Data in Brief 42 (2022) 108014

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Functional magnetic resonance imaging data for the association between polygenic risk scores for neuroticism and reward-punishment processing

Heekyeong Park[a,b,*], Katherine L. Forthman[a], Rayus Kuplicki[a], Teresa A. Victor[a], Tulsa 1000 Investigators[a,1], Hung-Wen Yeh[a,c], Wesley K. Thompson[d], Martin P. Paulus[a]

[a] *Laureate Institute for Brain Research*[2], *Tulsa, OK, USA*
[b] *University of North Texas at Dallas, Dallas, TX, USA*
[c] *Children's Mercy Hospital, Kansas City, MO, USA*
[d] *University of California, San Diego, CA, USA*

ARTICLE INFO

Article history:
Received 18 August 2021
Revised 25 February 2022
Accepted 28 February 2022
Available online 4 March 2022

Dataset link: N-PRS and reward processing_469 (Original data)

Keywords:
Neuroticism
Polygenic risk score
Reward
fMRI
Genetics

ABSTRACT

Neuroticism as a personality trait represents a heritable risk for psychiatric disorders. The polygenic risk score for neuroticism (N-PRS) is used to study genetic vulnerability to neuroticism. The current data present the association of the genetic risk for neuroticism to neural reward-punishment processing using functional magnetic resonance imaging. N-PRS was computed based on the individual's genotype information and a genome-wide association study on the UK Biobank data. While individuals performed a monetary incentive delay task, their neural activations for upcoming incentives (reward: gain, punishment: loss) were measured in blood oxygen level dependent (BOLD) signals during the delay phase. Multivariate ANCOVAs were used to analyze BOLD signals for finding the association between N-PRS and

DOI of original article: 10.1016/j.jad.2019.01.005
* Corresponding author at: Laureate Institute for Brain Research, Tulsa, OK, USA.
E-mail address: hpark@laureateinstitute.org (H. Park).
[1] The Tulsa 1000 Investigators include the following contributors: Robin Aupperle, Ph.D., Jerzy Bodurka, Ph.D., Salvador Guinjoan, M.D., Ph.D., Sahib S. Khalsa, M.D., Ph.D., Rayus Kuplicki, Ph.D., Martin P. Paulus, M.D., Jonathan Savitz, Ph.D., Jennifer Stewart, Ph.D., Teresa A. Victor, Ph.D.
[2] @LIBR_Tulsa

https://doi.org/10.1016/j.dib.2022.108014

reward-punishment processing by the incentive valence (Related research article: H. Park, K.L. Forthman, R. Kuplicki, T.A. Victor, Tulsa 1000 Investigators, H.W. Yeh, W.K. Thompson, M.P. Paulus, Polygenic risk for neuroticism modulates response to gains and losses in the amygdala and caudate: evidence from a clinical cohort. J. Affect. Disord. 293 (2021) 124–132. https://doi.org/10.1016/j.jad.2021.06.016). These data can be used as reference data for future studies examining the role of the genetic propensity for personality traits in the context of psychiatric disorders.

Specification Table

Subject	Biological Sciences Health and Medical Sciences
Specific subject area	Neuroscience: Biological Psychiatry Psychiatry and Mental Health
Type of data	Table
How data were acquired	The Magnetic Resonance Imaging (MRI) data were collected using the GE MR750 3T scanner with 8 RF coils for both anatomical and functional scans. Self-report measurements were electronically administered on Apple iPad using a secure web-based application for the electronic collection of research and clinical trial data (www.project-redcap.org).
Data format	Raw Derived
Parameters for data collection	The dataset is a part of the Tulsa 1000 project data. Anatomical images (T1-weighted 3D high resolution) were acquired in the MP-RAGE pulse sequence with scanning parameters of FOV 240 × 192 mm, TR/TE = 5/2.012 ms, and 186 axial slices. Functional images (T2*-weighted echo-planar images) were collected in 562 axial volumes (39 slices, 2.9 mm thick, 1.875^3 voxels) with the parameters of flip angle 78°, FOV 240 × 240 mm, and TR/TE = 2000/27 ms.
Description of data collection	N-PRS was computed based on the individual's genotype information and genome-wide association study (GWAS) summary statistics. MRI data were preprocessed and analysed with AFNI (Analysis of Functional NeuroImages software suite). PHQ-9 scores were collected through self-report.
Data source location	Institution: Laureate Institute for Brain Research City/Town/Region: Tulsa, Oklahoma Country: U.S.A.
Data accessibility	Repository name: Mendeley Data Data identification number: N/A Direct URL to data: PARK, HEEKYEONG (2021), "N-PRS and Reward Processing_469", Mendeley Data, V1, doi: 10.17632/n4534vcjkh.1 http://dx.doi.org/10.17632/n4534vcjkh.1
Related research article	**Co-submission:** H. Park, K.L. Forthman, R. Kuplicki, T.A. Victor, Tulsa 1000 Investigators, H.W. Yeh, W.K. Thompson, M.P. Paulus, Polygenic risk for neuroticism modulates response to gains and losses in amygdala and caudate: evidence from a clinical cohort, J. Affect. Disord. 293 (2021) 124–132. https://doi.org/10.1016/j.jad.2021.06.016

