tvst

Special Issue

Predicting High Coronary Artery Calcium Score From Retinal Fundus Images With Deep Learning Algorithms

Jaemin Son[1,*], Joo Young Shin[2,*], Eun Ju Chun[3], Kyu-Hwan Jung[1], Kyu Hyung Park[4], and Sang Jun Park[4]

[1] VUNO Inc., Seoul, Korea
[2] Department of Ophthalmology, Seoul National University College of Medicine, Seoul Metropolitan Government Seoul National University Boramae Medical Center, Seoul, Korea
[3] Department of Radiology, Seoul National University Bundang Hospital, Seongnam, Korea
[4] Department of Ophthalmology, Seoul National University College of Medicine, Seoul National University Bundang Hospital, Seongnam, Korea

Correspondence: Sang Jun Park, Department of Ophthalmology, Seoul National University Bundang Hospital, 82, Gumi-ro, 173 Beon-gil, Bundang-gu, Seongnam-si, Gyeonggi-do 13620, Korea. e-mail: sangjunpark@snu.ac.kr
Kyu-Hwan Jung, VUNO Inc., 6F, 507, Gangnam-daero, Seocho-gu, Seoul, Korea. e-mail: khwan.jung@vuno.co

Received: August 26, 2019
Accepted: March 6, 2020
Published: May 11, 2020

Keywords: coronary artery calcium score; retinal fundus images; deep learning

Citation: Son J, Shin JY, Chun EJ, Jung K-H, Park KH, Park SJ. Predicting high coronary artery calcium score from retinal fundus images with deep learning algorithms. Trans Vis Sci Tech. 2020;9(2):28, https://doi.org/10.1167/tvst.9.2.28

Purpose: To evaluate high accumulation of coronary artery calcium (CAC) from retinal fundus images with deep learning technologies as an inexpensive and radiation-free screening method.

Methods: Individuals who underwent bilateral retinal fundus imaging and CAC score (CACS) evaluation from coronary computed tomography scans on the same day were identified. With this database, performances of deep learning algorithms (inception-v3) to distinguish high CACS from CACS of 0 were evaluated at various thresholds for high CACS. Vessel-inpainted and fovea-inpainted images were also used as input to investigate areas of interest in determining CACS.

Results: A total of 44,184 images from 20,130 individuals were included. A deep learning algorithm for discrimination of no CAC from CACS >100 achieved area under receiver operating curve (AUROC) of 82.3% (79.5%–85.0%) and 83.2% (80.2%–86.3%) using unilateral and bilateral fundus images, respectively, under a 5-fold cross validation setting. AUROC increased as the criterion for high CACS was increased, showing a plateau at 100 and losing significant improvement thereafter. AUROC decreased when fovea was inpainted and decreased further when vessels were inpainted, whereas AUROC increased when bilateral images were used as input.

Conclusions: Visual patterns of retinal fundus images in subjects with CACS > 100 could be recognized by deep learning algorithms compared with those with no CAC. Exploiting bilateral images improves discrimination performance, and ablation studies removing retinal vasculature or fovea suggest that recognizable patterns reside mainly in these areas.

Translational Relevance: Retinal fundus images can be used by deep learning algorithms for prediction of high CACS.

Introduction

The retina is frequently examined for screening complications of systemic cardiovascular diseases (CVDs, e.g., hypertension and diabetes mellitus) owing to its uniqueness that blood vessels are directly visible without any radiation or invasive interventions. With high-resolution retinal fundus images, geometric factors of retinal vasculature such as diameter, curvature tortuosity, and branching angles can be analyzed to help assess coronary artery diseases,[1] CVDs,[2,3] or diabetes mellitus.[4] The diameters of the retinal vessels have also been shown to be associated with cognitive ability[5,6] and dementia[7] in an older population.

Coronary artery calcium score (CACS) was designed to quantify the degree of coronary artery calcium (CAC) algorithmically from cardiac computed

tomography (CT),[8] which requires cost and radiation. Among numerous methods for computing CACS, three methods are most popular—Agatston score (weighted sum of plaque attenuation),[9] calcium volume score, and relative calcium mass score.[10] CACS is often used to help stratify potential cardiovascular risks combined with the Framingham risk score[11] or other methods[12–15] to provide targeted guidelines for effective screening and treatment of asymptomatic patients. CACS has been used in clinical practice as a fixed cutoff value or adjusted by other risk factors such as age, sex, and ethnicity.[16,17] Factors such as age, sex, blood pressure, and smoking history are known to be associated with CVDs,[18,19] and a recent study[20] has shown that a deep learning algorithm[21] could predict these risk factors, as well as the future occurrence of CVDs from retinal fundus images.

