Associative Prediction of Carotid Artery Plaques Based on Ultrasound Strain Imaging and Cardiovascular Risk Factors in People Living With HIV and Age-Matched Control Subjects of the CHACS Cohort

Marie-Hélène Roy Cardinal, PhD,[a] Madeleine Durand, MD, MSc, FRCPC,[b,c] Carl Chartrand-Lefebvre, MD, MSc,[d,e] Gilles Soulez, MD, MSc,[d,e,f] Cécile Tremblay, MD,[g,h] and Guy Cloutier, Eng, PhD,[a,e,f] for the Canadian HIV and Aging Cohort Study

Background: There is a need for a specific atherosclerotic risk assessment for people living with HIV (PLWH).

Setting: A machine learning classification model was applied to PLWH and control subjects with low-to-intermediate cardiovascular risks to identify associative predictors of diagnosed carotid artery plaques. Associations with plaques were made using strain elastography in normal sections of the common carotid artery and traditional cardiovascular risk factors.

Methods: One hundred two PLWH and 84 control subjects were recruited from the prospective Canadian HIV and Aging Cohort Study (57 ± 8 years; 159 men). Plaque presence was based on clinical ultrasound scans of left and right common carotid arteries and internal carotid arteries. A classification task for identifying subjects with plaque was defined using random forest (RF) and logistic regression models. Areas under the receiver operating characteristic curves (AUC-ROCs) were applied to select 5 among 50 combinations of 4 or less features yielding the highest AUC-ROCs.

Results: To retrospectively classify individuals with and without plaques, the 5 most discriminant combinations of features had AUC-ROCs between 0.76 and 0.79. AUC-ROCs from RF were statistically significantly higher than those obtained with logistic regressions (P = 0.0001). The most discriminant features of RF classifications in PLWH were age, smoking status, maximum axial strain and pulse pressure (equal weights), and sex and antiretroviral therapy exposure (equal weights). When considering the whole population, the HIV status was identified as a cofactor associated with carotid artery plaques.

Conclusions: Strain elastography adds to traditional cardiovascular risk factors for identifying individuals with carotid artery plaques.

Key Words: carotid atherosclerosis, carotid artery plaque, ultrasound imaging, ultrasound strain elastography, cardiovascular risk factors

(*J Acquir Immune Defic Syndr* 2022;91:91–100)

INTRODUCTION

Cardiovascular diseases (CVDs) are still the leading cause of death worldwide.[1] In the case of people living with HIV (PLWH), the risk of developing CVDs is 2 times higher than in the general population.[2] This elevated risk remains after accounting for common risk factors in PLWH such as smoking, dyslipidemia, or hypertension.[3] As the global burden of HIV-associated CVD tripled from 1990 to 2015,[2] a scientific statement of the American Heart Association has established the need for an atherosclerotic risk assessment and prevention approach that is specific to PLWH.[3]

The common carotid intima-media thickness (cIMT) evaluated in B-mode ultrasound (US) has long been used to assess subclinical atherosclerosis. cIMT was shown to be

Received for publication January 4, 2022; accepted April 26, 2022.

From the [a]Laboratory of Biorheology and Medical Ultrasonics, University of Montreal Hospital Research Center, Montréal, Québec, Canada; [b]Department of Internal Medicine, University of Montreal Hospital, Montréal, Québec, Canada; [c]Department of Medicine, University of Montreal, Montréal, Québec, Canada; [d]Department of Radiology, University of Montreal Hospital, Montréal, Québec, Canada; [e]Department of Radiology, Radio-oncology, and Nuclear Medicine, University of Montreal, Montréal, Québec, Canada; [f]Institute of Biomedical Engineering, University of Montreal, Montréal, Québec, Canada; [g]Department of Microbiology and Infectious Diseases, University of Montreal Hospital, Montréal, Québec, Canada; and [h]Department of Microbiology, Infectious Diseases, and Immunology, University of Montreal, Montréal, Québec, Canada.

This study was initiated through the support of the Canadian Institute of Health Research (CIHR, Canadian HIV and Aging Cohort Study, group Grant #284512); it is now supported by CIHR group Grant #398643 and project Grant #399544.

C.C.-L. has research collaboration with Siemens Healthineers; G.S. has research grants from Siemens Healthineers, Vitaa Medical, Starpax Medical, and Cook Medical that are not related to this work; and G.C. has research grants from Siemens Healthineers and Vitaa Medical that are not related to this work. The other authors have no conflicts of interest to disclose.

M.-H.R.C. performed strain imaging, machine learning, and statistical analyses, and drafted the manuscript; M.D. is managing the CHACS cohort; C.C.-L. contributed in clinical Doppler examination reporting and is managing the cardiovascular component of CHACS; G.S. contributed in clinical Doppler examination reporting and co-pioneered the development of the ultrasound strain imaging method; C.T. co-conceived the study; and G.C. conceived the study, pioneered the development of the ultrasound strain imaging method, and drafted the manuscript. All authors contributed to and approved the final manuscript.

Correspondence to: Guy Cloutier, Eng, PhD, Laboratory of Biorheology and Medical Ultrasonics, University of Montreal Hospital Research Center, 900 Saint-Denis, room R11-464 Montréal, Québec H2X 0A9, Canada (e-mail: guy.cloutier@umontreal.ca).

associated with future vascular events in a nonlinear way[4] and to add in risk stratification for the occurrence of CVDs in asymptomatic individuals at intermediate cardiovascular risk.[5] However, it was also shown that adding cIMT to traditional cardiovascular risk prediction models did not statistically significantly increase the model performance[6] or that the improvement was unlikely to be of clinical importance.[7] Also, it is unclear if cIMT progresses differently in PLWH than in control subjects with reports showing similar progression on one hand[8] or greater progression on the other hand.[9] Carotid artery plaque assessment has also been used to characterize atherosclerosis because a different pathophysiology underlies plaque development compared with cIMT thickening.[10] Carotid plaque presence, a more advanced stage in the pathogenesis of atherosclerosis,[11] was shown to be stronger than cIMT in predicting strokes[12] and coronary artery disease events.[13] It was also shown that increased cIMT did not predict new plaque development in a prospective analysis after accounting for vascular risk factors.[14] In the case of PLWH, higher rates of progression of carotid artery plaque and thicker cIMT were found when compared with control subjects; however, over time, when cIMT progression did not differ between groups, new focal carotid artery plaque occurrence was higher in PLWH.[15] In individuals with low-to-moderate cardiovascular risk, higher carotid plaque occurrence was associated with HIV status when no difference was detected in cIMT measurements.[16]

