

Supplementary Text

Multi-trait colocalization and fine-mapping

To fine-map genome-wide association study (GWAS) summary statistics simultaneously across multiple studies, we applied a Bayesian non-parametric clustering algorithm to jointly analyse over 5,000 GWAS at each of 1703 non-overlapping regions of the autosomal genome (excluding the HLA region). The regions were chosen to minimise the linkage disequilibrium between them¹. In each region, the algorithm is based on a model which assumes that each study has either 0 or 1 causal variants. We refer to a set of studies which are assumed to have the same causal variant as a cluster. The algorithm uses Gibbs sampling to iteratively: 1) calculate the posterior distribution on the position of the causal variant for each cluster; 2) update the assignment of each study to either the null cluster (which assumes no causal variants), one of the existing clusters, or to a new cluster. The algorithm is implemented so that the prior distribution on the number and size of the clusters follows that of a Dirichlet Random Measure (also known as a Chinese Restaurant Process)². The prior probability of there being a causal variant for a given study in the region was set to 0.1. The prior probability on each variant being causal variant for a given cluster was uniform on the number of variants in the region. The prior probability that a study belongs to a new cluster was set to 0.001. The algorithm was run for 100 iterations. The output of the algorithm is a posterior probability of cluster membership for each study (cross-study meta-analysis), and the probability that each variant is causal in each of the clusters (cross-study fine-mapping). If the posterior probability of at least one study belonging to that cluster was greater than 0.5, then the variant with the highest posterior probability was selected for inclusion in the set of variants with good evidence for a causal effect on one or more trait.

Combination of CAD PRS and PCE into an Integrated Risk Tool

Our integrated risk tool (IRT) is based on the combined odds of a CAD event in the next 10 years, $Odds_{PCE + PRS}$, taking into account information from both standard risk calculators and PRS. This was found by converting the PRS for each individual into an odds ratio, $OR_{PRS|PCE}$, and multiplying this with the odds, $Odds_{PCE}$, under the PCE-only model:

$$Odds_{PCE + PRS} = Odds_{PCE} * OR_{PRS|PCE} \quad .$$

To capture statistical interaction, the odds ratio term $OR_{PRS|PCE}$ was estimated from a logistic regression model that included a PRS*PCE interaction term, and which was trained on incident CAD outcomes in the Group II dataset (where PRS is standardised to mean 0 and variance 1, and PCE is the logit-transform of estimated 10-year risk):

$$\text{logit}(CAD) \sim \alpha + \beta_1 PRS + \beta_2 PCE + \beta_3 PRS * PCE \quad .$$

The $OR_{PRS|PCE}$ term was then estimated in Group III as:

$$OR_{PRS|PCE} = e^{\beta_1*PRS+\beta_3*PRS*PCE} .$$

The training procedure was performed separately on men and women (following the separate PCE calculations applied to men and women).

Supplementary Tables

Supplementary Table S1: Definition of variables used in two standard risk calculators

PCE and QRISK3 variables	
Total cholesterol	UKB Field: 30690
HDL cholesterol	UKB Field: 30760
Systolic blood pressure	UKB Field: 4080
Diabetes	Type 1 diabetes was defined following a strict definition previously used in Sharp et al (2019) ³ . Type 2 diabetes was defined as probable T2D following decision tree in Eastwood et al (2016) ⁴ .
Smoking status	Smoking status was assessed using UKB Field: 90 combined with UKB field: 3456. Current smokers with no info on field 3456 were set as missing.
Hypertension medication	UKB Fields: 6177, 6153 and 20003
Additional QRISK3 variables	
Townsend index	UKB Field: 189
Smoking status	Smoking status was assessed using UKB Field: 90 combined with UKB field: 3456. Current smokers with no values on field 3456 were set as missing. Former and never smokers with values on field 3456 were set as missing.
Chronic kidney disease	Biochemical measurements of urine microalbumin and creatinine (UKB Field: 30500 and 30510) were used to estimate UACR. Serum creatinine and cystatin C (UKB Fields: 30700 and 30720) were used to estimate eGFR measurements. The UACR and eGFR estimates were used to define CKD stage3+ patients. We also used ICD10 codes N180,N183-185.
Family history*	UKB Fields: 20107, 20110, 20111.
SD of systolic blood pressure	We used the previously reported averages of 9.3 for women and 9.9 for men ⁵ , as this data is not present in UKB.
Atrial fibrillation	Self-reported atrial fibrillation (UKB Field: 20002, code 1471). ICD10 codes under I48. Set to missing when date of diagnosis unclear.

