Data Supplement for Axelrud et al., Polygenic Risk Score for Alzheimer's Disease: Implications for Memory Performance and Hippocampal Volumes in Early Life. Am J Psychiatry (doi: 10.1176/appi.ajp.2017.17050529)

SUPPLEMENTARY METHODS

CONTENTS

1. Psychometric steps for cognitive tests 2
1.1. ROCFT 3
Figure S1. Test Information Function for immediate and delayed recall 4
1.2. Reading and writing abilities 5
Figure S2. Test information curves for reading and writing 5
2. Exploratory and sensitivity analyses 8
2.1. Regressions for the Brazilian non-Caucasian subsample 8
Table S1. Brazilian total sample description 8
2.2. Regressions using other PRS thresholds 9
Figure S3. Associations between immediate recall and AD-PRS for multiple thresholds 10
Figure S4. Associations between delayed recall and AD-PRS for multiple thresholds 11
Figure S5. Associations between reading and AD-PRS for multiple thresholds 12
Figure S6. Associations between writing and AD-PRS for multiple thresholds 13
Figure S7. Associations between executive function and AD-PRS for multiple thresholds 14
Figure S8. Associations between left hippocampus and AD-PRS for multiple thresholds 15
Figure S9. Associations between right hippocampus and AD-PRS for multiple thresholds 16
2.3. Quantile regressions 17
Figure S10. Quantile regressions using the AD-PRS as the response variable 17
Table S2. Brazilian total sample description by quintiles of AD-PRS 19
2.4. Hippocampal subregions 19
Table S3. Associations between AD-PRS and hippocampal subregion volumes 20
3. Associations of cognition and hippocampal volumes with APOE 21
4. Assessment of executive functions 22
5. SNPs comprising AD-PRS and enrichment analysis for genes and pathways 23
Figure S11. Associations between AD-PRS genes and immediate and delayed recall 24
Table S4. Top 10 enriched genes 25
Table S5. Top 10 enriched pathways 25

.

1. Psychometric steps for cognitive tests

We investigated the unidimensional model of both ROCFT scores (i.e. immediate and delayed score) and education attainment scores (i.e. reading and writing). First, missing data was cleaned and processed using imputation by chained equations implemented in the "mice" package from R (1).

Second, we used Confirmatory Factor Analysis (CFA) and Item Response Theory (IRT) analyses to investigate the model fit of a unidimensional model of ROCFT scores (i.e. immediate and delayed scores) and education attainment scores (i.e. reading and writing scores). The CFA models were fitted to polychoric correlations using mean and variance adjusted weighted least squares (WLSMV) estimator implemented by the "lavaan" package from R (2). Model fit was judged to be good if CFI (Comparative Fit Index) and TLI (Tucker-Lewis Index) $\geq$ .95 and if RMSEA (Root Mean Square Error of Approximation) $\leq$. 06 and was judged to be acceptable if CFI and TLI $\geq$. 90 and RMSEA $\leq$ .08 (3,4). We estimated reliability indices using omega coefficients (5).

Third, we performed unidimensional Item Response Theory (IRT) analyses using two-parameter Graded Response Model (6) to calculate each item's discrimination and difficulty parameter. IRT analyses were fitted using the "ltm" package from R (7). We used the following interpretation to the discrimination parameters: 0 (none), .01-.34 (very low), .35-.64 (low), .65-1.34 (moderate), 1.35-1.69 (high) and higher than 1.7 (very high) (8). Moreover, we used test information curves to compare on which range each score is able to capture information.

Fourth, we investigated the linear, quadratic and cubic tendencies of these associations using mixed-effect models (9,10). We adjusted for test evaluator in the ROCFT scores analyses, for subjects' school in the education attainment analyses and for site in both analyses. Finally, we used the normalized residuals of the best fitting mixed effect model (i.e. linear for ROCFT scores and cubic for education attainment scores) to perform our analyses.

1.1. ROCFT

Scoring method

The Rey–Osterrieth Complex Figure Test (ROCFT) (11,12) is a well-recognized test used to assess visuospatial memory (13). Individuals are asked to draw the ROCFT in three steps: copy, in which they can look at the stimulus card, Immediate (IR) and Delayed Recall (DR), in which they are instructed to draw what they remember after 3 and 30 minutes, respectively.

In the Quantitative Scoring System (12,14), the ROCFT is divided in 18 items and each item receives a score from 0 to 2. Most studies use the sum of these scores (raw score), while others use the recall scores as a proportion of the copy score (percent-retained score) (Gallagher & Burke, 2007). These methods present several limitations. The raw score is highly influenced by the constructional, perceptual, and attentional components of the copy score and might not capture accurately the memory construct (15). Although the percent-retained score is independent of the copy score, it does not account for the fact that subjects might receive a better score in an item in the recall task than in the copy task.

In order to obtain a purer measure of the memory construct, we created a new score to assess the ROCFT. We assessed the immediate and delayed recall scores as a proportion of the copy score to each item separately and excluded the items in which the recall score outperformed the copy score. The items were given the following scores: 0, if the individual could not remember the item; 0.5, if the individual remembered partially the shape and/or location of the item; and 1, if the individual remembered the shape and the location of the item as they draw in the copy task. We also excluded subjects who draw less than 50% of the ROCFT items in the copy task (n = 6). Missing data was replaced by multiple imputation by chained equations using the "mice" package from R (1). Finally, we used the normalized residuals of the best fitting mixed effect models (i.e. linear) adjusted for the age of the participants, since previous research showed a significant effect of age on the ROCFT (15).

.

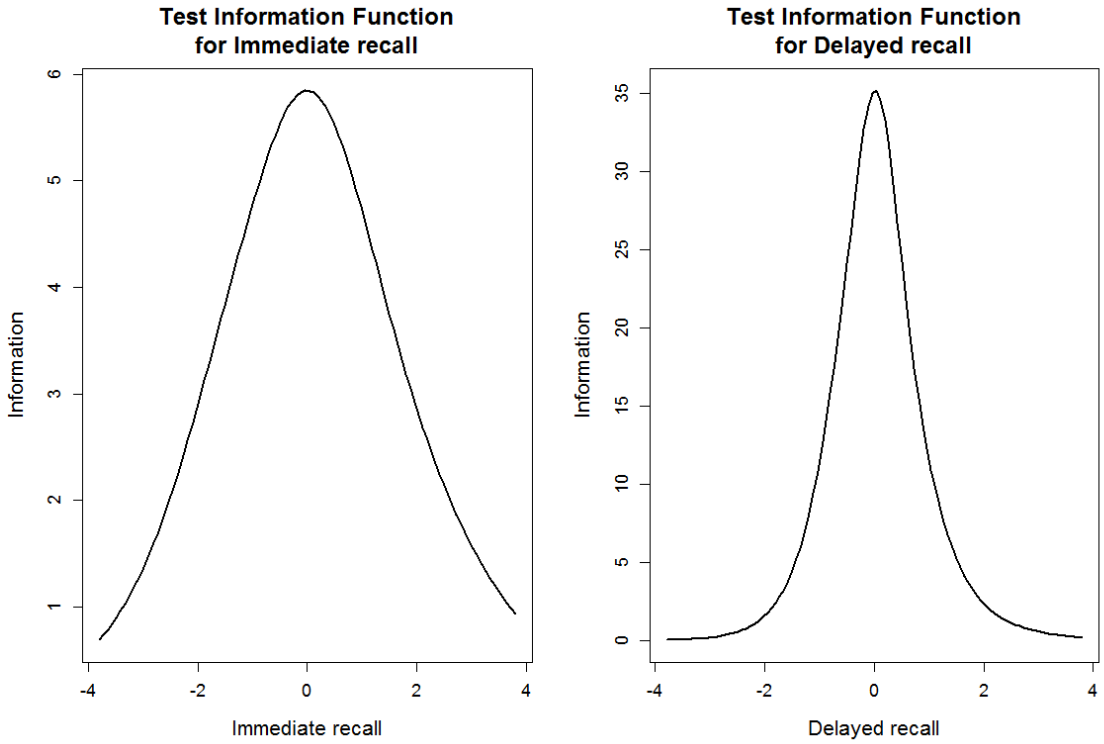
Confirmatory Factor Analysis (CFA)

CFA showed acceptable fit to the unidimensional model of the IR [*RMSEA=.057 (CI90% = .054-.060); CFI=.943; TLI=.936]* and DR scores [*RMSEA=.052 (CI90% = .049-.055); CFI=.952; TLI=.946]*. We also found good reliabilities for both IR (omega = 0.81) and DR scores (omega = 0.81).

