university of groningen

University Medical Center Groningen

University of Groningen

Polygenic risk score and coronary artery disease

Agbaedeng, Thomas A; Noubiap, Jean Jacques; Mofo Mato, Edith Pascale; Chew, Derek P; Figtree, Gemma A; Said, M Abdullah; van der Harst, Pim

Published in:
ATHEROSCLEROSIS

DOI:
10.1016/j.atherosclerosis.2021.08.020

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2021

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Agbaedeng, T. A., Noubiap, J. J., Mofo Mato, E. P., Chew, D. P., Figtree, G. A., Said, M. A., & van der Harst, P. (2021). Polygenic risk score and coronary artery disease: A meta-analysis of 979,286 participant data. *ATHEROSCLEROSIS*, *333*, 48-55. https://doi.org/10.1016/j.atherosclerosis.2021.08.020

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 05-06-2022

Atherosclerosis 333 (2021) 48–55

Contents lists available at ScienceDirect

Atherosclerosis

journal homepage: www.elsevier.com/locate/atherosclerosis

Polygenic risk score and coronary artery disease: A meta-analysis of 979,286 participant data

Thomas A. Agbaedeng [a,b,*], Jean Jacques Noubiap [a], Edith Pascale Mofo Mato [c], Derek P. Chew [d,e], Gemma A. Figtree [f], M. Abdullah Said [g], Pim van der Harst [h]

[a] Centre for Heart Rhythm Disorders, The University of Adelaide, Adelaide, Australia
[b] Centre for Cardiometabolic Experimentation and Innovation, Adelaide, Australia
[c] Molecular and Clinical Pharmacology Research Laboratory, University of Kwazulu Natal, South Africa
[d] College of Medicine and Public Health, Flinders University, Australia
[e] Heart Health Theme, South Australian Health and Medical Research Institute, Australia
[f] Kolling Institute and Charles Perkins Centre, The University of Sydney, Australia
[g] University of Groningen, University Medical Center Groningen, Department of Cardiology, 9700, RB, Groningen, the Netherlands
[h] Department of Cardiology, Heart and Lung Division, University Medical Center Utrecht, Utrecht, the Netherlands

ARTICLE INFO

Keywords:
Polygenic risk score
Coronary artery disease
Myocardial infarction
Genome-wide association study
Single-nucleotide polymorphism

ABSTRACT

Background and aims: Coronary artery disease (CAD) is a complex disease with a strong genetic basis. While previous studies have combined common single-nucleotide polymorphisms (SNPs) into a polygenic risk score (PRS) to predict CAD risk, this association is poorly characterised. We performed a meta-analysis to estimate the effect of PRS on the risk of CAD.

Methods: Online databases were searched for studies reporting PRS and CAD. PRS computation was based on log-odds (PRS_{LN}), pruning or clumping and thresholding ($PRS_{P/C+T}$), Lassosum regression ($PRS_{Lassosum}$), LDpred (PRS_{LDpred}), or metaGRS ($PRS_{metaGRS}$). The reported odds ratio (OR), hazard ratio (HR), C-indexes and their corresponding 95% confidence interval (95% CI) were pooled in a random-effects meta-analysis.

Results: Forty-nine studies were included (979,286 individuals). There was a significant association between 1-standard deviation [SD] increment in PRS and adjusted risks of both incident and prevalent CAD (OR [95% CI]: 1.67 [1.57–1.77] for $PRS_{metaGRS}$, 1.46 [1.26–1.68] for PRS_{LDpred}). The risk of incident CAD was highest for $PRS_{P/C+T}$ (HR [95% CI]: 1.49 [1.26–1.78]), $PRS_{metaGRS}$ (1.37 [1.27–1.47]), and PRS_{LDpred} (1.36 [1.31–1.42]). Analysis of model performance demonstrated that PRS predicted incident CAD with C-index of up to 0.71. Importantly, addition of PRS to clinical risk scores resulted in modest but statistically significant improvements in CAD risk prediction, with 1.5% observed for $PRS_{P/C+T}$ ($p < 0.001$) and 1.6% for PRS_{LDpred} ($p < 0.001$).

Conclusions: Polygenic risk score is strongly associated with increased risks of CAD. Future prospective studies should explore the usefulness of polygenic risk scores for identifying individuals at a high risk of developing CAD.

1. Introduction

Coronary artery disease (CAD) remains the leading global cause of multi-morbidity and mortality [1,2]. Thus, strategies for accurate identification of individuals at high-risk of developing CAD are paramount for individualized primary prevention. Current clinical risk scores, incorporating several traditional risk factors, have been shown to predict CAD risk. However, the increased susceptibility to CAD in younger and healthier patients underscores a contribution of factors beyond the traditional risk factors. Indeed, data from large prospective cohorts [3] and registries [4] have demonstrated that parental history of premature CAD is an independent predictor of future CAD in offspring. Consequently, familial aggregation of CAD risk is reported to explain up to 60% variation in the heritability of CAD [5].

The genetic basis of CAD can roughly be divided into its monogenic and polygenic components. The monogenic component, namely, high-impact genetic mutations that follow a classical Mendelian pattern, is best exemplified by as familial hypercholesterolemia (FH) [6–8]. While important for carriers and their families, FH has low prevalence in the general population (~1 in 200–400), and thus can only explain a small

* Corresponding author. School of Medicine, The University of Adelaide, Level 8 North, SAHMRI, 2 North Terrace, Adelaide, SA, 5000, Australia.
E-mail address: thomas.agbaedeng@outlook.com (T.A. Agbaedeng).

https://doi.org/10.1016/j.atherosclerosis.2021.08.020
Received 5 June 2021; Received in revised form 4 August 2021; Accepted 11 August 2021
Available online 12 August 2021
0021-9150/© 2021 Elsevier B.V. All rights reserved.

proportion of CAD risk in the population [8–10]. A far larger proportion of CAD risk is explained by genetic susceptibility conferred by polygenic risk: the aggregate contributions of many common (minor allele frequency (MAF) ≥1% or so), low-impact genetic variants [11]. Over 1,790 of such CAD loci have been identified to date by large-scale genome--wide association studies (GWAS) [12,13]. However, the use of genetic information for CAD risk stratification remains challenging. Recent data suggests that aggregation of these common variants into a risk score, termed genetic risk score or polygenic risk score (PRS), can improve CAD risk prediction and stratification [14,15].

