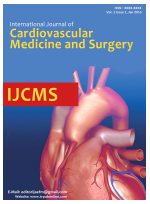
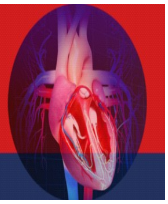
Nguyen T. Minh / Afj.Card.Cardvm. 2(1) (2026) 10-18
https://doi.org/10.62587/AFJCCM.2.1.2026.10-18 ISSN: 3136-9294



African Journal of Cardiology and Cardiovascular Medicine

Journal homepage: https://www.afjccm.com/



Research Article **Open Access**

Polygenic risk score integration with conventional risk factors for primary prevention of coronary artery disease in a South Asian Cohort

Nguyen T. Minh[1*]

[1]Cardiovascular Center, Vietnam Mekong University, Lotus Research Avenue, Hanoi, Vietnam.

Article Info

Volume 2, Issue 1, January 2026
Received : 12 November 2025
Accepted : 02 January 2026
Published : 25 January 2026
doi: 10.62587/AFJCCM.2.1.2026.10-18

Abstract

Objective: To evaluate if a Polygenetic Risk Score (PRS) combined with traditional risk factors (TNFRs) is more strongly associated with an incident Coronary Artery Disease (CAD) in a South Asian patient cohort for primary prevention. **Methods:** We performed a genome-wide polygenic risk (PRS) and conventional risk factor characterisation of 6800 South Asian adults without CAD at baseline and followed them for incident CAD. Harrell C-statistic, net reclassification improvement/integrated discrimination improvement with model reclassification of intermediate-risk individuals, and calibration were calculated to evaluate added value of the score above traditional risk factors while the impact of score in association to incident disease was assessed by Cox regression. **Results:** 631 participants (9.3%) had developed coronary artery disease after a median follow-up of 7-8 years. The PRS proved to be independent risk factor for the occurrence of the disease (HR per SD 1.61; 95% CI 1.49 to 1.74) with 4.49 folds greater risk among those in the highest 10th percentile of the PRS (aka decile) than in the lowest 10th percentile. Added this score to conventional risk factors increased the C-statistic from 0.738 to 0.765 (change 0.027) and resulted in a net reclassification improvement of 0.084; and reclassified 12.1% of those who were intermediate risk to high risk. **Conclusion:** In the South Asian population, a moderate but significant benefit was observed for the use of a PPRS in combination with traditional risk factors to predict incident coronary artery disease, and in the reclassification of a large number of intermediate-risk participants, suggesting a potential use of PPRS in modified primary prevention in this high-risk population.

Keywords: *Polygenic risk score, Coronary artery disease, Primary prevention, Risk prediction, Genetic epidemiology*

© 2026 Nguyen T. Minh. This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

1. Introduction

The greatest opportunity for primary prevention of CAD is to assess an individual's risk for disease prior to that occurrence and target preventive measures, especially statins, to those who are at sufficiently high risk. It

* *Corresponding author: Nguyen T. Minh, Cardiovascular Center, Vietnam Mekong University, Lotus Research Avenue, Hanoi, Vietnam.*

3136-9294/© 2026 Nguyen T. Minh. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

is traditionally estimated using a few risk factors: age, sex, diabetes, hypertension, smoking and lipid concentrations in a risk score, like the pooled cohort equations (Goff *et al.*, 2014; Arnett *et al.*, 2019). Each of these factors has strong evidence to support it (most have been established by large international trials of myocardial infarction), but these factors only account for a portion of the risk involved, with a significant amount of variance in who develops coronary disease unexplained by these factors alone (Yusuf *et al.*, 2004).

Some of the unexplained variation is genetic. The heritability of coronary artery disease lies in many thousands of common DNA variants, all of which have a small effect. As these small effects add together, it is possible to calculate a sum (known as a polygenetic risk score), which quantifies inherited risk and is fixed at birth, thus largely independent from conventional mid-life risk factors (Khera *et al.*, 2018; Inouye *et al.*, 2018). Those who score high on such a test have several-fold greater risk of coronary disease, as great as that from a rare monogenic disease, and they respond more to statin therapy (Mega *et al.*, 2015; Natarajan *et al.*, 2017).