Item Response Theory

IRT analyses showed most IR adjusted items informed "moderate" (0.782-1.497, median=0.958), while most DR adjusted items (1.452-4.609, median=1.573) informed "very high". Each item's discrimination and difficulty parameter is available upon request. Test information curves showed all scores are able to capture variability on the range from -2 to +2 z-scores.

FIGURE S1. Test Information Function for immediate and delayed recall



1.2. Reading and writing abilities

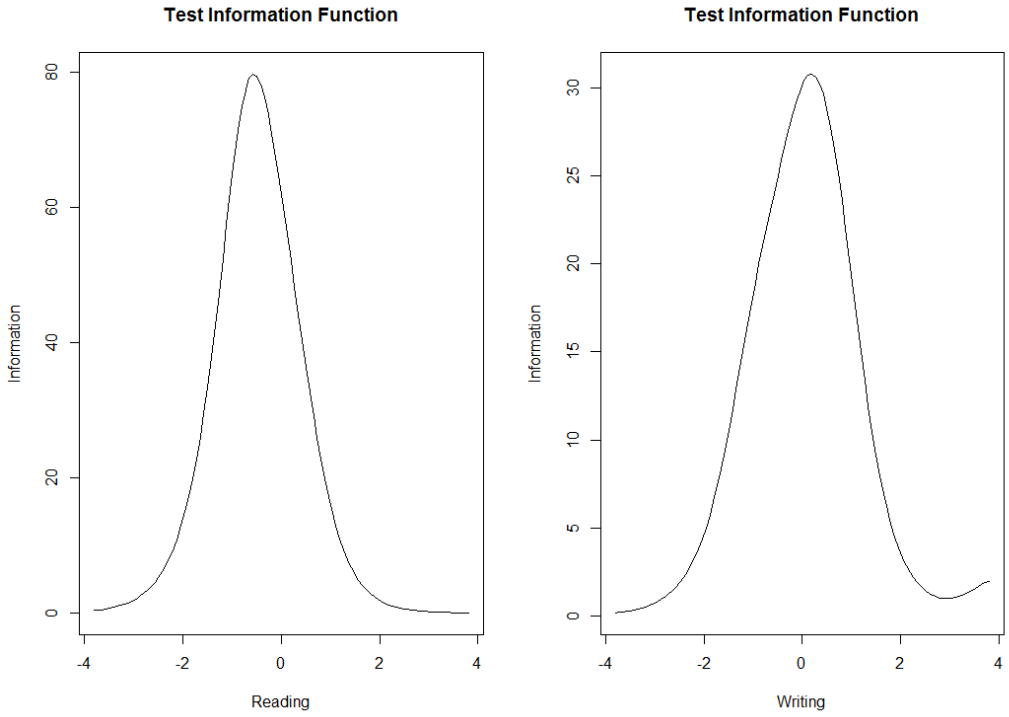
Confirmatory Factor Analyses

CFA showed good fit to the unidimensional model for both reading [*RMSEA=.0019 (CI90% = .0018-.0020); CFI=1.000; TLI=1.000]* and writing [*RMSEA=.031 (CI90% = .030-.033); CFI=.997; TLI=.997]*. We also found good reliability for both models (omega reading =0.987; omega writing=0.958).

Item Response Theory

IRT showed most items informed "very high" for both reading (1.570-3.495, median=2.483) and writing (0.866-2.980, median=2.377). Each item's discrimination and difficulty parameter is available upon request. Test information curves showed all scores are able to capture variability on the range from -2 to +2 z-scores.

FIGURE S2. Test information curves for reading and writing



References

1. Buuren S, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. J Stat Softw. 2011;45(3).

2. Rosseel Y. lavaan: An R package for structural equation modeling. J Stat Softw. 2012;48(2):1–36.

3. Cook KF, Kallen MA, Amtmann D. Having a fit: impact of number of items and distribution of data on traditional criteria for assessing IRT's unidimensionality assumption. Qual Life Res. 2009;18(4):447–60.

4. Hu L, Bentler PM. Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. Psychol Methods. 1998;3(4):424.

5. McDonald RP. Test theory: A unified approach. Mahwah NJ Lawrence Earlbaum. 1999;

6. Samejima F. Graded response model. In: Handbook of modern item response theory. Springer; 1997. p. 85–100.

7. Rizopoulos D. ltm: An R package for latent variable modeling and item response theory analyses. J Stat Softw. 2006;17(5):1–25.

8. Baker FB. The basics of item response theory. College Park, MD: ERIC Clearinghouse on Assessment and Evaluation; 2001.

9. Laird NM, Ware JH. Random-effects models for longitudinal data. Biometrics. 1982 Dec;38(4):963–74.

10. Verbeke G, Molenberghs G. Linear mixed models for longitudinal data. Springer Science & Business Media; 2009.

11. Osterrieth PA. Le test de copie d'une figure complexe. Arch Psychol. 1944;30:206–356.

12. Rey A. L'examen psychologique dans les cas d'encéphalopathie traumatique.(Les problems.). Arch Psychol. 1941;

13. Shin M-S, Park S-Y, Park S-R, Seol S-H, Kwon JS. Clinical and empirical applications of the Rey-Osterrieth Complex Figure Test. Nat Protoc. 2006;1(2):892–9.

14. Meyers JE, Meyers KR. Rey Complex Figure Test and recognition trial professional manual. Psychological Assessment Resources; 1995.

15. Gallagher C, Burke T. Age, gender and IQ effects on the Rey-Osterrieth Complex Figure Test. Br J Clin Psychol. 2007 Mar;46(Pt 1):35–45.

16. Chakravarty MM, Steadman P, van Eede MC, Calcott RD, Gu V, Shaw P, et al. Performing label-fusion-based segmentation using multiple automatically generated templates. Hum Brain Mapp. 2013 Oct;34(10):2635–54.

.

17. Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science. 1993 Aug 13;261(5123):921–3.

18. Myers RH, Schaefer EJ, Wilson PW, D'Agostino R, Ordovas JM, Espino A, et al. Apolipoprotein E epsilon4 association with dementia in a population-based study: The Framingham study. Neurology. 1996 Mar;46(3):673–7.

19. Corbo RM, Scacchi R. Apolipoprotein E (APOE) allele distribution in the world. Is APOE*4 a "thrifty" allele? Ann Hum Genet. 1999 Jul;63(Pt 4):301–10.

20. Wechsler D, Simões M, Ferreira C. WISC-III: escala de inteligência de Wechsler para crianças III. 2002.

21. Vandierendonck A, Kemps E, Fastame MC, Szmalec A. Working memory components of the Corsi blocks task. Br J Psychol Lond Engl 1953. 2004 Feb;95(Pt 1):57–79.

22. Hogan AM, Vargha-Khadem F, Kirkham FJ, Baldeweg T. Maturation of action monitoring from adolescence to adulthood: an ERP study. Dev Sci. 2005 Nov;8(6):525–34.

23. Bitsakou P, Psychogiou L, Thompson M, Sonuga-Barke EJS. Inhibitory deficits in attention-deficit/hyperactivity disorder are independent of basic processing efficiency and IQ. J Neural Transm. 2008;115(2):261–8.

24. Toplak ME, Tannock R. Tapping and anticipation performance in attention deficit hyperactivity disorder. Percept Mot Skills. 2005;100(3):659–75.

2. Exploratory and sensitivity analyses

2.1 Regressions for the Brazilian non-Caucasian subsample

In our main analyses, we used a sample from 668 to 716 children and adolescents, which is described in Table S1. Given that most studies using polygenic risk score for Alzheimer's disease (AD-PRS) use only Caucasian samples, we also performed regressions for a Brazilian Caucasian subsample (n=428) and a non-Caucasian subsample (n=288).