Herein, we conducted a systematic review and meta-analysis to provide a comprehensive evaluation of available data from validation and replication studies associating PRS of GWAS-derived SNPs and CAD risk. Our objectives were to provide a comparative assessment of the associations between PRS and CAD risks. We also aimed to evaluate the various methodologies for deriving PRS and their impact on PRS and CAD relationship.

2. Materials and methods

This meta-analysis was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines and the HuGENet™ HuGE Review Handbook for systematic reviews and meta-analyses of genetic association studies. The protocol was registered on PROSPERO (ID: CRD42020190305).

2.1. Search strategy

Relevant studies were identified by database searches performed on PubMed, EMBASE, Web of Science Core Collection, and Ovid MEDLINE from inception through 12 December 2020, using a combination of terms related to: genetic risk score, polygenic risk score, genome-wide association study, and coronary artery disease. The full search strategies are available in the supplement (Supplementary Method 1). The reference list of eligible articles and relevant reviews were also scrutinized to identify additional sources of information.

2.2. Study selection criteria

Initial screening of retrieved references was conducted by two investigators (TAA and JJN) independently, at the titles and abstracts stages. Then, the full texts of all potentially eligible articles were obtained and further assessed for final inclusion. Disagreements regarding study inclusion were resolved by consensus. Studies were included if they fulfilled the following criteria: (1), were published in English; (2), participants were aged 18 years or older; (3), reported on CAD as pre-existing or endpoint outcomes; (4), calculated PRS containing SNPs from ≥2 loci. Studies were excluded, if they: (1), reported only monogenic associations; (2), were of critically poor design; (3), had limited power (sample size ≤100 individuals); (3) twin or linkage studies; (4), were abstracts not yet published as full manuscripts, reviews, editorials, case reports, and letters to the editor.

2.3. Data extraction and outcomes

Data extraction was done independently by two investigators (TAA and EPMM), using a priori data abstraction form. Disagreements were resolved by consensus. The following data were extracted: first author's name, publication year, recruitment period, country of primary cohort, study design, number of participants, mean or median age, proportion of males, CAD risk factors, follow-up time, incidence/prevalence/recurrence of any CAD, genotype information, number of SNPs and modality for PRS computation, risk estimates (odds ratio [OR] or hazard ratio [HR]), and their 95% confidence intervals (95% CI). For details on PRS calculation, see Supplementary Method 2. Briefly, PRS was presented as a simple count, additive sum (uwPRS), or weighted score. Weighted methods were by effect size (natural log odds, PRS_{LN}), PRS_{LN} rescaled by number of SNPs ($PRS_{LN\text{-}rescaled}$), linkage disequilibrium (LD) pruning/clumping and/or thresholding ($PRS_{P/C\ +\ T}$), genome-wide LD clumping (metaGRS: $PRS_{metaGRS}$), penalised regression (i.e., Lassosum: $PRS_{Lasso\text{-}sum}$), and the Bayesian "Linkage Disequilibrium prediction" (LDpred: PRS_{LDpred}) models.

Composite endpoints were defined as the diagnosis of fatal or nonfatal MI, coronary interventions, ischaemic stroke, and/or cardiovascular death. CAD was defined as any diagnosis of myocardial infarction (MI), coronary intervention, and/or CAD/MI-related death. CAD was further defined as early onset (EOCAD), if it occurred before 55 years of age; late onset (LOCAD), if it occurred >65 years of age; or any type (AnyCAD), if it occurred at any age.

Details on the clinical risk scores (CRS) used in the current analyses are described and reported in Supplementary Methods 3 and Supplementary Table S1.

2.4. Risk of bias assessment

Risk of bias was assessed using the *Quality of Genetic Studies* (Q-Genie) tool. Studies reporting on prediction models of PRS for CAD were further assessed for potential bias in this domain and applicability by the Prediction model Risk Of Bias ASsessment Tool (PROBAST). Disagreements were resolved by consensus.

2.5. Data analysis

Analyses were based on multivariate estimates. We performed random-effects meta-analysis using the inverse method (DerSimonian-Laird estimator), expressing them as OR or HR and 95% CI. Where CAD risk was presented per categories of PRS, these were converted to risk per standard deviation in PRS (provided there were ≥3 categories), with covariance approximated using the method of Greenland and Longnecker [16]. The degree of heterogeneity in the comparisons was assessed by examination of forest plots and by the Higgins' I [2]-statistic, and small-study effect by funnel plots. We also pooled area under the receiver operating characteristic curve (AUC) or C-index and corresponding standard error to evaluate the discriminative ability of PRS models, and improvement of clinical risk scores (CRS) in a combined PRS plus CRS (PRS + CRS) model. Data were pooled separately for each PRS algorithm. We identified statistical significance using a two-*tailed p* value of ≤0.05. Analyses were performed in R version 3.6.2 (R Core Team for Statistical Computing, Vienna, Austria).

3. Results

3.1. Synthesis of literature search and study characteristics

A total of 1,679 articles were retrieved from electronic databases and supplementary searches. Of these, 214 were available for full text review, of which 49 studies were included (Fig. 1). Included studies contributed to a total of 977,716 participants from 110 cohorts. Participants were recruited between 1971 and 2016, mostly from Europe (53.1%) and North America (24.5%), and studies were published between 2007 and 2021 (Supplementary Tables S2–S5). Overall, the quality of the included studies was good. Using the Q-Genie tool, most studies (n = 30) were rated at low risk of bias (32.7%) or moderate (53.1%) risk of bias (Supplementary Table S6).

