Identifying individuals with high risk of Alzheimer's disease using polygenic risk scores is most accurate when using all genetic information.

Ganna Leonenko
Cardiff University
Emily Baker
Cardiff University https://orcid.org/0000-0001-5691-597X
Josua Stevenson-Hoare
Cardiff University
Annerieke Sierksma
VIB Center for the Biology of Disease, Leuven, Belgium Center for Human Genetics, Universitaire ziekenhuizen and LIND, KU, Leuven, Belgium. https://orcid.org/0000-0001-9233-972X
Mark Fiers
KU Leuven https://orcid.org/0000-0001-5694-2409
Julie Williams
Cardiff University https://orcid.org/0000-0002-4069-0259
Bart De Strooper
UCL https://orcid.org/0000-0001-5455-5819
Valentina Escott-Price (EscottPriceV@cardiff.ac.uk)
Cardiff University https://orcid.org/0000-0003-1784-5483

Article

Keywords: Alzheimer's disease (AD), genetic risk scores, polygenic risk scores (PRS).

DOI: https://doi.org/10.21203/rs.3.rs-137252/v1

Abstract

There is little agreement regarding the approach and optimal p-value threshold of SNPs to calculate genetic risk scores for Alzheimer's disease (AD). This reflects a fundamental underlying debate on the polygenic versus oligogenic disease architecture. We re-investigated the assumptions underlying the choice of specific p-value thresholds defining genetic loci used to determine polygenic risk scores (PRS). We find the optimal p-value threshold for SNP selection is 0.1, which supports the polygenic architecture of AD. We found that previous studies supporting an oligogenic model of AD did not take account of the reduction of *APOE*-ε4 allele frequency in older individuals, which skewed the results towards lower p-value thresholds and eclipsed the contribution of genes associated to AD with higher p-values. The polygenic approach to AD is also effective to identify individuals at high or low AD risk, when only *APOE*-ε3 homozygous individuals are considered. We also introduce the standardisation of PRS against a population data which ensures comparability of the PRS between studies. In conclusion, our work demonstrates that AD is fundamentally a polygenic disease and that stratifying populations for AD risk best takes the full PRS score into account.

1. Introduction

Alzheimer's disease (AD) is the most common type of dementia, and mainly affects the elderly population. AD is a progressive condition, which means that clinical features develop gradually over many years before diagnosis[1]. The ability to predict AD risk before disease onset is of great importance for personalised prevention and intervention therapies, stratifying people for clinical trials, or the selection of candidates for functional experimental studies.

The common genome-wide significant variants discovered though GWAS have small individual effects, with the exception of *APOE*-ε4[2]. Findings of the optimal p-value threshold (pT) of SNP association with AD, for inclusion in the polygenic risk score (PRS) suggest a range of thresholds from pT≤5e-8 to pT≤0.5 [3–6]. PRS is used as a global term for any number of SNPs included in a risk score. Although, it is clear that many genes are involved in disease development and progression, there is no agreement in the field as to whether AD is polygenic or oligogenic. We define oligogenic risk score (ORS) any PRS that include AD risk associated SNPs only (pT≤1e-5) as in Zhang et al. 2020[7]. The debate in the field became quite heated with recent papers strongly arguing in favour of an oligogenic view of AD[7,8]. However, substantial evidence suggests that the risk of AD is polygenic, similar to other major neurological disorders[3,6]. Given the important implications for the field, we set out to re-investigate the various methodologies and assumptions used for the different calculations, to see whether the disparate viewpoints can be reconciled and to provide all arguments from both sides to a broader readership, so they can make an informed decision on the use of PRS in their studies. We further present an approach to standardise PRS using population based datasets to facilitate it's utility for research and future clinical applications.

About 35% of life-time risk of dementia is modifiable by factors such as education, vascular aspects, social deprivation, etc.[9], which has potentially led to the decrease in incidence of dementia over the last decades [10]. As a consequence, while AD cases are (relatively) easy to detect by clinical assessment (although dementia has also other causes), comparative control samples are likely to be enriched with future AD cases who are yet to show symptoms. Furthermore, if controls are enrolled from a population and/or are younger than cases, not only are a large proportion of them likely to develop AD given time[11], but also the genes with small effect sizes associated with AD due to age related pathological changes[12] can be overlooked. For example, the ε4 allele is associated with earlier age at onset[13] but the ε4 allele frequency in a population decreases from 0.18 to 0.09 with increasing age[14]. A recent study comparing AD cases with relatively young age at onset as extreme cases, with centenarian-controls[15]

observed that the GWAS significant SNPs' effect sizes in this study are on average twice as high as those identified by the original GWAS studies, which confirms the importance of controls being age-matched to or even older than cases. It has been shown that the PRS contribution to AD risk differs with age and *APOE*-ε4 allele status. For example, the effect of PRS ($pT \leq 0.5$) is more pronounced in older people[16], and the effect of oligogenic risk scores constructed using SNPs with an association $pT \leq 1e\text{-}5$ is greater in ε4 homozygotes[8]. Based upon these observations, we hypothesised that unaccounted age-related genetic differences (in particular the *APOE*-ε4 age-dependent frequency) lead to the disagreement about the optimal p-value threshold and the consequent debate about oligogenic (ORS) vs polygenic (PRS) disease models. We tested this hypothesis in simulated data for 10,000 cases and 10,000 controls, varying the ε4 frequencies and the PRS distribution parameters with age, and confirmed this in a real dataset of 549 AD cases and controls.

The accuracy of predicting disease in the individuals at the extremes of the PRS distribution is high[17]. However, the choice of the PRS calculation methodology may lead to identification of different sets of individuals with high/low risk in these extremes of the distribution curve. All methods for PRS calculations attempt to reduce the signal to noise ratio by including fewer SNPs while keeping the most informative ones; of these PRS(P + T) is the simplest one. Bayesian-based methods use all SNPs, and offer strategies to adjust the effect sizes for LD, instead of LD-pruning[18–21]. Functionally informed Bayesian approaches vary the strength of LD-adjustment for each SNP based on its functional annotations. As a result, SNPs with low or medium functional annotation score will have their effect sizes directly scaled down, whereas SNPs with high effect sizes in large LD blocks will be adjusted for LD, but will still be promoted due to less penalisation compared to other SNPs. These methods may reach higher prediction accuracy in a population, but the posterior SNP effect sizes will differ from the "true" effect sizes if they were obtained from, e.g., multivariate regression[22]. Having in mind the goal of robust identification of AD PRS extremes, we explore a variety of PRS generation approaches.

Finally, to choose individuals with high and low PRS in a reliable, replicable and comparable manner, we investigated and PRS standardisation against a population.

2. Materials And Methods

2.1 Data sets and Quality Control

The **1000 Genomes** (1000G) Project[23] applied whole genome sequencing to individuals from different populations in order to compile a detailed resource of common human genetic variation. In this study we only consider individuals from a European population, N = 503.

