medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Development and validation of a deep learning algorithm based on fundus photographs for estimating the CAIDE dementia risk score

Rong Hua, Bachelor [1–3]; Jianhao Xiong, PhD[4]; Gail Li, MD, PhD[5]; Yidan Zhu, PhD[1–3]; Zongyuan Ge, PhD[4]; Yanjun Ma, Bachelor [1–3]; Meng Fu, MS[4]; Chenglong Li, Bachelor [1–3]; Bin Wang, MS[4]; Li Dong, MD[6]; Xin Zhao, MS[4]; Zhiqiang Ma, MD[7]; Jili Chen, Bachelor[8]; Chao He, PhD[4]; Zhaohui Wang, MD[7]; Wenbin Wei, PhD[6]; Yuzhong Chen, PhD[4]; Wuxiang Xie, PhD[1–3]

R. Hua and J. Xiong contributed equally to this work.

Affiliations:

[1] Peking University Clinical Research Institute, Peking University First Hospital, Beijing, China;

[2] PUCRI Heart and Vascular Health Research Center at Peking University Shougang Hospital, Beijing, China;

[3] Key Laboratory of Molecular Cardiovascular Sciences (Peking University), Ministry of Education, Beijing, China;

[4] Beijing Airdoc Technology Co., Ltd. Beijing, China;

[5] Departments of Psychiatry and Behavioral Sciences and Division of Gerontology and Geriatric Medicine, University of Washington, Seattle, WA, USA.

[6] Beijing Tongren Eye Center, Beijing Tongren Hospital, Beijing, China;

[7] iKang Guobin Healthcare Group Co., Ltd, Beijing, China;

[8] Shibei Hospital Jingan District, Shanghai, China.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Corresponding Authors:

Wuxiang Xie, PhD, Peking University Clinical Research Institute, Peking University First Hospital, No. 38 Xueyuan Road, Haidian District, 100191, Beijing, China. Telephone: +86-10-82805564-622; Fax: +86-10-62015547; E-mail: xiewuxiang@hsc.pku.edu.cn.

Yuzhong Chen, Beijing Airdoc Technology Co., Ltd. Room 21, Floor 4, Building 2, Courtyard A 2, North West Third Ring Road, Haidian District, 100089, Beijing, China; Telephone: +86-10-82362300; Fax: +86-10-82362300; E-mail: chenyuzhong@airdoc.com.

Total word count of the manuscript: 3412 words. 3 tables, 2 figures, and 1 supplementary file.

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Key Points

Question: Can a deep learning algorithm based on fundus images estimate the CAIDE dementia risk score and identify individuals with high dementia risk?

Findings: The algorithm developed by fundus photographs from 258,305 check-up participants could well identify individuals with high dementia risk, with area under the receiver operating characteristic curve of 0.931 in internal validation and 0.782 in external validation dataset, respectively. Besides, the estimated CAIDE dementia risk score generated from the algorithm exhibited significant association with cognitive function.

Meaning: The deep learning algorithm based on fundus photographs has potential to screen individuals with high dementia risk in population-based settings.

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Abstract

Importance: The Cardiovascular Risk Factors, Aging, and Incidence of Dementia (CAIDE) dementia risk score is a recognized tool for dementia risk stratification. However, its application is limited due to the requirements for multidimensional information and fasting blood draw. Consequently, effective, convenient and noninvasive tool for screening individuals with high dementia risk in large population-based settings is urgently needed.

Objective: To develop and validate a deep learning algorithm using retinal fundus photographs for estimating the CAIDE dementia risk score and identifying individuals with high dementia risk.

Design: A deep learning algorithm trained via fundus photographs was developed, validated internally and externally with cross-sectional design.

Setting: Population-based.

Participants: A health check-up population with 271,864 adults were randomized into a development dataset (95%) and an internal validation dataset (5%). The external validation used data from the Beijing Research on Ageing and Vessel (BRAVE) with 1,512 individuals.

Exposures: The estimated CAIDE dementia risk score generated from the algorithm.

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Main Outcome and Measure: The algorithm's performance for identifying individuals with high dementia risk was evaluated by area under the receiver operating curve (AUC) with 95% confidence interval (CI).

Results: The study involved 258,305 participants (mean aged 42.1 ± 13.4 years, men: 52.7%) in development, 13,559 (mean aged 41.2 ± 13.3 years, men: 52.5%) in internal validation, and 1,512 (mean aged 59.8 ± 7.3 years, men: 37.1%) in external validation. The adjusted coefficient of determination (R^2) between the estimated and actual CAIDE dementia risk score was 0.822 in the internal and 0.300 in the external validations, respectively. The algorithm achieved an AUC of 0.931 (95%CI, 0.922–0.939) in the internal validation group and 0.782 (95%CI, 0.749–0.815) in the external group. Besides, the estimated CAIDE dementia risk score was significantly associated with both comprehensive cognitive function and specific cognitive domains.

Conclusions and Relevance:

The present study demonstrated that the deep learning algorithm trained via fundus photographs could well identify individuals with high dementia risk in a population-based setting. Our findings suggest that fundus photography may be utilized as a noninvasive and more expedient method for dementia risk stratification.