For the Brazilian non-Caucasian subsample, we also found associations of AD-PRS with immediate recall (β=-0.204, p = 0.033) and delayed recall (β=-0.249, p=0.008). We did not find associations with right (β=-0.009, p=0.915) and left hippocampal (β=0.072, p=0.387) volumes. Regressions for the Brazilian Caucasian subsample are available on the main text.

TABLE S1. Brazilian total sample description

	Immediate and delayed recall (n=668)	Reading and writing (n=716)	Executive function (n=677)	Hippocampal volumes (n=670)
		Mean (SD)		
Age	10.19 (1.83)	10.10 (1.85)	10.19 (1.83)	10.19 (1.83)
IQ	101.14 (16.27)	101.13 (16.26)	101.14 (16.27)	101.18 (16.29)
Family income	3207.83 (2288.67)	3188.39 (2279.90)	3207.82 (2288.67)	3208.60 (2294.57)
		%		
Female	45.80%	45.90%	45.80%	45.80%
Ethnicity				
Caucasian	60.30%	60.24%	60.30%	60.30%
Black	9.90%	10.04%	9.90%	10.10%
Multiracial	29.00%	28.94%	29.00%	28.70%
Indigenous	0.60%	0.64%	0.60%	0.60%
Asian	0.10%	0.14%	0.10%	0.10%

2.2. Regressions using other PRS thresholds

We also performed multiple regressions using other thresholds for AD-PRS for the Brazilian total sample. Most important results for other phenotypes are described on the main text. We found associations with immediate recall for threshold $p<0.0119$ ($p=0.008$) and with delayed recall for threshold $p<0.02025$ ($p=0.002$). We also found associations with reading ($p=0.013$) and writing ($p=0.033$) for threshold $p<0.026$. For left hippocampal volume, maximum explanation of the phenotype was found at threshold $p<0.1318$ ($p=0.044$), with several other thresholds also explaining statistically significant levels of variance. For right hippocampal volume, associations were found for threshold $p<0.1185$ ($p=0.009$). No associations were found for executive function. Results are depicted in Figures S3 to S9.

.

FIGURE S3. Associations between immediate recall and AD-PRS for multiple thresholds

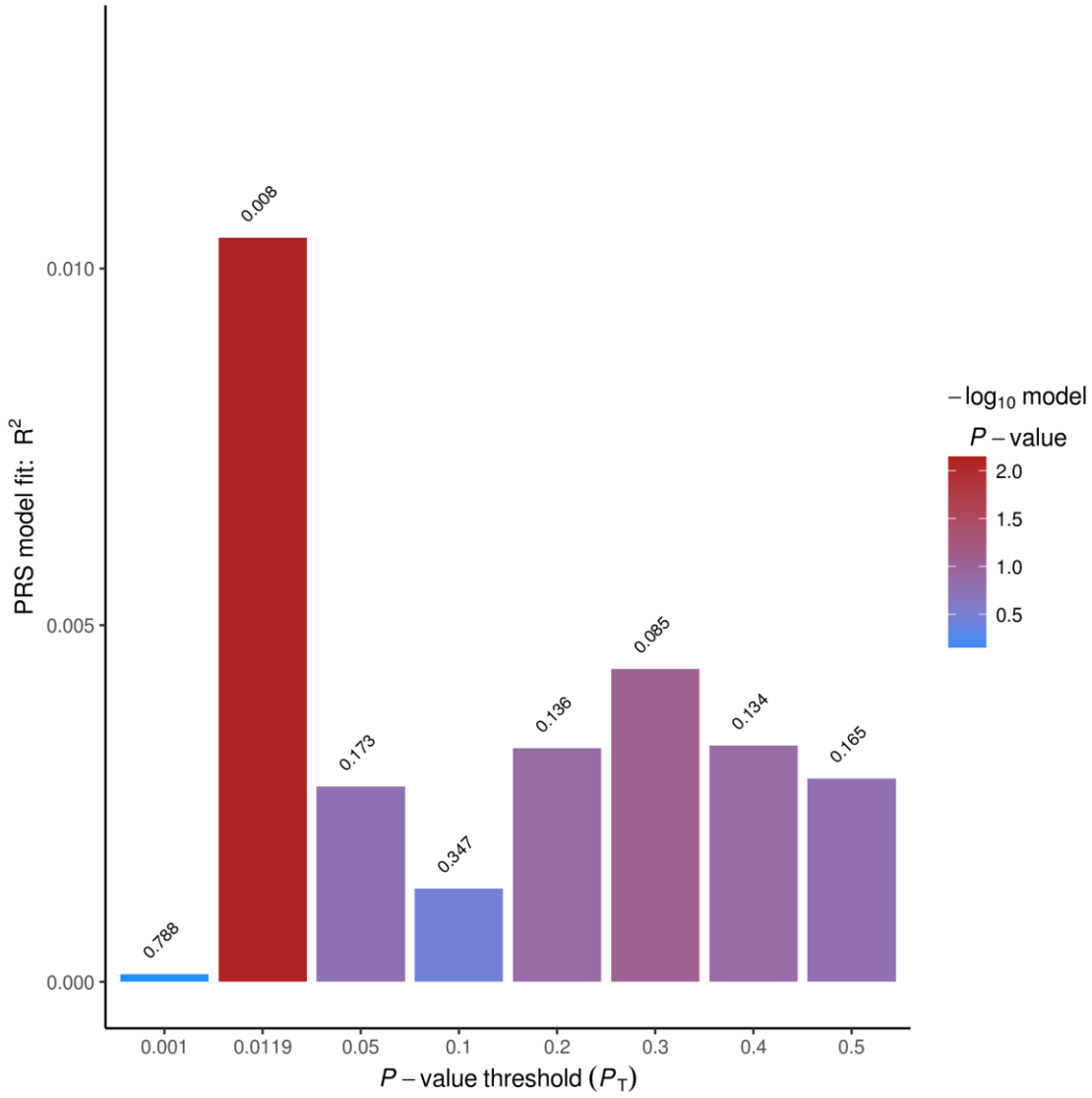


FIGURE S4. Associations between delayed recall and AD-PRS for multiple thresholds

