



A low-cost machine learning-based cardiovascular/stroke risk assessment system: integration of conventional factors with image phenotypes

Ankush Jamthikar¹, Deep Gupta¹, Narendra N. Khanna², Luca Saba³, Tadashi Araki⁴, Klaudija Viskovic⁵, Harman S. Suri⁶, Ajay Gupta⁷, Sophie Mavrogeni⁸, Monika Turk⁹, John R. Laird¹⁰, Gyan Pareek¹¹, Martin Miner¹², Petros P. Sfikakis¹³, Athanasios Protogerou¹⁴, George D. Kitas¹⁵, Vijay Viswanathan¹⁶, Andrew Nicolaides¹⁷, Deepak L. Bhatt¹⁸, Jasjit S. Suri¹⁹

¹Department of Electronics and Communication Engineering, Visvesvaraya National Institute of Technology, Nagpur, Maharashtra, India; ²Department of Cardiology, Indraprastha Apollo Hospitals, New Delhi, India; ³Department of Radiology, University of Cagliari, Cagliari, Italy; ⁴Division of Cardiovascular Medicine, Toho University, Tokyo, Japan; ⁵Department of Radiology and Ultrasound, University Hospital for Infectious Diseases Croatia, Zagreb, Croatia; ⁶Department of Neuroscience, Brown University, Providence, RI, USA; ⁷Department of Radiology, Weill Cornell Medicine, New York, NY, USA; ⁸Cardiology Clinic, Onassis Cardiac Surgery Center, Athens, Greece; ⁹Department of Neurology, University Medical Centre Maribor, Maribor, Slovenia; ¹⁰Heart and Vascular Institute, Adventist Health St. Helena, St. Helena, CA, USA; ¹¹Minimally Invasive Urology Institute, Brown University, Providence, RI, USA; ¹²Men's Health Center, Miriam Hospital Providence, Providence, RI, USA; ¹³Rheumatology Unit, ¹⁴Department of Cardiovascular Prevention & Research Unit Clinic & Laboratory of Pathophysiology, National and Kapodistrian University of Athens, Athens, Greece; ¹⁵R & D Academic Affairs, Dudley Group NHS Foundation Trust, Dudley, UK; ¹⁶M.V. Hospital for Diabetes and Professor M. Viswanathan Diabetes Research Centre, Chennai, India; ¹⁷Vascular Screening and Diagnostic Centre and University of Nicosia Medical School, Nicosia, Cyprus; ¹⁸Brigham and Women's Hospital Heart & Vascular Center, Harvard Medical School, Boston, MA, USA; ¹⁹Stroke Monitoring and Diagnostic Division, AtheroPoint™, Roseville, CA, USA

Contributions: (I) Conception and design: A Jamthikar, D Gupta, HS Suri, JS Suri, L Saba; (II) Administrative support: J Suri, D Gupta; (III) Provision of study materials or patients: T Araki; (IV) Collection and assembly of data: T Araki, JS Suri; (V) Data analysis and interpretation: A Jamthikar, JS Suri; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Jasjit S. Suri, MS, PhD, MBA, Fellow AIMBE. Stroke Monitoring and Diagnostic Division, AtheroPoint™, Roseville, CA 95661, USA. Email: jasjit.suri@atheropoint.com.

Background: Most cardiovascular (CV)/stroke risk calculators using the integration of carotid ultrasound image-based phenotypes (CUSIP) with conventional risk factors (CRF) have shown improved risk stratification compared with either method. However such approaches have not yet leveraged the potential of machine learning (ML). Most intelligent ML strategies use follow-ups for the endpoints but are costly and time-intensive. We introduce an integrated ML system using stenosis as an endpoint for training and determine whether such a system can lead to superior performance compared with the conventional ML system.

Methods: The ML-based algorithm consists of an offline and online system. The offline system extracts 47 features which comprised of 13 CRF and 34 CUSIP. Principal component analysis (PCA) was used to select the most significant features. These offline features were then trained using the event-equivalent gold standard (consisting of percentage stenosis) using a random forest (RF) classifier framework to generate training coefficients. The online system then transforms the PCA-based test features using offline trained coefficients to predict the risk labels on test subjects. The above ML system determines the area under the curve (AUC) using a 10-fold cross-validation paradigm. The above system so-called "AtheroRisk-Integrated" was compared against "AtheroRisk-Conventional", where only 13 CRF were considered in a feature set.

Results: Left and right common carotid arteries of 202 Japanese patients (Toho University, Japan) were retrospectively examined to obtain 395 ultrasound scans. AtheroRisk-Integrated system [AUC =0.80, P<0.0001, 95% confidence interval (CI): 0.77 to 0.84] showed an improvement of ~18% against AtheroRisk-Conventional ML (AUC =0.68, P<0.0001, 95% CI: 0.64 to 0.72).

Conclusions: ML-based integrated model with the event-equivalent gold standard as percentage stenosis is powerful and offers low cost and high performance CV/stroke risk assessment.

Keywords: Atherosclerosis; conventional risk factors (CRF); carotid ultrasound (CUS); carotid intima-media thickness (cIMT); carotid stenosis; cardiovascular disease (CVD); stroke; 10-year risk; machine learning (ML)

Submitted Jul 22, 2019. Accepted for publication Aug 22, 2019.

doi: 10.21037/cdt.2019.09.03

View this article at: <http://dx.doi.org/10.21037/cdt.2019.09.03>

Introduction

Annually, about 17.7 million people are affected by cardiovascular (CV) diseases including heart attack and stroke events (1). Atherosclerosis is the major contributor to such CV/stroke events (2). One way of predicting the occurrence of these events is by performing risk assessment using conventional risk factors (CRF) that are responsible for the growth of atherosclerosis (3). However, CRF alone does not explain the elevated risk of CV/stroke events (4). This is because of the morphological variations in the atherosclerotic plaque that cannot be captured using CRF alone but which can easily be assessed using imaging modalities (5,6). Thus, there is a need to look beyond the scope of CRF and search for preventive healthcare solutions that can provide an accurate routine risk assessment at an affordable cost.

Imaging plays a vital role in developing a comprehensive preventive strategy to combat stroke and heart disease. Ultrasound and in particular carotid ultrasound (CUS) screening can easily be adapted in routine clinical practice compared to its other non-invasive counterparts such as computed tomography and magnetic resonance imaging (4,5). Even in the low-resolution images, CUS can capture the image-based phenotypes that reflect morphological variations in atherosclerotic plaque (4). Furthermore, the carotid ultrasound image-based phenotypes (CUSIP) such as carotid intima-media thickness (cIMT), carotid plaque (CP), and carotid artery stenosis are considered the most significant biomarkers of CV/stroke events (7-10). Thus, integrating these CUSIP with CRF can convey a larger power of risk assessment (11-13). Such risk assessment systems are affordable but do not incorporate the full capabilities of deep learning algorithms which are now emerging as powerful and novel tools in healthcare.

Advancements in artificial intelligence (AI) are gaining popularity due to its ability to provide more accurate risk assessment in routine clinical practice (4). Specifically,

machine learning (ML) algorithms have shown promising results in atherosclerotic plaque characterization and CV/stroke risk stratification (4). Use of ML algorithms for accurate risk prediction requires the primary endpoints (CV events or cerebrovascular events) obtained from longitudinal trials which are expensive and require a large sample size (3,14,15). Furthermore, such kinds of trials are not feasible for all routine CV/stroke risk assessment. Thus, there is a need to find a low-cost alternative to the primary endpoints without compromising the accuracy of risk prediction (3,14,15).

