



Cardiovascular disease detection using machine learning and carotid/femoral arterial imaging frameworks in rheumatoid arthritis patients

George Konstantonis¹ · Krishna V. Singh² · Petros P. Sfikakis¹ · Ankush D. Jamthikar^{3,4} · George D. Kitas^{5,6} · Suneet K. Gupta⁷ · Luca Saba⁸ · Kleio Verrou⁹ · Narendra N. Khanna¹⁰ · Zoltan Ruzsa¹¹ · Aditya M. Sharma¹² · John R. Laird¹³ · Amer M. Johri¹⁴ · Manudeep Kalra¹⁵ · Athanasios Protogerou¹⁶ · Jasjit S. Suri¹⁷

Received: 25 October 2021 / Accepted: 29 November 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

The study proposes a novel machine learning (ML) paradigm for cardiovascular disease (CVD) detection in individuals at medium to high cardiovascular risk using data from a Greek cohort of 542 individuals with rheumatoid arthritis, or diabetes mellitus, and/or arterial hypertension, using conventional or office-based, laboratory-based blood biomarkers and carotid/femoral ultrasound image-based phenotypes. Two kinds of data (CVD risk factors and presence of CVD—defined as stroke, or myocardial infarction, or coronary artery syndrome, or peripheral artery disease, or coronary heart disease) as ground truth, were collected at two-time points: (i) at visit 1 and (ii) at visit 2 after 3 years. The CVD risk factors were divided into three clusters (conventional or office-based, laboratory-based blood biomarkers, carotid ultrasound image-based phenotypes) to study their effect on the ML classifiers. Three kinds of ML classifiers (Random Forest, Support Vector Machine, and Linear Discriminant Analysis) were applied in a two-fold cross-validation framework using the data augmented by synthetic minority over-sampling technique (SMOTE) strategy. The performance of the ML classifiers was recorded. In this cohort with overall 46 CVD risk factors (covariates) implemented in an online cardiovascular framework, that requires calculation time less than 1 s per patient, a mean accuracy and area-under-the-curve (AUC) of 98.40% and 0.98 ($p < 0.0001$) for CVD presence detection at visit 1, and 98.39% and 0.98 ($p < 0.0001$) at visit 2, respectively. The performance of the cardiovascular framework was significantly better than the classical CVD risk score. The ML paradigm proved to be powerful for CVD prediction in individuals at medium to high cardiovascular risk.

Keywords Cardiovascular risk estimation · Cardiovascular disease · Three-year follow-up · Conventional risk factors · Ultrasound · And machine learning

Abbreviations

ANOVA	Analysis of variance	Cluster 3	Fusion of office-based biomarker, laboratory-based biomarker, and carotid ultrasound image phenotypes
ASCVD	Atherosclerotic cardiovascular disease	CUSIP	Carotid ultrasound image phenotype
AUC	Area-under-the-curve	CV	Cross-validation
BMI	Body mass index	CVD	Cardiovascular disease
CAD	Coronary artery disease	CVD-3YFU	Cardiovascular disease risk-three-year follow-up
CCVRC	Conventional cardiovascular risk calculators	CVD-CR	Cardiovascular disease-current risk
Cluster 1	Conventional office-based biomarkers	CVE	Cardiovascular events
Cluster 2	Fusion of office-based biomarker and laboratory-based biomarkers	DM	Diabetes mellitus
		FH	Family history
		FNR	False-negative rate
		FPR	False-positive rate
		FRS	Framingham risk score
		HTN	Hypertension

✉ Jasjit S. Suri
jasjit.suri@atheropoint.com

Extended author information available on the last page of the article

IPN	Intraplaque neovascularization
LBBM	Laboratory-based biomarker
LDA	Linear discriminant analysis
ML	Machine learning
MPH	Maximum plaque height
NPV	Negative predictive value
OBBM	Office-based biomarker
PE	Performance evaluation matrices
PPV	Positive predictive value
RA	Rheumatoid arthritis
RF	Random forest
ROC	Receiver operating-characteristics
RRS	Reynolds risk score
SCORE	Systematic coronary risk evaluation
SMOTE	Synthetic minority over-sampling technique
SMOTE 5X	Five times synthetic minority over-sampling
SVMrbf	Support vector machine with the radial basis function
TPA	Total plaque area
TRF	Traditional risk factors
WHO	World Health Organization

Introduction

Individuals with rheumatoid arthritis (RA), diabetes mellitus (DM), and arterial hypertension are at increased cardiovascular disease (CVD) risk due to the accumulation of disease-associated or conventional CVD risk factors (e.g., obesity, dyslipidemia) [1, 2]. The patients with RA have a two- to three-fold increased risk of cardiovascular events (CVE) compared to the normal population [3, 4]. Further, RA and CVD share common risk factors such as gender, hyperlipidemia, hypertension, diabetes, body mass index, physical inactivity, and smoking, known as traditional risk factors (TRF). Because of this fact, most of the conventional cardiovascular risk calculators (CCVRC) have been adopted to predict the CVD risk in RA patients.

CCVRC are used to predict the CVD risk in non-RA and RA patients, however they often either under-estimate or over-estimate the CVD risk in patients [5–7]. This is mainly due to their dependence on TRF, which does not fully explain the increased CVD risk, especially in RA [3, 8–11]. Subclinical atherosclerosis is a common early phenomenon in both RA and non-RA individuals, described by the growth of atherosclerotic plaque [12], and can be used in CVD risk stratification beyond TRF.

The use of non-invasive and economical imaging modalities such as carotid and femoral ultrasound can capture the growth of atherosclerotic plaque. Two popular carotid ultrasound image-based phenotypes (CUSIP) such as carotid intima-media thickness and carotid plaque area

are considered as the surrogate indicator of coronary artery disease (CAD) [13–20]. For example, RA patients reportedly have higher values of this CUSIP compared to non-RA patients and, therefore, these atherosclerotic plaque-based phenotypes could be used for an accurate CVD risk assessment [21–24]. There are two major drawbacks with CCVRC (i) they do not integrate CUSIP in their risk prediction model and (ii) they do not handle the complex nonlinear association between several CVD risk factors and the CVE endpoints [12]. Lastly, since CCVRC is regression-based, they have limited capability to handle a large number of risk predictors.

ML-based systems have reported better and promising risk assessment compared to existing CCVRC in non-RA studies [12, 25]. We, therefore, hypothesize, that the ML-based systems can also provide better CVD risk assessment in RA/non-RA patients compared to CCVRC. We further hypothesize that this relation holds when TRF is fused with CUSIP in the ML framework instead of using TRF alone. The objective of the proposed study is to perform ML-based CVD risk stratification (so-called “ML-effect”) of both RA and non-RA patients while handling the class imbalance using synthetic minority over-sampling technique (SMOTE), referred as “SMOTE-effect”, between binary classes (such as CVD and No-CVD). Furthermore, this study examined the performance of ML-based systems at two distinct time points (i) visit 1 and (ii) visit 2 (after 3-year, against three popular CCVRC (such as FRS, ASCVD, and SCORE), referred to as the “Follow-up effect”. Further, for lower cost reasons, especially in low-income countries, it is important to provide CVD risk assessment using the TRF. Therefore, another objective of the proposed study is to investigate the performance of ML-based CVD risk assessment systems considering (i) office-based risk biomarkers (OBBM) alone, (ii) using the combination of OBBM and laboratory-based biomarkers (LBBM), and (iii) and further using the combination of OBBM, LBBM, and CUSIP (referred as “Cluster-effect”). Figure 1 shows the global architecture of the proposed online ML-based CVD risk stratification system. The CVD risk assessment on a test patient is predicted by fusing OBBM, LBBM, CUSIP, and transforming these by the offline training ML-based model. The test-classifier can be the same as the online classifier. The predicted CVD risk is depicted by the binary colour code (red-CVD or green-No CVD).

The four main innovations of this study are as follows: (i) **Cluster Effect:** For the first time, three diverse types of risk predictors such as OBBM, LBBM, and CUSIP are used under the ML framework to predict the CVD risk in RA participants. (ii) **SMOTE Effect:** To handle the problem of unequal classes (also categorically called class-imbalance), this study also shows the performance of an ML-based system with an imbalanced and balance dataset using the SMOTE

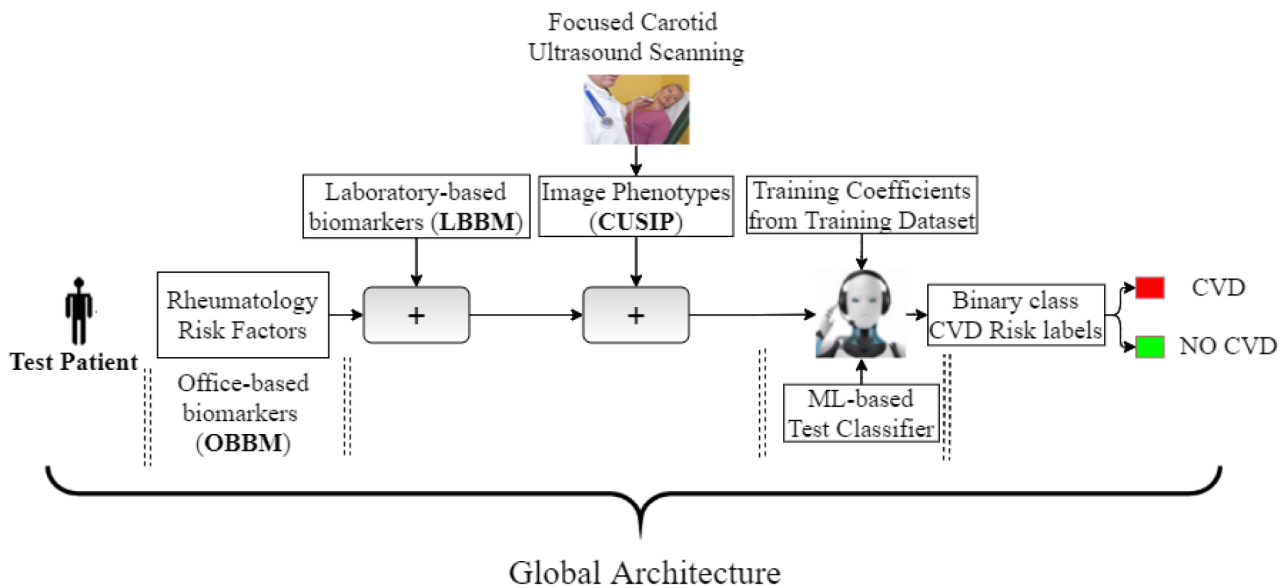


Fig. 1 The global architecture of the proposed ML-based CVD risk stratification system. *CVD* cardiovascular disease, *ML* machine learning

algorithm. (iii) **Follow-up Effect:** This is the first study that provided ML-based CVD risk stratification for RA patients at two distinct time points (visit 1: baseline (CVD-CR) and visit 2: 3-year follow-up (CVD-3YFU). Three popular ML-based algorithms (random forest (RF), support vector machine (SVM), and linear discriminant analysis (LDA)) are used to investigate the CVD risk presence of participants with and without RA. (iv) **ML Effect:** the performance of the ML-based CVD risk assessment system is benchmarked against three popular CCVRC (FRS, SCORE, and ASCVD).

Methodology

Patient demographics

This is a single-center cross-sectional design study at two different time points (visit 1 and visit 2). Consecutive consenting to participate individuals that were examined at the outpatient rheumatology, diabetes and hypertension clinics of the “LAIKO” general hospital in Athens Greece were recruited. All recruited participants were examined at that cardiovascular laboratory of our department, to have ultrasound tests and to optimize their cardiovascular risk stratification. This study used a total of 542 Greek European patients, out of which—at visit 1: baseline—535 individuals patients were free of established CVD and 7 patients had CVD (CVD-CR), such as stroke, or myocardial infarction, or coronary artery syndrome, or peripheral artery disease, or coronary heart disease, or death, collectively referred as CVE. At visit 2, after a 3-year follow-up (3YFU), out

of the 542 patients, 8 patients had established CVD. The mean age of the study participants with CVD-CR was 53.07 ± 13.5 years (ranging between 15 and 90 years), mean weight was 77.27 ± 17.0 Kgs, 37.5% had hypertension, 76% were smokers, and 27.1% had hyperlipidemia.

Ultrasound image acquisition

Carotid and femoral atherosclerotic plaque burden is a surrogate marker for CVE, which is increased in individuals with high CVD risk [24, 26–29]. The carotid and femoral ultrasound examination in all study participants were performed using GE Vivid (Vivid E9 Ultrasound System, GE Healthcare), equipped with a linear probe-type M12L (5.6–14 MHz), as presented in our previous study [30]. The American Society of Echocardiography (ASE) guidelines were used for the ultrasound image acquisition. A dedicated and very experienced (more than 4000 patients examined) trained technician performed all the measurements. CUSIP measurements were recorded from ultrasound scans [30, 31]. The American Society of Echocardiography (ASE) Task Force [32] recommendations were used for image acquisition. The patient was made to lie in the supine position with the head tilted backwards for examining the carotid arteries. Two-step standardized image acquisition protocol was adapted (i) carotid arteries were identified in a transverse position orthogonal to the blood flow, and (ii) probe was tilted by 90° to acquire the far wall of the longitudinal scans (parallel to the blood flow) of the carotid arteries. The above procedure was used and discussed in our previous studies [33–36].

Hypothesis

We hypothesize that (i) Cluster effect: the effect of CUSIP as a biomarker in ML framework can improve the CVD detection in RA and non-RA participants and we believe that the accuracy should follow the pattern $(OBBM + LBBM + CUSIP) > (OBBM + LBBM) > OBBM$. (ii) Follow-up effect: using ML, we were able to predict CVD-3YFU, and we believe that the accuracy obtained for CVD-3YFU is better than or equivalent to CVD-CR. (iii) SMOTE effect: Since AI requires a larger sample size for better training and accuracy prediction, we use SMOTE. Therefore, we hypothesize that the ML risk prediction using SMOTE 5X is superior to ML risk prediction without SMOTE. In continuation, we believe that the accuracy using SMOTE 5X for CVD-CR is superior to accuracy without SMOTE for CVD-CR. Further, the accuracy using SMOTE 5X for CVD-3YFU is superior to accuracy without SMOTE for CVD-3YFU. Finally, the accuracy using SMOTE 5X for ML-CVRC is superior to accuracy without SMOTE for ML-CCVRC. (iv) ML effect: lastly, we hypothesize that ML-based calculators for CVD-CR and CVD-3YFU scenarios are better than CCVRC.

Data preparation using SMOTE

CVD risk detection is a two-class problem. The distribution of participants was imbalanced. Considering CVE as an endpoint, the numbers of patients at visit 1 and visit 2 (CVD-CR and CVD-3YFU) setups were only seven and eight, respectively. This imbalanced dataset in terms of endpoint would result in bias of predicting boundary separation towards the majority class and this affects the performance of ML classifiers. So, to overcome this imbalance class issue, we used a well-published and standardized technique called “synthetic minority oversampling technique (SMOTE)” [37]. SMOTE algorithm used for balancing the risk classes and for data augmentation followed a nearest-neighbour technique that generated non-overlapping synthetic samples for the minority class. After using five times SMOTE (SMOTE 5X) ($n = 2710$), the dataset was partitioned in two halves using a standardized two-fold cross-validation technique to train and evaluate the ML-based system.