US-based strain imaging provides novel biomarkers that quantify the motion and deformation of the carotid artery wall induced by blood pulsation.[17,18] Strain imaging has a wide variety of applications in CVD assessment: aortic strain has been used to evaluate early vascular changes,[19] whereas myocardial strain and aortic stiffness have been associated with heart failure prognosis.[20,21] Carotid strain imaging is noninvasive and produces local measurements related to arterial elasticity compared with pulse wave velocity that assesses arterial elasticity globally.[22] Strains can be estimated within internal and common carotid artery walls and plaques.[23] Smaller translation and strain of the carotid artery wall were measured using this method in PLWH compared with age-matched controls, indicating premature arterial stiffening in the former population.[16] Other US-based approaches for indirectly assessing vascular wall elasticity have also found increased arterial rigidity in PLWH.[24–26]

Because US strain imaging could detect premature atherosclerosis better than cIMT in PLWH and other populations,[16,17] we investigated if it could be associated with plaque presence in PLWH having low-to-moderate cardiovascular risk. For this purpose, a machine learning strategy was used because it has been shown to significantly improve the accuracy of cardiovascular risk prediction compared with traditional statistical models.[27] We retrospectively classified individuals with and without plaque with features that would be available for prospective evaluations. A control HIV negative population was used for comparison. Traditional cardiovascular risk factors, clinical US measures, and US strain imaging biomarkers were used as inputs in classification models. Machine learning based on random forest (RF) and logistic regression (LR) classifications were tested.

METHODS

Study Participants

One hundred two PLWH and 84 control subjects were prospectively recruited from the Canadian HIV and Aging Cohort Study.[28] Inclusion criteria consisted of (1) being at least 40 years of age for both groups or having lived with HIV infection for at least 15 years for PLWH subjects, (2) having a life expectancy of at least 1 year, (3) having a low-to-intermediate cardiovascular risk defined as a 10-year Framingham score between 5% and 20%, and (4) being free of clinically overt CVD.[29] The life expectancy criterion was left at the judgment of the recruiting physician and it was included to consider the prospective nature of the Canadian HIV and Aging Cohort Study cohort with follow-up visits. This study has received approval from the institutional review board on ethics of the Centre Hospitalier de l'Université de Montréal. Written informed consent was obtained from all participants.

All recruited individuals had standardized interviews and blood sample collection during their research visit. Eighteen clinical variables were retained for the current retrospective machine learning study: age, sex, HIV status, Framingham risk score, creatinine level, low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, heart rate during US scan, pulse pressure, systolic and diastolic blood pressures, high waist circumference, diabetes status, hypertension status, statin usage, smoking status, pack-years (pack-years is the product of the number of packs of cigarettes smoked per day by the number of years the person has smoked), and physical activity level. Covariables specific to HIV were also considered, namely antiretroviral therapy (ART) exposure and duration, HIV duration, $CD4^+$ T-cell count, and ratio of $CD4^+$ to $CD8^+$ cell count.

Clinical US and Strain Imaging

Each participant had a standard clinical duplex color and pulsed Doppler US examination, followed by US radiofrequency (US_{RF}) image acquisitions for strain imaging. Clinical scans on longitudinal views of left and right common carotid arteries (CCAs) and internal carotid arteries (ICAs) were performed using a IU22 US system (Philips Healthcare, 11-L-3 probe) to identify diseased artery segments and grade the severity of atherosclerotic plaques. Acquired image features were CCA and ICA Doppler velocities at peak systole and end-diastole. The resistivity index[30] and stenosis percentage[31] were computed from Doppler velocity waveforms. The resistivity index was averaged from measurements on both sides for CCA and ICA. The maximum stenosis percentage grade ($<$50%, 50%–69%, $>$70%) from both sides of ICA was reported.

US_{RF} acquisitions were performed on arterial walls of left and right CCA and ICA resulting in 4 cine-loops for each participant. Longitudinal views of vessels were acquired with an Aixplorer system (SuperSonic Imagine) using a 256-element SuperLinear SL15-4 probe. Participants were asked to hold breath during acquisitions. The frame rate was set to 50 frames per second, and 5-second cine-loops were

recorded. In postprocessing, contours of a normal portion of the CCA far wall (free of plaque or calcium) were delineated in each image of the US_{RF} cine-loop using a semi-automatic segmentation method.[32] The cIMT was computed from automated segmentations as the mean distance between contours over the whole acquisition, and averaged over left and right sides. When plaques were present, they were delineated using the same semi-automatic segmentation method to report means from both sides and maximum thicknesses. When several plaques were present, the mean thickness of each plaque was first computed; these mean thicknesses were then averaged. The maximum thickness of each plaque was also measured; the highest value of all plaques was reported. The categorical variable defining the presence of plaque in the machine learning classification was a portion of the vessel wall thicker than 1 mm and protruding inside the lumen on near and/or far walls of any artery segment (left and right CCA and ICA).

Elastography cine-loops of axial and lateral translation motion maps, strains, and shear strain maps were computed from the US_{RF} data of the segmented normal wall section of the CCA on left and right sides.[33] These images quantified the motion and deformation induced by the artery pulsation. Axial was defined as parallel to the US beam direction, whereas lateral referred to motion and deformation perpendicular to it. Translations correspond to rigid displacements of carotid artery wall tissue components occurring during cardiac cycles. Strain represents either compression or dilation of the carotid artery wall, whereas shear strain can be viewed as an angular deformation or change in shape of carotid artery wall components. Six elastography features, namely the maximum axial strain (MaxAS) and cumulated axial strain (CAS), maximum and cumulated axial shear strain magnitude (Max|ShS| and C|ShS|), cumulated axial translation (CAT), and cumulated lateral translation (CLT), were extracted from elastograms.[16] Elastography features were averaged over left and right sides.

Associative Prediction Models

A dichotomous classification task for identifying subjects with one or more plaques was defined. Three US image features (resistivity indices in CCA and ICA, and cIMT), 6 elastography features on normal CCA (MaxAS, CAS, Max|ShS|, C|ShS|, CAT, and CLT), and 18 clinical covariates defined above were selected as inputs into 2 association prediction models: RF classification and LR. The same task was also defined using the subset of PLWH, where 5 supplementary clinical biomarkers were available as inputs (ART exposure, ART duration, HIV duration, $CD4^+$, and $CD4^+/CD8^+$).

RF Classification

Classification of the presence of plaque in CCA or ICA was first performed with the RF classification model.[34] Steps to realize classification with the selected framework were similar to those used in Ref. 35. To reduce data dimensionality and prevent overfitting, features selection was performed using the G-mean metric (G-mean = $\sqrt{\text{sensitivity} \times \text{specificity}}$). The G-mean has been recommended in the case of imbalance datasets as expected in the case of carotid plaque presence. The G-mean was evaluated from classification results of all possible combinations of 4 or less features among all available ones. The maximum number of features was determined empirically to minimize dimensionality, generalization error, and computation time. Fifty combinations of features with highest G-mean values were selected for computation of receiver operating characteristic (ROC) curves. For each combination, the ROC curve was computed using stratified "0.632+" bootstrap samples,[36] which yielded areas under each ROC curve (AUC). Participant data from bootstrap samples were used to train the classifier, whereas data excluded from the bootstrap sample were part of test sets. Five combinations of features with highest AUC-ROC were finally selected.