Percentage stenosis is a clinically well-established biomarker which if left untreated may lead to CV/stroke events (10,16). Furthermore, percentage stenosis detected using CUS can assist physicians in making a judgment about stroke management practices such as endarterectomy or stenting (16-18). Thus, in this study, carotid artery stenosis was used as an event-equivalence gold standard (EEGS). The proposed low-cost EEGS can be used as an alternative to the primary endpoints for retrospective studies. The rationale for using stenosis as EEGS is discussed in the next section.

The objective and novelty of this retrospective study is to provide an accurate and low-cost ML-based system with stenosis as EEGS that can be employed for the routine CV/stroke risk assessment of patients (*Figure 1*). Another novelty is to investigate the effect of integrated risk factors (a combination of CUSIP and CRF) against CRF standalone for CV/stroke risk assessment.

Since the CUSIP have the potential to reflect variations in atherosclerotic plaque compared to CRF, we hypothesize that the proposed low-cost ML system developed using integrated risk factors (so-called "AtheroRisk-Integrated") with EEGS can perform better compared with the ML system developed using CRF alone (so-called "AtheroRisk-Conventional"). The proposed integrated ML-based risk stratification system is the first of its kind that evaluates

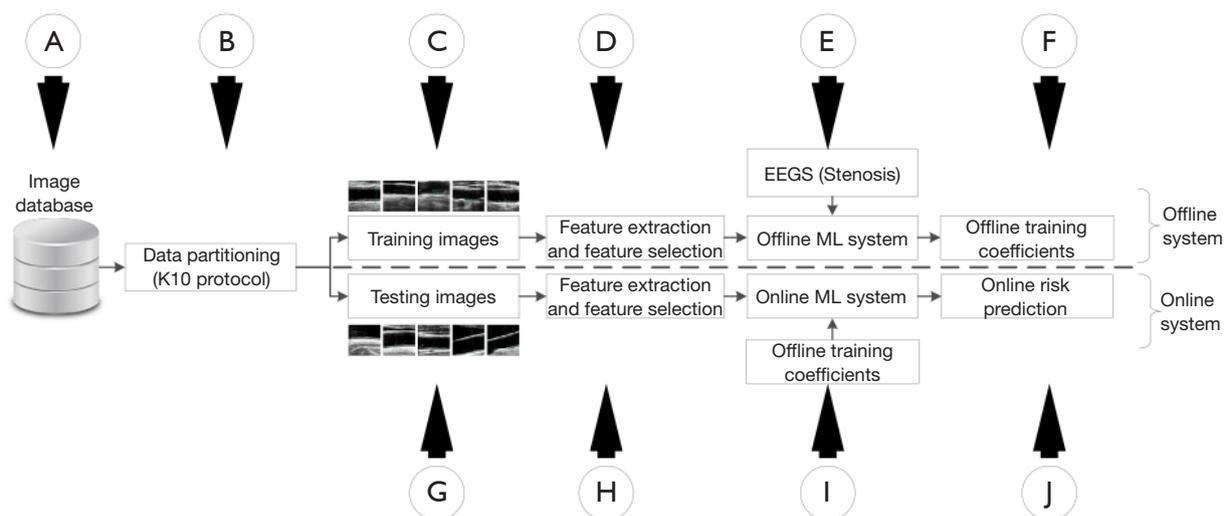


Figure 1 The generalized framework for risk stratification of patients using a ML system. EEGS, event-equivalence gold standard; ML, machine learning.

the risk of cardiovascular disease (CVD)/stroke using percentage stenosis as an EEGS.

EEGS

Primary gold standards or endpoints such as CV or cerebrovascular events require longitudinal trials which are expensive and time-consuming (3,14,15). Thus, it is important to look for surrogate markers that can mimic the characteristics of the primary endpoints at minimal cost (14,15). Such types of surrogate markers of CV/stroke events are also termed as EEGS. The EEGS needs to be evaluated using a small sample size, with lower-cost, and for short duration (14,15). According to Boissel *et al.* (19), surrogate markers should be (I) reproducible and convenient to compare with primary endpoints, (II) they must have a clear link with primary endpoints, and (III) they should provide clinically relevant benefits. Carotid artery stenosis is a well-established atherosclerosis-driven CV/stroke biomarker (10,16). It is generally accepted that the higher the luminal stenosis, the higher the risk of CV/stroke events (17,20). The annual risk of stroke increases to 1–2% in patients with asymptomatic yet significant (>50%) carotid artery stenosis (17,20). The risk of stroke events is moderate if the stenosis ranges between 30% and 69%, and more significant if the stenosis ranges between 70% and 99% (18). In clinical practice, knowing the accurate

percentage of carotid artery stenosis aids physicians to decide the management of stroke events either by using carotid endarterectomy or by using the appropriate medications (18). Thus, carotid stenosis is an important surrogate indicator of CV/stroke events and can be considered as EEGS.

Methods

Study population and image acquisition

A cohort of 202 patients was recruited for this retrospective study and an ultrasound examination was conducted between July 2009 and December 2010. The patients were approved by the institutional review board of Toho University, Japan and written consent was obtained from all the study participants. Both left and right carotid arteries were examined using B-mode ultrasonography scanner (Aplio XG, Xario, Aplio XV, Toshiba Inc., Tokyo, Japan) and a total of 404 CUS scans (202 patients × 2 CUS scans) were obtained. Nine CUS scans were excluded from this study due to its non-suitability. Thus, a total of 395 CUS scans were used to test the hypothesis of this study. Section A of the *Figure 1* indicates this source of 395 CUS scans. Mean pixel resolution was 0.0529 mm-per-pixel. Two operators (with 15 years of experience) analyzed the scans. We used the CUS image acquisition protocol presented before (11,12,21,22). The analytical approach and study

design for this manuscript are unique with respect to previously published reports using this cohort (11,12,21,22).

Data partitioning for ML algorithm

Section B of the *Figure 1* indicates the data partitioning used in this study. In general, any data partitioning protocol divides the input dataset into two parts: (I) training dataset (section C of the *Figure 1*) and (II) testing dataset (section G of the *Figure 1*). The proposed study uses a 10-fold data partitioning (or cross-validation) protocol (23-25), where the input image dataset set is divided into ten equal independent parts. The 10-fold cross-validation protocol is also termed as K10 protocol where 10 indicate the number of total partitions designed during ML-based training model (typically using 90% of the dataset for training). Out of 10 parts, at any time, nine parts were used for training the ML-based system while the remaining one part was used for validating the predictions of the system. Section B of the *Figure 1* indicates the data partitioning using K10 protocol. All our previous ML-based studies (23-25) reported a better performance while using K10 protocol compared to all the other data partitioning protocols, we have thus used K10 protocol in our proposed study.

Feature extraction: image-based phenotypes and CRF

A total of 47 risk factors were used to define the risk profile of the patients, of which, 34 were image-based phenotypes and 13 CRF, obtained from the combination of patients' demographics, blood biomarkers, along with a unity intercept term. Section D of *Figure 1* shows the feature extraction from both the CRF and CUSIP during the training phase of the ML system. Similarly, section H of *Figure 1* feature extraction from both the CRF and CUSIP during the testing phase of the ML system. These features were then fed into the feature selection module to optimize the ML paradigm.

