -0.109 |
| N=229 | N=137 | N=20 | 0 vs 2 | 0.78 | 0.027 | N=214 | N=110 | N=12 | 0 vs 2 | 0.13 | -0.388 |
| 11:113365084:C:T\_C(rs4936275) | **0.021** | m=711 | m=481 | m=312 | 0 vs 1 | 0.040 | 0.210 | 0.80 | m=474 | m=437 | m=555 | 0 vs 1 | 0.65 | 0.031 |
| N=152 | N=179 | N=55 | 0 vs 2 | 0.012 | 0.342 | N=154 | N=141 | N=41 | 0 vs 2 | 0.75 | -0.074 |

Abbreviations: 0=no minor alleles, 2 reference alleles; 1=1 minor allele, 1 reference allele (heterozygote variant); 2=2 minor alleles (homozygote variant); *d*=Cohen’s d; N=number of participants per group; m=component loadings group mean.

*Additional information:*

Component loadings were compared across three groups, i.e. homozygote for reference allele (group 0), heterozygote for the variant (group 1), and homozygote for the variant (group 2), for the 47 SNPs that were included in the PRS-MDD PT=5×10-8. Statistical differences were calculated using ANCOVA (covariates: age and five PCs) and post-hoc tests, man and women separately. Significant differences at p<0.05 were found for five SNPs.

For one variant on *SGIP1* (rs6656912) loadings per group were in opposite direction in men compared to women, and post-hoc analysis revealed that homozygotes groups (group 0 versus 2) were significantly different in both men (Cohen’s d, *d*=-0.435; p=0.007) and women (*d*=0.310; p=0.041).

**Fig. S1.** Distribution of individual fICA component 13 loadings per dataset

|  |  |
| --- | --- |
| *Chart, histogram  Description automatically generated* | *a.*N=1,123; mean=503; SD=1,138. |
| *Chart, histogram  Description automatically generated* | *b.*N=193; mean=130; SD=1,003. |
| *Chart, histogram  Description automatically generated* | *c.*N=535; mean=-282; SD=877. |

The distribution curves and statistics (mean component loading and standard deviation, SD) of the three independent datasets: dataset 1, the ‘discovery cohort’ (a), dataset 2, the rTMS cohort (b), and dataset 3, the antidepressant medication cohort (c).

**Table S3.** Correlation between the LORETA-ICA component and baseline characteristics

|  |  |
| --- | --- |
| **Baseline characteristic** | **Effect size of Pearson correlation with independent component 13** |
| *Women TMS* | *Men TMS* | *Women/men sertraline* |
| *Age* | *r*=0.035 (p>0.05) | *r*=-0.069 (p>0.05) | *r*=-0.038 (p>0.05) |
| *Years of education* | *r*=-0.043 (p>0.05) | *r*=-0.163 (p>0.05) | *r*=0.085 (p>0.05) |
| *BDI-II/QIDS* | *r*=0.056 (p>0.05) | *r*=-0.086 (p>0.05) | *r*=0.060 (p>0.05) |
| *DASS-anxiety* | *r*=0.210 (p=0.047) | *r*=0.121 (p>0.05) | *r*=0.046 (p=0.047) |
| *DASS-stress* | *r*=0.054 (p>0.05) | *r*=0.069 (p>0.05) | N/A |
| *DASS-depression* | *r*=0.088 (p>0.05) | *r*=-0.058 (p>0.05) | N/A |
| *FAA (eyes closed)* | N/A | N/A | *r*=0.022 (p>0.05) |
| *FAA (eyes open)* | N/A | N/A | *r*=0.088 (p>0.05) |

Abbreviations: BDI-II=Beck Depression Inventory, second edition; QIDS=Quick Inventory of Depressive Symptomatology; DASS=Depression Anxiety Stress Scales; FAA=Frontal Alpha Asymmetry; TMS=transcranial magnetic stimulation; N/A=data for analysis not available.