Value of the Data

- The data can be used to replicate the results of [1].
- The data contribute to understanding the genetic impact of neuroticism in increasing vulnerability for psychiatric disorders in public health.
- The data can be used for future studies examining the relationship between genetic risks of personality traits and functional neural markers for psychiatric disorders.

1. Data Description

This paper includes the data regarding the association between polygenic risk scores for neuroticism (N-PRS) and reward-punishment processing on the Monetary Incentive Delay task (MID) from a functional MRI experiment [2]. The present dataset is based on 469 individuals' data on N-PRS and BOLD signals by the incentive during the MID task. Table 1 displays BOLD signals showing the relationship between N-PRS and neural activity for reward-punishment processing after covarying out the participant's psychiatric diagnosis (depression, anxiety, substance use, eating disorders, as well as healthy control) in the statistical analysis. Table 2 presents BOLD signal changes associated with N-PRS but not with depression severity on the Patient Health Questionnaire (PHQ-9) in individuals with major depression disorders [3]. For both tables, the first column denotes an arbitrary number assigned to each participant. In the tables, the abbreviation 'A' represents the amygdala, while 'I' and 'P' represent the insula and the precuneus, respectively. In Table 2, 'C' represents the tail of the caudate.

2. Experimental Design, Materials and Methods

2.1. Participants

Participants were drawn from the first 500 individuals (T500) in the Tulsa 1000 project, a naturalistic study following 1000 individuals with mood, anxiety, substance use, and/or eating disorders as well as healthy volunteers [4]. The eligibility criteria for the clinical population were (1) Patient Health Questionnaire (PHQ-9) $\geq$ 10 [3]; (2) Overall Anxiety Severity and Impairment Scale (OASIS) $\geq$ 8 [5]; (3) Drug Abuse Screening Test (DAST-10) score $\geq$ 3 [6]; and/or (4) Eating Disorder Screen (SCOFF) score $\geq$ 2 [7]. Healthy volunteers screened negative for the above scales. Exclusion criteria included: positive results on a drug screening test; lifetime bipolar, schizophrenia spectrum, antisocial personality, or obsessive-compulsive disorders; active suicidal ideation; moderate to severe traumatic brain injury; severe or unstable medical conditions; change in psychiatric medication dose within the last 6 weeks; and MRI contraindications. All procedures were conducted following the study protocol approved by the Western Institutional Review Board. Prior to the study, all participants gave informed consent, and they were remunerated for their participation. Participants who had excessive head motions or incomplete data were excluded, resulting in 469 participants.

2.2. Polygenic risk score for neuroticism (N-PRS)

Participants' blood samples were genotyped using Illumina Infinium Global Screening Array-24 (v.2.0) BeadChip arrays by RUDCR Infinite Biologics. The genotyped data then underwent three rounds of quality control. In the first round of quality control, we checked for mismatches in the strand, ID names, position, alleles, and ref/alt assignments. We used the McCarthy Group Tool (HRC or 1000G Imputation preparation and checking, v.4.2.11) to compare

