



## OPEN ACCESS

## EDITED BY

Carole Sudre,  
University College London, United Kingdom

## REVIEWED BY

Simon Body,  
Boston University, United States  
Venkataraghavan Ramamoorthy,  
Baptist Health South Florida, United States

## \*CORRESPONDENCE

Cuiping Pan  
✉ pancuiping@aipm-gba.org.cn  
Shilong Zhong  
✉ gdph\_zhongsl@gd.gov.cn;  
✉ zhongsl@hotmail.com

RECEIVED 19 September 2023

ACCEPTED 31 January 2024

PUBLISHED 13 February 2024

## CITATION

Qin M, Wu Y, Fang X, Pan C and Zhong S  
(2024) Polygenic risk score predicts all-cause  
death in East Asian patients with prior coronary  
artery disease.  
Front. Cardiovasc. Med. 11:1296415.  
doi: 10.3389/fcvm.2024.1296415

## COPYRIGHT

© 2024 Qin, Wu, Fang, Pan and Zhong. This is  
an open-access article distributed under the  
terms of the [Creative Commons Attribution  
License \(CC BY\)](#). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that the  
original publication in this journal is cited, in  
accordance with accepted academic practice.  
No use, distribution or reproduction is  
permitted which does not comply with these  
terms.

# Polygenic risk score predicts all-cause death in East Asian patients with prior coronary artery disease

Min Qin<sup>1,2</sup>, Yonglin Wu<sup>2,3</sup>, Xianhong Fang<sup>4</sup>, Cuiping Pan<sup>3,5\*</sup> and Shilong Zhong<sup>1,2,6\*</sup>

<sup>1</sup>School of Medicine, South China University of Technology, Guangzhou, China, <sup>2</sup>Department of Pharmacy, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, China, <sup>3</sup>Center for Intelligent Medicine Research, Greater Bay Area Institute of Precision Medicine (Guangzhou), School of Life Sciences, Fudan University, Guangzhou, China, <sup>4</sup>Department of Cardiology, Guangdong Cardiovascular Institute, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Guangzhou, China, <sup>5</sup>Center for Evolutionary Biology, Fudan University, Shanghai, China, <sup>6</sup>Guangdong Provincial Key Laboratory of Coronary Heart Disease Prevention, Guangdong Cardiovascular Institute, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Guangzhou, China

**Introduction:** Coronary artery disease (CAD) is a highly heritable and multifactorial disease. Numerous genome-wide association studies (GWAS) facilitated the construction of polygenic risk scores (PRS) for predicting future incidence of CAD, however, exclusively in European populations. Furthermore, identifying CAD patients with elevated risks of all-cause death presents a critical challenge in secondary prevention, which will contribute largely to reducing the burden for public healthcare.

**Methods:** We recruited a cohort of 1,776 Chinese CAD patients and performed medical follow-up for up to 11 years. A pruning and thresholding method was used to calculate PRS of CAD and its 14 risk factors. Their correlations with all-cause death were computed via Cox regression.

**Results and discussion:** We found that the PRS for CAD and its seven risk factors, namely myocardial infarction, ischemic stroke, angina, heart failure, low-density lipoprotein cholesterol, total cholesterol and C-reaction protein, were significantly associated with death ( $P \leq 0.05$ ), whereas the PRS of body mass index displayed moderate association ( $P < 0.1$ ). Elastic-net Cox regression with 5-fold cross-validation was used to integrate these nine PRS models into a meta score, metaPRS, which performed well in stratifying patients at different risks for death ( $P < 0.0001$ ). Combining metaPRS with clinical risk factors further increased the discerning power and a 4% increase in sensitivity. The metaPRS generated from the genetic susceptibility to CAD and its risk factors can well stratify CAD patients by their risks of death. Integrating metaPRS and clinical risk factors may contribute to identifying patients at higher risk of poor prognosis.

## KEYWORDS

coronary artery disease, polygenic risk scores, all-cause death, prognostic model, secondary prevention

## Introduction

Coronary artery disease (CAD) is the leading cause of death worldwide and associated with 17.8 million deaths annually (1–4). The number of people suffering from CAD has increased substantially over the past three decades. There are more than 197 million people living with CAD globally (5), resulting in huge expenditures on medical treatment, as well as secondary and tertiary preventions. Therefore, early identification of CAD patients at elevated risk of death permits precise delivery of more intense medical interventions and effective lifestyle changes, which are essential to decreasing mortality and reducing healthcare burden.

CAD is highly heritable with heritability estimated at 40%–60% (6). As such, genetic susceptibility profiles may confer the potential for early risk screening. During the last two decades, numerous genome-wide association studies (GWAS) were devoted to identifying single nucleotide variants (SNVs) associated with CAD and its risk factors by aggregating millions of samples (7–10). As germline DNA variation is quantifiable at the time of birth, which is before the onset of the disease or symptoms, and is therefore attractive for predicting the incidence of various cardiovascular diseases, including CAD (11), type 2 diabetes (12), atrial fibrillation (13), and myocardial infarction (14). Polygenic risk score (PRS), which integrates inherited risk inferred by GWAS into a single quantitative metric, has displayed its potential in various cardiovascular diseases to classify patients by risks of future incidence (15–17). Due to the broad availability of the GWAS data, PRS is readily applicable to many diseases. However, most studies

thus far mainly focused on future incidence, i.e., primary prevention, in individuals of European ancestry without pre-existing CAD (15, 16, 18). Only a few studies explored whether the genetic susceptibility to CAD and its risk factors could predict risks of adverse outcomes in patients with pre-existing CAD (19, 20). Few examined that in patients of East Asian ancestry.