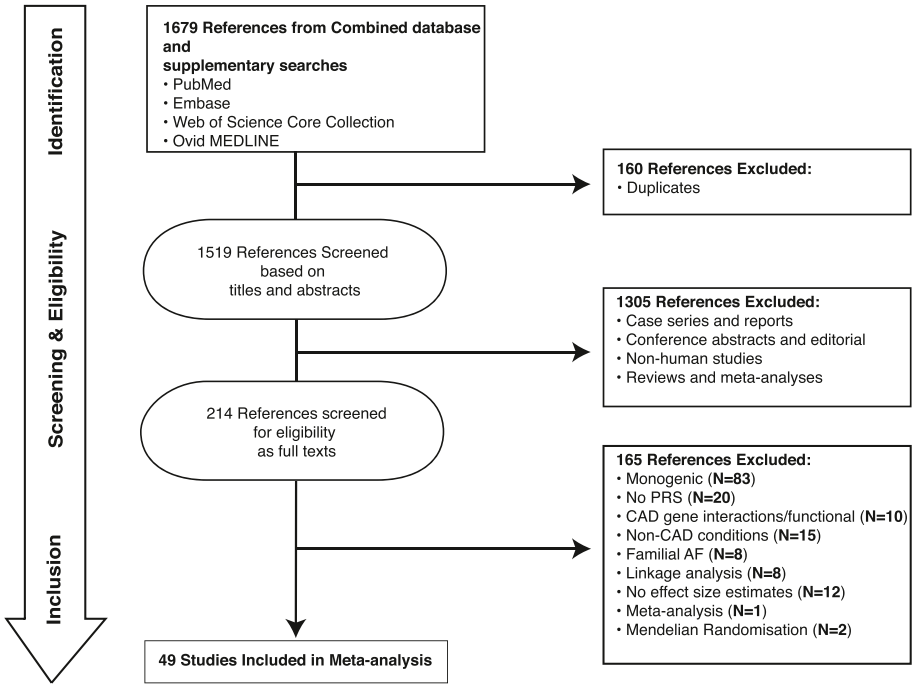

Fig. 1. PRISMA flowchart of literature search strategy.

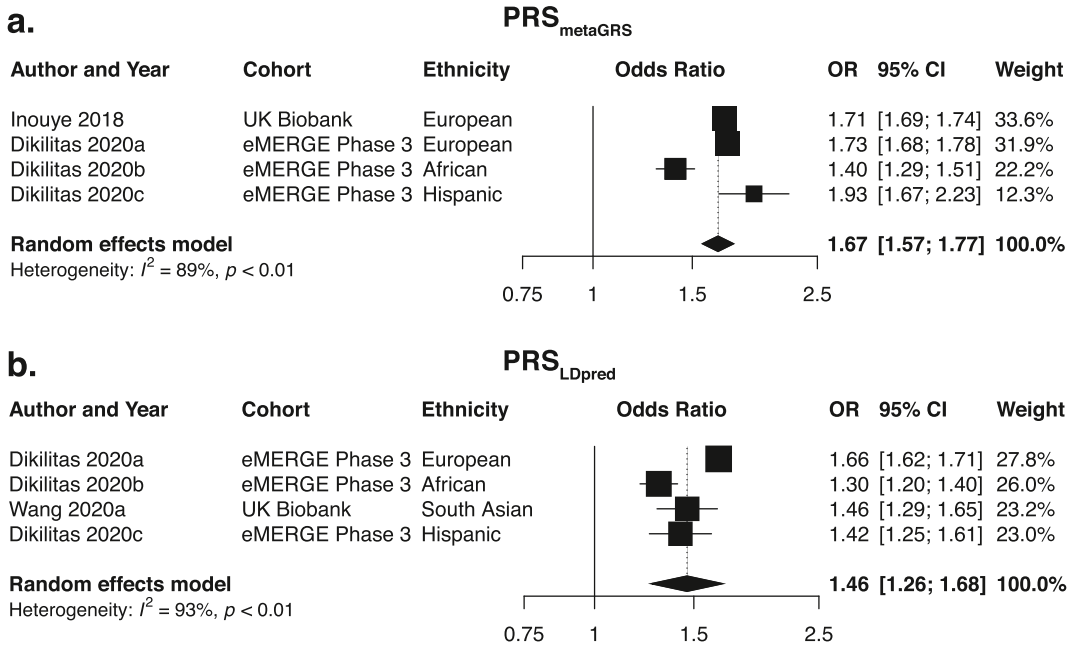

Fig. 2. Association between PRS and all coronary artery disease (CAD).
Forest plot showing pooled risk estimate associating PRS with the risk of incident and prevalent CAD. The summarised results were pooled using unadjusted estimates from the included studies. Data were pooled separately for each PRS algorithm and summarised on the same figure, without combining them. OR, odds ratio; 95% CI, 95% confidence interval.

3.2. Meta-analysis: PRS and CAD

Results of the association between PRS and risks of combined incident and prevalent CAD are presented in Fig. 2. Data were available for $PRS_{metaGRS}$ and PRS_{LDpred} and were pooled in a meta-analysis. The best association was seen with $PRS_{metaGRS}$, at 67% higher odds of any CAD (OR: 1.67 [1.57–1.77]). PRS_{LDpred} was also highly associated with CAD (OR: 1.46 [1.26–1.68]).

Results of the associations for incident CAD risks are shown in Fig. 3, Table 1, and Supplementary Table S8. PRS significantly predicted CAD incidence, ranging from 8% to 49% risks. The best result was noted for $PRS_{P/C + T}$ (HR: 1.49 [1.26–1.78]), though the analysis had high heterogeneity (I [2] 95%). $PRS_{metaGRS}$ and PRS_{LDpred} were also highly predictive of incident CAD (HR: 1.37 [1.27–1.47]; HR: 1.36 [1.31–1.42]). Compared to unweighted PRS based on SNPs previously associated CAD, weighted PRS had 10%–41% higher risks of CAD. We performed analysis