Whether adding a polygenic risk score to the conventional factors improves risk prediction enough to be useful has been debated. In large European cohorts the improvement in discrimination has been real but modest, while the reclassification of individuals at intermediate risk, the group in whom the treatment decision is genuinely uncertain, has been more substantial (Elliott *et al.*, 2020; Sun *et al.*, 2021; Mars *et al.*, 2020). Two considerations make the question particularly relevant to South Asians. First, they bear a high and early burden of coronary disease that the conventional scores, derived largely in other populations, tend to underestimate (Volgman *et al.*, 2018). Second, most polygenic scores were developed in individuals of European ancestry and transfer imperfectly to other groups, so their value in South Asians cannot be assumed and must be shown (Martin *et al.*, 2019; Wang *et al.*, 2020).

The way risk estimates are used in practice makes the intermediate-risk group central to this question. Prevention guidelines translate a continuous estimate of risk into a small number of categories, and the decision to start a statin is straightforward at the extremes but genuinely uncertain in the middle, where the estimated risk lies near the treatment threshold. It is in this intermediate band that additional information has the greatest capacity to change a decision, and it is therefore against the reclassification of these individuals, rather than against a change in a global measure of discrimination alone, that the value of a polygenic score is most usefully judged.

This study evaluated whether integrating a polygenic risk score with conventional risk factors improves the prediction of incident coronary artery disease for primary prevention in a South Asian cohort. In a prospectively followed cohort free of coronary disease at baseline, it assessed the association of the score with incident disease, its independence from the conventional factors, and the improvement in discrimination, reclassification and calibration that it added, with particular attention to the individuals at intermediate conventional risk.

2. Materials and methods

2.1. Study design and setting

This was a prospective cohort study conducted in the department of preventive cardiology of a tertiary care hospital, in a South Asian population. Adults free of coronary artery disease at enrolment were characterised by conventional risk factors and a polygenic risk score and followed for the development of coronary artery disease. The study was designed and reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology guidelines (von Elm *et al.*, 2007).

2.2. Participants

The primary-prevention cohort included individuals of South Asian descent who had no history of Coronary Artery Disease (CAD), nor any clinical evidence of CAD at baseline. Those with a prior history of coronary disease, or with a familial disorder of lipids whose management was specifically determined by the genotypic results, or with insufficient genotyping and/or follow-up data were excluded. 6800 participants entered and were followed.

2.3. Conventional risk factors

At baseline the conventional cardiovascular risk factors were recorded, comprising age, sex, diabetes mellitus,

hypertension, current smoking, the low-density lipoprotein cholesterol concentration and a family history of premature coronary disease. These variables constituted the conventional risk model against which the polygenic score was compared.

2.4. Polygenic risk score

A genome-wide polygenic risk score for coronary artery disease was computed for each participant from genotyping data, as the weighted sum of the risk alleles across the variants of the score, using published variant weights. The score was standardised to a mean of zero and a standard deviation of one within the cohort, and participants were also grouped into deciles and tertiles of the score for the analyses of the risk gradient. To place the score on an interpretable scale, the risk associated with each standard deviation and each quantile was expressed relative to the remainder of the cohort, and the deciles and tertiles were defined within the cohort so that the risk gradient reflected the distribution of the score in this population rather than in an external reference. The score was computed without knowledge of the participants' outcomes, and the outcomes were ascertained without knowledge of the score.

2.5. Outcome and follow-up

The endpoint was incident coronary artery disease which was determined as first myocardial infarction, coronary revascularisation, or death from coronary artery disease during the follow-up. Participants were followed up from enrolment to death, loss to follow-up or the end of the study (first event).