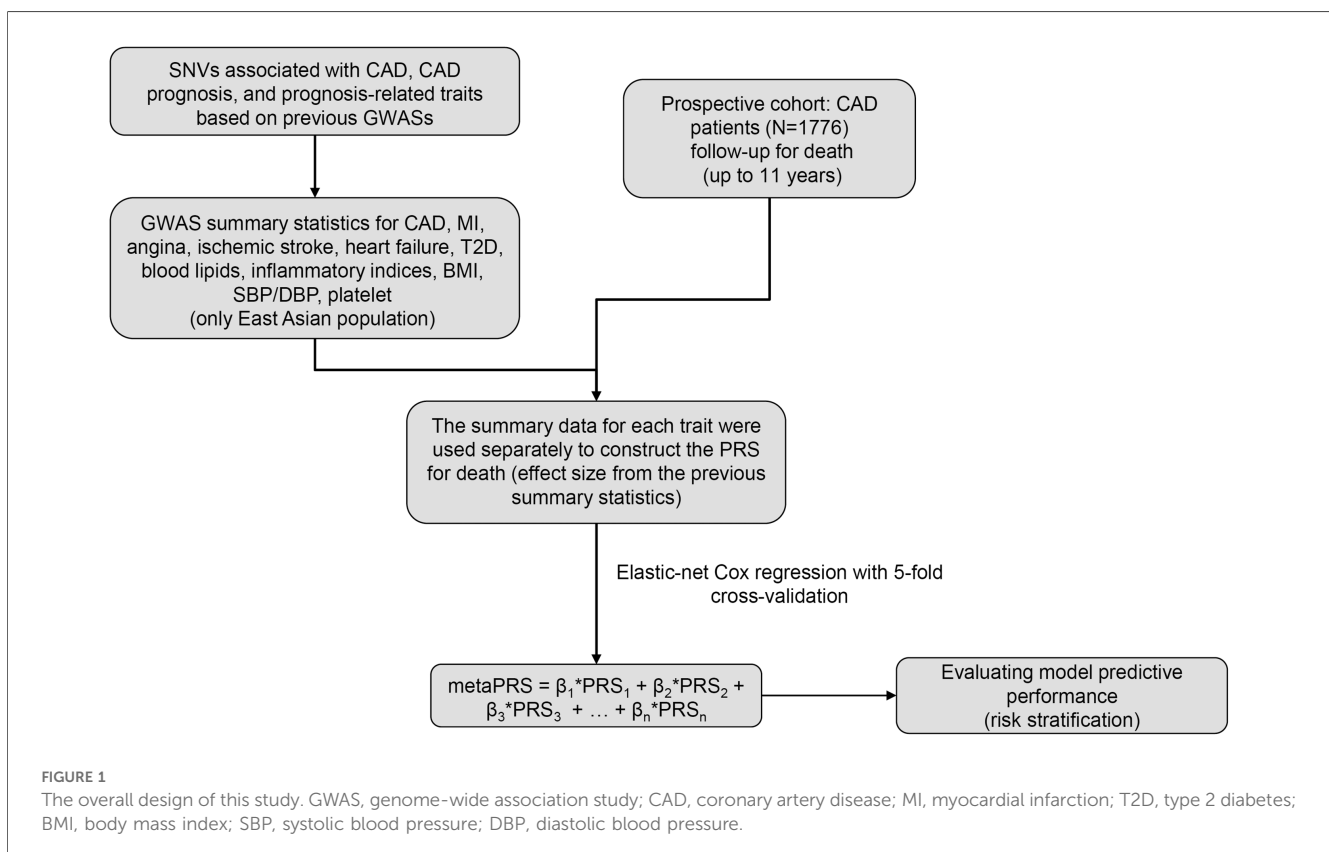
To address these needs, we recruited a prospective cohort consisting of 1,776 Chinese patients with pre-existing CAD and carried out medical follow-ups for up to 11 years. We aimed to leverage GWAS summary statistics derived from large cohorts of East Asians, such as Biobank Japan (9, 10), and develop the PRS models in our cohort for learning and adaptation. Our goal was to derive a meta PRS model for CAD patients of East Asian ancestry, which combined genetic profiles of CAD and its risk factors and was capable to predict the risk of poor prognosis.

## Methods

### Study population enrollment and baseline characteristics collection

This study was approved by the Medical Research Ethics Committee of Guangdong Provincial People's Hospital and complied with the Declaration of Helsinki. All patients provided written informed content.

The overall design of this study was depicted in Figure 1. We recruited 1,776 patients diagnosed with CAD from Guangdong



Provincial People's Hospital between January 2010 and December 2013 to form a prospective cohort. All individuals had stenosis of  $\geq 50\%$  in at least 1 major coronary artery by coronary angiography and/or a diagnosis of CAD based on the World Health Organization Criteria (21). Meanwhile, we obtained baseline characteristics, including demographics, medical history, biochemical measurements, and medication, from the hospital information database. In this study, subjects who met any of the following criteria were excluded: (1) renal dysfunction (defined as serum creatinine concentration  $>2$  times the upper limit of normal (ULN) (230  $\mu\text{mol/L}$ ), or with history of renal transplantation or dialysis); (2) hepatic dysfunction [defined as serum transaminase concentration  $>2$  times the ULN (80 U/L), or a diagnosis of cirrhosis]; (3) being pregnant or lactating; (4) diagnosed with advanced cancer or receiving hemodialysis; (5) history of thyroid disease and related medication; (6) blood sample and coronary angiography image unavailable; (7) missing follow-up information.

The endpoint of this study was an all-cause death event occurred during the follow-up period. All participants were followed prospectively and continuously by telephone contact with patients or their families every six months to collect the study's endpoint. The average follow-up time was 5 years and the maximum follow-up time was approximately 11 years.