In this study, the ability of a deep learning algorithm, which has shown tremendous potential in analysis of medical imageries,[22,23] to discriminate high CACS with retinal fundus images was assessed to evaluate the feasibility as an inexpensive and radiation-free screening method for patients at risk of CVD. We also evaluated which areas in the retinal fundus images were of most interest in the prediction of CACS, while also experimenting with whether considering both eyes improves the performance of this algorithm.

Methods

Study Population

Individuals who underwent nonmydriatic macula-centered retinal fundus photography for both eyes and CACS assessment on the same day for screening purposes on their demand at the Health Screening Center at Seoul National University Bundang Hospital between June 1, 2003, and June 3, 2016, were identified.[24,25] This study was approved by the institutional review board of Seoul National University Bundang Hospital (IRB No: B-1508-312-107) and conducted in accordance with the tenets of the Declaration of Helsinki. Any subjects with retinal, glaucomatous, or disc abnormalities found in their retinal fundus images were excluded. To assess CACS, individuals underwent cardiac CT on a 64-slice multi-detector CT (Brilliance 64; Philips Medical Systems, Best, The Netherlands) with 64 × 0.625 mm collimation and 420 ms rotation time. The scanning protocol for CACS included 120 kV tube voltage, 200 mAs tube current, and 2.5 mm scan thickness. CACS was calculated according to the Agatston score using a threshold of 130 Hounsfield units with a commercial workstation (Brilliance Workspace; Phillips Medical Systems). Retinal fundus images were acquired using two fundus cameras (Canon CR6-45NM, Kowa nonmyd7) with varying sizes of 1536 × 1024 to 2464 × 1632. The presence of hypertension along with demographic information, including age and sex, was collected by review of medical records.

Experimental Settings

Because retinal fundus images do not provide direct cues about the conditions of coronary arteries, the exact estimation of the CACS, which is a volumetric measure computed from stable CT scans, is prohibitive. CACS is usually used in clinical practice as a fixed cutoff value or in combination with other risk factors to stratify potential cardiovascular risk. Therefore it would be clinically more meaningful to have a deep learning model that discriminates individuals with high CACS from no CACS, rather than estimating the actual value of CACS from fundus images, which would be more challenging and inefficient. For this reason, the deep learning algorithm aims to classify retinal fundus images of patients with high CACS from those with no CAC. To the best of our knowledge, no strict standard for abnormal CACS exists, and the threshold for high CACS may vary depending on practitioners. To find the threshold value of abnormal CACS above which a deep learning algorithm can recognize patterns in retinal fundus images apart from those with no CAC, we set various cutoff values for high CACS, and the discriminative performances were measured for each threshold value. The threshold was set at 0, 100, 200, 300, and 400, and a neural network was trained to classify whether a given retinal fundus image belongs to a subject with no CAC or with high CACS above the specified threshold. Retinal fundus images with CACS that did not belong to either category were excluded from training, as well as from the validation dataset.

Along with original retinal fundus images, images with obscured retinal vessels or fovea were used as inputs to evaluate the attribution to these retinal landmarks in estimating CACS. Vessels were segmented with a previously-reported deep learning algorithm,[26] and pixels with a probability of being vessels above 0.2 were erased and inpainted with the biharmonic algorithm. Similarly, another previously-reported deep learning algorithm[27] was used to localize the coordinates of the fovea and the optic disc, and then a circular region centered at the fovea with a radius of 0.3 times the distance between the fovea and the optic disc was erased. For inpainting, *restoration* module in Python *scikit-image* package (v.0.13.0) was used. Examples of inpainted retinal fundus images

used in the experiments are shown in Supplementary Figure S1.

For evaluating whether including images from both eyes increases the performance, neural networks were trained to take images of both eyes of each subject for all threshold values of high CACS, and the performance was compared with when only a single eye image was used.

The performance of the deep learning algorithm using retinal fundus images was also compared with that of logistic regression models that additionally used other previously known cardiovascular risk factors, namely, age, gender, and hypertension. Information about sex and the presence of hypertension were encoded with binary values (0 or 1). Architectural details of these models are given in Supplementary Table S2. *linear_model* module in Python (http://www.python.org) *scikit-learn* package (https://scikit-learn.org) was used to implement the logistic regression models.