The **UK Biobank** (UKBB) is a large prospective cohort of approximately 500,000 individuals from the UK containing extensive phenotypic and genotypic data which is still being collected[24]. Participants recruited were aged 39–73 years with a mean age of 56.8. The data here were used under UKBB approval for application 15175 "Further defining the genetic architecture of Alzheimer's disease" and contain 443,018 individuals after Quality Control (QC) analysis. Additional information can be found at (Data Citation1).

HipSci (Human Induced Pluripotent Stem Cell Initiative)[25] is an initiative which is generating a large, high quality reference panel of human iPSC lines for the research community. These are created from tissue donations from both healthy volunteers and patients from particular rare disease communities. There were 1,228 samples from healthy volunteers available from this study.

ADNI (Alzheimer's Disease Neuroimaging Initiative) is a longitudinal study that was developed for the early detection of AD with the use of clinical, genetic and imaging data[26]. The data was collected from 900 participants between ages 55–90. Initially, participants were followed for 2–3 years with repeated imaging scans and psychometric measurements (ADNI1). The study was subsequently extended with the addition of new participants (ADNI-GO and ADNI2). Longitudinal data contained information on clinical assessments from the first, baseline visit to the latest available visit with mean follow up time approximately 5 years. Genetic data was available for 770 participants who provided written consent. More information can be found at (Data Citation2).

ROSMAP - Religious Orders Study (ROS) and the Rush Memory and Aging Project (MAP) are both ongoing longitudinal clinical-pathologic cohort studies of aging and AD. Older participants were recruited without dementia and multi-layer data were collected that includes structural and functional neuroimaging, quantitative clinical phenotypes, neuropathologic and neurobiological traits, multi-level omics and genetics[27–29]. The data were downloaded from (Data Citation3, Data Citation4), 1,196 samples with available genetic information. ROSMAP data can be requested at (Data Citation5).

MSBB (The Mount Sinai Brain Bank) study generated gene expression, genomic variant, proteomic and neuropathological data from brain specimens. Clinical dementia rating scale (CDR) was conducted for assessment of dementia and cognitive status[30]. The data was downloaded (Data Citation6, Data Citation7), resulting in 349 samples with available genetic information.

MAYO - Mayo Clinic Brain Bank is a post-mortem cohort that contains neuropathological, genetic, biochemistry, cell biology data. The samples that are used here are described in MAYO eGWAS[31]. Data was available to download (Data Citation8, Data Citation9), resulting in 349 samples with available genetic information.

All standard Quality-Control (QC) steps were performed separately in each dataset using PLINK[32](Data Citation10), see Supplementary Sect. 1. To gain more power we combined and harmonised ADNI, ROSMAP, MSBB and MAYO studies, removed overlapping samples that were used in the Kunkle et al GWAS study[2], leaving 271 AD cases and 278 controls with 6,077,045 SNPs for the remaining analysis (see details in Supplementary Table 1 and Supplementary Table 2). This data will be referred to as the case-control dataset for the remainder of the manuscript.

2.2. Primary PRS calculation (P + T)

For the PRS calculation we used the summary statistics from the largest available clinically assessed case-control GWAS study on AD[2] (N = 63,926) to generate genetic scores for all participants in the cohorts described above as the weighted sum of the risk alleles. PRS were generated with the PLINK genetic data analysis toolset[32] (Data Citation10) for pT ≤ 5e-8, 1e-5, 0.1, 0.5 on LD-pruned SNPs by retaining the SNP with the smallest p-value excluding variants with $r^2 > 0.1$ in a 1000-kb window. PRS.no.APOE was calculated excluding the *APOE* region (chromosome 19:44.4 Mb 19:46.5 Mb) due to the high LD in this region. PRS.AD was calculated as a weighted sum of PRS.no.APOE and *APOE*(ε2 + ε4), where *APOE* effects were weighted with effect sizes (B(ε2)=-0.47 and B(ε4) = 1.12) as in Kunkle et al. 2019[2]. Prior to any analyses, all derived scores were adjusted for PCs and then standardised a) within the sample and b) against population cohorts. For the latter, the dataset was merged with the population data, PCs were derived on the merged data, then the data was standardised using the mean and standard deviation (SD) from the population subsample. Table 1 details the description for each of the PRS models used throughout this manuscript.

Table 1
Model description for the PRS models presented in the manuscript

Model Name	Model Description
ORS.full	PRS including SNPs with a pT ≤ 1e-5
ORS.no.APOE	PRS including SNPs with a pT ≤ 1e-5 and excluding SNPs in the *APOE* region (chr19:44.4–46.5 Mb)
PRS.full	PRS including SNPs with a pT≤ 0.1 (unless otherwise specified)
PRS.no.APOE	PRS including SNPs with a pT≤ 0.1 and excluding SNPs in the *APOE* region (chr19:44.4–46.5 Mb) (unless otherwise specified)
PRS.AD	PRS calculated as a weighted sum of PRS.no.APOE (including SNPs with a pT≤ 0.1, unless otherwise specified) and *APOE*(ε2 + ε4), where *APOE* effects were weighted with effect sizes (B(ε2)=-0.47 and B(ε4) = 1.12) as in Kunkle et al. 2019

2.3. Other methods of PRS calculation

We computed PRS using a number of different methods, in particular PRSice, LDpred-inf, PRS-CS, LDAK and SBayesR taking the effect sizes from the Kunkle summary statistics[2]. To maintain a fair comparison, all PRS methods are applied to an identical dataset containing the same set of thresholded SNPs (pT $\leq$ 5e-8, 1e-5, 0.1, 0.5). We also computed PRS using the whole genome data without any prior pruning and thresholding with LDpred-inf, PRS-CS and LDAK, but software issues prevented us from being able to run this with SBayesR. The traditional approach (PRS(P + T))[33] requires additional LD pruning. PRSice[34] is a software which implements the PRS(P + T) method automatically and so the same LD-pruning parameters were specified for this approach. LDAK[18] does not require LD-pruning and calculates PRS adjusting SNP effect sizes for LD by reducing the contribution of SNPs in regions of high LD. LDpred-inf[19], PRS-CS[20] and SBayesR[21] are all Bayesian approaches which use estimates of SNP effect sizes based on SNP-based heritability and also account for regional LD structure. LD was estimated using the case-control dataset for LDpred-inf and SBayesR and the 1000 Genomes data for PRS-CS (as this was the only option available in the PRS-CS software). All methods were otherwise implemented using default options. The PRS generated were standardised against the 1000 Genomes population data.

