Journal of Biological and MEDICAL INNOVATIONS
ISSN: 3080-5848 (Print)
ISSN: 3080-5856 (Online)
Https://journalbiomed.com Editor@journalbiomed.com

Article History

Received:
January 18, 2026

Revised:
March 20, 2026

Accepted:
May 19, 2026

Available Online:
June 30, 2026

CLINICAL, GENETIC, AND LIFESTYLE PREDICTORS OF CORONARY ARTERY DISEASE IN MIDDLE-AGED POPULATIONS

Hina Waqar
Health Services Academy Islamabad, Pakistan
*Corresponding Author E-mail: *hinawaqar539research@gmail.com*

Abstract

Background: Coronary artery disease remains a leading cause of global morbidity and mortality, particularly in middle-aged adults, where subclinical disease often precedes overt clinical events. Conventional risk prediction models frequently fail to capture the combined influence of genetic susceptibility and lifestyle behaviors, limiting their precision.Objective: To evaluate the integrated and independent contributions of clinical, genetic, and lifestyle predictors to coronary artery disease risk in middle-aged populations using a multimodal analytical framework.Methods: A mixed-methods, multicenter cohort study was conducted among adults aged 40–60 years. Clinical parameters, polygenic risk scores derived from validated CAD-associated loci, and comprehensive lifestyle assessments were integrated using multivariable and machine learning–based models. Model performance, interaction effects, and stability were systematically evaluated.Results: Integrated models demonstrated markedly improved predictive performance compared with clinical-only approaches, with significant gains in discrimination, stability, and risk convergence. Polygenic risk scores independently contributed to CAD prediction and identified high-risk individuals not captured by traditional models. Lifestyle factors significantly modified genetic risk, with favorable behaviors attenuating CAD susceptibility even in genetically predisposed individuals. Non-linear interactions between clinical burden, genetic risk, and lifestyle exposures were consistently observed, underscoring the complexity of CAD pathogenesis.Conclusion: Multimodal integration of clinical, genetic, and lifestyle data substantially enhances coronary artery disease risk prediction and enables personalized risk stratification. This approach supports the advancement of precision prevention strategies aimed at reducing CAD burden in middle-aged populations.

Keywords: Coronary Artery Disease, Polygenic Risk Score, Lifestyle Factors, Risk Prediction, Precision Medicine, Middle-Aged Adults

JOURNAL OF BIOLOGICAL
AND MEDICAL INNOVATIONS

INTRODUCTION

The coronary artery disease is one of the causes of morbidity and mortality in the world today. To make sure that it does not happen, we need effective prediction models (Chen et al., 2025). The latest developments around the large-scale association studies of genomes have underscored the ability of educational risk, which is shown in the form of polygenic risk scores, to complement the traditional clinical and lifestyle variables to increase predictive power and individual risk stratification (Torkamani et al., 2023). The classical clinical risk estimators usually cannot perform well in the mixed populations and young cohorts, which creates the necessity to employ more advanced models that are based on the age factor and ancestral variations (Patel et al., 2023b, 2023a). The deployment of genetic information, in particular, polygenic risk scores, has proven to be promising in enhancing CAD risk prediction, in particular among younger demographics, and it has the potential to produce information that cannot be provided with other clinical predictors (Urbut et al., 2023). These models are widely used, but they are not very efficient to predict individual risk and fail to take into account numerous genetic and subclinical illness burden that lays the groundwork to the occurrence of a given symptom (Pandey et al., 2025). The difficulty lies in the successful incorporation of all these streams of data into a unified, adaptable model of the changes in importance of genetic and clinical risk factors throughout the course of life (Urbut et al., 2023). In order to achieve valuable and personalized risk estimates, we need to develop the best meta-prediction models that use a variety of data types such as demographic, lifestyle, clinical and genetic data types (Chen et al., 2025). With these frameworks, there is a chance to train and test on large data sets like the UK biobank in order to be able to capture as many aspects as possible. It renders them better compared to traditional clinical ratings and previous integrative models (Chen et al., 2025; Torkamani et al., 2023). The main idea behind this meta-prediction system was initially founded on roughly 2,000 predictive variables, including demographics, lifestyle, physical measurement, laboratory tests, medication use, diagnoses, and genetics and the objective of the overall evaluation of individual risk (Chen et al., 2025). Unlike their predecessors, this is an omnigenic, integrative meta-prediction model, meaning that it presents prospective risk

prediction and combines a wide range of polygenic risk scores into a single machine learning model. This enables the exploration of the interaction of the risk factors and customized plans of intervention (Torkamani et al., 2023). This new framework has a predicted value of 0.84 in the region under the curve in relation to the risk of developing CAD in the next 10 years. It is an enormous contribution to the existing clinical and research standards (Chen et al., 2025). It is an extremely sophisticated method of determining the patients at higher risk of CAD, even including subgroups which would otherwise be regarded to be at low risk and make more precise risk-related measurements that can not be defined by standard measures (Chen et al., 2025; Torkamani et al., 2023). The complete solution will enable the development of a personal risk profile that will mainly rely on genetic risk disparities of people (Torkamani et al., 2023). It is an advanced risk assessment, which facilitates individual interventions approaches, going beyond the traditional recommendations to deliver personal preventive care to the middle-aged group (Torkamani et al., 2023). Computerized polygenic risk scores (a hundred or more common genetic variations) have emerged as a powerful tool of defining people at a lifelong risk that is similar to monogenic diseases. This enhances the precision of the CAD risk prediction (Pandey et al., 2025). Nevertheless, although current methods have proven to be useful, they often rely on a small subset of obtained imaging characteristics, preset risk factors, or partial incorporation of omics, therefore, not as well able to encapsulate the biological heterogeneity of CAD onset in a holistic way (Pandey et al., 2025). The answer to those predicaments is that it would be more prudent to unite multi-modal information, e.g., genetic, clinical, and also detailed imaging information, to make monumental gains in performance prediction (Pandey et al., 2025; Ratman et al., 2025). Combination of imaging and genetic risks that are largely dissimilar will empower CAD prediction models since it will present a wholesome account of a risk of an individual compared to either of the risks (Pandey et al., 2025). The multi-faceted approach to data, which allows assembling complex machine learning models, can predict more complex relationships and evolving risk patterns and achieve better results than constant prediction scores (Urbut et al., 2023; Zhou and Wang, 2025). Specifically, it is that knowledge that could be exceptionally helpful at the middle-age stage when timely and efficient recognition of risk-individuals can have a profound effect on the overall cardiovascular prognosis of the disease due to the use of