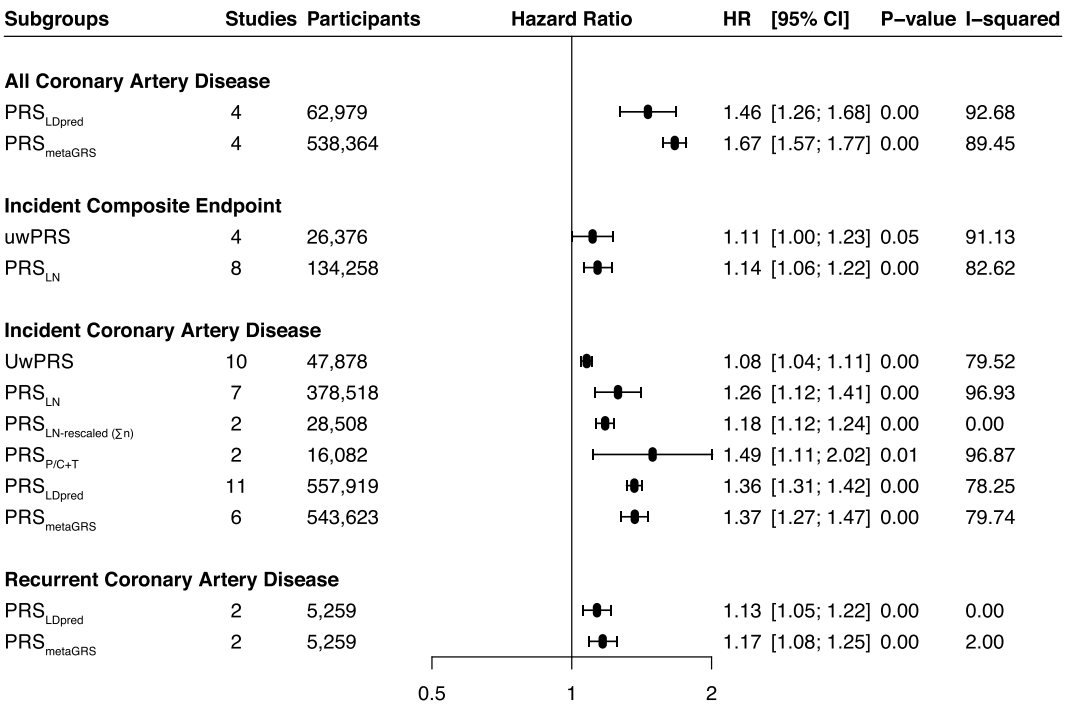

Fig. 3. Association between weighted PRS and incident CAD.
Forest plot showing pooled risk estimate associating 1-SD increment in PRS with the risk of CAD using estimates from fully adjusted models. Estimates are grouped by the weighting of GRS. CAD, coronary artery disease; PRS, genetic risk score; PRS_{LN}, PRS weighted by natural log-odds of the effect size; PRS_{LDpred}, PRS computed by LDpred; $PRS_{metaGRS}$, PRS computed by metaGRS.

Table 1
Summary association of weighted polygenic risk score and incident CAD.

Subgroup	Studies (N)	Participants	SNP (Range)	Associated risk		Heterogeneity		Egger's test (*p*-value)
				HR	95% CI (LL–UL)	I^2	*p*-value	
PRS_{LN}								
Overall	7	378,518	9–858	1.26	1.12–1.41	96.9%	<.001	0.518
European GWAS	4	42,744	9–858	1.28	1.12–1.45	93.9%	<.001	0.265
Trans-ethnic GWAS	2	330,875	30–169	1.24	0.90–1.72	96.8%	<.001	ND
Cohort study design	6	376,679	13–858	1.27	1.11–1.44	97.4%	<.001	0.542
$PRS_{LN\text{-}rescaled\ (\sum n)}$								
Overall	2	28,508	70–257	1.18	1.12–1.24	0.0%	1.000	ND
PRS_{LDpred}								
Overall	11	554,919	5.6–6.6 million	1.36	1.30–1.42	78.3%	<.001	0.022
European GWAS	3	55,735	6.6 million	1.27	1.05–1.53	88.9%	<.001	0.218
Trans-ethnic GWAS	5	148,134	5.6–6.6 million	1.29	1.17–1.43	82.2%	<.001	0.069
European testing cohorts	8	857,745	5.6–6.6 million	1.36	1.30–1.43	78.5%	<.001	0.022
Non-European testing cohorts	3	53,733	6.6 million	1.37	1.16–1.62	85.0%	0.001	0.666
PRSmetaGRS								
Overall	6	543,623	1.7 million	1.37	1.27–1.47	79.7%	<.001	0.066
European testing cohorts	4	533,533	1.7 million	1.39	1.29–1.50	82.4%	<.001	0.100
Non-European testing cohorts	2	10,090	1.7 million	1.34	1.13–1.59	52.5%	0.147	ND

CAD, coronary artery disease; PRS, polygenic risk score; PRS_{LN}, polygenic risk score weighted by natural log-odds of the effect size; PRS_{LDpred}, polygenic risk score computed by LDpred; $PRS_{metaGRS}$, polygenic risk score computed by genome-wide linkage disequilibrium pruning; GWAS, genome-wide association study; HR, hazard ratio; ND, not determined; CI: confidence interval; RCT, randomised clinical trial; SNP, single-nucleotide polymorphism.

by the ancestry of the primary GWAS of the SNPs used for deriving PRS; PRS_{LN} from European GWAS predicted a 28% higher risk of incident CAD compared 24% of *trans*-ethnic GWAS. However, for PRS_{LDpred}, PRS derived from either European or *trans*-ethnic lower than the overall pooled analysis but was still significant. When we evaluated the influence of the testing cohort ancestry, no difference was noted for PRS_{LDpred}. Compared to non-European cohorts, $PRS_{metaGRS}$ tested in European cohorts demonstrated a 5% higher incident CAD risk.

We tested PRS as prognostic marker in patients who have had CAD in past and present the results in Table 1 and Supplementary S2. Both $PRS_{metaGRS}$ and PRS_{LDpred} significantly predicted elevated risks of recurrent coronary events. However, the predicted risks were much less (HR: 1.17 [1.08–1.25] for $PRS_{metaGRS}$; HR: 1.13 [1.05–1.22] for PRS_{LDpred}) than for incident events in previously undiagnosed individuals.