CUSIP

The 34 CUSIP were derived using six steps: (I) Initially six current image-based phenotypes such as average cIMT (cIMTave), maximum cIMT (cIMTmax), minimum cIMT (cIMTmin), variability of cIMT (cIMTV), morphological CP area (also called as total plaque area or TPA), and normalized TPA (nTPA) were measured using an AtheroEdge (AtheroPoint, Roseville, USA) (22,26-28). (II) In the second step, six integrated 10-year image-based

phenotypes also called as CUSIP_{10yr} such as cIMTave_{10yr}, cIMTmax_{10yr}, cIMTmin_{10yr}, cIMTV_{10yr}, TPA_{10yr}, and nTPA_{10yr} were computed using the mathematical formulation provided in our previous study (12). The subscript "10 year" indicated the 10-year measurement. (III) CUS image-based AtheroEdge Composite Risk Scores were then computed for both (i) six current and (ii) six 10-year integrated phenotypes (13). Thus, in total 14 image-based phenotypes were measured which comprised of seven current CUSIP (CUSIP_{curr}) and seven 10-year CUSIP (CUSIP_{10yr}). (IV) The fourth step was to compute the 14 squared harmonics terms for the corresponding seven CUSIP_{curr} and seven CUSIP_{10yr}. (V) Five risk factors were derived from the age-adjusted grayscale median (AAGSM) which is a recently proposed biomarker that indicates the symptomatic and asymptomatic nature of the CP (29). (VI) Finally, a carotid plaque score (PS) (30) was also added as an image-based feature, thus totaling feature set of 34 risk factors. All the listed CUSIP indicated the wall variability and morphology of the atherosclerotic plaque.

CRF

In addition to CUSIP, 13 CRF were also obtained from patients' demographics and blood serum. The CRF consisted of age, sex, glycated hemoglobin (HbA1c), fasting blood sugar (FBS), low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), total cholesterol (TC), a ratio of TC and HDL-C*, hypertension (HT), smoking, family history (FH), and triglyceride (TG).

Feature selection

A feature selection technique called principal component analysis (PCA) was used to select the dominant features from (I) a pool of 13 risk factors for the design of conventional ML system (AtheroRisk-Conventional) and (II) 47 risk factors for the design of integrated ML system (AtheroRisk-Integrated). The polling-based strategy of PCA (Supplementary file 1) was used to decide the best cutoff point for PCA algorithm to maximize the risk stratification accuracy (31,32) by finding the best combination of CRF and CUSIP. Out of 13 CRF, the PCA-based polling strategy selected 9 dominant features (age, LDL-C, FH, HbA1c, HT, smoking, TG, FBS, and TC/HDL ratio) which were then applied to AtheroRisk-Conventional ML system. Similarly, out of 47 integrated features, the PCA-based polling strategy selected 18 dominant features

Table 1 Baseline characteristics of the patients divided into low-risk and high-risk classes.

SN	Parameters	Overall	High-risk [#]	Low-risk [#]	P value
R1	Total, n	202	12	190	–
R2	Male, n (%) [†]	156 (77.23)	7 (4.49)	149 (95.51)	0.11
R3	Age (years) [†]	68.97±10.96	76.33±9.27	68.50±10.91	0.02
R4	HbA1c (%)	6.28±1.11	6.38±0.84	6.27±1.13	0.73
R5	FBS (mg/dL)	121.21±34.81	137.58±53.15	120.18±33.26	0.09
R6	LDL-C (mg/dL)	100.75±31.48	99.75±36.91	100.81±31.22	0.91
R7	HDL-C (mg/dL)	50.49±14.97	45.92±18.09	50.77±14.76	0.28
R8	TC (mg/dL)	174.33±36.73	171.00±38.07	174.54±36.73	0.75
R9	TC/HDL	3.65±1.01	4.08±1.36	3.62±0.98	0.13
R10	HT, n (%) [†]	147 (72.77)	8 (5.44)	139 (94.56)	0.63
R11	SBP (mmHg) [†]	134.55±8.92	133.33±9.85	134.63±8.89	0.63
R12	DBP (mmHg) [†]	87.28±4.46	86.67±4.92	87.32±4.44	0.63
R13	Smoking, n (%)	81 (40.10)	4 (4.94)	77 (95.06)	0.62
R14	FH, n (%) [†]	24 (11.88)	0 (0.00)	24 (100.00)	0.19
R15	PS	9.09 (5.31)	14.75±6.47	8.74±5.04	<0.0001

[†], significant confounding factors; [#], stenosis was used for risk stratification. SN, Serial number; HbA1c, glycated hemoglobin; LDL-C, low-density lipoprotein cholesterol; HDL-C, high-density lipoprotein cholesterol; TC, total cholesterol; SBP, systolic blood pressure; DBP, diastolic blood pressure; FH, family history; PS, plaque score.

(HbA1c, PS, cIMT_{ave}, age, LDL-C, FBS, HT, TG, TC/HDL ratio, smoking, square harmonic of cIMT_{ave}, HDL-C, FH, TC, cIMT_{ave}_{10yr}, cIMTV_{10yr}, difference between average lumen diameter and cIMT_{max}, and total plaque pixels) which were then applied to AtheroRisk-Integrated ML system.

ML algorithm for risk stratification

The architecture of the ML system used in the current study is presented in *Figure 1*. The ML system is generally divided into two sub-types: (I) offline (training stage) and (II) online (testing stage) ML systems, respectively. Offline ML system extracts 47 novel features from the training CUS images, followed by dominant features selected using PCA algorithm (see *Figure 1*). Using these features and the EEGS, offline ML system learns to identify the risk profile of patient (i.e., high-risk or low-risk). This process is called training the ML system. Once the ML system is trained, the training coefficients can be used as input to transform the online test features yielding the ML-based risk. Such online ML systems can then be employed in routine clinical

practice without any requirement of EEGS. A 10-fold cross-validation protocol was used for data-partitioning and offline training of the ML system. Since the focus of this study was to develop a cost-effective, efficient ML system, a random forest (RF) classifier (Supplementary file 2) was incorporated in the ML system for risk stratification of patients (33). Two types of ML systems were designed (28) based on RF classifier: one with 13 CRF-called AtheroRisk-Conventional and a second with 47 integrated risk factors-called AtheroRisk-Integrated. The objective of this study is to compare the performance of AtheroRisk-Integrated against AtheroRisk-Conventional.

Statistical analysis

SPSS23.0 was used to perform statistical analysis. The baseline characteristics (*Table 1*) are presented as mean ± standard deviation (SD) for the continuous variables and as a percentage for the categorical variables. Independent sample *t*-test was used for the continuous variables and the chi-squared test was for the categorical variables. Carotid artery stenosis measured using the North American

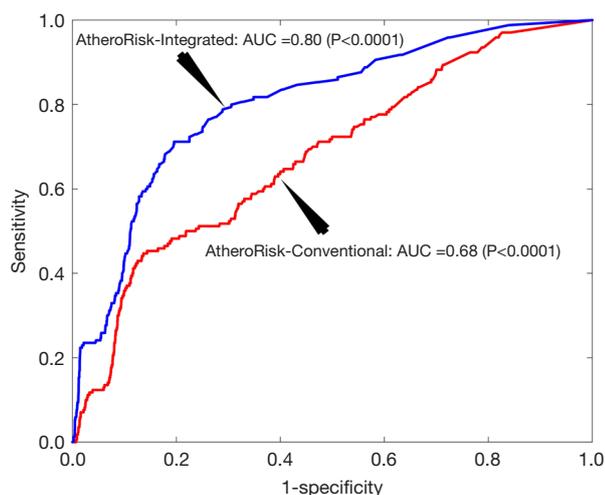


Figure 2 ROCs for AtheroRisk-Integrated ML system benchmarked against AtheroRisk-Conventional. ROC, receiver operating characteristic; ML, machine learning; AUC, area under the curve.