Logistic Regression

The same framework as described above was used with LRs. Classifications aimed to predict the presence of plaque using the same set of 27 or 32 (for PLWH) features. Feature selection was applied before computing the AUC-ROCs of 50 combinations of 4 or less features with highest G-means. ROC curves were also estimated with the "0.632+" bootstrap method.[36] Five combinations of features with highest AUC-ROCs were then selected.

Data Imputation

A proportion of subjects in the database had incomplete data (at least one feature value missing). A complete case analysis consists in restricting the dataset to participants with complete data. Multiple imputation, which is a simulation-based method for replacing missing data, was alternatively used to include data from all participants. Multiple imputation has been shown to provide unbiased estimates of features[37] and to measure outcome effects that are closer to those measured on a "true" full-dataset than on the complete case dataset.[38] Five different datasets were thus imputed using the multivariate imputation by chained equations (MICE) method.[39] Imputed datasets were identical for nonmissing data entries but differed in imputed values. Moreover, different types of variables (continuous or categorical) were handled with MICE. For each missing entry of a given feature, imputated values were randomly drawn from the posterior distribution of missing data, conditional on observed data. All features except plaque presence and HIV status were used in distribution models. The classification model was applied independently to the complete case and to 5 imputed datasets.

Statistical Analyses

Identification of subjects with plaques was performed with both classification models using all participants or only PLWH. Statistical analyses, data imputations, and classifications were conducted with software *R* (version 3.6.2, 2019; *R* Foundation for Statistical Computing, Vienna, Austria, https://www.r-project.org). Data imputation and RF classification

used packages "MICE" (version 3.13.0, 2021)[39] and "randomForest" (version 4.6–14, 2018),[40] respectively. A 3-way analysis of variance with factors of classification model (RF or LR), database (all subjects or PLWH), and dataset (complete or imputed) was also performed.

RESULTS

Figures 1 and 2 each shows a B-mode image of a CCA with superimposed axial strain and axial shear strain magnitude maps (2 examples among 6 elastography maps used to quantify the vessel wall mechanical behavior). For each strain map, average and cumulated time-varying curves over 3 cardiac cycles are also displayed. Carotid artery strains displayed in Figure 1 had lower magnitude than the ones displayed in Figure 2; lower strains are associated with stiffer arteries.

As presented in Table 1, carotid artery plaques were observed in 61 among 186 participants; plaque prevalence was thus 33%. One subject with a plaque had a stenosis between 50% and 69%, whereas remaining participants with plaques had stenoses below 50% (one missing data). The age range was 40–75 years for PLWH and 40–74 years for control subjects. In PLWH (Table 2), 43 of 102 participants had plaques (prevalence of 42%). All features used in classification models for groups of subjects with and without plaques are detailed in Tables 1 and 2. There were statistically significantly higher age, HIV positive number, Framingham risk score, current and former smokers, pack-years, cIMT, and C|ShS| in individuals with plaques (Table 1). For clinical biomarkers specific to PLWH, HIV duration was statistically significantly higher in individuals with plaques (Table 2). A statistically significantly higher feature corresponds to a higher average value or occurrence (for categorical variables) with a significance level of $P < 0.05$. At least one feature in Tables 1 and 2 was missing for 38 subjects (20%), resulting in a complete case dataset for 148 participants overall (69 PLWH and 79 control subjects). For each feature, the number of subjects with missing data is shown in the fourth column of Tables 1 and 2.

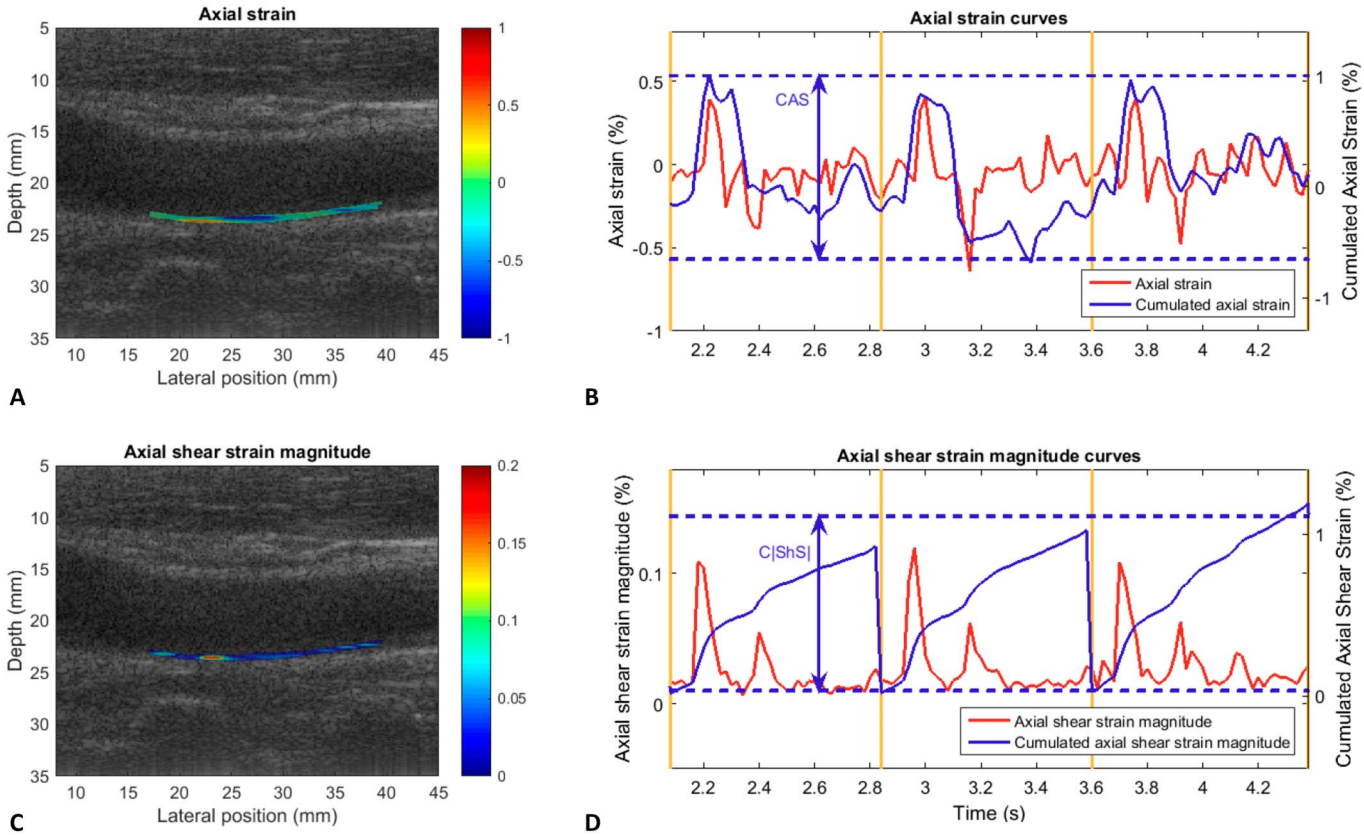

FIGURE 1. Right common carotid artery of a 60-year-old HIV-infected man with low strains. Two-dimensional longitudinal B-mode image reconstructed from radiofrequency data with axial strain superimposed on the segmented far wall of the carotid artery (A), and axial strain time-varying curves over 3 cardiac cycles (B). The MaxAS is computed from the maximum of each cycle (red curve), whereas the CAS is displayed in (B). Panels (C and D) show the axial shear strain magnitude superimposed on the B-mode image, and time-varying axial shear strain magnitude curves. The Max|ShS| is computed from the maximum of each cycle (red curve), whereas the C|ShS| is displayed in (D). Strains and shear strains were cumulated over each cardiac cycle independently. Cardiac cycles are delimited by yellow lines on panels (B and D). Color maps depict strain and shear strain in percentage (%). full color online