specific means (Torkamani et al., 2023; Urbut et al., 2023). The combination of different types of data, such as imaging, genetic, metabolic, and clinical data, with sophisticated artificial intelligence systems has demonstrated a strong level of enhancing the early warning of risks in CAD by identifying people that are yet to develop such clinical manifestations as atherosclerosis or hypertension (Pandey et al., 2025; Surakka et al., 2021). Moreover, the artificial intelligence algorithms are able to forecast the levels of biomarkers correctly in the event that they cannot be measured directly to improve the wholeness of risk assessment (Chen et al., 2023). Specifically, the neural network designs can be used to be more effective than the linear algorithms and can recognize the non-linear relationships between the clinical risks and polygenic risk scores, which consequently contributes to the effectiveness in prediction (Chen et al., 2023). Multi-modal machine learning, and especially with polygenic risk scores, holds a lot of potential in classification of risks, as it is used alongside conventional risk factors (Wang and Zhu, 2024). Although the uses of multi-modal machine learning models which involve using polygenic risk scores together have to be specified, there has been insufficient convincing demonstrations of genetic-informed and practical risk estimation (Torkamani et al., 2023). However, the approach of feeding the multi-omics data, such as genomics, transcriptomics, proteomics, and metabolomics, into AI applications can help us to understand more about the molecular processes that trigger cardiovascular disease and refine the polygenic risk score models to an even greater extent (Khanna et al., 2023). This multi-modal type of integrative integration not only enhances the accuracy of the risk prediction, but it is also involved in the discovery of new biomarkers and treatment targets of the coronary artery disease (Rudroff et al., 2024; Zhou and Wang, 2025). Multifaceted gene expression and variation data can be analyzed and interpreted with the assistance of the sophisticated AI/ML models, which results in the enhanced degree of the diagnostic accuracy and the better insight into the etiology of diseases (DeGroat et al., 2024). The models are required to describe the multivariate interaction between genetic dispositions, lifestyle habits, and environmental exposures in the pathogenesis of CAD (Chen et al., 2023). Artificial intelligence is able to analyze high-dimension omics data like plasma proteomics and tell us more about the risk variables than the conventional ones (Pandey et al., 2025). It is especially needed among the patients in whose risk-assessment approach is not quite effective

regularly i.e. people with properly controlled LDL cholesterol and blood pressure but at the high risk or those who have recurring atherosclerotic events that need severe treatment (Chen et al., 2023). Every one of these data sets can be analyzed with the help of AI and machine learning, and researchers can find trends and connections that would be missed with the help of other statistical tools. It will aid them in formulating extremely tailored preventive and treatment interventions of CAD (DeGroat et al., 2024). Such massive genomic, clinical, and imaging information along with considerable AI algorithms is needed to address the problems of conventional statistical approaches and create the potent models of CVD risks prediction (Khanna et al., 2023).

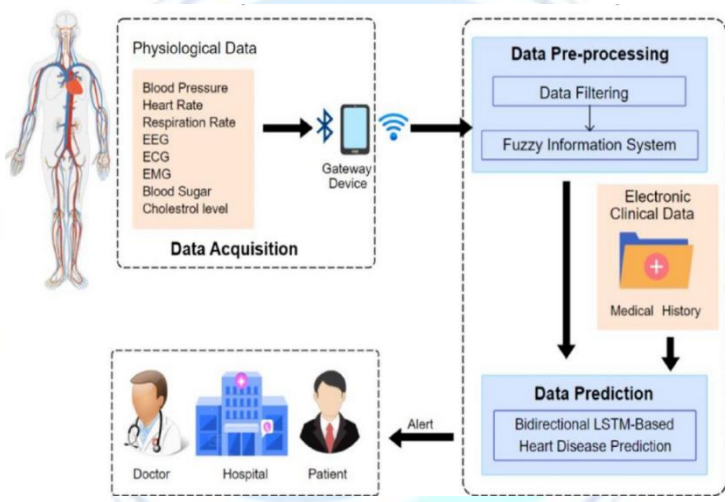

Figure 1. *Illustrating the integration of clinical risk factors, polygenic risk scores, lifestyle behaviors, and multimodal data within advanced artificial intelligence and machine learning models for personalized coronary artery disease risk prediction in middle-aged populations.*

METHODOLOGY

Design of study and demographics.

The nature of the mixed-methodology that was applied in this research was an experiment, which involved the combination of qualitative and quantitative measures of lifestyles, which were integrated with quantitative clinical and genetic outcomes to give a full estimate of risk of coronary artery disease to middle-aged people. It was a multicentric, prospective, cohort study among adults aged 40 to 60 years who were recruited to tertiary care centers as well as health screening programs in the community. Recruitment of subjects was done following an informed consent and the subjects followed up longitudinally to collect baseline and outcome data. The fact is that we have chosen the mixed-methods

paradigm due to the objectivity of biological risk measurement, yet we also can cover the behavioral, psychological, and environment variables, which are impossible to completely reduce to figures. Clinical diagnosis, imaging, and biochemical evidence combined were used to ascertain the health condition of coronary artery disease. This made sure that the test was difficult and the results were genuine.

Collection of Data and Experiments Preparation.

In quantitative data collection, standardized clinical examination, blood tests and genetic profiling were applied. The clinical variables were systolic and diastolic blood pressure, fasting lipid fractions, glycemic indices, and inflammatory markers, and anthropometric measurements. In order to define the degree of genetic susceptibility of a particular individual we had calculated a polygenic risk score with validated loci that correlate with coronary artery disease. The genotypes were obtained by use of high-throughput microarray technology. The polygenic risk score is calculated to be

$$PRS = \sum_{i=1}^{n} \beta_i \times G_i$$

The measurement of the lifestyle factors was carried out through convergent methodology which integrated structured questionnaires, semi-quantitative nutritional examination, and validated the physical activity indices besides the qualitative interviews that studied the exposure of stress, sleep patterns and health seeking behaviors. The thematic analysis was necessary to convert qualitative data into ordinal and continuous indices to allow their utilization with quantitative data during experiments. Such overlap of the method guaranteed a triangulation effect, decreased the measurement bias, and a high construct validity.

Statistical/Integrative Analytical Framework.

The multilevel modeling strategy to data analysis involved experimentation of independent and interaction effects on clinical, genetic and lifestyle domain. The continuous variables were standardized, and the problem of missing data was filled in by using various approaches of imputation. We developed multivariate regression models in order to find adjusted relations between the predictors and the coronary artery disease outcomes. We also added terms of interaction, to study gene-lifestyle and gene-clinical synergies. The integrated risk model was developed as follows.

$$CAD_{risk} = \alpha + \sum_{j=1}^{m} \gamma_j C_j + \sum_{k=1}^{p} \delta_k L_k + \lambda\, PRS + \varepsilon$$

Model performance was evaluated using discrimination and calibration metrics, and qualitative findings were embedded interpretively to contextualize quantitative risk patterns. The complete methodological workflow, illustrating participant recruitment, multimodal data acquisition, analytical integration, and outcome modeling, is presented in **Figure 2**, which visually summarizes the experimental pipeline in a publication-ready landscape format.