Overall machine learning architecture

ML-based algorithms can learn and capture the non-linear, and more complex “information” which is available in the risk features and risk predictors (consisting of OBBM, LBBM, and CUSIP) when together taken as part of the input cohort features. ML uses the gold standards (CVE) along with the covariates (risk factors) to learn the behaviour and

patterns in an offline mode to generate ML-based models. These are then used to transform the test patient covariates into a predicted risk. Such a framework is adopted for CVD-CR setup and CVD-3YFU setup. Since some of the covariates (features) are more powerful than other covariates, therefore, ML-based framework has the ability to orderly select features to increase the overall performance. Such ML-based calculators are benchmarked against CCVRC.

The ML architecture follows the standardized architecture where the ML offline training model is generated using the combination of training data and CVE as ground truth [38]. The CVD presence risk is predicted during the online process where the test patient data are transformed by the ML offline model. The accuracy of such a system is estimated using the K2 cross-validation (CV) protocol, where the data sets are divided into two parts, 50% used for training and the remaining 50% used for testing. This is implemented in the cyclic process where the test set is unique. Before CV execution and data partitioning, the system undergoes preprocessing of the data set using SMOTE algorithm. For optimization and preprocessing, we had used the Standard Scalar function for mapping the independent covariates in the range 0 to 1. Principal component analysis [39] with pooling (PCA pooling) was used to select the most significant features. PCA pooling selects covariates according to the magnitude (from high to low in absolute values) of their coefficients. Followed by this, we used Label Encoder to encode class values as integers. The performance of the system is computed using the following parameters such as sensitivity, specificity, positive predictive value, negative predictive value, false-positive rate, false-negative rate, accuracy, and area under curve score. We benchmarked our ML-based calculator against CCVRC. The overall system architecture is shown in.

Feature selection using two methods: PCA pooling and mutual information

PCA pooling is a well-known feature selection technique [39, 40] used for minimizing the dimensionality of the data. On the other hand, mutual information [41] is a feature selection technique that shows which covariate is highly significant in predicting the target ground-truth label. The PCA pooling technique is used to increase the performance of the CVD-CR and CVD-3YFU by finding the best cluster order among the OBBM, LBBM, and CUSIP. The decreasing order of selecting features using PCA pooling is shown in Table A1 of Appendix A for both CVD-CR and CVD-3YFU. The top covariate is the highest predictive power (LCB PLQ) and the bottom one is the least power (Total number of plaques).

Table 1 Cluster Effect: without SMOTE for CVD-CR and CVD-3YFU

SN	Clusters	CVD-CR			CVD-3YFU		
		LDA	SVMrbf	RF	LDA	SVMrbf	RF
1	Cluster 1	82.10 ± 14.94%	98.71 ± 0.18%	98.71 ± 0.32%	83.03 ± 13.65%	98.52 ± 0.00%	98.52 ± 0.00%
2	Cluster 2	84.69 ± 14.21%	98.71 ± 0.18%	98.71 ± 0.32%	95.02 ± 3.51%	98.52 ± 0.00%	98.52 ± 0.00%
3	Cluster 3	86.53 ± 11.99%	98.71 ± 0.18%	98.71 ± 0.32%	95.39 ± 2.77%	98.52 ± 0.00%	98.52 ± 0.00%

RF random forest; SVMrbf support vector machine with radial basis function; LDA linear discriminant analysis; CVD-CR cardiovascular disease-current risk; CVD-3YFU cardiovascular disease risk-three-year follow-up

Training/Testing classifiers and performance evaluation metrics

We used 3 types of classifiers such as random forest (RF) [42], support vector machine with the radial basis function (SVMrbf) [43, 44], and linear discriminant analysis (LDA) [45, 46] to perform CVD-CR and CVD-3YFU using the combination of 46 covariates and CVE (as ground truth) in the ML framework. The theory and the optimization parameters of these classifiers are shown in Appendix B. The performance evaluation metrics are shown in Appendix C. Further, we compare the ML calculators (RF, SVMrbf, and LDA) against the conventional calculators (FRS, SCORE, and ASCVD) [23, 47].

Experimental protocol

Experiment 1: CVD risk prediction using cross-validation (PCA vs. MI)

The objective of this experiment is to predict the CVD risk using the K2 type of CV protocol. In this protocol, the patient data are equally divided into 50% training and 50% testing; keep the ratio of CVD to non-CVD patients. The training model is generated (as per Fig. 2) using PCA training-based features, and then the model is applied to the test features to predict the CVD risk. A confusion matrix is determined along with the AUC as part of the performance. The objective is to compare and contrast the feature effect of the PCA vs. MI feature selection methods using the K2 CV protocol.

Experiment 2: Study the effect of features on CVD risk prediction (Cluster Effect)

The objective of this experiment is to study the effect of the features on CVD risk. The accuracy, performance evaluation matrices, and ROC are performed for all three clusters using PCA pooling, K2 CV protocol, with and without SMOTE 5X for both CVD-CR and CVD-3YFU. The accuracy of all three clusters was compared to study the effect of the feature.

Results

This section presents the accuracy results of the three classifiers for the CVD-CR and CVD-3YFU experiment while using PCA and MI-based feature extraction techniques. In the second part of the results, we present the cluster effect in SMOTE framework and compare it with without SMOTE. The second part of Sect. 5 presents the PE for (a) generation of performance matrix parameters and (b) ROC analysis. Finally, part three of the section demonstrate the statistical tests such as Tukey [48] and Shapiro-Wilks [49] for validating the significant covariates computed from Chi square (χ^2) and ANOVA test) [50].

Baseline characteristics

In each participant, 46 variable or CVD risk factors were assessed and were partitioned into 3 clusters. (i) cluster 1 included OBBM (see Table A2 of Appendix A) consisting of baseline characteristics from row number R1 to row number R13. This consisted of risk factors such as sex, age, weight, height, body mass index (BMI), average systolic blood pressure (AvSBP), average diastolic blood pressure (AvDBP), heart rate (HR), family history of coronary artery disease (Family His CAD), current smoker, hypertension, diabetes, and hyperlipidemia. (ii) cluster 2 included both OBBM and LBBM (see Table A2 baseline characteristics from row number R1 to row number R34), and (iii) cluster 3 included OBBM, LBBM, and CUSIP (Table A2 baseline characteristics from row number R1 to row number R46). There are two columns in the participant's characteristics: Part A consists of data recorded at visit 1 and Part B consist of the same variable recorded at visit 2 (after 3-year follow-up). Statistical tests were used to compute the p value. Note that in our cohort, we observed that only one more patient was added to the CVD pool from visit 1 to visit 2.

A. Significant CVD risk factors and established CVD as ground truth at visit 1

We observed five significant covariates which had p value < 0.05 in visit 1 with the presence of CVD events as ground truth (CVD-CR): (R10) current smoker with

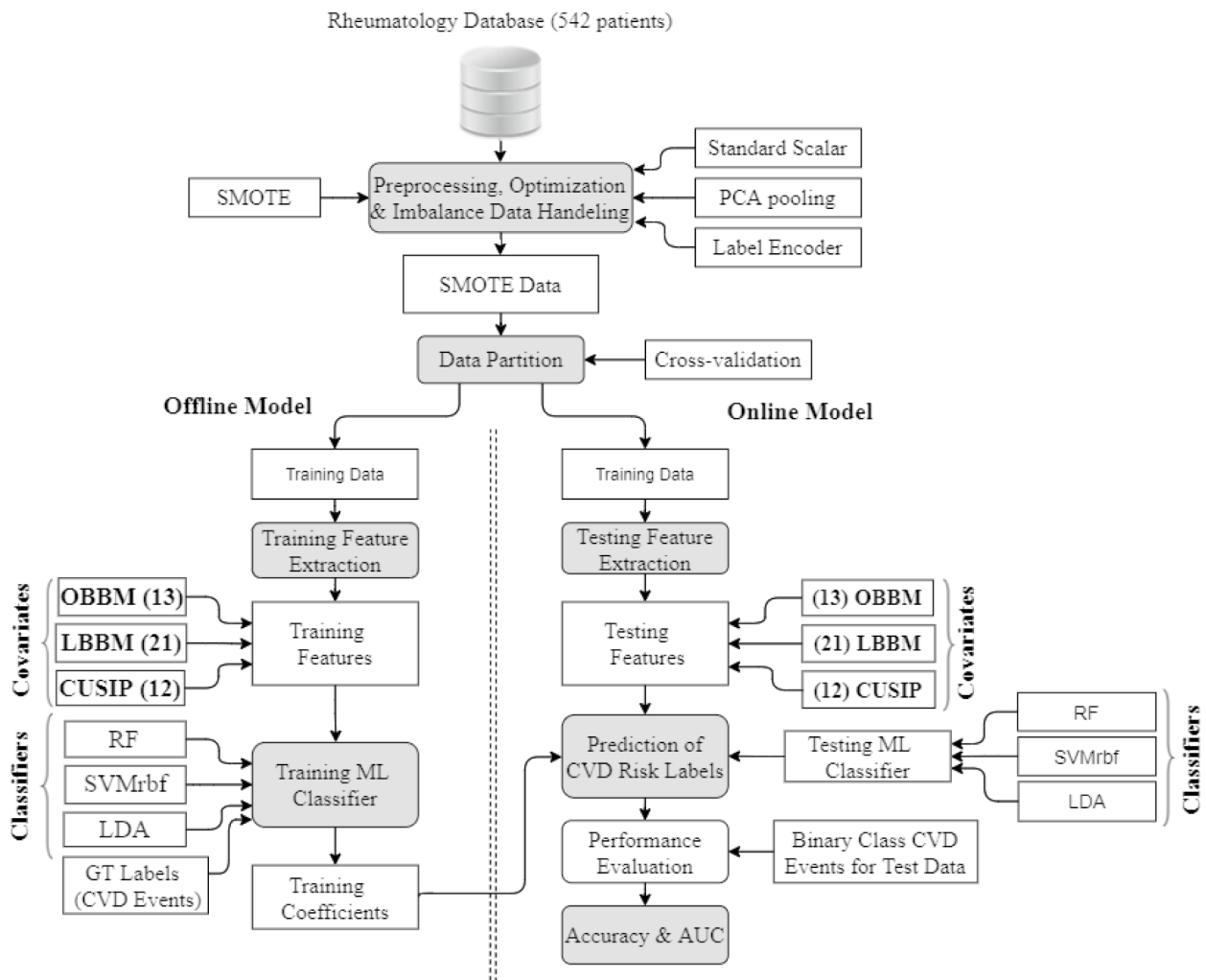


Fig. 2 Overall system architecture for ML-based CVD risk stratification system. *ML* machine learning; *AUC* area-under-the-curve; *OBBM* office-based biomarkers; *LBBM* laboratory-based biomarker; *CUSIP* carotid ultrasound image phenotype; *RF* random forest;

SVMrbf support vector machine with radial basis function; *LDA* linear discriminant analysis; *GT* ground truth; *CVD* cardiovascular disease

p value 0.006, (R16) c-reactive protein (CRP) with p value 0.018, (R17) cholesterol (Chol) with p value 0.023, (R32) low-density lipoprotein(LDL) with p value 0.026 and (R33) urea (UA) with p value 0.033.

B. Significant CVD risk factors and established CVD as ground truth at visit 2

There were six significant covariates which had p value < 0.05 in visit 2 with the presence of CVD as ground truth (CVD-CR): (R10) current smoker with p value 0.007, (R16) CRP with p value 0.008, (R17) Chol with p value 0.010, (R32) LDL with p value 0.017, (R33) UA with p value 0.040, and (R34) thyroid-stimulating hormone (TSH) with p value 0.013.

Accuracy for CVD-CR and CVD-3YFU (MI vs. PCA pooling)

The accuracy comparison of Mutual Information vs. PCA pooling is shown in Table A3 of Appendix A. The mean accuracy for PCA pooling (97.98%) is superior to the mean accuracy of MI (97.06%) for CVD-CR and the mean accuracy of PCA pooling (97.48%) is also superior to the mean accuracy of MI (97.34%) for CVD-3YFU. All the three classifiers had greater accuracy for PCA pooling than MI for both the experiments (CVD-CR and CVD-3YFU). Therefore, we had proved that feature selection using PCA pooling is better than MI.

Fig. 3 Mean Accuracy (%) using SMOTE 5X for CVD-CR. *RF*: random forest, *SVMrbf*: support vector machine with radial basis function, *LDA*: linear discriminant analysis, *Cluster 1*: OBBM, *Cluster 2*: fusion of OBBM, and LBBM, and *Cluster 3*: fusion of OBBM, LBBM, and CUSIP

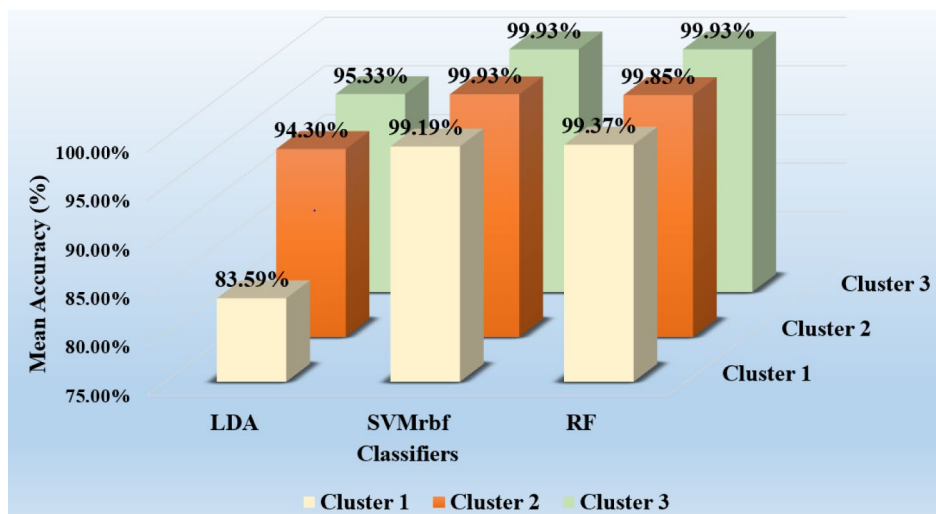
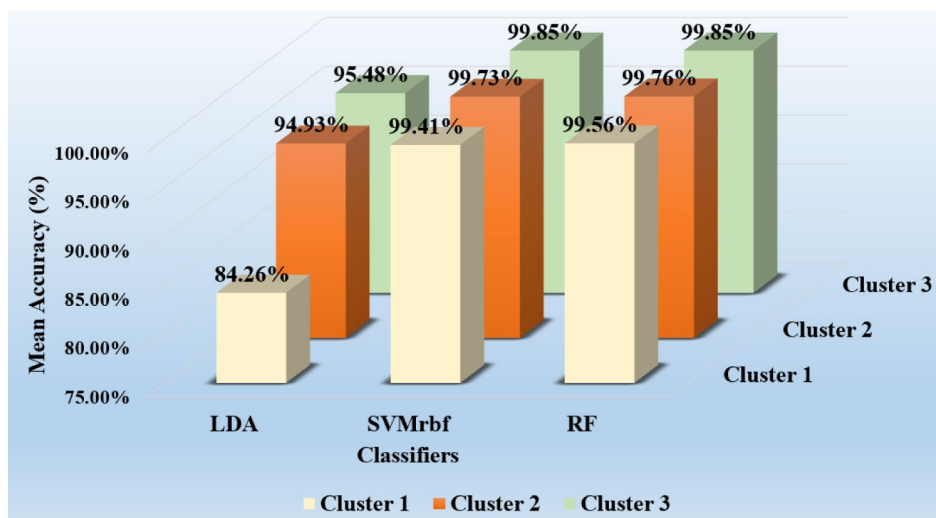


Fig. 4 Mean Accuracy (%) using SMOTE 5X for CVD-3YFU. *RF* random forest, *SVMrbf* support vector machine with radial basis function, *LDA* linear discriminant analysis, *Cluster 1* OBBM, *Cluster 2* fusion of OBBM, and LBBM, and *Cluster 3* fusion of OBBM, LBBM, and CUSIP



Effect of clusters on CVD-CR and CVD-3YFU

This section presents the effect of the features when taken in the form of clusters. In Fig. 3, the bar chart shows the mean accuracy (in %) using SMOTE 5X for CVD-CR setup, while Fig. 4 shows the mean accuracy (in %) using SMOTE 5X for CVD-3YFU setup. Both bar charts show the distribution of accuracy (z-axis) for the three classifiers (X-axis), for the corresponding clusters (Y-axis). Note that cluster 1 is OBBM, cluster 2 is OBBM combined with LBBM, while cluster 3 is OBBM, LBBM, and CUSIP. The bar chart for CVD-CR in Fig. 3 clearly shows the increase in the accuracy due to an increase in covariate strengths (cluster 3 > cluster 2 > cluster 1). This has the same behaviour in Fig. 4 for CVD-3YFU. This validates our hypothesis that with an increase in information due to covariate, all classifiers show better performance. Table 1 shows the cluster effect under non-SMOTE conditions, while Table 2 shows the cluster effect under SMOTE conditions.

