# Association of Cardiovascular Mortality and Deep Learning-Funduscopic Atherosclerosis Score derived from Retinal Fundus Images

JOOYOUNG CHANG, AHRYOUNG KO, SANG MIN PARK, SEULGGIE CHOI, KYUWOONG KIM, SUNG MIN KIM, JAE MOON YUN, UK KANG, IL HYUNG SHIN, JOO YOUNG SHIN, TAEHOON KO, JINHO LEE, BAEK-LOK OH, AND KI HO PARK

• PURPOSE: The prediction of atherosclerosis using retinal fundus images and deep learning has not been shown possible. The purpose of this study was to develop a deep learning model which predicted atherosclerosis by using retinal fundus images and to verify its clinical implications by conducting a retrospective cohort analysis.

• DESIGN: Retrospective cohort study.

• METHODS: The database at the Health Promotion Center of Seoul National University Hospital (HPC-SNUH) was used. The deep learning model was trained using 15,408 images to predict carotid artery atherosclerosis, which was named the deep-learning funduscopic atherosclerosis score (DL-FAS). A retrospective cohort was constructed of participants 30-80 years old who had completed elective health examinations at HPC-SNUH. Using DL-FAS as the main exposure, participants were followed for the primary outcome of death due to CVD until Dec. 31, 2017.

• RESULTS: For predicting carotid artery atherosclerosis among subjects, the model achieved an area under receiver operating curve (AUROC) and area under the precision-recall curve (AUPRC), accuracy, sensitivity, specificity, positive and negative predictive values of 0.713, 0.569, 0.583, 0.891, 0.404, 0.465, and 0.865 respectively. The cohort consisted of 32,227 participants, 78 cardiovascular disease (CVD) deaths, and 7.6-year median follow-up visits. Those with DL-FAS greater than 0.66 had an increased risk of CVD deaths compared to those with DL-FAS $<0.33$ (hazard ratio: 8.33; 95% confidence interval [CI], 3.16-24.7). Risk association was significant among intermediate and high Framingham risk score (FRS) subgroups. The DL-FAS improved the concordance by 0.0266 (95% CI, 0.0043-0.0489) over the FRS-only model. The relative integrated discrimination index was 20.45% and net reclassification index was 29.5%.

• CONCLUSIONS: A deep learning model was developed which could predict atherosclerosis from retinal fundus images. The resulting DL-FAS was an independent predictor of CVD deaths when adjusted for FRS and added predictive value over FRS. (Am J Ophthalmol 2020;217:121–130. © 2020 Elsevier Inc. All rights reserved.)

AJO.com Supplemental Material available at AJO.com.

Accepted for publication Mar 17, 2020.

From the Department of Biomedical Sciences (J.C., S.M.P., S.C., K.K., S.M.K.), Seoul National University Graduate School, Seoul, South Korea; Department of Family Medicine (A.K., S.M.P., J.M.Y.), Seoul National University Hospital, Seoul, South Korea; InTheSmart Co., Ltd. (U.K., I.H.S.), Seoul, South Korea; Seoul National University Hospital Biomedical Research Institute (U.K.), Seoul, South Korea Department of Ophthalmology (J.Y.S.), Seoul Metropolitan Government Seoul National University Boramae Medical Center, Seoul, Korea; Office of Hospital Information (T.K.), Seoul National University Hospital, Seoul, South Korea; Department of Ophthalmology (J.L.), Hallym University Chuncheon Sacred Heart Hospital, Chuncheon, Korea; and the Department of Ophthalmology (J.L., B-L.O., K.H.P.), Seoul National University Hospital, Seoul, South Korea

Inquiries to: Sang Min Park, Department of Family Medicine and Biomedical Sciences, College of Medicine, Seoul National University, 101 Daehak-ro, Jongno-gu, Seoul, South Korea; e-mail: smpark.snuh@gmail.com

CARDIOVASCULAR DISEASE (CVD) IS THE MOST COMmon cause of death worldwide and accounts for 32% of all deaths.[1] In 2015, CVD affected more than 400 million patients and caused more than 17 million deaths worldwide.[2] Thus, the assessment of CVD risk and the prevention thereof is of clinical importance.

Retinal fundus imaging contains valuable information regarding vascular health,[3–6] and with the emergence of computer-assisted retinal imaging, various measurements that use this modality have been shown to correlate with the severity of heart failure,[7] with certain stoke subtypes,[8] and with hypertension.[9]

Recent developments in deep learning have revealed that CVD risk factors such as age, sex, and smoking status could be predicted by using retinal fundus images.[10] However, the prediction of atherosclerosis, a subclinical marker of CVD, using retinal fundus imaging and deep learning has not been shown possible yet. Furthermore, the clinical implications of prediction of cardiovascular-related risk factors by retinal fundus imaging have not been addressed in terms of reclassification of patients at risk of CVD and time-to-event analysis. A formal analysis of additional benefits to the risk stratification of patients and time-to-event analysis can provide clinicians evidence to consider retinal fundus imaging for assessing patients of borderline CVD-risk.

This study developed and validated a deep model which used retinal fundus images to predict whether a patient has

0002-9394/$36.00
https://doi.org/10.1016/j.ajo.2020.03.027

atherosclerosis. The predicted value was named the deep-learning funduscopic atherosclerosis score (DL-FAS). Furthermore, analysis determined whether DL-FAS added value to the prediction of cardiovascular death relative to that of the Framingham risk score (FRS), and a retrospective cohort of more than 30,000 patients was conducted for incident cardiovascular deaths.

## SUBJECTS AND METHODS

• **STUDY POPULATION:** Data of participants who had completed medical examinations at the Health Promotion Center of Seoul National University Hospital (HPC-SNUH),[11] between January 2005 and December 2016 and received a retinal fundus image examination, were used for this study. HPC-SNUH offers elective medical health examinations including a survey, physical examinations, laboratory testing, and medical imaging. Participants must subscribe to one of many medical examination packages available at HPC-SNUH and are not offered based on any indication. Retinal fundus imaging is included in all packages; however, carotid artery sonography is only offered for more expensive packages. Participant data were merged with the National Death Certificate database to ascertain the death status and cause of death up to December 31, 2017. Patients were anonymized before analysis was performed, and the need for patient consent was waived by the institutional review board at SNUH (IRB: H-1703-044-837).