Development and Evaluation of Deep Learning Algorithm

The black background of the retinal fundus images was cropped and resized to 512 × 512, which is a resolution that has demonstrated reliable performance to predict abnormalities in retinal fundus images.[28] The dataset was split into five folds for cross-validation without any duplicate of subjects. Each fold was used to measure the performance of the model that had been trained with the other four folds. Models with the best AUROC on each validation fold were chosen for further evaluations. As the neural network architecture, ImageNet-pretrained inception-v3,[29] which has been widely used in the literature, including the previous related work,[20] was used. Performance did not differ significantly with use of various other networks, which is presented in detail in Supplementary Table S3. In the original ImageNet-pretrained inception-v3 architecture, the 1000-dimensional fully connected layer was modified to have a 1-dimensional output that was later fed into a sigmoid function. For the network that takes bilateral retinal fundus images, the same network was used for each eye, and the last convolutional feature maps were concatenated, and a fully connected layer was added before the final output layer. Details of the architecture are shown in Supplementary Table S1. The network was trained up to 50 epochs with class-balanced mini-batches (the ratio of no CAC to high CACS was roughly 1:1), with the size of 32 using the binary cross entropy loss. The learning rate was set to 1e-3 for the initial 30 epochs, 1e-4 for the following 10 epochs, and 1e-5 for the last 10 epochs. Training fundus images were augmented randomly by flip, rotation, and resize.

The performance was measured with the area under the receiver operating curve (AUROC) with a 95% confidence interval computed by averaging the results for the validation sets in the five-fold cross-validation. Also, the sensitivity and specificity were computed with a threshold of 0.5 for the output of the algorithm. Metrics module in Python (http://www.python.org) *scikit-learn* package (https://scikit-learn.org) was used to compute AUROC, and a confusion matrix was used to compute the sensitivity and specificity. For assessing the statistical significance of difference in AUROCs in different experimental settings, the Wilcoxon signed-rank test was used by the stats module in *scipy* package (https://www.statsmodels.org) in Python language.

We also plotted heatmaps that visualized where the deep learning algorithms attended to in the input fundus image. For the heatmap analyses, a 16 × 16 activation map in the neural network was superimposed onto the input image (512 × 512) after resizing. Values in the activation map were normalized to [0, 1] before resizing and smoothed with Gaussian filter with a radius of 32.

Results

A total of 44,184 images of 20,130 eligible individuals (8,517 women [42.3%]) who underwent examinations for bilateral retinal fundus images and CACS on the same day were collected. Mean age ± standard deviation was 48.9 ± 8.9 years (19–84 years). Of these, 2151 individuals (10.7%) had hypertension. The distribution of age, sex, and hypertension according to CACS in the study population is shown in Table 1.

Figure 1 shows the trends of AUROCs with varying CACS thresholds and different types of inputs. Receiver operating curves of various threshold values for abnormal CACS and various input types are shown in Figure 2. Empirical mean and standard variation for each experimental case are summarized in Table 2. For binary classification of no CAC (CACS = 0) versus positive CACS (CACS > 0), neural networks showed moderate discriminative performance, achieving mean AUROC about 75%. AUROC increased significantly by about 7 percentage points when neural networks were trained to discriminate no CAC versus CACS>100 for all input types ($P < 0.05$). No significant improvement was observed in AUROC when the threshold for high CACS was increased to 200, 300, and 400 for all input types. Compared with when a

Table 1. Demographics of the Dataset

	Range of CACS					
Age group	CACS = 0	0 <CACS ≤ 100	100 < CACS ≤ 200	200 < CACS ≤ 300	300 < CACS ≤ 400	400 < CACS
10 ≤ x < 20	3	0	0	0	0	0
20 ≤ x < 30	127	0	0	0	0	0
30 ≤ x < 40	2,891	106	1	3	2	1
40 ≤ x < 50	6,985	764	76	13	14	18
50 ≤ x < 60	4,856	1,315	191	83	39	87
60 ≤ x < 70	1,261	707	159	65	40	89
70 ≤ x < 80	74	79	27	14	6	24
80 ≤ x < 90	3	5	1	1	0	0
Total subjects	16,200	2,976	455	179	101	219
Age	47.36 (8.32)	54.15 (8.07)	57.55 (7.85)	58.97 (8.15)	58.06 (8.3)	59.87 (7.76)
Female subjects	7,667 (47.33%)	696 (23.39%)	89 (19.56%)	30 (16.76%)	14 (13.86%)	20 (9.13%)

Age is shown in the format of mean (standard deviation).