Performance evaluation

This section has two parts: Part A presents the performance metrics of three classifiers corresponding to the three clusters in SMOTE 5X framework for CVD-CR and CVD-3YFU, respectively. Part B shows the ROC curves for six classifiers (three ML and three CCVRC) in SMOTE 5X framework for cluster 1, cluster 2, and cluster 3 under CVD-CR condition and CVD-3YFU.

Table 3 shows the performance matrices (PE) for ML vs. CCVRC calculators corresponding to three clusters. The PE metrics consists of the following attributes: sensitivity, specificity, false-positive rate (FPR), false-negative rate (FNR), positive predictive value (PPV), negative predictive value (NPV), accuracy, and area-under-curve (AUC) along with their p value and its ranges. This is represented by column C1 to column C11 in Table 3. Note that these metrics are computed in SMOTE 5X framework for the CVD-CR

condition, while Table 4 shows the computation in SMOTE 5X for CVD-3YFU.

ROC comparison between three types of ML vs. three types of CCVRC for all three clusters

A comparison between all three ML classifiers (LDA, SVMrbf, and RF) against three CCVRC (FRS, SCORE, and ASCVD) for cluster 1 (OBBM), cluster 2 (OBBM+LBBM), and cluster 3 (OBBM+LBBM+CUSIP) using SMOTE 5X framework are shown in Fig. 5(a, b, c) corresponding to CVD-CR). Similar ROC curves can be seen in Fig. 6(a, b, c) for CVD-3YFU. The ML-based classifiers performed better than the CCVRC and the AUC score of all three ML-based classifiers ($0.98, p < 0.001$) were greater than the AUC score ($0.76, p < 0.001$) of all three CCVRCs for all clusters.

Validation of significant covariates using Tukey/Shapiro–Wilk Tests for CVD risks

In baseline characteristics, Chi square (χ^2) test [50] was used to compare categorical covariates and the ANOVA test [51] was used for continuous covariates. Tukey's range test, also known as "Tukey's honestly significant difference (HSD) posthoc test" [48] was used to study the multiple comparisons of CVD risk for both CVD-CR and CVD-3YFU (see section S1 of the Supplementary material). Tukey's test always used only those covariates that showed a highly significant association with CVD-CR and CVD-3YFU. Shapiro–Wilk test [49] (Table A4), a nonparametric test was used to study the normality of the continuous risk covariates. All the continuous covariates were manifest as mean \pm SD and categorical covariates were expressed as percentages. All the statistical tests were two-tailed as the significance cut-off is < 0.05 . There were five significant covariates for the baseline characteristics in CVD-CR and six significant covariates in CVD-3YFU. Shapiro–Wilk tests were needed to assure that all the significant covariates from baseline characteristics should also be in the Shapiro–Wilk test. Tukey HSD posthoc is needed to show that the p value from baseline should be the same as p values from Tukey's test for the binary-class CVD risk. The validation of significant

covariates using Tukey HSD post hoc and Shapiro–Wilk test is shown in Table 5 (for CVD-CR) and Table 6 (for CVD-3YFU), which signifies that the covariates which had significant cut-off $p < 0.05$ from baseline characteristics are highly significant and validated.

Scientific validation

The validation of the CVD risk prediction system can be done by considering a validation database whose results are known prior. For our system, we used a coronary artery database consisting of 500 subjects having 39 covariates. The test participants had a carotid ultrasound which was conducted at the time of the angiogram [52]. There were 298 participants who had an angiographic score of high-risk and 202 participants had no-risk. The validation dataset for all participants had a mean age of 64.49 ± 10.6 years (ranging from 29 to 95 years), 69.8% were males, 67.59% hypertension, and 57.59% hyperlipidemia. The baseline characteristic of the validation coronary dataset for 500 test participants is shown in Table A5. There were 12 significant covariates that had a p value < 0.05 . We had used a two-fold cross-validation protocol (K2) along with PCA with pooling for best feature selection.

The performance of the cross-validation accuracy for all three classifiers (LDA, SVMrbf, and RF), for two-fold cross-validation (K2) on the validation coronary dataset, is shown in Table 7. In the validation dataset, the ML-based classifiers are also superior to the CCVRC calculators. The brief accuracy comparison between without SMOTE and with SMOTE 5X is shown in Table 7. In the validation coronary dataset, SVMrbf is shown the highest performance of 84.76% with SMOTE 5X. A brief comparison of the AUC score is shown in Table 8. Figure 7 shows the ROC curve for SVMrbf and RF in SMOTE 5X framework, both having the same value of 0.92.

Our validation study had similar behaviour as the experimental database showing: (i) superior performance of the ML-based calculators against CCVRC calculators (Fig. 7), (ii) the performance of ML-based calculators with SMOTE 5X is superior to the performance of the ML-based calculators without SMOTE (Table 7), (iii) AUC score of ML-based

Table 2 Cluster Effect: with SMOTE 5X for CVD-CR and CVD-3YFU

SN	Clusters	CVD-CR			CVD-3YFU		
		LDA	SVMrbf	RF	LDA	SVMrbf	RF
1	Cluster 1	83.59 \pm 0.19%	99.19 \pm 0.22%	99.37 \pm 0.56%	84.26 \pm 0.26%	99.41 \pm 0.07%	99.56 \pm 0.00%
2	Cluster 2	94.30 \pm 1.48%	99.93 \pm 0.00%	99.85 \pm 0.00%	94.93 \pm 2.19%	99.73 \pm 0.07%	99.76 \pm 0.04%
3	Cluster 3	95.33 \pm 1.48%	99.93 \pm 0.07%	99.93 \pm 0.07%	95.48 \pm 1.78%	99.85 \pm 0.07%	99.85 \pm 0.07%

RF random forest; SVMrbf support vector machine with radial basis function; LDA linear discriminant analysis; CVD-CR cardiovascular disease-current risk; CVD-3YFU cardiovascular disease risk-three-year follow-up

Table 3 Performance metrics: ML (RF, SVMrbf, and LDA) vs. CCVRC (FRS, SCORE, ASCVD) calculators in the three clusters for CVD-CR

C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11		
											Risk calculator	Sensitivity (%)
Cluster 1 (OBMM)												
	RF	99.85	98.89	98.9	99.85	1.11	0.15	99.37	1	<0.0001	1	1
	SVMrbf	100	98.37	98.4	100	1.63	0	99.19	1	<0.0001	1	1
	LDA	88.89	78.3	80.38	87.57	21.7	11.11	83.59	0.9	<0.0001	0.89	0.91
	FRS	84.44	59.63	67.66	79.31	40.37	15.56	72.04	0.79	<0.0001	0.77	0.81
	SCORE	73.56	58.67	64.02	68.93	41.33	26.44	66.11	0.75	<0.0001	0.74	0.77
	ASCVD	95.33	53.19	67.07	91.93	46.81	4.67	74.26	0.78	<0.0001	0.76	0.8
	Mean ± SD	90.36 ± 10.27	74.51 ± 20.52	79.41 ± 15.93	87.93 ± 12.16	25.49 ± 20.52	9.66 ± 10.27	82.43 ± 14.21	0.87 ± 0.11	-	-	-
Cluster 2 (OBMM + LBBM)												
	RF	99.93	99.78	99.78	99.93	0.22	0.07	99.85	1	<0.0001	1	1
	SVMrbf	100	99.85	99.85	100	0.15	0	99.93	1	<0.0001	1	1
	LDA	100	88.59	89.76	100	11.41	0	94.3	0.98	<0.0001	0.97	0.98
	FRS	81.93	60.59	67.52	77.02	39.41	18.07	71.26	0.79	<0.0001	0.77	0.81
	SCORE	71.26	59.48	63.75	67.42	40.52	28.74	65.37	0.72	<0.0001	0.71	0.74
	ASCVD	93.11	50.59	65.33	88.02	49.41	6.89	71.85	0.77	<0.0001	0.76	0.79
	Mean ± SD	91.04 ± 11.98	76.48 ± 22.12	81.0 ± 17.38	88.73 ± 13.94	23.52 ± 22.12	8.96 ± 11.98	83.76 ± 15.92	0.88 ± 0.13	-	-	-
Cluster 3 (OBMM + LBBM + CUSIP)												
	RF	99.93	99.93	99.93	99.93	0.07	0.07	99.93	1	<0.0001	1	1
	SVMrbf	100	99.7	99.7	100	0.3	0	99.93	1	<0.0001	1	1
	LDA	100	90.67	91.46	100	9.33	0	95.33	0.99	<0.0001	0.98	0.99
	FRS	81.93	60.67	67.56	77.05	39.33	18.07	71.3	0.79	<0.0001	0.77	0.81
	SCORE	71.26	59.48	63.75	67.42	40.52	28.74	65.37	0.72	<0.0001	0.71	0.74
	ASCVD	93.11	50.74	65.4	88.05	49.26	6.89	71.93	0.77	<0.0001	0.76	0.79
	Mean ± SD	91.04 ± 11.98	76.87 ± 22.32	81.30 ± 17.54	88.74 ± 13.93	23.14 ± 22.32	8.96 ± 11.98	83.97 ± 16.06	0.88 ± 0.13	-	-	-

SVMrbf support vector machine with radial basis function; RF random forest; LDA linear discriminant analysis; FRS Framingham risk score; SCORE systematic coronary risk evaluation; ASCVD atherosclerotic cardiovascular disease; FPR false-positive rate; FNR false-negative rate; AUC area under the curve; SD standard deviation; CI-C11 column 1 to column 11; PPV positive predictive value; NPV negative predictive value; OBMM office-based biomarker; laboratory-based biomarker; CUSIP carotid ultrasound image phenotype

Table 4 Performance matrices: ML (RF, SVMrbf, and LDA) vs. CCVRC (FRS, SCORE, ASCVD) calculators in the three clusters for CVD-3YFU

C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	p value	AUC	Lower	Upper
Cluster 1 (OBMM)														
	RF	99.78	99.33	99.34	99.78	0.67	0.22	99.56	1	<0.0001	1	1	1	1
	SVMrbf	99.98	98.81	98.83	100	1.19	0	99.41	1	<0.0001	1	1	1	1
	LDA	88.96	79.56	81.31	87.82	20.44	11.04	84.26	0.92	<0.0001	0.9	0.93	0.93	0.93
	FRS	87.33	59.04	68.07	82.33	40.96	12.67	73.19	0.79	<0.0001	0.78	0.81	0.78	0.81
	SCORE	69.56	57.48	62.06	65.37	42.52	30.44	63.52	0.74	<0.0001	0.72	0.75	0.72	0.75
	ASCVD	95.04	52.15	66.51	91.31	47.85	4.96	73.59	0.76	<0.0001	0.74	0.78	0.74	0.78
	Mean ± SD	90.11 ± 11.37	74.40 ± 21.27	79.35 ± 16.57	87.77 ± 12.94	25.61 ± 21.27	9.88 ± 11.38	82.26 ± 14.87	0.87 ± 0.12	—	—	—	—	—
Cluster 2 (OBMM + LBBM)														
	RF	99.98	99.93	99.93	100	0.07	0	99.96	1	<0.0001	1	1	1	1
	SVMrbf	99.98	99.85	99.85	100	0.15	0	99.93	1	<0.0001	1	1	1	1
	LDA	95.96	89.85	90.79	100	10.15	0	94.93	0.98	<0.0001	0.97	0.99	0.97	0.99
	FRS	85.93	58.74	67.56	80.67	41.26	14.07	72.33	0.8	<0.0001	0.78	0.82	0.78	0.82
	SCORE	71.11	58.37	63.07	66.89	41.63	28.89	64.74	0.73	<0.0001	0.71	0.75	0.71	0.75
	ASCVD	93.78	49.7	65.09	88.87	50.3	6.22	71.74	0.77	<0.0001	0.75	0.79	0.75	0.79
	Mean ± SD	91.12 ± 11.08	76.07 ± 22.95	81.04 ± 17.69	89.41 ± 13.57	23.93 ± 22.95	8.19 ± 11.56	83.94 ± 16.03	0.88 ± 0.13	—	—	—	—	—
Cluster 3 (OBMM + LBBM + CUSIP)														
	RF	99.98	99.7	99.7	100	0.3	0	99.85	1	<0.0001	1	1	1	1
	SVMrbf	100	99.7	99.7	100	0.3	0	99.85	1	<0.0001	1	1	1	1
	LDA	98.29	90.96	91.71	100	9.04	0	95.48	0.98	<0.0001	0.98	0.99	0.98	0.99
	FRS	85.93	58.74	67.56	80.67	41.26	14.07	72.33	0.8	<0.0001	0.78	0.82	0.78	0.82
	SCORE	71.11	58.37	63.07	66.89	41.63	28.89	64.74	0.73	<0.0001	0.71	0.75	0.71	0.75
	ASCVD	93.78	49.7	65.09	88.87	50.3	6.22	71.74	0.77	<0.0001	0.75	0.79	0.75	0.79
	Mean ± SD	91.52 ± 11.33	76.20 ± 23.01	81.14 ± 17.72	89.41 ± 13.57	23.81 ± 23.01	8.20 ± 11.56	83.99 ± 16.07	0.88 ± 0.13	—	—	—	—	—