The deep learning model was validated using a 2-phase approach. First, during the training phase, participants with retinal fundus examinations plus carotid artery sonography were used to develop the deep model for prediction of atherosclerosis (n = 6,597). These patients were divided into training, validation, and testing sets, where the testing set patients were used to determine whether the model accurately predicted atherosclerosis. The deep model makes predictions for 1 eye image at a time, and the final averaged score of both eyes was named the DL-FAS. Second, during the cohort phase, those with only retinal fundus examinations but without carotid artery sonography (n = 32,227) were used to validate whether the DL-FAS could predict future cardiovascular deaths. The 2-phase validation approach is essential, first, because of the ability to predict atherosclerosis must be verified with cardiovascular mortality studies to draw meaningful clinical implications and contribute to clinical decision making, and second, a single-phase approach to directly predict cardiovascular death would ignore time-to-event analysis and provide no meaningful insights about causation.

• **CAROTID ARTERY ATHEROSCLEROSIS MEASUREMENT:** The carotid artery intima media thickness (CIMT) and existence of carotid artery plaque was used as the proxy marker for atherosclerosis. CIMT was measured using ultrasonography by averaging 3 measurements made 10 mm proximal to the bifurcation.[12] The far-wall intima media thickness (IMT) was identified as the region between the lumen-intima interface and the media-adventitia interface. Those with CIMT measurements of 0.9 mm or more[13] or carotid artery plaque[14] were considered to have atherosclerosis. Although thresholds of higher than 0.9 mm are better correlated to CVD outcomes,[15] a conservative threshold of 0.9 mm was used because atherosclerosis develops gradually over time and retinal features may be present in earlier stages of disease. Carotid artery plaque was defined as a focal increase in thickness of 0.5 mm or 50% of the surrounding CIMT value. Both the left and the right carotid arteries were measured. The carotid artery findings were based on the sonography reports written by board-trained radiologists. For confirmation and data cleaning, all ultrasonographic images and corresponding reports were reviewed by 4 board-trained family medicine physicians.

• **ACQUISITION OF RETINAL FUNDUS IMAGES:** A CR-2 model (Canon, Tokyo, Japan) digital nonmydriatic fundus camera[16] was used to obtain retinal fundus images. Patients' pupils were not dilated, and 1 color fundus picture was taken of each eye. The field of view included both the disc and the macula and was limited to a 45° angle of view.

• **DEEP LEARNING MODEL DEVELOPMENT AND VALIDATION:** Patients with both retinal fundus imaging and carotid artery sonography on the same health examination were used to train and tune the deep model. The training, validation, and testing sets were divided by patient level as outlined in Supplemental Figure 2. A total of 15,408 images were used in the training process of the model. A total of 5,296 patients were used for training, 647 patients were used for tuning, and 654 patients were used for testing (Supplemental Table 1). Each set had a similar proportion of atherosclerosis-positive images (Supplemental Table 1).

Training a deep learning model was carried out in 3 basic steps. First, the retinal fundus image was put into the deep learning model, which produces a prediction score between 0 and 1. Second, the prediction score is compared to the correct label of either atherosclerosis (1) or no atherosclerosis (0), measured by carotid artery sonography. Third, the model's weights are adjusted to minimize its prediction error.

The Xception model (Cornell University, Cornell, New York)[17] was used as the feature extractor followed by two fully connected layers. To speed up training, transfer learning was used with Keras software library (version 2.3.1; Burlingame, California), where the model weights are initialized with pretrained weights learned from ImageNet (Stanford, California).[18] Image pre-processing such as random zoom or horizontal flip were used to prevent over-fitting.

During validation, the deep-learning model produces a prediction ranging between 0 and 1 for each eye, which

