

Emily Simmonds, Anna McKeever, Ümran Yaman, Ganna Leonenko, Philip Taylor, Siddharthan Chandran, Mina Ryten, Baljit Khakh, John T O'Brien, Tara Spires-Jones, Dervis A Salih, Valentina Escott-Price
Valentina Escott-Price Lab

BACKGROUND

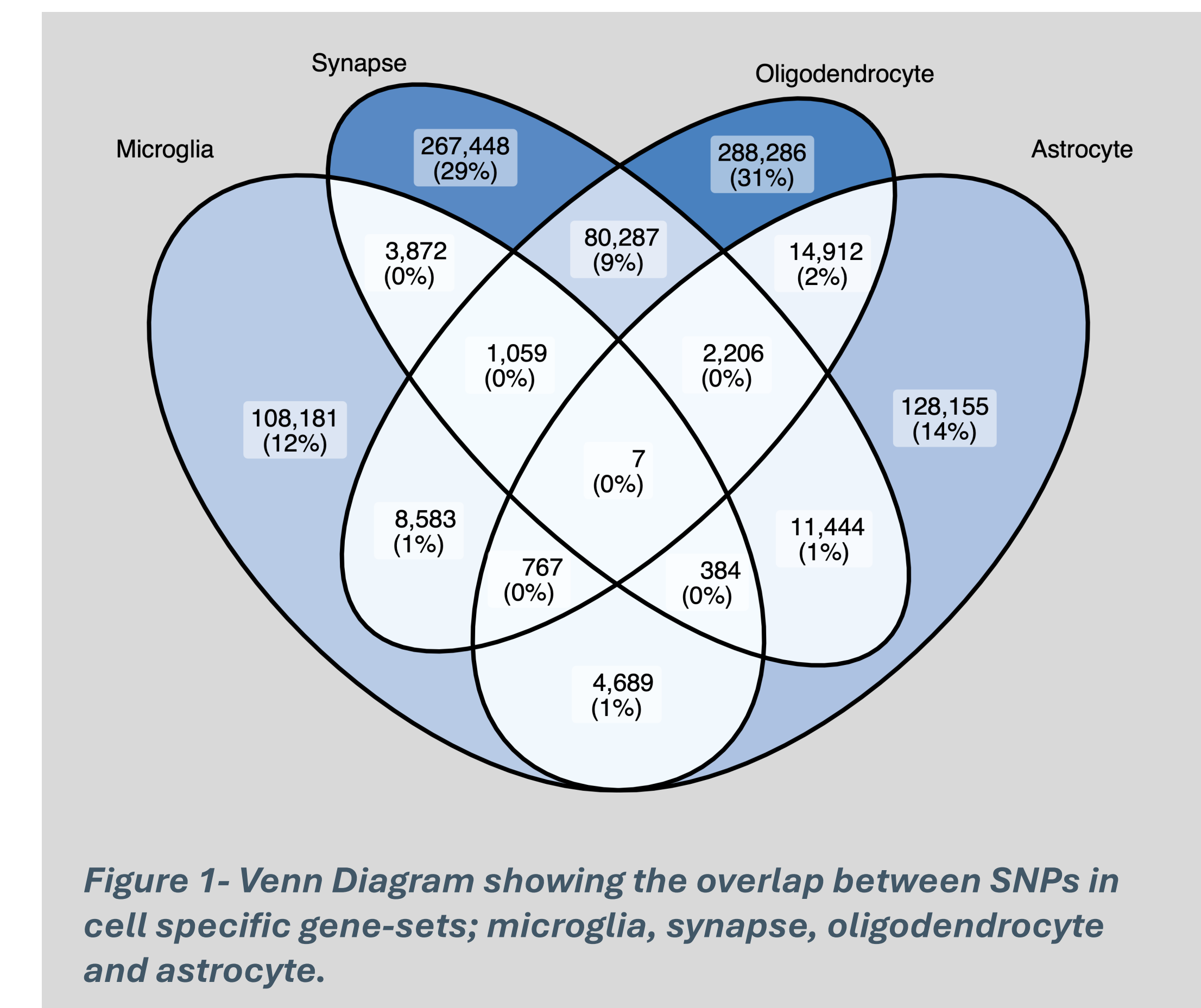
Postmortem brain evaluations in aged individuals frequently reveal multiple coexisting neuropathologies. When combined with clinical diagnostic inaccuracy, this substantially complicates the genetic disentanglement of neurodegenerative diseases (NDDs). This study aims to investigate whether disease-specific and/or cell-type specific PRS have the advantage of improved disease prediction accuracy or between disease specificity, we test whether NDD phenotypes can be predicted by PRS generated with SNP effect sizes from other NDDs.