SVMrbf support vector machine with radial basis function; RF random forest; LDA linear discriminant analysis; FRS Framingham risk score; SCORE systematic coronary risk evaluation; ASCVD atherosclerotic cardiovascular disease; FPR false-positive rate; FNR false-negative rate; AUC area under the curve; SD standard deviation; CI-C11 column 1 to column 11; PPV positive predictive value; NPV negative predictive value; OBMM office-based biomarker; laboratory-based biomarker; CUSIP carotid ultrasound image phenotype

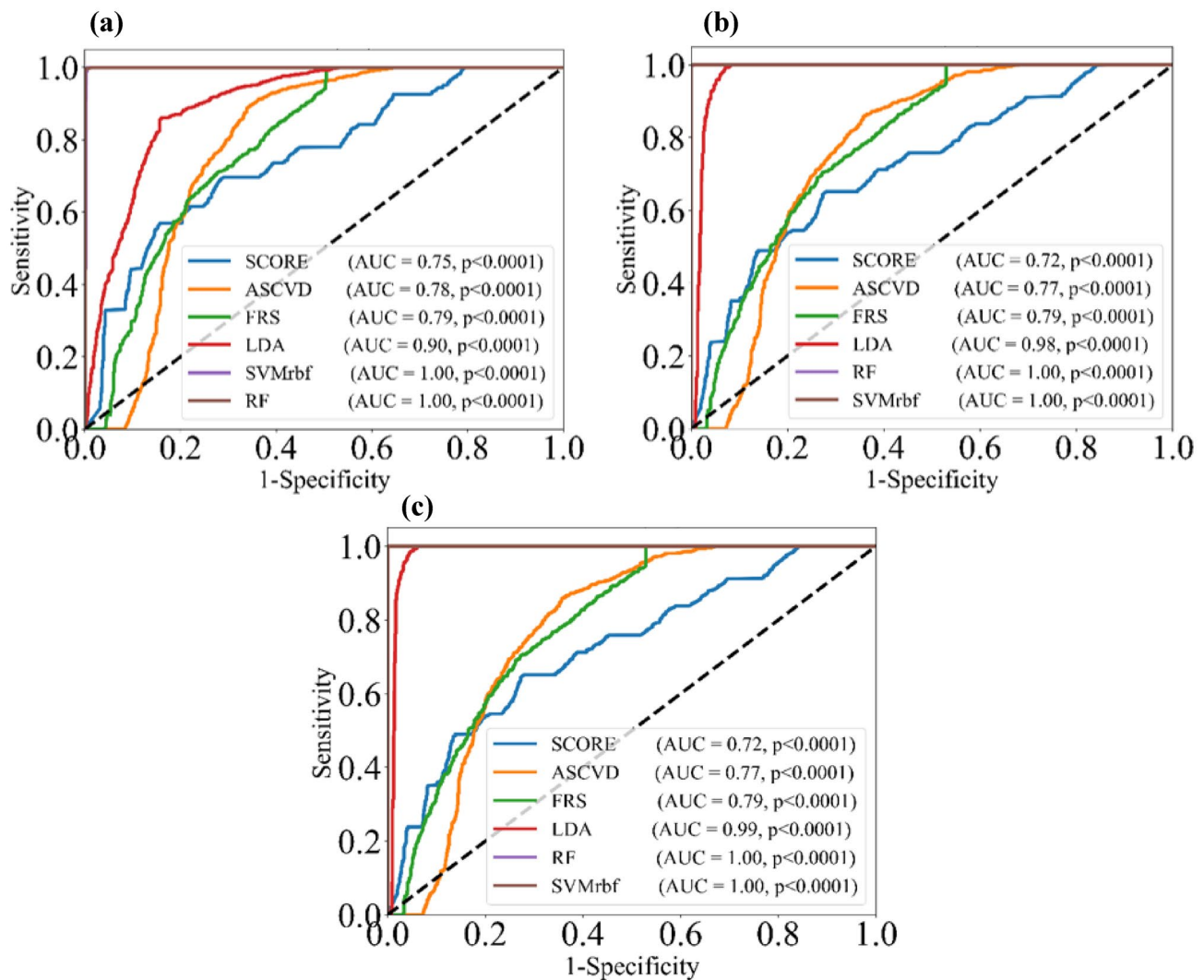


Fig. 5 AUC curve for (a) Cluster 1, (b) Cluster 2, and (c) Cluster 3, respectively for CVD-CR; *RF* random forest; *SVMrbf* support vector machine with radial basis function; *LDA* linear discriminant analysis;

AUC area-under-curve; *FRS* Framingham Risk Score; *SCORE* systematic coronary risk evaluation, *ASCVD* atherosclerotic cardiovascular disease

calculators with SMOTE 5X is superior to ML-based calculators without SMOTE (Table 8). Thus, the validation proved our hypothesis to be true.

Discussions

This is the first pilot study of its kind where the CVD presence can be detected using current covariates using two-time points (i) visit 1 current CVE as ground truth, and (ii) visit 2 (3-year follow-up) CVE as ground truth. The AthroEdge™ 3.0 (ML) used 46 covariates having both non-RA and RA patients. Using SMOTE framework and feature extraction paradigms such as PCA pooling or MI, our system

was able to generate superior performance with cross-validation K2 protocol while validating the following hypothesis: (i) *Cluster Effect*: The ML classifiers showed superior performance when carotid image phenotypes were fused with conventional and laboratory-based risk factors, unlike without carotid image phenotypes (Table 9); (ii) *SMOTE Effect*: CVD detection was superior in SMOTE framework, unlike in non-SMOTE framework, both when using current CVE (Table 10) or 3YFU CVE (Table 11) (iii) *Follow-up Effect*: CVD risk prediction using 3-year CVE showed equal or superior to CVD risk prediction with current CVE (Table 12); (iv) *ML Effect*: ML-based calculators showed superior performance compared to CCVRC (Table 13). The overall ML-based framework is automated and uses the

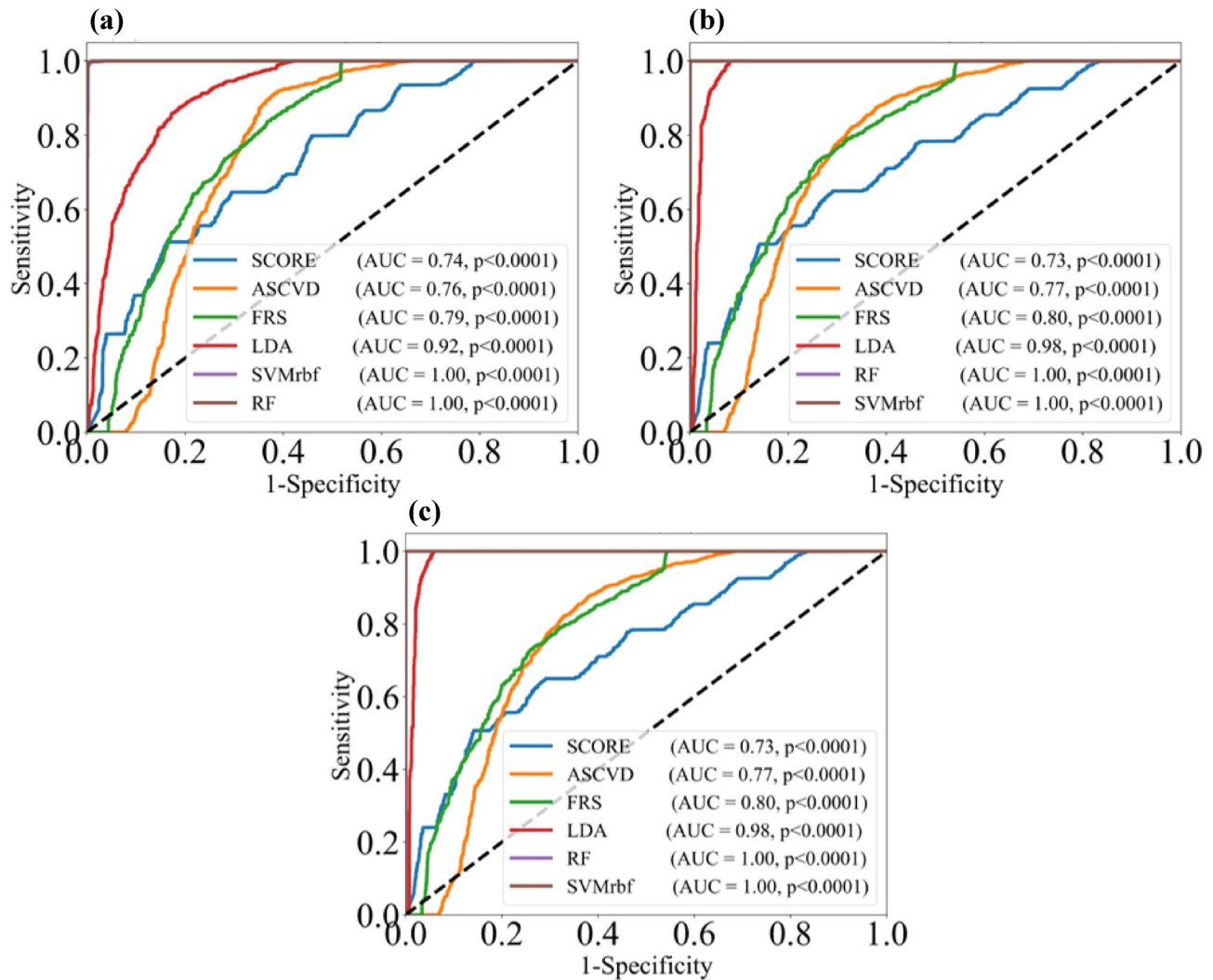


Fig. 6 AUC curve for (a) Cluster 1, (b) Cluster 2, and (c) Cluster 3, respectively for CVD-3YFU. RF: random forest; SVMrbf: support vector machine with radial basis function; LDA: linear discriminant

analysis; AUC: area-under-curve; FRS Framingham Risk Score; SCORE systematic coronary risk evaluation, ASCVD atherosclerotic cardiovascular disease

Table 5 Validating significant covariates using Tukey's HSD & Shapiro-Wilk test for CVD-CR

SN	Covariate Type	Tukey's Test	Shapiro-Wilk Test	Baseline (Chi-square and ANOVA)	Validated
1	Curr-Smoker	0.006	<0.0001	0.006	✓
2	CRP	0.018	<0.0001	0.018	✓
3	Chol	0.023	<0.0001	0.023	✓
4	LDL	0.026	<0.0001	0.026	✓
5	UA	0.033	<0.0001	0.033	✓

Curr-Smoker Current Smoker; CRP C-reactive protein; Chol cholesterol; LDL low-density lipoprotein; UA uric acid

Scikit-learn technique [53], which is an open-source library using Python. Our ML-based framework is user-friendly, feasible and very flexible to run.

Benchmarking

There are not many AI techniques for the CVD risk assessment in cohorts that include rheumatology patients.

Table 6 Validating significant covariates using Tukey’s HSD & Shapiro–Wilk test for CVD-3YFU

SN	Covariate type	Tukey’s test	Shapiro–Wilk test	Baseline (Chi square and ANOVA)	Validated
1	Curr-Smoker	0.007	<0.0001	0.007	✓
2	CRP	0.008	<0.0001	0.008	✓
3	Chol	0.010	<0.0001	0.010	✓
4	LDL	0.017	<0.0001	0.017	✓
5	UA	0.040	<0.0001	0.040	✓
6	TSH	0.013	<0.0001	0.013	✓

Curr-Smoker Current Smoker; *CRP* C-reactive protein; *Chol* cholesterol; *LDL* low-density lipoprotein; *UA* uric acid

Table 7 CV accuracy (%) for coronary data without and with SMOTE 5X for K2

SN	Classifiers	Without SMOTE	With SMOTE 5X	% Improvement
1	LDA	64.80 ± 0.00%	70.48 ± 0.95%	0.08
2	SVMrbf	65.20 ± 3.20%	84.76 ± 1.59%	0.23
3	RF	65.20 ± 3.20%	81.75 ± 1.27%	0.20

RF random forest; *SVMrbf* support vector machine with radial basis function; *LDA* linear discriminant analysis; *5X* five times

Table 8 AUC score for coronary data without and with SMOTE 5X for K2

SN	Classifiers	Without SMOTE	With SMOTE 5X	% Improvement
1	LDA	0.67	0.76	0.11
A2	SVMrbf	0.53	0.92	0.42
3	RF	0.64	0.92	0.30

RF random forest; *SVMrbf* support vector machine with radial basis function; *LDA* linear discriminant analysis; *5X* five times

However, AI techniques have been used to assess the severity of RA without considering the CVD risk [54–56]. Some of the reasons were that there is not much trend between people to estimate rheumatoid arthritis in patients due to bone join, and there is not much direct link between heart diseases on bone problems. This kind of research needs very dedicated follow-up and more money is needed for bone pain than for heart diseases. Because of that weak link, there are very few papers on rheumatology patients, that is why there are only seven patients in our rheumatology database who had CVD risk assessment [17, 57].

Six out of ten authors have used SVM classifiers for the prediction of CVD/stroke risk assessment. Further, five out of ten authors used follow-up CVE. Seven out of ten had used conventional cardiovascular risk factors as feature types, and three out of ten had used PCA-based technique for the selection of features. The similarity between other authors’ studies and our study is that we had used follow-up

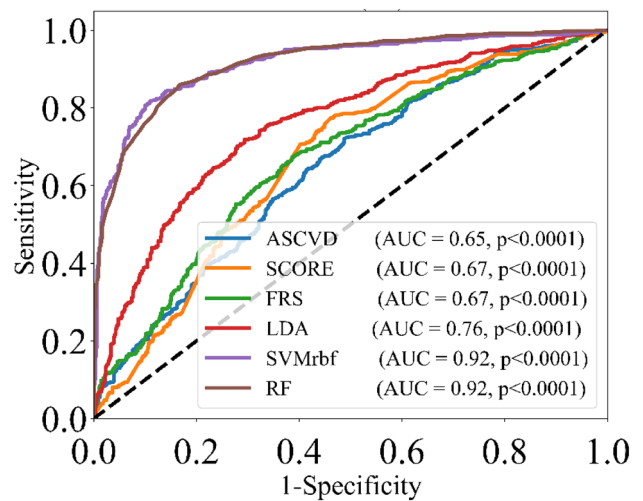


Fig. 7 AUC curve for K2 and SMOTE 5X for the validation coronary dataset. RF: random forest; SVMrbf: support vector machine with radial basis function; LDA linear discriminant analysis; AUC area-under-curve; FRS Framingham Risk Score; SCORE systematic coronary risk evaluation; ASCVD atherosclerotic cardiovascular disease

CVE, PCA-based technique for features selection, and SVM classifier for the prediction of CVD risk assessment (see Table 14, benchmarking).