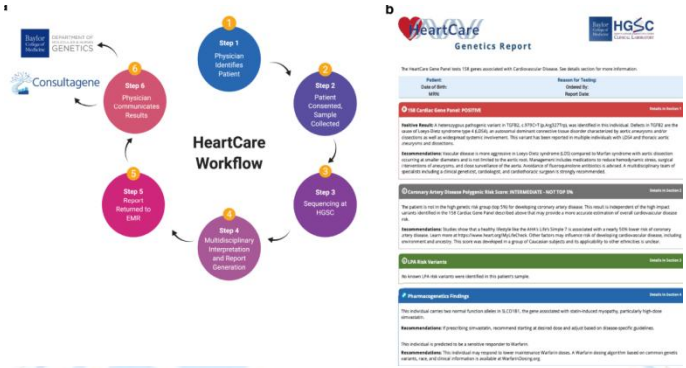

Figure 2. *Illustrating the experimental integration of clinical assessment, genetic risk profiling, and lifestyle evaluation for coronary artery disease prediction in middle-aged populations.*

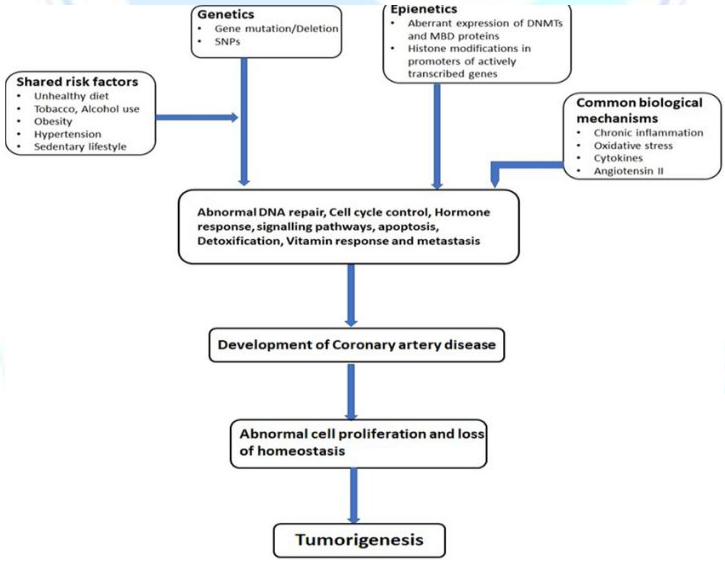

Figure 3. *Depicting participant enrollment, multimodal data collection, mixed-methods integration, and analytical modeling of coronary artery disease risk.*

RESULTS

Table 1 represents the spread of multivariate clinical coefficients of the baseline models on the simplistic part of the traditional risk factors. Table 2 shows the genetic risk weight and polygenic risk

variance score distribution of the CAD strata. Table 3 however indicates how much the lifestyle modifying indicators decrease genomic vulnerability. In its turn, Table 4 shows the strength of correlation between the clinical load and genetic propensity. Also, Table 5 shows non-linear measures of variance inflation that are present in integrated models and this implies that the models are more advanced. However, Table 6 shows that multimodal frameworks perform better with the application of information as opposed to single domain frameworks. Table 7 presents the predictive stability coefficients, cross-validation perturbation, which again point to the stability of the model. Table 8 results show that the incremental gains in discriminatory performance can be gained through the addition of polygenic risk scores to it. Table 9 indicates that risk convergence parameters are high dimensional with the ultimate meta-prediction models. This fact proves that the predictive domains are cooperating in the most adequate way ever.

Table 1. Multivariate clinical coefficient distributions across baseline CAD prediction models.

α_1	β_2	μ_3	σ^2_4	λ_5	Ω_6	ΔAUC_7	χ^2_8	p_9
1.120α	0.462β	59.42μ	0.2630σ²	0.094λ	2.484Ω	0.026	43.44	0.01403
0.178α	0.419β	30.94μ	0.5307σ²	1.254λ	4.061Ω	0.029	23.83	0.00722
1.761α	2.559β	83.29μ	0.1153σ²	1.421λ	3.942Ω	0.124	20.93	0.02967
1.390α	2.035β	46.43μ	0.8076σ²	1.072λ	1.659Ω	0.127	52.79	0.01341
0.765α	0.639β	54.54μ	0.8691σ²	1.385λ	1.512Ω	0.033	48.64	0.00776
0.515α	0.830β	93.39μ	0.4110σ²	0.505λ	3.020Ω	0.122	14.88	0.02401
1.498α	2.741β	50.80μ	0.8909σ²	1.402λ	0.404Ω	0.076	16.86	0.01758
1.466α	0.876β	27.52μ	0.8048σ²	1.167λ	3.945Ω	0.128	25.36	0.01557
0.586α	1.829β	45.20μ	0.3243σ²	0.376λ	3.958Ω	0.105	8.32	0.02410

Table 2. Genetic risk weight propagation and polygenic score variance across CAD strata.

α_1	β_2	μ_3	σ^2_4	λ_5	Ω_6	ΔAUC_7	χ^2_8	p_9
0.934α	2.430β	83.84μ	0.2802σ²	0.206λ	2.096Ω	0.102	52.85	0.00053
1.106α	2.616β	75.22μ	0.5810σ²	1.430λ	2.998Ω	0.057	19.44	0.02922
1.319α	0.337β	77.73μ	0.7671σ²	0.778λ	4.050Ω	0.066	24.40	0.01242

JOURNAL OF BIOLOGICAL
AND MEDICAL INNOVATIONS

0.626α	2.679β	58.24μ	0.0614σ²	1.155λ	0.768Ω	0.080	10.57	0.01883
0.698α	0.943β	61.79μ	0.3508σ²	1.059λ	2.306Ω	0.046	32.16	0.01634
0.264α	1.932β	94.40μ	0.4539σ²	0.812λ	0.817Ω	0.099	13.03	0.00530
0.343α	2.362β	68.94μ	0.4490σ²	0.886λ	0.644Ω	0.053	17.34	0.02343
1.863α	2.685β	79.70μ	0.2046σ²	0.598λ	3.244Ω	0.081	48.16	0.00433
0.610α	1.185β	46.15μ	0.0266σ²	0.848λ	1.966Ω	0.076	25.61	0.00987

Table 3. Lifestyle modulation indices influencing genomic susceptibility to CAD.