2.4 Statistical analysis

The case-control association analysis was performed using logistic regression with the glm() function in R (Data Citation11). The prediction accuracy was estimated in terms of a) area under the receiver operating characteristic curve (AUC) and b) R^2, the proportion of the variance explained by the regression model. The extremes at ± 2 SD were compared in terms of OR with 95% Confidence Intervals (CI), AUC, cases and controls at each tail of the PRS distribution, and pairwise overlap between the extremes for all methods. For the PRS extremes we compare the results of ORS (pT ≤ 1e-5) and PRS (pT ≤ 0.1), including the PRS.AD model. We used the Haldane correction[35] in instances when cell counts were zero in the 2 × 2 contingency table.

2.5 Simulation study

Independent genotypes were simulated in a sample of 10,000 cases and 10,000 controls. *APOE*-ε4 allele frequency was set at 0.142 in controls and 0.356 in cases[36]. For simplicity, we assumed that the age of cases is above the late onset (e.g., over 85) but the age of the controls is below the average early onset (e.g. below 60 years). To estimate the number of controls who will develop the disease at 84, we used results from[13]: showing a frequency of 91% of

donors with AD of ε4ε4 homozygotes, and a mean age of onset of 68 years of age, for ε4 heterozygotes this is 47% at 76 years, and 20% of ε4 non-carriers at 84 years of age, suggesting there to be about 28% “hidden” or putative cases among the controls. Then we re-simulated ε4 genotypes with slightly reduced allele frequency (f = 0.355) in cases, slightly elevated allele frequency (f = 0.36) for putative controls, so the joint allele frequency is ~ 0.356 for the true cases (10,000 + 2,800) and for the 10,000 young population controls matching the distribution of ε4 frequency by age[14]. We set frequency of ε4 allele to 0.142 for the rest of controls[2]. ORS.full based upon 68 LD-pruned SNPs with pT ≤ 1e-5, including one of the most significant *APOE* variant (rs429358), with frequencies and effect sizes as reported in summary statistics Kunkle et al. 2019[2]. The PRS.full was calculated for 10,068 SNPs, where 68 SNPs had the effect sizes as above and 10,000 SNPs were simulated with minor allele frequencies uniformly distributed between 0.01 and 0.45 (70%/30% of SNPs with minor/major risk allele) and effect sizes decreasing from OR = 1.005 to 1 and from OR = 1.003 to 1 for the cases and putative controls, respectively, see R script (Data Citation 12).

3. Results

3.1 Optimal p-value threshold

In our earlier work on PRS in AD[3,37] we have observed that using the directly genotyped *APOE* isoforms ε2 and ε4 as separate terms in the regression model in addition to the PRS excluding the *APOE* region (PRS.AD), provides higher prediction accuracy than modelling the *APOE* region as part of a full PRS. In the case-control dataset presented here, we observed that the optimal p-value threshold for the PRS depends upon how the *APOE* effect is accounted for. Table 2 presents the AUC and R^2 in the case-control dataset in three scenarios with four SNP p-value thresholds (pT ≤ 5e-8, 1e-5, 0.1, 0.5). The first section of the table shows the model with the PRS used as one variable (PRS.full). In the second section the PRS was calculated excluding the *APOE* region (PRS.no.APOE). The third section shows the model with two independent variables i.e. PRS.no.APOE and *APOE*(ε2 + ε4) (PRS.AD).

Table 2
PRS prediction accuracy for the AD case-control dataset using different p-value thresholds and methods to model APOE.

pT	PRS.full				PRS.no.APOE				PRS.AD		
	N SNPs	AUC (%)	R^2	OR (95% CI)	N SNPs	AUC (%)	R^2	OR (95% CI)	AUC (%)	R^2	OR (95% CI)
APOE (ε2 + ε4)	2	70.0	0.18	2.2 (1.8,2.7)	-	-	-	-	70.0	0.18	2.2(1.8,2.7)
5e-8	65	69.8	0.16	2.2 (1.8, 2.7)	17	55.7	0.02	1.2 (1.0, 1.5)	71.4	0.19	2.4 (2.0, 3.0)
1e-5 (ORS)	126	69.4	0.16	2.2 (1.8, 2.7)	66	56.7	0.02	1.2 (1.1, 1.5)	72.0	0.20	2.4 (2.0, 3.0)
0.1	68,681	64.9	0.09	1.8 (1.5, 2.2)	68,516	61.3	0.06	1.6 (1.3, 1.9)	74.1	0.24	2.8 (2.2, 3.4)
0.5	203,950	62.6	0.07	1.7 (1.4, 2.0)	203,710	60.5	0.05	1.5 (1.3, 1.8)	73.7	0.23	2.7 (2.2, 3.4)

Legend: PRSs were calculated on a case-control cohort (271 clinically defined AD cases and 278 cognitively normal controls) using Kunkle et al. (2019) summary statistics for pT ≤ 5e-8, 1e-5, 0.1, 0.5 LD-pruned SNPs and APOE(ε2 + ε4). The number of SNPs (NSNPs) in each risk score are reported. Three PRS models were considered: PRS.full calculated on the full summary statistics; PRS.no.APOE where the APOE region was excluded (chr19:44.4–46.5 Mb); PRS.AD which is calculated as a weighted sum of PRS.no.APOE and APOE(ε2 + ε4), where APOE effects were weighted with effect sizes (B(ε2)=-0.47 and B(ε4) = 1.12) as in Kunkle et al (2019). The number of SNPs for PRS.AD models is always two more than for PRS.no.APOE. Prediction was estimated in terms of AUC, R^2 and OR with 95% Confidence Intervals (CI).

The best prediction accuracy for the PRS.full model is achieved using genome-wide significant SNPs, pT ≤ 5e-8, (AUC = 69.8%), but this is not better than *APOE(ε2 + ε4)* alone (AUC = 70.0%). When more risk genes are included by relaxing the p-value threshold the AUC decreases to 62.6% (first section of Table 1). The lowest prediction accuracy is observed with the PRS.no.APOE model excluding the *APOE* locus. The prediction accuracy does, however, increase from AUC = 55.7% for pT ≤ 5e-8 to 61.3% for pT ≤ 0.1. Note that the results do not change much between pT ≤ 0.1 and pT ≤ 0.5, despite the inclusion of 3 times as many SNPs at pT ≤ 0.5. The best prediction accuracy (AUC = 74.1%) and the highest variance explained (R^2 = 0.24) is achieved by the PRS.AD model where PRS.no.APOE is combined with *APOE(ε2 + ε4)* (last section of Table 1), using pT ≤ 0.1. The results of the PRS.full model, conversely, show a rather paradoxical trend, i.e. that the prediction accuracy decreases when including more risk SNPs i.e. by relaxing the pT threshold. These opposing results reflect very well the current controversies in the field. To investigate why such contradictory conclusions may be drawn from the same data, we set up a simulation study.