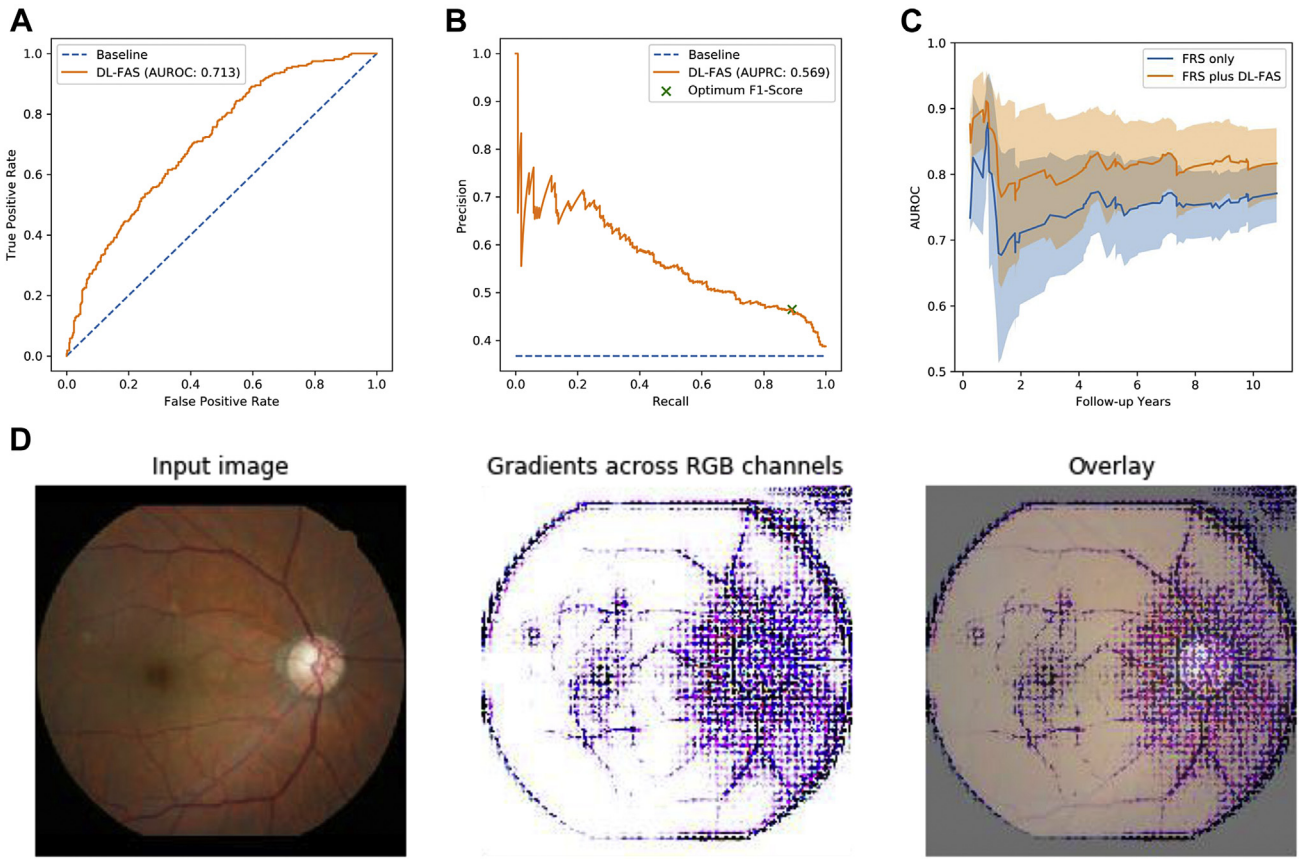


**FIGURE 1. The performance metric curves and saliency map of deep-learning funduscopic atherosclerosis score for prediction of atherosclerosis and Cox model fit statistics beyond Framingham risk score. (A) The ROC for prediction of atherosclerosis using DL-FAS among the test set patients. (B) The PRC for prediction of atherosclerosis using DL-FAS among the test set patients. (C) The time-dependent AUROC of FRS only and FRS plus DL-FAS Cox models among cohort patients calculated by Uno's method based on 50 perturbed samples. (D) Saliency map representation using guided backpropagation for contributions toward positive DL-FAS. AUROC = area under receiver operating curve; DL-FAS = deep-learning funduscopic atherosclerosis score; FRS = Framingham risk score; PRC = precision recall curve; ROC = receiver operating curve.**

results in 2 prediction values for each patient per visit. To prevent duplicating patients, the DL-FAS was calculated for each patient visit and defined as the average of predictions for both eyes. The area under receiver operating curve (AUROC) and area under the precision-recall curve (AUPRC) were calculated. The accuracy, sensitivity, specificity, positive predictive value, and negative predictive value were calculated using a threshold resulting in the maximum F1 score. The F1 score is calculated as [2 × (precision × recall)/(precision + recall)]. We identified saliency features using guided backpropagation (Supplemental Figure 1).[19,20]

• **COHORT CONSTRUCTION:** A retrospective cohort study was performed using patients 30-80 years old, with the index date set to their first medical health examination between 2005 and 2016. Participants who were used for training the deep model were excluded in the cohort study. Those with missing nonsurvey covariates were excluded. These patients were referred to as the cohort set and were mutually exclusive to the training, validation, and test sets.

The primary exposure was the DL-FAS, determined by using retinal fundus images taken on the day of enrollment. Because the DL-FAS ranged between 0 and 1, the model thresholds were chosen so that each score category had an equal range, namely 0-0.33, 0.33-0.67, and 0.67-1.0. For the stratified analysis, score categories were used resulting in an equal number of people per category (ie, terciles).

The primary endpoint of the study was cardiovascular mortality, defined according to the International Classification for Diseases, 10th revision, codes I00 to I99. Patients were followed until December 31, 2017. The secondary endpoint of the study was all-cause mortality.

• **EXPOSURE AND COVARIATES:** For adjustment of conventional risk factors, FRS was calculated using age, sex, high-density lipoprotein cholesterol, systolic blood pressure, and current smoker status.[21] For stratified analysis, high-, intermediate-, and low-risk patients were defined

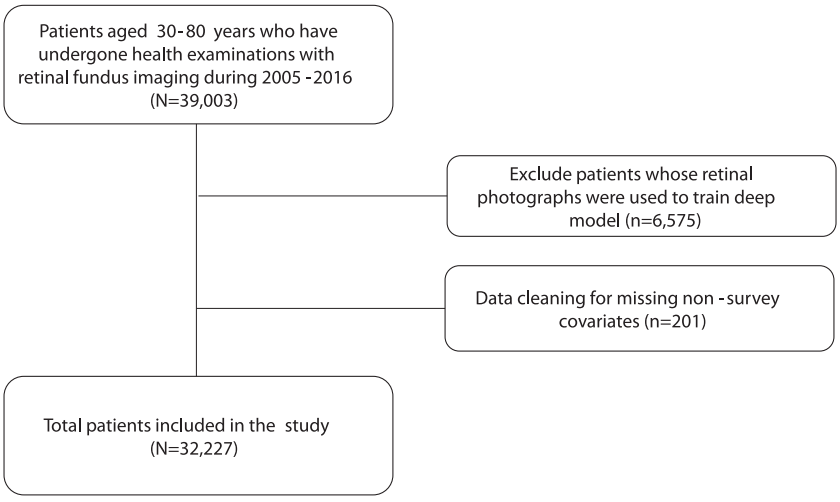


**FIGURE 2. Design of cohort caption.**

as FRS ≥20%, FRS 10%-19%, and FRS <10%, respectively.

