Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Comparative Validation of ASCVD, Framingham Risk Score, Globorisk, and PREVENT for 10-Year Cardiovascular Risk Prediction in the UK Biobank

Anagha N
Forus Health Private Limited
Shreen Anith Kumar Jeyalakshmi
Forus Health Private Limited
Midhula Vijayan
Forus Health Private Limited
Deepthi K Prasad
deepthi.prasad@forushealth.com

Forus Health Private Limited

Research Article

Keywords: Cardiovascular risk prediction, UK Biobank, ASCVD, Framingham Risk Score, Globorisk, PREVENT, Decision Curve Analysis, Calibration, Discrimination

Posted Date: July 28th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10275384/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Cardiovascular Disorders on October 28th, 2026. See the published version at https://doi.org/10.1186/s12872-026-06773-x.

Abstract

Background

Cardiovascular disease (CVD) remains the leading cause of global mortality. Accurate and generalizable risk prediction models are essential for effective primary prevention. Widely used models such as the Atherosclerotic Cardiovascular Disease (ASCVD) pooled cohort equations, Framingham Risk Score (FRS), Globorisk, and the more recent Predicting Risk of cardiovascular disease EVENTs (PREVENT) model have demonstrated variable performance across populations.

Objective

To assess and compare the discrimination, calibration, and clinical utility of the ASCVD, FRS, Globorisk, and PREVENT models for predicting 10-year cardiovascular risk in a large contemporary UK cohort.

Methods

A total of 183,124 participants without prior cardiovascular disease were selected from the UK Biobank cohort and followed for up to 10 years for incident cardiovascular events. Risk scores were computed using baseline clinical variables as specified by each model: ASCVD (2013 ACC/AHA Pooled Cohort Equations, Goff et al.), FRS (2008 general cardiovascular risk model, D'Agostino et al.), Globorisk (UK-specific calibration, Hajifathalian et al., 2015), and PREVENT (2023 AHA equations, Khan et al.). Model performance was evaluated using receiver operating characteristic (ROC) analysis, area under the curve (AUC) with 95% confidence intervals, calibration analysis, sensitivity, specificity, positive and negative predictive values, Decision Curve Analysis (DCA), and Net Reclassification Improvement (NRI).

Results

Over the 10-year follow-up period, 10,560 incident cardiovascular events were recorded (5.8% event rate). ASCVD demonstrated the best overall discrimination (AUC = 0.717, 95% CI: 0.712–0.723), closely followed by FRS (AUC = 0.715, 95% CI: 0.709–0.721). PREVENT achieved intermediate performance (AUC = 0.668, 95% CI: 0.662–0.674), while Globorisk showed the lowest discrimination (AUC = 0.605, 95% CI: 0.599–0.611). ASCVD demonstrated the best calibration across risk strata. Decision Curve Analysis indicated that ASCVD provided the highest net clinical benefit across most decision thresholds. NRI analysis confirmed that no single model provided consistent reclassification benefit over others. Model performance declined across all models in older adults (> 59 years).

Conclusion

ASCVD and FRS demonstrate acceptable discrimination and calibration for 10-year CVD risk prediction in a contemporary UK cohort, with ASCVD showing comparatively superior clinical utility on Decision Curve Analysis. PREVENT, though designed with broader applicability, showed intermediate performance. Globorisk demonstrated the lowest discriminatory ability in this cohort. Performance declines in older adults across all models underscore the need for age-specific recalibration and population-tailored risk prediction frameworks.

1 Introduction

Cardiovascular diseases (CVDs) remain the leading cause of mortality worldwide, responsible for approximately 19.8 million deaths in 2022 [1]. A substantial proportion of these events are preventable through early identification of individuals at elevated risk and timely implementation of preventive interventions [2]. Consequently, accurate cardiovascular risk prediction models have become a cornerstone of primary prevention strategies and clinical decision-making [3, 4].

Several multivariable cardiovascular risk prediction models have been developed over the past decades to estimate an individual's probability of experiencing a cardiovascular event within a defined time horizon. Among the most widely used are the Framingham Risk Score (FRS) [5], the ASCVD pooled cohort equations [6], Globorisk [7], and the more recently introduced PREVENT model [8].

The Framingham Risk Score (FRS, 2008 general CVD model) was developed from the Framingham Heart Study and estimates 10-year CVD risk using age, sex, systolic blood pressure, total and HDL cholesterol, diabetes, smoking, and blood pressure treatment status [5]. The ASCVD pooled cohort equations (2013 ACC/AHA, Goff et al.) extended this approach by incorporating race-specific equations and recalibration for atherosclerotic CVD endpoints [6]. Globorisk was developed to address the need for internationally applicable tools through country-specific calibration methods, enabling adjustment for national disease incidence and risk factor distributions [7]. The PREVENT (Predicting Risk of CVD EVENTs) model, introduced by the American Heart Association in 2023 (Khan et al.), incorporates additional metabolic parameters including kidney function markers and is designed with broader demographic applicability [8].

Despite the availability of these models, discrepancies in predictive performance have been reported across demographic groups [9]. Several studies suggest that traditional risk scores may underestimate risk in women and certain ethnic populations, potentially contributing to disparities in diagnosis and treatment [2, 10]. Age-related differences in predictive accuracy have also been documented, with declining discrimination observed among older individuals where competing health risks and multimorbidity become more prominent [11, 12].

The UK Biobank [13], a large population-based cohort comprising over 500,000 participants aged 40–69 years, provides a valuable contemporary dataset for evaluating cardiovascular risk prediction models. Despite widespread clinical use, traditional models rely on static baseline variables and may not capture dynamic or population-specific risk patterns, particularly in diverse cohorts such as the UK Biobank.

While individual models have been evaluated previously, a systematic four-way comparison-including the most recent PREVENT model-within the same large contemporary UK cohort, incorporating Decision Curve Analysis and NRI, has not been reported.