We make the assumption that the population controls are younger than cases for our simulations, as this is often observed in real studies. This implies that some of the control population have not reached the age of disease onset yet. Based upon ε4 frequency and studies of ε4 dependent age at onset[13], we estimate that 28% of them will develop AD (see Methods). Accounting for the prevalence of cases[2] (34%), using reported *APOE-ε4* allele frequency in the whole sample 0.216 and OR = 3.326, we calculate the allele frequencies in cases and controls as 0.356 and 0.142, respectively. To allow for differential allele frequency by age in cases, we simulated genotypes with *ε4* allele frequencies of 0.355 in the 10,000 cases, 0.36 in the 28% of controls predicted to develop AD, so the total allele frequency in cases remains 0.356=(0.34*10,000 + 0.36*2,800)/12,800, as in ORS.full included only the 68 SNPs with frequencies as reported in Kunkle et al. 2019[2] (including the *APOE-ε4*). Then we simulated the PRS including all SNPs (see Methods section for details), and analysed ORS.full, PRS.full and PRS.no.APOE combined with a separate variable *APOE-ε4*. This results to the "contradictory" pattern of AD risk prediction, similar to that observed in other AD PRS studies[7,8] (Supplementary Fig. 1). In particular, ORS.full has advantage over the PRS.full, however when *APOE* is accounted for separately in addition to PRS.noAPOE, the PRS.AD has the best prediction accuracy (AUC) and the variance explained (R^2).

Informed by the simulation results, we have explored the ε4 allele frequencies in the case-control dataset with age (see Fig. 1 (A), and Supplementary Table 2). As reported in other studies, the ε4 allele frequency in this data set decreases with age, the ε3 frequency increases and ε2 frequency remains approximately the same. Figure 1 (B) and (C) shows that ε4 frequency reduces faster in cases than in controls (pink line). The oligogenic risk score (ORS.no.APOE) (based on SNPs with $pT \leq 10^{-5}$) also decreases in cases with age but is on average higher than in controls, with the highest being in ORS.no.APOE for ε44 cases as reported in[8]. Contrary to ORS.no.APOE, the mean of PRS.no.APOE (blue line) is higher in older cases and lower in older controls[16]. Thus, because of the changing allelic frequencies of *APOE* genotypes over age, it is clear that the *APOE* genotype by itself and the ORS.no.APOE become much less accurate predictors in older cases, while the reverse is seen with the PRS.no.APOE score. Clearly, *APOE* and ORS will serve as better predictors of AD risk at younger ages. The PRS increases with age, whether this is a true effect or is due to random variation, requires further investigation and replication. Figure 1 shows that the net age effect for the *sum* of ORS and PRS is smaller than the separate score changes with age. Since these changes are in opposite directions, they cancel each other out if taken as a sum. Moreover, the net effect is approximately the same in cases and in controls. This net effect corresponds to the model that is referred to as "polygenic" in the field and leads to conclusion in favour of an "oligogenic" model. However, the differential age effect, leveraging the polygenic disease architecture, can only be discovered when considering *APOE* (and/or ORS) and PRS.no.APOE separately. Adjusting the combined score for age only corrects for the small net effect. Thus, these sample and simulation data demonstrate that even though the ORS is a good predictor for AD at younger ages, it is mainly driven by the age-specific APOE allele frequency distribution.

3.2 Comparison of PRS calculation approaches

Until now, we have used the PRS(P + T) method for the calculations of PRS. Calculation of PRS is based on different assumptions, and an important consideration is what are the most reliable methods to predict the right patients versus controls with maximal accuracy. Figure 2 shows the results of prediction accuracy of ORS and PRS for six different methods of PRS calculation, namely PRS(P + T), PRSice, LDpred-inf, PRS-CS, LDAK and SBayesR as discussed in the introduction. The highest prediction accuracy was found in our case-control sample for both ORS.full and PRS.full using PRS(P + T) with AUC = 65–70% (R^2 = 0.09 – 0.016), and lowest for SBayesR with AUC = 54–61% (R^2 = 0.01–0.05). It should be noted that LDpred-inf, PRS-CS, LDAK and SBayesR do not require p-value thresholding. Therefore, we computed PRS in the full SNP set for LDpred-inf, PRS-CS and LDAK (SBayesR would not

run for all chromosomes for all SNPs) and results were similar to those from PRS using thresholded SNPs with pT≤0.1 (AUC = 59.3, 69.6 and 59.7% respectively).

PRS(P + T) and PRSice showed very similar results across all prediction metrics; which is in fact anticipated as both methods use the same approach with PRSice performing an automatic filtering of SNPs that may differ from PRS(P + T). We computed the PRS.AD model with each method. In line with the earlier conclusions, both prediction metrics (AUC, R^2) are better when *APOE* is modelled separately and subsequently added to the PRS.no.APOE for all methods (AUC = 73–74%, R^2 = 0.22–0.24). The detailed results can be seen in Supplementary Table 3.

3.3. Population-based standardisation

We compared AD ORS and PRS distributions, the latter with and without *APOE* calculated with the PRS(P + T) approach in two European populations; UKBB (N = 364,236) and 1000 Genomes (N = 503). Both populations are European, however, vary by sample size and genotyping platform. When comparing PRS(P + T) distributions for 1000 Genomes and UKBB, it can be observed that the two distributions are very similar at a p-value threshold of pT ≤ 5e-8, 1e-5 and 0.1, see Supplementary Fig. 2. More differences can be observed though at pT ≤ 0.5, where the UKBB PRS distribution has the mean slightly shifted to the left and a smaller standard deviation than that of the 1000 Genomes. The shift of the mean can be explained by the fact that UKBB participants are reporting fewer illnesses, higher education and occupation than the UK general population[38], which are also modifying life-time risk of AD[9]. The smaller SD of the single-country UKBB-PRS (based on large number of SNPs) is also expected when compared to a sample comprising individuals from a number of European countries (1000 Genomes). For SNPs with an AD risk association p-value below the threshold (pT ≤ 0.1) the AD PRS distribution parameters are sufficiently similar, and for ease-of-use reasons we therefore decided to work with the 1000 Genomes hereafter.

When comparing the PRS.AD distributions of the case-control dataset standardised a) within the dataset and b) against 1000 Genomes (Supplementary Fig. 3, Supplementary Table 4), it can be clearly seen that, as expected, the PRS distribution of the population lies between controls (shifted to the left) and cases (shifted to the right). In addition, the population-based standardisation increases the variation in the case-control sample, implying more cases and controls falling above and below, respectively, a predefined PRS cut-off (e.g. 2SD).