Symptomatic Carotid Endarterectomy Trial (NASCET) criteria (18) was used as EEGS for the training of the ML system, given the selected features using polling-based PCA. The validity of the recruited sample size was tested using power analysis. The resultant sample size with a 95% confidence interval (CI) and a 5% margin of error was 384 samples. Thus, the recruited sample size (395 scans) in our study was about 3% higher compared with the required sample size of 384.

Results

Baseline characteristics

Table 1 indicates the baseline characteristics of the Japanese cohort. The cohort of 202 patients (156 were males and 46 were females) was analyzed in this study. Similarly, 147 patients suffered from HT, and 49 patients were diabetic. The criterion for HT was systolic blood pressure (SBP) ≥ 130 mg/dL and diastolic blood pressure (DBP) ≥ 80 mg/dL or treatment with antihypertensive medications (34). The criterion for diabetes was HbA1c $\geq 6.5\%$ or the treatment with hypoglycemic agents. The average age of the patients was 68.97 ± 10.96 years (ranging between 29 and 88 years), HbA1c was $6.28\% \pm 1.11\%$ (ranging between 4.80% and 13%), FBS was 121.21 ± 34.81 mg/dL (ranging between 64 and 255 mg/dL), LDL-C was 100.75 ± 31.48 mg/dL (ranging between 24 and 193 mg/dL), HDL-C was 50.49 ± 14.97 mg/dL (ranging between 18 and 115 mg/dL), and

TC was 174.33 ± 36.73 mg/dL (ranging between 61 and 255 mg/dL). PS was the strong significant risk factor. In order to risk stratify the patients into low-risk or high-risk class, stenosis was used as an EEGS with a threshold value of 40%. The justification for the 40% threshold is presented in the “discussion” section.

Effect of integrated risk factors on the performance of ML system

The performance of AtheroRisk-Integrated was evaluated against the AtheroRisk-Conventional using the area under the curve (AUC). Since AUC reflects the trade-off between sensitivity and specificity values, it was used as a primary performance evaluation metric. Comparative results between AtheroRisk-Integrated and AtheroRisk-Conventional ML systems using sensitivity, specificity, and risk stratification accuracy are presented in the Supplementary file 3 of this manuscript. The AUC value for AtheroRisk-Integrated (AUC = 0.80, $P < 0.0001$, 95% CI: 0.77 to 0.84) was ~18% higher compared with the AUC value for AtheroRisk-Conventional system (AUC = 0.68, $P < 0.0001$, 95% CI: 0.64 to 0.72). The results clearly validated our hypothesis that the integrated risk factors were more effective in CV/stroke risk stratification compared to CRF alone. Figure 2 indicates the receiver operating characteristic (ROC) plot for the two ML systems: AtheroRisk-Conventional (red color) and AtheroRisk-Integrated system (blue color).

Visual depiction of patient’s carotid scans using AtheroRisk-Integrated

Using an integrated feature-based AtheroRisk-Integrated ML system, all the patients were risk stratified into two risk classes which were determined using a stenosis threshold value of 40%. Figure 3 indicates the outcome of the risk stratification process for two types of CUS scans using AtheroRisk-Integrated ML system. It should be noted that the baseline stenosis values for Figure 3A,B (input and output) and Figure 3C,D (input and output) were 10.50% and 69.87%, respectively. This indicated baseline low-risk and high-risk nature for Figure 3A,B and Figure 3C,D, respectively is correctly captured by the ML-based system with integrated features (AtheroRisk-Integrated).

Discussion

In this study, we demonstrated a CV/stroke risk

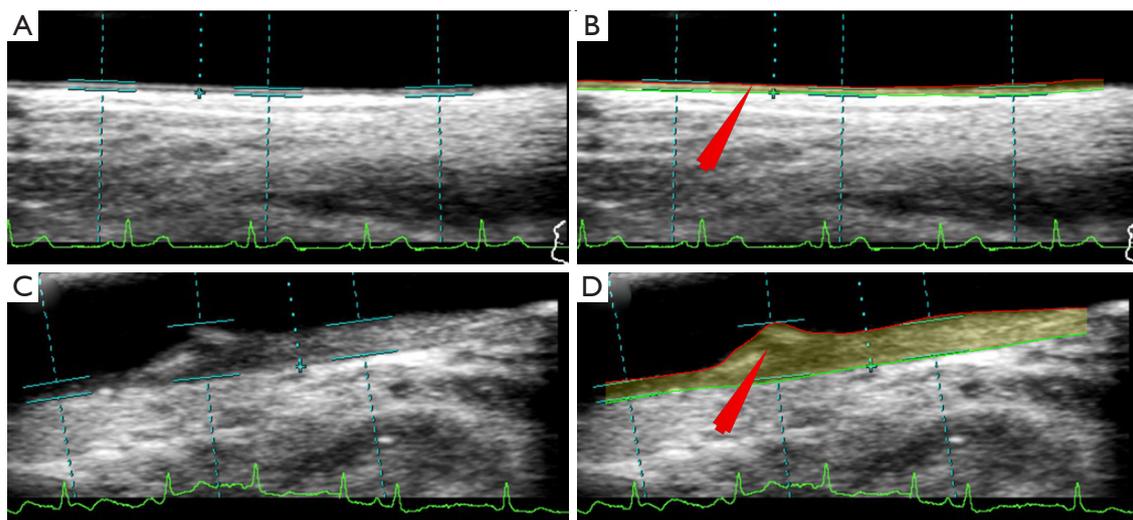


Figure 3 Risk stratification based on automated AtheroRisk-Integrated system. Row 1—patient 109L (low-risk): (A) original image; (B) processed image using AtheroEdge™ 2.0; stenosis: 10.5%. Row 2—patient 10L (high-risk): (C) original image; (D) processed image using AtheroEdge™ 2.0; stenosis: 69.87%.

stratification approach that introduced the concept of a low-cost system that incorporated carotid stenosis as EEGS in an ultrasound framework. It should be noted that the ML systems proposed in this study investigate the risk of CVD/stroke using a surrogate biomarker or EEGS. Second, the proposed study further showed the concept of the effect of integration of 13 CRF with 34 CUSIP features thereby performing a comparison between CRF alone and an integrated approach also taking into account plaque features. The third concept was to adapt a polling-based PCA model for the best combination of feature selection while performing the RF-based classification paradigm for risk prediction and stratification. Thus, the overall system can be characterized with three major contributions: (I) introduction of EEGS in a low-cost ML system design; (II) integration of CRF with image phenotypes and (III) incorporation of ML intelligence embedded with an efficient paradigm for feature selection using a PCA-based polling strategy. Using the above novel combination, AtheroRisk-Integrated showed an improvement in AUC by 18% over AtheroRisk-Conventional for the Japanese diabetic cohort. The system was generalized with stable performance using cross-validation protocol, validating the hypothesis. Note that, even though, American College of Cardiology (ACC) recommended restricted measurement of cIMT within the 10 mm region of CCA (35), our study used AtheroEdge (AtheroPoint, Roseville, USA) for the full-length CUSIP

measurements (36) following the spirit for the usage of CUSIP for CV/stroke risk assessment (37,38).

