Measuring Heritable Contributions to AD: Polygenic Risk Score Analysis in Biometric, SNP-Based & AUC Models with Twins

Ida K. Karlsson[1,2], Valentina Escott-Price[3], Margaret Gatz[2,4], John Hardy[5,6], Nancy L. Pedersen[2,4], Maryam Shoai[5,6], Chandra A. Reynolds[7]

(1)Jönköping University, Jönköping, Sweden, (2)Karolinska Institutet, Stockholm, Sweden, (3)MRC Centre for Neuropsychiatric Genetics & Genomics, Institute of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, United Kingdom, (4)University of Southern California, Los Angeles, CA, USA, (5)UK Dementia Research Institute at UCL, London, United Kingdom, (6)UCL Queen Square Institute of Neurology, London, United Kingdom, (7)University of California Riverside, Psychology, Riverside, CA, USA

Introduction

The polygenic contribution to AD risk varies across study designs and the comparability of estimates remains unclear. AD polygenic risk sores (PRS) capture much of the common genetic influences contributing to risk prediction, with maximum area under the curve (AUC) estimates approaching .90 and inferred heritability estimates of .27 to .55 (Escott-Price et al, 2017). Studies estimating SNP-based heritability report estimates of .24 to .53 (e.g., Ridge et al, 2016). On the other hand, twin studies of AD risk suggest a median heritability of .52, and a maximum value of .79 (Gatz et al, 2014).

Aim

We compare multiple methods all within twin samples and ask -- how do AD PRS contributions vary across methods and what does AD PRS contribute beyond *APOE*?

Methods

Sample

Swedish Twin Registry (STR) samples with clinically-based dementia and AD diagnoses (c.f., Gatz 2006):

- Cases = 430, Controls = 1154, from 1135 twin pairs (449 complete)
- *M*age = 85.29, *SD* = 7.02 years; 44% male

Age distributions across cases and controls are listed in Table 1, and are comparable. Age is coded as last age (age at last follow-up, death, or AD onset).

AD Status	N	LastAge	SD	Min	Max
0	1154	84.67	7.31	66.08	104.39
1	430	86.97	5.88	66.05	102.26
Total	1584	85.29	7.02	66.05	104.39

Analysis

AD PRS scores derived from the recent IGAP2 update (Kunkle et al, 2019). PRSs were adjusted for the first four ancestry PCs, and standardized within SNP array (PsychChip, Omni Express).

GLMM: PRS effects were tested using GLMM (*lme4* in R; Bates et al, 2015) controlling for LastAge and sex, with random effects estimated for pair to account for sibling dependencies.

Biometric: PRS contributions were tested using Mplus 8.4 (Muthen & Muthen, 2017) with Additive genetic (A) and common (C) and person-specific Environmental (E) variance (see Figure 1) with expected correlations: $r_{MZ} = a^2 + p^2 + c^2$ and $r_{DZ} = \frac{1}{2}a^2 + \frac{1}{2}p^2 + c^2$.

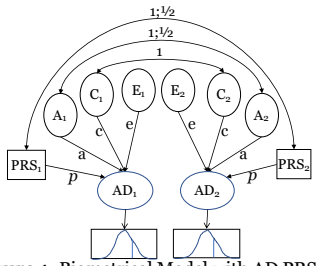

Figure 1. Biometrical Model with AD PRS.

Results

GLMM. The strongest prediction was at a PRS threshold of $p<1x10^{-4}$ (7.3%) but nearly identical to $p<1x10^{-5}$ (7.2%). The prediction of the PRS without the *APOE* region was 1.6% at both thresholds. AUC values ranged from .93 to .96. When *APOE* SNPs were tested, the directly genotyped ε2 and ε4 SNPs contributed 8.8% and the residual PRS without the *APOE* region captured an additional 1.4% at both thresholds (see Table 2). The PRS distribution at $p<1x10^{-5}$ is shown in Figure 2, adjusted for the first four ancestry PCs and array type.

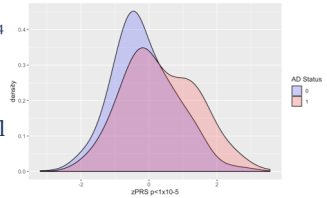

Figure 2. PRS distribution in cases (1) & controls (0) at $p<1x10^{-5}$.