2.6. Statistical analysis

The association of the polygenic risk score with incident coronary artery disease was assessed by Cox proportional-hazards regression, with the hazard ratio expressed per standard deviation of the score and, separately, for the top compared with the bottom decile and tertile, both unadjusted and adjusted for the conventional risk factors. The incremental value of the score was quantified by the change in the Harrell C-statistic between the conventional model and the model with the score added, with a bootstrap confidence interval, and by the net reclassification improvement and the integrated discrimination improvement (Uno *et al.*, 2011; Pencina *et al.*, 2008). Calibration was assessed by comparing the predicted and observed risks across deciles of predicted risk, and cumulative incidence was estimated by the Kaplan-Meier method. The proportional-hazards assumption was examined for the polygenic score and the conventional covariates, and the association of the score with incident disease was confirmed to be consistent across the follow-up period.

3. Results

3.1. Cohort and incident disease

This group involved 6800 middle-aged and older South Asians without coronary artery disease at baseline. During a median follow-up period of 7.8 years, 631 (9.3%) of the participants developed new coronary artery disease. The level of prevalence of diabetes at the baseline was high which is consistent with the metabolic profile of the population and the age range and conventional risk range among the cohort was broad, allowing the range of risk with which a polygene score can be assessed. Nearly all of the subjects were followed up and events were determined in a similar fashion in all three parts of the score, so that differential follow-up did not distort the comparisons between the tertiles.

3.2. Polygenic risk score and baseline characteristics

When participants were grouped into tertiles of the polygenic risk score, the conventional risk factors were similar across the tertiles, as shown in Table 1. This similarity confirms that the score is largely independent of the conventional factors, measuring an inherited component of risk that they do not capture, and it is the reason the score can add predictive information beyond them.

3.3. Association with incident coronary artery disease

The polygenic risk score was strongly associated with incident coronary artery disease. In the model adjusted for the conventional risk factors, each standard deviation increase in the score raised the hazard of coronary disease (hazard ratio 1.61, 95% CI 1.49 to 1.74), and the association was graded across the deciles of the score,

Table 1: Baseline characteristics by tertile of the polygenic risk score

Characteristic	Low	Intermediate	High	p
Age (years)	52.0 ± 8.6	52.1 ± 8.8	52.0 ± 8.8	0.88
Male sex	54%	56%	56%	0.33
Diabetes mellitus	22%	20%	20%	0.29
Hypertension	35%	35%	34%	0.68
Current smoking	22%	21%	23%	0.25
LDL cholesterol (mmol/L)	3.3 ± 0.8	3.3 ± 0.8	3.3 ± 0.8	0.31
Family history of CAD	18%	18%	18%	0.92

Table 2: Adjusted hazard ratios for incident coronary artery disease

Variable	Hazard ratio	95% CI	p
Age (per year)	1.09	1.08–1.10	<0.001
Male sex	1.41	1.20–1.66	<0.001
Diabetes mellitus	2.41	2.04–2.84	<0.001
Hypertension	1.50	1.28–1.75	<0.001
Current smoking	1.70	1.43–2.01	<0.001
LDL cholesterol (per mmol/L)	1.53	1.39–1.69	<0.001
Family history of CAD	2.00	1.68–2.38	<0.001
Polygenic risk score (per SD)	1.61	1.49–1.74	<0.001

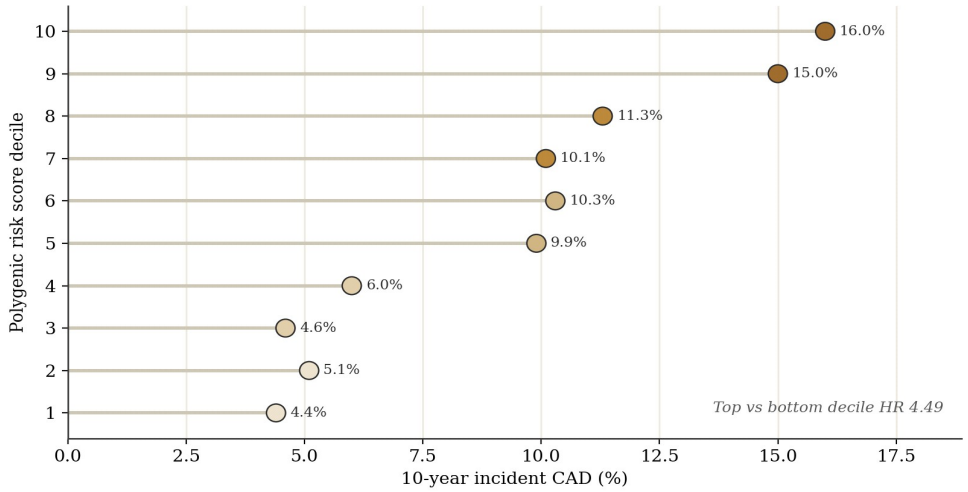


Figure 1: Incident coronary artery disease across deciles of the polygenic risk score