## Single nucleotide variants genotyping and imputation

DNA was extracted on the hemocytes samples by TGuideM16 automatic nucleic acid extractor (Cat.NO. OSE-M16) with a matching extraction kit (Cat.NO. OSR-M102) of TIANGEN. Each individual was genotyped by the Global Screening Array (GSA) bead chip of Illumina, according to the standard protocol for Illumina Infinium HTS Assay. GenomeStudio software and the calling algorithm of Illumina were used to normalize the signal intensity. Finally, 700,078 SNVs were genotyped for each individual. Before genotype imputation, a series of strict quality controls were used to derive high-quality SNVs, with the following exclusion criteria: (1) samples with SNV genotyping rate less than 95%; (2) SNVs genotyped in less than 95% samples; (3) SNVs with a  $P$ -value of Hardy-Weinberg smaller than  $1\text{E-}06$ ; (4) SNVs with minor allele frequency (MAF) less than 5%.

Genotype imputation was performed against the China Metabolic Analytics Project (ChinaMAP, <http://www.mbiobank.com>), a large-scale and high-depth whole-genome sequencing dataset based on the Chinese population. This reference panel was constructed based on the ChinaMAP phase1 dataset, which includes 136,745,826 SNVs in 10,588 individuals. Our clean genotyped data were imputed by the ChinaMAP Imputation Server. After imputation, SNVs with the  $R^2$  smaller than 0.3 and/or SNVs with the  $P$ -value of Hardy-Weinberg smaller than  $1 \times 10^{-06}$  were excluded. At last, 4,933,061 high-quality SNVs of 1,776 CAD patients passed the quality assessment.

## Generation of polygenic risk score for all-cause death

GWAS summary statistics for 15 traits, including CAD, myocardial infarction, ischemic stroke, heart failure, angina, type 2 diabetes (T2D), body mass index (BMI), high-density lipoprotein cholesterol (HDL), low-density lipoprotein cholesterol (LDL), total cholesterol (CHOL), triglycerides (TG), systolic blood pressure (SBP), diastolic blood pressure (DBP), C-reaction protein (CRP), and platelet levels, were downloaded from Biobank Japan (Supplementary Table S1). The GWAS effect size were transferred to our data to construct PRS scores for our patients via a pruning and thresholding approach implemented in PRSice2 (22). Briefly, clumping was performed on our patients' genotypes to remove SNVs in linkage disequilibrium with the most significant SNVs ( $r^2 > 0.1$ ) within a 250 kb window. Subsequently, under each trait, different PRS models were constructed based on different SNVs, which were selected by incrementing  $P$ -value threshold from  $5 \times 10^{-8}$  to 0.5 in a few hundred steps. Thus, hundreds of PRS models were generated under each trait. Next, to identify the optimal PRS model per trait for predicting death, under each model, all PRS scores were used as independent variables for fitting a logistic regression with all-cause death. Among hundreds of regression results, the model with the maximum fitted  $R$ -square was defined as the optimal PRS model for that trait for predicting death.

## Construction of metaPRS

To construct the metaPRS, the optimal PRS models having significant association ( $P < 0.1$ ) with all-cause death were identified by Cox regression, adjusting for age, sex, medication including proton pump inhibitors (PPI), angiotensin-converting enzyme inhibitors (ACEI),  $\beta$ -blockers (BB), and calcium channel blockers (CCB), and the first 10 principal components (PCs) of the genotypes. Next, Elastic-net Cox regression with 5-fold cross-validation was used to construct the metaPRS by combining all significant PRS models using the R package "glmnet", which was proven effective in the presence of correlation between each PRS (23). A series of different metaPRS models were generated by tuning the parameter  $\lambda$  with different penalties, conferring different weights to each independent variable – in this case, the selected optimal PRS models. The performance of each metaPRS model was evaluated by the area under the operator curves (AUC), and the effect size of each optimal PRS model was determined when the AUC reached its maximum value among all penalties. At last, the metaPRS for death was determined from the model with the largest AUC value.

## Statistical analysis

In this study, the baseline characteristics of patients were presented as mean  $\pm$  standard deviation (SD) for continuous

variables and count (percent) for categorical variables. A two-tailed Students' *t*-test was used to identify the difference of continuous variables between patients with or without death,  $\chi^2$  test was used to detect the difference between two groups with normally distributed data, and the Mann-Whitney *U*-test was used to identify the difference between two groups with non-normally distributed data. A *P*-value less than 0.05 was considered statistically significant. Pearson correlation analysis was used to identify the relationship between each optimal PRS. These patients were divided into three groups with low [ $\leq$ quartile (Q) 1], intermediate ( $>Q1$  and  $<Q3$ ), and high risk of death ( $\geq Q3$ ) based on the quartiles of the calculated metaPRS for all individuals. The AUC values of all predictive models were calculated using the R package "timeROC". Sensitivity and specificity of the prediction models were calculated based on the function of confusionMatrix implemented in the R package "InformationValue".

## Results

### Baseline characteristics of the study population

The overall study design was illustrated in [Figure 1](#). We utilized a prospective cohort design and recruited 1,776 patients with CAD. During a follow-up period of up to 11 years, 171 patients experienced all-cause death, and the death rate of patients in this study was illustrated in [Supplementary Figure S1](#). Characteristics of the patients at enrollment, i.e., baseline, were displayed in [Table 1](#). The patients were predominantly male (79.56%), both in the death group and the survival group. The mean age of patients at enrollment was significantly older in the death group than in the survival group ( $P = 5.37 \times 10^{-18}$ ). In addition, patients in the death group were inclined for diabetes mellitus, heart failure, and arrhythmia, and their blood levels of aspartate aminotransferase, creatine kinase, creatine, and glucose were higher ( $P < 0.05$ , [Table 1](#)). Overall, baseline characteristics of these patients were consistent with the epidemiologic findings.