A series of ML-based publications were proposed by Suri and Jamthikar, which are being summarized as follows: Jamthikar et al. [58] showed the usage of ML-based system (called AtheroRisk-integrated from AtheroPoint, Roseville, CA) performs better than the conventional calculator using PCA-based feature extraction system in K10 paradigm. The database used 13 current risk factors and 34 carotid image-based phenotypes, totalling 47 covariates. The authors used event-equivalent gold standard (EEGS) as percentage stenosis in the carotid artery. The ML-framework used RF classifier leading an AUC of ML system: 0.80 (95% CI: 0.77–0.84) against AUC for CCVRC: 0.68 (95% CI: 0.64–0.72), showing an improvement of 18%. The methodology neither used SMOTE, nor the ground truth had CVE. Further, the system did not consist of RA patients. In another

Table 9 Hypothesis: Cluster effect (Cluster3 > Cluster2 > Cluster1) using accuracy with SMOTE 5X

Clusters	CVD-CR			CVD-3YFU		
	LDA	SVMrbf	RF	LDA	SVMrbf	RF
Cluster 1 (C1)	83.59 ± 0.19%	99.19 ± 0.22%	99.37 ± 0.56%	84.26 ± 0.26%	99.41 ± 0.07%	99.56 ± 0.00%
Cluster 2 (C2)	94.30 ± 1.48%	99.93 ± 0.00%	99.85 ± 0.00%	94.93 ± 2.19%	99.73 ± 0.07%	99.76 ± 0.04%
Cluster 3 (C3)	95.33 ± 1.48%	99.93 ± 0.07%	99.93 ± 0.07%	95.48 ± 1.78%	99.85 ± 0.07%	99.85 ± 0.07%
Hypothesis (C3 > C2 > C1)	✓	✓	✓	✓	✓	✓

RF random forest; SVMrbf: support vector machine with radial basis function; LDA linear discriminant analysis; CVD-CR cardiovascular disease-current risk; CVD-3YFU cardiovascular disease risk-three-year follow-up

Table 10 Hypothesis: SMOTE 5X performance is superior to w/o SMOTE for CVD-CR

Clusters	W/o SMOTE			With SMOTE 5X			Hypothesis
	LDA	SVMrbf	RF	LDA	SVMrbf	RF	
Cluster 1	82.10 ± 14.94%	98.71 ± 0.18%	98.71 ± 0.32%	83.59 ± 0.19%	99.19 ± 0.22%	99.37 ± 0.56%	✓
Cluster 2	84.69 ± 14.21%	98.71 ± 0.18%	98.71 ± 0.32%	94.30 ± 1.48%	99.93 ± 0.00%	99.85 ± 0.00%	✓
Cluster 3	86.53 ± 11.99%	98.71 ± 0.18%	98.71 ± 0.32%	95.33 ± 1.48%	99.93 ± 0.07%	99.93 ± 0.07%	✓

RF random forest; SVMrbf support vector machine with radial basis function; LDA linear discriminant analysis; CVD-CR cardiovascular disease-current risk; CVD-3YFU cardiovascular disease risk-three-year follow-up

Table 11 Hypothesis: SMOTE 5X performance are superior to w/o SMOTE for CVD-3YFU

Clusters	W/o SMOTE			With SMOTE 5X			Hypothesis
	LDA	SVMrbf	RF	LDA	SVMrbf	RF	
Cluster 1	83.03 ± 13.65%	98.52 ± 0.00%	98.52 ± 0.00%	84.26 ± 0.26%	99.41 ± 0.07%	99.56 ± 0.00%	✓
Cluster 2	95.02 ± 3.51%	98.52 ± 0.00%	98.52 ± 0.00%	94.93 ± 2.19%	99.73 ± 0.07%	99.76 ± 0.04%	✓
Cluster 3	95.39 ± 2.77%	98.52 ± 0.00%	98.52 ± 0.00%	95.48 ± 1.78%	99.85 ± 0.07%	99.85 ± 0.07%	✓

RF random forest; SVMrbf support vector machine with radial basis function; LDA linear discriminant analysis; CVD-CR cardiovascular disease-current risk; CVD-3YFU cardiovascular disease risk-three-year follow-up

Table 12 Hypothesis: performance of CVD-3YFU is better or equivalent to CVD-CR

Clusters	CVD-CR			CVD-3YFU			Hypothesis
	LDA	SVMrbf	RF	LDA	SVMrbf	RF	
Cluster 1	83.59 ± 0.19%	99.19 ± 0.22%	99.37 ± 0.56%	84.26 ± 0.26%	99.41 ± 0.07%	99.56 ± 0.00%	✓
Cluster 2	94.30 ± 1.48%	99.93 ± 0.00%	99.85 ± 0.00%	94.93 ± 2.19%	99.73 ± 0.07%	99.76 ± 0.04%	✓
Cluster 3	95.33 ± 1.48%	99.93 ± 0.07%	99.93 ± 0.07%	95.48 ± 1.78%	99.85 ± 0.07%	99.85 ± 0.07%	✓

RF random forest; SVMrbf support vector machine with radial basis function; LDA linear discriminant analysis; CVD-CR cardiovascular disease-current risk; CVD-3YFU cardiovascular disease risk-three-year follow-up

Table 13 Hypothesis: ML-based calculators are better than CCVRC calculators (using AUC score)

Clusters	ML-based calculators			CCVRC calculators			Hypothesis
	LDA	SVMrbf	RF	FRS	SCORE	ASCVD	
CVD-CR	0.99	1.00	1.00	0.79	0.72	0.77	✓
CVD-3YFU	0.98	1.00	1.00	0.80	0.73	0.77	✓

RF random forest; SVMrbf support vector machine with radial basis function; LDA linear discriminant analysis; CVD-CR cardiovascular disease-current risk; CVD-3YFU cardiovascular disease risk-three-year follow-up

study, Jamthikar et al. [59] showed the use of an ML-based classifier called Random Forest for CVD risk prediction using 38 covariates that consisted of 13 conventional risk factors and 25 CUSIP (6 types of current CUSIP, 6 types of 10-year CUSIP, 12 types of quadratic CUSIP (harmonics),

and age-adjusted grayscale median). Using RF classifier and Jack Knief (JK) protocol, the ML-based CVD risk assessment system showed an improvement of (~57%) against the conventional system (AUC = 0.99, p value < 0.001

Table 14 Benchmarking against ML-based CVD/stroke risk assessment in non-RA cohorts

C1	C2	C3	C4	C6	C7	C8	C9	C10
SN	FA(Year)	Nop	FT	Selected Features	Classifiers	GS/GT Label	PE	Benchmarking Against
R1	Gastouniotti et al. [61]	56	Kinematics Features	FDR, PCA, WRS	SVM	FU data labels	ACC: 88%	DT, K-NN, DA, PNN
R2	Unnikrishnan et al. [62]	2406	CCVR-F	NA	SVM	FU data labels	Se: 68.2%, Sp: 85.9%, AUC: 0.71	FRS
R3	Weng et al. [63]	378,256	CCVR-F	-	RFC, LR, GBM, ANN	FU data labels	AUC: 0.764	PCRS
R4	Venkatesh et al. [64]	6814	CCVR-F, I-bf, & Serum Biomarkers	MDMST	RFC, Cox, LASSO-cox, AIC-Cox backward regression	FU data labels	C-Index: 0.81, BS: 0.083	PCRS & FRS
R5	Araki et al. [65]	204	I-bf	Statistical Test	SVM	LD-based risk labels	ACC: NW: 95.08%, FW: 93.47%	-
R6	Banchhor et al. [69]	22	Texture & wall-based features	PCA	SVM	Carotid plaque burden	ACC: 91.28% AUC: 0.91	-
R7	Kakadiaris et al. [70]	6459	CCVR-F	-	SVM	FU data labels	Se: 86%, Sp: 95%, AUC: 0.92	PCRS
R8	Jamthikar et al. [58]	202	CCVR-F and CUS I-bf	PCA polling	RF	Carotid stenosis Surrogate endpoint of CVD	AUC of ML system: 0.80 (95% CI: 0.77–0.84) AUC for CCVRC: 0.68 (95% CI: 0.64–0.72)	-
R9	Jamthikar et al. [59]	202	CCVR-F and CUS I-bf	Logistic Regression	RF	LD as Surrogate endpoint of CVD	AUC: 0.99 ($p < 0.001$)	-
R10	Jamthikar et al. [60]	202	CCVR-F and CUS I-bf	-	SVM	Surrogate endpoint of CVD	AUC: 0.88 ($p < 0.001$)	13 CCVRC
R11	Proposed Study	542	OBBM, LBBM, CUSIP	PCA pooling & MI	RF, SVMrbf, LDA	CVE: (MI, or stroke, or Death), and FU data labels	ACC: 98.40% for CVD-CR and 98.38% for CVD-3YFU, AUC: 0.96 ($p < 0.0001$)	CCVRC (FRS, SCORE, and ASCVD)

SN serial num, Nop number of patients, FU follow-up, CVD cardiovascular disease, CUS carotid ultrasound, LD lumen diameter, LR logistic regression, FDR Fisher discriminant ratio, WRS: Wilcoxon rank-sum, GS/GT labels: gold standard/GT labels, PCA: principal component analysis, MDMST: minimal depth of maximal subtree, SVMC support vector machine classifier, GMM: Gaussian mixture model, DA discriminant analysis, RBPNN: Radial basis probabilistic neural network, I-bf: Image-based features, DT Decision tree, K-NN K-Nearest Neighbors, NB Naïve Bays, FC Fuzzy Classifier, FA first author, QNN: quantum neural network, MLP: multilayer perceptron, RFC: random forest classifier, SOM self organization map, IGR: information gain ranking, DB: database, CCVRC: conventional cardiovascular risk calculators, ANN: artificial neural network, DWT: discrete wavelet transform, HoS higher-order spectra, FT: features types, CCVR-F: conventional cardiovascular risk factors, ACC: accuracy, Se: sensitivity, Sp: specificity, AUC: area under the curve, BS Brier score, PCRS: pooled cohort risk score, and FRS: Framingham risk score, C-Index: Concordance index, NW: near walls, FW: far walls

vs. AUC = 0.63, p value < 0.001), while considering 202 patients.

Jamthikar et al. [60], further showed the advantage of fusing carotid image-based phenotypes in ML framework using SVM. The authors compared their ML system against 13 types of conventional CVD risk calculators having the following performances (AECRS2.0 (AUC 0.83, p < 0.001), QRISK3 (AUC 0.72, p < 0.001), WHO (AUC 0.70, p < 0.001), ASCVD (AUC 0.67, p < 0.001), FRS-cardio (AUC 0.67, p < 0.01), FRS-stroke (AUC 0.64, p < 0.001), MSRC (AUC 0.63, p = 0.03), UKPDS56 (AUC 0.63, p < 0.001), NIPPON (AUC 0.63, p < 0.001), PRO-CAM (AUC 0.59, p < 0.001), RRS (AUC 0.57, p < 0.001), UKPDS60 (AUC 0.53, p < 0.001), and SCORE (AUC 0.45, p < 0.001). The cohort consisted of 202 Japanese patients (156 M/46F) having a mean age of 69 ± 11 years. Their system showed an improvement of 42% increase in performance having an AUC of 0.88 (p < 0.001). Gastouniotti et al. [61] showed the advantage of kinematic features of the atherosclerotic plaque for 56 patients from two different hospitals with 15 computer-aided diagnoses (CAD) schemes in an ML-based framework with several cross-validation strategies. They used Fisher Discriminant Ratio (FDR), PCA, and Wilcoxon Rank-Sum (WRS) for selecting features with follow-up ground truth for the prediction of CVD risk assessment and they got 88% accuracy. Their methodology neither used CCVRC calculators, nor they had shown a feature effect. Further, the system did not consist of RA patients.

Unnikrishnan et al. [62] showed the study of cardiovascular risk assessment in 2406 participants with conventional cardiovascular risk factors in ML-based framework using Support Vector Machine classifier (SVM) classifier, benchmarking against Framingham risk score. Using SVM, the ML-based CVD risk assessment showed an AUC of 0.71, 68.02% sensitivity (Se), and 85.9% specificity. Their methodology neither used feature selection technique nor had shown the effect of different clusters. Weng et al. [63] showed the improvement of cohort study for cardiovascular risk prediction in 378,256 patients from UK family practices using random forest classifier (RFC), logistic regression (LR), gradient boosting machines (GBM), benchmarking against pooled cohort risk score (PCRS), got an overall AUC of 0.764. They neither used feature selection technique nor had benchmarking ML-based calculators against CCVRC. Further, their framework did not consist of RA patients.

Venkatesh et al. [64] predicted the six cardiovascular outcomes in comparison to standard cardiovascular risk scores and they had used 6814 test patients from the MESA (Multi-Ethnic Study of Atherosclerosis), aged between 45 and 84 years, from four different ethnicities, and six different centres across the United States. Using ML-based framework with Minimal Depth of Maximal Subtree as feature selection, random survival forests technique performed

better than established risk scores with increased prediction accuracy, concordance index (C-Index) of 0.81, and Brier Score (BS) of 0.083. Araki et al. [65] showed the study of both the near and far walls of the carotid artery in Stroke risk assessment for 204 patients with Image-based features (I-bf) using Statistical Test as feature selection technique. Using SVM, they got an accuracy of 95.08% in near walls, 93.47% in far walls, respectively.

A special note on SMOTE

SMOTE [37, 66], also known as synthetic minority oversampling technique, is a well-known oversampling method where the synthetic samples were generated for the minority class. It is used for augmentation and for solving imbalance problems and is preferred in a generalized framework. While using SMOTE 5X, there is an increase of performance accuracy for (i) LDA classifier from 86.53% to 95.33%, (ii) SVMrbf classifier from 98.71% to 99.93%, (iii) RF classifier from 98.71% to 99.93% for the current CVD-CR set. The same behaviour was observed when using the CVD-3YFU setup, where the accuracy increased in (i) LDA classifier from 95.39% to 95.48%, (ii) SVMrbf classifier increased from 98.52% to 99.85%, and (iii) RF classifier increased from 98.52% to 99.85%. This is because SMOTE creates the sample of the minority class, augment the dataset, and this technique helped to overcome the overfitting problem.

A special note on CVD risk assessment during follow-up

Even though the performance of classifiers in the CVD-3YFU set-up is equal to or better than CVD-CR, the number of CVE in both scenarios is very low. Only 7 out of 542 had CVE in the CVD-CR setup, while 8 out of 542 had CVE in the CVD-3YFU set-up. This puts a strain on the ML classification system even if the SMOTE is activated. Note that in most of the clinical trials the CVE is typically < 5%, however, in our case, this is < 1.5%. This is one of the major weaknesses of our study. One solution would be to partner with other clinical centres which have a larger number of CVE. The second solution would be having a longer period of follow-up, however, this puts the constraint on the economics of healthcare. From the ML perspective, we successfully demonstrated that ML is stable and comparable in both scenarios (CVD-CR and CVD-3YFU) and easily extendable for larger period follow-up.

Strengths/Weakness/Extensions

ML-based CVD risk prediction is superior to CCVRC and can handle the non-linearity between the covariates, CVE

in CVD-CR setup and CVD-3YFU setu This puts ML-based designs very effective and powerful when carotid ultrasound image-based phenotypes are integrated with the conventional and laboratory-based biomarkers. Further, SMOTE adds a powerful incentive when the sample size is small. The proposed system demonstrated a clear design where the future follow-ups CVD risks can be computed using the ML framework. While the four hypotheses were validated statistically, our dataset had limited cases of CVE cases during CVD-CR and CVD-3YFU setups. This was partially compensated by the SMOTE-based strategy and the SMOTE-based methodology was better in all the classification models. Since this is the first pilot study, there is a clear need for extensions to our models. A larger number of classification models can be tried, optimize the SMOTE paradigms, better cross-validation frameworks can be applied, and ensemble methods can be developed for the rheumatology framework. There is surely a potential of collecting multicenter RA and non-RA datasets to mix-match the training model designs and even try the AI under unseen paradigms in big data framework [65]. Lastly, the machine learning models can be extended to deep learning frameworks [67, 68].