Stenotic threshold selection for optimal EEGS design and its sensitivity analysis

The risk stratification threshold used for EEGS plays a major role while initiating the preventive measures for CV/stroke events. However, the choice of the threshold depends upon the types of risk factors (covariates or features) included in the risk stratification model and the patients' baseline characteristics. The European Carotid Surgery Trial (ECST) and NASCET have reported the highest incidence of stroke events when the stenosis is $\geq 70\%$ (18,36). Studies have also indicated the use of 50% stenosis threshold for moderate risk of CV/stroke events (17,39). Since the prevalence of stenosis increases with age (17), such moderate-risk patients, if left untreated may lead to severe stenosis (i.e., $\geq 70\%$) and may further be qualified for carotid endarterectomy procedures. Furthermore, the risk of CV/stroke events increases if the patients have one or more risk factors (40). The Japanese cohort used in our current study has mean stenosis of $21.15\% \pm 10.19\%$ (ranging between 5% and 67%) which is, according to ECST trial (36), categorized into a moderate-risk category. Thus, for this retrospective study, an equivalence stenotic threshold of 40% was selected for CV/

stroke risk stratification (17,39). A similar stenotic threshold of 40% was also used by Prati *et al.* (35) to investigate the stenotic and non-stenotic nature of the CP. The sensitivity of the selected threshold was also analyzed by varying the EEGS threshold by 1% in both directions. This has resulted in an overall change in AUC value by less than 5%. This has indicated a stable and reliable threshold for the stenosis for the ML design for our cohort.

Ranking of dominant features

The PCA-based polling strategy with a cutoff value of 0.99 selected 18 dominant features out of 47 input risk factors to train the AtheroRisk-Integrated ML system. HbA1c was at first place followed by PS, cIMTave, and age. It should be noted that, in a pool of 202 patients, 49 were diabetic (HbA1c $\geq 6.5\%$) and 59 patients were pre-diabetic (HbA1c $\geq 6\%$ and HbA1c $\leq 6.4\%$). This may be the reason for HbA1c to be the first during the ranking of covariates. PS and cIMTave secured the second and third position in the ranking list, respectively. It means these two CUSIP are stronger compared to others and showed high contributions towards the risk of CV/stroke events compared to other CRF. These ranking results are in-line with the recently published study by Cuadrado-Godia *et al.* (41).

Therapeutic implications of ML-based risk stratification

Risk stratification of patients assists the physicians in recommending either the surgical procedures or the use of medications for preventing the occurrence of CV/stroke events. Compared to statistical risk prediction models, ML-based risk assessment systems are becoming better in terms of risk prediction capability (4,42). Statins are generally used as a primary treatment to control lipids thereby lowering the risk of CV/stroke event (43). ML-based risk assessment systems help the physicians in deciding the statin eligibility of patients. A recent study by Kakadiaris *et al.* (42) had recommended the use of statins to 11.1% of their study population (AUC =0.92) using ML-based risk stratification model. In contrast, the statistically derived ACC/AHA calculator recommended the use of statins to 46% (AUC =0.76) of the study population. This clearly indicates the influential role of ML-based system in the risk stratification of patients (4). Thus, in comparison to these conventional risk assessment tools, ML systems may be employed and preferred for routine risk assessment in clinical settings (4).

Study limitations, and future scope

We believe that the AtheroRisk-Integrated ML system is efficient, accurate, and affordable. Risk stratification performed using EEGS is the first step to prevent higher costs and simplicity in design (4,11,12,29,41). Although the proposed study has clearly met the pre-defined hypothesis, more modifications can be possible to improve the CV/stroke risk assessment such as conducting a multicenter study. In addition, though we believe the evidence supports the use of carotid stenosis as a robust EEGS metric, we acknowledge that increasingly plaque composition determined by imaging, beyond the lumen stenosis measurements alone, will be important to incorporate in future ML-based studies using plaque phenotypes. Further, risk factors like inflammatory markers (i.e., erythrocyte sedimentation rate and high sensitivity C-reactive protein), renal disease markers (i.e., uric acid and estimated glomerular filtration rate), and arterial/vascular age (44,45) can also be integrated in the future to evaluate the ML system.

Conclusions

This study is focused on the design of a CV/stroke risk stratification keeping three concepts in mind: (I) low-cost by incorporating event-equivalent gold standard in ultrasound framework; (II) integration of 13 CRF with 34 image-based phenotypes; (III) usage of PCA with polling for feature selection followed by an intelligence-based paradigm by adapting simple and efficient classification framework such as RF. The system demonstrated an 18% improvement in integrated ML approach *vs.* the conventional ML approach. We incorporated image-phenotypes completely automatically from the CUS scans and believe that future studies are now warranted examining this integrated ML approach in larger cohorts to aid in improving our methods for the prevention of heart disease and stroke.

Acknowledgments

None.

Footnote

Conflicts of Interest: Dr. Suri is affiliated to AtheroPoint™, focused in the area of stroke and cardiovascular imaging. Dr. Bhatt discloses the following relationships—Advisory

Board: Cardax, Cereno Scientific, Elsevier Practice Update Cardiology, Medscape Cardiology, PhaseBio, Regado Biosciences; Board of Directors: Boston VA Research Institute, Society of Cardiovascular Patient Care, TobeSoft; Chair: American Heart Association Quality Oversight Committee; Data Monitoring Committees: Baim Institute for Clinical Research (formerly Harvard Clinical Research Institute, for the PORTICO trial, funded by St. Jude Medical, now Abbott), Cleveland Clinic (including for the ExCEED trial, funded by Edwards), Duke Clinical Research Institute, Mayo Clinic, Mount Sinai School of Medicine (for the ENVISAGE trial, funded by Daiichi Sankyo), Population Health Research Institute; Honoraria: American College of Cardiology (Senior Associate Editor, Clinical Trials and News, ACC.org; Vice-Chair, ACC Accreditation Committee), Baim Institute for Clinical Research (formerly Harvard Clinical Research Institute; RE-DUAL PCI clinical trial steering committee funded by Boehringer Ingelheim; AEGIS-II executive committee funded by CSL Behring), Belvoir Publications (Editor in Chief, Harvard Heart Letter), Duke Clinical Research Institute (clinical trial steering committees), HMP Global (Editor in Chief, Journal of Invasive Cardiology), Journal of the American College of Cardiology (Guest Editor; Associate Editor), Medtelligence/ReachMD (CME steering committees), Population Health Research Institute (for the COMPASS operations committee, publications committee, steering committee, and USA national co-leader, funded by Bayer), Slack Publications (Chief Medical Editor, Cardiology Today's Intervention), Society of Cardiovascular Patient Care (Secretary/Treasurer), WebMD (CME steering committees); Other: Clinical Cardiology (Deputy Editor), NCDR-ACTION Registry Steering Committee (Chair), VA CART Research and Publications Committee (Chair); Research Funding: Abbott, Amarin, Amgen, AstraZeneca, Bayer, Boehringer Ingelheim, Bristol-Myers Squibb, Chiesi, CSL Behring, Eisai, Ethicon, Ferring Pharmaceuticals, Forest Laboratories, Idorsia, Ironwood, Ischemix, Lilly, Medtronic, PhaseBio, Pfizer, Regeneron, Roche, Sanofi Aventis, Synaptic, The Medicines Company; Royalties: Elsevier (Editor, Cardiovascular Intervention: A Companion to Braunwald's Heart Disease); Site Co-Investigator: Biotronik, Boston Scientific, St. Jude Medical (now Abbott), Svelte; Trustee: American College of Cardiology; Unfunded Research: FlowCo, Fractyl, Merck, Novo Nordisk, PLx Pharma, Takeda. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The patients were approved by the institutional review board of Toho University, Japan and written consent was obtained from all the study participants. The study was approved by Toho University Japan (Ohashi Ethics Committee, Authorization No. 13-56).