In this study, we perform a comparative evaluation of ASCVD, FRS, Globorisk, and PREVENT using data derived from the UK Biobank. The analysis encompasses discrimination, calibration, sensitivity, specificity, predictive values, Decision Curve Analysis, and Net Reclassification Improvement. By systematically comparing these models within the same population, this study aims to provide insights into their relative performance and implications for cardiovascular risk stratification in contemporary clinical practice.

2 Methods

2.1 Study Population

Participants were selected from the UK Biobank cohort, which includes approximately 500,000 individuals aged 40–69 years recruited between 2006 and 2010 [13]. Individuals with a history of cardiovascular disease at baseline were excluded to ensure that only participants free of cardiovascular events at study entry were included. Participants were required to have sufficient clinical and demographic data to calculate all four cardiovascular risk scores. Individuals with incomplete data for risk score calculation were excluded.

To ensure adequate follow-up for outcome assessment and minimise survivorship bias in the event-free subgroup, participants whose follow-up duration was less than 10 years and who had not experienced a cardiovascular event during that time were excluded. This approach is consistent with the methodology of original model development studies [5, 6, 7, 8], which defined the outcome horizon as 10 years. Participants who experienced a cardiovascular event within the 10-year follow-up period were classified as cases; those who completed 10 years of follow-up without an event were classified as non-cases. After applying all inclusion and exclusion criteria, the final analytical cohort consisted of 183,124 participants (Fig. 1).

2.2 Data Extraction and Cardiovascular Outcome Definition

Data were extracted from the UK Biobank database for eligible participants meeting the inclusion criteria. Cardiovascular events were identified using predefined ICD-10 codes, encompassing fatal and non-fatal myocardial infarction and stroke, consistent with the outcome definitions used in the development of the ASCVD, FRS, and PREVENT models. The follow-up period was defined from the baseline assessment date to the date of the first recorded cardiovascular event or the end of the 10-year observation window.

2.3 Data Variables

Table 1 summarises the input variables used by each of the four risk prediction models. Common predictors across all models include age, sex, systolic blood pressure, smoking status, diabetes status, and total cholesterol. The ASCVD model additionally incorporates HDL cholesterol, blood pressure treatment status, and race/ethnicity adjustment [6]. FRS includes HDL cholesterol and blood pressure treatment status [5]. Globorisk applies UK-specific calibration factors to adjust predicted risks for national disease incidence [7]. PREVENT incorporates kidney function markers and broader cardiovascular endpoints including heart failure [8].

Table 1
Variables used for calculating 10-year cardiovascular risk scores in ASCVD, FRS, Globorisk, and PREVENT models.

Variable	ASCVD	FRS	Globorisk	PREVENT
Age	Yes	Yes	Yes	Yes
Sex	Yes	Yes	Yes	Yes
Systolic Blood Pressure	Yes	Yes	Yes	Yes
Smoking Status	Yes	Yes	Yes	Yes
Diabetes	Yes	Yes	Yes	Yes
Total Cholesterol	Yes	Yes	Yes	Yes
HDL Cholesterol	Yes	Yes	No	No
Blood Pressure Treatment	Yes	Yes	No	No
Race / Ethnicity	Yes	No	No	No
Kidney Function (eGFR)	No	No	No	Yes
Statin Use	No	No	No	Yes
Country-specific Calibration	No	No	Yes (UK)	No

2.4 Cardiovascular Risk Prediction Models

The ASCVD pooled cohort equations (2013 ACC/AHA; Goff et al. [6]) use sex- and race-specific logistic regression equations to estimate 10-year risk of atherosclerotic CVD events. Individuals with a predicted risk ≥ 7.5% are classified as intermediate or higher risk per guideline thresholds. The FRS (2008 general CVD risk model; D'Agostino et al. [5]) employs separate sex-specific Weibull survival models. Risk categories are defined as low (< 10%), intermediate (10–20%), and high (> 20%). Globorisk (Hajifathalian et al., 2015 [7]) was developed through pooled analysis of prospective cohorts and applied with UK-specific recalibration coefficients to adjust for country-level CVD incidence and risk factor distributions. PREVENT (Khan et al., 2023 AHA [8]) was developed using contemporary US population data and employs age- and sex-specific equations that additionally incorporate kidney function (eGFR) and statin

use, with risk categories defined as low (< 5%), intermediate (5-<15%), and high (≥ 15%) for total CVD events.

Specific model equations and coefficients are provided in Supplementary Material S1. Equations were implemented as described in the original publications, with no modifications to the model coefficients.

2.5 Statistical Analysis

All statistical analyses were performed using Python (version 3.12) with standard scientific libraries including pandas, NumPy, SciPy, scikit-learn, and dcurves. Continuous variables were summarised as median with interquartile range (IQR); categorical variables as frequencies and percentages.

Discriminative performance was evaluated using ROC curve analysis and AUC with 95% confidence intervals computed using the Hanley-McNeil method. Classification metrics including sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were derived at the optimal risk threshold identified by the Youden Index (maximising sensitivity + specificity − 1). Calibration was evaluated using calibration plots comparing mean predicted probabilities against observed event rates across deciles of predicted risk.

Decision Curve Analysis (DCA) was performed to evaluate the net clinical benefit of each model across a range of decision thresholds from 1% to 50%, using the dcurves Python library. Net benefit represents the trade-off between true positives and weighted false positives for a given threshold probability, providing a clinically interpretable measure of model utility that transcends simple discrimination metrics [14].

Net Reclassification Improvement (NRI) was calculated for all pairwise model comparisons at a classification threshold of 7.5%, consistent with guideline-recommended thresholds for initiating preventive therapy [6]. NRI quantifies the proportion of individuals appropriately reclassified (upward among events, downward among non-events) when moving from one model to another.