Conclusions

The proposed study is the first pilot study that used an ML-based system for CVD detection that used data at two different time points to establish associations between characteristics and current cardiovascular events or 3-year follow-up cardiovascular events. Three kinds of classifiers—Random Forest, Support Vector Machine, and Linear Discriminant Analysis were designed in SMOTE framework and compared against conventional cardiovascular disease CVD risk prediction calculators such as FRS, SCORE, and ASCVD. Our system demonstrated through-and-through the validation of the 4-hypothesis labelled as cluster-effect, SMOTE-effect, ML-effect, and follow-up effect. While using 46 covariates in 542 patients, the CVD-CR system showed the mean accuracy and AUC as 98.40% and 0.98 ($p < 0.0001$), while the CVD-3YFU showed the mean accuracy and ACU of 98.39% and 0.98 ($p < 0.0001$). The ML system was effective, clinically reliable, and reasonably fast (less than one second) during online prediction.

Appendix A

See Tables 15, 16, 17, 18, 19.

Table 15 Ordered list of covariates (selected feature) using PCA pooling

CVD-CR		CVD-3YFU	
Covariates	New SN (Old SN)	Covariates	New SN (Old SN)
LCB PLQ	1 (38)	K	1 (24)
Height	2 (3)	Tg	2 (18)
RICA PLQ	3 (42)	CPK	3 (30)
CRP	4 (32)	RCCA PLQ	4 (40)
K	5 (24)	LCB PLQ	5 (38)
CPK	6 (30)	WBC	6 (21)
RCCA PLQ	7 (40)	ESR	7 (23)
Tg	8 (18)	HDL	8 (17)
WBC	9 (21)	Height	9 (3)
Gluc	10 (13)	RICA PLQ	10 (42)
LICA PLQ	11 (39)	RCB PLQ	11 (41)
Na	12 (25)	Na	12 (25)
Urea	13 (27)	LCCA PLQ	13 (37)
ESR	14 (23)	BMI	14 (4)
PL	15 (22)	RCFA PLQ	15 (44)
Diabetes T2D	16 (11)	LICA PLQ	16 (39)
HR	17 (7)	Crp	17 (32)
RCB PLQ	18 (41)	Age	18 (1)
HDL	19 (17)	Sex	19 (0)
LCCA PLQ	20 (37)	Hypertension	20 (10)
Hg	21 (20)	HR	21 (7)
RCFA PLQ	22 (43)	PLT	22 (22)
Create	23 (26)	Hg	23 (20)
BMI	24 (4)	Gluc	24 (13)
Age	25 (1)	HCT	25 (19)
AvDBP	26 (6)	Create	26 (26)
HCT	27 (19)	Urea	27 (27)
TSH	28 (33)	Diabetes T2D	28 (11)
Hypertension	29 (10)	LCFA PLQ	29 (43)
UA	30 (31)	Alt	30 (29)
AvSBP	31 (5)	AvDBP	31 (6)
LCFA PLQ	32 (43)	UA	32 (31)
Alt	33 (29)	TSH	33 (33)
Current Smoker	34 (9)	AST	34 (28)
Sex	35 (0)	² PLQ	35 (34)
Hyperlipidemia	36 (12)	¹ FH CAD	36 (8)
AST	37 (28)	Weight	37 (2)
² PLQ	38 (34)	AvSBP	38 (5)
¹ FH CAD	39 (8)	Current Smoker	39 (9)
LDL	40 (16)	HbA1c	40 (14)
HbA1c	41 (14)	Ldl	41 (16)
⁴ C-PLQ	42 (35)	Hyperlipidemia	42 (12)
Weight	43 (2)	⁴ C-PLQ	43 (35)
Chol	44 (15)	Chol	44 (15)
⁵ F-PLQ	45 (36)	⁵ F-PLQ	45 (36)
³ TNPLQ	46 (45)	³ TNPLQ	46 (45)

Table 15 (continued)

SN: Serial Number

¹FH CAD: family history of coronary artery disease²PLQ: plaques (Y/N)³TNPLQ: total number of plaques⁴C-PLQ: carotid plaque⁵F-PLQ: femoral plaque

AvSBP average systolic blood pressure; AvDBP average diastolic blood pressure; HR heart rate; LCCA PLQ left common carotid artery plaque; LCB PLQ left carotid bulb; LICA PLQ left internal carotid artery; RCCA PLQ right common carotid artery plaque; RCB PLQ right carotid bulb; RICA PLQ right internal carotid artery; Gluc glucose; Chol cholesterol; LDL low-density lipoprotein; HDL high-density lipoprotein; HbA1c hemoglobin A1c; PLT platelet count test; ESR erythrocyte sedimentation rate; K potassium; Na Sodium; AST aspartate aminotransferase; ALT alanine aminotransferase; CRP C-reactive protein; TSH thyroid-stimulating hormone; UA uric acid

Table 16 Baseline characteristics for the rheumatology data for CVD-CR and CVD-3YFU

SN	Parameter	Overall (542)	CVD-CR classes (No CVD and CVD)			CVD-3YFU classes (No CVD and CVD)		
			No CVD (535)	CVD (7)	<i>p</i> value	No CVD (534)	CVD (8)	<i>p</i> value
Cluster 1: Baseline parameters (OBBM)								
R1	Sex (male)	311(57.38%)	307(98.71)	4(1.29)	0.831	306(98.39)	5(1.61)	0.934
R2	Age (years)	53.07 ± 13.5	52.99 ± 13.5	58.86 ± 11.3	0.253	53.00 ± 13.5	57.75 ± 11.0	0.323
R3	Weight (kg)	77.27 ± 17.0	77.26 ± 17.1	78.04 ± 14.4	0.904	77.24 ± 17.1	79.47 ± 14.0	0.713
R4	Height (meter)	1.65 ± 0.1	1.65 ± 0.1	1.63 ± 0.1	0.568	1.65 ± 0.1	1.63 ± 0.1	0.538
R5	BMI (kg/m ²)	28.17 ± 5.3	28.16 ± 5.3	29.14 ± 4.3	0.626	28.15 ± 5.3	29.71 ± 4.3	0.406
R6	AvSBP (mmHg)	130.61 ± 19.1	130.60 ± 19.2	130.86 ± 14.4	0.972	130.60 ± 19.2	130.62 ± 13.5	0.998
R7	AvDBP (mmHg)	78.44 ± 11.0	78.44 ± 11.1	78.86 ± 8.1	0.920	78.43 ± 11.1	79.25 ± 7.7	0.835
R8	HR (bpm)	68.31 ± 10.5	68.37 ± 10.4	64.14 ± 11.7	0.290	68.38 ± 10.5	63.75 ± 11.0	0.215
R9	¹ FH CAD (%)	72(13.3%)	71(98.6%)	1(1.4%)	0.630	70(97.2%)	2(2.8%)	0.646
R10	Current Smoker †	0.76 ± 0.8	0.75 ± 0.8	1.57 ± 0.7	0.006	0.75 ± 0.8	1.50 ± 0.7	0.007
R11	Hypertension (mmHg)	203(37.5%)	198(97.5%)	5(2.5%)	0.140	197(97.0%)	6(3.0%)	0.065
R12	Diabetes T2D (mg/dl)	0.36 ± 0.7	0.36 ± 0.7	0.29 ± 0.5	0.785	0.36 ± 0.7	0.25 ± 0.4	0.663
R13	Hyperlipidemia (%)	147(27.1%)	143(97.3%)	4(2.7%)	0.171	142(96.6%)	5(3.4%)	0.062
Cluster 2: OBBM+LBBM								
R14	Gluc (mmol/l)	102.61 ± 31.5	102.55 ± 31.6	107.14 ± 28.2	0.703	102.58 ± 31.6	104.75 ± 27.2	0.847
R15	HbA1c (mg/dl)	5.94 ± 0.9	5.93 ± 0.9	6.23 ± 0.7	0.372	5.93 ± 0.9	6.14 ± 0.7	0.511
R16	Chol † (mg/dl)	202.93 ± 37.9	203.36 ± 37.6	170.57 ± 46.7	0.023	203.44 ± 37.6	168.88 ± 43.9	0.010
R17	LDL † (mg/dl)	125.02 ± 32.1	125.37 ± 31.9	98.14 ± 34.9	0.026	125.42 ± 31.9	98.25 ± 32.7	0.017
R18	HDL (mg/dl)	57.23 ± 15.2	57.30 ± 15.2	52.29 ± 16.6	0.386	57.33 ± 15.1	50.75 ± 16.1	0.224
R19	Tg (mg/dl)	110.10 ± 57.2	110.17 ± 57.4	104.71 ± 39.9	0.802	110.21 ± 57.4	102.62 ± 37.7	0.710
R20	Hct (mg/dl)	41.27 ± 3.5	41.29 ± 3.5	39.84 ± 2.4	0.285	41.29 ± 3.6	39.92 ± 2.2	0.281
R21	Hg (mg/dl)	13.69 ± 1.7	13.70 ± 1.7	12.89 ± 1.0	0.220	13.70 ± 1.7	12.99 ± 1.0	0.252
R22	WBC (K/ul)	6836.75 ± 1766.9	6838.22 ± 1777.5	6724.29 ± 483.4	0.866	6830.06 ± 1769.1	7283.75 ± 1547.7	0.472
R23	PLT (mg/dl)	248.89 ± 68.0	248.69 ± 68.0	264.29 ± 69.4	0.548	248.82 ± 68.0	253.62 ± 70.8	0.843
R24	ESR (mm/hr)	15.46 ± 13.1	15.34 ± 12.6	24.71 ± 33.8	0.060	15.35 ± 12.6	23.00 ± 31.9	0.102
R25	K (mEq/l)	4.38 ± 0.3	4.38 ± 0.3	4.31 ± 0.2	0.595	4.38 ± 0.3	4.39 ± 0.2	0.939
R26	Na (mEq/l)	140.70 ± 2.0	140.70 ± 2.0	141.10 ± 1.7	0.592	140.69 ± 2.0	141.34 ± 1.7	0.360
R27	Create (mEq/l)	0.85 ± 0.2	0.85 ± 0.2	0.90 ± 0.2	0.536	0.85 ± 0.2	0.89 ± 0.2	0.631
R28	Urea (mg/dl)	33.36 ± 9.4	33.30 ± 9.4	37.93 ± 9.7	0.194	33.29 ± 9.4	37.44 ± 9.2	0.215
R29	AST (u/l)	21.25 ± 7.9	21.23 ± 7.9	22.61 ± 3.0	0.646	21.25 ± 7.9	21.29 ± 4.5	0.990

Table 16 (continued)

SN	Parameter	Overall (542)	CVD-CR classes (No CVD and CVD)			CVD-3YFU classes (No CVD and CVD)		
			No CVD (535)	CVD (7)	<i>p</i> value	No CVD (534)	CVD (8)	<i>p</i> value
R30	ALT (u/l)	24.51 ± 14.5	24.54 ± 14.6	22.71 ± 4.1	0.742	24.55 ± 14.6	21.75 ± 4.6	0.589
R31	CPK (mg/l)	95.62 ± 67.7	95.94 ± 68.1	71.07 ± 13.4	0.335	95.96 ± 68.1	73.19 ± 13.8	0.346
R32	UA [†] (mg/l)	4.60 ± 1.2	4.61 ± 1.2	3.62 ± 0.3	0.033	4.61 ± 1.2	3.72 ± 0.4	0.040
R33	CRP [†] (mg/l)	4.27 ± 5.5	4.20 ± 5.4	9.19 ± 11.4	0.018	4.19 ± 5.4	9.41 ± 10.7	0.008
R34	TSH (mIU/l)	1.62 ± 0.7	1.61 ± 0.7	1.78 ± 0.3	0.558	1.61 ± 0.7	2.27 ± 1.3	0.013
Cluster 3: OBBM+LBBM+CUSIP								
R35	² PLQ (mm ²)	260(48.0%)	255(98.1%)	5(1.9%)	0.384	254(97.7%)	6(2.3%)	0.236
R36	⁴ C-PLQ (mm ²)	187(34.5%)	183(97.9%)	4(2.1%)	0.385	182(97.3%)	5(2.7%)	0.192
R37	⁵ F-PLQ (mm ²)	190(35.1%)	187(98.4%)	3(1.6%)	0.971	186(97.9%)	4(2.1%)	0.604
R38	LCCA PLQ (mm ²)	15(2.8%)	15(100.0%)	0(0.0%)	0.478	15(100.0%)	0(0.0%)	0.545
R39	LCB PLQ (mm ²)	119(22.0%)	117(98.3%)	2(1.7%)	0.973	117(98.3%)	2(1.7%)	0.825
R40	LICA PLQ (mm ²)	39(7.2%)	39(100.0%)	0(0.0%)	0.996	38(97.4%)	1(2.6%)	0.917
R41	RCCA PLQ (mm ²)	11(2.0%)	11(100.0%)	0(0.0%)	0.334	11(100.0%)	0(0.0%)	0.394
R42	RCB PLQ (mm ²)	132(24.4%)	129(97.7%)	3(2.3%)	0.481	128(97.0%)	4(3.0%)	0.198
R43	RICA PLQ (mm ²)	25(4.6%)	25(100.0%)	0(0.0%)	0.748	25(100.0%)	0(0.0%)	0.824
R44	LCFA PLQ (mm ²)	139(25.6%)	137(98.6%)	2(1.4%)	0.797	136(97.8%)	3(2.2%)	0.715
R45	RCFA PLQ (mm ²)	164(30.3%)	162(98.8%)	2(1.2%)	0.752	161(98.2%)	3(1.8%)	0.951
R46	³ TNPLQ (mm ²)	1.19 ± 1.6	1.19 ± 1.6	1.29 ± 0.9	0.867	1.18 ± 1.6	1.62 ± 1.2	0.429

[†] Significant cutoff of risk covariate is *p* < 0.05

¹ FH CAD family history of coronary artery disease

² PLQ plaques (Y/N)

³ TNPLQ total number of plaques

⁴ C-PLQ: carotid plaque

⁵ F-PLQ: femoral plaque

AvSBP: average systolic blood pressure; AvDBP: average diastolic blood pressure; HR: heart rate; LCCA PLQ: left common carotid artery plaque; LCB PLQ: left carotid bulb; LICA PLQ: left internal carotid artery; RCCA PLQ: right common carotid artery plaque; RCB PLQ: right carotid bulb; HbA1c hemoglobin A1C; RICA PLQ: right internal carotid artery; Gluc glucose; Chol: cholesterol; LDL low-density lipoprotein; HDL high-density lipoprotein; Tg triglyceride; Hct hematocrit; HbA1c hemoglobin A1c; PLT platelet count test; ESR erythrocyte sedimentation rate; K potassium; Na Sodium; AST aspartate aminotransferase; ALT alanine aminotransferase; CRP C-reactive protein; UA uric acid