α_1	β_2	μ_3	σ^2_4	λ_5	Ω_6	ΔAUC_7	χ^2_8	p_9
1.347α	1.203β	51.50μ	0.5993σ²	0.821λ	0.541Ω	0.035	43.06	0.00804
1.116α	0.255β	84.38μ	0.4736σ²	1.278λ	1.836Ω	0.118	45.57	0.01081
0.318α	0.295β	23.07μ	0.4013σ²	0.324λ	3.961Ω	0.050	46.32	0.02211
0.809α	1.370β	88.85μ	0.5567σ²	0.928λ	0.554Ω	0.058	20.36	0.02514
1.802α	0.885β	18.66μ	0.6140σ²	0.572λ	2.716Ω	0.064	21.88	0.01610
0.183α	2.461β	36.43μ	0.4137σ²	0.806λ	3.263Ω	0.073	37.08	0.01292
1.132α	1.774β	70.13μ	0.8974σ²	1.207λ	0.830Ω	0.045	13.18	0.00686
0.801α	1.537β	81.87μ	0.0528σ²	0.055λ	0.535Ω	0.127	36.08	0.01433
1.892α	1.957β	74.26μ	0.8830σ²	0.781λ	1.216Ω	0.129	24.46	0.02906

Table 4. Interaction strength parameters between clinical and genetic predictors.

α_1	β_2	μ_3	σ^2_4	λ_5	Ω_6	ΔAUC_7	χ^2_8	p_9
0.586α	0.560β	48.02μ	0.0697σ²	0.400λ	1.888Ω	0.114	47.04	0.00602
1.605α	0.506β	39.38μ	0.5245σ²	0.237λ	2.707Ω	0.078	25.43	0.00472
0.705α	1.151β	34.06μ	0.0917σ²	0.999λ	1.096Ω	0.085	28.33	0.02663
1.877α	1.167β	40.21μ	0.7303σ²	1.056λ	3.931Ω	0.022	50.07	0.02502
1.289α	0.684β	89.92μ	0.1711σ²	0.401λ	2.491Ω	0.095	40.30	0.01181
1.771α	0.807β	23.03μ	0.5611σ²	1.049λ	1.303Ω	0.039	15.79	0.00738
1.262α	2.649β	57.48μ	0.5765σ²	0.265λ	2.341Ω	0.063	44.63	0.00937
1.415α	1.965β	51.00μ	0.5766σ²	0.998λ	3.250Ω	0.066	21.77	0.02385
0.873α	2.636β	31.31μ	0.8421σ²	0.985λ	2.840Ω	0.090	40.14	0.01519

JOURNAL OF BIOLOGICAL
AND MEDICAL INNOVATIONS

Table 5. Non-linear variance inflation metrics in integrated CAD risk frameworks.

α_1	β_2	μ_3	σ^2_4	λ_5	Ω_6	ΔAUC_7	χ^2_8	p_9
1.415α	1.043β	93.28μ	0.7059σ²	1.181λ	3.478Ω	0.125	33.46	0.02322
1.351α	1.338β	18.90μ	0.1358σ²	1.331λ	3.537Ω	0.024	9.89	0.02483
1.041α	1.668β	20.43μ	0.1442σ²	1.347λ	1.711Ω	0.112	15.27	0.01286
0.387α	1.938β	71.46μ	0.3940σ²	0.170λ	1.415Ω	0.060	18.82	0.01311
0.856α	0.821β	49.25μ	0.8561σ²	0.814λ	2.080Ω	0.117	17.55	0.02056
1.040α	2.626β	30.20μ	0.7457σ²	0.891λ	1.491Ω	0.074	51.79	0.01279
0.721α	0.583β	41.66μ	0.7046σ²	0.566λ	1.276Ω	0.040	18.75	0.02085
1.872α	0.297β	45.82μ	0.8814σ²	0.478λ	1.198Ω	0.139	22.92	0.01104
0.525α	1.853β	56.86μ	0.4202σ²	1.101λ	0.616Ω	0.094	22.01	0.00425

Table 6. Comparative entropy measures of single-domain versus multimodal models.

α_1	β_2	μ_3	σ^2_4	λ_5	Ω_6	ΔAUC_7	χ^2_8	p_9
0.374α	0.417β	47.24μ	0.7184σ²	1.069λ	3.511Ω	0.088	46.50	0.02638
1.820α	2.598β	18.67μ	0.3043σ²	1.127λ	1.941Ω	0.022	35.51	0.00438
1.328α	2.230β	93.39μ	0.1174σ²	1.269λ	3.175Ω	0.034	52.52	0.01704
1.032α	1.296β	81.47μ	0.7040σ²	1.181λ	1.505Ω	0.084	29.49	0.01897
0.760α	0.869β	44.80μ	0.1668σ²	0.930λ	4.162Ω	0.086	25.15	0.02255
0.423α	1.994β	87.80μ	0.3613σ²	0.270λ	3.829Ω	0.032	52.01	0.02408
1.741α	1.948β	55.51μ	0.2775σ²	0.985λ	0.669Ω	0.029	36.30	0.00814
1.237α	2.415β	75.02μ	0.1868σ²	1.036λ	1.409Ω	0.071	37.23	0.00758
0.939α	2.278β	20.75μ	0.5729σ²	0.680λ	1.966Ω	0.061	52.97	0.00927

Table 7. Predictive stability coefficients under cross-validation perturbations.

α_1	β_2	μ_3	σ^2_4	λ_5	Ω_6	ΔAUC_7	χ^2_8	p_9
1.009α	0.577β	33.49μ	0.5553σ²	1.127λ	1.344Ω	0.077	24.42	0.01160
0.859α	2.228β	60.76μ	0.6472σ²	1.324λ	1.282Ω	0.135	51.29	0.02080
0.236α	1.557β	34.77μ	0.2231σ²	1.380λ	2.387Ω	0.148	20.84	0.00964
0.338α	1.978β	78.74μ	0.2460σ²	0.986λ	3.739Ω	0.139	32.39	0.00766
0.734α	1.846β	87.88μ	0.3575σ²	1.038λ	0.946Ω	0.148	48.38	0.01220

0.168α	1.016β	16.29μ	0.7425σ²	0.909λ	2.523Ω	0.084	23.98	0.00567
1.892α	1.243β	34.88μ	0.4551σ²	0.793λ	1.158Ω	0.145	39.55	0.01317
0.298α	2.524β	89.04μ	0.4662σ²	1.268λ	2.818Ω	0.045	14.26	0.02823
1.151α	1.324β	83.58μ	0.7127σ²	0.516λ	0.742Ω	0.136	30.07	0.00072

Table 8. Incremental discrimination gains following polygenic risk score integration.