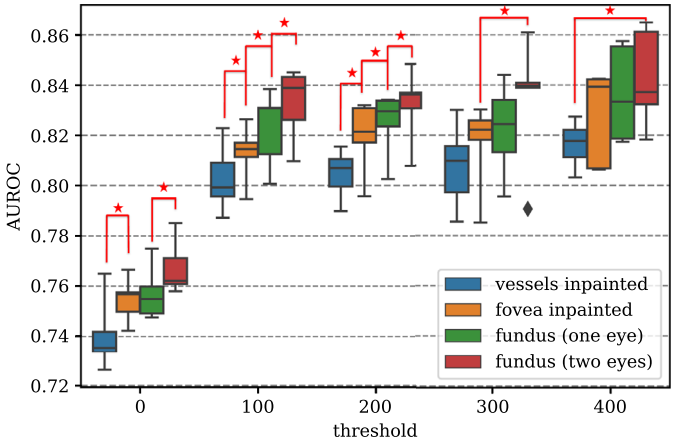


Figure 1. Comparison of area under the receiver operating curve (AUROC) with different input types and varying thresholds for high coronary artery calcium score (CACS). While the threshold of high CACS was set to 0, 100, 200, 300, and 400, normal CACS was defined as CACS = 0. AUROC increased monotonically as vessels-inpainted < fovea-inpainted < fundus (one eye) < fundus (two eyes) at all threshold values, and statistical significance was maintained for distant pairs of input types. For instance, if AUROC differed significantly between vessels-inpainted images and fovea-inpainted images, AUROC of vessel-inpainted images also differed significantly compared with that of unilateral fundus images as well as bilateral fundus images. For the sake of visual clarity, the nearest pairs are marked with asterisks when there exist statistically significant margins between two different input types ($P < 0.05$).

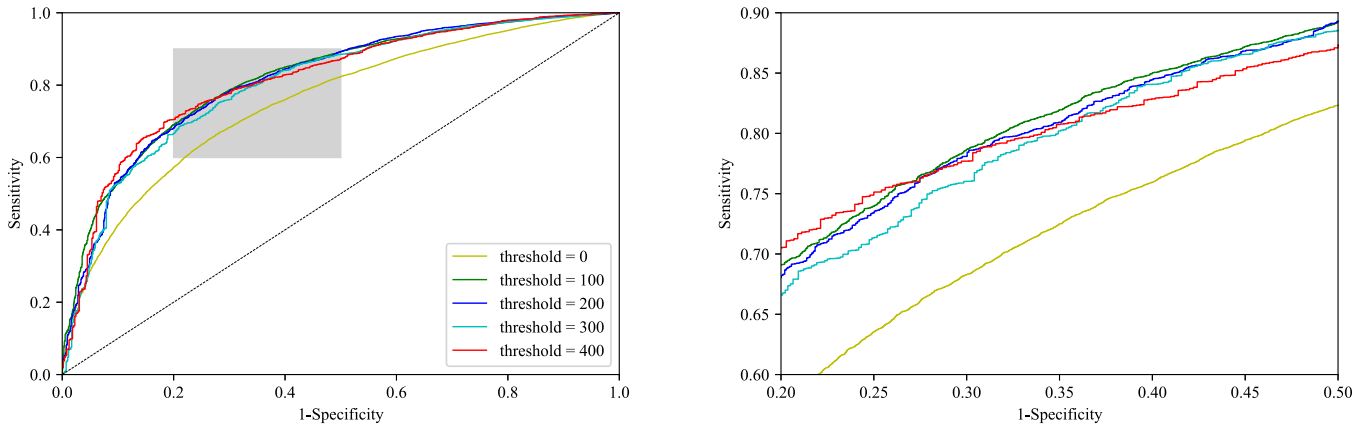
unilateral fundus image was given as input, AUROC of vessel-inpainted images dropped significantly by about 2 percentage points ($P < 0.05$) at the thresholds of 0, 100, 200, and AUROC of fovea-inpainted images dropped about 1 % point ($P < 0.05$) at the thresholds of 100 and 200, whereas AUROC of bilateral images significantly increased about 1% point ($P < 0.05$) at the threshold of 0, 100, 200. AUROC of bilateral fundus images was better than that of vessels-inpainted images at all thresholds.

We investigated where neural networks attend to for making the binary decision by visualizing activation maps (Fig. 3). Neural networks that discriminate no CACS from CACS higher than 100 were used, which was the threshold at which AUROC seemed to saturate. The heatmaps show that neural networks mainly attend to temporal retinal vasculature for decision making.