Agreement between predicted risk scores was assessed using Spearman rank correlation. All analyses were stratified by sex (male/female) and age group (mid-aged: ≤59 years; older adults: >59 years). Two-sided p-values < 0.05 were considered statistically significant; all results are reported with 95% confidence intervals.

3 Results

3.1 Participant Characteristics

The final analytical cohort comprised 183,124 participants (Table 2). The median age was 58 years (IQR 51–63), with females comprising 55.8% of the cohort (n = 102,152) and males 44.2% (n = 80,972). Over the 10-year follow-up period, 10,560 incident cardiovascular events were recorded, representing an overall event rate of 5.8%. The event rate was substantially higher among males (8.3%; n = 6,714) than females (3.8%; n = 3,846). Diabetes was present in 4.6% overall, with higher prevalence in males (6.0%)

compared with females (3.4%). Smoking was reported in 45.2% of participants (males 50.7%, females 40.7%). The median systolic blood pressure was 138 mmHg (IQR 126–151) overall, with males demonstrating higher values (141 mmHg) compared with females (135 mmHg).

Table 2
Baseline demographic and clinical characteristics of the study population.

Variable	Overall (N = 183,124)	Female (N = 102,152)	Male (N = 80,972)	p-value*
Age, years - median (IQR)	58 (51–63)	58 (51–63)	58 (50–63)	-
Female	102,152 (55.8%)	-	-	-
Male	80,972 (44.2%)	-	-	-
Incident CVD events	10,560 (5.8%)	3,846 (3.8%)	6,714 (8.3%)	< 0.001
Diabetes	8,424 (4.6%)	3,473 (3.4%)	4,951 (6.0%)	< 0.001
Smoking status	82,841 (45.2%)	41,576 (40.7%)	41,265 (50.7%)	< 0.001
Systolic BP, mmHg - median (IQR)	138 (126–151)	135 (122–149)	141 (130–153)	< 0.001

** Chi-squared test for categorical variables; Mann-Whitney U test for continuous variables.*

Figure 1. *Study population flowchart showing derivation of the final analytical cohort of 183,124 participants with all four risk scores available from the UK Biobank.*

3.2 Comparison of Risk Prediction Models

3.2.1 Calibration Analysis

Calibration analysis evaluated the agreement between predicted and observed 10-year CVD risk for all four models (Fig. 2). The dashed diagonal represents perfect calibration. The ASCVD model showed the closest alignment with the ideal calibration line across most predicted risk levels, demonstrating relatively better agreement between predicted probabilities and observed outcomes. The FRS model systematically underestimated CVD risk, particularly at moderate and higher predicted risk levels. PREVENT showed reasonable calibration at low predicted risk ranges (0–5%) but overestimated risk at intermediate levels. Globorisk demonstrated reasonable calibration at lower predicted risk levels but diverged progressively at higher risk thresholds. In the low-risk range (0–5%), all models showed minimal deviation from perfect calibration; differences became more pronounced at predicted risks exceeding 10–15%, where ASCVD maintained the best calibration and FRS showed the greatest degree of systematic underestimation.

Figure 2. *Calibration of 10-year CVD risk prediction models. The dashed line represents perfect calibration. ASCVD shows the best agreement with observed event rates across predicted risk strata.*

3.2.2 ROC Curve Analysis and Discrimination

The discriminative performance of all four models, evaluated using ROC curve analysis across the overall cohort and stratified by sex and age group, is presented in Tables 3 and 4. Overall, ASCVD (AUC = 0.717, 95% CI: 0.712–0.723) and FRS (AUC = 0.715, 95% CI: 0.709–0.721) demonstrated comparable and superior discrimination. PREVENT achieved intermediate performance (AUC = 0.668, 95% CI: 0.662–0.674), while Globorisk showed the weakest discrimination (AUC = 0.605, 95% CI: 0.599–0.611), though all models performed better than chance. All p-values were < 0.001.

The FRS demonstrated the highest overall sensitivity (75.5%) at its Youden-optimal threshold of 11.5%, though this was accompanied by lower specificity (55.7%), reflecting a higher rate of false-positive classifications. ASCVD offered a more balanced profile (sensitivity 62.2%, specificity 68.5%) at a threshold of 5.58%. PREVENT showed moderate sensitivity (56.5%) and specificity (66.6%) at a lower threshold of 3.7%. Globorisk exhibited high specificity (73.9%) but substantially lower sensitivity (38.4%). Positive predictive values were low across all models (8.3–10.8%), reflecting the relatively low event rate in the cohort, while negative predictive values were consistently high (95.1–97.4%), indicating strong ability to identify low-risk individuals.

Table 3
Model discrimination and classification metrics - overall and sex-stratified analysis.

Model	AUC (95% CI)	Sensitivity	Specificity	PPV	NPV	Cutoff	Subgroup
ASCVD	0.717 (0.712–0.723)	62.2%	68.5%	10.8%	96.7%	5.58%	Total
FRS	0.715 (0.709–0.721)	75.5%	55.7%	9.4%	97.4%	11.5%	Total
Globorisk	0.605 (0.599–0.611)	38.4%	73.9%	8.3%	95.1%	6.41%	Total
PREVENT	0.668 (0.662–0.674)	56.5%	66.6%	9.4%	96.2%	3.70%	Total
ASCVD	0.672 (0.665–0.679)	76.0%	48.2%	11.7%	95.7%	9.26%	Men
FRS	0.677 (0.669–0.684)	88.9%	32.4%	10.6%	97.0%	14.4%	Men
Globorisk	0.647 (0.639–0.654)	26.3%	86.4%	14.9%	92.8%	5.62%	Men
PREVENT	0.673 (0.666–0.680)	50.8%	71.8%	14.0%	94.2%	3.50%	Men
ASCVD	0.709 (0.700-0.718)	38.2%	83.9%	8.5%	97.2%	3.57%	Women
FRS	0.695 (0.685–0.704)	52.2%	73.3%	7.1%	97.5%	7.70%	Women
Globorisk	0.671 (0.661–0.680)	59.4%	64.5%	6.1%	97.6%	8.51%	Women
PREVENT	0.701 (0.691–0.710)	66.5%	62.7%	6.5%	98.0%	4.40%	Women

AUC = Area Under the Curve; PPV = Positive Predictive Value; NPV = Negative Predictive Value. Cutoffs determined by Youden Index. All AUCs $p < 0.001$ vs. null.