3.4. Individuals at the extreme tails of the PRS distribution

We next investigated to what extent the PRS score can be used to identify, with good confidence, individuals with high and low risk of AD. We define PRS extremes as individuals with a score exceeding ± 2 SD from the data mean or from the population mean, depending on the method of standardisation. We assess the effects of 1000G-based standardization on a human iPSC resource, i.e. HipSci, which is population based, as well as on a case-control dataset. For the PRS.AD model when the HipSci sample is standardised within the sample, 11 positive and 2 negative extremes are observed. When standardised against the 1000G population cohort there are 6 positive and 5 negative extremes. It appears that standardisation of the HipSci data against the population provides no advantage above considering them internally, which is not surprising as the PRS distributions in the population and in the population based HipSci should be the same.

The results become much more interesting when considering a case-control dataset. Now the number of positive and negative extremes is greater for PRS standardised against the population than within the sample (see Table 3 and Supplementary Table 4). The highest OR and prediction accuracy is observed with PRS.AD (OR = 124, AUC = 88.2) and the lowest with ORS.full (OR = 10, AUC = 74.6). Often, when selecting individuals at the extremes of risk for AD, researchers may want to understand risk beyond *APOE*. Thus, in Table 3 we also present the results for extremes selection in the ε3 homozygotes using a score excluding the *APOE* region. As expected, the number of extremes is

lower when *APOE* is excluded, but the accuracy remains high with PRS.no.APOE (OR = 95, AUC = 95.7). The ORS.no.APOE accuracy for ε33 carriers drops to AUC = 56.3 with an OR smaller than 1, showing that the prediction is in the wrong direction. Therefore, the oligogenic model is not useful for discrimination between ε33 cases and controls in these data.

Table 3
Number of ORS/PRS extremes in the case-control dataset standardised within the sample and against 1000 Genomes European population.

Sample	Risk Score	Tail	In-sample standardisation				Population-based standardisation			
			N cases	N controls	OR (95% CI)	AUC	N cases	N controls	OR (95% CI)	AUC
All	ORS.full	Positive	18	2	9 (0.4, 207)	84.2	33	5	10 (1, 75)	74.6
		Negative	1	1			2	3		
	PRS.full	Positive	11	2	20 (3, 145)	81.3	19	3	32 (6, 180)	83.1
		Negative	3	11			3	15		
	PRS.AD	Positive	21	1	100 (3, 2989)	84.5	33	3	124 (6, 2707)	88.2
		Negative	0	3			0	6		
ε3ε3	ORS.no.APOE	Positive	1	3	1.7 (0.1, 38)	43.8	1	2	0.6 (0.03, 14)	56.3
		Negative	1	5			1	3		
	PRS.no.APOE	Positive	4	1	39 (1, 1191)	100	7	2	95 (3, 2683)	95.7
		Negative	0	6			0	10		

Legend: In case-control dataset the number of cases (N cases) and controls (N controls) in PRS were identified and the prediction accuracy of these extremes was assessed with AUC and OR (95% Confidence Intervals) when standardised a) using sample mean and SD b) using mean and SD from 1000 Genomes data. We define PRS extremes as individuals with a score exceeding ± 2 SD from the data mean or population mean. Three models were used for the whole dataset (549 individuals): ORS.full (pT≤ 1e-5), PRS.full (pT≤ 0.1) and PRS.AD (pT≤ 0.1) and two models were used for ε3 homozygotes individuals (N = 267): ORS.no.APOE and PRS.no.APOE. ORS.no.APOE and PRS.no.APOE exclude the APOE region and PRS.AD models APOE separately and subsequently adds this to PRS.no.APOE.

Finally, we examined whether the individuals in the extremes are the same across all different PRS methods for both ORS and PRS for positive and negative extremes separately (see pairwise visualisation plot in Supplementary Fig. 4). It can be observed that the greatest number of shared extremes is between PRS(P + T) and PRSice, which again, was anticipated given the methodological similarities of these approaches. The smallest number of shared identifications is between SBayesR and other methods. Overall, the individuals identified with LDpred-Inf, PRS(P + T), PRSice and PRS-CS overlap considerably, in contrast to LDAK and SBayesR.

It can be seen that there are fewer negative extremes identified by ORS than by PRS in all methods. This is explained by the fact that ORS is predominantly driven by *APOE*-ε4 with the consequence that ORS is not very good at identifying negative extremes. Additional plots for mapping 5 top and 5 bottom PRS.no.APOE extremes in ε33 individuals across different methods are presented in Supplementary Fig. 5. The individuals with the most extreme PRS in both the positive and negative tails are consistent between PRS(P + T) and PRSice, while the identified extremes may differ substantially across the other different PRS methods. We advise using PRS(P + T) or PRSice for the selection of individuals at risk because the SNPs contributing to the PRS can be easily identified which is crucial for future experiments.

4. Discussion

PRS could be useful to identify individuals at risk of disease development, however, the accuracy of current methods for the distribution as a whole, precludes the use of PRS in the clinic (too many false positives and false negatives). Results of this and other studies[17] confirm that identification based on having a PRS above/below a certain threshold provides much better prediction accuracy than attempting to classify all individuals in a dataset.

In this study we provide ample evidence that AD should be modelled as a polygenic disease. In fact, risk of AD is not different from other diseases where liability to disease is continuous, and disease becomes evident after a threshold has been passed (the liability threshold model). In the threshold model, liability for a genetic disorder is (normally) distributed across the population and polygenic risk scores are a measure of disease liability[39]. The relative contributions of alleles of various effect sizes and frequencies are not fully resolved; while common alleles of small risk, captured by genome-wide association study arrays, capture between a third and a half of the genetic variance in liability, *APOE* alone substantially increases risk for the disorder[40]. A major problem with AD and using PRS to categorise people at risk, is the age of the study participants. Here we show that *APOE*-ε4 carriers have a lower burden of common AD risk alleles of small effect, implying that under the liability threshold model, the *APOE* risk is substantial enough to develop the disease with a lesser burden of common risk alleles with small effects. Since allelic variation at the *APOE* locus impacts survival altering the age at onset of AD and risk of other conditions (hyperlipidaemia, atherosclerosis, cardiovascular disease[41–46]), the frequency of *APOE*-ε4 goes down with age whereas genetic liability to AD measured by the PRS increases. In other polygenic diseases like schizophrenia, penetrance of the phenotype is mostly complete at 40 years of age[47], while for AD even at 80 there are still individuals at risk but who have not yet developed AD. Looking at the means of the oligogenic risk score and the polygenic risk score across age groups, we found that following the pattern of *APOE*-ε4 frequency, the ORS decreased with age in cases but was on average higher than in controls. Conversely, PRS increased with age in cases, but decreased in controls. This can be explained if *APOE* and most of the GWAS significant SNPs used to calculate oligogenic scores point to genes which are in the same or overlapping pathways[48–52]. This would also explain why adding the oligogenic scores to the calculation do not improve prediction very much compared to *APOE* genotype alone (see Table 2). An important point here is that ORS is likely not very suitable to identify genes that provide protection, while PRS becomes lower in controls.