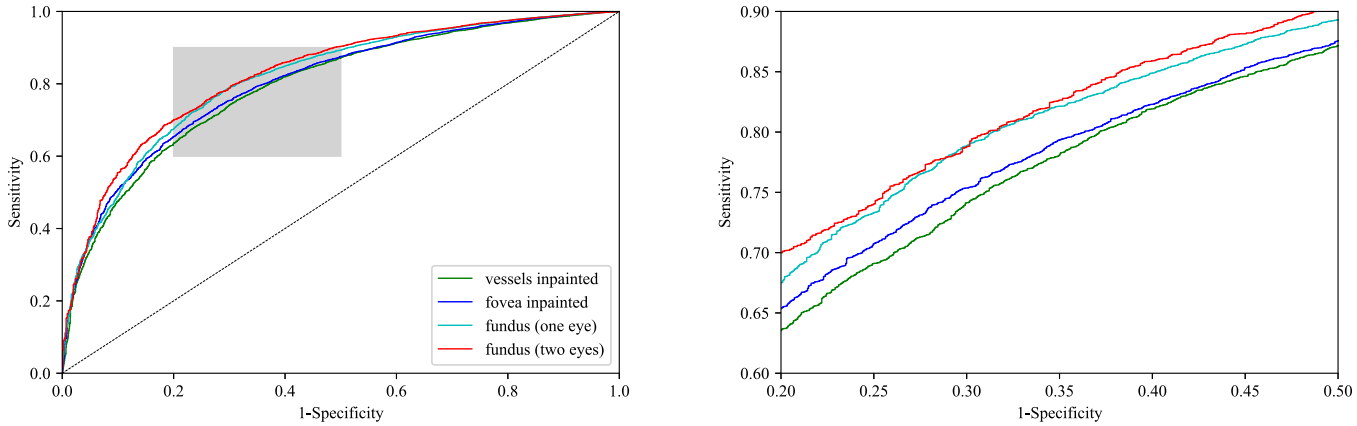
The performance of the deep learning algorithm using retinal fundus images was compared with algorithms using other previously known cardiovascular risk factors. The AUROC for predicting high CACS (CACS > 100) from no CACS was 82.3 (79.5–85.0) with unilateral fundus images and 83.2 (80.2–86.3) with bilateral images. AUROC of 0.82.8 (80.9–84.8) could be achieved using age only and improved further by adding other cardiovascular risk factors; the AUROC was 84.4 (81.6–87.1) with age and presence of hypertension, 87.5 (86.3–88.6) with age and gender, 88.3 (86.9–89.7) with age, presence of hypertension, and sex, and 88.6 (87.1–90.2) with age, presence of hypertension, gender, and retinal fundus images.

Discussion

Researchers have historically sought for a relevant, less-invasive, radiation-free biomarker that can predict future cardiovascular accidents.[30–33] The present study suggests that retinal fundus images might be a promis-



(a) Various threshold values for abnormal CACS (fundus image)



(b) Various input types (no CACS vs. CACS > 100)

Figure 2. Receiver operating curve curves for (a) various threshold values for high coronary artery calcium score (CACS) when input is retinal fundus image, and (b) various input types when high CACS threshold was set to 100 (no CACS vs. CACS > 100). Discriminating regime is magnified.

ing modality for predicting cardiovascular accidents by reflecting the CAC status. The deep learning algorithm moderately discriminates between individuals having high CACS with their retinal fundus images and performs better when the threshold for abnormal CACS is increased. Interestingly, the performance increased rapidly when the threshold was changed from 0 to 100, whereas above 100, the improvements

Table 2. Performance of the Deep Learning Algorithm With Varying Thresholds for High Coronary Artery Calcium Score (CACS) and Different Input Types (Vessel-Inpainted Images, Fovea-Inpainted Images, Intact Fundus Images [One Eye], Fundus Images [Two Eyes])

	Input Type			
Threshold	Vessel Inpainted	Fovea Inpainted	Fundus (One Eye)	Fundus (Two Eyes)
> 0	74.0 (71.5–76.6)	75.5 (73.9–77.1)	75.7 (73.8–77.7)	76.8 (74.7–78.8)
> 100	80.3 (77.9–82.7)	81.3 (79.2–83.3)	82.3 (79.5–85.0)	83.2 (80.2–86.3)
> 200	80.5 (78.7–82.2)	81.9 (79.4–84.5)	82.5 (80.2–84.8)	83.2 (80.7–85.8)
> 300	80.8 (77.8–83.8)	81.6 (78.5–84.8)	82.2 (78.9–85.5)	83.4 (78.4–88.3)
> 400	81.6 (80.0–83.3)	82.7 (79.4–86.1)	83.6 (80.3–87.0)	84.3 (80.4–88.2)

Area under the receiver operating curve (AUROC) and 95% confidence interval is given in a format of mean (lower bound – upper bound).