### Association between death and polygenic risk scores of CAD and its risk factors

We evaluated whether genetic susceptibility to CAD and its risk factors could predict all-cause death in patients with prior CAD. As such, we chose to study CAD and its 14 risk factors, namely myocardial infarction, ischemic stroke, heart failure, angina, T2D, BMI, HDLC, LDLC, CHOL, TG, SBP, DBP, CRP, and platelet levels. For each trait, we transferred the GWAS effect sizes, i.e., the beta values, learned in the East Asians in Biobank Japan, to compute PRS scores for our patients. For each trait, we used different significance thresholds (*P*-values ranged from  $5 \times 10^{-08}$  to 0.5) to select different risk alleles and computed the PRS, thus deriving a series of PRS scores. We then performed

TABLE 1 Baseline characteristics of the participants in the CAD prognostic cohorts.

| Characteristics                | Death (N = 171)        | Survival (N = 1,605)  | P-value |
|--------------------------------|------------------------|-----------------------|---------|
|                                | N (%) or mean $\pm$ SD | N(%) or mean $\pm$ SD |         |
| <b>Demographic data</b>        |                        |                       |         |
| Age (year)                     | 69 $\pm$ 9             | 62 $\pm$ 10           | <0.0001 |
| Male                           | 142 (83.04)            | 1,271 (79.09)         | 0.28    |
| <b>Medical history</b>         |                        |                       |         |
| Diabetes mellitus              | 72 (42.11)             | 420 (26.22)           | <0.0001 |
| Hypertension                   | 108 (63.16)            | 934 (58.27)           | 0.25    |
| Heart failure                  | 45 (26.32)             | 194 (12.10)           | <0.0001 |
| Arrhythmia                     | 28 (16.37)             | 48 (7.08)             | 0.0002  |
| <b>Biomedical measurements</b> |                        |                       |         |
| ALT, U/L                       | 29 $\pm$ 17            | 27 $\pm$ 14           | 0.13    |
| AST, U/L                       | 35 $\pm$ 66            | 26 $\pm$ 11           | <0.0001 |
| LDLC, mmol/L                   | 2.61 $\pm$ 0.97        | 2.67 $\pm$ 0.99       | 0.45    |
| HDLC, mmol/L                   | 0.97 $\pm$ 0.24        | 0.99 $\pm$ 0.24       | 0.31    |
| Triglycerides, mmol/L          | 1.60 $\pm$ 1.20        | 1.64 $\pm$ 1.12       | 0.69    |
| ApoA, g/L                      | 1.02 $\pm$ 0.28        | 1.07 $\pm$ 0.26       | 0.01    |
| TC, mmol/L                     | 4.25 $\pm$ 1.14        | 4.34 $\pm$ 1.25       | 0.38    |
| LPA, mg/dl                     | 333.51 $\pm$ 350.81    | 290.69 $\pm$ 315.30   | 0.13    |
| CK, U/L                        | 150.10 $\pm$ 591.76    | 111.16 $\pm$ 106.83   | 0.02    |
| CKMB, U/L                      | 8.5 $\pm$ 8.1          | 8.0 $\pm$ 6.1         | 0.39    |
| Creatinine, $\mu$ mol/L        | 102 $\pm$ 38           | 86 $\pm$ 39           | <.0001  |
| Glucose, mmol/L                | 7.11 $\pm$ 2.68        | 6.54 $\pm$ 2.61       | 0.007   |
| <b>Medication</b>              |                        |                       |         |
| PPI                            | 104 (60.82)            | 863 (53.77)           | 0.09    |
| ACEI                           | 105 (61.40)            | 949 (59.13)           | 0.62    |
| BB                             | 151 (88.30)            | 1,416 (88.22)         | 0.99    |
| CCB                            | 70 (40.94)             | 429 (26.73)           | 0.0001  |

ALT, alanine aminotransferase; AST, aspartate aminotransferase; LDLC, low-density lipoprotein cholesterol; HDLC, high-density lipoprotein cholesterol; ApoA, apolipoprotein A; TC, total cholesterol; LPA, Lipoprotein (a); CK, creatine kinase; CKMB, creatine kinase MB; PPI, proton pump inhibitors; ACEI, angiotensin-converting enzyme inhibitors; BB,  $\beta$ -blockers; CCB, calcium channel blockers.

logistic regression, treating each PRS score as an independent variable and adjusting for covariates including sex, age, medications, and the first 10 principal components (PC) of the genotypes, for identifying which PRS model of that particular trait was best associated with all-cause death in our cohort. The optimal PRS was defined as having the biggest effect size in the association tests. As such, we derived 15 optimal PRS models and their scores for each patient in our cohort ([Supplementary Table S1](#)). We discovered that the PRS scores of seven traits, i.e., CAD, IS, MI, HF, Angina, TC, and LDLC, showed significant association with all-cause death ( $P < 0.05$ , [Figure 2](#)), and CRP and BMI showed weak significance with death ( $P < 0.1$ , [Figure 2](#)). In total, nine PRS models displayed significant association with all-cause death with the direction of effect consistent with the epidemiologic findings.

Next, we explored the overlap of variants used to compose the nine PRS models and their pairwise correlations. The number of variants in each model ranged from 1,026 to 51,982 ([Figure 3A](#)), exhibiting varying degrees of overlap. PRS of CAD overlapped the most with myocardial infarction,



## Predicting death risk by a meta polygenic risk score

Given that the nine PRS models correlated with each other to different degrees, we adopted the elastic-net Cox regression approach with 5-fold cross-validation to integrate them all into a metaPRS model, which consisted of 193,312 unique SNVs. As expected, the metaPRS had the strongest association with all-cause death (HR: 1.85; 95% CI: 1.50–2.28) compared to the single PRS models (Figure 2).