References

1. World Health Organization. Cardiovascular diseases (CVDs): key facts. Available online: [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))
2. Libby P. History of discovery: inflammation in atherosclerosis. *Arterioscler Thromb Vasc Biol* 2012;32:2045-51.
3. Thompson JB, Blaha M, Resar JR, et al. Strategies to reverse atherosclerosis: an imaging perspective. *Curr Treat Options Cardiovasc Med* 2008;10:283-93.
4. Jamthikar A, Gupta D, Khanna NN, et al. A special report on changing trends in preventive stroke/cardiovascular risk assessment via B-mode ultrasonography. *Curr Atheroscler Rep* 2019;21:25.
5. Tarkin JM, Dweck MR, Evans NR, et al. Imaging atherosclerosis. *Circ Res* 2016;118:750-69.
6. Boi A, Jamthikar AD, Saba L, et al. A survey on coronary atherosclerotic plaque tissue characterization in intravascular optical coherence tomography. *Curr Atheroscler Rep* 2018;20:33.
7. Polak JF, Pencina MJ, Pencina KM, et al. Carotid-wall intima-media thickness and cardiovascular events. *N Engl J Med* 2011;365:213-21.
8. Park HW, Kim WH, Kim KH, et al. Carotid plaque is associated with increased cardiac mortality in patients with coronary artery disease. *Int J Cardiol* 2013;166:658-63.
9. den Hartog AG, Achterberg S, Moll FL, et al. Asymptomatic carotid artery stenosis and the risk of ischemic stroke according to subtype in patients with clinical manifest arterial disease. *Stroke* 2013;44:1002-7.
10. Kwon H, Kim HK, Kwon SU, et al. Risk of major adverse cardiovascular events in subjects with asymptomatic mild carotid artery stenosis. *Sci Rep* 2018;8:4700.
11. Khanna NN, Jamthikar AD, Gupta D, et al. Performance evaluation of 10-year ultrasound image-based stroke/cardiovascular (CV) risk calculator by comparing against ten conventional CV risk calculators: a diabetic study.

- Comput Biol Med 2019;105:125-43.
12. Khanna NN, Jamthikar AD, Araki T, et al. Nonlinear model for the carotid artery disease 10-year risk prediction by fusing conventional cardiovascular factors to carotid ultrasound image phenotypes: a Japanese diabetes cohort study. *Echocardiography* 2019;36:345-61.
 13. Khanna NN, Jamthikar AD, Gupta D, et al. Effect of carotid image-based phenotypes on cardiovascular risk calculator: AECRS1.0. *Med Biol Eng Comput* 2019;57:1553-66.
 14. Revkin JH, Shear CL, Pouleur HG, et al. Biomarkers in the prevention and treatment of atherosclerosis: need, validation, and future. *Pharmacol Rev* 2007;59:40-53.
 15. Bots ML, Evans GW, Tegeler CH, et al. Carotid intima-media thickness measurements: relations with atherosclerosis, risk of cardiovascular disease and application in randomized controlled trials. *Chin Med J (Engl)* 2016;129:215-26.
 16. Orrapin S, Rerkasem K. Carotid endarterectomy for symptomatic carotid stenosis. *Cochrane Database Syst Rev* 2017;6:CD001081.
 17. de Weerd M, Greving JP, de Jong AW, et al. Prevalence of asymptomatic carotid artery stenosis according to age and sex: systematic review and meta-regression analysis. *Stroke* 2009;40:1105-13.
 18. North American Symptomatic Carotid Endarterectomy Trial. Methods, patient characteristics, and progress. *Stroke* 1991;22:711-20.
 19. Boissel JP, Collet JP, Moleur P, et al. Surrogate endpoints: a basis for a rational approach. *Eur J Clin Pharmacol* 1992;43:235-44.
 20. Raman G, Moorthy D, Hadar N, et al. Management strategies for asymptomatic carotid stenosis: a systematic review and meta-analysis. *Ann Intern Med* 2013;158:676-85.
 21. Molinari F, Meiburger KM, Saba L, et al. Automated carotid IMT measurement and its validation in low contrast ultrasound database of 885 patient Indian population epidemiological study: results of AtheroEdge® software. In: Saba L, Sanches JM, Pedro LM, et al. editors. *Multi-modality atherosclerosis imaging and diagnosis*. New York: Springer, 2014:209-19.
 22. Molinari F, Pattichis CS, Zeng G, et al. 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* 2012;21:1211-22.
 23. Araki T, Ikeda N, Shukla D, et al. 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* 2016;128:137-58.
 24. Shrivastava VK, Londhe ND, Sonawane RS, et al. A novel and robust Bayesian approach for segmentation of psoriasis lesions and its risk stratification. *Comput Methods Programs Biomed* 2017;150:9-22.
 25. Shrivastava VK, Londhe ND, Sonawane RS, et al. Computer-aided diagnosis of psoriasis skin images with HOS, texture and color features: a first comparative study of its kind. *Comput Methods Programs Biomed* 2016;126:98-109.
 26. Molinari F, Meiburger KM, Zeng G, et al. 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* 2012;31:42-53.
 27. Saba L, Mallarini G, Sanfilippo R, et al. Intima media thickness variability (IMTV) and its association with cerebrovascular events: a novel marker of carotid atherosclerosis? *Cardiovasc Diagn Ther* 2012;2:10-8.
 28. Molinari F, Zeng G, Suri JS. Intima-media thickness: setting a standard for a completely automated method of ultrasound measurement. *IEEE Trans Ultrason Ferroelectr Freq Control* 2010;57:1112-24.
 29. Kotsis V, Jamthikar AD, Araki T, et al. Echolucency-based phenotype in carotid atherosclerosis disease for risk stratification of diabetes patients. *Diabetes Res Clin Pract* 2018;143:322-31.
 30. Hirata T, Arai Y, Takayama M, et al. Carotid plaque score and risk of cardiovascular mortality in the oldest old: results from the TOOTH study. *J Atheroscler Thromb* 2018;25:55-64.
 31. Shrivastava VK, Londhe ND, Sonawane RS, et al. Reliable and accurate psoriasis disease classification in dermatology images using comprehensive feature space in machine learning paradigm. *Expert Systems with Applications* 2015;42:6184-95.
 32. Song F, Guo Z, Mei D. Feature selection using principal component analysis. In 2010 international conference on system science, engineering design and manufacturing informatization. *IEEE* 2010;1:27-30.
 33. Ho TK. Random decision forests. In proceedings of 3rd international conference on document analysis and recognition. *IEEE* 1995;1:278-82.
 34. Whelton PK, Carey RM, Aronow WS, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in