3.3 Sex-Stratified Analysis

Sex-stratified ROC analysis revealed notable differences in model performance between males and females. Among females, ASCVD achieved the highest discrimination (AUC = 0.709), followed by PREVENT (AUC = 0.701), FRS (AUC = 0.695), and Globorisk (AUC = 0.671). Among males, FRS achieved slightly higher discrimination (AUC = 0.677), closely followed by PREVENT (AUC = 0.673) and ASCVD (AUC = 0.672), with Globorisk lowest (AUC = 0.647).

A notable finding was that PREVENT performed comparably to ASCVD among females (AUC 0.701 vs. 0.709) while maintaining a more balanced sensitivity-specificity profile-demonstrating higher sensitivity

(66.5% vs. 38.2%) at a lower Youden-optimal threshold (4.4% vs. 3.57%). This suggests that PREVENT may be more sensitive for identifying high-risk women at lower absolute risk thresholds. Among males, all four models showed relatively lower AUC values (0.647–0.677) compared with the female subgroup, reflecting the higher event rate and compressed dynamic range of risk score differentiation in this group.

Across both sexes, FRS exhibited the highest sensitivity (90% in men; 52.2% in women) at the cost of substantially lower specificity (32.4% and 73.3%, respectively), reflecting its higher classification thresholds. Globorisk showed a reversed pattern-high specificity particularly in males (86.4%) but very low sensitivity (26.3%). ASCVD and PREVENT offered intermediate trade-offs across both sexes.

3.4 Age-Stratified Analysis

Age-stratified analyses demonstrated consistent declines in discriminative performance with increasing age across all four models (Table 4). Among mid-aged adults (≤ 59 years), ASCVD and FRS achieved the highest AUCs (both 0.726), followed by PREVENT (0.663) and Globorisk (0.571). In the older adult subgroup (> 59 years), AUC values declined substantially: ASCVD to 0.651, FRS to 0.646, PREVENT to 0.576, and Globorisk to 0.527-approaching near-random discrimination for Globorisk in this age group.

Positive predictive values increased with age for all models, reflecting the higher absolute event rate (8.5%) in older adults compared with mid-aged adults (3.6%). Conversely, negative predictive values remained high (> 91%) across all age-model combinations. Optimal risk thresholds were uniformly higher in older adults, consistent with the age-related increase in predicted risk under all scoring frameworks.

Table 4
Model discrimination and classification metrics - age-stratified analysis (all p-values < 0.001).

Model	AUC (95% CI)	Sensitivity	Specificity	PPV	NPV	Cutoff	Age Group
ASCVD	0.726 (0.716–0.735)	29.0%	90.5%	10.2%	97.1%	3.35%	Mid-aged (≤ 59)
FRS	0.726 (0.717–0.736)	57.1%	74.9%	7.8%	97.9%	8.40%	Mid-aged (≤ 59)
Globorisk	0.571 (0.561–0.580)	14.9%	89.8%	5.2%	96.6%	6.25%	Mid-aged (≤ 59)
PREVENT	0.663 (0.653–0.673)	19.8%	91.9%	8.4%	96.8%	2.50%	Mid-aged (≤ 59)
ASCVD	0.651 (0.644–0.659)	80.1%	39.2%	10.9%	95.5%	10.64%	Older (> 59)
FRS	0.646 (0.638–0.653)	85.4%	30.0%	10.2%	95.7%	16.8%	Older (> 59)
Globorisk	0.527 (0.520–0.534)	51.0%	52.8%	9.1%	92.1%	11.89%	Older (> 59)
PREVENT	0.576 (0.569–0.584)	76.3%	33.0%	9.6%	93.7%	5.70%	Older (> 59)

Mid-aged: n = 102,311, events = 3,694 (3.6%); Older adults: n = 80,813, events = 6,866 (8.5%). Cutoffs determined by Youden Index.

3.5 Decision Curve Analysis

Decision Curve Analysis evaluated the net clinical benefit of each model across decision thresholds ranging from 1% to 50% (Fig. 3). ASCVD demonstrated the highest net benefit across the widest range of clinically relevant thresholds (approximately 5–20%), indicating superior clinical utility for guiding preventive treatment decisions in this range. PREVENT provided comparable net benefit to ASCVD at lower thresholds (< 8%) but declined more steeply at intermediate thresholds. FRS showed negative net benefit across thresholds of approximately 8–25%, indicating that applying FRS-based classification in this threshold range would result in net harm relative to treating no one or treating everyone. Globorisk likewise showed reduced net benefit at thresholds above 8%, with negative values at intermediate ranges, reflecting its lower discriminative performance.

These findings suggest that for clinical decision-making contexts where treatment is considered at thresholds of approximately 5–20%, ASCVD provides the most favourable clinical utility profile among the four models evaluated.

Figure 3. *Decision Curve Analysis. Net benefit is plotted against decision threshold probability for ASCVD, FRS, Globorisk, and PREVENT. Higher net benefit across a wider range of thresholds indicates greater clinical utility.*

3.6 Net Reclassification Improvement

NRI analysis was performed for all six pairwise model comparisons at a classification threshold of 7.5% (Table 5). When ASCVD was compared with FRS as the reference, the Event NRI was positive (0.241), indicating that FRS reclassified a greater proportion of true cases upward relative to ASCVD. However, the Nonevent NRI was negative (− 0.273), reflecting that FRS also reclassified more non-events upward (false upward reclassifications), resulting in an overall negative NRI (− 0.032). This pattern is consistent with FRS's higher sensitivity but lower specificity profile.