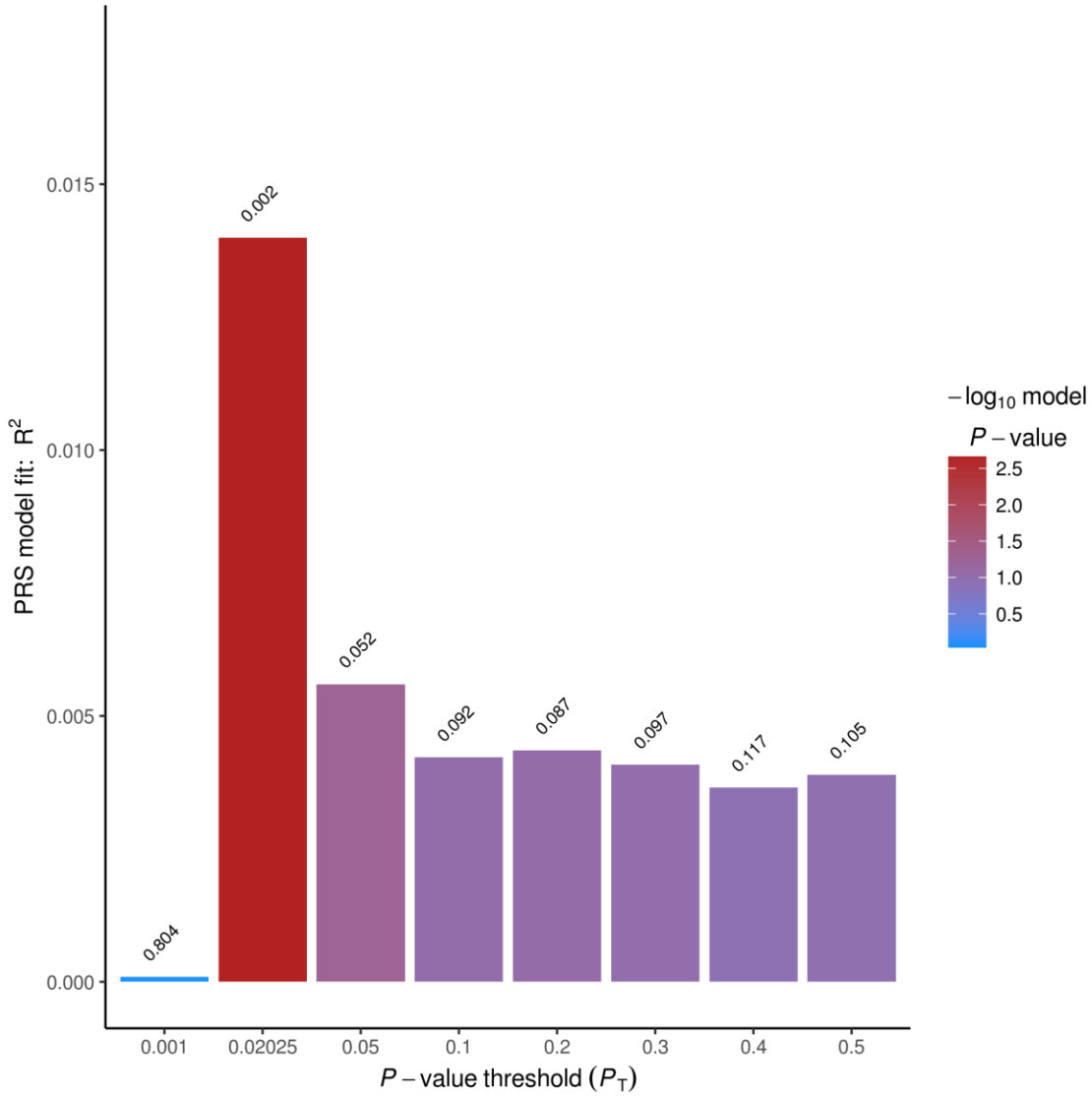


FIGURE S5. Associations between reading and AD-PRS for multiple thresholds

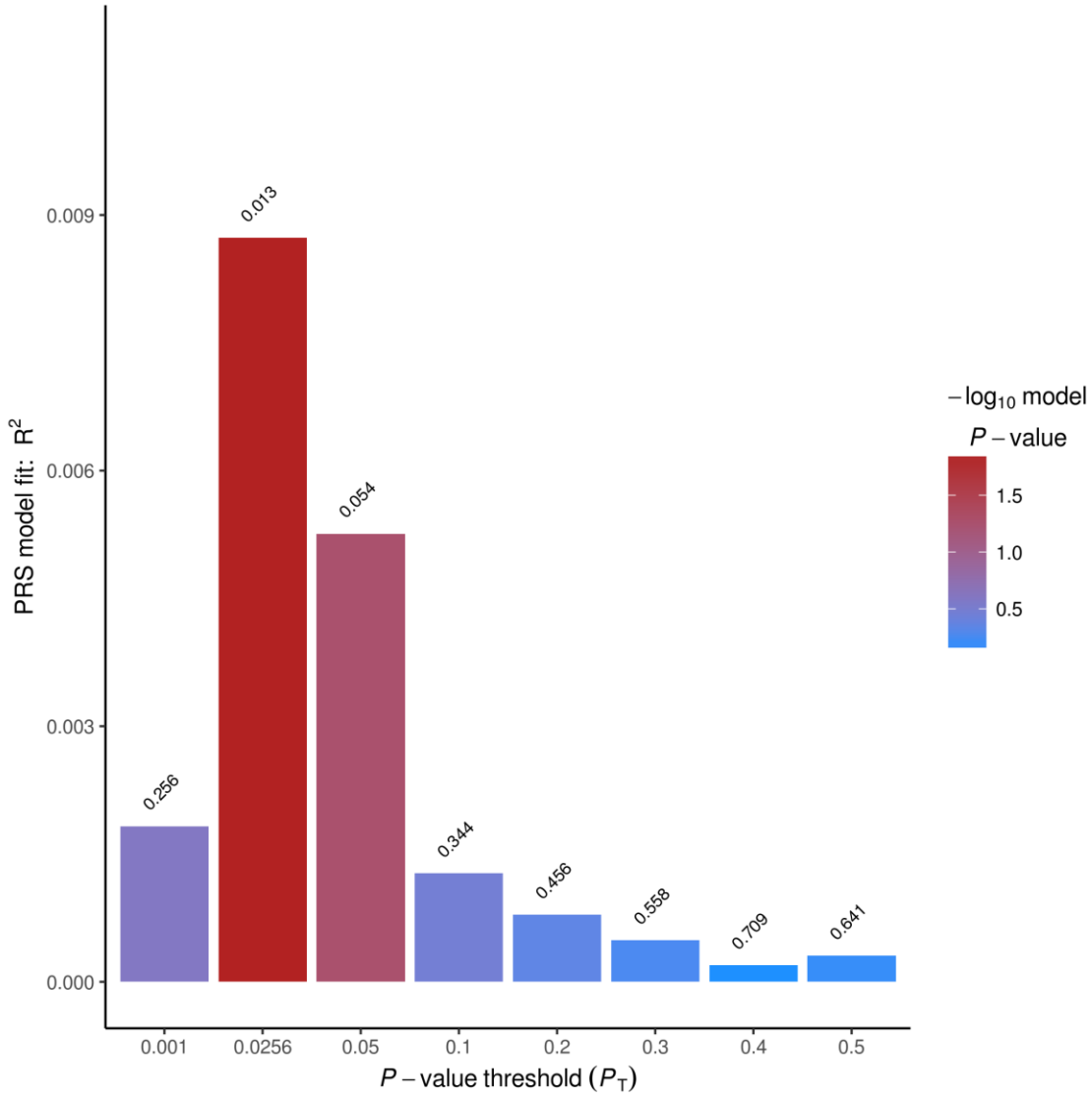


FIGURE S6. Associations between writing and AD-PRS for multiple thresholds