For assessing its ability in risk stratification, we divided our patients into three groups by the quartiles of metaPRS scores, with the scores less than the first quartile ( $\leq Q1$ ) and greater than the third quartile ( $\geq Q3$ ) as the low and the high risk groups, respectively, and the scores between first and third quartiles ( $>Q1$  and  $<Q3$ ) as the intermediate risk group. We found that the cumulative incidence of all-cause death was significantly greater in CAD patients in the third quartile than in the first quartile, with the HR of 3.99 (95% CI: 2.40–6.64) per SD increment ( $P = 9.10 \times 10^{-8}$ ), and the cumulative incidence of all-cause death in CAD patients with intermediate metaPRS was also significantly greater than in the first quartile, with the HR of 2.18 (95% CI: 1.33–3.59) per SD increment ( $P = 2.10 \times 10^{-3}$ ) (Figure 4).

## Predicting death risk by polygenic and clinical risk scores

Other than genetic susceptibility, environmental factors and lifestyles also contribute to the incidence and progression of cardiovascular diseases. For example, age is a known strong predictor. Therefore, we explored whether combining metaPRS

and clinical risk factors could improve the prediction power for all-cause death in CAD. First, we evaluated single risk factors and found that age had a good predictability with an AUC of 0.69 (Supplementary Table S2), followed by smoking (AUC = 0.58) and sex (AUC = 0.54). As a reference, the metaPRS model achieved an AUC of 0.63. Next, we combined the three clinical factors and achieved an AUC of 0.74. Combining metaPRS with the three clinical risk factors resulted in a further increase of AUC to 0.76.

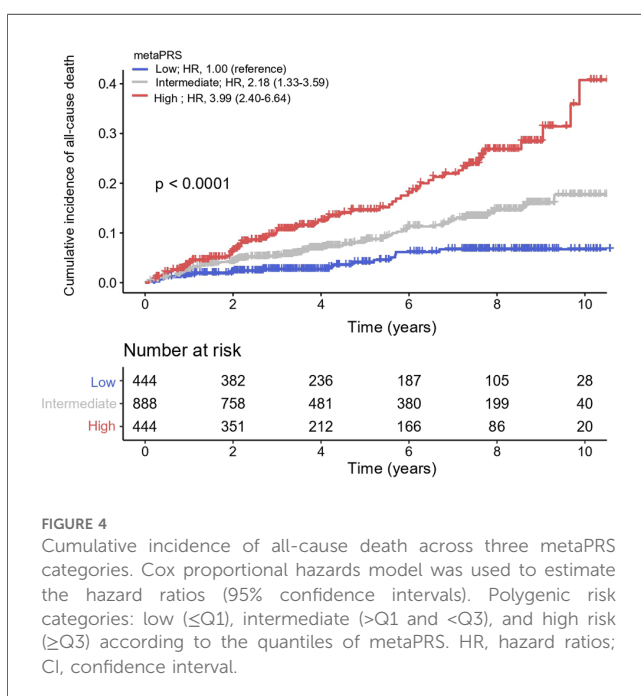
Close examination of the prediction scores revealed that combining metaPRS with the clinical risk factors, which we termed as the combined model, significantly improved prediction than using the clinical risk factors alone (Figure 5). While no difference was observed for the survival group between the combined model and the clinical risk model, an increase in prediction scores was observed in the death group using the combined model. Furthermore, we observed a 4% increase in sensitivity after adding metaPRS to the clinical risk model, while the specificity remained unchanged (Table 2). Taken together, these results indicate that combining genetic and clinical risk factors is most advantageous in identifying patients at elevated risk of death, which is critical to enhancing secondary preventions in CAD patients.

## Discussion

In this study, we evaluated whether genetic susceptibility to CAD and its risk factors could predict future death in CAD patients. Our results showed that the PRSs for CAD and its eight risk factors were significantly associated with death. Meanwhile, we developed a metaPRS model for predicting death, which could well stratify CAD patients of low, intermediate, and high risks of death. In addition, we found that combining metaPRS with clinical risk factors moderately increased the prediction power compared with using clinical risk factors alone.

Previous large-scale GWAS studies were predominantly conducted in European ancestry populations and focused on the incidence of CAD (15, 16), leading to an insufficient discovery both in East Asian ancestry and adverse outcomes of CAD. The large-scale meta-GWAS analysis of CAD and its risk factors conducted by Biobank Japan (9, 10) allowed us to apply their learning in the East Asian population to study the impact of genetics on CAD progression in our Chinese patients. The good stratification of patients at low, intermediate, and high risk of death by our proposed metaPRS model supports the notion that genetics not only contributes to the incidence but also progression of CAD. Our metaPRS model provides a valid quantification for such genetic contribution.

The prognosis of CAD is a complex conversation driven by both genetic and nongenetic factors, such as environmental exposure and lifestyles (24), presenting a challenge in predicting future adverse outcomes. In this study, we found that PRS of CAD had a positive association with all-cause death in CAD patients, after adjusting for traditional risk factors and medication usage. This was consistent with previous studies, such