For additional adjustment, body mass index, alcohol consumption, exercise frequency, diabetes, hypertension, and dyslipidemia were considered. Diabetes was defined as self-reported diabetes, self-reported diabetic medication history, fasting blood glucose >126 mg/dL, or HbA1c ≥ 6.5%; hypertension as self-reported antihypertensive medication history; and dyslipidemia as low-density lipoprotein >160 mg/dL or self-reported medication history. Missing survey variables were considered a category of their own.

• **STATISTICAL ANALYSIS:** The discrimination, calibration, and reclassification of predicted atherosclerosis scores over FRS risk levels was assessed for predicting cardiovascular deaths among cohort-phase patients. Logistic regression was used to model cardiovascular deaths using, first, FRS risk levels or, second, FRS risk levels plus DL-FAS. The relative integrated discrimination improvement (IDI),[22] category-free net reclassification improvement (NRI),[22] and Hosmer-Lemeshow chi-squared test with 10 groups[23] was performed by comparing the logits of these 2 models.

For a time-to-event analysis, a multivariate Cox regression model was used to estimate the hazard ratios (HR) and 95% confidence intervals (CI). To show whether the DL-FAS significantly added to the CVD mortality prediction of FRS risk scores,[21] the C-statistic estimates were calculated for Cox regression models using FRS risk levels versus FRS risk levels plus DL-FAS. The differences between these concordances were calculated to see whether the addition of the DL-FAS significantly improved the concordance estimate. Concordance estimates and differences between concordance estimates were calculated using Uno's method.[24] All statistical analyses were performed using SAS version 9.0 software (SAS Institute, Cary, North Carolina).

## RESULTS

AMONG 751 TESTING SET PATIENT VISITS IN WHICH 276 PATIENTS tested positive for carotid artery atherosclerosis (prevalence, 0.368), the model was able to predict the sonographically confirmed carotid artery atherosclerosis with an AUROC of 0.713 and an AUPRC of 0.569 (Figure 1A and B). The model achieved an F1 score of 0.611 at the optimal DL-FAS threshold of 0.368. The accuracy, sensitivity, specificity, positive predictive value, and negative predictive value were 0.583, 0.891, 0.404, 0.465, and 0.865, respectively, indicating a predictive model driven by low specificity. Saliency maps positively identified the retinal vessels contributing to positive atherosclerosis prediction as well as pathologic findings including disc rim narrowing, increased cup-to-disc ratio, peripapillary atrophy, and cotton wool spots (Figure 1D and Supplemental Figure 1).

For the validation of the atherosclerosis score to predict cardiovascular mortality, the deep model was used to predict the atherosclerosis of cohort set patients and conducted prediction analysis and a time-to-event analysis. The cohort consisted of 32,227 patients (Figure 2) with varying characteristics with respect to their DL-FAS (Table 1). Of 32,227 patients in the cohort, only 74 patients (0.23%) had only 1 eye imaged, and 99.8% had images for both eyes. The resulting DL-FAS variations were similar, 0.025 and 0.021, respectively. There were strong associations between the DL-FAS and age, sex, and FRS risk level. The median follow-up for the study was 7.6 years, with 78 incident CVD deaths.

The prediction modeling was performed for cardiovascular mortality of a logistic model using only FRS risk levels versus a model using FRS risk levels plus DL-FAS. The IDI analysis showed that the model using FRS risk levels plus DL-FAS had IDI of 0.0007 ($P = 0.008$) and relative

**TABLE 1.** Characteristics of Cohort Participants 30-80 Years Old Whose Retinal Fundus Photographs Were Not Used for Training the Deep Model

| | | DL-Funduscopic Atherosclerosis Score | | |
|---|---|---|---|---|
| | Total | 0-0.33 | 0.33-0.67 | 0.67-1.0 |
| Number of participants | 32,227 | 13,057 | 18,310 | 860 |
| Mean age, y (SD) | 52.6 (10.6) | 44.2 (7.0) | 57.6 (8.3) | 71.7 (6.3) |
| Males, % | 49.5 | 44.8 | 52.6 | 57.0 |
| Body mass index, % | | | | |
| <23 kg/m$^2$ | 41.9 | 46.6 | 38.6 | 40.9 |
| 23-25 kg/m$^2$ | 25.9 | 23.9 | 27.3 | 25.0 |
| ≥25 kg/m$^2$ | 32.2 | 29.5 | 34.1 | 34.1 |
| FRS risk level, %[a] | | | | |
| Low | 58.9 | 82.5 | 44.1 | 15.0 |
| Intermediate | 26.3 | 14.8 | 34.2 | 32.4 |
| High | 14.8 | 2.6 | 21.8 | 52.6 |
| Cigarette Smoking, % | | | | |
| Never | 47.1 | 48.9 | 46.1 | 43.5 |
| Past | 20.3 | 15.8 | 23.1 | 29.0 |
| Current | 16.6 | 19.4 | 14.9 | 9.8 |
| Missing | 16.0 | 16.0 | 15.9 | 17.8 |
| Alcohol Consumption, % | | | | |
| Nondrinker | 41.3 | 41.8 | 40.4 | 51.6 |
| Current drinker | 37.4 | 35.4 | 39.2 | 29.5 |
| Missing | 21.3 | 22.8 | 20.3 | 18.8 |
| Exercise frequency, % | | | | |
| Regular | 31.6 | 35.9 | 28.7 | 27.6 |
| None | 39.2 | 35.0 | 42.4 | 36.4 |
| Missing | 29.2 | 29.1 | 29.0 | 36.0 |
| Diabetes, %[b] | | | | |
| Yes | 19.1 | 12.2 | 23.2 | 37.1 |
| Hypertension, %[c] | | | | |
| Yes | 17.3 | 6.1 | 23.7 | 49.7 |
| Dyslipidemia, %[d] | | | | |
| Yes | 31.8 | 33.0 | 30.8 | 33.8 |

DL = deep learning; FRS = Framingham risk score.

[a]Calculated using age, sex, high-density lipoprotein cholesterol, systolic blood pressure, and smoker.