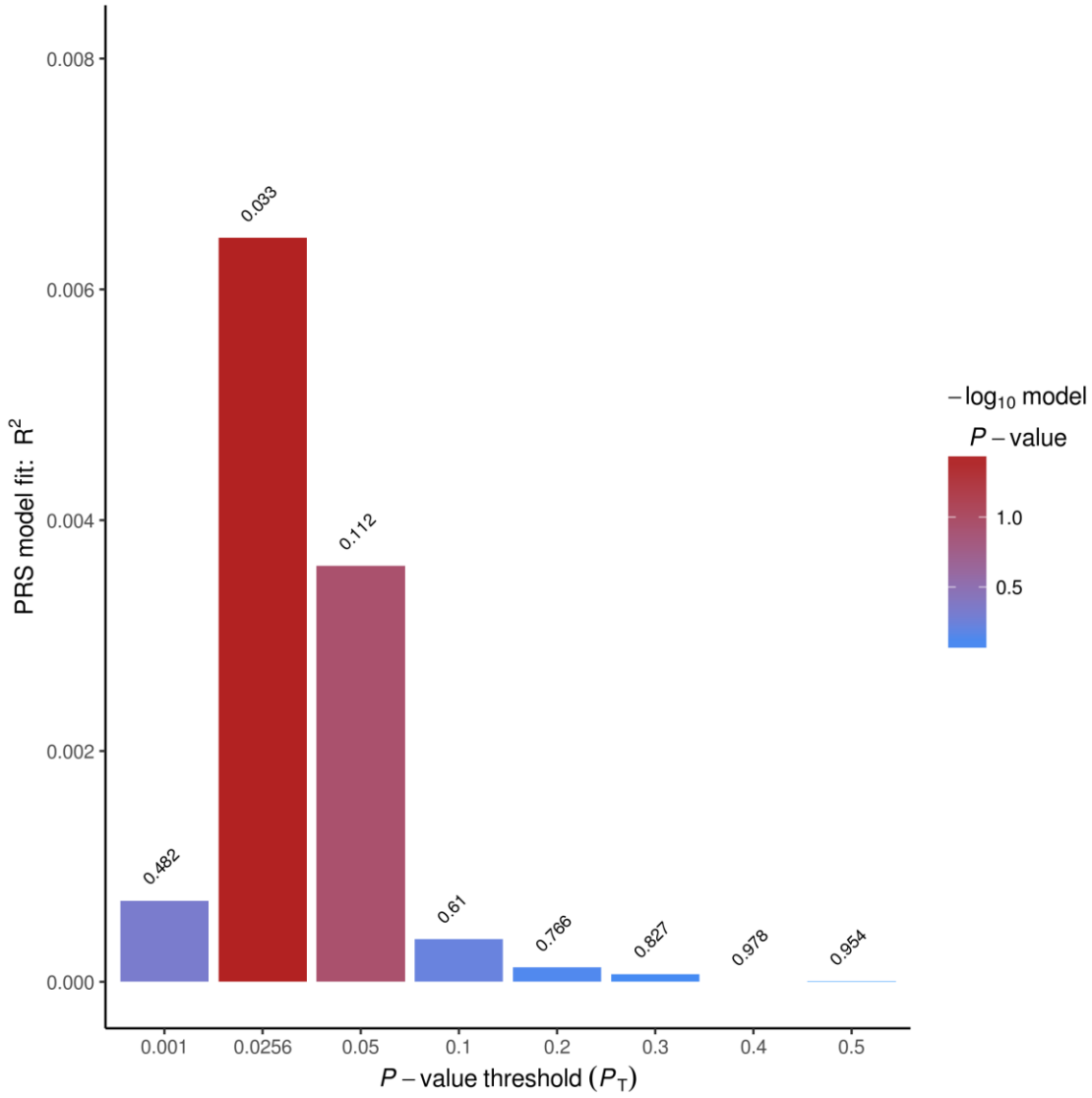


FIGURE S7. Associations between executive function and AD-PRS for multiple thresholds

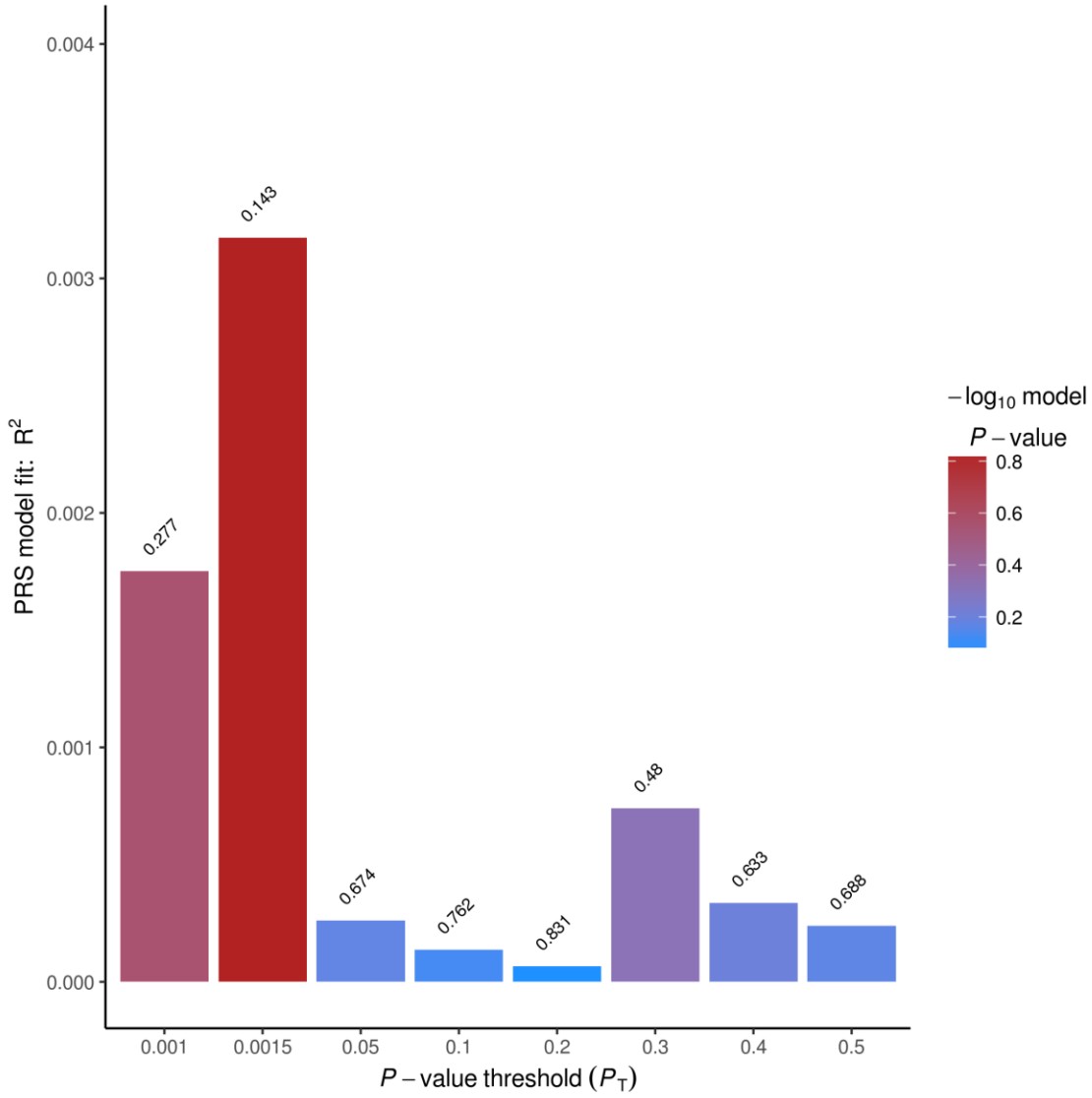


FIGURE S8. Associations between left hippocampal volume and AD-PRS for multiple thresholds