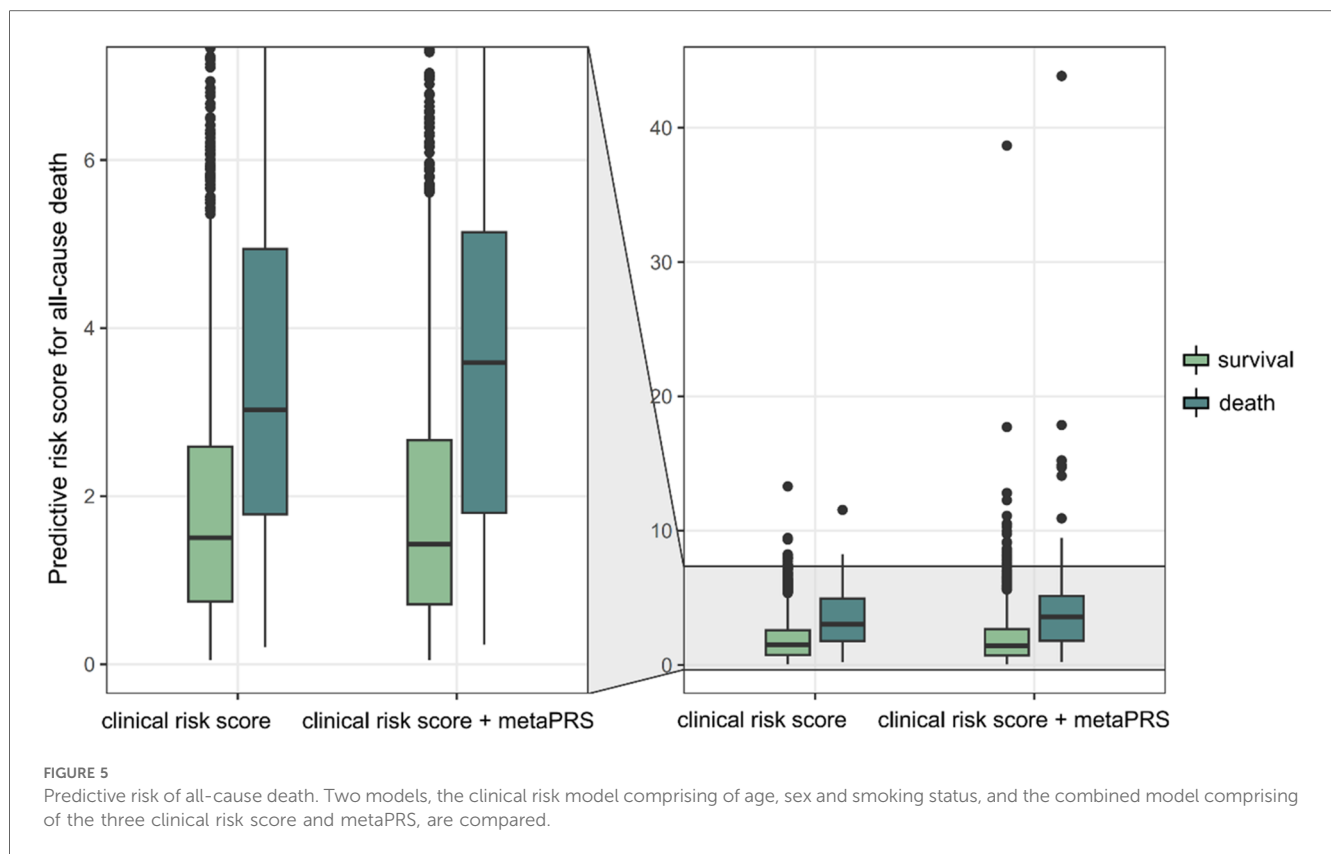


TABLE 2 Sensitivity analysis for metaPRS, clinical model and the combination of clinical and metaPRS model.

| Predictive   | Reference | metaPRS  |       | Clinical model |       | Clinical + metaPRS model |       |
|--------------|-----------|----------|-------|----------------|-------|--------------------------|-------|
|              |           | Survival | Death | Survival       | Death | Survival                 | Death |
| Survival     |           | 954      | 70    | 1,146          | 66    | 1,136                    | 60    |
|              |           | (TN)     | (FN)  | (TN)           | (FN)  | (TN)                     | (FN)  |
| Death        |           | 651      | 101   | 459            | 105   | 469                      | 111   |
|              |           | (FP)     | (TP)  | (FP)           | (TP)  | (FP)                     | (TP)  |
| Totals       |           | 1,605    | 171   | 1,605          | 171   | 1,605                    | 171   |
| Sensitivity  |           |          | 59%   |                | 61%   |                          | 65%   |
| Specificity  |           | 59%      |       | 71%            |       | 71%                      |       |
| Cutoff value |           | 1.095    |       | 2.4            |       |                          |       |

TN, true negative; TP, true positive; FN, false negative; FP, false positive.

as Odyssey Outcome Trial and Fourier Trial that CAD patients of higher PRS would benefit better from the medical treatments from death and major adverse cardiovascular events (25–28). Meanwhile, we noticed that all PRS scores for severe and even fatal events such as myocardial infarction, heart failure, and ischemic stroke harbored greater associations with death in the CAD patients than the PRS of CAD itself, suggesting that genetic susceptibility to CAD reflected broader spectrum of disease conditions but was less effective in predicting adverse outcomes, whereas genetic susceptibility to more extreme conditions confer greater risk to the patients. Thus, our study highlights the importance of considering overall genetic profiles, particularly those associated with fatal conditions, when developing tools for the secondary prevention of CAD.

In addition to predicting the incidence of CAD (18), our study revealed that age was also the biggest contributor to predicting future death in CAD patients. Although the addition of metaPRS to the clinical risk factors achieved only a modest improvement in the AUC value, the median of the predictive risk scores had an obvious elevation in the death group, which in turn resulted in a 4% increase in sensitivity. Studies suggest that restoring a healthy lifestyle and increasing physical activity may reduce the risk of cardiovascular death and all-cause death, even in patients with the highest genetic risks (24, 29, 30). Thus, predicting a long-term risk solely based on age may lead to an inaccurate estimation. Genetic risk supplements the picture by providing an independent component, which can be detected early in life and has demonstrated promises in predicting future outcomes (31,

32). Our results showed that combining genetic risk with clinical risk factors achieved the best prediction power. Indeed, the combined model increased the risk scores for patients in the death group while maintaining the scores at similar levels for patients in the survival group, suggesting the genetic risk harbored specificity and did not broadly inflate the risk scores. Such specificity confers the combined model the capability to identify patients at higher risk for adverse outcomes.