[b]Self-reported diabetes, self-reported diabetic medication history, concentrations of fasting blood glucose >126 mg/dL or HbA1c ≥6.5%.

[c]Self-reported history of antihypertensive medication.

[d]Low-density lipoprotein >160 mg/dL or self-reported history of the medication.

IDI of 20.45% relative to the model using FRS risk levels alone. Category-free NRI was 29.5% ($P$ = 0.009). Hosmer-Lemeshow chi-squared test results showed a $P$ value of 0.427 for the FRS model and $P$ = 0.609 for the FRS plus DL-FAS model, indicating no evidence for poor fitting.

For the time-to-event analysis, the HR for each DL-FAS score group was calculated for CVD mortality, adjusting for FRS and other baseline risk factors (Table 2). Compared to the lowest atherosclerosis score group, those with scores 0.33-0.67 and 0.67-1.00 had significantly higher risk of CVD mortality (HR, 8.83; 95% CI, 1.41-6.15; and HR, 8.83; 95% CI, 3.16-24.7, respectively). This positive association showed a significant trend ($P$ for trend <.001). DL-FAS was associated with CVD mortality when adjusted for all baseline covariates in their single covariate forms (Supplemental Table 2). Among individual covariates in the multivariate model, only age, systolic blood pressure, and smoking status were significantly associated with CVD mortality (Supplemental Table 2). The DL-FAS was also associated with all-cause mortality (Supplemental Table 3).

Because DL-FAS were highly correlated with FRS, age, and sex, a stratified analysis was performed of subgroups based on age, sex, and FRS levels (Table 3). For subgroups of low, intermediate, and high risk, the highest tercile of DL-FAS had significantly higher risk of CVD mortality than the lowest tercile (HR, 4.76; 95%CI, 1.05-21.63; HR, 3.14; 95% CI, 1.04-9.47; and HR, 5.11; 95% CI,

**TABLE 2.** Risk of CVD Mortality According To Predicted Atherosclerosis Score Using a Multivariate Cox Regression Model

| | DL-Funduscopic Atherosclerosis Score Ranges | | | |
|---|---|---|---|---|
| | 0-0.33 | 0.33-0.67 | 0.67-1 | *P* Trend |
| CVD mortality | | | | |
| Cases | 10 | 59 | 9 | |
| Person-years, $10^3$ | 118.7 | 123.2 | 4.0 | |
| Incidence rate, /$10^3$ person-years | 0.08 | 0.48 | 2.25 | |
| aHR (95% CI)[a] | 1 (ref.) | **2.94 (1.41-6.15)** | **8.83 (3.16-24.7)** | **<.001** |

aHR = adjusted hazards ratio; CI = confidence interval; CVD = cardiovascular disease; DL = deep learning.
Statistically significant values are given in bold.
[a]Adjusted for Framingham risk score 10-year CVD risk (including age, sex, high-density lipoprotein cholesterol, total cholesterol, systolic blood pressure, smoker), body mass index, alcohol consumption, exercise frequency, diabetes, hypertension, and dyslipidemia.

1.94-13.5, respectively) with significant trend (*P* for trend: .038, .032, and .001, respectively).

For subgroups of patients 30-49 years old, DL-FAS was not significantly associated with CVD mortality (HR, 1.69; 95% CI, 0.39-7.25; HR, 0.34; 95% CI, 0.03-4.00, respectively). Among those 50 years old or older, the highest and middle terciles had significantly higher risk of CVD mortality than those in the lowest tercile (HR, 2.66; 95% CI, 1.15-6.17; HR, 5.09; 95% CI, 2.25-11.6, respectively) with significant trend (*P* for trend, <.001). Among both male and female participants, the highest tercile of DL-FAS was associated with a higher risk of CVD mortality than those in the lowest tercile (HR, 3.03; 95% CI, 1.14-8.09; HR, 9.61; 95% CI, 1.95-47.3, respectively).

The concordance estimates for the Cox regression model fitted on FRS only had a concordance of 0.78 (95% CI, 0.73-0.82); the model fitted on FRS plus DL-FAS had a concordance of 0.81 (95% CI, 0.76-0.85). Figure 1C shows the concordance of both models over the follow-up period. The improvement in concordance was 0.0266 with a *P* difference of .020 (Supplemental Table 4).

## DISCUSSION

THE PURPOSE OF THIS STUDY WAS TO DEVELOP A MEASUREment using retinal fundus images which could predict atherosclerosis and stratify the cardiovascular risk of patients over conventional risk factors such as FRS, diabetes, hypertension, dyslipidemia, and health habits. A deep model was trained which predicted atherosclerosis with moderate predictive performance, and the resulting DL-FAS was significantly associated with an increased hazard for CVD mortality among a cohort of otherwise healthy participants after adjustment for FRS. Furthermore, the significant risk association was evident in the stratified analysis of intermediate-risk participants, and the addition of the DL-FAS significantly improved the concordance estimate over the model using FRS only.

This work is novel in several points. First, the prediction of atherosclerosis using retinal fundus images and deep learning has not been done. Second, this work provides validation using not only a cross-sectional analysis but also a longitudinal retrospective cohort for CVD mortality outcomes. Using a cohort to verify newly developed deep-learning measurements is uncommon. Third, this work shows that the DL-FAS is an independent predictor of CVD mortality compared to conventional risk estimates such as the FRS. Some studies used deep learning to predict cardiovascular risk factors such as age, sex, and blood pressure but did not show the added diagnostic value of their models compared with conventional risk-estimate models.[10]

• **PREDICTIVE VALUE OF DL-FAS:** Our results show not only that retinal fundus images may be used to predict the atherosclerosis of the carotid arteries but also that this prediction may add to conventional risk-stratification scores such as the Framingham risk score for longitudinal outcomes of cardiovascular mortality. Previous meta-analyses and cohort studies analyzed the added predictive value of C-reactive protein,[25] CIMT,[26] computed tomography coronary artery calcium score,[27,28] and the ankle-brachial index,[29,30] beyond FRS. Improvement of C-statistics above conventional risk factors and Framingham risk score, determined by the difference between concordance using conventional risk factors and concordance using conventional risk factors plus the additional measurement, were insignificant for CIMT (0.00-0.002),[26,27] ankle-brachial index (0.00-0.002),[27,30] and C-reactive protein (0.00),[27] while improvements were mostly significant for computed tomography coronary artery calcium score (0.02-0.13).[27,28] In our work, the improved concordance estimate of DL-FAS for cardiovascular mortality was small but significant (HR, 0.027; 95% CI, 0.004-0.049), and relative IDI and NRI measurements were also significant at 20.45% and 29.5%, respectively.