Introduction

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Worldwide, the number of people have dementia is projected to triply increase to 152 million by 2050, given the dramatic rise in ageing populations, yet there are no curative therapeutics available.[1] Dementia has a long preclinical phase when no symptomatic cognitive impairments, but neurodegenerative progressions are occuring.[2] Early identification of high-risk individuals is essential for preventing dementia, which efficiently targets participants who could benefit most from more intensive examinations and interventions.[3]

The Cardiovascular Risk Factors, Aging, and Incidence of Dementia (CAIDE) dementia risk score was a recognized model to predict 20-year dementia risk, which based on multidimensional risk factors: age, sex, educational level, physical inactivity, systolic blood pressure (SBP), total cholesterol (TC), and body mass index (BMI). It was also highly predictive in external validation of a large multiethnic population,[4-6] and adopted in Finnish Geriatric Intervention Study (FINGER) to select eligible at-risk participants. However, the CAIDE dementia risk score entails measurements by questionnaire inquiry, physical examinations and fasting blood draw, these procedures are time-consuming or invasive for participants, also increase the labor costs of healthcare practitioners and produce biohazardous waste. Consequently, effective, convenient and noninvasive tool to screen individuals with high dementia risk in large population-based settings is warranted.

Vascular disease, especially microvasculature damage in the brain, is recognized as a major contributor to dementia.[1,7] Anatomically and

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

developmentally, the retina shares homology with the brain.[8] The retina is an exceptional site where the microcirculation can be handily and noninvasively visualized by fundus photography,[9] thus providing insights into the brain microvasculature. Large population studies have demonstrated the correlations between various retinal microvascular abnormalities (such as retinopathy, arteriolar narrowing and venular dilation) and increased risk of dementia.[10-12] Moreover, The emerging artificial intelligence technique, especially deep learning, has realized integrating multiple retinal features from fundus photographs, to provide estimation on vascular risk factors, and prediction on cardiovascular diseases.[13-15] However, to our knowledge, this method has not been investigated on dementia.

Herein, we hypothesized that the deep learning algorithm trained via fundus photographs might help to dementia risk stratification. Due to the insufficient time length to occur enough dementia events in our dataset, the present study aimed to train a deep learning algorithm for estimating the CAIDE dementia risk score thus identifying individuals with high dementia risk, and we proposed that the estimated score generated from the algorithm associated with the cognitive function.

Methods

Study design

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

This was a cross-sectional study. A deep learning algorithm based on fundus photographs for estimating the CAIDE dementia risk score was developed and internally validated by a medical check-up dataset. Additionally, by an independent cohort dataset, we externally validated the algorithm's discrimination on individuals with high dementia risk, and further explored the association between the estimated CAIDE dementia risk score and cognitive function.

Participants and datasets

For the algorithm development, a dataset from 271,864 participants from Tongren Hospital in Beijing, Shibei Hospital in Shanghai, and iKang Healthcare Group who attending medical check-up in 19 province-level administrative regions of China during September 2018 to December 2019, were randomly divided into development (95%) and internal validation (5%) components. This dataset contained retinal fundus images and routine medical information, including age, sex, SBP, TC, and BMI. The use of the dataset for the algorithm training was approved by Tongren Hospital Institutional Review Board, Shibei Hospital Institutional Review Board, and iKang Healthcare Group Institutional Review Board with a waiver of informed consent. The algorithm's performance was further externally validated using the baseline data from Beijing Research on Ageing and VEssel (BRAVE), a community-based cohort collecting fundus images and health information of middle-aged and older adults in Shijingshan District, Beijing in 2019.[16] The BRAVE was approved by the Peking University

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

School Institutional Review Board (ethical review approval number: IRB0001052–19060), all participants have given written informed consent.

A variety of digital nonmydriatic fundus cameras were adopted to obtain fundus images, including Canon CR1/CR2 and Crystalvue FundusVue/ TonoVue in the check-up dataset, and Centervue DRS in the BRAVE. All images were captured using 45º fields of view. Both datasets calculated the CAIDE dementia risk score based on the function proposed by Kivipelto et al.[4] However, educational level and physical inactivity were not collected in the check-up dataset. Considering the importance of education on dementia, we imputed the risk score of educational level to the algorithm based on the Sixth National Census,[17] according to the average risk score of educational level among the corresponding sex and age group of the individual.

Development of the algorithm

The development dataset consists of a training dataset and a tuning dataset. The training dataset was used to update model parameters during the training stage, and the tuning set was used for model selection. The label for training and testing of the network is given as $y_{\text{CAIDE Score}}$ which is the score summation of risk factors according to the CAIDE dementia risk model.[4]

A convolutional neural network, Inception-Resnet-v2,[18] was deployed as the backbone of the deep learning architecture. The input size of fundus image was 299×299. The return of the backbone is 8×8×1536 dimension vector. This vector was forwarded to 8×8 size average pooling layer and further been

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

flatten into a vector with a size of 1×1536. Subsequently, a dimension of 1×32 fully connected layer was applied and connected to the output $y_{\text{CAIDE Score}}$ estimation. The loss function of score estimation was defined as MAE (mean absolute error). The gradient for model parameter updating from the loss function was calculated by Adam method with an initial learning rate at 0.001.[19] The model development and validation were performed with a batch size of 64.