We tested PRS as diagnostic marker in case-control datasets and present the results in Supplementary Tables S7–S8 and Supplementary S3. PRS was significantly associated with prevalent CAD, ranging from 24% to 69% odds of CAD presence. The best result was noted for $PRS_{metaGRS}$ (OR: 1.69 [1.58–1.81]). In subgroup analyses, PRS were more strongly associated with CAD if they were computed using SNPs from European and *trans*-ethnic GWAS and tested in European and younger populations. Moreover, PRS demonstrated the strongest

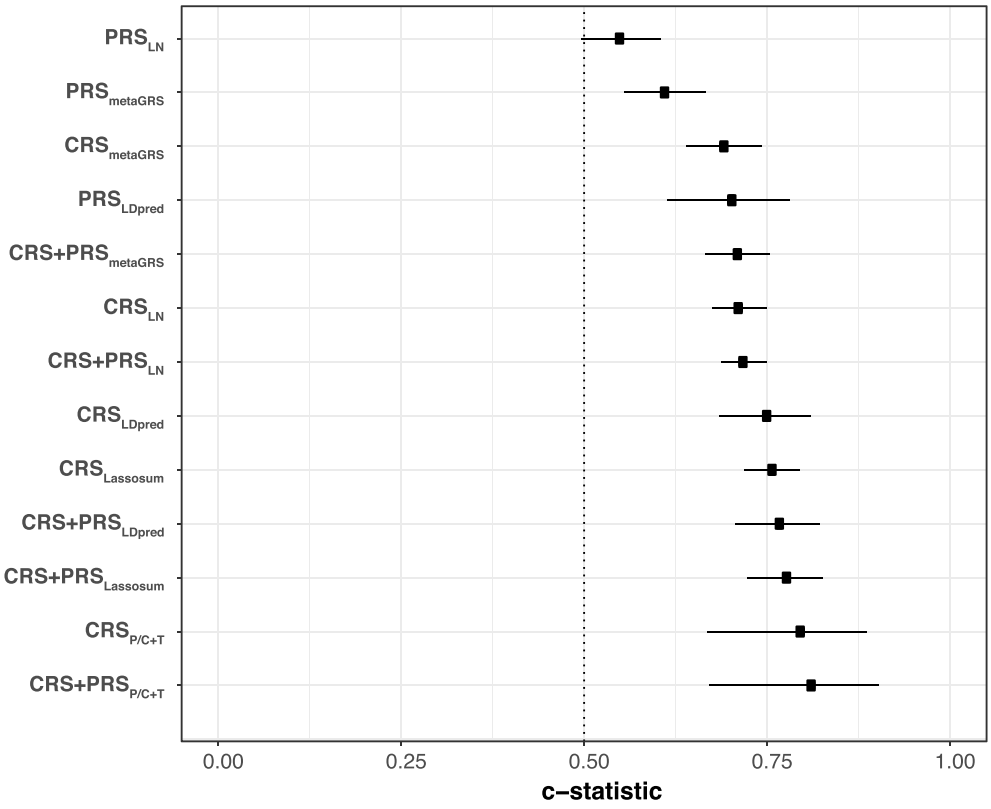

Fig. 4. Summarised performances of CAD risk prediction models.

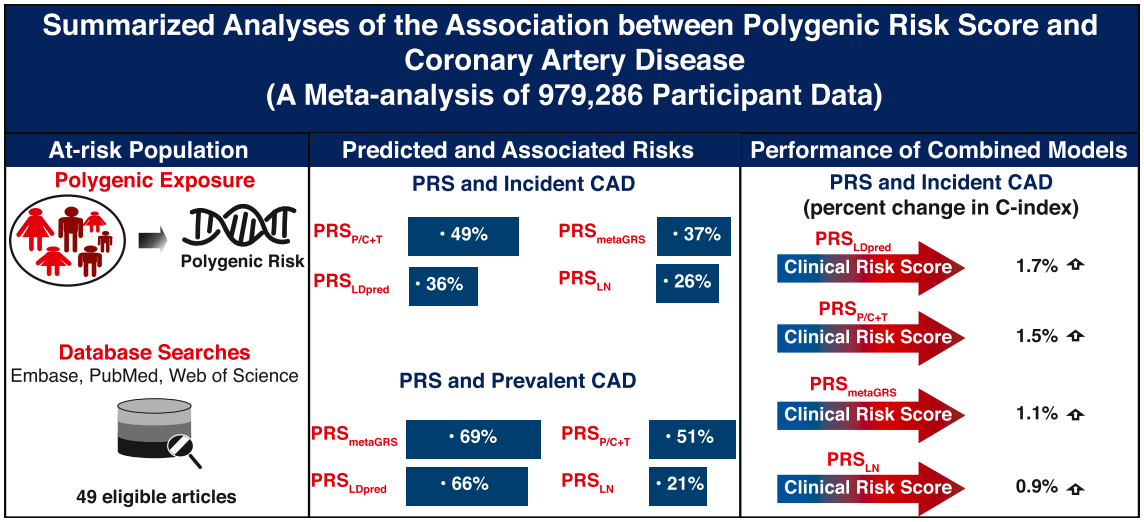

Fig. 5. Summarised analyses of the association between polygenic risk score and coronary artery disease.

association with EOCAD compared to AnyCAD or LOCAD (PRS$_{LN}$ p < 0.001, uwPRS p 0.006).

3.3. Performance of genetic prediction models for CAD

Results of the performance of GRS-based prediction models are presented in Table 2, Fig. 4, and Supplementary Figs. S5–S10. PRS predicted incident CAD with AUC ranging from 0.55 to 0.71, with the highest C-index seen for PRS$_{LDpred}$. Almost all PRS modestly but significantly improved prediction models for incident CAD when included in the models containing clinical risk scores (CRS). The pooled improvements in the combined PRS plus CRS models were: 0.9% for PRS$_{LN}$ (p < 0.001), 1.5% for PRS$_{P/C+T}$ (p < 0.001), 1.6% for PRS$_{LDpred}$ (p < 0.001), 1.1% for and PRS$_{metaGRS}$ (p 0.005), respectively, Supplementary Tables S9–S10. For specific CRS, we noted the biggest improvement of 1.7% for PRS$_{LDpred}$ added to ACC/AHA pooled cohort equation (PCE, p < 0.001).