the incidence rising steadily from the lowest to the highest decile, in which the risk was 4.49 times that of the lowest (Figure 1). The adjusted hazard ratios for the conventional factors and the score are given in Table 2. That the score retained its association after adjustment for the conventional factors, with little attenuation, confirms that it acts independently of them.

3.4. Cumulative incidence by polygenic risk

The cumulative incidence of coronary artery disease separated clearly between the tertiles of the polygenic risk score, the curves diverging from early in follow-up and continuing to separate thereafter, so that participants in the high tertile reached an incidence several times that of the low tertile by the end of follow-up (hazard ratio for the high versus the low tertile 3.35, 95% CI 2.68 to 4.18), as shown in Figure 2. The incidence in the highest decile of the score reached 16 percent, against 4.4% in the lowest, so that the score identified, from a blood sample taken at a single point in time, a subgroup whose risk approached that at which intensive preventive treatment is unequivocally indicated.

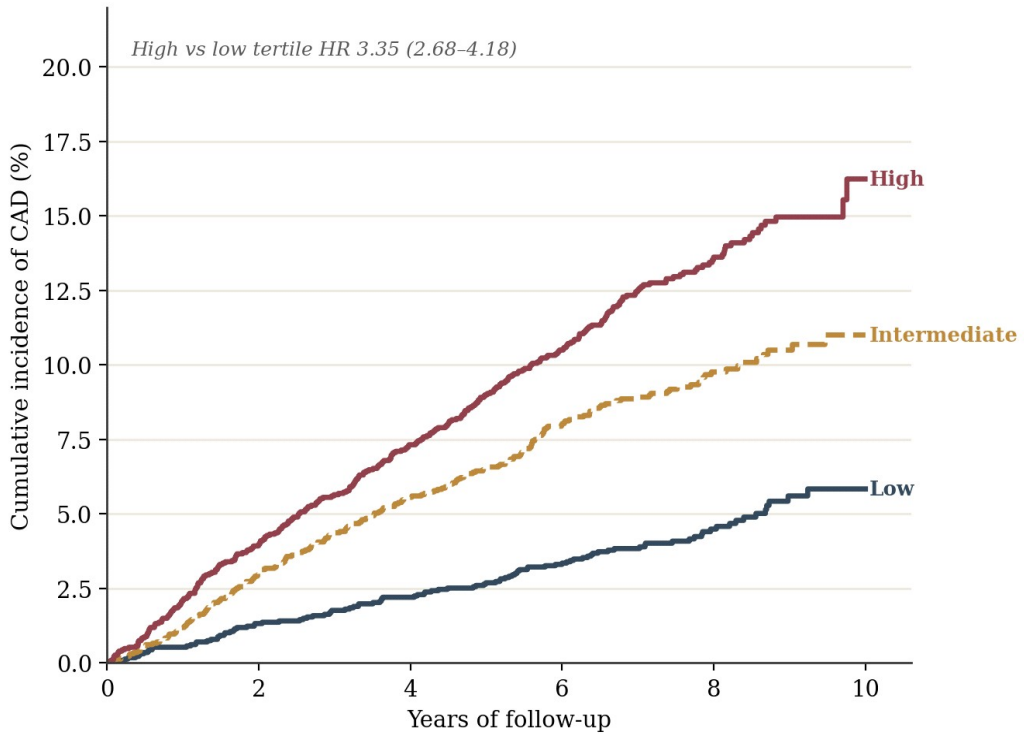


Figure 2: Cumulative incidence of coronary artery disease by tertile of the polygenic risk score