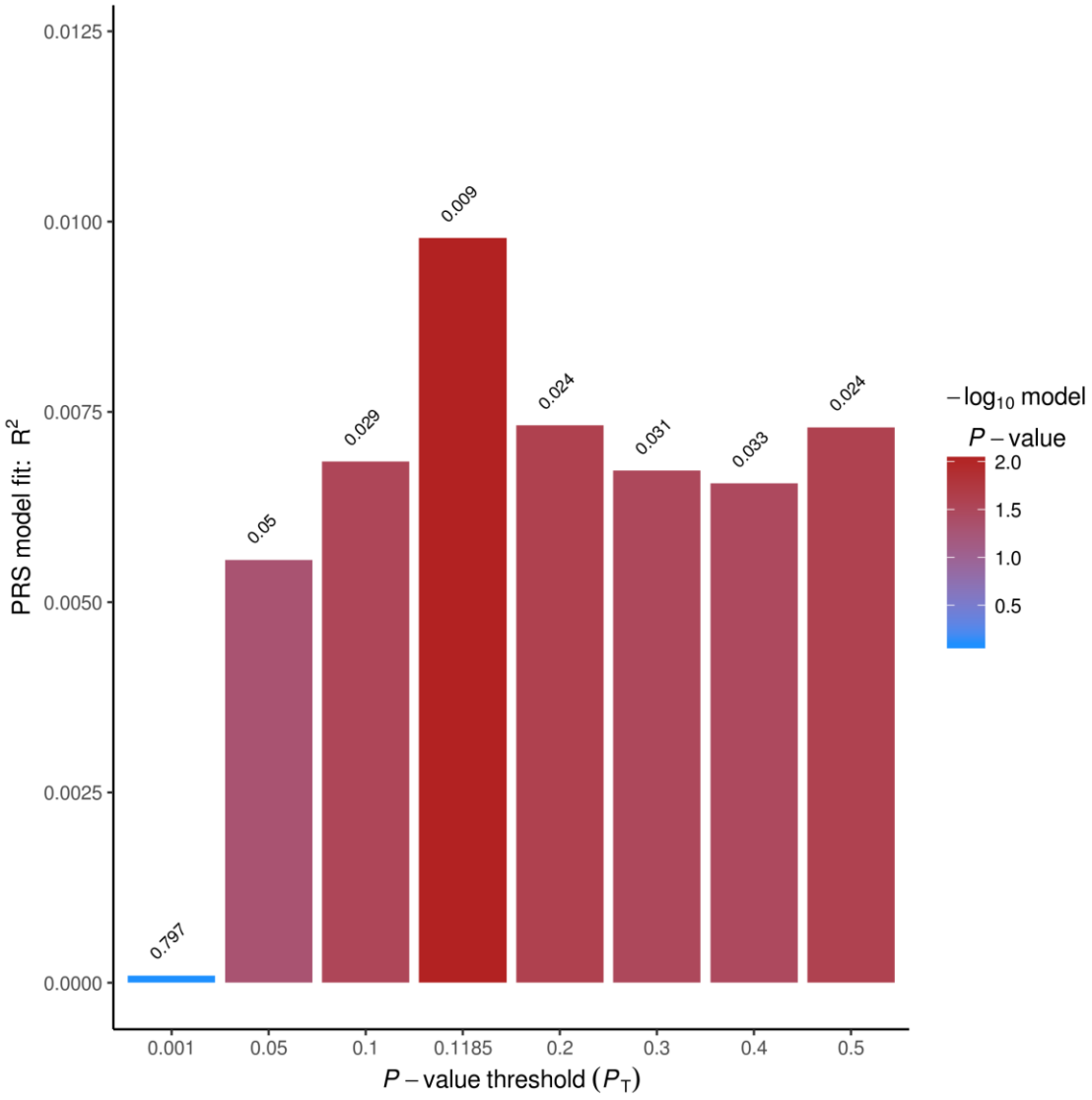
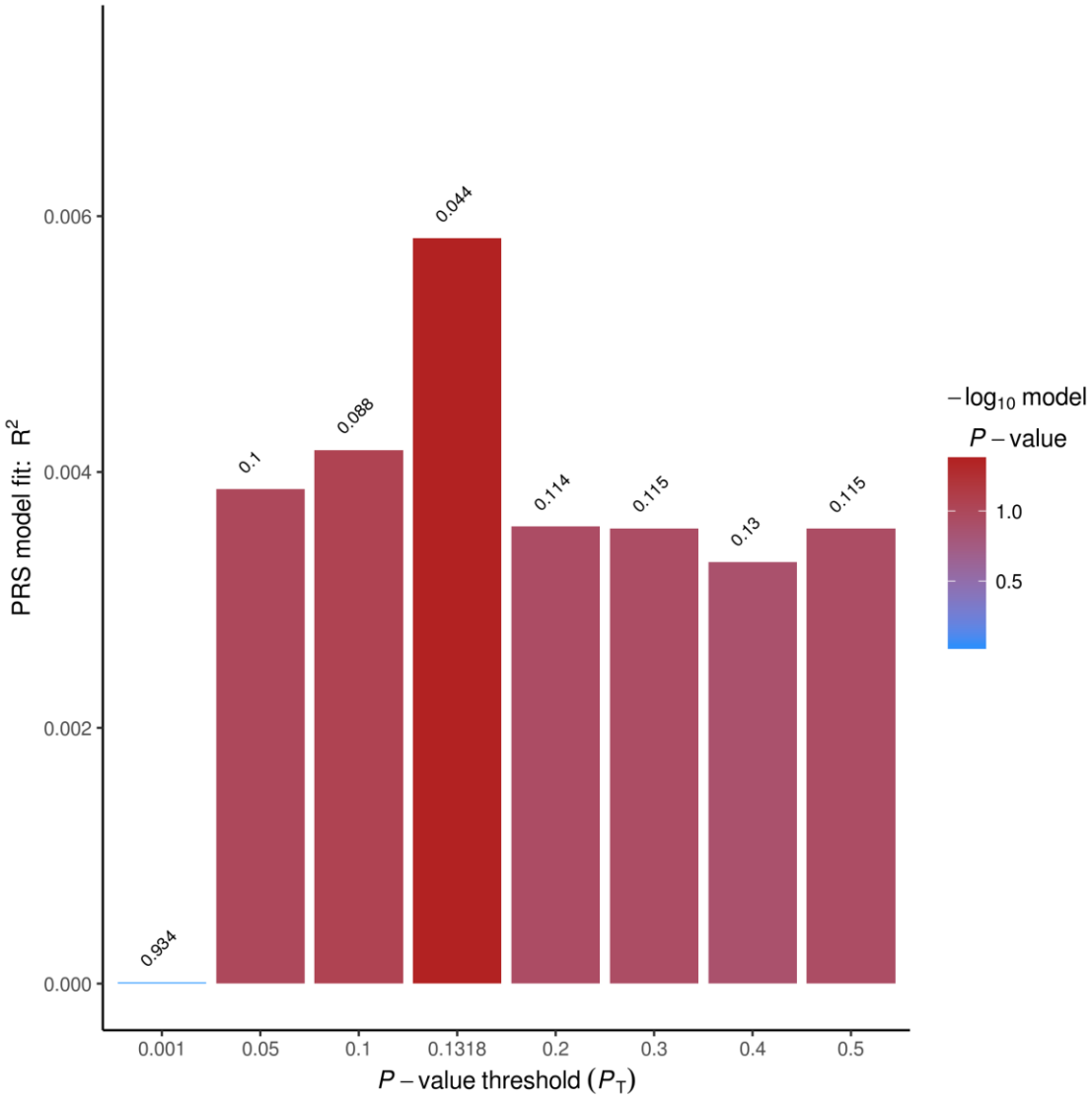


FIGURE S9. Associations between right hippocampal volume and AD-PRS for multiple thresholds

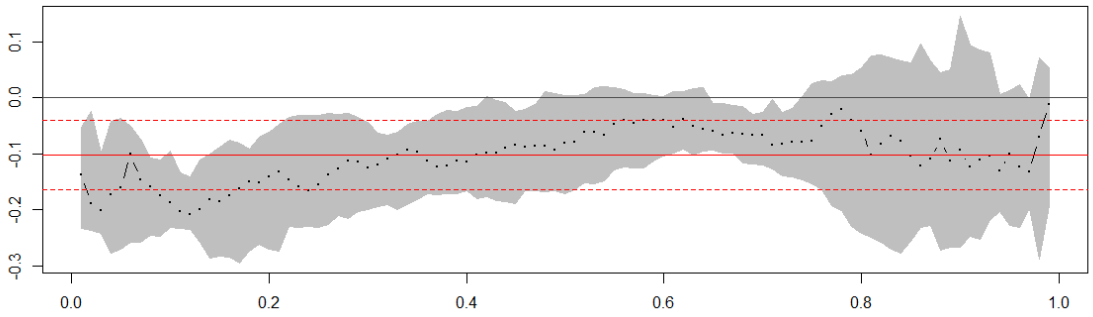


2.3. Quantile regressions

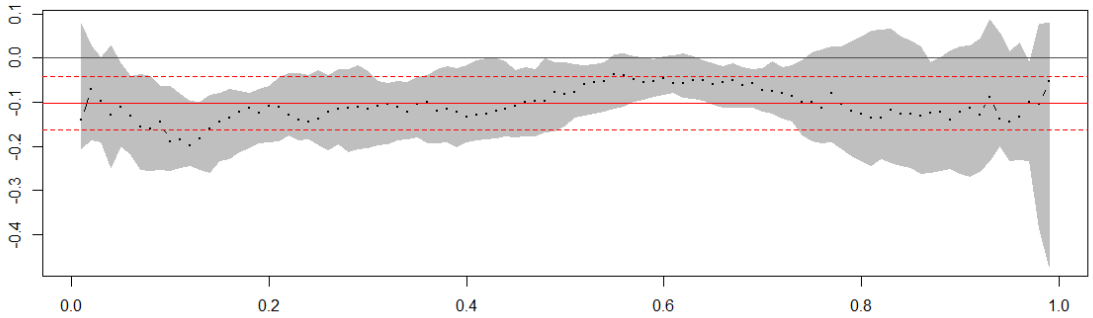
We also performed quantile regressions for the Brazilian total sample to investigate whether associations between AD-PRS and other variables change within different levels of the response variable. In quantile regression, quantiles of the distribution of the response variable are expressed as functions of observed covariates. Quantile regressions using the AD-PRS as the response variable are depicted in Figure S10. Main results for hippocampal volumes are described in the main text. Brazilian total sample description by AD-PRS quintiles can be found in Table S2.