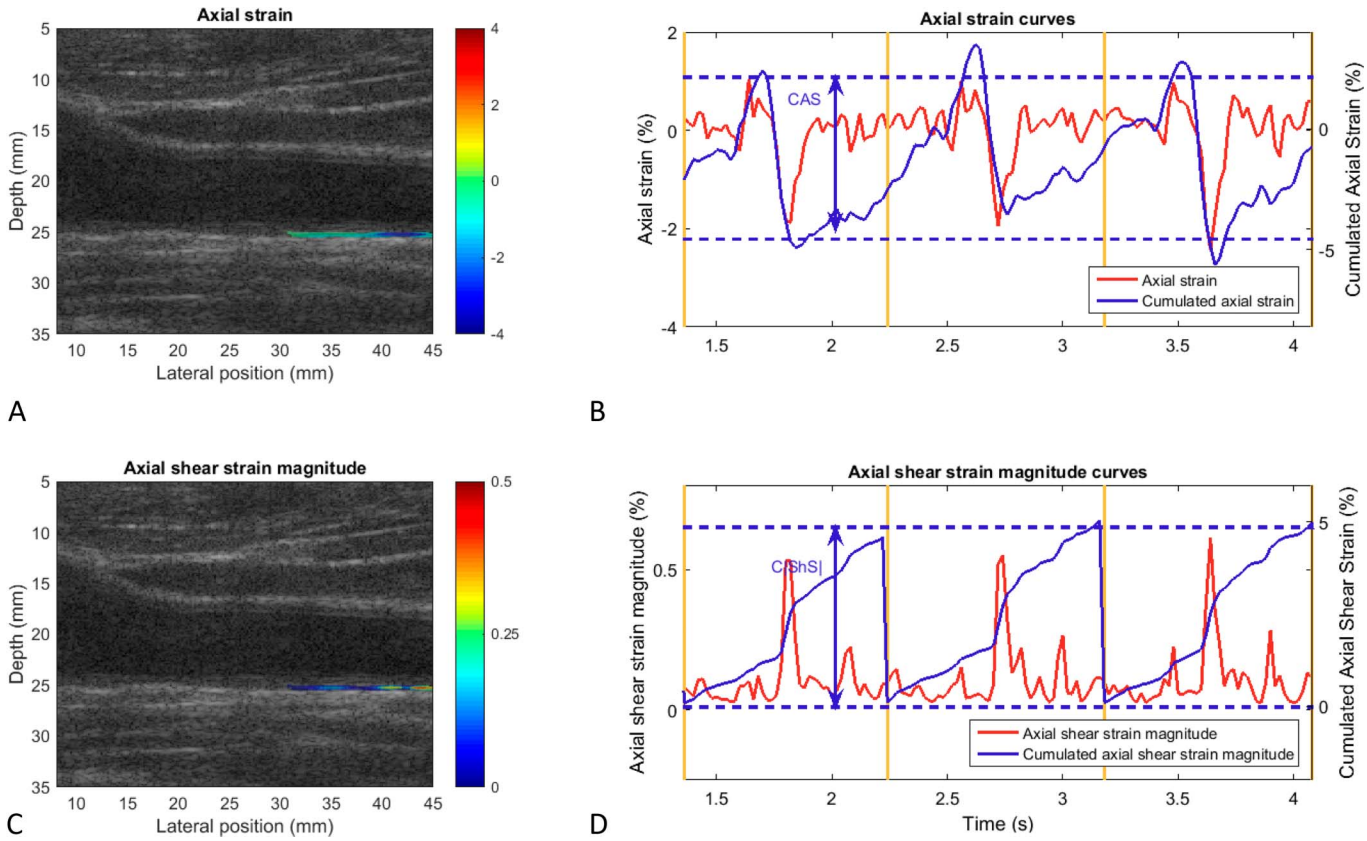

FIGURE 2. Right common carotid artery of a 59-year-old male control participant with high strains. Two-dimensional longitudinal B-mode image reconstructed from radiofrequency data with axial strain superimposed on the segmented far wall of the carotid artery (A), and axial strain time-varying curves over 3 cardiac cycles (B). The MaxAS is computed from the maximum of each cycle (red curve), whereas the CAS is displayed in (B). Panels (C and D) show the axial shear strain magnitude superimposed on the B-mode image, and time-varying axial shear strain magnitude curves. The Max|ShS| is computed from the maximum of each cycle (red curve), whereas the C|ShS| is displayed in (D). Strains and shear strains were cumulated over each cardiac cycle independently. Cardiac cycles are delimited by yellow lines on panels (B and D). Color maps depict strain and shear strain in percentage (%). full color online

Classification Models

To classify individuals with and without plaques, 50 combinations of 4 or less features were tested for each of the 6 datasets (5 imputations and 1 complete dataset) for a total of 300 combinations. Thirty combinations, corresponding to 5 best combinations for each of the 6 datasets, provided highest AUC-ROCs ranging between 0.77 and 0.79 for the RF classification model. In the case of LR classification, highest AUC-ROCs were between 0.76 and 0.78. Using only PLWH, highest AUC-ROCs to classify individuals with and without plaques ranged from 0.77 to 0.81 for RF and from 0.72 to 0.77 for LR. Figure 3 shows ROC curves with highest AUCs of all 6 datasets for each classification model, and for each of the 2 databases (all subjects and PLWH). Figure 4 shows the 5 highest AUC-ROCs obtained with imputed and complete datasets for both populations.

The 3-way analysis of variance showed statistically significant effects of the classification model (P = 2e-16), database (P = 1.9e-15), and dataset (P = 2e-16). AUC-ROCs were statistically significantly higher when using the RF classification model. Higher AUC-ROCs were also obtained with the whole database (largest sample size). They were also higher when using the complete dataset compared with imputations; but there was no statistically significant difference between classification performances using the different imputated datasets. Table 3 shows averaged AUC-ROC values for both classification models and databases, with P-values of Tukey multiple comparisons test of classification model factor.