An image normalization which maps the image pixel values to a certain distribution was applied in the development and validation stages of the neural network.[20] During the training of network, data augmentation of random cropping, random rotation (±30°) and random horizontal flipping were used to improve model robustness. The network is developed and tested on computer environment with Keras platform v2.2.2[21] and Python scikit-learn package 0.22.2.

Validation of the algorithm

The estimated CAIDE dementia risk score of the participants deprived from mean estimated $y_{\text{CAIDE Score}}$ of both eyes, and the actual dementia risk score was calculated according to the CAIDE model. The goodness of fit of the algorithm was assessed by adjusted coefficient of determination (R^2) in the internal validation dataset and the BRAVE. Besides, the algorithm's discrimination on identifying individuals with high dementia risk was evaluated by area under the receiver operating curve (AUC) with 95% confidence interval (CI) by the pROC package version 1.16.2.[22] Consistent with Sindi et al,

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

dementia risk score ≥10 points was recognized as high dementia risk.[6] The maximum Youden index was applied to determine the optimal cut-off point.[23]

Cognitive assessments

We further explored the associations between the estimated CAIDE dementia risk score and cognitive function based on the BRAVE dataset. The primary cognitive measurement in the BRAVE was the Chinese version of Montreal Cognitive Assessment (MoCA) Basic, a sensitive and validated cognitive test battery to comprehensively assess nine cognitive domains.[24,25] In addition, the BRAVE also supplemented three tests to further assess specific cognitive domains. Specifically, the memory function was measured by immediate and delayed recall of a list of ten unrelated words, and the total score ranged from 0 to 20.[26] The language and executive function was assessed by a verbal fluency test, which requiring participants to speak names of animals as many as possible within 1 minute, and the total number of animal names (excluding repetitive names) was count as the test score.[27,28] The attention function and executive function were evaluated by the Chinese version of Trails Making Test (TMT),[29] which asking individuals to draw a line through 25 numbers consecutively in ascending order, and as fast as they could. The TMT included two tasks, the TMT-A comprised numbers from 1 to 25, while the TMT-B was different in 25 numbers enclosed in squares from 1 to 12 and circles from 1 to 13. The TMT-A evaluated processing speed and visual attention, and the TMT-B assessed executive function by measuring cognitive alternation ability.

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

In both tests of memory and verbal fluency, the higher score indicated better cognitive performance, while in the TMT, the longer time manifested worse performance.

Statistical analysis

The results were presented using percentage for categorical variables and means ± standard deviations (SD) or the median with interquartile range (IQR) for continuous variables. We ran multiple linear regression models to examine the associations between the estimated CAIDE dementia risk score and different cognitive assessments. The first model adjusted for age and sex, while the second model adjusted for multiple covariates, which contained age, sex, BMI (kg/m^2), TC (mmol/L), educational level, marriage status, drinking status, smoking status, physical inactivity, depressive symptoms, *APOE* ε4 status, and chronic diseases status. Specifically, high level of education was defined as educational year ≥10 years. Marriage status indicated currently married or not. Participants were divided into non-smokers (including ex-smokers) and current smokers. Alcohol consuming was defined as drinking at least once per week over the past one year. Subjects who took participant in moderate or vigorous physical activity at least twice a week with more than 20 minutes each time, were recognized as active, and others were regarded as physical inactive. The BRAVE employed the ten-item version of the Center for Epidemiologic Studies Depression Scale (CES-D) to assess depressive symptoms, with a summed score ranged from 0 to 30. According to the prior

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

study, a score ≥12 was defined as having depressive symptoms in our study.[30] Individuals were divided into *APOE* ε4 carriers (indicated the presence of one or two ε4 alleles) and noncarriers. Hypertension was defined as systolic blood pressure of ≥140 mm Hg, or diastolic blood pressure of ≥90 mm Hg, or self-reported treatment with anti-hypertensive medication. Diabetes was defined as HbA_{1c} ≥6.5% or fasting blood glucose ≥7.0 mmol/L, or self-reported current use of anti-diabetic therapy. Chronic disease measures also included self-reported physician-diagnosed coronary heart disease, cancer, stroke, and chronic obstructive pulmonary disease. Besides, we also employed analysis of covariance to compare cognitive performance between quartiles of the estimated dementia risk score, with the lowest quartile as the reference. Linear trend was also tested by including risk score quartiles as numerical variables.