No pairwise comparison demonstrated a positive overall NRI, indicating that no single model provided consistent net reclassification benefit over another across both events and non-events at the 7.5% threshold. PREVENT showed positive Nonevent NRI relative to both FRS and Globorisk (0.469 and 0.302, respectively), suggesting that PREVENT correctly reclassifies more non-events downward-a clinically meaningful attribute for reducing unnecessary treatment in low-risk individuals.

Table 5
Net Reclassification Improvement (NRI) - pairwise model comparisons at threshold 7.5%.

Reference Model	Comparison Model	Event NRI	Nonevent NRI	Overall NRI
ASCVD	FRS	0.241	−0.273	−0.032
ASCVD	Globorisk	−0.058	−0.106	−0.164
ASCVD	PREVENT	−0.377	0.196	−0.180
FRS	Globorisk	−0.298	0.167	−0.132
FRS	PREVENT	−0.617	0.469	−0.148
Globorisk	PREVENT	−0.319	0.302	−0.017

Positive Event NRI: new model reclassifies more events upward. Positive Nonevent NRI: new model reclassifies more non-events downward. Threshold = 7.5% (ACC/AHA guideline threshold for statin consideration).

3.7 Agreement Between Risk Scores

Spearman correlation analysis demonstrated very strong agreement between ASCVD and FRS ($\rho = 0.95$, $p < 0.001$), indicating near-equivalent ranking of individuals according to predicted cardiovascular risk. ASCVD and PREVENT showed strong correlation ($\rho = 0.80$), as did PREVENT and Globorisk ($\rho = 0.81$). FRS and PREVENT correlated moderately strongly ($\rho = 0.76$). Globorisk showed only moderate correlations with ASCVD ($\rho = 0.51$) and FRS ($\rho = 0.49$), suggesting substantially different risk classification relative to the other three models (Fig. 4).

Figure 4. *Spearman rank correlation matrix. Cell values represent pairwise Spearman rho coefficients. All $p < 0.001$.*

4 Discussion

4.1 Principal Findings

In this large population-based analysis of 183,124 UK Biobank participants followed over 10 years, we evaluated and compared the discrimination, calibration, and clinical utility of four widely used cardiovascular risk prediction models: ASCVD, FRS, Globorisk, and PREVENT. Our principal findings are: (i) ASCVD and FRS demonstrated the best overall discrimination (AUC ≈ 0.72), substantially outperforming Globorisk (AUC ≈ 0.61); (ii) PREVENT achieved intermediate discrimination (AUC = 0.668), with particularly competitive performance in women and at lower risk thresholds; (iii) ASCVD showed the best calibration and the highest net clinical benefit on DCA across most decision thresholds; (iv) no model demonstrated consistent NRI benefit over others at the 7.5% threshold; and (v) discriminative performance declined substantially in older adults across all four models.

These findings support the continued use of ASCVD and FRS as primary tools for 10-year CVD risk stratification in clinical practice, while highlighting important limitations in model calibration and subgroup-specific performance that have direct implications for cardiovascular risk management.

4.2 Comparison with Existing Literature

The AUC values obtained for ASCVD (0.717) and FRS (0.715) are consistent with prior external validations of these models in European populations. Parsons et al. [15] validated QRISK3 in the UK Biobank and reported AUCs in the range of 0.71–0.74, comparable with our findings for ASCVD and FRS. Our Globorisk AUC (0.605) is lower than values reported in some non-UK populations [7], which is likely attributable to Globorisk's country-specific calibration parameters being optimised for overall CVD mortality rather than incident non-fatal events as defined here.

The intermediate performance of PREVENT (AUC = 0.668) in this cohort warrants discussion. PREVENT was developed using contemporary US National Health and Nutrition Examination Survey (NHANES) and other US-based data, and its additional inputs-including eGFR and statin use-may be less discriminating in the UK Biobank cohort, where statin use prevalence and CKD distributions differ from the US development population [8]. Notably, PREVENT showed competitive performance in women (AUC = 0.701), which may reflect its design to overcome historical underestimation of CVD risk in women seen with older models [16].

The calibration findings-specifically the systematic underestimation of risk by FRS and the overestimation by PREVENT at intermediate ranges-are consistent with patterns reported when models derived in different demographic settings are applied to contemporary European populations [17]. This underscores the well-recognised need for country-specific recalibration when applying risk models across populations.

Our DCA results extend prior work by demonstrating that ASCVD maintains positive net clinical benefit across the broadest range of clinically relevant thresholds (5–20%), which spans the guideline-recommended treatment decision zone [6]. The negative net benefit of FRS at intermediate thresholds (8–25%) is an important clinical finding: it indicates that using FRS-based thresholds in this range could lead to net harm through unnecessary treatment of misclassified individuals. This is consistent with FRS's known tendency to overestimate absolute risk in contemporary European cohorts [17].

The NRI results, showing no positive overall NRI for any pairwise comparison, align with the observation that none of the four models represents a clear step-change improvement over the others in terms of categorical reclassification at the 7.5% threshold. The positive Nonevent NRI for PREVENT compared with FRS (0.469) suggests that PREVENT may more accurately identify truly low-risk individuals-an important property for avoiding overtreatment, which aligns with PREVENT's design intent [8].