Associative Predictions

For both RF and LR, the 50 combinations with highest AUC-ROCs contained only 12 of the 27 features selected in the classification models. Table 4 shows these 12 features and their occurrence in the 50 top combinations to classify subjects with plaques from the database including all participants. Traditional cardiovascular risk factors of age, smoking status, pulse pressure, and HIV status were most often selected as predictors of carotid plaque presence with RF (Table 4, left column). Four US strain elastography parameters contributed to the highest AUC-ROC classifications; the most frequent being the MaxAS.

In LR, the 5 clinical features associated with highest AUC-ROCs were age, HIV status, smoking status, pack-years,

TABLE 1. Clinical Biomarkers, Clinical Ultrasound Imaging Features, and Ultrasound Strain Imaging Features Considered in Machine Learning Models for Individuals With and Without Plaques

	With Plaque (61)	No Plaque (125)	Missing N	*P*
Clinical features				
Age (yr)	61 ± 8	55 ± 7	0	**7.8e-7***
Male, n (%)	55 (90%)	104 (83%)	0	0.30
HIV positive, n (%)	43 (70%)	59 (47%)	0	**0.005***
Framingham risk score	14.3 ± 7.8	10.6 ± 5.3	14	**0.0028***
Creatinine	81 ± 15	80 ± 15	5	0.53
Low-density lipoprotein cholesterol	2.80 ± 1.06	3.01 ± 0.85	13	0.084
High-density lipoprotein cholesterol	1.38 ± 0.38	1.32 ± 0.40	9	0.30
Heart rate (beats per min)	62 ± 6	64 ± 8	0	0.51
Pulse pressure (mm Hg)	47 ± 11	46 ± 8	1	0.35
Systolic blood pressure (mm Hg)	121 ± 13	122 ± 11	1	0.55
Diastolic blood pressure (mm Hg)	75 ± 9	76 ± 9	1	0.27
High waist circumference, n (%)	17 (28%)	35 (28%)	1	1.00
Diabetes, n (%)	1 (2%)	6 (5%)	0	0.51
Hypertension, n (%)	19 (31%)	30 (24%)	1	0.19
Statin usage, n (%)	19 (31%)	23 (18%)	0	0.078
Smoking status, n (%)			2	**0.005***
Current smoker	18 (30%)	16 (13%)		
Former smoker	28 (46%)	48 (38%)		
Never smoker	15 (25%)	59 (47%)		
Pack-years	21 ± 25	11 ± 19	6	**0.0007***
Physical activity, n (%)			6	0.060
30 min, daily	28 (46%)	47 (38%)		
30 min, 3 times a wk	6 (10%)	34 (27%)		
30 min, weekly	4 (7%)	7 (6%)		
No physical activity	21 (34%)	33 (26%)		
Clinical ultrasound imaging features				
Resistivity index CCA	0.75 ± 0.05	0.75 ± 0.05	1	0.68
Resistivity index ICA	0.64 ± 0.05	0.63 ± 0.05	9	0.23
cIMT (mm)	0.57 ± 0.08	0.54 ± 0.06	0	**0.002***
Plaque thickness			0	
Average (mm)	1.25 ± 0.34	NA		
Maximum (mm)	1.98 ± 0.73	NA		
Stenosis percentage, n (%)			1	
<50%	59	NA		
50%–69%	1	NA		
≥70%	0	NA		
Strain imaging features				
MaxAS (%)	0.53 ± 0.19	0.50 ± 0.20	0	0.18
CAS (%)	3.16 ± 1.42	2.77 ± 1.11	0	0.14
Max|ShS| (%)	0.19 ± 0.07	0.19 ± 0.09	0	0.93
C|ShS| (%)	2.73 ± 0.99	2.48 ± 1.14	0	**0.016***
CAT (%)	0.28 ± 0.10	0.27 ± 0.10	0	0.72
CLT (%)	0.43 ± 0.19	0.43 ± 0.17	0	0.90

Column "Missing N" shows the number of subjects with missing data.
Pulse pressure: systolic minus diastolic blood pressure.
High waist circumference: over 102 cm in males or over 88 cm in females.
*Statistically significant at $P < 0.05$ (unpaired *t*-tests for continuous normally distributed variables, Mann–Whitney tests for continuous nonnormally distributed variables, and χ^2 goodness-of-fit tests for dichotomous variables).

and sex (Table 4, right column). The most contributing US strain imaging feature in the LR model was the Max|ShS|. The cIMT had the same occurrence as Max|ShS|. Diabetes, pulse pressure, and CAS were also part of feature combinations yielding the highest AUC-ROCs.

For the database containing exclusively PLWH, combinations of features yielding the highest AUC-ROCs contained 13 and 15 of the 32 features for RF and LR classification, respectively; these features and their occurrence are shown in Table 5. All US strain imaging features occurred in some of the highest AUC-ROC combinations of features. Features specific to PLWH were used in classification models: ART exposure and duration were part of the best AUC-ROC combinations for RF, and ART exposure and HIV duration were retained biomarkers for LR. In both classification models, the 5 most occurring features to classify all subjects were also in the most occurring ones for PLWH (Tables 4 and 5).

DISCUSSION

In this cross-sectional study, carotid US strain imaging of a plaque-free arterial wall segment was performed in PLWH and HIV-negative subjects. Vascular elasticity measures provide information about arterial wall hardening associated with CVD risk and events[41–43] and with HIV infection.[24,25] Because there is currently no specific predictive score of ischemic stroke risk for PLWH,[44] machine learning was used to find associations between plaque presence in carotid arteries and novel strain imaging features. Among tested traditional cardiovascular risk factors and cIMT measurements, HIV status and strain imaging features were associated with the presence of carotid artery plaques. Subjects in this study had low-to-moderate cardiovascular risk (Framingham score) and had a carotid plaque prevalence of 33% (42% in PLWH) based on wall thickening >1 mm. These plaques were mostly associated with low-grade stenoses (Table 1).

TABLE 2. Clinical Biomarkers Related to PLWH That Were Used in Machine Learning Models for Individuals With and Without Plaques

	With Plaque (43)	No Plaque (59)	Missing N	*P*
Clinical features				
HIV duration (yr)	21 ± 8	18 ± 8	1	**0.017***
ART duration (yr)	16 ± 7	14 ± 7	7	0.24
Any exposure to ART	39 (91%)	56 (95%)	2	0.21
$CD4^+$ T-cell count (cells/mm^3)	582 ± 221	598 ± 286	5	0.99
Ratio $CD4^+/CD8^+$	0.97 ± 0.52	0.91 ± 0.48	5	0.64

*Statistically significant at $P < 0.05$ (unpaired *t*-tests for continuous normally distributed variables, Mann–Whitney tests for continuous nonnormally distributed variables, and χ^2 goodness-of-fit tests for dichotomous variables).