Since 1) age is the major confounding factor, 2) *APOE* is strongly associated with the age at onset, and 3) it is difficult to disentangle the aging and disease pathogenic components, we suggest to model *APOE* and PRS.no.APOE as two independent predictors or to use PRS.no.APOE as a predictor in subsamples stratified by the *APOE* genotype. In this study, the results show that the prediction accuracy of the oligogenic risk score was not better than using the effect of the *APOE* gene alone. The best performance overall was found here (and in our earlier study[3]) using the model with two variables (i.e. PRS.AD), *APOE* and PRS at $pT \leq 0.1$, which excludes the *APOE* region (PRS.no.APOE).

These differences in prediction modelling also explain why different optimal pTs may have been observed in other studies[5,7,8].

We also looked at individuals at the extremes of the PRS distributions (above and below 2SDs) and found that both OR and AUC are very high in the whole sample (OR = 124, 95%CI=[6, 2707]) and for the ε3-homozygous individuals (OR = 95, 95%CI=[3, 2683]) using the proposed approach. The confidence intervals for the ORs are of course broad, as the sample size is small when looking at the extremes, but the accuracy remains high. The ORs for the extremes identified by ORS were smaller (OR = 10, 95%CI=[1, 75]) and the ORs had narrower CIs, suggesting that this model identifies a greater number of extremes than the polygenic model, but with poorer accuracy. The oligogenic score was not suitable to identify the extremes in the ε33 individuals with OR = 0.6, i.e. misclassifying high ORS cases as controls and vice versa.

Notably, the prediction accuracies using p-value thresholds of 0.1 and 0.5 (the latter reported in earlier work by us and others[36]) were similar. The reduction of the optimal pT from 0.5 to 0.1 is likely due to the improved estimation of SNP effect sizes, imputation quality and increased GWAS sample size in the latest GWAS[2] in comparison to the earlier GWAS study[53]. Similar findings have been observed for other polygenic disorders, e.g. in Schizophrenia and Bipolar datasets of the Psychiatric Genetic Consortium[54].

Comparing six PRS calculation methods, we conclude that the prediction accuracy in the whole sample is very similar, however, the individuals' scores differ. The choice of the individuals at the extremes of the PRS distribution were concordant with PRSice, LDpred-inf, PRS-CS and PRS(P + T). There were more differences shown between LDAK and SBayesR. Due to lack of transparency of the Bayesian approaches, it is difficult to explain why certain individuals are at high polygenic risk whereas others are not, compared to PRS (P + T), where the SNP effect sizes and the LD pruning parameters are traceable. All these reasons allow us to conclude that for AD, PRS(P + T) is the method of choice.

An interesting and important conclusion of our study is that projecting a relatively small case/control sample onto the general population, results in a much better representation of risk in the study. Since case-control samples are enriched for cases as compared to the general population, the PRS distribution of the former is a mixture of two distributions (cases and controls) with distinct means. The PRS distribution for a population sample is likely to have a mean between the means of cases and controls, and a smaller variance (and hence, standard deviation) than that of the combined case-control sample. Standardising the case-control sample to the population sample will result in the shift of the individual scores in the case-control samples to the positive or negative side of the population mean. This makes the detection of more patients at high risk or with high protection possible. Increasing the size of the population sample will provide better estimates of the population PRS mean and SD (since the standard errors of these estimates will decrease as N increases). Note that including a larger population sample will proportionally increase the total number of people in the above and below 2SD categories in the joint (population plus case-control) sample, but that this will not necessarily be enriched by the individuals from the case-control sample. Hence the use of the 1000G population is an easy and straight-forward way to obtain this beneficial effect.

In conclusion, identifying individuals at high and low polygenic risk is very important for further work to understand how genetic risk translates into mechanisms of disease[40]. It might also become very relevant for drug development efforts targeting precise mechanisms of disease, as the PRS scores could be used to select small samples of patients in which proof of concept for the treatment can be obtained before testing the drug in larger cohorts. We show here that for AD, the optimal p-value threshold is $pT \leq 0.1$, and the PRS calculation should account for the age-

related genetic differences in cases and controls either by modelling *APOE* separately to the PRS or matching cases and controls for age and *APOE* status. This adjustment will be refined when we have a better idea of which genes are contributing to the disease aetiology via aging and which are directly on the pathology pathway.

Declarations

Contributions

GL, EB, JSH – performed statistical analysis and wrote the manuscript; AS, MF-reviewed and approved the manuscript and add valuable comments; JW, BS, VEP-conceived and designed the study and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Data Citation:

1. https://www.ukbiobank.ac.uk/
2. http://adni.loni.usc.edu/
3. https://www.synapse.org/syn3191087
4. https://www.synapse.org/syn10901595
5. https://www.radc.rush.edu
6. https://www.synapse.org/syn6101474
7. https://www.synapse.org/syn10901600
8. https://www.synapse.org/syn3817650
9. https://www.synapse.org/syn10901601
10. https://www.cog-genomics.org/plink2/
11. https://www.R-project.org/
12. https://github.com/DRI-Cardiff/APOE-modelling

References

1. Sperling, R. a *et al.* Toward defining the PRECLINICAL stages of Alz Disease: recommendations from the NIAGING-ALZ ASSO workgroups on DX guidelines for AD. *Alzheimers. Dement.* (2011). doi:10.1016/j.jalz.2011.03.003
2. Kunkle, B. W. *et al.* Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. *Nat. Genet.* (2019). doi:10.1038/s41588-019-0358-2
3. Escott-Price, V. *et al.* Common polygenic variation enhances risk prediction for Alzheimer's disease. *Brain* **138**, 3673–3684 (2015).
4. Cruchaga, C. *et al.* Polygenic risk score of sporadic late-onset Alzheimer's disease reveals a shared architecture with the familial and early-onset forms. *Alzheimer's Dement.* (2018). doi:10.1016/j.jalz.2017.08.013
5. Itziar de Rojas, Sonia Moreno-Grau, Niccolò Tesi Benjamin Grenier-Boley, Victor Andrade, Iris Jansen, Nancy L. Pedersen, Najada Stringa, Anna Zettergren, Isabel Hernández Laura Montrreal, Carmen Antúnez, A. Common

variants in Alzheimer ' s disease : Novel association of six genetic variants with AD and risk stratification by polygenic risk scores. *MedRxiv* 1–16 (2020).