To test the robustness of the algorithm, we evaluated the performance of the algorithm using 9 points as the cut-off score of high dementia risk, in consistent with a previous study.[31] We further tested the ability of the algorithm to identify participants eligible for multidomain intervention, since the FINGER trial adopted CAIDE score ≥6 points as one of the inclusion criteria to select eligible at-risk participants among the general population.[32] In addition, we conducted subgroup analyses according to sex, age group (<60 years and ≥60 years), respectively, based on the external validation dataset. For algorithm performance in identifying high risk individuals (with CAIDE score ≥10 points), we used Delong test to compare the AUC between subgroups. For the

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

association with cognitive function, we respectively included the interaction terms of estimated dementia risk score with sex, as well as age group in multivariate linear regression models.

All statistical analyses were performed by SAS 9.4 (SAS Institute, Cary, NC), and R language 4.0.0 (R Foundation, Vienna, Austria), with two-tailed alpha value of 0.05 as the statistically significant level.

Results

Study population

The characteristics of individuals in the development dataset, internal validation dataset, and the BRAVE were summarized in Table 1.

Among the 271,864 check-up participants, we randomly divided 95% (258,305 participants, mean aged 42.1 ± 13.4 years, men: 52.7%) into the development group and 5% (13,559 participants, mean aged 41.2 ± 13.3 years, men: 52.5%) into the internal validation group (eFigure 1a). These two groups shared similar baseline characteristics as shown in Table 1. Besides, the characteristics of participants in the training and tuning groups were displayed in eTable 1. Among 1,554 individuals taking participant in the baseline survey of BRAVE, 1,512 participants (mean aged 59.8 ± 7.3 years, men: 37.1%) had fundus photographs and complete information for calculating CAIDE dementia risk score and thus were included in the external validation group (eFigure 1b). Compared with check-up participants, individuals in the BRAVE were older,

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

had a higher proportion of female, with higher SBP and BMI. Respectively, 200 (1.5%) individuals in the internal validation dataset and 159 (10.5%) in the BRAVE were in high dementia risk, with their CAIDE dementia risk score ≥10 points.

Algorithm performance

The adjusted R^2 between the estimated and actual CAIDE dementia risk score was 0.822 in the internal validation dataset and 0.300 in the BRAVE (Figure 1). As shown in Figure 2, the algorithm achieved an AUC of 0.931 (95%CI, 0.922–0.939) in the internal validation dataset and 0.782 (95%CI, 0.749–0.815) in the BRAVE for identifying individuals with high dementia risk. The maximum Youden index on the two receiver operating characteristic curves were 0.780 with the sensitivity of 0.985 and specificity of 0.795, corresponded to the optimal cut-off point of 5.843 in the internal validation dataset, and 0.453 with the sensitivity of 0.811 and specificity of 0.642, corresponded to the optimal cut-off point of 7.305 in the BRAVE, respectively.

The estimated score and cognitive function

Linear regression analyses found that the estimated CAIDE dementia risk score (as continuous variable) was significantly associated with the score of MoCA. As shown in Table 2, 1-point increment of estimated CAIDE dementia risk score was significantly associated with −0.324 (95%CI, −0.510 to −0.138) increment of the MoCA score after multivariable adjustment, which manifested worse comprehensive cognitive performance. Similarly, the higher estimated

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

CAIDE dementia risk score was significantly associated with lower score of memory and verbal fluency test, which indicated poorer performance of memory, language and executive function. The higher estimated score was also significantly associated with longer TMT-A and TMT-B time, which represented worse attention and executive function. The analysis of covariance found that after full adjustment, compared with the lowest quartile, the second, third, and highest quartiles were associated with worse comprehensive cognitive function, with lower MoCA score by −0.544 (95%CI, −1.062 to −0.046), −0.711 (95%CI, −1.281 to −0.141), and −0.984 (95%CI, −1.625 to −0.343), respectively (*P* for linear trend = 0.005, Table 3). Similar trends were also observed in performance of memory test, verbal fluency test, TMT-A and TMT-B, while the associations between quartiles of estimated CAIDE dementia risk score and verbal fluency test score were not significant.

Sensitivity analysis

As shown in eFigure 2, the algorithm still performed well in screening individuals with high dementia risk when the cut-off score changed to 9 points, with an AUC of 0.949 (95%CI, 0.944–0.953) in the internal validation dataset and 0.751 (95%CI, 0.722–0.780) in the BRAVE. Besides, as shown in eFigure 3, the algorithm exhibited moderate performance in identifying participants eligible for multidomain intervention, with an AUC of 0.978 (95%CI, 0.976–0.980) in the internal validation dataset and 0.738 (95%CI, 0.711–0.766) in the BRAVE. In addition, eFigure 4 summarized the algorithm performance in

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

subgroups of the BRAVE. The algorithm presented a higher AUC in female (0.812 vs 0.738, *P* = 0.045), as well as in participants <60 years (0.802 vs 0.704, *P* = 0.019). As eFigure 5 presented, we found no interaction effect of sex or age group on the associations between estimated CAIDE dementia risk score and the score of MoCA, or other specific cognitive functions.