We highlight the strengths and limitations of this study. For the strengths, first, we recruited a prospective CAD cohort with detailed baseline information and up to 11 years of follow-up on all-cause death, which allowed us to explore comprehensively the genetic profile of CAD and its risk factors and their ability to predict a 10-year death risk in the CAD patients. Second, this study filled a gap in utilizing PRS for secondary prevention of CAD in East Asian populations. However, our study also has limitations. First, the cohort enrolled in this study has a small sample size relative to biobank studies such as UK Biobank. Second, this study did not include an external validation cohort to verify the performance of metaPRS, due to the lack of CAD patient cohorts with long-term follow-up. Future studies are warranted for validation.

## Conclusions

In summary, our study developed a polygenic risk score with good promising in facilitating the identification of patients at higher risk of developing adverse outcomes. Incorporating polygenic risk scores into clinical care may provide a valuable guidance on risk stratification for early identification of patients who would benefit from intensive lifestyle changes and drug treatment.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: EMBL-EBI [Project: PRJEB42554; Analyses: ERZ1714343].

## Ethics statement

The studies involving humans were approved by the Medical Research Ethics Committee of Guangdong Provincial People's Hospital. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

## Author contributions

MQ: Conceptualization, Writing – review & editing, Formal Analysis, Methodology, Visualization, Writing – original draft.

YW: Methodology, Writing – review & editing, Investigation. XF: Methodology, Writing – review & editing, Data curation. CP: Methodology, Writing – review & editing, Conceptualization, Project administration, Supervision. SZ: Conceptualization, Project administration, Supervision, Writing – review & editing, Data curation, Funding acquisition, Resources.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article.

This study was supported by the National Natural Science Foundation of China (No. 82274016, 32270626, 81872934, 82304623), the Key-Area Research and Development Program of Guangdong Province, China (No. 2019B020229003), Greater Bay Area Research Institute of Precision Medicine (Guangzhou) Research Grants (I0005, R2001), National key research and development program (No. 2017YFC0909301) and the Science and Technology Program of Guangzhou (No. 2023B03J1251).

## Acknowledgments

We thank the CAD patients for their willingness to consent and participate in this scientific study. We thank the medical staff of the Guangdong Provincial People's Hospital, particularly the nurses, for their diligent and careful follow-up of the patients. We thank the Biobank Japan (BBJ) Project consortium for access to GWAS summary statistics data.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcvm.2024.1296415/full#supplementary-material>