FIGURE S10. Quantile regressions using the AD-PRS as the response variable

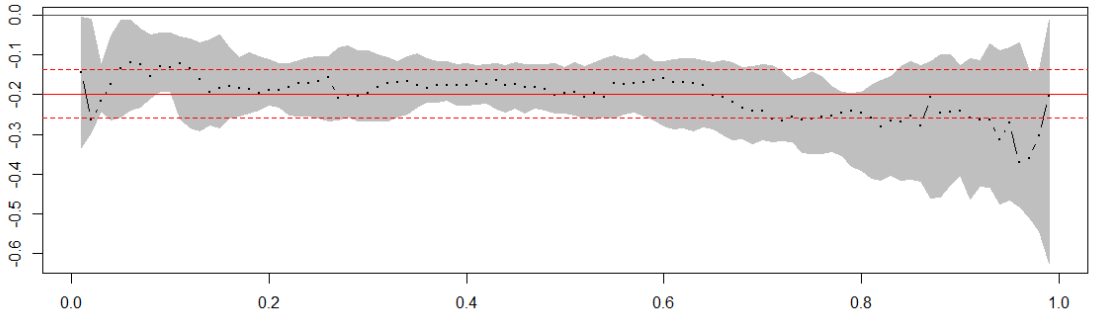
a. Immediate recall



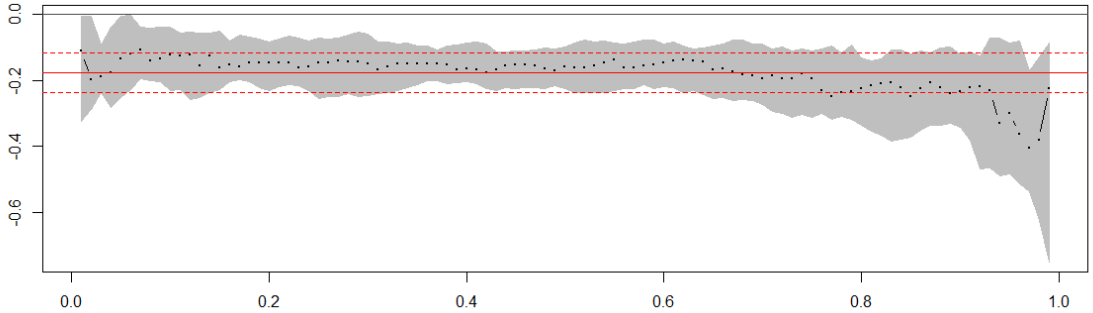
b. Delayed recall



c. Reading



d. Writing



e. Executive function

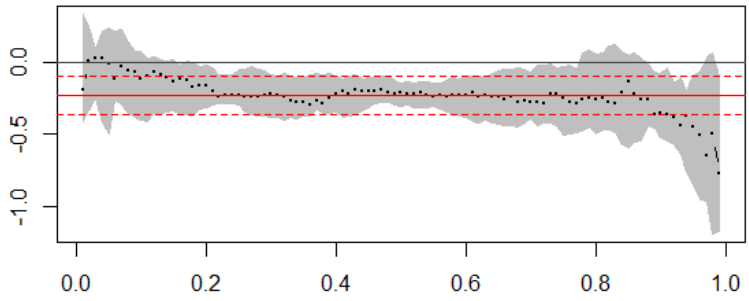


TABLE S2. Brazilian total sample description by quintiles of AD-PRS

	1st quintile	2nd quintile	3rd quintile	4th quintile	5th quintile
			Mean (SD)		
Age	10.30 (1.93)	10.16 (1.99)	9.93 (1.86)	9.93 (1.70)	10.17 (1.79)
IQ	105.47 (17.46)	103.83 (16.57)	99.22 (14.69)	99.06 (15.72)	97.75 (15.44)
Family income	3464.80 (2949.86)	3470.03 (2214.46)	3203.30 (2502.16)	2775.30 (1602.60)	3011.20 (1803.11)
			%		
Female	46.50%	44.80%	42%	50.30%	46.20%
Ethnicity					
Caucasian	85.20%	75.60%	63.70%	51.50%	22.30%
Black	2.10%	3.70%	4.40%	10.30%	30.80%
Multiracial	12%	20%	31.90%	35.30%	46.90%
Indigenous	0%	0.7%	0%	2.20%	0%
Asian	0%	0%	0%	0.70%	0%

2.4. Hippocampal subregions

We performed multiple regressions using hippocampal sub-regions (1). We found associations of AD-PRS with right Ammon's horn (CA) 4 and dentate gyrus (DG) for both Brazilian samples and with left CA 4 and DG, right CA 1 and left subiculum only for the Brazilian discovery sample. We did not find associations with other sub-regions. Results of these regressions are described in Table S3.

.

TABLE S3. Associations between AD-PRS and hippocampal subregion volumes

	Brazilian discovery sample (n=364)			Brazilian replication sample (n=352)		
	β	T	p-value	β	T	p-value
Right CA1	0.214	2.617	0.008*	0.064	0.656	0.5116
Left CA1	0.032	0.082	0.6924	0.079	0.798	0.425
Right CA2 CA3	0.016	0.181	0.8563	0.159	0.090	0.0781
Left CA2 CA3	-0.088	-1.040	0.2984	0.064	0.679	0.4973
Right CA4 DG	0.237	2.797	0.0051*	0.191	1.990	0.0465*
Left CA4 DG	0.176	1.973	0.0485*	0.122	1.291	0.1965
Right subiculum	0.130	1.622	0.1047	-0.008	-0.080	0.9361
Left subiculum	0.2084	2.535	0.0112*	0.128	1.355	0.1753
Right stratum	0.168	1.901	0.0573	0.064	0.678	0.498
Left stratum	0.047	0.562	0.5741	0.068	0.709	0.4784

CA –Ammon's horn; DG – Dentate gyrus.

References

1. Chakravarty MM, Steadman P, van Eede MC, Calcott RD, Gu V, Shaw P, et al. Performing label-fusion-based segmentation using multiple automatically generated templates. Hum Brain Mapp. 2013 Oct;34(10):2635–54.

3. Associations of cognition and hippocampal volumes with APOE

Apolipoprotein E (APOE) is implicated in plasma lipoprotein metabolism and lipid transport within tissues. There are three common polymorphisms for APOE (2, 3 and 4), and the type 4 allele (APOE-epsilon 4) is a major risk factor for Alzheimer's disease (AD) (1). Previous studies showed APOE epsilon 4 to increase the odds of developing AD by up to four fold in heterozygous and up to 30 fold in homozygous compared to non-carriers (2). The frequency of the E4 polymorphisms is estimated in 15% in Caucasians and in 25% in African Americans (3). The frequency of the alleles in the Brazilian total sample was 2.4% (n=17) for 4/4. 19.6% (n=140) for 3/4, 1.8% (n=13) for 2/4, 65.8% (n=469) for 3/3, 10% (n=71) for 3/2, and 0.4% (n=3) for 2/2.

Given the association between APOE-epsilon 4 and AD, we investigated the associations between APOE polymorphisms and memory, verbal abilities, executive function and hippocampal volumes. We adjusted all analyses for the four principal components of genotyping. The results of these analyses appear in the article main text.

References

1. Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science. 1993 Aug 13;261(5123):921–3.

2. Myers RH, Schaefer EJ, Wilson PW, D'Agostino R, Ordovas JM, Espino A, et al. Apolipoprotein E epsilon4 association with dementia in a population-based study: The Framingham study. Neurology. 1996 Mar;46(3):673–7.

3. Corbo RM, Scacchi R. Apolipoprotein E (APOE) allele distribution in the world. Is APOE*4 a "thrifty" allele? Ann Hum Genet. 1999 Jul;63(Pt 4):301–10.

.

4. Assessment of executive functions

Brazilian participants performed six tasks to assess working memory, inhibitory control, and time processing. To evaluate working memory, participants performed the backward digital span (1) and backward Corsi blocks (2). In the first, they were asked to repeat sequences of numbers, in the order stated or in the reverse order. In the second, they were asked to repeat a spatial sequence tapped in nine identical blocks. To evaluate inhibitory control, participants performed the conflict control task (3) and the go/no go task (4). The conflict control task involves indicating either the actual (congruent trials) or opposite (incongruent trials) direction towards which an arrow pointed, with percentage of correct incongruent responses indicating performance. The go/no go task required participants to inhibit the tendency to press a button indicating the directing of arrows when a double-headed arrow appeared, with the percentage of failed inhibitions indicating performance. The time processing task (5) involved trials where participants anticipated the appearance of a visual stimulus after 400ms and 2,000ms, with the percentage of hits indicating performance.