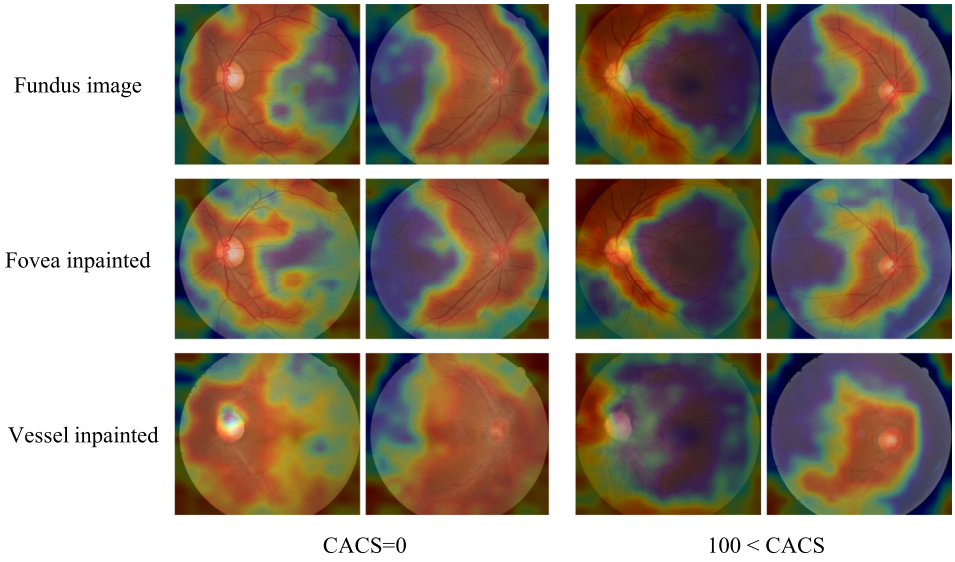


Figure 3. Exemplar heatmaps for deep learning algorithms that discriminate coronary artery calcium score (CACS) = 0 from CACS > 100 with different input types. Heatmaps for fovea-inpainted and vessels-inpainted images were compared with those of intact fundus images. Regardless of input types, the deep learning algorithms mainly attend to central main retinal branches to make the binary decision with respect to the abnormality in CACS. The attention tends to be more diffused when vessels were removed and inpainted.

were marginal. This implies that discriminative features related to high CACS exist in retinal fundus images and that neural networks can perceive when the CACS is above 100. Through the analysis of performances with inpainted images and the activation maps, it is conjectured that such visual patterns reside mainly in the configuration of temporal retinal vasculature and to some part in the fovea.

This study aligns with previous efforts to estimate cardiovascular risks from retinal fundus images. It has been reported that the deep learning algorithm is capable of learning visual patterns from retinal fundus images to predict major adverse cardiovascular events (MACE) in 5 years with a moderate performance (AUROC about 0.7),[20] let alone other CVD risk factors such as age and sex. Whereas MACE is a binary label, CACS is a continuous variable that is often used in clinical practice as a stratified categorical value. Several guidelines exist to help interpret CACS in the context of cardiovascular risk of a patient. A widely accepted guideline of CACS does not use it as its actual value, but uses it as a categorical value to divide into different groups that require different types of care and interventions. For example, the guideline on the primary prevention of cardiovascular disease suggested by the American Heart Association states that CACS >400 with other complications would indicate very high risk of cardiovascular failure whereas CACS <100 translates to the low risk and 100 < CACS <399 means intermediate risk.[34] At least when dividing the CACS value into categories, dividing by 0, 1–100, and 101–300 (or 400) has some consensus.[35] In addition, many studies have shown that patients with CACS >100 are more likely to develop cardiovascular disease in the future and undergo cardiovascular intervention or surgery. According to Jung et al.,[36] the CACS cut-off value for predicting coronary revascularization was 111.0 among Koreans. Therefore various thresholds for high CACS were tested in this study to evaluate whether CACS could be predicted using deep learning algorithms, for which performance seemed to saturate at 100, which can also be considered a clinically significant value according to the previously mentioned studies.

Age, sex, and the presence of hypertension, which are CVD risk factors, are well known to have a moderate relationship with CACS (e.g., an elderly man with hypertension would have higher CACS). In this study, retinal fundus images had an approximately similar predictive performance as age in predicting high CACS from no CACS. Also, a combination of known cardiovascular risk factors showed higher AUROC, as can be expected from the well-known relationship of these factors with CACS. However, retinal fundus images also had an additive effect on increasing AUROC to nearly 90% when combined with age, sex, and presence of hypertension, suggesting retinal fundus images may provide additional information unrelated to these previously known risk factors.

Retinal funduscopic examination has been considered as a method of detecting various eye diseases rather than CVD. However, retinal fundus images provide an overview of retinal vasculature, which could exhibit the signs of apparent cardiovascular burdens, possibly associated with high accumulation of CAC. High CACS would incur hypertensive vascular changes such as increases in stiffness in retinal vasculature and alternation in venules, which may not be readily identified in fundus images of subtle cases. Deep learning algorithms, which operate on a pixel level, may perform superiorly to human experts when it comes to capturing such minute changes that would have otherwise failed to attract humans' attention. The algorithms seem to inspect retinal vasculature based on the observed heatmaps around the main retinal vessel branches as shown in Figure 3, and perhaps the algorithms can perceive visual characteristics that reside in the eyes of patients with high CACS other than the well-known features of hypertension. Interestingly, the performance exacerbated markedly when the retinal vasculature was erased from the input images, while performance decreased by a lesser degree when the fovea was erased. These results support researchers' long belief that the retinal vasculature provides meaningful features assessing cardiovascular burden, including CACS.[20] The fovea, which is known to play a key role in predicting sex with retinal fundus images,[20] may not act as a clue in assessing cardiovascular burden. Still, further investigation is warranted to specify which visual patterns are recognized by the algorithms.