Discussion

To the best of our knowledge, the present study is the first investigation on developing a deep learning algorithm based on fundus photographs for identifying individuals with high dementia risk, with an AUC of 0.931 (95%CI, 0.922–0.939) in the internal validation, and 0.782 (95%CI, 0.749–0.815) in the external validation dataset. Moreover, the estimated CAIDE dementia risk score exhibited significant associations with both comprehensive and specific domains of cognitive function, which further supported the reasonability of the algorithm. Taken together, our study clarified the feasibility of adopting deep learning algorithm based on fundus photographs to screen individuals with high dementia risk in population-based settings.

The rationale of our work based on the concept that, the retina shares similar morphological features and physiological properties with the brain, and hence provide a unique site to detect changes in microvasculature related to the development of dementia.[8] Previous studies have investigated the associations between a spectrum of retinal vascular abnormalities measured

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

via fundus photography and the risk of dementia.[10-12,33] However, most studies measured retinal signs by semi-automated software, requiring human identification on the basis of prespecified protocols, which might introduce intra- and inter-variability. Besides, recent systematic reviews indicated that combination of multiple retinal vascular parameters, rather than individual marker, might provide higher prognostic value.[34,35] The present study utilized artificial intelligence technique, which might exhibit notable advantages in these issues. Artificial intelligence operates in absence of human assessment, and even performs superiorly to ophthalmologists in capturing subtle retinal changes that would otherwise fail to attract human attention.[36] With faster, easier, more consistent and precise output, the artificial intelligence reduces variability and human cost, thus enhancing the clinical utility of retinal photography.[37,38] Moreover, artificial intelligence is able to fully extract and integrate multiple retinal features (including information beyond human existed perception or understanding) that are related to dementia risk.

Participants in the BRAVE were much older, and had a larger proportion of female. The significant demographic heterogeneity between the development dataset and the external validation dataset suggested the algorithm's robustness and promising wider utility. One application scenario for the algorithm is screening individuals with high dementia risk in community. Traditional dementia prediction models requiring cognitive tests or multidimensional risk factors increased application difficulties in

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

population-based settings. By contrast, fundus photography is easy to implement and timesaving. According to our practical experience in BRAVE, an investigator with no background on ophthalmology could take fundus photographs within one minute after a few hours of training. Besides, compared with risk factors like blood lipids or glucose, the retinal images have no requirement for fasting status, with less fluctuation and can be obtained noninvasively, thus facilitating the acceptability and convenience of participants. In addition, the algorithm could also be recommended as an add-on to routine screening for diabetic retinopathy, given that patients with diabetes were significantly associated with higher risk of cognitive decline and dementia.[28] Moreover, our algorithm has potential utility in assessing pre-test dementia probability for further diagnostic tests in outpatient clinics. Last but not the least, this algorithm might also be adopted in dementia clinical trials, incorporated as inclusion criteria to efficiently target eligible participants, or surrogate outcome which could be observed expediently.[32,39]

Previous studies have investigated deep learning algorithm based on fundus photographs for screening cardiovascular diseases and anaemia,[13-15,40] our study added novel evidence regarding dementia in this field, potentially facilitating the eventual application of fundus photography for simultaneous screening of multiple diseases in large population-based settings. The foremost strength of the present work was employing convolutional neural network to deal with large dataset of fundus images. The development dataset

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

contained 579,880 fundus images of 258,305 individuals from 19 province-level administrative regions of China, the convolutional neural network exhibited distinct advantages in processing such large dataset, by extracting information from images with a deep architecture, which was similar to image process in human brain.[41] Another strength was incorporating external validation cohort with varied demographic characteristics and comprehensive cognitive tests, the results externally validated the performance and further supported the scientificalness of the algorithm.

There were, however, also limitations in our study. First, the CAIDE dementia risk score was derived from cross-sectional data, future investigations based on incident dementia events in longitudinal settings are warranted to further verify the predictive performance of the algorithm. Second, the R^2 in the external validation dataset was relatively small, probably due to the distinct age difference between the development and external validation dataset, given that age is the most important factor for dementia and cognitive function. Another possible explanation could be the absence of physical inactivity and educational level in the development dataset, which was conducted in medical check-up settings mainly collecting health measurements rather than lifestyle or socioeconomic factors. Despite that, we have imputed the educational level by national census data, and the algorithm performed well in identifying individuals with high dementia risk. Third, the present study only included Chinese participants, which might limit the

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

generalization of our algorithm to other ethnicities.

Conclusions

The present study demonstrated that a deep learning algorithm based on fundus photographs could well identify individuals with high dementia risk, and hold promise for wider application in community-based screening or clinic. Future studies based on longitudinal settings are warranted to further validate the predictive performance of the algorithm.

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Author Contributors: Wuxiang Xie and Yuzhong Chen had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Wuxiang Xie.

Acquisition, analysis, or interpretation of data: Rong Hua, Jianhao Xiong, Zongyuan Ge, Meng Fu, Bin Wang, Xin Zhao, Chao He, Yuzhong Chen, Li Dong, Zhiqiang Ma, Zhaohui Wang, Wenbin Wei, Wuxiang Xie.

Drafting of the manuscript: Rong Hua, Jianhao Xiong.

Critical revision of the manuscript for important intellectual content: Gail Li, Yuzhong Chen, Wuxiang Xie.