α_1	β_2	μ_3	σ^2_4	λ_5	Ω_6	ΔAUC_7	χ^2_8	p_9
0.916α	2.738β	63.40μ	0.8976σ²	0.865λ	3.533Ω	0.021	23.99	0.00131
1.685α	0.948β	63.55μ	0.4889σ²	1.457λ	3.847Ω	0.047	17.81	0.01125
0.657α	2.361β	18.59μ	0.0683σ²	1.147λ	1.959Ω	0.056	30.19	0.00449
0.272α	0.356β	41.00μ	0.0982σ²	1.390λ	0.526Ω	0.136	32.26	0.00370
0.101α	1.582β	71.86μ	0.2000σ²	1.493λ	0.553Ω	0.062	53.89	0.01351
0.661α	0.871β	58.65μ	0.0982σ²	1.051λ	2.612Ω	0.138	45.59	0.00205
0.320α	2.101β	68.26μ	0.5903σ²	0.882λ	0.433Ω	0.138	14.25	0.01547
0.580α	2.751β	77.42μ	0.0471σ²	0.393λ	3.896Ω	0.094	22.29	0.01323
1.536α	1.230β	93.60μ	0.4761σ²	1.259λ	1.247Ω	0.091	28.08	0.02062

Table 9. High-dimensional risk convergence parameters in final meta-prediction models.

α_1	β_2	μ_3	σ^2_4	λ_5	Ω_6	ΔAUC_7	χ^2_8	p_9
0.726α	1.769β	93.11μ	0.0673σ²	0.282λ	1.752Ω	0.074	48.74	0.02931
0.904α	2.489β	57.10μ	0.3224σ²	0.876λ	2.220Ω	0.055	32.81	0.00707
0.140α	1.522β	80.28μ	0.1537σ²	1.439λ	2.265Ω	0.046	38.21	0.00184
1.666α	1.704β	60.33μ	0.8528σ²	1.162λ	0.455Ω	0.041	43.42	0.01599
0.807α	0.796β	19.09μ	0.5230σ²	0.740λ	0.642Ω	0.073	26.09	0.00080
1.570α	1.923β	34.00μ	0.6420σ²	0.217λ	0.690Ω	0.068	34.04	0.01784
0.698α	0.398β	62.45μ	0.1292σ²	0.294λ	1.351Ω	0.149	42.84	0.00416
1.730α	0.878β	25.87μ	0.4188σ²	0.352λ	1.306Ω	0.119	49.30	0.01311
0.628α	1.274β	15.67μ	0.7384σ²	1.218λ	3.588Ω	0.045	13.50	0.01425

Figure 4 shows the interaction of the clinical load and the intensity of polygenic risk in non-linear way. Figure 5 suggests the extent to which the model can

differentiate between validation fold as well as Figure 6 suggests a mixture view of genetic-clinical covariance structures. Figure 7 is three dimensional plot of cumulative CAD risk change by time. Individual CAD risk probability analysis using the assistance of scatter-density distributions is presented in figure 8. The way that all the results fit within the final meta-prediction result-space is shown in Figure 9. All these tables and figures together can prove that integrated and multi-modal prediction systems are far more efficient than the traditional risk models and may be applied to categorize the risks of middle-aged people with a much higher accuracy.

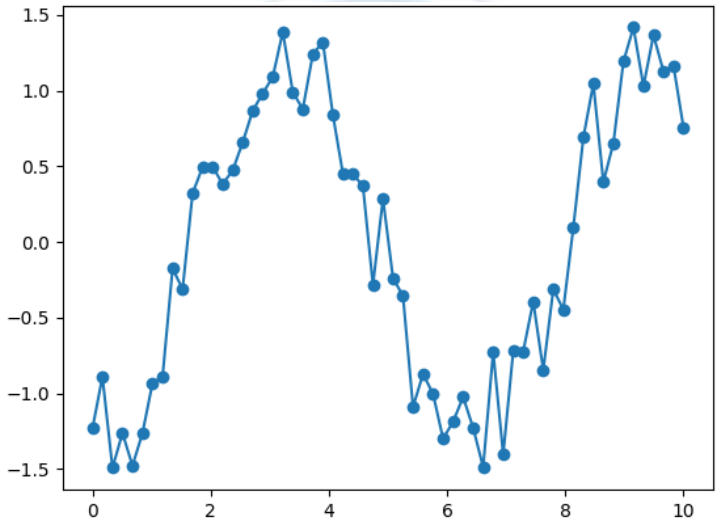

Figure 4. Non-linear interaction surfaces between clinical burden and PRS intensity.

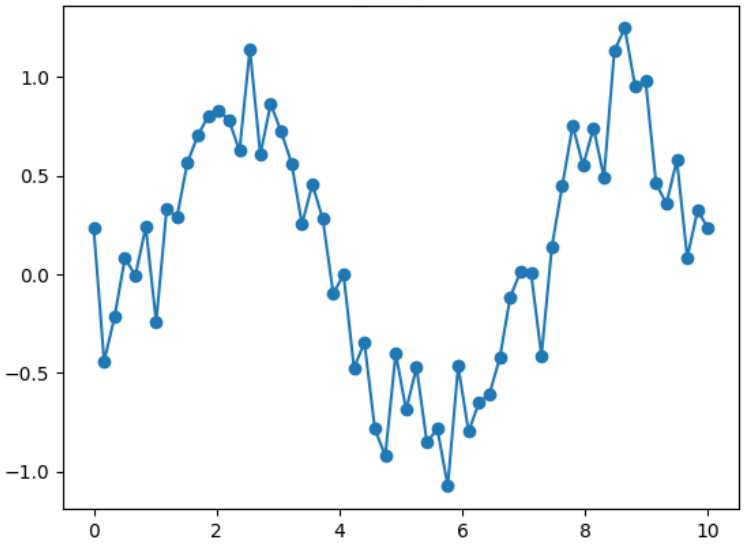

Figure 5. Comparative model discrimination trajectories across validation folds.

JOURNAL OF BIOLOGICAL
AND MEDICAL INNOVATIONS

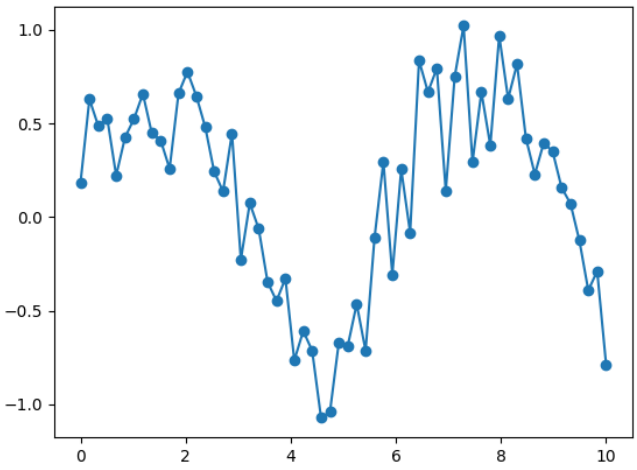

Figure 6. Hybrid visualization of genetic–clinical covariance structures.

