MACHINE LEARNING

Eyeing cardiovascular risk factors

Smoking status, blood pressure, age and other cardiovascular risk factors can be predicted from retinal images by using deep learning.

Daniel Shu Wei Ting and Tien Yin Wong

The retina is the only 'window' into the body that allows physicians to directly visualize non-invasively the local microvasculature. The measurement of retinal vascular parameters (such as the diameters of arteries and veins) from retinal fundus images by using semi-automated software has shown that these parameters are associated with the risk of major cardiovascular disease (CVD) outcomes[1], including stroke[2], heart disease, chronic kidney disease[3] (CKD) and mortality[4]. For example, people with wider retinal venules and narrower arterioles are associated with an increased risk of CVD[5]. Moreover, for diabetic patients, the use of measurements based on retinal imaging along with traditional risk factors (such as blood pressure) and contemporary biomarkers (such as serum levels of C-reactive protein) improved the prediction of the risk of CVD[1]. Also, in these patients, the prediction of the risk of stroke on the basis of established risk factors improved by about 12% when measurements from retinal images were included[2]. In population studies, retinal-imaging parameters may predict the onset of CKD and the worsening of kidney function[3] (people with both abnormal retinal parameters and CKD have a much higher risk of mortality[6]).

Therefore, populations could in principle be stratified according to CVD risk on the basis of information from retinal images. However, the use of retinal fundus images to predict CVD risk in a clinical setting is limited by the need for expert assessors to evaluate the images using currently available semi-automated software. Some software requires 30–45 min to just correctly classify arteries and veins, to measure specific segments of the vessels and to adjudicate variations. Writing in *Nature Biomedical Engineering*, Lily Peng and colleagues now show that a suitably trained deep-learning algorithm can predict a range of cardiovascular risk factors — including age, gender, current smoking status (smoker or non-smoker), systolic blood pressure and body mass index (BMI) — from retinal fundus photographs[7].

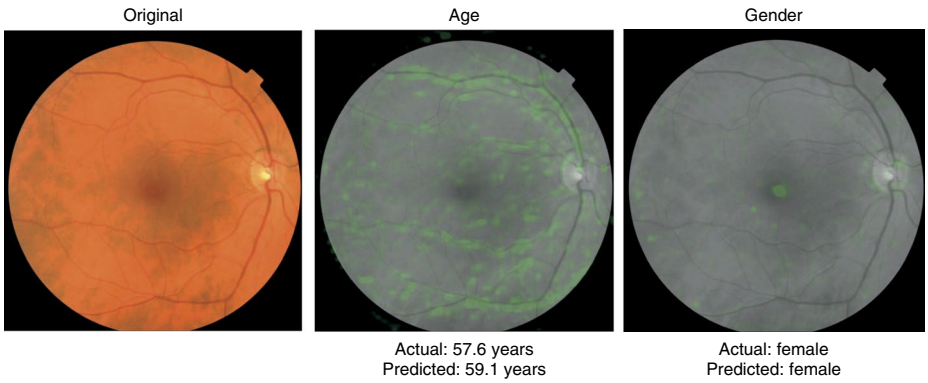

Fig. 1 | Anatomical regions in retinal fundus images associated with predictions from the deep-learning algorithm. Predictions of age and gender rely on, respectively, the vasculature and the optic disc (indicated in green in the images). Figure reproduced from ref. [7], Macmillan Publishers Ltd.

Deep learning — the most used machine-learning method for data extraction and transformation, consisting of a neural network of interconnected layers of nonlinear processing units — has sparked tremendous interest in medicine[8], in particular in imaging applications such as the detection of diabetic retinopathy and related retinal diseases[9–11], tuberculosis[12,13], melanoma[14] and lymph-node metastases secondary to breast cancer[15]. Such an interest stems from significant improvements in the underlying mathematical models and in data availability, as well as from robust outcomes. In fact, deep learning has achieved comparable, if not superior, performance to the current standards of care (typically, assessment by human experts) in all these conditions.

Peng and co-authors' deep-learning algorithm was trained on data collected from 284,335 patients included in the UK Biobank (48,101 patients) and the United States EyePACS database (236,234 patients), and validated on two independent datasets of 12,026 patients (from the UK Biobank) and 999 patients (from the EyePACS database). The researchers demonstrate that the deep-learning algorithm accurately predicted the CVD risk factors: age within 3.26 years, gender with an area under the curve (AUC) of 0.97, current smoking status with an AUC of 0.71, systolic blood pressure within 11.23 mmHg, and BMI within 3.29 units (Table 1). The algorithm also predicted the risk of a major adverse cardiac event (MACE; such as myocardial infarction and cardiac shock) within 5 years, with an AUC of 0.70. Importantly, some of the predictions of the deep-learning algorithm could be associated to specific regions and anatomical features (such as blood vessels or the optic disc) in the fundus photographs (Fig. 1). Such associations should have biological relevance and thus help physicians identify pathological phenomena.

During the peer review of Peng and co-authors' work, the authors were asked to test the performance of the deep-learning algorithm in a small dataset (239 patients) from a randomly selected Asian database. Comparison of the predictions with the ground truth revealed that the performance of the algorithm for all predicted CVD risk factors across all datasets was similar (Table 1). This indicates that the predictions of the algorithm are robust and that they should be generalizable to Asian populations (the UK Biobank and EyePACS datasets consist of primarily white patients).

The use of large and diverse datasets for the training of the deep-learning algorithm and for the validation of the predictions is a strength of Peng and co-authors' work.