However, this study has several limitations. Future external validations using heterogeneous datasets acquired from various retinal fundus cameras are warranted. Furthermore, ethnicity in our study is heavily skewed towards Asian populations, and the same trends may not be observed in other ethnicities. Also, there were a relatively smaller number of cases with high CACS, although more than 20,000 individuals were included in this study, which also raises the need for further evaluation in larger studies. Constraints in resources rendered more rigorous statistical analyses infeasible, with a small sample size limiting statistical power for the comparison of performance among settings with the Wilcoxon signed rank test. Finally, the question of precisely what features predict CACS still remains unanswered. For such reasons, our findings should be carefully interpreted with deliberation. Despite these limitations, the strength of the present study lies in the novelty of the unprecedented dataset, unique in its composition and scale. We evaluated more than 40,000 retinal fundus images of more than 20,000 individuals, along with CACS evaluation from CT scans, all acquired on the same day. With the availability of retinal fundus images for both eyes, we observed empirical advantages of using two eyes for the prediction of high CACS.

We demonstrated that visual patterns of retinal fundus images in subjects with high CAC could be recognized by deep learning algorithms, and such patterns may reside in retinal vasculature rather than the fovea through ablation tests and heatmap analysis. Ideally, a deep learning algorithm for predicting high CACS could be run at the time a patient takes an automatic retinal fundus photograph, because only a few additional computations are required. We believe that our experiments are an early beginning in paving the way to implementing an additional screening method for the risk of CVD in an accessible and affordable way. However, the current performance is incomplete for the system to be deployed in clinical settings and warrants improvement and rigorous validation in heterogeneous external datasets. Additionally, the prediction of CACS in a direct manner from retinal fundus images also seems to be an interesting research direction that deserves further investigation.

Acknowledgment

Supported by the Research Grant for Intelligence Information Service Expansion Project, which is funded by National IT Industry Promotion Agency (NIPA-C0202-17-1045) and the Small Grant for Exploratory Research of the National Research Foundation of Korea (NRF), which is funded by the Ministry of Science, ICT, and Future Planning (NRF-2018R1D1A1A09083241). The sponsors or funding organizations had no role in the design or conduct of this research.

Disclosure: **J. Son,** VUNO (E); **J.Y. Shin,** None; **E.J. Chun,** None; **K.-H. Jung,** VUNO (I, E); **K.H. Park,** None; **S.J. Park,** VUNO (I)

* JS and JYS contributed equally to this work.

References

1. Wang SB, Mitchell P, Liew G, et al. A spectrum of retinal vasculature measures and coronary artery disease. *Atherosclerosis*. 2018;268:215–224.
2. Seidelmann SB, Claggett B, Bravo PE, et al. Retinal Vessel Calibers in Predicting Long-Term Cardiovascular Outcomes: The Atherosclerosis Risk in