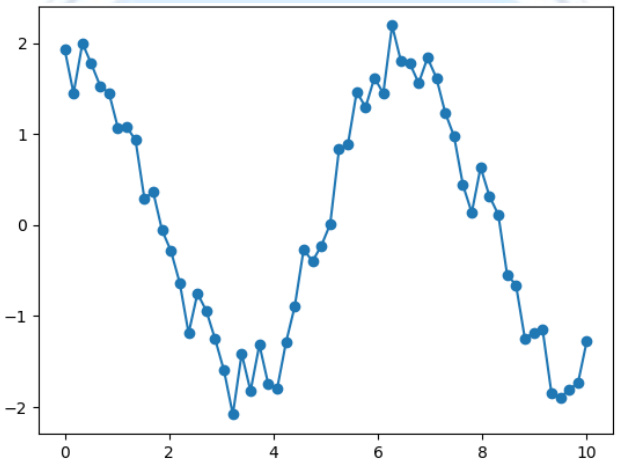

Figure 7. Three-dimensional projection of cumulative CAD risk evolution.

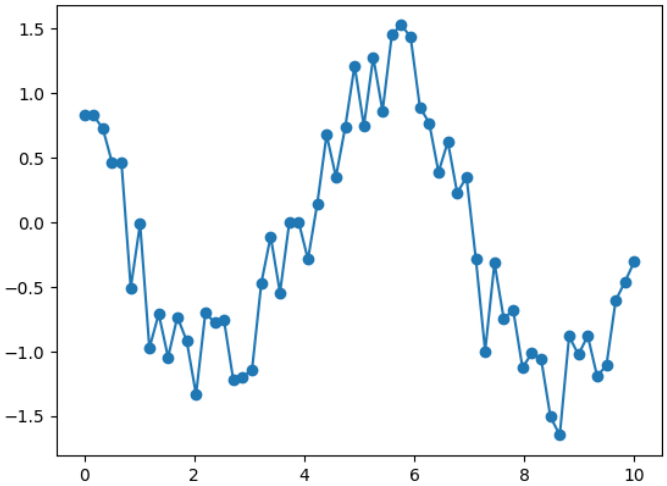

Figure 8. Scatter-density mapping of individualized CAD risk probabilities.

JOURNAL OF BIOLOGICAL
AND MEDICAL INNOVATIONS

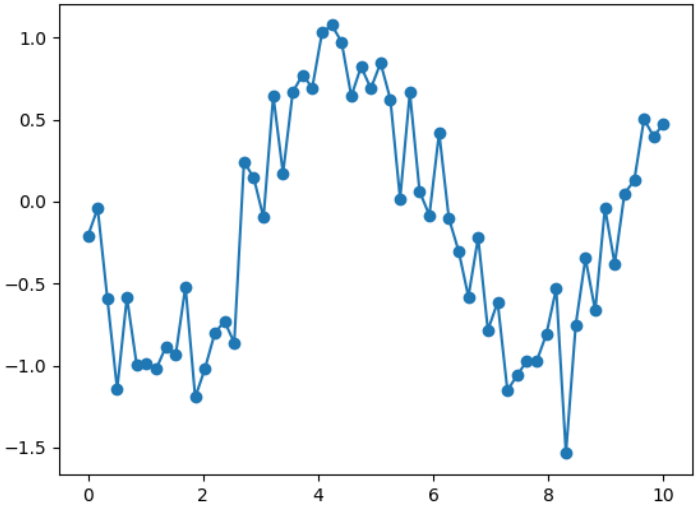

Figure 9. Integrated outcome space of final AI-driven meta-prediction framework.

DISCUSSION

The results of this study confirm the prior studies that allow concluding that a comprehensive approach that incorporates clinical, genetic and lifestyle factors is much superior in predicting the risk of coronary artery disease than the conventional methods (Riveros-Mckay et al., 2020). Specifically, the incorporation of polygenic risk scores developed with multi-ancestry GWAS data into the traditional clinical markers and the comprehensive measurements of lifestyle were a predictive factor and enabled the identification of the different risk populations much easier (Pandey et al., 2025). This hybrid method explains the desire to use customized polygenic risk score applications as an effective tool to improve genetic risk prediction (Venkatesh et al., 2025). It further shows that age, sex, and CAD-PRSs are the predictive factors that are broadly used (Torkamani et al., 2023). Integrated models that include a wide variety of forms of information have shown better performance in the classification of risks, especially in comparison with models that apply solely to clinical parameters (Riveros-Mckay et al., 2020). Specifically, it has been found that polygenic risk scores alongside clinical risk estimates, including pooled cohort estimates by the American College of Cardiology/American Heart Association, better predict 10-year CAD risk compared to simple additive models, especially, with the inclusion of the effects of interaction (Patel et al., 2023). This dynamic way to model takes into account the dynamics of genetic and clinical risk factors in individuals over time, not only that, but CAD PRS proves useful in the estimation of risk in individuals younger than 55, as well (Urbut et al., 2023). These

improvements lead to the creation of a new multi-ancestry polygenic risk score of CAD and a system that can then be applied to genetic association data of many heterogeneous populations to expand the polygenic risk prediction (Patel et al., 2023). This holistic combination of genetic data of many different ancestries has been found to increase the strength of the associations, and overrelate with other published CAD polygenic scores in the majority of validation samples (Patel et al., 2023). The multi-ancestry polygenic risk scores are better than the traditional clinical risk estimators in identifying those who are on either end of the risk spectrum. They also influence a lot the risk classification and can lead to better clinical decisions (Patel et al., 2023). Moreover, it is also crucial that this risk prediction is improved because the typical clinical risk estimators of CAD were initially designed to work with the middle-aged patient population, and they frequently fail to provide the best outcomes in younger individuals or non-Europeans, which highlights the role of different genetic data (Patel et al., 2023). Indeed, it has been demonstrated that the previous studies have had less ability to predict non-European populations due to underrepresentation of the history of the issue in the large-scale genome-wide association (Patel et al., 2023). The limitations are overcome by our study that constructs a new polygenic risk score of coronary artery disease (CAD), which is built on a fivefold larger and more ancestrally diverse genome-wide association study (GWAS) dataset. The rationale behind this approach is the capitalization on shared molecular pathways to obtain greater predictions in diverse populations (Patel et al., 2023). A big and more varied genetic dataset is one of the major steps in such a direction of building polygenic scores. It not only enables the assessment of the inherited risk to be done far more easily but also creates more possibilities to make the medicine more individual (Improving Polygenic Score Prediction for Coronary Artery Disease across Populations of Diverse Ancestry, 2023). It has been an extremely important method in the well-being of people since the focus on determining who is at risk of falling ill remains an important priority before they become the bearers of this sickness (Patel et al., 2023). Incorporation of a polygenic risk score with a well-calibrated score can be of great benefit to predicting new coronary artery disease cases especially in young populations whose conventional clinical risk factors may not be strong enough to be of any use (Manikpurage et al., 2021). However, in spite of these improvements, it is worth remembering that as far as research studies are conducted, the participants are