3.5. Incremental discrimination and reclassification

The polygenic risk score was found to further enhance the predictive ability of the conventional risk factors for the occurrence of the disease. As summarized in Table 3, the C-statistic increased by 0.027 (0.018 to 0.036) when the integrated model was used compared with the conventional model, the NRI was 0.084, and the IDI was 0.035. Although the magnitude of gain in discrimination—as for other cohorts—was not large, the impact on individuals at intermediate risk was more substantial: of the 2429 who were intermediate risk with the conventional model, 12.1% moved up to high-risk and 26.7% down to low-risk. This reclassification was also reflected by an integrated discrimination improvement that represents the improvement in separation of the risks predicted by the model in those who had and hadn't developed disease; the gain was not uniform across all patients but rather focused on those that crossed the threshold into treatment who are at intermediate risk of having the disease.

3.6. Calibration

The integrated model was well calibrated, the predicted ten-year risks agreeing closely with the observed incidence across the range of predicted risk, as shown in Figure 3.

Table 3: Incremental value of the polygenic risk score over conventional risk factors

Model	C-statistic	Change in C (95% CI)	NRI	IDI
Conventional risk factors	0.738	Reference	n/a	n/a
Conventional + PRS	0.765	+0.027 (0.018–0.036)	0.084	0.035

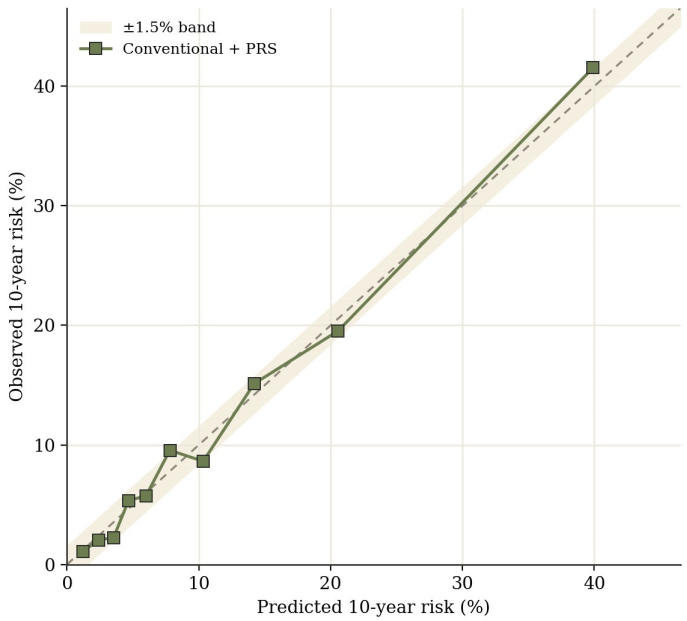


Figure 3: Calibration of the integrated risk model against the observed incidence

4. Discussion

A polygenetic risk score for coronary artery disease was a strong and independent predictor of incident disease in this prospective South Asian cohort, above and beyond conventional risk factors, and therefore was associated with a greater risk of developing coronary artery disease. The small reduction in overall discrimination is accompained by a reclassification of a sufficiently large number of individuals in the intermediate conventional risk group into high risk and a good calibration.

The hazard ratio for the score, and its residual impact after accounting for the conventional factors, aligns with the effect sizes of genome-wide scores for South Asian populations which tend to be a little smaller but are still significant, and with the overall trend that the additive risk captured by the score is largely independent of the additive risk captured by the conventional factors (Patel *et al.*, 2023). The variation from decile to decile represented as grades is consistent with and is several times the difference exhibited by the extremes for which the clinical application of the scores rely on this risk gradient. The independence of the score from the conventional factors is why the score can provide additional information because a marker that restated the conventional risk would not introduce any extra information in prediction. The absolute risk associated with the score, one should observe, in this population. This is because a lot more people have coronary disease in a population of South Asians than in one with a lower incidence: a fixed hazard ratio therefore generates a greater difference in the absolute amount of disease risk, and the same score would move people in the extremes of the range further from a fixed level of risk and hence would be more likely to move people across a fixed treatment threshold. This interplay of the relative impact of the score versus the incidence or risk of the population is part of the explanation of why the clinically meaningful reclassification observed here in the high incidence cohort did not translate into a big gain in overall discrimination.