- adults: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *J Am Coll Cardiol* 2018;71:e127-248.
35. Prati P, Vanuzzo D, Casaroli M, et al. Prevalence and determinants of carotid atherosclerosis in a general population. *Stroke* 1992;23:1705-11.
 36. MRC European Carotid Surgery Trial: interim results for symptomatic patients with severe (70-99%) or with mild (0-29%) carotid stenosis. European Carotid Surgery Trialists' Collaborative Group. *Lancet* 1991;337:1235-43.
 37. Stein JH, Korcarz CE, Hurst RT, et al. Use of carotid ultrasound to identify subclinical 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* 2008;21:93-111; quiz 189-90.
 38. Greenland P, Alpert JS, Beller GA, et al. 2010 ACCF/AHA guideline for assessment of cardiovascular risk in asymptomatic adults: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol* 2010;56:e50-103.
 39. Goessens BM, Visseren FL, Kappelle LJ, et al. Asymptomatic carotid artery stenosis and the risk of new vascular events in patients with manifest arterial disease: the SMART study. *Stroke* 2007;38:1470-5.
 40. Cohen SN, Hobson RW 2nd, Weiss DG, et al. Death associated with asymptomatic carotid artery stenosis: long-term clinical evaluation. VA Cooperative Study 167 Group. *J Vasc Surg* 1993;18:1002-9; discussion 1009-11.
 41. Cuadrado-Godia E, Jamthikar AD, Gupta D, et al. Ranking of stroke and cardiovascular risk factors for an optimal risk calculator design: Logistic regression approach. *Comput Biol Med* 2019;108:182-95.
 42. Kakadiaris IA, Vrigkas M, Yen AA, et al. Machine learning outperforms ACC/AHA CVD risk calculator in MESA. *J Am Heart Assoc* 2018;7:e009476.
 43. Goff DC Jr, Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol* 2014;63:2935-59.
 44. D'Agostino RB Sr, Vasan RS, Pencina MJ, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. *Circulation* 2008;117:743-53.
 45. Stein JH, Fraizer MC, Aeschlimann SE, et al. Vascular age: integrating carotid intima-media thickness measurements with global coronary risk assessment. *Clin Cardiol* 2004;27:388-92.
 46. Maniruzzaman M, Jahanur Rahman M, Ahammed B, et al. Statistical characterization and classification of colon microarray gene expression data using multiple machine learning paradigms. *Comput Methods Programs Biomed* 2019;176:173-93.
 47. Maniruzzaman M, Rahman MJ, Al-Mehedi Hasan M, et al. Accurate diabetes risk stratification using machine learning: role of missing value and outliers. *J Med Syst* 2018;42:92.
 48. Dimitriadis SI, Liparas D, Alzheimer's Disease Neuroimaging Initiative. How random is the random forest? Random forest algorithm on the service of structural imaging biomarkers for Alzheimer's disease: from Alzheimer's disease neuroimaging initiative (ADNI) database. *Neural Regen Res* 2018;13:962-70.
 49. Marchese Robinson RL, Palczewska A, Palczewski J, et al. Comparison of the predictive performance and interpretability of random forest and linear models on benchmark data sets. *J Chem Inf Model* 2017;57:1773-92.
 50. Breiman L. Random forests. *Machine learning* 2001;45:5-32.
 51. Quinlan JR. Induction of decision trees. *Machine learning* 1986;1:81-106.
 52. Araki T, Jain PK, Suri HS, et al. Stroke risk stratification and its validation using ultrasonic echolucent carotid wall plaque morphology: a machine learning paradigm. *Comput Biol Med* 2017;80:77-96.
 53. Shrivastava VK, Londhe ND, Sonawane RS, et al. Exploring the color feature power for psoriasis risk stratification and classification: a data mining paradigm. *Comput Biol Med* 2015;65:54-68.
 54. Araki T, Ikeda N, Shukla D, et al. A new method for IVUS-based coronary artery disease risk stratification: A link between coronary & carotid ultrasound plaque burdens. *Comput Methods Programs Biomed* 2016;124:161-79.

Cite this article as: Jamthikar A, Gupta D, Khanna NN, Saba L, Araki T, Viskovic K, Suri HS, Gupta A, Mavrogeni S, Turk M, Laird JR, Pareek G, Miner M, Sfikakis PP, Protogerou A, Kitas GD, Viswanathan V, Nicolaides A, Bhatt DL, Suri JS. A low-cost machine learning-based cardiovascular/stroke risk assessment system: integration of conventional factors with image phenotypes. *Cardiovasc Diagn Ther* 2019;9(5):420-430. doi: 10.21037/cdt.2019.09.03

Supplementary file 1: PCA-based feature selection

Feature selection techniques minimize the redundancy in the feature set and select only the dominant features which improves the efficiency of the ML algorithm. PCA is one of the most widely accepted and efficient feature selection techniques. The individual application of PCA reduces the dimensionality of the feature space by altering the feature values. In order to preserve the feature values a polling-based PCA strategy was recently presented (32) which extracts the indices of the dominant features. A detailed PCA polling strategy has been already discussed in our previous studies (25).

Supplementary file 2: role of classifier and choice of RF

The role of ML-based classifier is to categorize the input data into predefined labels or classes. For example, in a CVD/stroke event prediction task, using the input features classifier predicts either “event” or “no-event” category. In this study, the ML-based classifier identifies one of the two risk profiles of the patients: (I) low-risk or (II) high-risk. Since the focus of this study was to develop a cost-effective, efficient ML system, a RF classifier was incorporated in the ML system for risk stratification of the patients (33). Our previously published studies reported a better performance while using RF for classification tasks (46,47). Furthermore, RF is a commonly used algorithm which reported higher predictive ability compared to other ML-based algorithms (48,49). Thus, RF classifier was selected in our proposed study to risk stratify the patients.

The term RF was first coined by Tin Kan Ho from Bell labs in 1995 (33). RF is a type of ensemble learning algorithm which is used in classification or regression applications (50). The term ensemble learning indicates a combination of several decision tree (DT) classifiers that provide a voting-based final decision to perform a classification or regression task. DT is a fundamental building block of RF classifier (51). As the name indicates, DT is a set of multiple decisions that are required to perform the classification and regression task (51). RF is a combination of multiple such DT classifiers. In our current study, RF was used for classifying Japanese patients into two risk categories: low-risk and high-risk. The main advantage of RF is that it is a better fit for the categorical data after obtaining the final solution in the majority voting system, where the result of each tree is judged. In this study, a total of the 400 trees were used in an RF algorithm.

Supplementary file 3: performance evaluation using RF

The performance of the two types of ML-based was further

evaluated using the risk stratification accuracy, sensitivity, and specificity. *Figures S1-S4* shows the bar charts for four performance evaluation metrics such as accuracy, AUC, sensitivity, and specificity. Bar charts were plotted for all the 10-trials. Each value of bar-chart is the mean of 10 different combinations of 10-fold cross-validation. Note that here each trial contains 10 independent combinations of training and testing sets. The entire dataset gets shuffled from one trial to another trial. The legends in each bar chart represent the mean overall the 10 trials. All these four types of performance evaluation metrics were also tabulated in *Table S1*.