Statistical analysis: Rong Hua, Yidan Zhu, Yanjun Ma, Chenglong Li.

Obtained funding: Wuxiang Xie.

Administrative, technical, or material support: Wuxiang Xie.

Supervision: Wuxiang Xie.

Conflict of Interest Disclosure: Jianhao Xiong, Zongyuan Ge, Meng Fu, Bin Wang, Xin Zhao, Chao He, and Yuzhong Chen are employees of Beijing Eaglevision Technology Development Co., Ltd. All other authors declare no competing interests.

Funding/Support: The present study was supported by the National Natural Science Foundation of China (project no. 81974489) and 2019 Irma and Paul Milstein Program for Senior Health Research Project Award.

Role of the Funder/Sponsor: The funding sources had no role in the design

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Acknowledgments: We thank all participants in the health checkup population involved in the present study and the Beijing Research on Ageing and Vessel.

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Reference

1. Livingston G, Huntley J, Sommerlad A, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet.* 2020;396(10248):413-446.

2. Sperling RA, Aisen PS, Beckett LA, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & dementia : the journal of the Alzheimer's Association.* 2011;7(3):280-292.

3. Kivipelto M, Mangialasche F, Ngandu T. Lifestyle interventions to prevent cognitive impairment, dementia and Alzheimer disease. *Nature Reviews Neurology.* 2018;14(11):653-666.

4. Kivipelto M, Ngandu T, Laatikainen T, Winblad B, Soininen H, Tuomilehto J. Risk score for the prediction of dementia risk in 20 years among middle aged people: a longitudinal, population-based study. *Lancet Neurol.* 2006;5(9):735-741.

5. Exalto LG, Quesenberry CP, Barnes D, Kivipelto M, Biessels GJ, Whitmer RA. Midlife risk score for the prediction of dementia four decades later. *Alzheimers Dement.* 2014;10(5):562-570.

6. Sindi S, Calov E, Fokkens J, et al. The CAIDE Dementia Risk Score App: The development of an evidence-based mobile application to predict the risk of dementia. *Alzheimers Dement (Amst).*

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

2015;1(3):328-333.

7. Cheung CYL, Ikram MK, Chen C, Wong TY. Imaging retina to study dementia and stroke. *Prog Retin Eye Res.* 2017;57:89-107.

8. Patton N, Aslam T, Macgillivray T, Pattie A, Deary IJ, Dhillon B. Retinal vascular image analysis as a potential screening tool for cerebrovascular disease: a rationale based on homology between cerebral and retinal microvasculatures. *J Anat.* 2005;206(4):319-348.

9. Bernardes R, Serranho P, Lobo C. Digital Ocular Fundus Imaging: A Review. *Ophthalmologica.* 2011;226(4):161-181.

10. Lesage SR, Mosley TH, Wong TY, et al. Retinal microvascular abnormalities and cognitive decline The ARIC 14-year follow-up study. *Neurology.* 2009;73(11):862-868.

11. de Jong FJ, Schrijvers EM, Ikram MK, et al. Retinal vascular caliber and risk of dementia: the Rotterdam study. *Neurology.* 2011;76(9):816-821.

12. Deal JA, Sharrett AR, Albert M, et al. Retinal signs and risk of incident dementia in the Atherosclerosis Risk in Communities study. *Alzheimers Dement.* 2019;15(3):477-486.

13. Poplin R, Varadarajan AV, Blumer K, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nature Biomedical Engineering.* 2018;2(3):158-164.

14. Cheung CY, Xu D, Cheng C-Y, et al. A deep-learning system for the assessment of cardiovascular disease risk via the measurement of

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

retinal-vessel calibre. *Nature Biomedical Engineering.* 2021;5(6):498-508.

15. Rim TH, Lee CJ, Tham Y-C, et al. Deep-learning-based cardiovascular risk stratification using coronary artery calcium scores predicted from retinal photographs. *The Lancet Digital Health.* 2021;3(5):e306-e316.
16. Ma YN, Xie WX, Hou ZH, et al. Association between coronary artery calcification and cognitive function in a Chinese community-based population. *J Geriatr Cardiol.* 2021;18(7):514-522.
17. Tubulation on the 2010 population census of the People's Republic of China. http://www.stats.gov.cn/english/Statisticaldata/CensusData/rkpc2010/indexch.htm. Published 2010. Accessed July 22, 2021.
18. Szegedy C, Loffe S, Vanhoucke V, AA A. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. Thirty-First Aaai Conference on Artificial Intelligence; 2017.
19. DeHoog E, Schwiegerling J. Optimal parameters for retinal illumination and imaging in fundus cameras. *Appl Optics.* 2008;47(36):6769-6777.
20. Xiong JH, He AW, Fu M, et al. Improve Unseen Domain Generalization via Enhanced Local Color Transformation. International Conference on Medical Image Computing and Computer-Assisted Intervention; 2020.
21. Keras. https://keras.io. Accessed Feb 29, 2020.
22. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

R and S plus to analyze and compare ROC curves. *Bmc Bioinformatics.* 2011;12.