The pattern of the improvement is more substantial for discrimination down from intermediate to less than intermediate-risk, without any corresponding change in the overall discrimination, a sort of reclassification,

which is the most common finding across the literature and has made the score particularly interesting for intermediate-risk individuals. These individuals are those in whom the decision to initiate preventive therapy is unclear; the individuals who would be shifted an appreciable fraction out of the category where treatment was not indicated would have a clear and significant clinical need for which a small change in a global measure of discrimination was inadequate value. It is particularly relevant in South Asians, where the classic score underestimated risk, and where the earlier onset of disease allows more opportunity for the risk to be reclassified resulting in altered management (Aragam *et al.*, 2020).

The importance of this reclassification lies in what can be done to those who are pushed to the high-risk category. Even if the genetic risk represented by the score is fixed, so are the consequences—the genetic risk is modifiable—and those with high polygenic risk have been found to benefit more than others from statin therapy, as well as from more intensive lipid-lowering in secondary prevention, making the identification of such individuals not just prognostic but actionable (Damask *et al.*, 2020). Reclassification to a category in which statin therapy is clearly indicated thus leads more directly to a treatment decision, and that is the only way the reclassification matters, beyond the small magnitude of change in the global discrimination. The polygen score is also complementary to the other potential method of stratification that is coronary artery calcium scanning, as each assesses a different aspect of the disease—the former the disease already present, the latter the inherited susceptibility to the disease; using one might be most advantageous when the other is available early, before calcium is likely to have reached measurable levels, to identify those at highest risk of developing the disease.

These findings should be interpreted in the light of limitations of polygen scores. Many scores were developed in populations of European ancestry, and scores developed by ancestry matched training data are likely to be more accurate than those developed in other ancestry groups, which is why the effect reported above is likely to be an underestimate of what might be achieved with a score developed with sample specific to South Asian ancestry; indeed, the development of scores trained in South Asian data is a priority. Furthermore, the score is pre-established, cannot be altered by therapeutic intervention and should be interpreted more as a look at refining the traditional scoring rather than supplanting it; hence its value is more in the "timing" and "targeting" of prevention than as a tool for monitoring.

5. Conclusion

In this prospective South Asian cohort, the polygenetically derived CAD PolyGS was independently associated with onset of CAD and when combined with the traditional risk factors, modestly and significantly enhanced prediction of CAD and reclassification of a relevant fraction of intermediate-risk individuals into the high-risk category without compromising the calibration. In South Asians, a group of people who are at high and often underestimated risk, adding a polygenetic risk score to traditional risk factors may help to tailor primary prevention targets, although scores developed using South Asian data and evidence that acting based on the score will benefit outcome are required before widespread adoption.

Ethical approval

The study was approved by the Institutional Ethics Committee and was conducted in accordance with the Declaration of Helsinki. Written informed consent, including consent for genetic analysis, was obtained from every participant.

Conflict of interest

The authors declare that no conflict of interest exists in relation to this work.

Funding

This work received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data availability

The data supporting the findings of this study are available from the corresponding author on reasonable request.

References

Aragam, K.G., Dobbyn, A., Judy, R., Chaffin, M., Chaudhary, K., Hindy, G. *et al.* (2020). Limitations of contemporary guidelines for managing patients at high genetic risk of coronary artery disease. *J Am Coll Cardiol*, 75(22): 2769-80.

Arnett, D.K., Blumenthal, R.S., Albert, M.A., Buroker, A.B., Goldberger, Z.D., Hahn, E.J. *et al.* (2019). 2019 ACC/AHA guideline on the primary prevention of cardiovascular disease. *Circulation*, 140(11): e596-646.

Damask, A., Steg, P.G., Schwartz, G.G., Szarek, M., Hagstrom, E., Badimon, L. *et al.* (2020). 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*, 141(8): 624-36.

Elliott, J., Bodinier, B., Bond, T.A., Chadeau-Hyam, M., Evangelou, E., Moons, K.G.M. *et al.* (2020). Predictive accuracy of a polygenic risk score-enhanced prediction model vs a clinical risk score for coronary artery disease. *JAMA*, 323(7): 636-45.