References

1. Martel MM, Pan PM, Hoffmann MS, Gadelha A, do Rosario MC, Mari JJ, et al. A general psychopathology factor (P factor) in children: Structural model analysis and external validation through familial risk and child global executive function. J Abnorm Psychol. 2017 Jan;126(1):137–48.
2. Wechsler D, Simões M, Ferreira C. WISC-III: escala de inteligência de Wechsler para crianças III. 2002.
3. Vandierendonck A, Kemps E, Fastame MC, Szmalec A. Working memory components of the Corsi blocks task. Br J Psychol Lond Engl 1953. 2004 Feb;95(Pt 1):57–79.
4. Hogan AM, Vargha-Khadem F, Kirkham FJ, Baldeweg T. Maturation of action monitoring from adolescence to adulthood: an ERP study. Dev Sci. 2005 Nov;8(6):525–34.
5. Bitsakou P, Psychogiou L, Thompson M, Sonuga-Barke EJS. Inhibitory deficits in attention-deficit/hyperactivity disorder are independent of basic processing efficiency and IQ. J Neural Transm. 2008;115(2):261–8.

.

5. SNPs comprising AD-PRS and enrichment analysis for genes and pathways

Using the 5116 SNPs from AD-PRS ($p<0.01$ threshold), we performed an association study for memory performance and hippocampal volumes adjusting for the four principal components. We found no significant associations between each individual SNP from AD-PRS ($p<0.01$ threshold) and memory performance or hippocampal volumes after corrections for multiple comparisons ($p<9.7\ 10^{-6}$) (Fig. S11), which suggests it is the aggregate weighted risk conferred by the SNPs included in the score and not specific associations that are driving the results.

In order to investigate the genes and pathways associated with AD-PRS, we used MAGMA tool (1) considering all SNPs available in our cohort but using the International Genomics of Alzheimer's Project (IGAP; web.pasteur-lille.fr/en/recherche/u744/igap/igap_download) GWAS p-values to find enriched genes and pathways (using Reactome as background). We found several enriched pathways, including immunoregulatory interactions between lymphoid and non-lymphoid cells (R-HSA-198933), VLDL assembly (R-HSA-8866423) and Netrin-1 signaling (R-HSA-373752). The genes and pathways are described in Tables S4 and S5, respectively.

.

FIGURE S11. Associations between AD-PRS genes and immediate and delayed recall

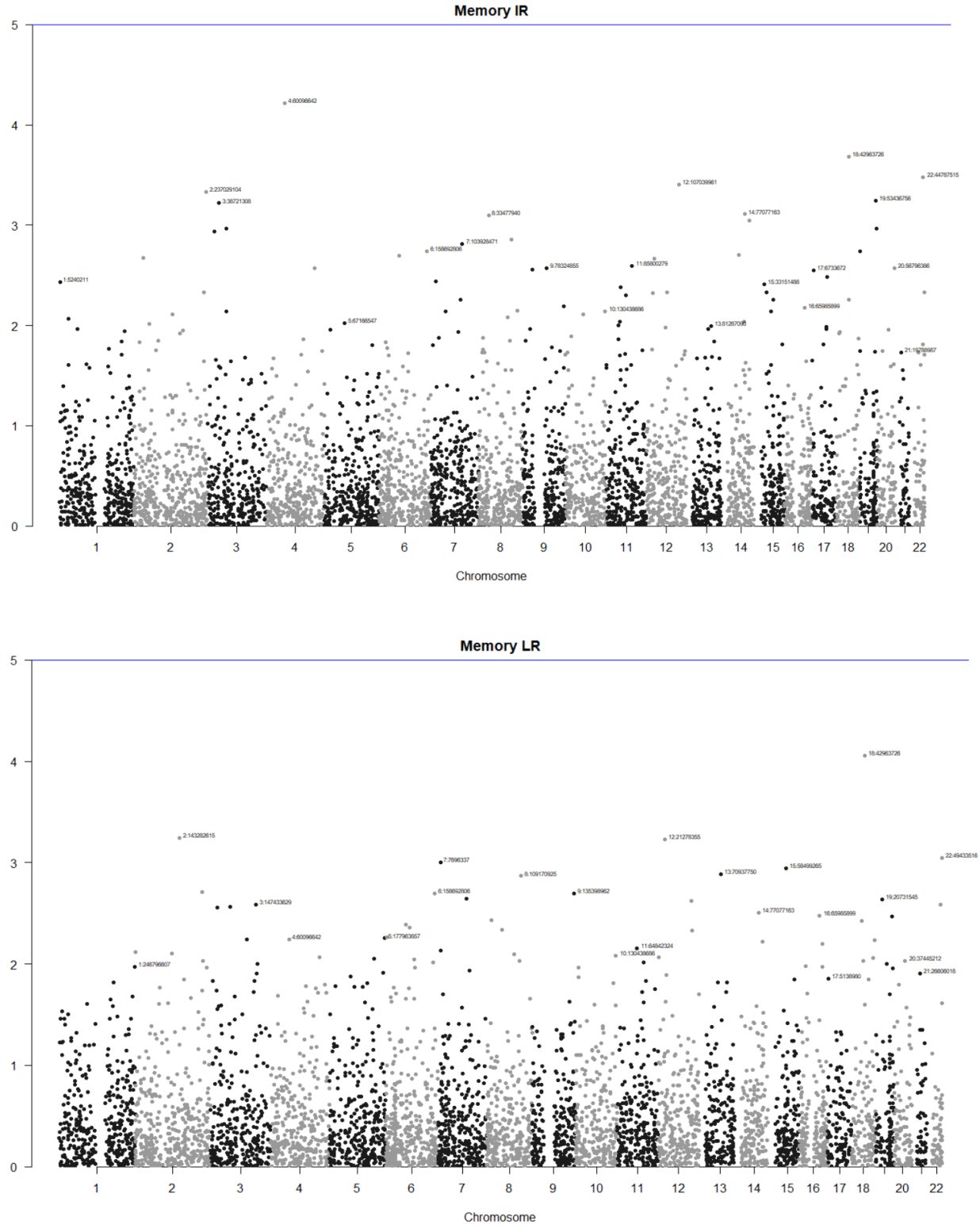


TABLE S4. Top 10 enriched genes

Gene	Chromosome	Start	Stop	Number of SNPS	Number of parameters	Z-value	p
BCAM	19	45311316	45325678	17	14	7.8922	1.48E-15
RELB	19	45503707	45542456	51	31	7.7305	5.36E-15
BCL3	19	45245070	45264301	28	20	7.4684	4.06E-14
CLU	8	27453434	27473328	37	23	7.4531	4.56E-14
CLPTM1	19	45456842	45497604	110	33	7.3626	9.02E-14
APOC4	19	45444495	45449753	10	7	7.1603	4.03E-13
CBLC	19	45280126	45304903	20	15	6.7522	7.28E-12
PVR	19	45146098	45170429	47	25	6.6207	1.79E-11
APOC2	19	45448239	45453822	22	12	6.5317	3.25E-11
ABCA7	19	1039102	1066571	99	55	6.5303	3.28E-11

TABLE S5. Top 10 enriched pathways

Pathway	Number of genes	p-value	Pathway description
R-HSA-198933	109	0.00042212	Immunoregulatory interactions between a Lymphoid and a non-Lymphoid cell
R-HSA-8866423	5	0.0028539	VLDL assembly
R-HSA-428790	13	0.0028873	Facilitative Na+-independent glucose transporters
R-HSA-373752	41	0.0035995	Netrin-1 signaling
R-HSA-428776	4	0.0038013	Class II GLUTs
R-HSA-2408508	8	0.0041715	Metabolism of ingested SeMet, Sec, MeSec into H2Se
R-HSA-428540	13	0.00476	Activation of Rac
R-HSA-1433559	16	0.0060558	Regulation of KIT signaling
R-HSA-156581	14	0.0080862	Methylation
R-HSA-2179392	9	0.0091603	EGFR transactvation by Gastrin

Reference

1. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput Biol. 2015 Apr;11(4):e1004219.

.