## References

- Ananth CV, Rutherford C, Rosenfeld EB, Brandt JS, Graham H, Kostis WJ, et al. Epidemiologic trends and risk factors associated with the decline in mortality from coronary heart disease in the United States, 1990–2019. *Am Heart J.* (2023) 263:46–55. doi: 10.1016/j.ahj.2023.05.006
- Li S, Liu Z, Joseph P, Hu B, Yin L, Tse LA, et al. Modifiable risk factors associated with cardiovascular disease and mortality in China: a PURE substudy. *Eur Heart J.* (2022) 43(30):2852–63. doi: 10.1093/eurheartj/ehac268
- Tsao CW, Aday AW, Almarzooq ZI, Anderson CAM, Arora P, Avery CL, et al. Heart disease and stroke statistics-2023 update: a report from the American heart association. *Circulation.* (2023) 147(8):e93–621. doi: 10.1161/CIR.0000000000001123
- Disease GBD, Injury I, Prevalence C. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017. *Lancet.* (2018) 392(10159):1789–858. doi: 10.1016/S0140-6736(18)32279-7
- Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, et al. Global burden of cardiovascular diseases and risk factors, 1990–2019: update from the GBD 2019 study. *J Am Coll Cardiol.* (2020) 76(25):2982–3021. doi: 10.1016/j.jacc.2020.11.010
- McPherson R, Tybjaerg-Hansen A. Genetics of coronary artery disease. *Circ Res.* (2016) 118(4):564–78. doi: 10.1161/CIRCRESAHA.115.306566
- Tcheandjieu C, Zhu X, Hilliard AT, Clarke SL, Napolioni V, Ma S, et al. Large-scale genome-wide association study of coronary artery disease in genetically diverse populations. *Nat Med.* (2022) 28(8):1679–92. doi: 10.1038/s41591-022-01891-3
- Aragam KG, Jiang T, Goel A, Kanoni S, Wolford BN, Atri DS, et al. Discovery and systematic characterization of risk variants and genes for coronary artery disease in over a million participants. *Nat Genet.* (2022) 54(12):1803–15. doi: 10.1038/s41588-022-01233-6
- Sakaue S, Kanai M, Tanigawa Y, Karjalainen J, Kurki M, Koshiba S, et al. A cross-population atlas of genetic associations for 220 human phenotypes. *Nat Genet.* (2021) 53(10):1415–24. doi: 10.1038/s41588-021-00931-x
- Ishigaki K, Akiyama M, Kanai M, Takahashi A, Kawakami E, Sugishita H, et al. Large-scale genome-wide association study in a Japanese population identifies novel susceptibility loci across different diseases. *Nat Genet.* (2020) 52(7):669–79. doi: 10.1038/s41588-020-0640-3
- Patel AP, Wang M, Ruan Y, Koyama S, Clarke SL, Yang X, et al. A multi-ancestry polygenic risk score improves risk prediction for coronary artery disease. *Nat Med.* (2023) 29(7):1793–803. doi: 10.1038/s41591-023-02429-x
- Ge T, Irvin MR, Patki A, Srinivasainagendra V, Lin YF, Tiwari HK, et al. Development and validation of a trans-ancestry polygenic risk score for type 2 diabetes in diverse populations. *Genome Med.* (2022) 14(1):70. doi: 10.1186/s13073-022-01074-2
- Miyazawa K, Ito K, Ito M, Zou Z, Kubota M, Nomura S, et al. Cross-ancestry genome-wide analysis of atrial fibrillation unveils disease biology and enables cardioembolic risk prediction. *Nat Genet.* (2023) 55(2):187–97. doi: 10.1038/s41588-022-01284-9
- Isgut M, Sun J, Quyyumi AA, Gibson G. Highly elevated polygenic risk scores are better predictors of myocardial infarction risk early in life than later. *Genome Med.* (2021) 13(1):13. doi: 10.1186/s13073-021-00828-8
- Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet.* (2018) 50(9):1219–24. doi: 10.1038/s41588-018-0183-z
- Inouye M, Abraham G, Nelson CP, Wood AM, Sweeting MJ, Dudbridge F, et al. Genomic risk prediction of coronary artery disease in 480,000 adults: implications for primary prevention. *J Am Coll Cardiol.* (2018) 72(16):1883–93. doi: 10.1016/j.jacc.2018.07.079
- Wang M, Menon R, Mishra S, Patel AP, Chaffin M, Tanneer D, et al. Validation of a genome-wide polygenic score for coronary artery disease in south Asians. *J Am Coll Cardiol.* (2020) 76(6):703–14. doi: 10.1016/j.jacc.2020.06.024
- Lu X, Liu Z, Cui Q, Liu F, Li J, Niu X, et al. A polygenic risk score improves risk stratification of coronary artery disease: a large-scale prospective Chinese cohort study. *Eur Heart J.* (2022) 43(18):1702–11. doi: 10.1093/eurheartj/ehac093
- Levin MG, Kember RL, Judy R, Birtwell D, Williams H, Arany Z, et al. Genomic risk stratification predicts all-cause mortality after cardiac catheterization. *Circ Genom Precis Med.* (2018) 11(11):e002352. doi: 10.1161/CIRCGEN.118.002352
- Yun JS, Jung SH, Shivakumar M, Xiao B, Khera AV, Park WY, et al. Associations between polygenic risk of coronary artery disease and type 2 diabetes, lifestyle, and cardiovascular mortality: a prospective UK biobank study. *Front Cardiovasc Med.* (2022) 9:919374. doi: 10.3389/fcvm.2022.919374
- None. Nomenclature and criteria for diagnosis of ischemic heart disease. Report of the joint international society and federation of cardiology/world health organization task force on standardization of clinical nomenclature. *Circulation.* (1979) 59(3):607–9. doi: 10.1161/01.CIR.59.3.607
- Choi SW, O'Reilly PF. PRSice-2: polygenic risk score software for biobank-scale data. *Gigascience.* (2019) 8(7):giz082. doi: 10.1093/gigascience/giz082
- Abraham G, Malik R, Yonova-Doing E, Salim A, Wang T, Danesh J, et al. Genomic risk score offers predictive performance comparable to clinical risk factors for ischaemic stroke. *Nat Commun.* (2019) 10(1):5819. doi: 10.1038/s41467-019-13848-1
- Khera AV, Emdin CA, Drake I, Natarajan P, Bick AG, Cook NR, et al. Genetic risk, adherence to a healthy lifestyle, and coronary disease. *N Engl J Med.* (2016) 375(24):2349–58. doi: 10.1056/NEJMoa1605086
- Mega JL, Stitzel NO, Smith JG, Chasman DI, Caulfield M, Devlin JJ, et al. Genetic risk, coronary heart disease events, and the clinical benefit of statin therapy: an analysis of primary and secondary prevention trials. *Lancet.* (2015) 385(9984):2264–71. doi: 10.1016/S0140-6736(14)61730-X
- Damask A, Steg PG, Schwartz GG, Szarek M, Hagstrom E, Badimon L, et al. Patients with high genome-wide polygenic risk scores for coronary artery disease may receive greater clinical benefit from alirocumab treatment in the ODYSSEY OUTCOMES trial. *Circulation.* (2020) 141(8):624–36. doi: 10.1161/CIRCULATIONAHA.119.044434
- Natarajan P, Young R, Stitzel NO, Padmanabhan S, Baber U, Mehran R, et al. Polygenic risk score identifies subgroup with higher burden of atherosclerosis and greater relative benefit from statin therapy in the primary prevention setting. *Circulation.* (2017) 135(22):2091–101. doi: 10.1161/CIRCULATIONAHA.116.024436
- Marston NA, Kamanu FK, Nordio F, Gurmu Y, Roselli C, Sever PS, et al. Predicting benefit from evolocumab therapy in patients with atherosclerotic disease using a genetic risk score: results from the FOURIER trial. *Circulation.* (2020) 141(8):616–23. doi: 10.1161/CIRCULATIONAHA.119.043805
- Said MA, Verweij N, van der Harst P. Associations of combined genetic and lifestyle risks with incident cardiovascular disease and diabetes in the UK biobank study. *JAMA Cardiol.* (2018) 3(8):693–702. doi: 10.1001/jamacardio.2018.1717
- Tikkanen E, Gustafsson S, Ingelsson E. Associations of fitness, physical activity, strength, and genetic risk with cardiovascular disease: longitudinal analyses in the UK biobank study. *Circulation.* (2018) 137(24):2583–91. doi: 10.1161/CIRCULATIONAHA.117.032432
- Aragam KG, Natarajan P. Polygenic scores to assess atherosclerotic cardiovascular disease risk: clinical perspectives and basic implications. *Circ Res.* (2020) 126(9):1159–77. doi: 10.1161/CIRCRESAHA.120.315928
- Torkamani A, Wineinger NE, Topol EJ. The personal and clinical utility of polygenic risk scores. *Nat Rev Genet.* (2018) 19(9):581–90. doi: 10.1038/s41576-018-0018-x