One important result of this study is that the DL-FAS showed significant association even among intermediate-

**TABLE 3.** Risk of CVD Mortality by Predicted Atherosclerosis Score Stratified by FRS Risk and Age

| | DL-Funduscopic Atherosclerosis Score | | | |
|---|---|---|---|---|
| | Q1 | Q2 | Q3 | *P* trend |
| FRS low risk | | | | |
| Quantile limits | 0.05, 0.25 | 0.25, 0.38 | 0.38, 0.8 | |
| Cases | 3 | 5 | 7 | |
| Person-years, $10^3$ | 61.1 | 47.9 | 35.3 | |
| Incidence rate, /$10^3$ person-years | 0.05 | 0.10 | 0.20 | |
| aHR (95% CI)[a] | 1 (ref.) | 2.25 (0.51-9.96) | **4.76 (1.05-21.6)** | **.038** |
| FRS intermediate risk | | | | |
| Quantile limits | 0.12, 0.38 | 0.38, 0.5 | 0.5, 0.81 | |
| Cases | 5 | 7 | 12 | |
| Person-years, $10^3$ | 27.1 | 21.4 | 16.3 | |
| Incidence rate, /$10^3$ PY | 0.18 | 0.33 | 0.74 | |
| aHR (95% CI)[a] | 1 (ref.) | 1.50 (0.46-4.84) | **3.14 (1.04-9.47)** | **.032** |
| FRS high risk | | | | |
| Quantile limits | 0.17-0.46 | 0.46-0.57 | 0.57-0.81 | |
| Cases | 6 | 15 | 18 | |
| Person-years, $10^3$ | 15.0 | 12.4 | 9.4 | |
| Incidence rate, /$10^3$ person-years | 0.40 | 1.21 | 1.91 | |
| aHR (95% CI)[a] | 1 (ref.) | **3.09 (1.18-8.06)** | **5.11 (1.94-13.5)** | **.001** |
| Age 30-49 y | | | | |
| Quantile limits | 0.05, 0.22 | 0.22, 0.3 | 0.3, 0.73 | |
| Cases | 3 | 6 | 1 | |
| Person-years, $10^3$ | 42.1 | 36.1 | 23.2 | |
| Incidence rate, /$10^3$ person-years | 0.07 | 0.17 | 0.04 | |
| aHR (95% CI)[a] | 1 (ref.) | 1.69 (0.39-7.25) | 0.34 (0.03-4.00) | .5342 |
| Age, 50 years or more | | | | |
| Quantile limits | 0.13, 0.4 | 0.4, 0.51 | 0.51, 0.81 | |
| Cases, n | 8 | 19 | 35 | |
| Person-years, $10^3$ | 60.9 | 45.7 | 37.2 | |
| Incidence rate, /$10^3$ person-years | 0.13 | 0.42 | 0.94 | |
| aHR (95% CI) [a] | 1 (ref.) | **2.66 (1.15-6.17)** | **5.09 (2.25-11.6)** | **< .001** |
| Males | | | | |
| Quantile limits | 0.07, 0.32 | 0.32, 0.46 | 0.46, 0.81 | |
| Cases | 7 | 10 | 37 | |
| Person-years, $10^3$ | 48.7 | 40.5 | 32.9 | |
| Incidence rate, /$10^3$ person-years | 0.14 | 0.25 | 1.12 | |
| aHR (95% CI) [a] | 1 (ref.) | 0.99 (0.36-2.75) | **3.03 (1.14-8.09)** | **.003** |
| Females | | | | |
| Quantile limits | 0.05, 0.29 | 0.29, 0.44 | 0.44, 0.81 | |
| Cases | 2 | 5 | 17 | |
| Person-years, $10^3$ | 51.3 | 40.7 | 31.7 | |
| Incidence rate, /$10^3$ person-years | 0.04 | 0.12 | 0.54 | |
| aHR (95% CI)[a] | 1 (ref.) | 2.63 (0.49-14.0) | **9.61 (1.95-47.3)** | **.001** |

aHR = adjusted hazards ratio; CI = confidence interval; CVD = cardiovascular disease; DL = deep learning; FRS = Framingham risk score; Q = quantile.

Statistically significant values are given in bold.

[a]Adjusted for Framingham risk score 10-year CVD risk (including age, sex, high-density lipoprotein cholesterol, total cholesterol, systolic blood pressure, and smoker), body mass index, alcohol consumption, exercise frequency, diabetes, hypertension, and dyslipidemia.

risk patients. The risk stratification of such patients has been an area of considerable research, as conventional risk estimates may underestimate risk of patients with evidence of asymptomatic preclinical atherosclerosis.[31] The American College of Cardiology Foundation/American Heart Association[32] and European Society of Cardiology[33] guidelines recommend further investigation for intermediate- and moderate-risk patients to search for target organ

damage including coronary artery calcium score, ankle-brachial index, and atherosclerotic plaque detection by carotid artery scanning. Conducting a retinal fundus image examination for all such patients may be premature at this this stage, but the present results show the added benefit of using the DL-FAS among intermediate risk patients. Because retinal fundus imaging is noninvasive compared to blood testing, DL-FAS may find new possibilities for use in stratifying intermediate risk patients.