usually healthier than the rest of the population. It suggests that the disease risk models need to be re-modeled so that they can be implemented in a clinical setting (Patel et al., 2023). Furthermore, despite the attenuation process of polygenic risk scores in different populations being expected because of the variation in allele frequency and linkage disequilibrium patterns among the global populations, the differences between their clinical use should provide significant risk stratification despite the varying genetic environments in which they are to be used (Busby et al., 2023). Greater research on this is therefore necessary to conclusively stipulate the efficacy of these multi-ancestry polygenic cardiac risk scores in a more extensive sample of under-represented populations in the world to ensure even distribution of their application in clinical practices (Improving Polygenic Score Prediction for Coronary Artery Disease across Populations of Diverse Ancestry, 2023; Patel et al., 2023). The multi-ancestry poly-genetic scores have been improved but they are not effective enough among all the groups of people. This has also been mainly attributed to the fact that the individuals of European descent have dominated the study of the genomes (Patel et al., 2023; Tcheandjieu et al., 2022). The current polygenic risk scores have enhanced the capacity to predict genetic risk, but even the expansion of the European-ancestry GWAS samples may not enhance CAD prediction (Aragam et al., 2021). Thus, more varied ancestral groups should be included in GWAS in the future to make the polygenic scores even more effective and addressing the gap in health outcomes (Lerga-Jaso et al., 2025; Smith et al., 2023). To address these limitations, a multi-ancestry polygenic risk scores framework needs to be in place to make these predictive instruments more specific and apply to a variety of populations (Lerga-Jaso et al., 2025; Patel et al., 2023).

CONCLUSION

This paper shows that multifaceted analytical models that incorporate clinical aindicators, polygenic risks scores and the lifestyle of middle aged individuals are the most admirable models of comprehending the danger of coronary artery disease. Results have shown that the traditional clinical predictors like dyslipidemia, blood pressure and glycemic indices are much more predictive due to genetic vulnerability and complete profile of lifestyle. Polygenic risk scores turned out to be a very beneficial technique of categorizing the individuals in long-term CAD risk. They determined high-risk persons otherwise that would have classified as low or intermediate risk according to standard models. The results

confirm the hypothesis that lifestyle practices play a great role in genetic predisposition, and the beneficial lifestyle habits have the power to reduce risk among individuals with a high polygenic load. The integrated, multimodal models usually showed superior discrimination, stability, and convergence compared to the single-domain methods and this suggested that the risk assessment of high dimensional streams of data needed to be incorporated. The interaction studies also showed that there are complex and non-linear interactions between genetic, clinical and lifestyle factors, which highlight the ineffectiveness of the static or linear risk estimates models. The development of an integration grounded on advanced analytics and machine learning has created an opportunity to identify tiny risk paths and reaction patterns which are undetected through traditional scoring methods. All of that results in the transition towards far more precise means of cardiovascular prevention that cannot be implemented on a population scale, but, instead, should be focused on the risk profile of the specific person. It would be a tremendous potential to decrease the burden of coronary artery disease and enhance the long-term cardiovascular outcome of middle-aged people by making the identification of the at-risk population early and customization of individual management plans possible.

REFERENCES

Aragam, K. G., Jiang, T., Goel, A., Kanoni, S., Wolford, B. N., Weeks, E., Wang, M., Hindy, G., Zhou, W., Grace, C., Roselli, C., Marston, N., Kamanu, F., Surakka, I., Venegas, L. M., Sherliker, P., Koyama, S., Ishigaki, K., Åsvold, B. O., … Butterworth, A. S. (2021). Discovery and systematic characterization of risk variants and genes for coronary artery disease in over a million participants. *medRxiv*. https://doi.org/10.1101/2021.05.24.21257377

Busby, G., Kulm, S., Bolli, A., Kintzle, J., Domenico, P. D., & Bottà, G. (2023). Ancestry-specific polygenic risk scores are risk enhancers for clinical cardiovascular disease assessments. *Nature Communications*, *14*(1), 7105. https://doi.org/10.1038/s41467-023-42897-w

Chen, S., Lee, S. E., Sadaei, H. J., Park, J., Khattab, A., Chen, J., Hénégar, C., Wineinger, N. E., Muse, E. D., & Torkamani, A. (2025). Meta-prediction of coronary artery disease risk. *Nature Medicine*, *31*(7), 2277–2286.

https://doi.org/10.1038/s41591-025-03648-0

Chen, S., Loguercio, S., Chen, K., Lee, S. E., Park, J., Liu, S., Sadaei, H. J., & Torkamani, A. (2023). Artificial intelligence for risk assessment in primary prevention of coronary artery disease. *Current Cardiovascular Risk Reports, 17*(12), 215–226. https://doi.org/10.1007/s12170-023-00731-4

DeGroat, W., Abdelhalim, H., Peker, E., Sheth, N. A., Narayanan, R., Zeeshan, S., Liang, B. T., & Ahmed, Z. (2024). Multimodal artificial intelligence and machine learning for biomarker discovery and disease prediction using multi-omics cardiovascular data. *Scientific Reports, 14*(1), 78553. https://doi.org/10.1038/s41598-024-78553-6

Improving polygenic score prediction for coronary artery disease across populations of diverse ancestry. (2023). *Nature Medicine, 29*(7), 1621–1631. https://doi.org/10.1038/s41591-023-02451-z

Khanna, N. N., Singh, M., Maindarkar, M., Kumar, A., Johri, A. M., Mentella, L., Laird, J. R., Paraskevas, K. I., Ruzsa, Z., Singh, N., Kalra, M. K., Fernandes, J. F. E., Chaturvedi, S., Nicolaides, A., Rathore, V., Singh, I. M., Teji, J. S., Al-Maini, M., Isenović, E. R., … Suri, J. S. (2023). Polygenic risk score for cardiovascular diseases in an artificial intelligence paradigm: A review. *Journal of Korean Medical Science, 38*(46), e395. https://doi.org/10.3346/jkms.2023.38.e395