**Table S4.** Correlation analysis adding frontal alpha asymmetry as covariate

|  |  |
| --- | --- |
| **QIDS change** | **Effect size of Pearson correlation with independent component 13 within the sertraline group** |
| *Without covariates* | *FAA eyes closed as covariate* | *FAA eyes open as covariate* |
| *Absolute change* | *r*=0.180 (p=0.015) | *r*=0.180 (p=0.015) | *r*=0.183 (p=0.017) |
| *Relative change* | *r*=0.182 (p=0.014) | *r*=0.183 (p=0.014) | *r*=0.186 (p=0.015) |

Abbreviations: QIDS=Quick Inventory of Depressive Symptomatology; FAA=Frontal Alpha Asymmetry.

*Additional information:*

Pearson correlation analyses within the rTMS and sertraline treatment group showed no significant correlations at p<0.01 with the obtained LORETA-ICA component and several baseline characteristics (see Table S3), including frontal alpha asymmetry (FAA) scores, obtained in a previous iSPOT-D study by Arns *et al*. *(33)*

Correlation analysis – within the sertraline iSPOT-D group – between the component and QIDS change yielded no difference in effect size (of absolute nor relative change) when FAA with 1) eyes open or 2) eyes closed were added as covariate (see Table S4). This indicates the effect on treatment outcome is not (solely) attributable to frontal asymmetry part of the component.

**Fig. S2.** ROC curves of the improved treatment prediction models for response



ROC (receiver operating characteristic) curves for the prediction of response by the prefrontal-sensorimotor (PF-SM) network loading and age as predictors. Red and blue lines represent outcome after rTMS treatment in women and men, respectively (dataset 2). Green lines represent outcome after sertraline treatment in women and men (dataset 3). The area under the curve (AUC) for each model is displayed in the figure.

**Table S5.** Stratification based on network loading cut-offs for remission

|  |  |  |  |
| --- | --- | --- | --- |
| ***Stratification*** | **rTMS 1 or 10 Hz** | **sertraline** | **esc/ven** |
| ♀¹ | ♂² | ♀+ ♂² | ♀² |
|   | obs. R | obs. NR | obs. R | obs. NR | obs. R | obs. NR | obs. R | obs. NR |
| *R-strat.* | **N=26** | **N=8** | **N=30** | **N=14** | **N=36** | **N=44** | *N=38* | *N=62* |
| *NR-strat.* | N=25 | N=34 | N=22 | N=27 | N=29 | N=79 | N=25 | N=50 |
| *non-strat.* | N=51 | N=42 | N=52 | N=41 | N=65 | N=123 | N=63 | N=112 |
|   | χ2 | V | χ2 | V | χ2 | V | χ2 | V |
| *statistic* | 10.126 | 0.330 | 5.099 | 0.234 | 6.691 | 0.189 | 0.405 | 0.048 |
| *p-value\** | 0.002 | 0.036 | 0.013 | 0.633 |
|   | non-strat. | strat. (PPV) | non-strat.  | strat. (PPV) | non-strat.  | strat. (PPV) | non-strat.  | strat. (PPV) |
| *remission* | 54.8% | **76.5%** | 55.9% | **68.2%** | 34.6% | **45.0%** | 36.0% | *38.0%* |
| *improvement* | *39.4%* | *21.9%* | *30.2%* | *5.6%* |

Abbreviations: rTMS=repetitive transcranial magnetic stimulation; esc/ven=escitalopram or venlafaxine R=remission; NR=non-remission; obs.=observed (true remission and non-remission numbers in the dataset); strat.=stratified (i.e. predicted, based on 2 cut-offs); χ2=Chi-Squared test (difference between 2 proportions in the crosstab); V=Cramérs V (effect size); PPV=positive predictive value (note that the remission rate within ‘R-strat.’ subgroup is equal to PPV of the EEG component).

¹cut-off point=254; ²cut-off point=-197; \*exact 2-sided p-value