• **MECHANISM AND FEATURES:** The mechanism by which the deep-learning model predicts atherosclerosis is not clear, but these authors believe that the deep model recognized features of the retinal microvasculature to predict atherosclerosis. Many studies have verified significant relationships between findings of retinal fundus images and CVD. Retinal vascular pathology is associated with cerebral small-vessel disease.[34] Arteriovenous nicking is associated with increased odds of cardiovascular mortality.[35] Retinal vascular caliber is associated with greater risk of death due to coronary heart disease.[36] Retinal microvascular hemorrhage, microaneurysms, soft exudates, and arteriovenous nicking are associated with an increased risk of stroke.[37] Tortuosity is associated with death due to ischemic heart disease.[38] These studies showed that the retinal fundus images hold valuable information regarding cardiovascular health, which the authors presume this study was able to extract.

The saliency maps using guided backpropagation showed that retinal vessels were making positive contributions to the atherosclerosis prediction. Previous works have used retinal fundus images to predict anemia[20] or other cardiovascular risk factors,[10] using deep learning and have provided similar saliency maps which identify vascular anatomy. The present work suggests that certain changes in the retinal vasculature may be biomarkers for atherosclerosis. Furthermore, some saliency maps identified pathologic findings related to glaucoma, such as increased cup-to-disc ratio, disc rim narrowing, peripapillary atrophy, and cotton wool spots, which suggests the possible association between glaucomatous and atherosclerotic changes through overlapping mechanisms of hypertension.

• **LIMITATIONS AND STRENGTHS:** This study must be interpreted considering the following limitations. First, the database was constructed using single-center data and consisted entirely of Korean nationals. The generalizability of these results may be limited because CVD risk is dependent on ethnicity.[39] Further research is warranted of the development of DL-FAS using multiethnic populations. Second, although the threshold-independent metric of DL-FAS for prediction of atherosclerosis was greater than that in the baseline, the DL-FAS had low accuracy and specificity at the designated threshold. At the current threshold, the DL-FAS may not be fit for specific detection of atherosclerosis. Third, the DL-FAS did not significantly increase the risk of cardiovascular death among those under 50 years old. Although this was most likely due to the small number of cardiovascular deaths, the application of the DL-FAS among younger age groups may not be appropriate. Fourth, incident CVD, myocardial infarction, or stroke information was not available for the current study, and the primary endpoint of the study was death due to CVDs. Hence, such low event rates limit the interpretability of event classification. These results may not accurately estimate risk of incident sudden CVDs such as stroke, myocardial infarction, or heart failure. Finally, the authors we not have access to medical charts to verify the CVD mortality outcomes in the death certificate database. Although by law only medical professionals can issue death certificates, the lack of a robust chart review may cause misclassification bias.

In this work, a classifier for atherosclerosis prediction was trained and then used for a time-to-event Cox analysis, but several alternative methods may improve the results. The use of Cox directly as a loss function to train the deep model or the incorporation of covariate factors as auxiliary inputs may further improve the predictive value of the deep model. Although these methods provide an opportunity for improved performance, the purpose of the study was not to produce the best model possible but, first, to predict and screen atherosclerosis by using retinal fundus images by deep learning and, second, to validate its clinical implications using analyses of risk stratification, association with cardiovascular mortality, and improvement beyond FRS. This purpose was achievable with a deep model using retinal fundus images alone. However, technical optimizations for improvements in predictive power merits future work.

The strengths of the present work include the cross-sectional and longitudinal cohort study design, the analysis of a novel measurement over conventional risk factors, and the comparison of concordance estimates of the cohort analysis. The DL-FAS was not only a good predictor of atherosclerosis at one point in time but also associated with incident cardiovascular deaths in the cohort analysis of 245,900 person-years. The adjustment for multiple conventional risk factors, namely, FRS, body mass index, diabetes, hypertension, dyslipidemia, and other health habits, indicates the DL-FAS is an independent predictor, even within the intermediate risk patient strata. Furthermore, the comparison of concordance estimates between FRS-only and FRS-plus DL-FAS shows the added benefit of adjusting for atherosclerosis score derived from retinal fundus images.

## CONCLUSIONS

THIS SINGLE-CENTER RETROSPECTIVE COHORT STUDY OF Koreans showed that a deep-learning model could be

used to predict atherosclerosis, using retinal fundus images. Furthermore, the study showed that the resulting DL-FAS was associated with CVD mortality after adjustment of conventional risk factors including FRS and increased the C-statistic of the Cox model beyond FRS.

ALL AUTHORS HAVE COMPLETED AND SUBMITTED THE ICMJE FORM FOR DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST and none were reported.

Funding/Support: This work was supported by the InTheSmart, Co. Ltd., grant 0620180650 through the Seoul National University Hospital Research Fund, and by Seoul National University Hospital Research Fund grant 0320190160.

Financial Disclosures: U.K. and I.H.S. are employees of InTheSmart, Co. Ltd. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose.