Table 17 Accuracy comparison MI (with top 50% MI scores) vs. PCA pooling

Classifiers	CVD-CR			CVD-3YFU		
	No FS	MI	PCA pooling	No FS	MI	PCA pooling
RF	97.26%	98.53%	98.71%	98.53%	98.53%	98.52%
SVMrbf	95.59%	96.32%	98.71%	97.06%	98.53%	98.52%
LDA	95.59%	96.32%	96.53%	94.85%	94.96%	95.39%
Mean	96.15%	97.06%	97.98%	96.81%	97.34%	97.48%
SD	0.01	0.01	0.001	0.02	0.02	0.02

FS feature selection, RF random forest; SVMrbf support vector machine with radial basis function; LDA linear discriminant analysis; CVD-CR cardiovascular disease-current risk; CVD-3YFU cardiovascular disease risk-three-year follow-up

Table 18 Shapiro Test p values for 46 covariates in CVD-CR and CVD-3YFU frameworks

SN	Covariates	CVD-CR			CVD-3YFU		
		Overall	No CVD	CVD	Overall	No CVD	CVD
R1	Sex	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R2	Age	<0.0001	<0.0001	0.400	<0.0001	<0.0001	0.700
R3	Weight	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R4	Height	<0.0001	<0.0001	0.500	<0.0001	<0.0001	0.400
R5	BMI	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R6	AvSBP	<0.0001	<0.0001	0.200	<0.0001	<0.0001	0.200
R7	AvDBP	<0.0001	<0.0001	0.200	<0.0001	<0.0001	0.300
R8	HR	<0.0001	<0.0001	0.300	<0.0001	<0.0001	0.200
R9	¹ FH CAD	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R10	Current Smoker	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R11	Hypertension	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R12	Diabetes T2D	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R13	Hyperlipidemia	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R14	Gluc	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R15	HbA1c	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R16	Chol	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R17	LDL	<0.0001	<0.0001	0.300	<0.0001	<0.0001	0.200
R18	HDL	<0.0001	<0.0001	0.900	<0.0001	<0.0001	0.900
R19	Tg	<0.0001	<0.0001	0.900	<0.0001	<0.0001	0.900
R20	HCT	<0.0001	<0.0001	0.500	<0.0001	<0.0001	0.400
R21	Hg	<0.0001	<0.0001	0.500	<0.0001	<0.0001	0.900
R22	WBC	<0.0001	<0.0001	0.200	<0.0001	<0.0001	<0.0001
R23	PLT	<0.0001	<0.0001	0.100	<0.0001	<0.0001	0.100
R24	ESR	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R25	K	<0.0001	<0.0001	0.100	<0.0001	<0.0001	0.200
R26	Na	<0.0001	<0.0001	0.300	<0.0001	<0.0001	0.200
R27	Create	<0.0001	<0.0001	0.100	<0.0001	<0.0001	<0.0001
R28	Urea	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R29	AST	<0.0001	<0.0001	0.200	<0.0001	<0.0001	0.500
R30	ALT	<0.0001	<0.0001	0.800	<0.0001	<0.0001	0.500
R31	CPK	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R32	UA	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R33	CRP	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R34	TSH	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R35	² PLQ	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R36	⁴ C-PLQ	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R37	⁵ F-PLQ	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R38	LCCA PLQ	<0.0001	<0.0001	1.000	<0.0001	<0.0001	1.000
R39	LCB PLQ	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R40	LICA PLQ	<0.0001	<0.0001	1.000	<0.0001	<0.0001	<0.0001
R41	RCCA PLQ	<0.0001	<0.0001	1.000	<0.0001	<0.0001	1.000
R42	RCB PLQ	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R43	RICA PLQ	<0.0001	<0.0001	1.000	<0.0001	<0.0001	1.000
R44	LCFA PLQ	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R45	RCFA PLQ	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
R46	³ TNPLQ	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.200

¹FH CAD: family history of coronary artery disease²PLQ: plaques (Y/N)³TNPLQ: total number of plaques; ⁴C-PLQ carotid plaque; ⁵F-PLQ femoral plaque; AvSBP average systolic blood pressure; AvDBP average diastolic blood pressure; HR heart rate; LCCA PLQ left common carotid artery Plaque; LCB PLQ left carotid bulb; LICA PLQ left internal carotid artery; RCCA PLQ right common carotid artery plaque; RCB PLQ right carotid bulb; RICA PLQ right internal carotid artery; Gluc glucose; Chol cholesterol; LDL low-density lipoprotein; HDL high-density lipoprotein; HbA1c hemoglobin A1c; PLT platelet count test; ESR erythrocyte sedimentation rate; K potassium; Na Sodium; AST: aspartate aminotransferase; ALT alanine aminotransferase; CRP C-reactive protein; TSH thyroid-stimulating hormone; UA: uric acid

Table 19 Baseline characteristics for the validation coronary database

SN	Parameter	Overall	AngioScore classes (2 class)		p value
			No-Risk	High Risk	
R1	Age [†]	64.49 ± 10.5	62.76 ± 10.9	65.67 ± 10.1	0.002
R2	Caucasian	486(97.2%)	195(40.1%)	291(59.9%)	0.641
R3	¹ ADba	76.74 ± 13.7	76.84 ± 14.0	76.67 ± 13.5	0.897
R4	Pre-Diabetic	20(4.0%)	7(35.0%)	13(65.0%)	0.787
R5	Diabetes T1D	5(1.0%)	3(60.0%)	2(40.0%)	0.660
R6	Diabetes T2D [†]	114(22.8%)	33(28.9%)	81(71.1%)	0.006
R7	Diabetes (any) [†]	118(23.6%)	36(30.5%)	82(69.5%)	0.016
R8	Hypertension	338(67.6%)	126(37.3%)	212(62.7%)	0.050
R9	Hyperlipidemia [†]	288(57.6%)	98(34.0%)	190(66.0%)	0.001
R10	² ASba	135.35 ± 21.4	135.01 ± 20.9	135.59 ± 21.8	0.767
R11	Current Smoker [†]	100(20.0%)	29(29.0%)	71(71.0%)	0.013
R12	Casual Smoker	15(3.0%)	5(33.3%)	10(66.7%)	0.765
R13	Previous Smoker	218(43.6%)	87(39.9%)	131(60.1%)	0.916
R14	Smoking Hx [†]	330(66.0%)	120(36.4%)	210(63.6%)	0.014
R15	Sex [†]	349(69.8%)	119(34.1%)	230(65.9%)	< 0.0001
R16	GFR [†]	78.96 ± 18.2	81.02 ± 17.6	77.57 ± 18.4	0.037
R17	Creatinine [†]	83.99 ± 22.6	80.89 ± 20.9	86.10 ± 23.4	0.011
R18	Obesity	215(43.0%)	98(45.6%)	117(54.4%)	0.050
R19	BMI	30.41 ± 6.3	30.82 ± 6.6	30.14 ± 6.0	0.228
R20	³ FHPCVD	146(29.2%)	62(42.5%)	84(57.5%)	0.614
R21	⁴ FHD	195(39.0%)	76(39.0%)	119(61.0%)	0.670
R22	Drinks/wk	4.86 ± 10.4	4.64 ± 9.2	5.00 ± 11.1	0.706
R23	Angina	124(24.8%)	47(37.9%)	77(62.1%)	0.584
R24	Family Hx of CVD	321(64.2%)	123(38.3%)	198(61.7%)	0.240
R25	⁵ HCRI [†]	272(54.4%)	94(34.6%)	178(65.4%)	0.005
R26	⁶ OAA	9(1.8%)	4(44.4%)	5(55.6%)	0.926
R27	ACE Inhibitors	191(38.2%)	72(37.7%)	119(62.3%)	0.382
R28	ARBs Angiotensis	45(9.0%)	13(28.9%)	32(71.1%)	0.136
R29	Alpha-Blockers	30(6.0%)	10(33.3%)	20(66.7%)	0.534
R30	Beta-Blockers [†]	236(47.2%)	77(32.6%)	159(67.4%)	0.001
R31	⁷ CCB	93(18.6%)	37(39.8%)	56(60.2%)	0.987
R32	⁸ AP/AC [†]	368(73.6%)	131(35.6%)	237(64.4%)	< 0.0001
R33	Diuretics	99(19.8%)	41(41.4%)	58(58.6%)	0.908
R34	⁹ AANSAIDS	81(16.2%)	29(35.8%)	52(64.2%)	0.425
R35	Insulin	38(7.6%)	10(26.3%)	28(73.7%)	0.095
R36	¹⁰ N-IDM	72(14.4%)	22(30.6%)	50(69.4%)	0.087
R37	MPH	2.64 ± 1.4	2.59 ± 1.3	2.67 ± 1.4	0.570
R38	TPA	47.68 ± 44.3	45.08 ± 41.7	49.45 ± 45.9	0.280
R39	IPN	1.16 ± 0.8	1.11 ± 0.8	1.19 ± 0.8	0.240

[†] Significant cutoff of risk covariate is p < 0.05 (12 covariates); ¹ADba: Ave Dias before angio; ²ASba: Avg Sys before angio; ³FHPCVD: Family Hx of Premature CVD; ⁴FHD: Family Hx of Diabetes; ⁵HCRI: HMG-Co Reductase Inhibitors (Statins); ⁶OAA: Other Antilipemic Agents (not statins); ⁷CCB: Calcium Channel Blockers; ⁸AP/AC: Anti-Platelet/Anti-Coagulants; ⁹AANSAIDS: Anti-Anginals and NSAIDS; ¹⁰ N-IDM: Non-Insulin Diabetes Medications; MPH: maximum plaque height; TPA: total plaque area; IPN: intra-plaque neovascularization, NSAIDS: non-steroidal anti-inflammatory drugs; ARB: angiotensin receptor blocker; ACE: angiotensin-converting enzyme; BMI: body mass index, T1D: Type 1 diabetes, T2D: Type 2 diabetes

Appendix B

Three Classifiers

A. *Random forest*

RF is an ensemble ML-based learning algorithm that is comprised of various decision trees (DT). Each DT in the RF classifier receives a complete covariate set and performs the classification of the CVD risk. It follows a voting-based method, where the independent risk predictions by all DT are combined, and final risk prediction is performed using majority voting. Table 20 shows the set of tuning parameters used in the RF classifier.

B. *Support Vector Machine with radial the basis function*

Support Vector Machine (SVM) splits the input dataset samples into two classes that maximize the margin between two classes using optimal decision boundary. The classification framework is binary and was performed using a “one-vs-all” method. When the input dataset is non-linear (linearly inseparable), we use kernel functions such as RBF (kernel = “RBF”). When the input dataset is linearly separable, we use the linear kernel function (kernel = “linear”). We used a “RandomizedSearchCV” algorithm with RBF kernel for the optimization of tuning parameters. The three optimized tuning parameters used for the SVM classifier with the “RBF” kernel are provided in Table 21.

C. *Linear discriminant analysis*

Linear discriminant analysis (LDA) is a dimensionality reduction method and is a generalization of Fisher LDA. This is an ML-based paradigm and is used to find the linear combination of covariates that classify into binary class or multiclass, where each class has equal covariance matrices. The set of optimized tuning parameters required for the LDA classifier are provided in Table 22.

Appendix C

Performance evaluation metrics

The performance of all three ML-based classifiers (LDA, SVMrbf, and RF) and all three CCVRC calculators (FRS, SCORE, and ASCVD) [23, 47] was assessed using various metrics such as sensitivity, specificity, false-positive rate (FPR), false-negative rate (FNR), positive predictive value (PPV), negative predictive value (NPV), accuracy and area-under-the-curve (AUC). The AUC score for all the matrices

Table 20 Optimized tuning parameters of the RF algorithm

SN	Tuning parameter	Optimized value
1	Total number of trees	617
2	Maximum depth	142
2	Minimum sample split	7
3	Minimum sample leaf	4
4	Criterion	Gini

Table 21 Three optimized tuning parameters of the SVM algorithm

SN	Tuning parameters	opti- mized value
1	Kernel	Radial basis func- tion (RBF)
2	C	2
3	Gamma	0.1

Table 22 Three optimized tuning parameters of the LDA algorithm

SN	Tuning Parameters	Optimized Value
1	Solver	Singular value decomposition
2	Tol	1.0e-4
3	Store Covariance	False

was computed using a 2X2 confusion matrix that contains false-positive (FP), false negative (FN), true-positive (TP), and true-negative (TN). The 2X2 confusion matrix compares the two categorical covariates— ground-truth labels and the predicted risk labels from the ML-based classifiers. The true-positive (TP) shows the total number of times the classifier has correctly recognized the positive class and the true-negative (TN) shows the total number of times the classifier has correctly recognized the negative class. The false-positive defines the total number of times the classifier has correctly allocated a positive label to the patient with respect to the actual negative ground-truth label. Similarly, the false-negative defines the total number of times the classifier has correctly allocated a negative label to the patient with respect to the actual positive ground-truth label. Since we have binary-class risk labels, we first compute confusion

Table 23 Mathematical expressions for performance evaluation metrics

SN	Performance measure	Mathematical expression
1	Sensitivity, %	$\text{Sensitivity} = \left[\frac{\text{TP}}{\text{TP}+\text{FN}} \right] \times 100$
2	Specificity, %	$\text{Specificity} = \left[\frac{\text{TN}}{\text{TN}+\text{FP}} \right] \times 100$
3	False Predictive Rate (FPR), %	$\text{FPR} = \left[\frac{\text{FP}}{\text{FP}+\text{TN}} \right] \times 100$
4	False Negative Rate (FNR), %	$\text{FNR} = \left[\frac{\text{FN}}{\text{FN}+\text{TP}} \right] \times 100$
5	Positive Predictive Value (PPV), %	$\text{PPV} = \left[\frac{\text{TP}}{\text{TP}+\text{FP}} \right] \times 100$
6	Negative Predictive Value (NPV), %	$\text{NPV} = \left[\frac{\text{TN}}{\text{TN}+\text{FN}} \right] \times 100$
7	Mean Accuracy (ACC), %	$\text{ACC} = \left[\frac{\text{TP}+\text{TN}}{\text{TP}+\text{FP}+\text{TN}+\text{FN}} \right] \times 100$

matrices and the corresponding four sets of TP, TN, FP and FN, then we use these four sets to find the performance matrices involving sensitivity, specificity, FPR, FNP, PPV, NPV, and accuracy (Table 23).

In this study, we also calculate the receiver operating characteristics (ROC) curve. It is a plot between the true-positive rate and the false-positive rate. In other word, it is a plot between sensitivity and 100-specificity of the probability predictions of the ML classifiers at various cut-off points. Every point on the ROC curve indicates the sensitivity/specificity pair at the cut-off point. The upper left corner of the plot shows 100% sensitivity and specificity.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00296-021-05062-4>.