Table 1 | Performance of the deep-learning algorithm in predicting cardiovascular risk factors for two independent datasets

CVD risk factors	Primary validation set[7] (mostly white patients)	Independent validation set (Asian patients)
Age (MAE in years)	3.26 (3.22,3.31)	3.42 (3.06,3.78)
Age (R^2)	0.74 (0.73,0.75)	0.79 (0.74,0.83)
Gender (AUC)	0.97 (0.966,0.971)	0.98 (0.96,0.99)
Current smoker (AUC)	0.71 (0.70,0.73)	0.79 (0.70,0.88)
HbA1c (MAE in %)	1.39 (1.29,1.50)	0.92 (0.83,1.00)
HbA1c (R^2)	0.09 (0.03,0.16)	0.24 (0.10,0.39)
Systolic BP (MAE in mmHg)	11.35 (11.18,11.51)	14.31 (12.96,15.66)
Systolic BP (R^2)	0.36 (0.35,0.37)	0.34 (0.25,0.44)
Diastolic BP (MAE in mmHg)	6.42 (6.33,6.52)	7.93 (7.25,8.61)
Diastolic BP (R^2)	0.32 (0.30,0.33)	0.36 (0.26,0.45)
BMI (MAE)	3.29 (3.24,3.34)	3.57 (3.21,3.94)
BMI (R^2)	0.13 (0.11,0.14)	0.16 (0.06,0.28)

The two datasets used are a primary dataset from UK and US patients[7] (HbA1c from the EyePACS database, and the rest of the risk factors from the UK Biobank), and a smaller dataset from Asian patients only. Values within brackets represent 95% confidence intervals. AUC, area under the curve; BP, blood pressure; MAE, mean absolute error; R^2, coefficient of determination.

Still, the datasets used by the authors to predict 5-year MACEs did not include data on relevant Framingham risk factors, such as glycosylated haemoglobin (HbA1c), and total-cholesterol and high-density-lipoprotein cholesterol levels. Without these classic risk factors, it is unclear whether the predictions from the fundus-photograph-based deep-learning algorithm can be directly compared to those from the conventional Framingham risk equation (a predictive equation borne out of the Framingham Heart Study, a long-term cardiovascular cohort study). In addition, some of the risk factors (such as blood pressure and BMI) in the UK Biobank and EyePACS datasets were measured with different protocols, which may also contribute to variability in the diagnostic performance of the deep-learning algorithm. Nevertheless, the risk predictions for MACEs from Peng and co-authors' deep-learning algorithm and from risk-assessment maps based on various non-invasive, easily measured risk factors (age, gender, blood pressure, BMI and smoking status) are comparable. It would thus be worthwhile to test whether a fundus photograph can replace some of the invasive blood tests (such as lipid and renal profiles) for the prediction of MACEs in patient populations. The translational value of deep learning in the prediction of CVD risk is the ability to replace the risk factors that are most difficult to measure (such as HbA1c and cholesterol profile). For this, and towards potential clinical adoption, the algorithm would need to be tested in larger, well-phenotyped and diverse prospective population cohorts.

Interestingly, that information in a retinal image can be used for the prediction of a person's gender is surprising and puzzling. This underscores the potential of artificial intelligence to revolutionize the way medicine is practiced and to help discover hidden associations. Given the increasing size of ageing populations, of chronic-disease burden and of healthcare costs, artificial intelligence and big-data analytics will need to play a critical role in disease prevention and diagnosis. ❐

Daniel Shu Wei Ting and Tien Yin Wong
Singapore Eye Research Institute, Singapore National Eye Center, Duke-NUS Medical School and National University of Singapore, Singapore, Singapore.
e-mail: daniel.ting.s.w@singhealth.com.sg; wong.tien.yin@singhealth.com.sg

Published online: 7 March 2018
https://doi.org/10.1038/s41551-018-0210-5

References
1. Ho, H. et al. *Sci. Rep.* **7**, 41492 (2017).
2. Cheung, C. Y. et al. *Stroke* **44**, 2402–2408 (2013).
3. Yip, W. et al. *Sci. Rep.* **7**, 9374 (2017).
4. Wong, T. Y., Knudtson, M. D., Klein, R., Klein, B. E. & Hubbard, L. D. *Am. J. Epidemiol.* **159**, 819–825 (2004).
5. McGeechan, K. et al. *Ann. Intern. Med.* **151**, 404–413 (2009).
6. Yip, W. et al. *Sci. Rep.* **6**, 27442 (2016).
7. Poplin, R. et al. *Nat. Biomed. Eng.* https://doi.org/10.1038/s41551-018-0195-0 (2018).
8. Brynjolfsson, E. & Mcafee, A. What's driving the machine learning explosion? *Harvard Business Review* (18 July 2017).
9. Ting, D. S. W. et al. *JAMA* **318**, 2211–2223 (2017).
10. Gulshan, V. et al. *JAMA* **316**, 2402–2410 (2016).
11. Wong, T. Y. & Bressler, N. M. *JAMA* **316**, 2366–2367 (2016).
12. Lakhani, P. & Sundaram, B. *Radiology* **284**, 574–582 (2017).
13. Ting, D. S. W., Yi, P. H. & Hui, F. *Radiology* **286**, 729–731 (2018).
14. Esteva, A. et al. *Nature* **542**, 115–118 (2017).
15. Ehteshami Bejnordi, B. et al. *JAMA* **318**, 2199–2210 (2017).

Competing interests
D.S.W.T. and T.Y.W. are co-developers of a deep-learning system for diabetic retinopathy, glaucoma and age-related macular degeneration.