Goff, D.C. Jr, Lloyd-Jones, D.M., Bennett, G., Coady, S., D'Agostino, R.B., Gibbons, R. *et al.* (2014). 2013 ACC/AHA guideline on the assessment of cardiovascular risk. *Circulation*, 129(25 Suppl 2): S49-73.

Inouye, M., Abraham, G., Nelson, C.P., Wood, A.M., Sweeting, M.J., Dudbridge, F. *et al.* (2018). Genomic risk prediction of coronary artery disease in 480,000 adults: Implications for primary prevention. *J Am Coll Cardiol*, 72(16): 1883-93.

Khera, A.V., Chaffin, M., Aragam, K.G., Haas, M.E., Roselli, C., Choi, S.H. *et al.* (2018). Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet*, 50(9): 1219-24.

Mars, N., Koskela, J.T., Ripatti, P., Kiiskinen, T.T.J., Havulinna, A.S., Lindbohm, J.V. *et al.* (2020). Polygenic and clinical risk scores and their impact on age at onset and prediction of cardiometabolic diseases and common cancers. *Nat Med*, 26(4): 549-57.

Martin, A.R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B.M. and Daly, M.J. (2019). Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet*, 51(4): 584-91.

Mega, J.L., Stitziel, N.O., Smith, J.G., Chasman, D.I., Caulfield, M.J., Devlin, J.J. *et al.* (2015). Genetic risk, coronary heart disease events, and the clinical benefit of statin therapy: An analysis of primary and secondary prevention trials. *Lancet*, 385(9984): 2264-71.

Natarajan, P., Young, R., Stitziel, N.O., Padmanabhan, S., Baber, U., Mehran, R. *et al.* (2017). Polygenic risk score identifies subgroup with higher burden of atherosclerosis and greater relative benefit from statin therapy in the primary prevention setting. *Circulation*, 135(22): 2091-101.

Patel, A.P., Wang, M., Ruan, Y., Koyama, S., Clarke, S.L., Yang, X. *et al.* (2023). A multi-ancestry polygenic risk score improves risk prediction for coronary artery disease. *Nat Med*, 29(7): 1793-803.

Pencina, M.J., D'Agostino, R.B. Sr, D'Agostino, R.B. Jr and Vasan, R.S. (2008). Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. *Stat Med*, 27(2): 157-72.

Sun, L., Pennells, L., Kaptoge, S., Nelson, C.P., Ritchie, S.C., Abraham, G. *et al.* (2021). Polygenic risk scores in cardiovascular risk prediction: A cohort study and modelling analyses. *PLoS Med*, 18(1): e1003498.

Uno, H., Cai, T., Pencina, M.J., D'Agostino, R.B. and Wei, L.J. (2011). On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Stat Med*, 30(10): 1105-17.

Volgman, A.S., Palaniappan, L.S., Aggarwal, N.T., Gupta, M., Khandelwal, A., Krishnan, A.V. *et al.* (2018). Atherosclerotic cardiovascular disease in South Asians in the United States: Epidemiology, risk factors, and treatments. *Circulation*, 138(1): e1-e34.

von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gotzsche, P.C. and Vandenbroucke, J.P. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *Lancet*, 370(9596): 1453-7.

Wang, M., Menon, R., Mishra, S., Patel, A.P., Chaffin, M., Tanneeru, D. *et al.* (2020). Validation of a genome-wide polygenic score for coronary artery disease in South Asians. *J Am Coll Cardiol*, 76(6): 703-14.

Yusuf, S., Hawken, S., Ounpuu, S., Dans, T., Avezum, A., Lanas, F. *et al.* (2004). Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study). *Lancet*, 364(9438): 937-52.

Cite this article as: Nguyen T. Minh (2026). Polygenic risk score integration with conventional risk factors for primary prevention of coronary artery disease in a South Asian Cohort. *African Journal of Cardiology and Cardiovascular Medicine*. 2(1), 10-18. doi: 10.62587/AFJCCM.2.1.2026.10-18.