4.3 Sex-Specific Considerations

The sex-stratified findings highlight important differential performance across models. Among women, ASCVD and PREVENT both achieved AUCs above 0.70, while PREVENT demonstrated higher sensitivity (66.5% vs. 38.2% for ASCVD), suggesting it may be better suited to detecting high-risk women who would otherwise be missed at conventional thresholds. FRS showed comparatively lower performance in women (AUC = 0.695) relative to ASCVD and PREVENT, consistent with previously reported concerns about FRS's performance in female populations [16].

Among men, all four models showed lower AUCs (0.647–0.677) relative to the female subgroup, which is partly attributable to the higher event rate among men (8.3%), which compresses discriminative capacity. The lower specificity of FRS in men (32.4%)-driven by its high classification threshold (14.4%)-implies that a substantial proportion of the male population would be flagged as high-risk, warranting careful consideration in resource-limited preventive care settings.

4.4 Age-Stratified Findings and Implications

The decline in AUC from mid-aged (AUC ≈ 0.73 for ASCVD/FRS) to older adults (AUC ≈ 0.65) is consistent with literature describing attenuated risk factor discrimination in older populations [12]. In older individuals, the predictive contribution of classical risk factors diminishes relative to accumulated multimorbidity, functional decline, and competing mortality risks, none of which are captured by these models. The near-random performance of Globorisk in older adults (AUC = 0.527) suggests it should not be used as a standalone tool in this age group within UK populations.

These findings underscore the need for age-specific risk prediction frameworks, potentially incorporating dynamic risk factors such as longitudinal blood pressure trajectories, multimorbidity indices, or frailty measures, particularly for individuals over 60 years of age.

4.5 Strengths and Limitations

Key strengths of this study include the large sample size (n = 183,124), the systematic four-way comparison of models within the same well-characterised cohort, and the comprehensive analytical framework incorporating discrimination, calibration, DCA, and NRI. The inclusion of PREVENT alongside established models in a large external validation represents a novel contribution.

Several limitations should be acknowledged. The UK Biobank exhibits healthy volunteer selection bias, which may result in lower overall event rates than seen in unselected clinical populations, potentially affecting calibration estimates. Risk factor measurements were obtained at a single baseline assessment; changes in risk profiles over the follow-up period were not incorporated. Participants with less than 10 years of event-free follow-up were excluded; while this was done to ensure complete outcome ascertainment consistent with the 10-year horizon of the models, this exclusion may reduce the representativeness of the cohort relative to an unselected clinical population, and residual censoring bias cannot be excluded. Furthermore, specific model equations were applied as published, without recalibration to the UK Biobank population; recalibration may improve absolute calibration for all models.

Finally, the UK Biobank cohort is predominantly White European, which limits the generalisability of findings, particularly for ASCVD's race-specific equations.

5 Conclusion

In this large contemporary UK cohort, ASCVD and FRS demonstrated the best overall discrimination and clinical utility for predicting 10-year cardiovascular risk. ASCVD exhibited superior calibration and the highest net clinical benefit on Decision Curve Analysis across clinically relevant treatment thresholds. PREVENT, a newer model with broader design intent, showed intermediate discrimination but competitive performance in women and improved non-event reclassification relative to FRS. Globorisk demonstrated the weakest performance in this cohort, particularly in older adults.

Across all four models, discrimination declined substantially in adults over 59 years, highlighting that static, baseline-variable models have important performance limitations in older populations. Future risk prediction frameworks should incorporate dynamic risk factors, longitudinal data, and population-specific calibration to improve accuracy across the full age spectrum. The present findings provide evidence to support model selection for cardiovascular risk stratification in contemporary UK clinical practice, and emphasise the value of comprehensive model evaluation-including DCA-as a standard component of clinical prediction model reporting.

Declarations

Funding

Not applicable.

Competing Interests

The authors declare no competing interests.

Ethics Approval and Consent to Participate

All methods were carried out in accordance with relevant guidelines and regulations. This study was conducted within UK Biobank project 968135 and was approved by the Swedish Ethical Review Authority (2019-02328 and 2021-05762-02). Informed consent was obtained from all subjects.

Consent for Publication

Not applicable. This study utilized de-identified data from UK Biobank project 968135. No identifying images, individual-level clinical details, or personal information that could compromise participant

anonymity are presented in this manuscript.

Data Availability

The UK Biobank is accessible to researchers worldwide at https://www.ukbiobank.ac.uk/. Derived variables used in this study are available from the corresponding author upon reasonable request.

Authors' Contributions

AN contributed to study design, data analysis, and manuscript drafting. SAKJ contributed to statistical analysis, interpretation of results, and manuscript revision. MV contributed to data extraction and model computation. DKP contributed to study design and critical revision. All authors approved the final manuscript.

Clinical Trial Number statement: Clinical trial number: not applicable.

References

1. World Health Organization. Cardiovascular diseases (CVDs). 2023. Available from: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds).
2. Update A. Heart disease and stroke statistics-2017 update. Circulation. 2017;135:e146-e603.
3. Yusuf S, et al. Modifiable risk factors, cardiovascular disease, and mortality in 155,722 individuals from 21 countries (PURE). Lancet. 2020;395:795-808.
4. Sun L, et al. Polygenic risk scores in cardiovascular risk prediction: a cohort study and modelling analyses. PLoS Med. 2021;18:e1003498.
5. D'Agostino RB Sr, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation. 2008;117:743-753.
6. Goff DC Jr, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk. J Am Coll Cardiol. 2014;63:2935-2959.
7. Hajifathalian K, et al. A novel risk score to predict cardiovascular disease risk in national populations (Globorisk). Lancet Diabetes Endocrinol. 2015;3:339-355.
8. Khan SS, et al. Development and validation of the American Heart Association's PREVENT equations. Circulation. 2024;149:430-449.
9. Marouco JL, et al. Race and ethnicity disparities in cardiovascular and cancer mortality. J Racial Ethn Health Disparities. 2025;12:285-297.
10. Meysamie A, et al. Cardiovascular risk assessment by FRS and SCORE in Iranian adult population. J Diabetes Metab Disord. 2017;16:35.
11. Wenger NK. Coronary heart disease: an older woman's major health risk. BMJ. 1997;315:1085-1090.