genotypes to the Haplotype Reference Consortium (HRC) release 1.1 reference panel. The McCarthy Group Tool outputs the following corrections: SNPs to exclude, name corrections, strand flips, and allele flips. The recommended corrections were applied using the software PLINK (v.1.9, https://www.cog-genomics.org/plink2; v.2.0, https://www.cog-genomics.org/plink/2.0). The second round of quality control applied several exclusion criteria for SNPs, including (1) call rates lower than 2%, (2) duplicated SNPs, and (3) violation of Hardy-Weinberg equilibrium. The third round of quality control applied exclusion criteria for participants in the sample, including (1) missingness greater than 2% and (2) close genetic relationship with any participant ($\hat{\pi} > 0.2$) within the sample. Both the second and third rounds of quality control were also conducted with PLINK. Genotype Imputation was done via the Michigan Imputation Server Pipeline using Minimac4, version 1.2.4 [8], with the options, (1) Reference Panel – HRC r1.1 2016 (GRCh37/hg19), (2) Array Build – GRCh37/hg19, (3) rsq Filter – off, (4) Phasing – Eagle v2.4 (phased output), (5) Population – Other/Mixed, (6) Mode – Quality Control & Imputation, (7) AES 256 encryption – unchecked. Genome information was imputed from 569,641 to 40,359,612 SNPs [8].

A genome-wide association study (GWAS) for neuroticism was taken from the Atlas of GWAS Summary Statistics (https://atlas.ctglab.nl). The selected GWAS was performed by the Center for Neurogenomics and Cognitive Research (CNCR) Complex Trait Genetics Laboratory (CTG Lab) on data from the UK Biobank, release 2. The GWAS summary file was downloaded from the CTG Lab website (https://atlas.ctglab.nl/traitDB/3990#:~:text=https%3A//ctg.cncr.nl/documents/p1651/sumstats_neuro_sum_ctg_format.txt.gz). The GWAS describes the strength and significance of the association of 10,846,943 SNPs with neuroticism measured on a 12-item sum scale from 380,060 participants [9]. N-PRS was computed using PRSice-2 [10]. We accounted for linkage disequilibrium (LD) by clumping in PRSice using standard parameters (250 kilobases, $r^2 = 0.1$, $p = 1$), resulting in 144,443 SNPs. Principal component analysis (PCA) was used to quantify population stratification (i.e., genetic ancestry) with 10 components. PCA was performed using FlashPCA2 [11]. The GWAS summary included p-values for the relationship between each SNP and neuroticism. To reduce noise, a *p*-value threshold was used to determine which SNPs to include in the calculation of the PRS. To choose the optimal p-value threshold, 14 *p*-value thresholds were selected (5×10^{-8}, 5×10^{-7}, 5×10^{-6}, 5×10^{-5}, 5×10^{-4}, 0.001, 0.005, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, & 1), and a PRS was created for the T500 cohort at each threshold. Each PRS was then compared to the T500 participant neuroticism score from the Big Five Inventory. For this comparison, two linear models were created: the first predicted neuroticism score using the PRS and the 10 principal components of population stratification, while the second predicted neuroticism score using only the principal components. The variance explained was quantified with the R^2 of the second model subtracted from the R^2 of the first model. The *p*-value thresholds from 0.05 to 1 explained the variance greater than 0.025. Because there is no a-priori method to determine the best threshold, and because the *p*-value thresholds between 0.05 and 1 shared approximately equivalent fit to the phenotype, we chose the most stringent threshold of 0.05 among the *p*-value thresholds that explained greater than 0.025 of the variance. For this study, N-PRS was standardized across participants with a mean of zero and a standard deviation of 1.

2.3. fMRI experiment

The MID task was used to examine neural processes associated with reward and punishment. The study design consisted of 2 levels of valence (gain, loss) and 3 levels of magnitude of the incentive (high, low, no), yielding 6 conditions, high-gain/high-loss (+$5/-$5), low-gain/low-loss (+$1/-$1), and no-gain/no-loss (+$0/-$0), with 15 trials per condition. A trial began with a task cue (an object) for 2 s indicating the valence (circle: gain, square: loss) and magnitude (top: $5, middle: $1, bottom: $0) of the incentive for the trial by the shape of the task cue and the location of a line in the cue, respectively. After a varied delay between 2.25 s and 3 s, a target of the trial (a triangle) was presented, prompting participants to press a button to gain reward or to avoid loss as early as they could. Participants received practice trials out of the scanner before the scan session. The mean response time of each participant collected from practice trials was