In both RF and LR models, risk factors of age, HIV status, smoking status, pack-years, and sex, that have been previously associated with CVD,[2,45] were found to detect plaque presence (Table 4). In RF classifications, the MaxAS was the fourth most frequent in highest AUC-ROC combinations. This arterial biomechanical parameter was more often associated with plaque presence than the HIV status, sex, and systolic blood pressure. The cIMT was absent in retained features of RF classifications but retained in the LR model. Different features were found in RF and LR combinations, possibly because RF also has the ability to find nonlinear associations.[46] RF provided the best performance in AUC-ROCs.

Associative Predictions in PLWH

In classification tests performed in PLWH, where the sample size was lower, a higher number of features appeared in best combinations. Both classification strategies seemed consistent in AUC-ROC performance. Indeed, the 5 most frequent features to classify the whole database in each model were also present to classify PLWH (excluding the HIV status that was not used in the PLWH database; Tables 4 and 5). More clinical and imaging features predicted the presence of plaque in PLWH compared with all subjects. Because traditional CVD risk prediction models tend to underestimate CVD risk among PLWH,[3] specific models considering US mechanical strain features could add valuable information in the CVD risk management in PLWH.

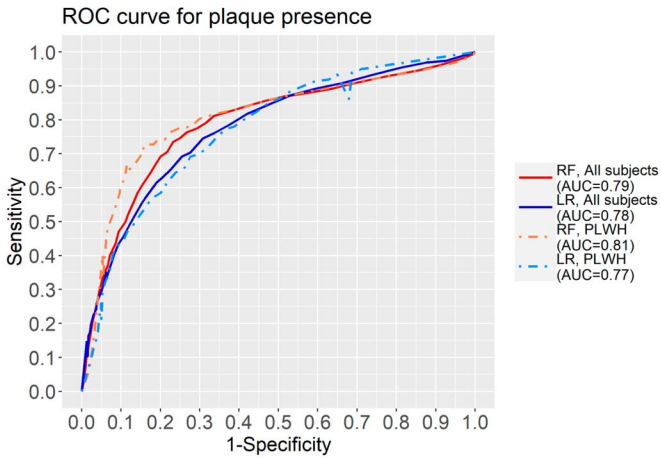

FIGURE 3. ROC curves with highest AUC of all 6 datasets (complete and imputation) for plaque presence detection with RF and LR for all subjects (N = 186) and for PLWH only (N = 102). full color online

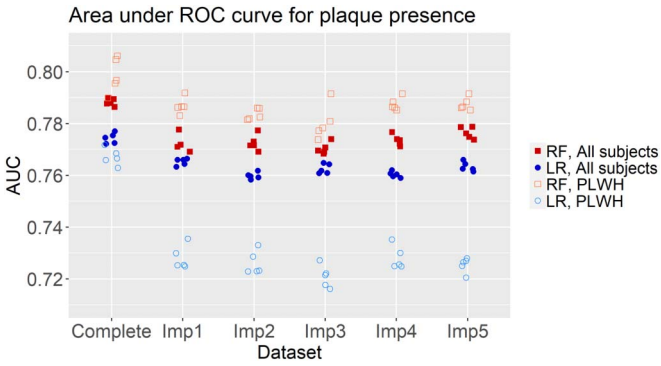

FIGURE 4. AUC-ROCs of the 5 most discriminant combinations of 4 features or less to identify the presence of plaque on complete and imputated datasets (Imp1–Imp5) using RF and LR classification models for all subjects (N = 186) and for PLWH only (N = 102). Points represent the AUC-ROC obtained from classification with different combinations of 4 features or less. Red square and blue circle markers correspond, respectively, to RF and LR classification models; filled and empty symbols correspond, respectively, to all subjects and PLWH only. The symbols with the highest value in each of the database and model (all datasets combined) correspond to ROC curves in Figure 3. full color online

US strain imaging features were measured in plaque-free segments of the CCA wall. Reduced arterial elasticity in PLWH was also observed using other vascular stiffness measurements, such as pulse wave velocity and distensibility measures[24,26]; however, these methods that do not locally assess carotid artery did not always find a difference between PLWH and control subjects.[47] Overall, in RF and LR models, strain parameters contributed to the best performing feature combinations at least as frequently as classic CVD risk factors such as sex, cholesterol, and systolic blood pressure that accounted for 7%–11% of the explained variance in carotid plaque areas.[48] Unsurprisingly, age and smoking status occurred most often in the best AUC-ROC combinations; they have indeed been shown to account for 14%–50% of the explained variance in carotid plaque areas.[48,49] MaxAS and Max|ShS| occurrences were consistent when analyzing either all participants or only PLWH in the RF and LR models. These features provided classification information in machine learning models even though, as shown in Table 1, they were not statistically significantly different between subjects with and without plaques. Their association with carotid plaque presence might be nonlinear or had interactions with other covariables.

Compared with traditional statistical models, machine learning can improve the accuracy of cardiovascular risk

TABLE 3. Mean AUC-ROC Obtained for Combinations of 4 or Less Features, for the Prediction of Carotid Atherosclerotic Plaque Presence

	RF AUC-ROC (n = 30)	LR AUC-ROC (n = 30)	*P*
Classification of all subjects	0.775 ± 0.007	0.764 ± 0.005	0.0001
Classification of PLWH	0.787 ± 0.007	0.733 ± 0.016	2e-16

P-values are for multiple comparisons of classification models with Tukey test for the 5 highest AUC-ROCs (of each dataset).

prediction.[27] Indeed, statistically significantly higher AUC-ROCs were obtained with RF classification compared with LR (Table 3). However, the difference was small when the database included all subjects. As shown in Figure 3, the 5 highest AUC-ROCs for each dataset were close. The improved performance of the RF classification model was more substantial when studying only PLWH, but, in these tests, the database had a lower sample size.

Data imputation enabled the usage of all participants in classification models even when data were missing (Table 1). Because combinations of 4 features or less were used in tested classification models as opposed to using all 27 parameters (or 32 for PLWH), the number of participants with at least one missing value in subsets of 4 or less features were often less than 20% of the whole population. For example, combinations of features from Table 4 had missing data in 3% of participants at most. Because features with high number of missing entries in Table 1, such as the Framingham risk score, were not part of combinations yielding the highest AUC-ROCs, the proportion of participants with complete data in subsets of 4 or less features from the imputed datasets was higher than for the same datasets including all features.

TABLE 4. Occurrence of Features in Combinations Yielding Highest Classification AUCs Using the Database Containing All Subjects (N = 186), for the Prediction of Carotid Atherosclerotic Plaque Presence

RF		LR	
Features	No. of Combinations	Features	No. of Combinations
Age	30	Age	30
Smoking status	27	HIV status	27
Pulse pressure	14	Smoking status	18
MaxAS	11	Pack-years	12
Sex	10	Sex	10
HIV status	8	Max|ShS|	8
Max|ShS|	5	cIMT	8
Systolic BP	5	Diabetes	5
CAT	3	CAS	1
C|ShS|	3	Pulse pressure	1

BP, blood pressure.