Communities Study. *Circulation*. 2016;134:1328–1338.
3. McGeechan K, Liew G, Macaskill P, et al. Risk prediction of coronary heart disease based on retinal vascular caliber (from the Atherosclerosis Risk In Communities [ARIC] Study). *Am J Cardiol*. 2008;102:58–63.
4. Liew G, Benitez-Aguirre P, Craig ME, et al. Progressive Retinal Vasodilation in Patients With Type 1 Diabetes: A Longitudinal Study of Retinal Vascular Geometry. *Invest Ophthalmol Vis Sci*. 2017;58:2503–2509.
5. Taylor AM, MacGillivray TJ, Henderson RD, et al. Retinal vascular fractal dimension, childhood IQ, and cognitive ability in old age: the Lothian Birth Cohort Study 1936. *PLoS One*. 2015;10:e0121119.
6. McGrory S, Taylor AM, Kirin M, et al. Retinal microvascular network geometry and cognitive abilities in community-dwelling older people: The Lothian Birth Cohort 1936 study. *Br J Ophthalmol*. 2017;101:993–998.
7. McGrory S, Cameron JR, Pellegrini E, et al. The application of retinal fundus camera imaging in dementia: A systematic review. *Alzheimers Dement (Amst)*. 2017;6:91–107.
8. Conti CR. Coronary Artery Calcium Scanning: Past, Present, and Future. *JACC Cardiovasc Imaging*. 2016;9:330–331.
9. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. *Journal of the American College of Cardiology*. 1990;15:827–832.
10. McCollough CH, Ulzheimer S, Halliburton SS, Shanneik K, White RD, Kalender WA. Coronary artery calcium: a multi-institutional, multimanufacturer international standard for quantification at cardiac CT. *Radiology*. 2007;243:527–538.
11. Greenland P, Smith Jr SC, Grundy SM. Improving coronary heart disease risk assessment in asymptomatic people: role of traditional risk factors and noninvasive cardiovascular tests. *Circulation*. 2001;104:1863–1867.
12. Arad Y, Goodman KJ, Roth M, Newstein D, Guerci AD. Coronary calcification, coronary disease risk factors, C-reactive protein, and atherosclerotic cardiovascular disease events: the St. Francis Heart Study. *J Am Coll Cardiol*. 2005;46:158–165.
13. Greenland P, LaBree L, Azen SP, Doherty TM, Detrano RC. Coronary artery calcium score combined with Framingham score for risk prediction in asymptomatic individuals. *JAMA*. 2004;291:210–215.
14. Nasir K, Clouse M. Role of nonenhanced multidetector CT coronary artery calcium testing in asymptomatic and symptomatic individuals. *Radiology*. 2012;264:637–649.
15. Becker A, Leber A, Becker C, Knez A. Predictive value of coronary calcifications for future cardiac events in asymptomatic individuals. *Am Heart J*. 2008;155:154–160.
16. McClelland RL, Chung H, Detrano R, Post W, Kronmal RA. Distribution of coronary artery calcium by race, gender, and age: results from the Multi-Ethnic Study of Atherosclerosis (MESA). *Circulation*. 2006;113:30–37.
17. Budoff MJ, Nasir K, McClelland RL, et al. Coronary calcium predicts events better with absolute calcium scores than age-sex-race/ethnicity percentiles: MESA (Multi-Ethnic Study of Atherosclerosis). *J Am Coll Cardiol*. 2009;53:345–352.
18. Tanuseputro P, Manuel DG, Leung M, Nguyen K, Johansen H, Canadian Cardiovascular Outcomes Research T. Risk factors for cardiovascular disease in Canada. *Can J Cardiol*. 2003;19:1249–1259.
19. Ueshima H, Sekikawa A, Miura K, et al. Cardiovascular disease and risk factors in Asia: a selected review. *Circulation*. 2008;118:2702–2709.
20. 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.
21. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521:436–444.
22. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. 2017;42:60–88.
23. Moccia S, De Momi E, El Hadji S, Mattos LS. Blood vessel segmentation algorithms—review of methods, datasets and evaluation metrics. *Comput Methods Programs Biomed*. 2018;158:71–91.
24. Park SJ, Shin JY, Kim S, Son J, Jung K-H, Park KH. A novel fundus image reading tool for efficient generation of a multi-dimensional categorical image database for machine learning algorithm training. *J Korean Med Sci*. 2018;33:e239.
25. Son J, Shin JY, Kim HD, Jung KH, Park KH, Park SJ. Development and validation of deep learning models for screening multiple abnormal findings in retinal fundus images. *Ophthalmology*. 2020;127:85–94.
26. Son J, Park SJ, Jung KH. Towards accurate segmentation of retinal vessels and the optic disc

in fundoscopic images with generative adversarial networks. *J Digit Imaging*. 2019;32:499–512.
27. Porwal P, Pachade S, Kamble R, et al. Indian Diabetic Retinopathy Image Dataset (IDRiD). *IEEE Dataport*. 2018.
28. Krause J, Gulshan V, Rahimy E, et al. Grader variability and the importance of reference standards for evaluating machine learning models for diabetic retinopathy. *Ophthalmology*. 2018;125:1264–1272.
29. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016:2818–2826.
30. 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.
31. Andrus B, Lacaille D. 2013 ACC/AHA guideline on the assessment of cardiovascular risk. *J Am Coll Cardiol*. 2014;63:2886.
32. Conroy RM, Pyorala K, Fitzgerald AP, et al. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J*. 2003;24:987–1003.
33. Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. *BMJ*. 2017;357:j2099.
34. Fihn SD, Gardin JM, Abrams J, et al. 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS guideline for the diagnosis and management of patients with stable ischemic heart disease: a report of the American College of Cardiology Foundation/American Heart Association task force on practice guidelines, and the American College of Physicians, American Association for Thoracic Surgery, Preventive Cardiovascular Nurses Association, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons. *Circulation*. 2012;126:3097–3137.
35. Cho I, Al'Aref SJ, Berger A, et al. Prognostic value of coronary computed tomographic angiography findings in asymptomatic individuals: a 6-year follow-up from the prospective multicentre international CONFIRM study. *Eur Heart J*. 2018;39:934–941.
36. Jung IH, Kim JS, Nam HJ, et al. Clinical outcomes according to coronary calcium scores in asymptomatic individuals undergoing coronary CT angiography. *Korean J Med*. 2010;78:466.