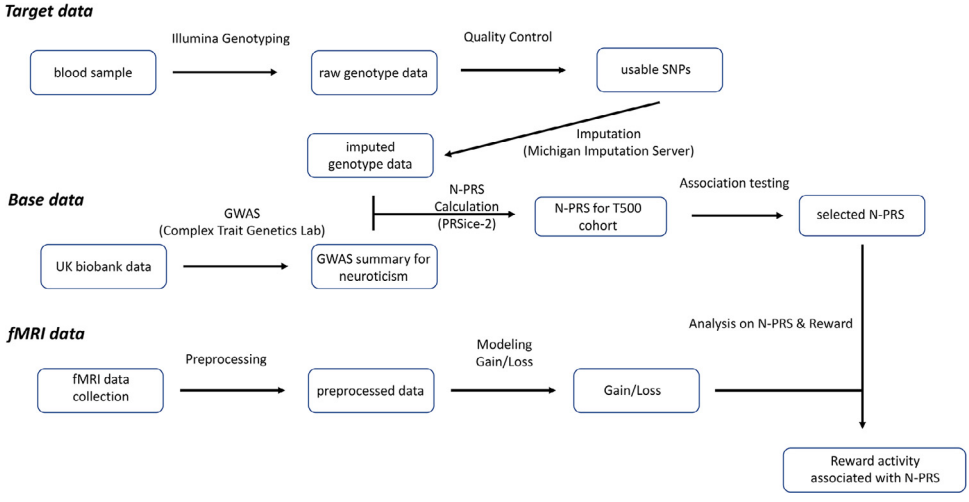

Fig. 1. The workflow diagram of the study.

used for adjusting the individualized target duration for approximately 66% of success on button press. After the target, the outcome with the amount of earned/lost money was presented for 2 s as feedback. The study lasted about 19 min in the scanner.

Both T1-weighted 3D high-resolution anatomical images (MP-RAGE pulse sequence, FOV 240 × 192 mm, TR/TE = 5/2.012 ms, 186 axial slices) and T2*-weighted echo-planar images (flip angle 78°, FOV 240 × 240 mm, TR/TE = 2000/27 ms, axial plane, 39 slices/volume, 2.9 mm thick, 1.875 × 1.875 mm voxels) were collected using two GE MR750 3T scanners equipped with 8 RF channel phased array coils at the Laureate Institute for Brain Research. fMRI data were acquired in two runs of 281 volumes while participants performed the MID task in the scanner.

2.4. fMRI data preprocessing and analysis

The Analysis of Functional NeuroImages software suite (AFNI, https://afni.nimh.nih.gov/) was used for preprocessing and analyzing imaging data. The first 3 volumes were discarded to avoid magnetic saturation effects. fMRI data were despiked, slice-time corrected to the first slice, co-registered to a T1-weighted anatomical image, and motion-corrected (ENORM > 0.3). Functional data were also normalized to the MNI space with resampling of isotropic 2 mm voxels and smoothed with an isotropic 4 mm FWHM Gaussian kernel. For analysis, 6 events were modeled for anticipatory reward-punishment processing (±$5, ±$1, ±$0) on a subject level. To capture motivational valence processing for an upcoming event, only anticipatory neural activity during the delay phase was included in the analysis. The BOLD response to an incentive cue was convolved with a 4-s boxcar function from the onset of the cue. All other events were considered as no-interest events including 6 motion parameters and the first 4 polynomial terms. The contrasts of incentive valence for gain (+$5 > +$0) and loss (−$5 > −$0) were constructed as the main events-of-interest. The $1 conditions were not used for the analysis. A multivariate ANCOVA model (3dMVM: valence (2) * N-PRS) was constructed to investigate the association between N-PRS and anticipatory reward-punishment processing in the brain with covariates of age, sex, and race, at the group level. The effect of PRS for neuroticism was examined for each valence separately and collapsed over the valence for estimating the main effect. The contrast showing the difference between gain versus loss by PRS was used for an interaction effect. A voxel-wise threshold of $p < .001$ was used for this study in conjunction with a cluster-extent threshold of $\alpha < .05$ ($k > 43$) based on the estimated ACF parameters of the group level error terms using

3dFWHMx and 3dClustSim. Significant clusters were further probed with beta coefficients in the clusters for follow-up comparisons. Self-reported PHQ-9 scale scores were used for estimating the severity of depressive symptoms Fig. 1. shows the diagram depicting the workflow of the study method.

Ethics Statement

All procedures were approved by the Western Institutional Review Board (NCT02450240). Participation in the study was completely consensual, anonymous, and voluntary. Participants provided written informed consent. Data collection was conducted according to the Declaration of Helsinki.