The sensitivity indicates the likelihood of detecting high-risk patients by an automated ML-based algorithm when EEGS also indicates the high-risk status for the same patient. Similarly, specificity indicates the likelihood of detecting the low-risk patients by the automated ML-based algorithm when EEGS also indicates the low-risk status for the same patient. Ideally, both sensitivity and specificity should be 100%. This indicates that all the high-risk and low-risk patients, respectively, are correctly identified by the ML-based system. The area under the ROCs curve indicates the trade-off between sensitivity and specificity values. *Figures S2,S4* represents mean sensitivity and mean specificity, respectively, for 10-fold cross-validation. It should be noted that the proposed ML-based system was highly specific to non-high-risk patients (or low-risk patients). It is indicated by the higher values of specificity for both AtheroRisk-Conventional (96.46%) and AtheroRisk-Integrated systems (99.15%). At the same time, a low sensitivity was observed for both types of ML-based systems. Low sensitivity indicates low predictive power for high-risk patients. This may be because of the sample number of high-risk patients (high-risk patients: 12) available in the proposed study. Furthermore, the difference in mean sensitivities between AtheroRisk-Conventional and integrated model is very low. This is primarily due to very low (~4%) high-risk samples. Typically, a good ML system behavior requires equal distribution of the risk classes. Due to imbalance, the noisy features are less likely to be in generalized pool for the risk stratification (23,46,52-54).

The AUC is a tradeoff between both sensitivity and specificity. This was the reason for projecting the overall analysis using only AUC in our proposed study. As shown in *Figure S1*, the mean AUC over 10-trials and 10-fold cross-validation were 0.68 for AtheroRisk-Conventional and 0.80 for AtheroRisk-Integrated system. Similarly, the overall risk stratification accuracy was reported to be 92.77% for AtheroRisk-Conventional and 95.15% for AtheroRisk-Integrated system.

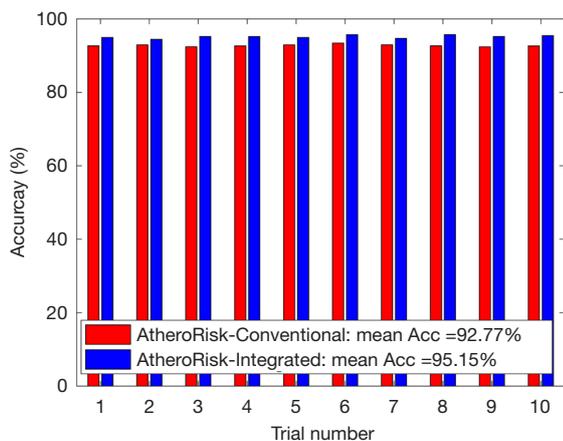


Figure S1 Bar chart showing the risk stratification accuracy plotted against the 10 trails for both conventional and integrated ML systems. ML, machine learning; Acc, accuracy.

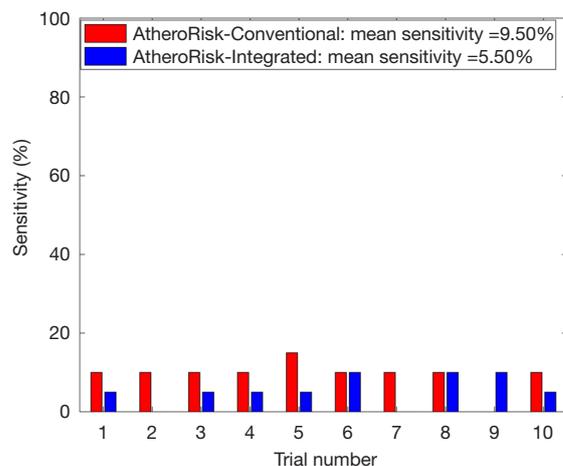


Figure S3 Bar chart showing the sensitivity against the 10 trails for both conventional and integrated ML systems. ML, machine learning.

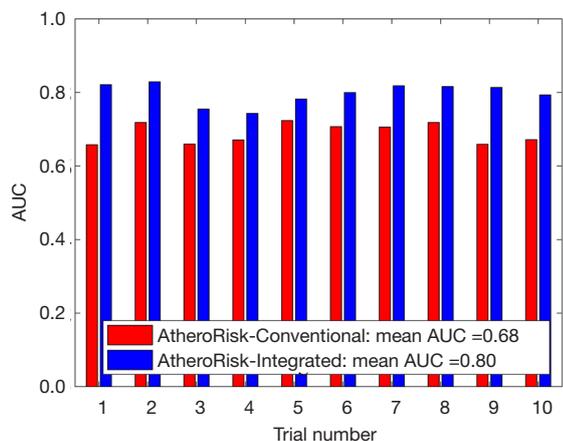


Figure S2 Bar chart showing the AUC against the 10 trails for both conventional and integrated ML systems. ML, machine learning; AUC, area under the curve.

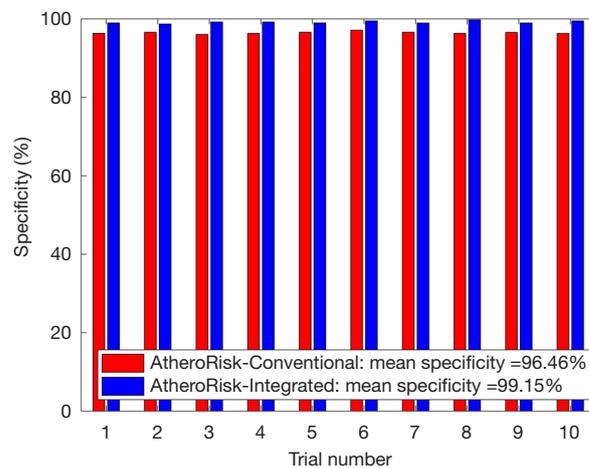


Figure S4 Bar chart showing the specificity against the 10 trails for both Conventional and integrated ML systems. ML, machine learning.

Table S1 Performance evaluation metrics for AtheroRisk-Conventional and AtheroRisk-Integrated ML based system with RF classifier and K10 protocol

Trial number	Accuracy (%)		AUC		Sensitivity (%)		Specificity (%)	
	Conventional	Integrated	Conventional	Integrated	Conventional	Integrated	Conventional	Integrated
1	92.67	94.94	0.66	0.82	10.00	5.00	96.30	98.95
2	92.94	94.44	0.72	0.83	10.00	0.00	96.56	98.68
3	92.42	95.20	0.66	0.75	10.00	5.00	96.03	99.20
4	92.65	95.19	0.67	0.74	10.00	5.00	96.29	99.20
5	92.94	94.94	0.72	0.78	15.00	5.00	96.57	98.95
6	93.42	95.70	0.71	0.80	10.00	10.00	97.11	99.47
7	92.95	94.69	0.71	0.82	10.00	0.00	96.57	98.93
8	92.67	95.71	0.72	0.82	10.00	10.00	96.32	99.74
9	92.40	95.20	0.66	0.81	0.00	10.00	96.55	98.94
10	92.65	95.46	0.67	0.79	10.00	5.00	96.28	99.47
Mean	92.77	95.15	0.69	0.80	9.50	5.50	96.46	99.15
SD	0.30	0.41	0.03	0.03	3.69	3.69	0.29	0.32

Each trial comprised of 10 independent combinations of training and testing datasets (K10 protocol). ML, machine learning; RF, random forest; AUC, area under the curve; SD, standard deviation.