BMI	UKB Field: 23104
Systemic lupus erythematosus	Self-reported lupus (UKB Field: 20002, code 1381). ICD10 codes under M32. Set to missing when date of diagnosis unclear.
Migraine	Self-reported migraine (UKB Field: 20002, code 1265). ICD10 codes under G43. Set to missing when date of diagnosis unclear.
Rheumatoid arthritis	Self-reported migraine (UKB Field: 20002, code 1464). ICD10 codes under M05,M06,M08. Set to missing when date of diagnosis unclear.
Severe mental illness	Self-reported severe mental illness (UKB Field: 20002, codes 1289,1291,1286 and UKB Field: 20126, codes 1,2,3,4). ICD10 codes under F20,F31, F331, F332, F333. Set to missing when date of diagnosis unclear.
Diagnosis of or treatment for erectile dysfunction (ED)	Self-reported ED (UKB Field: 20002, code 1518). ED medication under UKB Field: 20003. Set to missing when date of diagnosis unclear.
Atypical antipsychotic medication	UKB Fields: 20003
Regular steroid tablets	UKB Fields: 20003

*We note that there is no age of onset for 1st degree relatives in UKB and therefore we might be overestimating the risk for a proportion of individuals with family history where the age of onset was >60.

Supplementary Table S2: Follow-up time quartiles per Group and stratified by age and sex.

Population	Q1/Median/Q3 follow-up time			
	Group I	Group II	Group III	Group IV
Overall	7.47/8.19/8.96	7.61/8.28/8.85	7.57/8.21/8.81	7.51/8.18/8.82
Men 40-54	7.5/8.32/8.99	7.65/8.33/8.9	7.6/8.28/8.87	7.62/8.3/8.9
Women 40-54	7.47/8.17/8.95	7.67/8.32/8.88	7.62/8.3/8.85	7.68/8.37/8.95
Men 55-69	7.47/8.22/8.97	7.56/8.2/8.84	7.51/8.18/8.78	7.47/8.13/8.78
Women 55-69	7.44/8.14/8.94	7.58/8.21/8.81	7.55/8.19/8.79	7.52/8.2/8.82

Supplementary Table S3: Summary of prediction metrics for incident CAD, comparing QRISK3 only and QRISK3+PRS. We compared models using Harrell's C and the net reclassification improvement (NRI, based on a ten-year high/low risk threshold of 10% for estimated cardiovascular disease risk).

category	Harrell's C (95% CI)			NRI		
	QRISK3 only	Including PRS	Difference	Combined	Cases	Noncases
Overall	0.78 (0.78-0.79)	0.81 (0.80-0.81)	0.02 (0.01-0.03)	2.81 (1.65-3.96)	2.28 (1.14-3.42)	0.52 (0.36-0.69)
Men (40-54yo)	0.78 (0.76-0.79)	0.81 (0.79-0.83)	0.04 (0.01-0.06)	8.70 (4.59-12.82)	18.22 (14.12-22.31)	-9.52 (-9.95--9.08)
Women (40-54yo)	0.76 (0.73-0.80)	0.77 (0.74-0.81)	0.01 (-0.04-0.06)	9.70 (4.12-15.28)	11.88 (6.30-17.45)	-2.18 (-2.34--2.02)
Men (55-69yo)	0.67 (0.66-0.68)	0.71 (0.69-0.72)	0.04 (0.02-0.05)	8.03 (6.84-9.22)	-3.33 (-4.46--2.20)	11.35 (10.99-11.72)
Women (55-69yo)	0.72 (0.70-0.74)	0.74 (0.72-0.76)	0.03 (0.00-0.05)	2.67 (-0.19-5.54)	2.56 (-0.28-5.40)	0.11 (-0.22-0.44)

Supplementary Table S4: Summary of prediction metrics for incident CVD, comparing PCE only and PCE+PRS. We compared models using Harrell's C and the net reclassification improvement (NRI, based on a ten-year high/low risk threshold of 7.5% for estimated cardiovascular disease risk).

category	Harrell's C (95% CI)			NRI		
	PCE only	Including PRS	Difference	Combined	Cases	Noncases
Overall	0.73 (0.72-0.74)	0.74 (0.74-0.75)	0.01 (0.00-0.02)	1.92 (1.27-2.57)	2.23 (1.59-2.87)	-0.31 (-0.42--0.19)
Men (40-54yo)	0.71 (0.69-0.73)	0.73 (0.72-0.75)	0.03 (0.00-0.05)	8.60 (6.00-11.20)	13.50 (10.92-16.08)	-4.90 (-5.23--4.57)
Women (40-54yo)	0.70 (0.67-0.72)	0.70 (0.68-0.72)	0.00 (-0.03-0.04)	1.96 (0.38-3.54)	2.42 (0.84-4.00)	-0.46 (-0.53--0.39)
Men (55-69yo)	0.63 (0.62-0.64)	0.66 (0.65-0.66)	0.02 (0.01-0.04)	3.30 (2.56-4.04)	-1.71 (-2.40--1.03)	5.01 (4.73-5.29)
Women (55-69yo)	0.65 (0.64-0.66)	0.66 (0.65-0.68)	0.01 (0.00-0.03)	2.53 (1.12-3.94)	3.90 (2.51-5.29)	-1.37 (-1.60--1.14)