TABLE 5. Occurrence of Features in Combinations Yielding Highest Classification AUCs Using the Database Containing Exclusively PLWH (N = 102), for the Prediction of Carotid Atherosclerotic Plaque Presence

RF		LR	
Features	No. of Combinations	Features	No. of Combinations
Age	27	Age	26
Smoking status	24	Smoking status	21
MaxAS	11	Sex	23
Pulse pressure	11	Systolic BP	15
Sex	9	Pack-years	5
ART exposure	9	Max|ShS|	5
HDL-C	6	C|ShS|	5
ART duration	5	Diastolic BP	4
Creatinine	5	ART exposure	3
High waist circumference	2	HIV duration	2
Systolic BP	2	Framingham risk	2
Diabetes	2	RI ICA	2
CAT	1	CAS	1
		CLT	1
		LDL-C	1

HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; RI, resistivity index.

AUC-ROCs computed with the complete dataset were statistically significantly higher than those of imputated datasets. However, differences were small with 95% confidence intervals over the mean difference of 0.011 to 0.018 and 0.020 to 0.030 for RF and LR models, respectively. The difference of means between complete and imputated dataset results was the highest for the LR classification of PLWH exclusively (95% CI of 0.036 to 0.047). The smaller sample size of the PLWH database was thus more problematic for the LR classification where the generalization error differences were greater.

Study Limitations

This study from a prospective cohort was retrospective and based on robust machine learning with bootstrapping to improve consistency of predictions. Associations with plaque presence were found by considering traditional cardiovascular risk factors, conditions specific to PLWH, and novel US strain imaging mechanical biomarkers; but future plaque development or progression was not predicted. Prospective carotid artery follow-up studies would deserve to be conducted using the proposed elastography imaging parameters.

The relatively small sample size particularly in the case of associations found with the PLWH database is also a limitation. Quantitative measures of plaques such as their areas might be more sensitive in predicting CVD risk than the

presence of plaque.[10] With a larger sample size and the subsequent larger subset of subjects with plaques, associations of strain features with carotid plaque area could also be investigated. In the database of PLWH, the history of exposure to different ART regimens was not available; the association of different ART regimens to presence of carotid artery plaque should also be investigated.

Another limitation, related to the use of machine learning, is the lack of mechanistic interpretation to explain the association between the presence of plaque and selected features because linear relations between the model output and inputs are not measured with RF. Nonetheless, carotid artery translation and strain parameters used here were shown to be of value to assess premature atherosclerosis in other studies[16,17,50] and advanced plaques of symptomatic and asymptomatic patients with carotid artery stenoses.[51–53] The new US imaging arsenal proposed here for assessing artery wall biomechanics may be of value to study noninvasively plaque progression and stroke risk in PLWH (and in the general population). Associations made in this study need to be confirmed prospectively.

CONCLUSIONS

In summary, this study revealed that strain imaging features complemented traditional cardiovascular risk factors to classify individuals with carotid artery plaques whether in a group of control subjects and PLWH or in a group exclusively including PLWH. This study may guide establishing prediction models for prospective detection and prevention of early atherosclerosis and plaque progression in PLWH and thus improve the management of CVDs in this population.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Jean-Guy Baril and Dr. Benoit Trottier for referring HIV participants. They are also grateful to Dr. Louise Allard for recruiting PLWH and control subjects and Boris Chayer for managing US acquisitions.

REFERENCES

1. Benjamin EJ, Muntner P, Alonso A, et al. Heart disease and stroke statistics—2019 update: a report from the the American Heart Association. *Circulation*. 2019;139:e56–e528.
2. Shah ASV, Stelzle D, Lee KK, et al. Global burden of atherosclerotic cardiovascular disease in people living with HIV: systematic review and meta-analysis. *Circulation*. 2018;138:1100–1112.
3. Feinstein MJ, Hsue PY, Benjamin LA, et al. Characteristics, prevention, and management of cardiovascular disease in people living with HIV: a scientific statement from the American Heart Association. *Circulation*. 2019;140:e98–e124.
4. Lorenz MW, Markus HS, Bots ML, et al. Prediction of clinical cardiovascular events with carotid intima-media thickness: a systematic review and meta-analysis. *Circulation*. 2007;115:459–467.
5. Peters SA, den Ruijter HM, Bots ML, et al. Improvements in risk stratification for the occurrence of cardiovascular disease by imaging subclinical atherosclerosis: a systematic review. *Heart*. 2012;98:177–184.
6. van den Oord SC, Sijbrands EJ, ten Kate GL, et al. Carotid intima-media thickness for cardiovascular risk assessment: systematic review and meta-analysis. *Atheroscler*. 2013;228:1–11.
7. Den Ruijter HM, Peters SA, Anderson TJ, et al. Common carotid intima-media thickness measurements in cardiovascular risk prediction: a meta-analysis. *JAMA*. 2012;308:796–803.
8. Siedner MJ, Bibangambah P, Kim JH, et al. Treated HIV infection and progression of carotid atherosclerosis in rural Uganda: a prospective observational cohort study. *J Am Heart Assoc*. 2021;10:e019994.
9. Hsue PY, Scherzer R, Hunt PW, et al. Carotid intima-media thickness progression in HIV-infected adults occurs preferentially at the carotid bifurcation and is predicted by inflammation. *J Am Heart Assoc*. 2012;1:jah3, e000422(1)-(12).
10. Naqvi TZ, Lee MS. Carotid intima-media thickness and plaque in cardiovascular risk assessment. *JACC Cardiovasc Imaging*. 2014;7:1025–1038.
11. Hegele RA. The pathogenesis of atherosclerosis. *Clin Chim Acta*. 1996;246:21–38.
12. Mathiesen EB, Johnsen SH, Wilsgaard T, et al. Carotid plaque area and intima-media thickness in prediction of first-ever ischemic stroke: a 10-year follow-up of 6584 men and women: the Tromsø Study. *Stroke*. 2011;42:972–978.
13. Inaba Y, Chen JA, Bergmann SR. Carotid plaque, compared with carotid intima-media thickness, more accurately predicts coronary artery disease events: a meta-analysis. *Atherosclerosis*. 2012;220:128–133.
14. Rundek T, Gardener H, Della-Morte D, et al. The relationship between carotid intima-media thickness and carotid plaque in the Northern Manhattan Study. *Atherosclerosis*. 2015;241:364–370.
15. Hanna DB, Post WS, Deal JA, et al. HIV Infection is associated with progression of subclinical carotid atherosclerosis. *Clin Infect Dis*. 2015;61:640–650.
16. Roy Cardinal MH, Durand M, Chartrand-Lefebvre C, et al. Increased carotid artery wall stiffness and plaque prevalence in HIV infected patients measured with ultrasound elastography. *Eur Radiol*. 2020;30:3178–3187.
17. El Jalbout R, Cloutier G, Roy Cardinal MH, et al. The value of non-invasive vascular elastography (NIVE) in detecting early vascular changes in overweight and obese children. *Eur Radiol*. 2019;29:3854–3861.
18. Bjällmark A, Lind B, Peolsson M, et al. Ultrasonographic strain imaging is superior to conventional non-invasive measures of vascular stiffness in the detection of age-dependent differences in the mechanical properties of the common carotid artery. *Eur J Echocardiogr*. 2010;11:630–636.
19. Redheuil A, Yu WC, Wu CO, et al. Reduced ascending aortic strain and distensibility: earliest manifestations of vascular aging in humans. *Hypertension*. 2010;55:319–326.
20. Park JJ, Park JB, Park JH, et al. Global longitudinal strain to predict mortality in patients with acute heart failure. *J Am Coll Cardiol*. 2018;71:1947–1957.
21. Requena-Ibáñez JA, Santos-Gallego CG, Rodriguez-Cordero A, et al. Mechanistic insights of empagliflozin in nondiabetic patients with HFrEF: from the EMPA-TROPISM study. *JACC Heart Fail*. 2021;9:578–589.
22. Zhu ZQ, Chen LS, Wang H, et al. Carotid stiffness and atherosclerotic risk: non-invasive quantification with ultrafast ultrasound pulse wave velocity. *Eur Radiol*. 2019;29:1507–1517.
23. Mercure E, Destrempes F, Roy Cardinal MH, et al. A local angle compensation method based on kinematics constraints for non-invasive vascular axial strain computations on human carotid arteries. *Comput Med Imaging Graph*. 2014;38:123–136.
24. van Vonderen MG, Smulders YM, Stehouwer CD, et al. Carotid intima-media thickness and arterial stiffness in HIV-infected patients: the role of HIV, antiretroviral therapy, and lipodystrophy. *J Acquir Immune Defic Syndr*. 2009;50:153–161.
25. Baker JV, Duprez D, Rapkin J, et al. Untreated HIV infection and large and small artery elasticity. *J Acquir Immune Defic Syndr*. 2009;52:25–31.
26. Sun D, Wu Y, Yuan Y, et al. Is the atherosclerotic process accentuated under conditions of HIV infection, antiretroviral therapy, and protease inhibitor exposure? Meta-analysis of the markers of arterial structure and function. *Atherosclerosis*. 2015;242:109–116.
27. Weng SF, Reps J, Kai J, et al. Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS One*. 2017;12:e0174944.
28. Durand M, Chartrand-Lefebvre C, Baril JG, et al. The Canadian HIV and aging cohort study–determinants of increased risk of cardio-vascular