6. Altmann, A. *et al.* A comprehensive analysis of methods for assessing polygenic burden on Alzheimer's disease pathology and risk beyond APOE. *Brain Commun.* (2020). doi:10.1093/braincomms/fcz047
7. Zhang, Q. *et al.* Risk prediction of late-onset Alzheimer's disease implies an oligogenic architecture. *Nat. Commun.* (2020). doi:10.1038/s41467-020-18534-1
8. Fulton-Howard, B. *et al.* Greater effect of polygenic risk score for Alzheimer's disease among younger cases who are apolipoprotein E-ε4 carriers. *Neurobiol. Aging* (2020). doi:10.1016/j.neurobiolaging.2020.09.014
9. Livingston, G. *et al.* Dementia prevention, intervention, and care. *The Lancet* (2017). doi:10.1016/S0140-6736(17)31363-6
10. Satizabal, C. L. *et al.* Incidence of Dementia over Three Decades in the Framingham Heart Study. *N. Engl. J. Med.* (2016). doi:10.1056/nejmoa1504327
11. Escott-Price, V. *et al.* Genetic analysis suggests high misassignment rates in clinical Alzheimer's cases and controls. *Neurobiol. Aging* (2019). doi:10.1016/j.neurobiolaging.2018.12.002
12. Mann, D. M. A., Yates, P. O. & Marcyniuk, B. Alzheimer's presenile dementia, senile dementia of alzheimer type and down's syndrome in middle age form an age related continuum of pathological changes. *Neuropathol. Appl. Neurobiol.* (1984). doi:10.1111/j.1365-2990.1984.tb00351.x
13. Liu, C. C., Kanekiyo, T., Xu, H. & Bu, G. Apolipoprotein e and Alzheimer disease: Risk, mechanisms and therapy. *Nature Reviews Neurology* (2013). doi:10.1038/nrneurol.2012.263
14. McKay, G. J. *et al.* Variations in apolipoprotein e frequency with age in a pooled analysis of a large group of older people. *American Journal of Epidemiology* (2011). doi:10.1093/aje/kwr015
15. Tesi, N. *et al.* Centenarian controls increase variant effect sizes by an average twofold in an extreme case–extreme control analysis of Alzheimer's disease. *Eur. J. Hum. Genet.* (2019). doi:10.1038/s41431-018-0273-5
16. Bellou, E. *et al.* Age-dependent effect of APOE and polygenic component on Alzheimer's disease. *Neurobiol. Aging* (2020). doi:10.1016/j.neurobiolaging.2020.04.024
17. Sims, R., Hill, M. & Williams, J. The multiplex model of the genetics of Alzheimer's disease. *Nature Neuroscience* (2020). doi:10.1038/s41593-020-0599-5
18. Speed, D. & Balding, D. J. SumHer better estimates the SNP heritability of complex traits from summary statistics. *Nat. Genet.* (2019). doi:10.1038/s41588-018-0279-5
19. Vilhjalmsson, B. *et al.* Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. *Model. Link. Disequilibrium Increases Accuracy Polygenic Risk Scores* (2015). doi:10.1101/015859
20. Ge, T., Chen, C. Y., Ni, Y., Feng, Y. C. A. & Smoller, J. W. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat. Commun.* (2019). doi:10.1038/s41467-019-09718-5
21. Lloyd-Jones, L. R. *et al.* Improved polygenic prediction by Bayesian multiple regression on summary statistics. *Nat. Commun.* (2019). doi:10.1038/s41467-019-12653-0
22. Escott-Price V. & K, S. Challenges of adjusting SNP effect sizes for Linkage Disequilibrium, Human Heredity. *Hum. Hered.* (2020).
23. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
24. Sudlow, C. *et al.* UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLoS Med.* (2015). doi:10.1371/journal.pmed.1001779

25. Vigilante, A. *et al.* Identifying Extrinsic versus Intrinsic Drivers of Variation in Cell Behavior in Human iPSC Lines from Healthy Donors. *Cell Rep.* (2019). doi:10.1016/j.celrep.2019.01.094
26. Petersen, R. C. *et al.* Alzheimer's Disease Neuroimaging Initiative (ADNI): Clinical characterization. *Neurology* **74**, 201–209 (2010).
27. A. Bennett, D., A. Schneider, J., Arvanitakis, Z. & S. Wilson, R. Overview and Findings from the Religious Orders Study. *Curr. Alzheimer Res.* (2013). doi:10.2174/156720512801322573
28. Bennett, D. A. *et al.* Overview and findings from the rush Memory and Aging Project. *Curr. Alzheimer Res.* (2012).
29. Bennett, D. A. *et al.* Religious Orders Study and Rush Memory and Aging Project. *Journal of Alzheimer's Disease* (2018). doi:10.3233/JAD-179939
30. Morris, J. C. The clinical dementia rating (cdr): Current version and scoring rules. *Neurology* (1993). doi:10.1212/wnl.43.11.2412-a
31. Zou, F. *et al.* Brain expression genome-wide association study (eGWAS) identifies human disease-associated variants. *PLoS Genet.* (2012). doi:10.1371/journal.pgen.1002707
32. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
33. Purcell, S. M. *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748–52 (2009).
34. Euesden, J., Lewis, C. M. & O'Reilly, P. F. PRSice: Polygenic Risk Score software. *Bioinformatics* (2015). doi:10.1093/bioinformatics/btu848
35. Haldane, J. B. S. The Mean and Variance of | chi 2 , When Used as a Test of Homogeneity, When Expectations are Small. *Biometrika* (1940). doi:10.2307/2332614
36. Corder, E. H. *et al.* Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science (80-.).* (1993). doi:10.1126/science.8346443
37. Leonenko, G. *et al.* Genetic risk for alzheimer disease is distinct from genetic risk for amyloid deposition. *Ann. Neurol.* (2019). doi:10.1002/ana.25530
38. Fry, A. *et al.* Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants with Those of the General Population. *Am. J. Epidemiol.* (2017). doi:10.1093/aje/kwx246
39. So, H. C. & Sham, P. C. A unifying framework for evaluating the predictive power of genetic variants based on the level of heritability explained. *PLoS Genet.* (2010). doi:10.1371/journal.pgen.1001230
40. Sierksma, A., Escott-Price, V. & De Strooper, B. Translating genetic risk of Alzheimer's disease into mechanistic insight and drug targets. *Science* (2020). doi:10.1126/science.abb8575
41. Wilson, P. W. F. *et al.* Apolipoprotein E Alleles, Dyslipidemia, and Coronary Heart Disease: The Framingham Offspring Study. *JAMA J. Am. Med. Assoc.* (1994). doi:10.1001/jama.1994.03520210050031
42. Hofman, A. *et al.* Atherosclerosis, apolipoprotein E, and prevalence of dementia and Alzheimer's disease in the Rotterdam Study. *Lancet* (1997). doi:10.1016/S0140-6736(96)09328-2
43. Xu, M. *et al.* Apolipoprotein e Gene Variants and Risk of Coronary Heart Disease: A Meta-Analysis. *BioMed Research International* (2016). doi:10.1155/2016/3912175
44. El-Lebedy, D., Raslan, H. M. & Mohammed, A. M. Apolipoprotein E gene polymorphism and risk of type 2 diabetes and cardiovascular disease. *Cardiovasc. Diabetol.* (2016). doi:10.1186/s12933-016-0329-1
45. Lumsden, A. L., Mulugeta, A., Zhou, A. & Hyppönen, E. Apolipoprotein E (APOE) genotype-associated disease risks: a phenome-wide, registry-based, case-control study utilising the UK Biobank. *EBioMedicine* (2020).