Supplementary Table S5: Summary of prediction metrics for incident CVD, comparing QRISK3 only and QRISK3+PRS. We compared models using Harrell's C and the net reclassification improvement (NRI, based on a ten-year high/low risk threshold of 10% for estimated cardiovascular disease risk).

category	Harrell's C (95% CI)			NRI		
	QRISK3 only	Including PRS	Difference	Combined	Cases	Noncases
Overall	0.75 (0.75-0.76)	0.76 (0.75-0.76)	0.01 (0.00-0.02)	0.66 (0.05-1.27)	0.65 (0.05-1.25)	0.01 (-0.12-0.14)
Men (40-54yo)	0.74 (0.73-0.76)	0.76 (0.75-0.78)	0.02 (0.00-0.04)	4.78 (2.40-7.16)	8.93 (6.57-11.29)	-4.15 (-4.49--3.82)
Women (40-54yo)	0.72 (0.70-0.74)	0.72 (0.70-0.74)	0.00 (-0.03-0.03)	1.84 (0.10-3.58)	2.79 (1.06-4.53)	-0.95 (-1.07--0.84)
Men (55-69yo)	0.66 (0.65-0.67)	0.68 (0.67-0.69)	0.02 (0.00-0.03)	3.90 (3.18-4.62)	-1.53 (-2.19--0.87)	5.43 (5.14-5.71)
Women (55-69yo)	0.69 (0.68-0.70)	0.69 (0.68-0.70)	0.01 (-0.01-0.02)	-0.78 (-2.10-0.55)	0.00 (-1.30-1.30)	-0.78 (-1.03--0.53)

Supplementary Table S6: Prediction performance metrics (with 95% CI) for incident CAD outcomes in Group III, comparing PCE and IRT (PCE+PRS) models and stratifying into age-by-sex subgroups. Cases were right-censored at the 10 year mark instead of 7 year mark, in both the training and test sets.

category	Harrell's C			NRI (PCE vs IRT)		
	PCE	IRT	Difference	Full	Within-cases	Within-noncases
Overall	0.77 (0.76-0.78)	0.80 (0.79-0.81)	0.03 (0.02-0.04)	5.56 (4.38-6.74)	5.89 (4.72-7.05)	-0.33 (-0.49--0.17)
Men (40-54yo)	0.75 (0.73-0.77)	0.80 (0.78-0.82)	0.05 (0.02-0.075)	15.76 (11.86-19.67)	25.72 (21.83-29.60)	-9.96 (-10.37--9.54)
Women (40-54yo)	0.75 (0.72-0.79)	0.76 (0.73-0.80)	0.01 (-0.04-0.07)	8.72 (4.38-13.06)	9.90 (5.56-14.24)	-1.18 (-1.29--1.07)
Men (55-69yo)	0.65 (0.63-0.66)	0.69 (0.68-0.70)	0.05 (0.03-0.06)	7.73 (6.64-8.82)	-2.59 (-3.63--1.56)	10.32 (9.98-10.67)
Women (55-69yo)	0.68 (0.66-0.70)	0.72 (0.70-0.74)	0.04 (0.01-0.07)	9.77 (6.51-13.03)	12.08 (8.83-15.32)	-2.31 (-2.63--1.98)

Supplementary Table S7: Sensitivity, specificity and positive predictive values of PCE compared to IRT for prediction of CAD events in Group III, both overall and stratifying into age-by-sex subgroups, using 7.5% as the high/low 10-year risk threshold. Note that if we assume 10-year risk can take any value between 0 – 100%, then even a perfect test is not expected to have sensitivity, specificity or positive predictive values of 100%, as the threshold for high/low risk classification is itself an estimate of the probability of an event.

category	Sensitivity (%)		Specificity (%)		Positive predictive value (%)	
	PCE	IRT	PCE	IRT	PCE	IRT
Overall	76.28	82.27	63.93	63.62	3.09	3.29
Men (40-54yo)	46.34	74.65	82.45	71.39	3.30	3.26
Women (40-54yo)	9.38	18.12	99.16	98.11	3.38	2.92
Men (55-69yo)	97.82	95.16	8.92	19.76	3.63	3.99
Women (55-69yo)	55.42	67.17	69.77	67.70	1.78	2.01

References

- 1 Berisa T, Pickrell JK. Approximately independent linkage disequilibrium blocks in human populations. *Bioinformatics* 2016; **32**: 283–5.
- 2 Aldous DJ. Exchangeability and related topics. 1985: 1–198.
- 3 Sharp SA, Rich SS, Wood AR, *et al.* Development and standardization of an improved type 1 diabetes genetic risk score for use in newborn screening and incident diagnosis. In: *Diabetes Care*. 2019. DOI:10.2337/dc18-1785.
- 4 Eastwood S V, Mathur R, Atkinson M, *et al.* Algorithms for the capture and adjudication of prevalent and incident diabetes in UK Biobank. *PLoS One* 2016; **11**. DOI:10.1371/journal.pone.0162388.
- 5 Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. *Br Med J* 2017; **357**: j2099.