23. Lai C-Y, Tian L, Schisterman EF. Exact confidence interval estimation for the Youden index and its corresponding optimal cut-point. *Comput Stat Data Anal.* 2012;56(5):1103-1114.

24. Chen KL, Xu Y, Chu AQ, et al. Validation of the Chinese Version of Montreal Cognitive Assessment Basic for Screening Mild Cognitive Impairment. *J Am Geriatr Soc.* 2016;64(12):E285-E290.

25. Huang L, Chen KL, Lin BY, et al. Chinese version of Montreal Cognitive Assessment Basic for discrimination among different severities of Alzheimer's disease. *Neuropsych Dis Treat.* 2018;14:2133-2140.

26. Hua R, Ma YJ, Li CL, Zhong BL, Xie WX. Low levels of low-density lipoprotein cholesterol and cognitive decline. *Science Bulletin.* 2021;16:1684-1690.

27. Xie WX, Zheng FF, Yan L, Zhong BL. Cognitive Decline Before and After Incident Coronary Events. *Journal of the American College of Cardiology.* 2019;73(24):3041-3050.

28. Zheng FF, Yan L, Yang ZC, Zhong BL, Xie WX. HbA(1c), diabetes and cognitive decline: the English Longitudinal Study of Ageing. *Diabetologia.* 2018;61(4):839-848.

29. Wei MQ, Shi J, Li T, et al. Diagnostic Accuracy of the Chinese Version of the Trail-Making Test for Screening Cognitive Impairment. *J Am Geriatr*

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Soc. 2018;66(1):92-99.

30. Ma Y, Liang L, Zheng F, Shi L, Zhong B, Xie W. Association Between Sleep Duration and Cognitive Decline. *JAMA Netw Open.* 2020;3(9):e2013573-e2013573.

31. Kaffashian S, Dugravot A, Elbaz A, et al. Predicting cognitive decline: a dementia risk score vs. the Framingham vascular risk scores. *Neurology.* 2013;80(14):1300-1306.

32. Ngandu T, Lehtisalo J, Solomon A, et al. A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. *Lancet.* 2015;385(9984):2255-2263.

33. Schrijvers EMC, Buitendijk GHS, Ikram MK, et al. Retinopathy and risk of dementia The Rotterdam Study. *Neurology.* 2012;79(4):365-370.

34. McGrory S, Cameron JR, Pellegrini E, et al. The application of retinal fundus camera imaging in dementia: A systematic review. *Alzheimers Dement (Amst).* 2016;6:91-107.

35. Wagner SK, Fu DJ, Faes L, et al. Insights into Systemic Disease through Retinal Imaging-Based Oculomics. *Transl Vis Sci Technol.* 2020;9(2):6-6.

36. Son J, Shin JY, Kim HD, Jung KH, Park KH, Park SJ. Development and Validation of Deep Learning Models for Screening Multiple Abnormal

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Findings in Retinal Fundus Images. *Ophthalmology.* 2020;127(1):85-94.

37. Ting DSW, Cheung CYL, Lim G, et al. Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations With Diabetes. *Jama-J Am Med Assoc.* 2017;318(22):2211-2223.

38. Gulshan V, Peng L, Coram M, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *Jama-J Am Med Assoc.* 2016;316(22):2402-2410.

39. Reid CM, Storey E, Wong TY, et al. Aspirin for the prevention of cognitive decline in the elderly: rationale and design of a neuro-vascular imaging study (ENVIS-ion). *BMC Neurol.* 2012;12:3-3.

40. Mitani A, Huang A, Venugopalan S, et al. Detection of anaemia from retinal fundus images via deep learning. *Nature Biomedical Engineering.* 2020;4(1):18-27.

41. Anwar SM, Majid M, Qayyum A, Awais M, Alnowami M, Khan MK. Medical Image Analysis using Convolutional Neural Networks: A Review. *J Med Syst.* 2018;42(11):226.

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure legends:

Figure 1. Estimation of CAIDE dementia risk score in the internal (a) and external (b) validation.

External validation was conducted using data from the BRAVE. The adjusted R^2 (coefficient of determination) between the estimated CAIDE dementia risk score and actual CAIDE dementia risk score is presented.

Abbreviation: BRAVE = Beijing Research on Ageing and Vessel.

Figure 2. Algorithm performance for identifying participants with high dementia risk in the internal validation dataset (a) and in the BRAVE (b).

Individuals with high dementia risk were defined as CAIDE dementia risk score ≥10 points. The points on line indicate the maximum Youden index.

Abbreviation: AUC = area under the receiver operating characteristic curve.