Table 2. GLMM models with AD PRS at $p<1x10^{-5}$

Model	R^2 AD PRS	R^2 APOE ε2 & ε4	AUC
1. PRS p<1x10-5	.072	--	.941
2. PRS p<1x10-5 No APOE region	.016	--	.960
3. PRS No APOE & Direct APOE*	.014	.088	.929

*1151 controls, 428 cases with direct APOE genotyping

Biometric Twin Model. Twin models used complete pairs (190 MZ, 259 DZ). The strongest prediction was at a PRS threshold of $p<1x10^{-5}$. A baseline model suggested heritable influences of .645 (p = .010) or 64.5% of the liability to AD risk (see Table 3). Adding PRS at $p<1x10^{-5}$ contributed 15.9% (p = .041) and 2.6% to this background variation, with and without the APOE region, respectively. The contribution at $p<1x10^{-4}$ was lower at 15.5% (p = .031) and 2.3%, respectively.

Table 3. Twin Model: AD PRS at $p<1x10^{-5}$.

Model	VC	Est	se	*t*	*p*
0. Baseline	A	.645	.251	2.570	.010
	C	.095	.218	.437	.662
	E	.260	.072	3.623	.000
	PRS	--	--	--	--
1. PRS	A	.563	.291	1.937	.053
	C	.002	.240	.010	.992
	E	.276	.092	3.005	.003
	PRS	.159	.078	2.039	.041
2. PRS NO APOE	A	.555	.251	2.213	.027
	C	.159	.209	.762	.446
	E	.260	.083	3.135	.002
	PRS	.026	.026	1.026	.305

Note. VC= Variance Component. Adjusted for LastAge and LastAge^2; VC estimates were constrained equal across males and females.

Results (cont.)

In comparing the IGAP2 summary statistics (Kunkle et al, 2019) to GWAS results based on our current sample, the beta coefficients suggested similar effect sizes for those included in the PRS at $p<1x10^{-5}$ (r=.54, $p < 6.1e^{-08}$; N_{SNPs} = 89) and $p< 1x10^{-4}$ (r=.36, $p < 2.1e^{-08}$; N_{SNPs} = 233).

Conclusion

The estimates of AD PRS contribution to AD risk vary across methods (7.2% – 15.9%) even within the same sample, which may reflect assumptions among the methods about the underlying scale of AD risk. Nonetheless, the *APOE* region explains much of the measurable contribution to AD, with smaller polygenic contribution from other common genetic influences.

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1 - 48. doi:http://dx.doi.org/1.18637/jss.v067.i01

Escott-Price, V., Shoai, M., Pither, R., Williams, J., & Hardy, J. (2017). Polygenic score prediction captures nearly all common genetic risk for Alzheimer's disease. *Neurobiology of Aging*, 49, 214-e7.

Gatz, M., Reynolds, C. A., Fratiglioni, L., Johansson, B., Mortimer, J. A., Berg, S., Fiske, A., & Pedersen, N. L. (2006). Role of Genes and Environments for Explaining Alzheimer Disease. *Arch Gen Psychiatry*, 63(2), 168-174.â

Gatz, M., Jang, J. Y., Karlsson, I. K., Pedersen, N.L. (2014). Dementia: Genes, Environments, Interactions. In D. Finkel and C.A. Reynolds (Eds), *Behavior genetics of cognition across the lifespan*. (pp. 201-231). Advances in Behavior Genetics, vol 1. Springer, New York, NY.

Kunkle, B. W., Grenier-Boley, B., Sims, R., Bis, J. C., Damotte, V., Naj, A. C., ... & Pericak-Vance, M.A. (2019). Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. *Nature Genetics*, 51(3), 414.

Muthén, L.K. and Muthén, B.O. (1998-2017). Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén

Ridge, P. G., Hoyt, K. B., Boehme, K., Mukherjee, S., Crane, P. K., Haines, J. L., ... & Kauwe, J. S. (2016). Assessment of the genetic variance of late-onset Alzheimer's disease. *Neurobiology of Aging*, 41, 200-e13.

Acknowledgements

Funding:
NIH R01AG058068, R01AG060470