Declaration of Competing Interest

This work has been supported in part by The William K. Warren Foundation and the National Institute of General Medical Sciences Center Grant Award Number 1P20GM121312. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. MP is an advisor to Spring Care, Inc., a behavioral health startup, and he has received royalties for an article about methamphetamine in UpToDate.

Data Availability

N-PRS and reward processing_469 (Original data) (Mendeley Data)

Supplementary Materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.dib.2022.108014.

CRediT Author Statement

Heekyeong Park: Formal analysis, Writing – original draft; **Katherine L. Forthman:** Methodology, Writing – review & editing; **Rayus Kuplicki:** Methodology, Formal analysis, Writing – review & editing; **Teresa A. Victor:** Writing – review & editing; **Hung-Wen Yeh:** Writing – review & editing; **Wesley K. Thompson:** Writing – review & editing; **Martin P. Paulus:** Formal analysis, Writing – original draft.

References

[1] H. Park, K.L. Forthman, R. Kuplicki, T.A. Victor, H.W. Yeh, W.K. Thompson, M.P. Paulus, Tulsa 1000 Investigators, Polygenic risk for neuroticism modulates response to gains and losses in amygdala and caudate: evidence from a clinical cohort, J. Affect. Disord. 293 (2021) 124–132, doi:10.1016/j.jad.2021.06.016.
[2] B. Knutson, G.W. Fong, C.M. Adams, J.L. Varner, D. Hommer, Dissociation of reward anticipation and outcome with event-related fMRI, Neuroreport 12 (2001) 3683–3687, doi:10.1097/00001756-200112040-00016.
[3] K. Kroenke, R.L. Spitzer, J.B. Williams, The PHQ-9: validity of a brief depression severity measure, J. Gen. Intern. Med. 16 (2001) 606–613, doi:10.1046/j.1525-1497.2001.016009606.
[4] T.A. Victor, S.S. Khalsa, W.K. Simmons, J.S. Feinstein, J. Savitz, R.L. Aupperle, H.W. Yeh, J. Bodurka, M.P. Paulus, Tulsa 1000: a naturalistic study protocol for multilevel assessment and outcome prediction in a large psychiatric sample, BMJ Open 24 (2018) e016620, doi:10.1136/bmjopen-2017-016620.
[5] L. Campbell-Sills, S.B. Norman, M.G. Craske, G. Sullivan, A.J. Lang, D.A. Chavira, A. Bystritsky, A.J. Sherbourne, P. Roy-Byrne, M.B. Stein, Validation of a brief measure of anxiety-related severity and impairment: the Overall Anxiety Severity and Impairment Scale (OASIS), J. Affect. Disord. 112 (2009) 92–101, doi:10.1016/j.jad.2008.03.014.

[6] S.E. McCabe, C.J. Boyd, J.A. Cranford, J. Slayden, J.E. Lange, M.B. Reed, J.M. Ketchie, M.S. Scott, Alcohol involvement and participation in residential learning communities among first-year college students, J. Stud. Alcohol Drugs 68 (2007) 722–726, doi:10.15288/jsad.2007.68.722.
[7] J.F. Morgan, F. Reid, J.H. Lacey, The SCOFF questionnaire: a new screening tool for eating disorders, West. J. Med. 172 (2000) 164–165, doi:10.1136/ewjm.172.3.164.
[8] S. Das, L. Forer, S. Schonberr, C. Sidore, A.E. Locke, A. Kwong, C. Fuchsberger, et al., Next-generation genotype imputation service and methods, Nat. Genet. 48 (2016) 1284–1287, doi:10.1038/ng.3656.
[9] M. Nagel, K. Watanabe, S. Stringer, D. Posthuma, S. der van Sluis, Item-level analyses reveal genetic heterogeneity in neuroticism, Nat. Commun. 9 (2018) 905, doi:10.1038/s41467-018-03242-8.
[10] S.W. Choi, P.F. O'Reilly, PRSice-2: polygenic risk score software for Biobank-scale data, GigaScience 8 (2019) giz082, doi:10.1093/gigascience/giz082.
[11] G. Abraham, Y. Qiu, M. Inouye, FlashPCA2: principal component analysis of Biobank-scale genotype datasets, Bioinformatics 33 (2017) 2776–2778, doi:10.1093/bioinformatics/btx299.