METHODS

To improve the characterisation of genetic risk and to capture both disease-specific and cross-disorder genetic signatures, we analysed a spectrum of NDDs using disease-specific and cross-disease polygenic risk scores (PRS) in the UK Biobank (UKB), restricting analyses to genes reported as being expressed in major brain cells microglia, astrocytes, oligodendrocytes, and neurons.

Gosselin et al. defined microglia regions based on GWAS signatures and epigenetic/gene regulatory data. A synapse gene-set was defined by Koopmans et al. using Gene Ontology (GO) terms for synaptic locations and processes. An astrocyte gene-set was defined in mice by utilising gene expression and astrocyte morphology data and comparing these genes with astrocytic differentially expressed genes from AD mouse models. An oligodendrocyte gene-set list was based on annotations in ROSMAP data by Mathys et al. and Xiong et al. Genes preferentially expressed by oligodendrocytes in controls were used from the ROSMAP single-nucleus expression data, where genes with $p\text{-adj} < 0.05$ in oligodendrocytes compared to other cell types were used.

We used Cox regression for each NDD-PRS pair to investigate the strength of associations between the PRS and phenotype. We also employed Lasso regression to determine which combination of cell-type PRS best explain each phenotype.

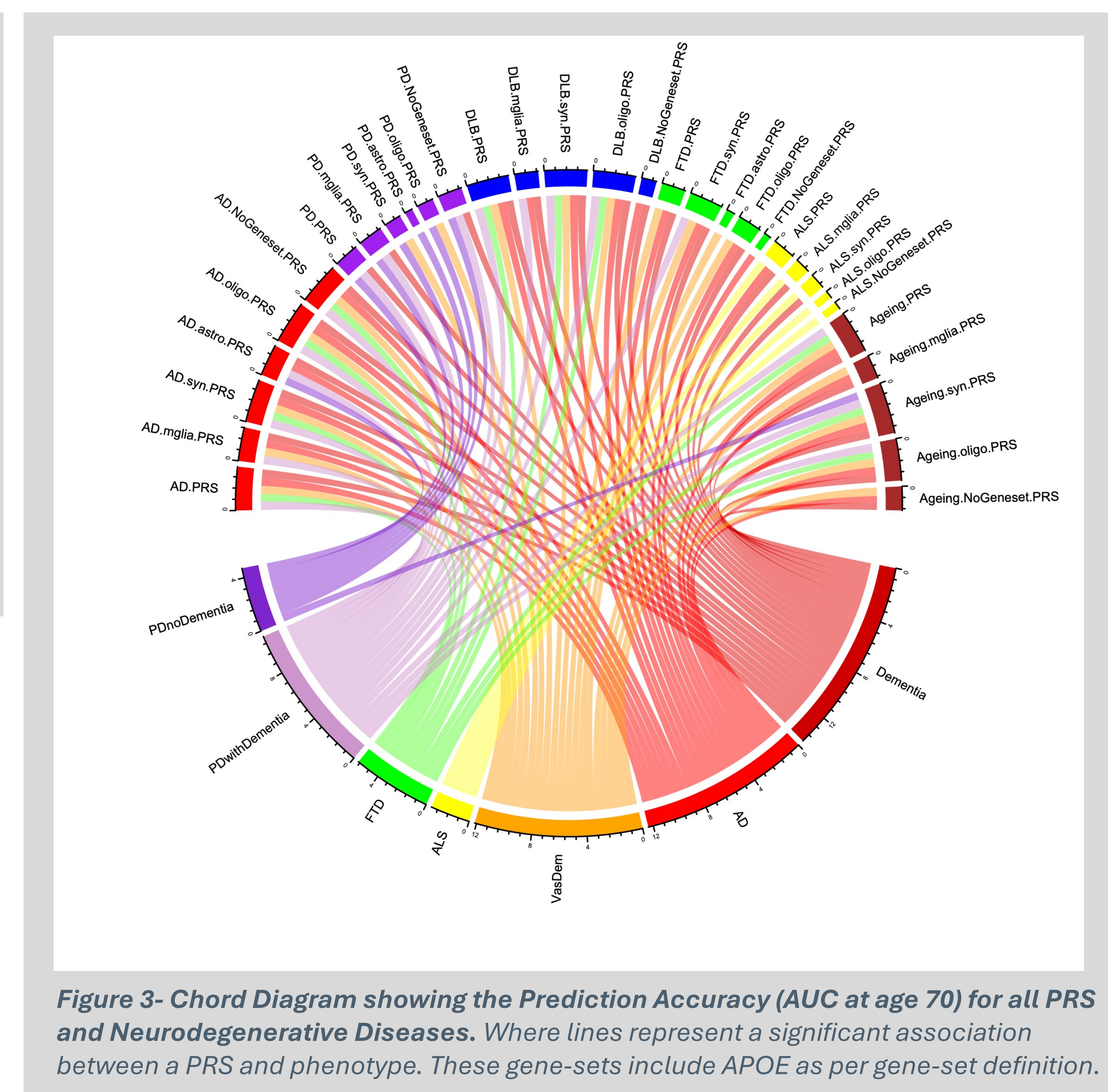
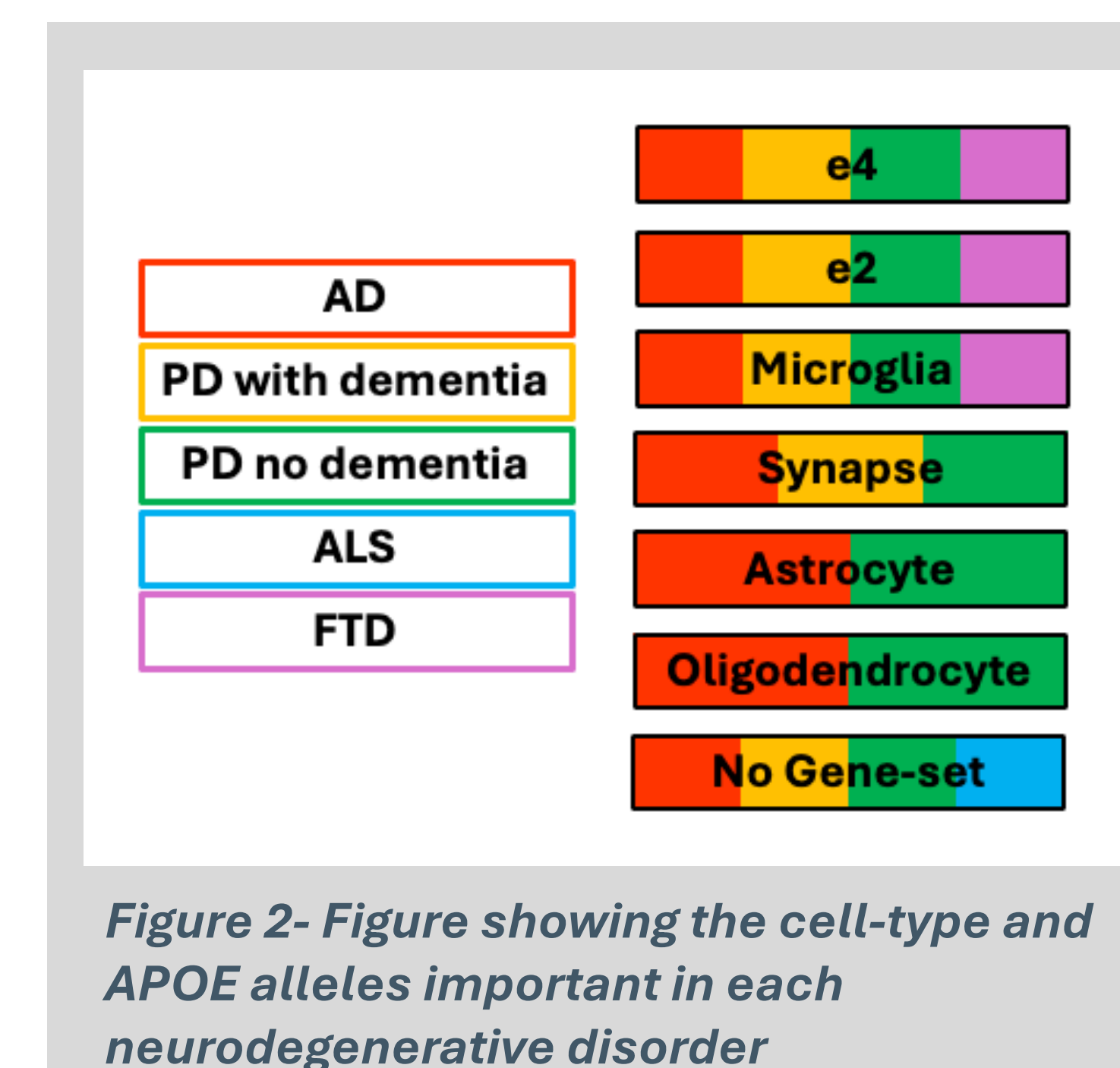