## REFERENCES

1. Wang H, Naghavi M, Allen C, et al. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 2016;388:1459–1544.
2. Roth GA, Johnson C, Abajobir A, et al. Global, regional, and national burden of cardiovascular diseases for 10 causes, 1990 to 2015. *J Am Coll Cardiol* 2017;70:1–25.
3. Bhargava M, Ikram M, Wong T. How does hypertension affect your eyes? *J Hum Hypertens* 2012;26:71.
4. Flammer J, Konieczka K, Bruno RM, Virdis A, Flammer AJ, Taddei S. The eye and the heart. *Eur Heart J* 2013;34:1270–1278.
5. Grassi G, Buzzi S, Dell'Oro R, et al. Structural alterations of the retinal microcirculation in the "prehypertensive" high-normal blood pressure state. *Curr Pharm Des* 2013;19:2375–2381.
6. Wong TY. Is retinal photography useful in the measurement of stroke risk? *Lancet Neurol* 2004;3:179–183.
7. Nagele MP, Barthelmes J, Ludovici V, et al. Retinal microvascular dysfunction in heart failure. *Eur Heart J* 2018;39:47–56.
8. Doubal FN, MacGillivray TJ, Hokke PE, Dhillon B, Dennis MS, Wardlaw JM. Differences in retinal vessels support a distinct vasculopathy causing lacunar stroke. *Neurology* 2009;72:1773.
9. Cheung CY, Ikram MK, Sabanayagam C, Wong TY. Retinal microvasculature as a model to study the manifestations of hypertension. *Hypertension* 2012;60:1094–1103.
10. Poplin R, Varadarajan AV, Blumer K, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat Biomed Eng* 2018;2:158–164.
11. Yoon C, Goh E, Park SM, Cho B. Effects of smoking cessation and weight gain on cardiovascular disease risk factors in Asian male population. *Atherosclerosis* 2010;208:275–279.
12. 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.
13. Mancia G, De Backer G, Dominiczak A, et al. 2007 Guidelines for the management of arterial hypertension: the Task Force for the Management of Arterial Hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology (ESC). *J Hypertens* 2007;25:1105–1187.
14. Members ATF, Mancia G, Fagard R, et al. 2013 ESH/ESC guidelines for the management of arterial hypertension: the task force for the management of arterial hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology (ESC). *Eur Heart J* 2013;34:2159–2219.
15. O'Leary DH, Polak JF, Kronmal RA, Manolio TA, Burke GL, Wolfson SK Jr. Carotid-artery intima and media thickness as a risk factor for myocardial infarction and stroke in older adults. Cardiovascular Health Study Collaborative Research Group. *N Engl J Med* 1999;340:14–22.
16. Panwar N, Huang P, Lee J, et al. Fundus photography in the 21st century—a review of recent technological advances and their implications for worldwide healthcare. *Telemed J E Health* 2016;22:198–208.
17. Chollet F. Xception: Deep Learning With Depthwise Separable Convolutions. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2017;:1251–1258.
18. Russakovsky O, Deng J, Su H, et al. Imagenet large scale visual recognition challenge. *Int J Comput Vis* 2015;115:211–252.
19.. Springenberg JT, Dosovitskiy A, Brox T, Riedmiller M. Striving for Simplicity: The All Convolutional Net 2014. arXiv e-prints. arXiv:1412.6806.https://ui.adsabs.harvard.edu/abs/2014arXiv1412.6806S. Accessed December 1, 2014.
20. Mitani A, Huang A, Venugopalan S, et al. Detection of anaemia from retinal fundus images via deep learning. *Nat Biomed Eng* 2020;4:18–27.
21. 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.
22. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med* 2008;27:157–172. discussion 207–12.
23. Hosmer DW, Hosmer T, Le Cessie S, Lemeshow s. A comparison of goodness-of-fit tests for the logistic regression model. *Stat Med* 1997;16:965–980.
24. Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Stat Med* 2011;30:1105–1117.
25. Kaptoge S, Di Angelantonio E, Lowe G, et al. C-reactive protein concentration and risk of coronary heart disease, stroke, and mortality: an individual participant meta-analysis. *Lancet* 2010;375(9709):132–140.
26. 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.
27. Ferket BS, van Kempen BJH, Hunink MGM, et al. Predictive value of updating Framingham risk scores with novel risk markers in the U.S. general population. *PLoS One* 2014;9: e88312.
28. Peters SA, den Ruijter HM, Bots ML, Moons KG. Improvements in risk stratification for the occurrence of cardiovascular disease by imaging subclinical atherosclerosis: a systematic review. *Heart* 2012;98:177–184.
29. Fowkes FG, Murray GD, Butcher I, et al. Ankle brachial index combined with Framingham Risk Score to predict cardiovascular events and mortality: a meta-analysis. *JAMA* 2008; 300:197–208.
30. Murphy TP, Dhangana R, Pencina MJ, D'Agostino RB Sr. Ankle-brachial index and cardiovascular risk prediction: an analysis of 11,594 individuals with 10-year follow-up. *Atherosclerosis* 2012;220:160–167.
31. Reiner Z, Catapano AL, De Backer G, et al. ESC/EAS Guidelines for the management of dyslipidaemias: the Task Force for the management of dyslipidaemias of the European Society of Cardiology (ESC) and the European Atherosclerosis Society (EAS). *Eur Heart J* 2011;32:1769–1818.
32. 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–e103.
33. Perk J, De Backer G, Gohlke H, et al. European guidelines on cardiovascular disease prevention in clinical practice (version 2012). The Fifth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of nine societies and by invited experts). *Eur Heart J* 2012;33:1635–1701.
34. Kwa VI, van der Sande JJ, Stam J, Tijmes N, Vrooland JL. Retinal arterial changes correlate with cerebral small-vessel disease. *Neurology* 2002;59:1536–1540.
35. Wong TY, Klein R, Nieto FJ, et al. Retinal microvascular abnormalities and 10-year cardiovascular mortality: a population-based case-control study. *Ophthalmology* 2003; 110:933–940.
36. Rochtchina E, Burlutsky G, Liew G, et al. Retinal vessel diameter and cardiovascular mortality: pooled data analysis from two older populations. *Eur Heart J* 2007;28: 1984–1992.
37. Wong TY, Klein R, Couper DJ, et al. Retinal microvascular abnormalities and incident stroke: the Atherosclerosis Risk in Communities Study. *Lancet* 2001;358:1134–1140.
38. Witt N, Wong TY, Hughes AD, et al. Abnormalities of retinal microvascular structure and risk of mortality from ischemic heart disease and stroke. *Hypertension* 2006;47:975–981.
39. DeFilippis AP, Young R, Carrubba CJ, et al. An analysis of calibration and discrimination among multiple cardiovascular risk scores in a modern multiethnic cohort. *Ann Intern Med* 2015;162(4):266–275.