diseases in HIV-infected individuals: rationale and study protocol. *BMC Infect Dis*. 2017;17:611.
29. Ford ES, Giles WH, Mokdad AH. The distribution of 10-Year risk for coronary heart disease among US adults: findings from the National Health and Nutrition Examination Survey III. *J Am Coll Cardiol*. 2004;43:1791–1796.
30. Nelson TR, Pretorius DH. The Doppler signal: where does it come from and what does it mean? *Am J Roentgenology*. 1988;151:439–447.
31. Grant EG, Benson CB, Moneta GL, et al. Carotid artery stenosis: gray-scale and Doppler US diagnosis—society of radiologists in ultrasound consensus conference. *Radiology*. 2003;229:340–346.
32. Destrempes F, Meunier J, Giroux MF, et al. Segmentation of plaques in sequences of ultrasonic B-mode images of carotid arteries based on motion estimation and a Bayesian model. *IEEE Trans Biomed Eng*. 2011; 58:2202–2211.
33. Schmitt C, Soulez G, Maurice RL, et al. Noninvasive vascular elastography: toward a complementary characterization tool of atherosclerosis in carotid arteries. *Ultrasound Med Biol*. 2007;33:1841–1858.
34. Breiman L. Random forests. *Machine Learn*. 2001;45:5–32.
35. Roy Cardinal MH, Destrempes F, Soulez G, et al. Assessment of carotid artery plaque components with machine learning classification using homodyned-K parametric maps and elastograms. *IEEE Trans Ultrason Ferroelectr Freq Control*. 2019;66:493–504.
36. Efron B, Tibshirani R. Improvements on cross-validation: the 632+ bootstrap method. *J Am Stat Assoc*. 1997;92:548–560.
37. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med*. 2011;30:377–399.
38. Perkins NJ, Cole SR, Harel O, et al. Principled approaches to missing data in epidemiologic studies. *Am J Epidemiol*. 2018;187:568–575.
39. Van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. *J Stat Softw*. 2011;45:1–67.
40. Liaw A, Wiener M. Classification and regression by randomForest. *R News*. 2002;2:18–22.
41. Svedlund S, Eklund C, Robertsson P, et al. Carotid artery longitudinal displacement predicts 1-year cardiovascular outcome in patients with suspected coronary artery disease. *Arterioscler Thromb Vasc Biol*. 2011; 31:1668–1674.
42. Niiranen TJ, Kalesan B, Mitchell GF, et al. Relative contributions of pulse pressure and arterial stiffness to cardiovascular disease. *Hypertension*. 2019;73:712–717.
43. van Sloten Thomas T, Sedaghat S, Laurent S, et al. Carotid stiffness is associated with incident stroke. *J Am Coll Cardiol*. 2015;66: 2116–2125.
44. Lin TC, Burton BN, Barleben A, et al. Association of HIV infection with age and symptomatic carotid atherosclerotic disease at the time of carotid intervention in the United States. *Vasc Med*. 2018;23: 467–475.
45. 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;117:743–753.
46. Schonlau M, Zou RY. The random forest algorithm for statistical learning. *Stata J*. 2020;20:3–29.
47. Echeverría P, Bonjoch A, Moltó J, et al. Pulse wave velocity as index of arterial stiffness in HIV-infected patients compared with a healthy population. *J Acquir Immune Defic Syndr*. 2014;65:50–56.
48. Herder M, Johnsen SH, Arntzen KA, et al. Risk factors for progression of carotid intima-media thickness and total plaque area. *Stroke*. 2012;43: 1818–1823.
49. Spence JD, Hegele RA. Noninvasive phenotypes of atherosclerosis. *Arterioscler Thromb Vasc Biol*. 2004;24:e188.
50. Çiftel M, Yılmaz O, Kardelen F, et al. Carotid intima media thickness and arterial stiffness in children with acute rheumatic fever. *Pediatr Cardiol*. 2014;35:16–21.
51. Cloutier G, Roy Cardinal MH, Ju Y, et al. Carotid plaque vulnerability assessment using ultrasound elastography and echogenicity analysis. *AJR Am J Roentgenol*. 2018;211:847–855.
52. Varghese T, Meshram NH, Mitchell CC, et al. Lagrangian carotid strain imaging indices normalized to blood pressure for vulnerable plaque. *J Clin Ultrasound*. 2019;47:477–485.
53. Golemati S, Patelaki E, Gastounioti A, et al. Motion synchronisation patterns of the carotid atheromatous plaque from B-mode ultrasound. *Sci Rep*. 2020;10:11221.