12. Bonanad C, et al. Cardiovascular prevention in elderly patients. J Geriatr Cardiol. 2022;19:377.
13. Sudlow C, et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases. PLoS Med. 2015;12:e1001779.
14. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making. 2006;26:565-574.
15. Parsons RE, et al. Independent external validation of the QRISK3 cardiovascular disease risk prediction model using UK Biobank. Heart. 2023;109:1690-1697.
16. St Pierre SR, et al. Sex-specific cardiovascular risk factors in the UK Biobank. Front Physiol. 2024;15:1339866.
17. Damen JA, et al. Prediction models for cardiovascular disease risk in the general population: systematic review. BMJ. 2016;353.
18. Arnett DK, et al. 2019 ACC/AHA guideline on the primary prevention of cardiovascular disease. J Am Coll Cardiol. 2019;74:e177-e232.
19. Grundy SM, et al. 2018 AHA/ACC guideline on the management of blood cholesterol. J Am Coll Cardiol. 2019;73:e285-e350.
20. Whelton PK, et al. 2017 ACC/AHA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults. J Am Coll Cardiol. 2018;71:e127-e248.

Figures

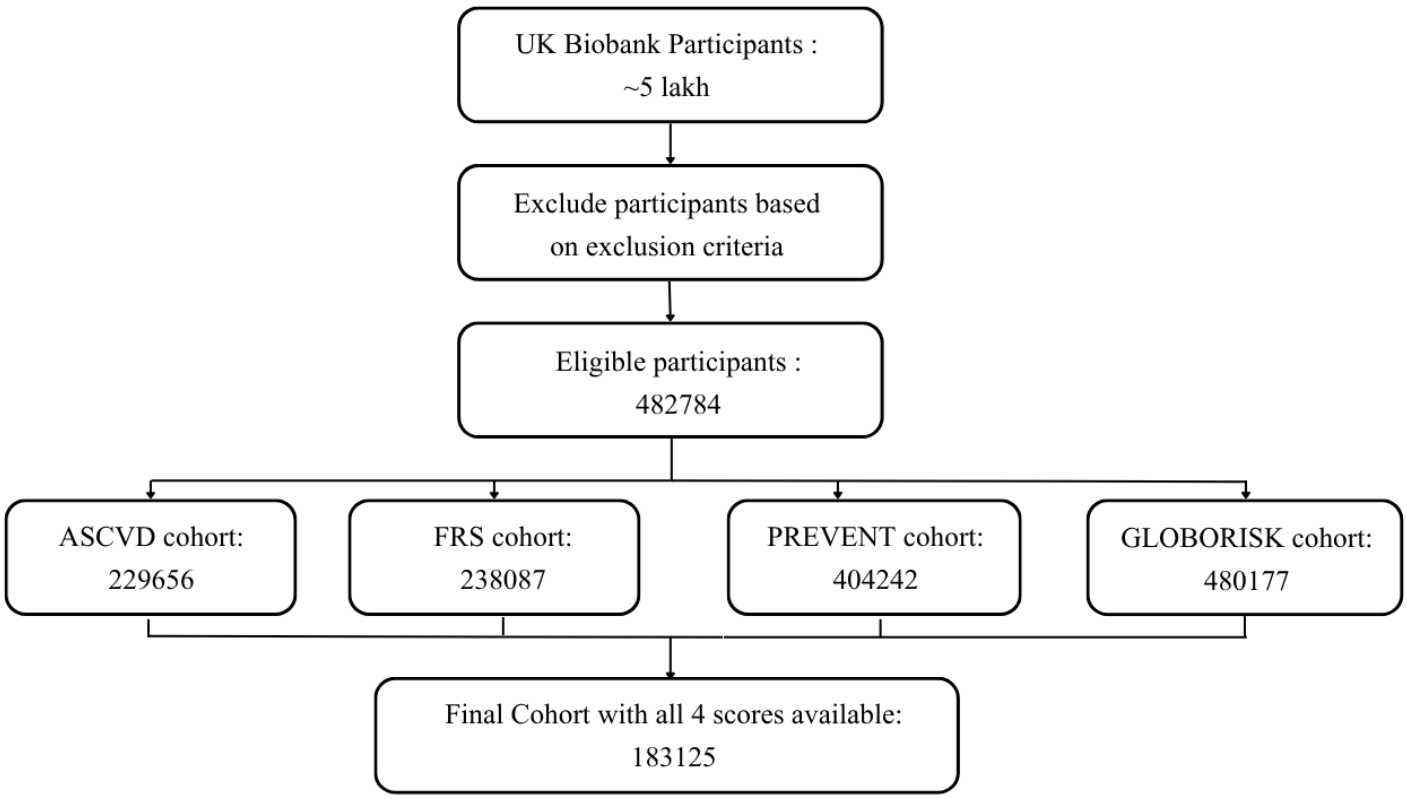


Figure 1

Study population flowchart.

Figure 1. Study population flowchart showing derivation of the final analytical cohort of 183,124 participants with all four risk scores available from the UK Biobank.