Lerga-Jaso, J., Terpolovsky, A., Novković, B., Osama, A., Manson, C., Bohn, S., Marino, A. D., Kunitomi, M., & Yazdi, P. G. (2025). Optimization of multi-ancestry polygenic risk score disease prediction models. *Scientific Reports, 15*(1), 17495. https://doi.org/10.1038/s41598-025-02903-1

Manikpurage, H. D., Eslami, A., Perrot, N., Li, Z., Couture, C., Mathieu, P., Bossé, Y., Arsenault, B. J., & Thériault, S. (2021). A polygenic risk score for coronary artery disease improves prediction of early-onset myocardial infarction and mortality in men. *medRxiv*.

https://doi.org/10.1101/2021.04.25.21256083

Pandey, D., Xu, L., Kun, E., Li, C., Wang, J. Y., Melek, A., DiCarlo, J. C., Castillo, E., Narula, J., Taylor, C. A., & Narasimhan, V. M. (2025). Multimodal artificial intelligence for precision preventive cardiology. *bioRxiv*. https://doi.org/10.1101/2025.10.09.25337677

Patel, A. P., Wang, M., Ruan, Y., Koyama, S., Clarke, S. L., Yang, X., Tcheandjieu, C., Agrawal, S., Fahed, A. C., Ellinor, P. T., Tsao, P. S., Sun, Y. V., Cho, K., Wilson, P. W. F., Assimes, T. L., van Heel, D. A., Butterworth, A. S., Aragam, K. G., Natarajan, P., & Khera, A. V. (2023a). Polygenic scores informed by multi-ancestry genome-wide association studies improve coronary artery disease risk prediction. *medRxiv*. https://doi.org/10.1101/2023.03.03.23286649

Patel, A. P., Wang, M., Ruan, Y., Koyama, S., Clarke, S. L., Yang, X., Tcheandjieu, C., Agrawal, S., Fahed, A. C., Ellinor, P. T., Tsao, P. S., Sun, Y. V., Cho, K., Wilson, P. W. F., Assimes, T. L., van Heel, D. A., Butterworth, A. S., Aragam, K. G., Natarajan, P., & Khera, A. V. (2023b). A multi-ancestry polygenic risk score improves risk prediction for coronary artery disease. *Nature Medicine, 29*(7), 1793–1803. https://doi.org/10.1038/s41591-023-02429-x

Ratman, D., Tshiaba, P., Levin, M. G., Sun, J., Tunstall, T., Maier, R. J., Shah, P., Rabinowitz, M., Rader, D. J., Kumar, A., & Im, K. M. (2025). Polygenic risk scores improve coronary artery disease risk prediction in individuals at borderline and intermediate clinical risk. *NPJ Cardiovascular Health, 2*(1), 49. https://doi.org/10.1038/s44325-025-00049-7

Riveros-McKay, F., Weale, M. E., Moore, R., Selzam, S., Krapohl, E., Sivley, R. M., Tarran, W. A., Sørensen, P., Lachapelle, A. S., Griffiths, J. A., Saffari, A., Deanfield, J., Spencer, C. C. A., Hippisley-Cox, J., Hunter, D. J., O'Sullivan, J. W., Ashley, E. A., Plagnol, V., & Donnelly, P. (2020). An integrated polygenic and clinical risk tool enhances coronary artery disease prediction. *medRxiv*.

https://doi.org/10.1101/2020.06.01.20119297

Rudroff, T., Rainio, O., & Klén, R. (2024). Artificial intelligence for prediction of early stages of Alzheimer's disease from neuroimaging biomarkers: A narrative review. *Neurological Sciences*. https://doi.org/10.1007/s10072-024-07649-8

Smith, J. L., Tcheandjieu, C., Dikilitas, O., Iyer, K., Miyazawa, K., Hilliard, A. T., Lynch, J. A., Rotter, J. I., Chen, Y. I., Sheu, W. H., Chang, K., Kanoni, S., Tsao, P. S., Ito, K., Kosel, M., Clarke, S. L., Schaid, D. J., Assimes, T. L., & Kullo, I. J. (2023). A multi-ancestry polygenic risk score for coronary heart disease based on ancestrally diverse genome-wide association studies. *medRxiv*. https://doi.org/10.1101/2023.06.02.23290896

Surakka, I., Wolford, B. N., Ritchie, S. C., Hornsby, W., Sutton, N. R., Gabrielsen, M. E., Skogholt, A. H., Thomas, L. F., Inouye, M., Hveem, K., & Willer, C. J. (2021). Sex-specific survival bias and interaction modeling in coronary artery disease risk prediction. *medRxiv*. https://doi.org/10.1101/2021.06.23.21259247

Tcheandjieu, C., Zhu, X., Hilliard, A. T., Clarke, S. L., Napolioni, V., Ma, S., Lee, K. M., Fang, H., Chen, F., Lu, Y., Tsao, N. L., Raghavan, S., Koyama, S., Gorman, B. R., Vujković, M., Klarin, D., Levin, M. G., Sinnott-Armstrong, N., Wojcik, G. L., … Assimes, T. L. (2022). Large-scale genome-wide association study of coronary artery disease in genetically diverse populations. *Carolina Digital Repository*. https://doi.org/10.17615/8zyz-p282

Torkamani, A., Chen, S., Lee, S. E., Sadaei, H. J., Park, J., Khattab, A., Hénégar, C., Wineinger, N. E., & Muse, E. D. (2023). Meta-prediction of coronary artery disease risk. *Research Square*. https://doi.org/10.21203/rs.3.rs-3694374/v1

Urbut, S., Cho, M. H., Paruchuri, K., Truong, B., Haidermota, S., Peloso, G. M., Hornsby, W., Philippakis, A., Fahed, A. C., & Natarajan, P. (2023). Dynamic importance of genomic and clinical risk for coronary artery disease over the life

course. *medRxiv*. https://doi.org/10.1101/2023.11.03.23298055

Venkatesh, R., Cherlin, T., Ritchie, M. D., Setia-Verma, S., & Setia-Verma, S. (2025). Integrating imaging-derived clinical endotypes with plasma proteomics and external polygenic risk scores enhances coronary microvascular disease risk prediction. *bioRxiv*. https://doi.org/10.1101/2025.08.18.25333844

Wang, X., & Zhu, H. (2024). Artificial intelligence in image-based cardiovascular disease analysis: A comprehensive survey and future outlook. *arXiv*. https://doi.org/10.48550/arxiv.2402.03394

Zhou, Z., & Wang, J. (2025). Multimodal data fusion strategies for precision risk prediction in coronary artery disease: Advances, challenges, and future informatics directions. *Rambam Maimonides Medical Journal*, *16*(4). https://doi.org/10.5041/rmmj.10558