Results of the reclassification of CAD risk by PRS are presented in Supplementary Table S11. As shown, the pooled NRI was significant for PRS$_{LN}$ and PRS$_{P/C+T}$, with the greatest NRI of 10.3% seen with PRS$_{P/C+T}$ compared to PCE (p < 0.001).

Table 2
Summary of C-indices of weighted polygenic risk score and incident CAD.

	PRS_{LN}							$PRS_{P/C+T}$	
	PRS	CRS			PRS + CRS			CRS	PRS + CRS
		Any	FRS	ACRS	Any	FRS	ACRS		
Pooled studies (N)	2	7	3	2	7	3	2	2	2
Participants (N)	358,402	104,932	58,231	10,881	104,932	58,231	10,881	32,164	32,164
C-index (95% CI)	0.551 (0.496–0.605)	0.713 (0.675–0.749)	0.710 (0.622–0.785)	0.739 (0.664–0.802)	0.720 (0.688–0.750)	0.719 (0.631–0.793)	0.747 (0.653–0.823)	0.798 (0.668–0.886)	0.813 (0.671–0.902)

ACRS, Atherosclerosis Risk in Community (ARIC) coronary heart disease (CHD) risk score; CRS, clinical risk score; FRS, Framingham Risk Score; PCE, pooled cohort equation. For other abbreviations, see Table 1.

3.4. Meta-analysis: PRS and composite endpoints

Results for studies reporting a composite endpoint of incident CAD, stroke, or CV death are shown in Supplementary Fig. S4. PRS_{LN} was significantly associated with a 14% increased risk (HR 1.14, $p < 0.01$) of incident composite outcomes compared to 11% by uwPRS. In the two studies reporting composite endpoints as prevalence of CAD, stroke, and/or CV death, the pooled adjusted estimate was 23% higher odds per unit allele increment in uwPRS (OR 1.23, $p < 0.01$).

4. Discussion

Despite the increasing detection of CAD-associated SNPs in large-scale GWAS, their incorporation into risk prediction models remains challenging. Herein, we summarised all the reported associations between PRS and CAD in a literature meta-analysis. Using data from 977,716 individuals, we demonstrate that weighted PRS (Fig. 5): independently predicted a greater risk of incident CAD; with more stringent weighting modalities, such as LDpred and metaGRS, showing the strongest associations with CAD risk. Moreover, PRS demonstrated the stronger association with EOCAD compared to AnyCAD or LOCAD Incorporation of PRS to clinical models modestly but significantly increased CAD risk prediction, with the best performance seen with PRS by lassosum, LDpred, and metaGRS.

Our findings highlight the importance of using weighted methods as opposed to unweighted methods for polygenic scoring and polygenic predictions of CAD. As a simple count score, the unweighted PRS sums all CAD-associated variants from discovery GWAS into an allele score. One major limitation with the uwPRS is the assumption of equal contributions from the risk variants included in the score. Consequently, uwPRS models have demonstrated poorer performance in several complex traits compared to wPRS models. In the present analysis, uwPRS predicted an 8% increased risk of incident CAD compared to 26%–49% for weighted PRS.

Additionally, our analysis demonstrated more rigorous weighting methodologies are needed for optimal performance of PRS. Earlier methods for polygenic scoring (i.e., PRS_{LN}) used strict p value thresholds and effect size estimates of few risk variants from discovery GWAS. These methods suffered from several limitations, among which included failure to account for smaller effect SNPs and the oversimplification of genetic heritability [17,18]. More recently, polygenic models using expanded PRS have achieved improved prediction of complex disorders [17,19]. Consistent with this, the current meta-analysis demonstrated the strongest associations with LDpred and metaGRS for incident and prevalent CAD. More complex methods, such as metaGRS and Bayesian regression by LDpred, owe their high performance to their ability to account for correlations between risk variants (via LD adjustments) [20,21] and do not suffer from model overfitting [22].

Furthermore, an important consideration for polygenic CAD risk prediction is the question of transethnic portability of PRS. Like most genetic studies, primary GWAS are predominantly performed in people of European ancestry. Thus, the derivation and validation of current PRS models reflect this dilemma, with non-European validation of PRS also performed using SNPs derived from European GWAS [23]. This has two-fold implications. First, it fails to account for the variable genetic aetiologies across multiple ethnicities [24]. Second, it grossly underestimates the differences in CAD incidence and mortality noted in contemporary observations. In a recent *trans*-ethnic meta-analysis of GWAS of Japanese and European populations, eight of the seventy-four identified SNPs were significantly different between two ethnicities [25]. In the first decade of PRS studies, 67% included European individuals and 19% East Asians, with only 3.8% using African, Hispanic, or Indigenous populations [26]. Consequently, the predictive performance was lower in non-European cohorts compared to European cohorts, with the least performance seen in African ancestry cohorts [26]. In the current analysis, association of PRS with CAD was greatest in European and mixed ancestry testing cohorts. Moreover, PRS from *trans*-ethnic and European GWAS demonstrated the greatest associations with CAD. These results underscore the need to diversify GWAS and PRS studies. It is expected that newer studies, such as the ongoing Multi-Ethnic New Zealand Study of Acute Coronary Syndromes (MENZACs) [27], would help clarify the transethnic translation of polygenic CAD risk prediction.