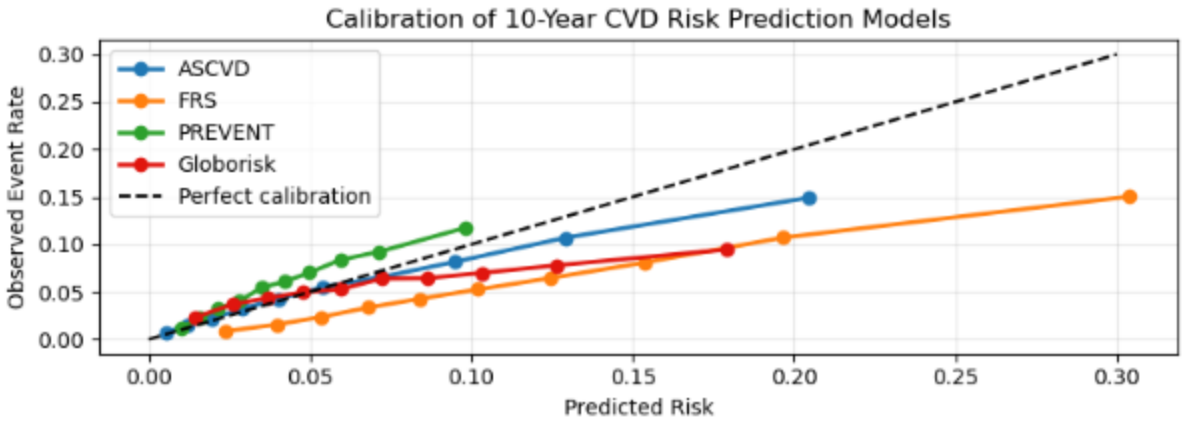


Figure 2

Calibration curves for the four cardiovascular risk prediction models.

Figure 2. Calibration of 10-year CVD risk prediction models. The dashed line represents perfect calibration. ASCVD shows the best agreement with observed event rates across predicted risk strata.

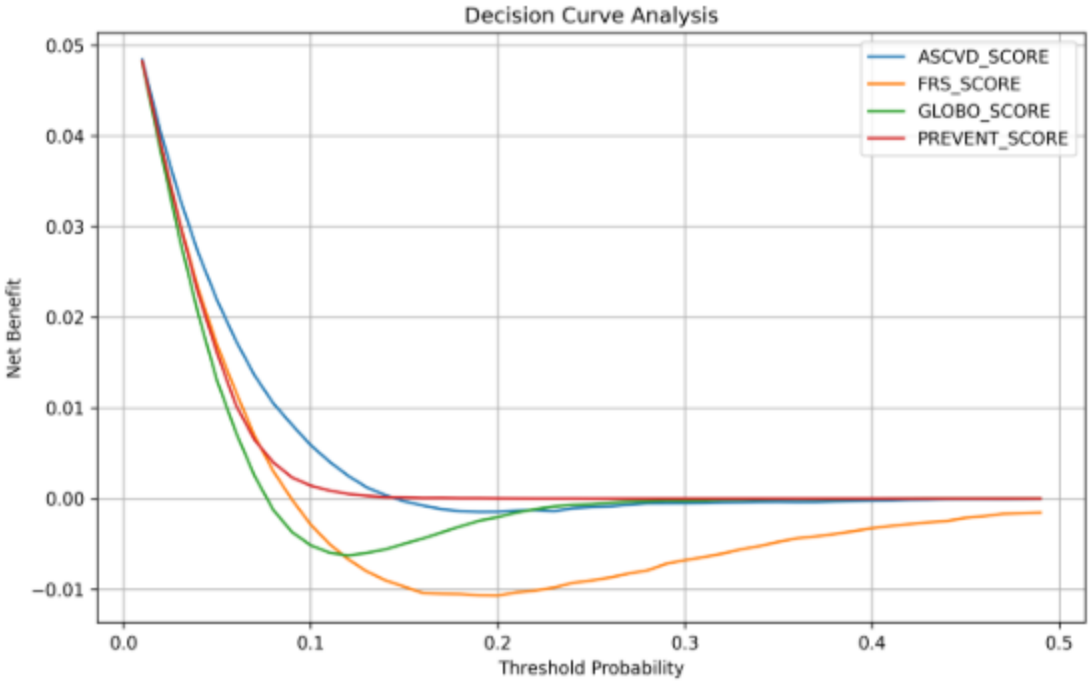


Figure 3

Decision Curve Analysis for the four cardiovascular risk prediction models.

Figure 3. Decision Curve Analysis. Net benefit is plotted against decision threshold probability for ASCVD, FRS, Globorisk, and PREVENT. Higher net benefit across a wider range of thresholds indicates greater clinical utility.

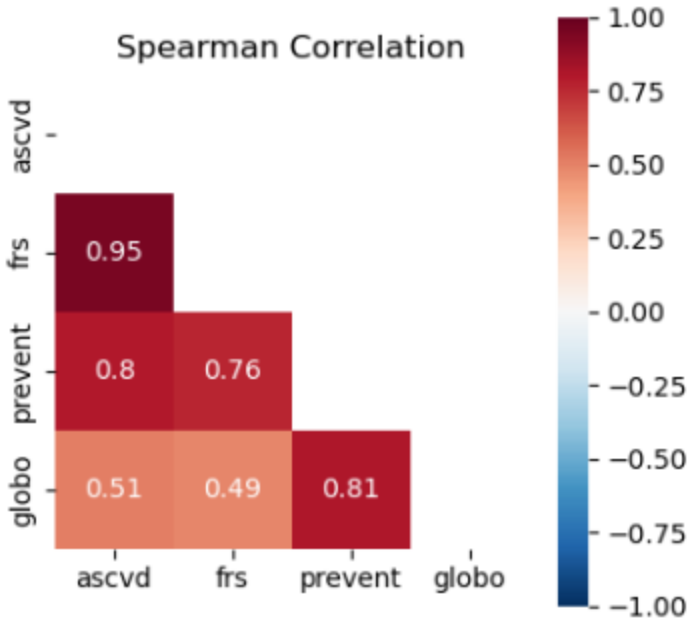


Figure 4

Spearman rank correlation matrix for the four risk prediction models.

Figure 4. Spearman rank correlation matrix. Cell values represent pairwise Spearman rho coefficients. All p < 0.001.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterial.docx