References














- Cross M et al (2014) The global burden of rheumatoid arthritis: estimates from the global burden of disease 2010 study. *Ann Rheum Dis* 73(7):1316–1322
- van Vollenhoven RF (2009) Sex differences in rheumatoid arthritis: more than meets the eye. *BMC Med* 7(1):12
- Crowson CS et al (2013) Rheumatoid arthritis and cardiovascular disease. *Am Heart J* 166(4):622–628
- Domingues VDS et al (2021) Increased short-term risk of cardiovascular events in inflammatory rheumatic diseases: results from a population-based cohort. *Rheumatol Int* 41(2):311–318
- Arts E, et al. (2014) Performance of four current risk algorithms in predicting cardiovascular events in patients with early rheumatoid arthritis. *Ann Rheum Dis* 2014: annrheumdis-2013–204024
- Crowson CS et al (2012) Usefulness of risk scores to estimate the risk of cardiovascular disease in patients with rheumatoid arthritis. *Am J Cardiol* 110(3):420–424
- Arts E, et al. (2015) Prediction of cardiovascular risk in rheumatoid arthritis: performance of original and adapted SCORE algorithms. *Ann Rheum Dis* 2015: annrheumdis-2014–206879
- Escalante A, Haas RW, Del Rincón I (2005) Paradoxical effect of body mass index on survival in rheumatoid arthritis: role of comorbidity and systemic inflammation. *Arch Intern Med* 165(14):1624–1629
- Urruela MA, Suarez-Almazor ME (2012) Lipid paradox in rheumatoid arthritis: changes with rheumatoid arthritis therapies. *Curr Rheumatol Rep* 14(5):428–437
- Gabriel SE (2010) Heart disease and rheumatoid arthritis: understanding the risks. *Ann Rheum Dis* 69(1):61–64
- del Rincón ID et al (2001) High incidence of cardiovascular events in a rheumatoid arthritis cohort not explained by traditional cardiac risk factors. *Arthritis Rheum* 44(12):2737–2745
- Jamthikar AD, et al. (2021) Multiclass machine learning vs. conventional calculators for stroke/CVD risk assessment using carotid plaque predictors with coronary angiography scores as gold standard: a 500 participants study. *Int J Cardiovasc Imag* 37(4): 1171–1187
- Araki T et al (2016) A new method for IVUS-based coronary artery disease risk stratification: a link between coronary and carotid ultrasound plaque burdens. *Comput Methods Programs Biomed* 124:161–179
- Biswas M et al (2021) A review on joint carotid intima-media thickness and plaque area measurement in ultrasound for cardiovascular/stroke risk monitoring: artificial intelligence framework. *J Digit Imaging* 34(3):581–604
- Cuadrado-Godia E et al (2018) Morphologic TPA (mTPA) and composite risk score for moderate carotid atherosclerotic plaque is strongly associated with HbA1c in diabetes cohort. *Comput Biol Med* 101:128–145
- Jamthikar A et al (2019) A low-cost machine learning-based cardiovascular/stroke risk assessment system: integration of conventional factors with image phenotypes. *Cardiovasc Diagn Ther* 9(5):420–430
- Jamthikar AD et al (2020) Cardiovascular risk assessment in patients with rheumatoid arthritis using carotid ultrasound B-mode imaging. *Rheumatol Int* 40(12):1921–1939
- Khanna NN et al (2019) Effect of carotid image-based phenotypes on cardiovascular risk calculator: AECRS10. *Med Biol Eng Comput* 57(7):1553–1566
- Puvvula A et al (2020) Morphological carotid plaque area is associated with glomerular filtration rate: a study of south asian indian patients with diabetes and chronic kidney disease. *Angiology* 71(6):520–535
- Viswanathan V et al (2020) Does the carotid bulb offer a better 10-Year CVD/stroke risk assessment compared to the common carotid artery? A 1516 ultrasound scan study. *Angiology* 71(10):920–933
- del Rincón I et al (2005) Relative contribution of cardiovascular risk factors and rheumatoid arthritis clinical manifestations to atherosclerosis. *Arthritis Rheum* 52(11):3413–3423
- del Rincón I et al (2015) Systemic inflammation and cardiovascular risk factors predict rapid progression of atherosclerosis in rheumatoid arthritis. *Ann Rheum Dis* 74(6):1118–1123
- Jamthikar A et al (2020) Ultrasound-based stroke/cardiovascular risk stratification using framingham risk score and ASCVD Risk Score based on “integrated vascular age” instead of “chronological age”: a multi-ethnic study of Asian Indian, Caucasian, and Japanese cohorts. *Cardiovasc Diagn Ther* 10(4):939–954
- Saba L et al (2019) Global perspective on carotid intima-media thickness and plaque: should the current measurement guidelines be revisited? *Int Angiol* 38(6):451–465
- Soriano-Valdez D et al (2021) The basics of data, big data, and machine learning in clinical practice. *Clin Rheumatol* 40:11–23
- Porcu M et al (2020) Carotid plaque imaging profiling in subjects with risk factors (diabetes and hypertension). *Cardiovasc Diagn Ther* 10(4):1005

27. Ikeda N et al (2015) Improved correlation between carotid and coronary atherosclerosis SYNTAX score using automated ultrasound carotid bulb plaque IMT measurement. *Ultrasound Med Biol* 41(5):1247–1262
28. Mach F et al (2019) 2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk: the task force for the management of dyslipidaemias of the european society of cardiology (ESC) and european atherosclerosis society (EAS). *Eur Heart J* 41(1):111–188
29. Knuuti J et al (2019) 2019 ESC Guidelines for the diagnosis and management of chronic coronary syndromes: the task force for the diagnosis and management of chronic coronary syndromes of the european society of cardiology (ESC). *Eur Heart J* 41(3):407–477
30. Johri AM et al (2016) Focused vascular ultrasound for the assessment of atherosclerosis: a proof-of-concept study. *J Am Soc Echocardiogr* 29(9):842–849
31. Johri AM et al (2013) Can carotid bulb plaque assessment rule out significant coronary artery disease? A comparison of plaque quantification by two- and three-dimensional ultrasound. *J Am Soc Echocardiogr* 26(1):86–95
32. Stein JH et al (2008) Use of carotid ultrasound to identify sub-clinical vascular disease and evaluate cardiovascular disease risk: a consensus statement from the American society of echocardiography carotid intima-media thickness task force. Endorsed by the society for vascular medicine. *J Am Soc Echocardiogr* 21(2):93–111
33. Molinari F et al (2012) Automated carotid IMT measurement and its validation in low contrast ultrasound database of 885 patient Indian population epidemiological study: results of AtheroEdge™ Software. *Int Angiol* 31(1):42–53
34. Molinari F et al (2012) Completely automated multiresolution edge snapper—a new technique for an accurate carotid ultrasound IMT measurement: clinical validation and benchmarking on a multi-institutional database. *IEEE Trans Image Process* 21(3):1211–1222
35. Molinari F, Zeng G, Suri JS (2010) Intima-media thickness: setting a standard for a completely automated method of ultrasound measurement. *IEEE Trans Ultrason Ferroelectr Freq Control* 57(5):1112–1124
36. Saba L et al (2012) Intima media thickness variability (IMTV) and its association with cerebrovascular events: a novel marker of carotid atherosclerosis? *Cardiovasc Diagn Ther* 2(1):10–18
37. Hao M, Wang Y, Bryant SH (2014) An efficient algorithm coupled with synthetic minority over-sampling technique to classify imbalanced PubChem BioAssay data. *Anal Chim Acta* 806:117–127
38. Jamthikar AD et al (2020) Cardiovascular risk assessment in patients with rheumatoid arthritis using carotid ultrasound B-mode imaging. *Rheumatol Internat* 2020:1–19
39. Jamthikar A et al (2019) A special report on changing trends in preventive stroke/cardiovascular risk assessment Via B-mode ultrasonography. *Curr Atheroscler Rep* 21(7):25
40. Saba L et al (2017) Plaque tissue morphology-based stroke risk stratification using carotid ultrasound: a polling-based PCA learning paradigm. *J Med Syst* 41(6):98
41. Jónsson H, G Cherubini G, Eleftheriou E (2020) Convergence behavior of DNNs with mutual-information-based regularization. *Entropy (Basel)* 22(7).
42. Jamthikar A et al (2020) Cardiovascular/stroke risk prevention: a new machine learning framework integrating carotid ultrasound image-based phenotypes and its harmonics with conventional risk factors. *Indian Heart J* 72(4):258–264
43. Acharya UR et al (2013) Understanding symptomatology of atherosclerotic plaque by image-based tissue characterization. *Comput Methods Programs Biomed* 110(1):66–75
44. Araki T et al (2016) PCA-based polling strategy in machine learning framework for coronary artery disease risk assessment in intravascular ultrasound: a link between carotid and coronary grayscale plaque morphology. *Comput Methods Programs Biomed* 128:137–158
45. Maniruzzaman M et al (2019) Statistical characterization and classification of colon microarray gene expression data using multiple machine learning paradigms. *Comput Methods Programs Biomed* 176:173–193
46. Maniruzzaman M et al (2017) Comparative approaches for classification of diabetes mellitus data: machine learning paradigm. *Comput Methods Programs Biomed* 152:23–34
47. Jamthikar A et al (2020) Cardiovascular/stroke risk predictive calculators: a comparison between statistical and machine learning models. *Cardiovasc Diagn Ther* 10(4):919–938
48. Fariniuk LF et al (2017) Efficacy of protaper instruments during endodontic retreatment. *Indian J Dent Res* 28(4):400–405
49. Fernandes Filho J et al (2017) Evaluation and comparison of five skinfold calipers. *Nutr Hosp* 34(1):111–115
50. Banchor SK, et al. (2017) Relationship between automated coronary calcium volumes and a set of manual coronary lumen volume, vessel volume and atheroma volume in japanese diabetic cohort. *J Clin Diagn Res* 11(6): Tc9-Tc14
51. Saba L et al (2016) Inter-observer variability analysis of automatic lung delineation in normal and disease patients. *J Med Syst* 40(6):142
52. Johri AM, et al. (2021) Role of artificial intelligence in cardiovascular risk prediction and outcomes: comparison of machine-learning and conventional statistical approaches for the analysis of carotid ultrasound features and intra-plaque neovascularization. *Int J Cardiovasc Imaging*.
53. Abraham A et al (2014) Machine learning for neuroimaging with scikit-learn. *Front Neuroinform* 8:14
54. Cruz GV et al (2021) Machine learning reveals the most important psychological and social variables predicting the differential diagnosis of rheumatic and musculoskeletal diseases. *Rheumatol Internat* 2021:1–10
55. Kataria S, Ravindran V (2018) Digital health: a new dimension in rheumatology patient care. *Rheumatol Int* 38(11):1949–1957
56. Schmajuk G, Yazdany J (2017) Leveraging the electronic health record to improve quality and safety in rheumatology. *Rheumatol Int* 37(10):1603–1610
57. Khanna NN et al (2019) Rheumatoid arthritis: atherosclerosis imaging and cardiovascular risk assessment using machine and deep learning-based tissue characterization. *Curr Atheroscler Rep* 21(2):7
58. Jamthikar A et al (2019) A low-cost machine learning-based cardiovascular/stroke risk assessment system: integration of conventional factors with image phenotypes. *Cardiovasc Diagn Therapy* 9(5):420
59. Jamthikar A, et al. (2020) Cardiovascular/stroke risk prevention: a new machine learning framework integrating carotid ultrasound image-based phenotypes and its harmonics with conventional risk factors. *Indian Heart J*
60. Jamthikar A et al (2020) Cardiovascular/stroke risk predictive calculators: a comparison between statistical and machine learning models. *Cardiovasc Diagn Therapy* 10(4):919–938
61. Gastouniotti A et al (2015) A novel computerized tool to stratify risk in carotid atherosclerosis using kinematic features of the arterial wall. *IEEE J Biomed Health Inform* 19(3):1137–1145
62. Unnikrishnan P, et al. (2016) Development of health parameter model for risk prediction of CVD using SVM. *Comput Mathemat Methods Med* 2016
63. Weng SF et al (2017) Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS ONE* 12(4):e0174944

64. Ambale-Venkatesh B, et al. (2017) *Cardiovascular event prediction by machine learning: the Multi-Ethnic Study of Atherosclerosis*. *Circulation Res*, Circresaha. 117:311312
65. Araki T et al (2017) Stroke risk stratification and its validation using ultrasonic echolucent carotid wall plaque morphology: a machine learning paradigm. *Comput Biol Med* 80:77–96
66. Nakamura M et al (2013) LVQ-SMOTE - learning vector quantization based synthetic minority over-sampling technique for biomedical data. *BioData Min* 6(1):16
67. Saba L et al (2021) Ultrasound-based internal carotid artery plaque characterization using deep learning paradigm on a super-computer: a cardiovascular disease/stroke risk assessment system. *Int J Cardiovasc Imaging* 37(5):1511–1528
68. Skandha SS et al (2020) 3-D optimized classification and characterization artificial intelligence paradigm for cardiovascular/stroke risk stratification using carotid ultrasound-based delineated plaque: Atheromatic™ 20. *Comput Biol Med* 125:103958
69. Banchhor SK et al (2017) Wall-based measurement features provides an improved IVUS coronary artery risk assessment when fused with plaque texture-based features during machine learning paradigm. *Comput Biol Med* 91:198–212
70. Kakadiaris IA et al (2018) Machine learning outperforms ACC/AHA CVD risk calculator in MESA. *J Am Heart Assoc* 7(22):e009476

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

George Konstantonis¹ · Krishna V. Singh²  · Petros P. Sfikakis¹  · Ankush D. Jamthikar^{3,4}  · George D. Kitas^{5,6}  · Suneet K. Gupta⁷  · Luca Saba⁸  · Kleio Verrou⁹ · Narendra N. Khanna¹⁰  · Zoltan Ruzsa¹¹  · Aditya M. Sharma¹²  · John R. Laird¹³  · Amer M. Johri¹⁴  · Manudeep Kalra¹⁵  · Athanasios Protogerou¹⁶  · Jasjit S. Suri¹⁷

¹ Rheumatology Unit, National Kapodistrian University of Athens, Athens, Greece

² Research Intern, AtheroPoint™, Roseville, CA, USA

³ Research Scientist, AtheroPoint™, USA, Roseville, CA, USA

⁴ Visvesvaraya National Institute of Technology, Nagpur, India

⁵ Academic Affairs, Dudley Group NHS Foundation Trust, Dudley, UK

⁶ Arthritis Research UK Epidemiology Unit, Manchester University, Manchester M13, UK

⁷ Department of Computer Science, Bennett University, Gr. Noida, India

⁸ Department of Radiology, University of Cagliari, Cagliari, Italy

⁹ Department of Medicine, National and Kapodistrian University of Athens, Athens, Greece

¹⁰ Department of Cardiology, Indraprastha Apollo Hospitals, New Delhi, India

¹¹ Department of Internal Medicines, Invasive Cardiology Division, University of Szeged, Szeged, Hungary

¹² Division of Cardiovascular Medicine, University of Virginia, Charlottesville, VA, USA

¹³ Heart and Vascular Institute, Adventist Health St. Helena, St Helena, CA, USA

¹⁴ Department of Medicine, Division of Cardiology, Queen's University, Kingston, ON, Canada

¹⁵ Department of Radiology, Massachusetts General Hospital, 55 Fruit Street, Boston, MA, USA

¹⁶ Cardiovascular Prevention Unit, Department of Pathophysiology, National Kapodistrian University of Athens, Athens, Greece

¹⁷ Stroke Monitoring and Diagnostic Division, AtheroPoint™, Roseville, CA 95661, USA