Current risk stratification systems, which are purely dependent on CRS such as the FRS or ACC/AHA PCE risk score, have a limited ability in identifying low-risk patients, [28,29], especially in non-European populations [30]. The recommendation for combined polygenic-clinical prediction models should rely on the evidence of a significant increment in prediction performance from clinical scores alone to combined genetic and clinical scores. Based on the results presented in the present analysis, we believe integrating PRS along with clinical risk scores in CAD risk prediction can improve performance. This will have far-reaching benefits, including early identification of patients at risk, early initiation of primary prevention, and better monitoring. Finally, the feasibility of using a combined prediction model should warrant further prospective investigations in future studies. These studies should focus on subgroups whose CAD risks are either underestimated or overestimated by conventional clinical risk scores, such as women, younger populations, lower-risk groups, and minority communities [31–34]. For instance, in a multi-ethnic cohort, the combination of polygenic and clinical prediction models had the best discrimination for incident CAD in the younger participants (AUC 0.80 $<$ 55 years *versus* AUC 0.74 $\geq$ 55 years) and women (AUC 0.76 *versus* AUC 0.71 in males) [35]. Additionally, there should be consideration for better phenotyping of control patients to avoid potential inclusion of subclinical atherosclerosis in both GWAS and PRS analyses; and triage for benefits in predicting CAD outcomes, identifying patients for plaque imaging, and guiding CAD treatment. Furthermore, to have utility in daily clinical practice, there would need to be increased availability, accessibility, and affordability of genomics and computational technology for routine care.

The present meta-analysis has several strengths and merits. The large number of patients provides a strong power for very robust CAD risk

$PRS_{Lassosum}$		PRS_{LDpred}					$PRS_{metaGRS}$		
CRS	PRS + CRS	PRS	CRS		PRS + CRS		PRS	CRS	PRS + CRS
			Any	PCE	Any	PCE			
2	2	6	8	4	8	4	3	4	4
356,869	356,869	475,291	531,769	340,396	531,769	340,396	484,276	538,364	538,364
0.759	0.779	0.705	0.752	0.729	0.770	0.748	0.613	0.694	0.712
(0.719–0.795)	(0.724–0.826)	(0.614–0.781)	(0.685–0.809)	(0.651–0.795)	(0.707–0.823)	(0.663–0.818)	(0.556–0.667)	(0.639–0.743)	(0.666–0.754)

estimation. The novelty in the present data is worth highlighting. We used studies with good methodological qualities and generally low risk of bias for our analyses. We highlight the importance of using stringent weighting methods and larger PRS. We have also shown that the usefulness of PRS is particularly noteworthy in patients with EOCAD than any CAD or late-onset CAD. However, these results should be interpreted with caution. First, the SNPs used for deriving PRS were from GWAS of individuals of European ancestry. Second, the PRS used herein is limited by the available GWAS studies, which may not have captured all CAD risk variants. Third, the lack of individual patient data is another limitation of the present study. Noteworthy, individual-patient data meta-analysis is shown to be superior in risk stratifications compared to aggregate analysis.

4.1. Conclusions

In conclusion, our results provide robust evidence associating genetic risk scores with CAD. PRS independently predicts increased risks of incident, prevalent, and recurrent CAD. Weighted scores are better associated with CAD risks compared to non-weighted polygenic scores, with the strongest associations seen with early-onset CAD. More importantly, CAD risk predictions were most robustly noted for scores with larger SNPs and complex weighting, such as metaGRS and LDpred-derived PRS. Future prospective studies should investigate the benefits of incorporating PRS for better prediction of CAD outcomes, identifying patients for plaque imaging, and guiding CAD treatment.

PROSPERO registration number

CRD42020190305.

Financial support

Dr Agbaedeng is supported by Postdoctoral Fellowships from The University of Adelaide, Australia. Dr Noubiap is supported by an Adelaide Scholarship International from The University of Adelaide, Australia.

CRediT authorship contribution statement

Thomas A. Agbaedeng: Conceptualization, Writing – original draft, Conception and Design, Search strategy, Studies selection, Data extraction, Data synthesis, Data interpretation, Manuscript drafting, Manuscript revision, Approval of the final manuscript. **Jean Jacques Noubiap:** Writing – original draft, Search strategy, Studies selection, Data synthesis, Data interpretation, Manuscript drafting, Manuscript revision, Approval of the final manuscript. **Edith Pascale Mofo Mato:** Writing – original draft, Data extraction, Data interpretation, Manuscript revision, Approval of the final manuscript. **Derek P. Chew:** Writing – original draft, Manuscript revision, Approval of the final manuscript. **Gemma A. Figtree:** Writing – original draft, Manuscript revision, Approval of the final manuscript. **M. Abdullah Said:** Writing – original draft, Manuscript revision, Approval of the final manuscript. **Pim van der Harst:** Writing – original draft, Manuscript revision, Approval of the final manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.atherosclerosis.2021.08.020.