RESULTS

The PRS for each disease is significantly associated with the disease, except for FTD, which nevertheless shows a positive association (Table 1). The lack of significant association is likely due to limited power as the FTD GWAS has one of the smallest sample sizes and the fewest cases in the UK Biobank. The prediction accuracy for PRS is highest for AD by AD-PRS with AUC=0.67.

Lasso regression was used to select which cell-type PRS and *APOE* best explain each phenotype, results are displayed in Figure 2. The microglia gene-set remains significantly associated with all diseases except for ALS, where only the gene-set containing SNPs not in the four gene-sets or the *APOE* region remains in the final model. For AD, PD with and without dementia and ALS, there is some genetic variation remaining in the PRS excluding all cell-type specific gene-sets. In contrast, for FTD, all variation appears to be explained by *APOE* and microglia.

Figure 3 shows all associations between NDDs and cross-disease PRS. Diseases which share dementia as a symptom are predicted by PRS based on multiple diseases. Dementia, AD, vascular dementia and FTD are all associated with AD-, ageing- and DLB- based PRS. PD with dementia is associated with AD-, ageing-, DLB- and PD-based PRS. Only two diseases, PD without dementia and ALS, are only associated with their own disease-corresponding PRS.



PHENOTYPE	MODEL	HR	SE	Padj	R2	AUC at age 70
AD	AD-PRS	1.93	0.02	<2.2e-308	0.369	0.67
ALS	ALS-PRS	1.21	0.04	9.5e-06	0.081	0.59
FTD	FTD-PRS	1.08	0.06	0.323	0.062	0.54
PD with Dementia	PD-PRS	1.14	0.03	7.4e-05	0.187	0.65
PD no Dementia	PD-PRS	1.23	0.02	6.8e-25	0.126	0.60

Table 1- Association and Prediction Accuracy for NDDs with disease corresponding whole genome PRS

DISCUSSION

This study demonstrates the value of cell-type specific polygenic risk scores in predicting neurodegenerative disease, while also highlighting challenges related to misdiagnosis, mixed dementia presentations and limited and inconsistent information about cell-type specificity reported in the public domain. Logically, disorders with dementia as a shared symptom may be predicted by genetic risk scores derived from different underlying diseases.

This study provides important conceptual groundwork in the field of dementia by supporting the idea that future therapeutic strategies will need to target specific brain cell-types.