doi:10.1016/j.ebiom.2020.102954

46. Shinohara, M. *et al.* Apoe2 is associated with longevity independent of alzheimer's disease. *Elife* **9**, 1–16 (2020).
47. Sham, P. C., MacLean, C. J. & Kendler, K. S. A typological model of schizophrenia based on age at onset, sex and familial morbidity. *Acta Psychiatr. Scand.* (1994). doi:10.1111/j.1600-0447.1994.tb01501.x
48. Keren-Shaul, H. *et al.* A Unique Microglia Type Associated with Restricting Development of Alzheimer's Disease. *Cell* (2017). doi:10.1016/j.cell.2017.05.018
49. Sala Frigerio, C. *et al.* The Major Risk Factors for Alzheimer's Disease: Age, Sex, and Genes Modulate the Microglia Response to Aβ Plaques. *Cell Rep.* (2019). doi:10.1016/j.celrep.2019.03.099
50. Krasemann, S. *et al.* The TREM2-APOE Pathway Drives the Transcriptional Phenotype of Dysfunctional Microglia in Neurodegenerative Diseases. *Immunity* (2017). doi:10.1016/j.immuni.2017.08.008
51. Griciuc, A. *et al.* TREM2 Acts Downstream of CD33 in Modulating Microglial Pathology in Alzheimer's Disease. *Neuron* (2019). doi:10.1016/j.neuron.2019.06.010
52. Andreone, B. J. *et al.* Alzheimer's-associated PLCγ2 is a signaling node required for both TREM2 function and the inflammatory response in human microglia. *Nat. Neurosci.* (2020). doi:10.1038/s41593-020-0650-6
53. Lambert, J. C. *et al.* Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* **45**, 1452–1458 (2013).
54. The Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).

Figures

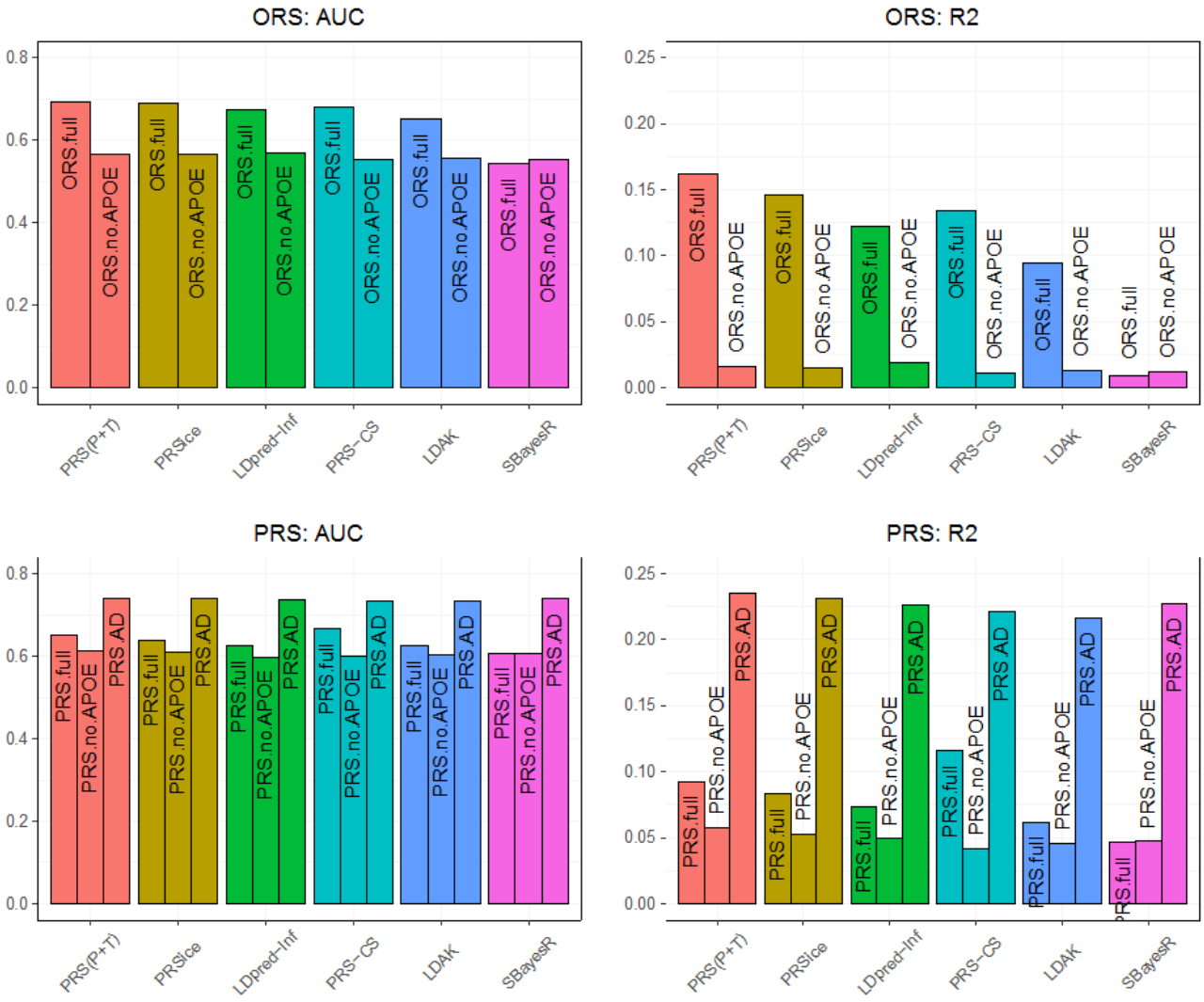

Figure 2

Prediction accuracy across different PRS methods (PRS(P+T), PRSice, LDpred-Inf, PRS-CS, LDAK and SBayesR) for ORS.full, ORS.no.APOE, PRS.full, PRS.no.APOE and PRS.AD. Bar plot for prediction accuracy (AUC and R2) across 6 PRS approaches: PRS(P+T), PRSice, LDpred-Inf, PRS-CS, LDAK and SBayesR. The colour of each PRS method is consistent across all plots. Upper figures represent ORS and lower figures represent PRS and PRS.AD models in the case-control dataset (271 cases and 278 controls). ORS.full includes SNPs with pT £ 1e-5 and PRS.full includes SNPs with pT £ 0.1, ORS.no.APOE and PRS.no.APOE exclude SNPs in the APOE region and PRS.AD models APOE separately and subsequently adds this to PRS.no.APOE.