References

[1] S.S. Virani, Alonso A, E.J. Benjamin, et al., Heart disease and stroke statistics—2020 update: a report from the American heart association, Circulation 141 (2020) e139–e596.
[2] G.A. Roth, C. Johnson, A. Abajobir, et al., Global, regional, and national burden of cardiovascular diseases for 10 causes, 1990 to 2015, J. Am. Coll. Cardiol. 70 (2017) 1–25.
[3] J.M. Schildkraut, R.H. Myers, L.A. Cupples, D.K. Kiely, W.B. Kannel, Coronary risk associated with age and sex of parental heart disease in the Framingham Study, Am. J. Cardiol. 64 (1989) 555–559.
[4] M. Nielsen, C. Andersson, T.A. Gerds, et al., Familial clustering of myocardial infarction in first-degree relatives: a nationwide study, Eur. Heart J. 34 (2013) 1198–1203.
[5] S. Zdravkovic, A. Wienke, N.L. Pedersen, M.E. Marenberg, A.I. Yashin, U. De Faire, Heritability of death from coronary heart disease: a 36-year follow-up of 20 966 Swedish twins, J. Intern. Med. 252 (2002) 247–254.
[6] J.C. Cohen, E. Boerwinkle, T.H. Mosley Jr., H.H. Hobbs, Sequence variations in PCSK9, low LDL, and protection against coronary heart disease, N. Engl. J. Med. 354 (2006) 1264–1272.
[7] N.S. Abul-Husn, K. Manickam, L.K. Jones, et al., Genetic identification of familial hypercholesterolemia within a single U.S. health care system, Science 354 (2016).
[8] A.V. Khera, H.H. Won, G.M. Peloso, et al., Diagnostic yield and clinical utility of sequencing familial hypercholesterolemia genes in patients with severe hypercholesterolemia, J. Am. Coll. Cardiol. 67 (2016) 2578–2589.
[9] M. Benn, G.F. Watts, A. Tybjærg-Hansen, B.G. Nordestgaard, Mutations causative of familial hypercholesterolaemia: screening of 98 098 individuals from the Copenhagen General Population Study estimated a prevalence of 1 in 217, Eur. Heart J. 37 (2016) 1384–1394.
[10] G.M. Peloso, P.L. Auer, J.C. Bis, et al., Association of low-frequency and rare coding-sequence variants with blood lipids and coronary heart disease in 56,000 whites and blacks, Am. J. Hum. Genet. 94 (2014) 223–232.
[11] T.W.T.C.C. Consortium, Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls, Nature 447 (2007) 661–678.
[12] P. van der Harst, N. Verweij, Identification of 64 novel genetic loci provides an expanded view on the genetic architecture of coronary artery disease, Circ. Res. 122 (2018) 433–443.
[13] M. Nikpay, A. Goel, H.H. Won, et al., A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease, Nat. Genet. 47 (2015) 1121–1130.
[14] M. Inouye, G. Abraham, C.P. Nelson, et al., Genomic risk prediction of coronary artery disease in 480,000 adults implications for primary prevention, J. Am. Coll. Cardiol. 72 (2018) 1883–1893.
[15] A.V. Khera, M. Chaffin, K.G. Aragam, et al., Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations, Nat. Genet. 50 (2018), 1219-+.
[16] S. Greenland, M.P. Longnecker, Methods for trend estimation from summarized dose-response data, with applications to meta-analysis, Am. J. Epidemiol. 135 (1992) 1301–1309.
[17] F. Dudbridge, Power and predictive accuracy of polygenic risk scores, PLoS Genet. 9 (2013), e1003348.

[18] N.R. Wray, J. Yang, B.J. Hayes, A.L. Price, M.E. Goddard, P.M. Visscher, Pitfalls of predicting complex traits from SNPs, Nat. Rev. Genet. 14 (2013) 507–515.
[19] M. Thomas, L.C. Sakoda, M. Hoffmeister, et al., Genome-wide modeling of polygenic risk score in colorectal cancer risk, Am. J. Hum. Genet. 107 (2020) 432–444.
[20] B.J. Vilhjálmsson, J. Yang, H.K. Finucane, et al., Modeling linkage disequilibrium increases accuracy of polygenic risk scores, Am. J. Hum. Genet. 97 (2015) 576–592.
[21] T.S.H. Mak, R.M. Porsch, S.W. Choi, X. Zhou, P.C. Sham, Polygenic scores via penalized regression on summary statistics, Genet. Epidemiol. 41 (2017) 469–480.
[22] P.J. Newcombe, C.P. Nelson, N.J. Samani, F. Dudbridge, A flexible and parallelizable approach to genome-wide polygenic risk scores, Genet. Epidemiol. 43 (2019) 730–741.
[23] P.G. Joseph, G. Pare, S. Asma, et al., Impact of a genetic risk score on myocardial infarction risk across different ethnic populations, Can. J. Cardiol. 32 (2016) 1440–1446.
[24] M. Scutari, I. Mackay, D. Balding, Using genetic distance to infer the accuracy of genomic prediction, PLoS Genet. 12 (2016), e1006288.
[25] H. Matsunaga, K. Ito, M. Akiyama, et al., Transethnic meta-analysis of genome-wide association studies identifies three New loci and characterizes population-specific differences for coronary artery disease, Circulation: Genomic and Precision Medicine 13 (2020), e002670.
[26] L. Duncan, H. Shen, B. Gelaye, et al., Analysis of polygenic risk score usage and performance in diverse human populations, Nat. Commun. 10 (2019) 3328.
[27] N.J. Earle, K.K. Poppe, R.N. Doughty, A. Rolleston, A.J. Kerr, M.E. Legget, Clinical characteristics and burden of risk factors among patients with early onset Acute coronary Syndromes: the ANZACS-QI New Zealand national cohort (ANZACS-QI 17), Heart Lung Circ. 27 (2018) 568–575.
[28] K. Nasir, M.S. Bittencourt, M.J. Blaha, et al., Implications of coronary artery calcium testing among statin candidates according to American college of cardiology/American heart association cholesterol management guidelines: MESA (Multi-Ethnic study of atherosclerosis), J. Am. Coll. Cardiol. 66 (2015) 1657–1668.
[29] P. Muntner, L.D. Colantonio, M. Cushman, et al., Validation of the atherosclerotic cardiovascular disease Pooled Cohort risk equations, Jama 311 (2014) 1406–1415.
[30] Y. Jiang, R. Ma, H. Guo, et al., External validation of three atherosclerotic cardiovascular disease risk equations in rural areas of Xinjiang, China, BMC Publ. Health 20 (2020) 1471.
[31] S.J. Baart, V. Dam, L.J.J. Scheres, et al., Cardiovascular risk prediction models for women in the general population: a systematic review, PloS One 14 (2019), e0210329.
[32] N.R. Cook, P.M. Ridker, Further insight into the cardiovascular risk calculator: the roles of statins, revascularizations, and underascertainment in the Women's Health Study, JAMA Intern Med 174 (2014) 1964–1971.
[33] J.S. Rana, G.H. Tabada, M.D. Solomon, et al., Accuracy of the atherosclerotic cardiovascular risk equation in a large contemporary, multiethnic population, J. Am. Coll. Cardiol. 67 (2016) 2118–2130.
[34] Z.M. Zhang, P.M. Rautaharju, R.J. Prineas, et al., Race and sex differences in the incidence and prognostic significance of silent myocardial infarction in the atherosclerosis risk in communities (ARIC) study, Circulation 133 (2016) 2141–2148.
[35] J. Elliott, B. Bodinier, T.A. Bond, et al., Predictive accuracy of a polygenic risk score-enhanced prediction model vs a clinical risk score for coronary artery disease, Jama 323 (2020) 636–645.