Table 1. Characteristics of individuals in development, internal validation, and external validation datasets

Characteristics	Development dataset		Internal validation dataset		External validation dataset	
	Men	Women	Men	Women	Men	Women
No. of images	306,795	273,085	16,124	14,246	1,122	1,902
No. of participants	136,157	122,148	7,120	6,439	561	951
Age (years)	42.3±13.6	41.8±13.8	42.2±13.6	41.9±13.9	60.7±7.1	59.3±7.3
Systolic blood pressure (mm Hg)	124.1±16.3	116.2±17.8	124.5±16.6	116.1±17.9	136.1±16.1	130.3±17.4
Total cholesterol (mmol/L)	4.9±0.9	4.9±1.0	4.9±1.0	4.9±1.0	4.8±0.9	5.0±0.9
Body mass index (kg/m^2)	25.0±3.4	22.8±3.4	25.0±3.5	22.7±3.4	26.2±3.2	25.8±3.5
Education ≥10 years (%) [a]	-	-	-	-	328 (58.5)	576 (60.6)
Physical inactive (%) [b]	-	-	-	-	500 (89.1)	739 (77.7)

Data are presented as mean ± SD or n (%).
[a] Educational level was not available in development or internal validation datasets.
[b] Physical inactive was not available in development or internal validation datasets.

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 2. Association between estimated CAIDE dementia risk score and different cognitive assessments: using multiple linear regression models

Cognitive assessments	Model 1[a]		Model 2[b]	
	***β* (95% CI)**	***P***	***β* (95% CI)**	***P***
MoCA score	−0.240 (−0.419, −0.060)	0.009	−0.324 (−0.510, −0.138)	0.001
Memory test score	−0.317 (−0.488, −0.146)	<0.001	−0.392 (−0.570, −0.214)	<0.001
Verbal fluency test score	−0.306 (−0.519, −0.092)	0.005	−0.355 (−0.577, −0.132)	0.002
TMT-A time (minutes)	0.039 (0.036, 0.139)	0.005	0.043 (0.015, 0.071)	0.003
TMT-B time (minutes)	0.048 (0.017, 0.079)	0.002	0.060 (0.028, 0.092)	<0.001

[a] Model 1 adjusted for age (years) and sex.
[b] Model 2 adjusted for age (years), sex, BMI (kg/m^2), TC (mmol/L), educational level, marriage status, drinking status, smoking status, physical inactivity, depressive symptoms, *APOE* ε4 status, hypertension, diabetes, and self-reported diagnosis of coronary heart disease, stroke, cancer and chronic obstructive pulmonary disease.
Note: The lower score of MoCA, memory test and verbal fluency indicate worse performance. The longer time of TMT-A, TMT-B represent worse performance.

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 3. Association between quartiles of estimated CAIDE dementia risk score and different cognitive assessments: using analysis of covariance

Estimated CAIDE dementia risk score	MoCA score		Memory test score		Verbal fluency test score		TMT-A time (minutes)		TMT-B time (minutes)	
	***β* (95% CI)**	***P***	***β* (95% CI)**	***P***	***β* (95% CI)**	***P***	***β* (95% CI)**	***P***	***β* (95% CI)**	***P***
Model 1[a]										
Quartile 1	0.000 (ref.)	/	0.000 (ref.)	/	0.000 (ref.)	/	0.000 (ref.)	/	0.000 (ref.)	/
Quartile 2	−0.314 (−0.825, 0.198)	0.229	−0.563 (−1.050, −0.075)	0.024	−0.150 (−0.760, 0.460)	0.630	−0.015 (−0.092, 0.063)	0.714	0.025 (−0.064, 0.114)	0.581
Quartile 3	−0.564 (−1.132, 0.004)	0.051	−0.466 (−1.007, 0.076)	0.092	−0.415 (−1.093, 0.262)	0.230	0.063 (−0.023, 0.149)	0.153	0.063 (−0.035, 0.162)	0.208
Quartile 4	−0.724 (−1.468, −0.208)	0.009	−0.979 (−1.579, −0.378)	0.001	−0.675 (−1.426, 0.077)	0.078	0.124 (0.028, 0.220)	0.011	0.118 (0.009, 0.227)	0.034
Test for linear trend	−0.274 (−0.472, −0.075)	0.007	−0.271 (−0.460, −0.081)	0.005	−0.235 (−0.472, 0.001)	0.051	0.049 (0.019, 0.080)	0.001	0.041 (0.006, 0.075)	0.021
Model 2[b]										
Quartile 1	0.000 (ref.)	/	0.000 (ref.)	/	0.000 (ref.)	/	0.000 (ref.)	/	0.000 (ref.)	/
Quartile 2	−0.554 (−1.062, −0.046)	0.032	−0.731 (−1.217, −0.244)	0.003	−0.340 (−0.949, 0.269)	0.274	0.092 (−0.059, 0.095)	0.648	0.075 (−0.012, 0.162)	0.092
Quartile 3	−0.711 (−1.281, −0.141)	0.015	−0.565 (−1.110, −0.019)	0.043	−0.416 (−1.099, 0.267)	0.232	0.062 (−0.024, 0.148)	0.159	0.078 (−0.019, 0.176)	0.117
Quartile 4	−0.984 (−1.625, −0.343)	0.003	−1.077 (−1.691, −0.463)	0.001	−0.654 (−1.423, 0.114)	0.095	0.111 (0.014, 0.208)	0.025	0.119 (0.010, 0.229)	0.033
Test for linear trend	−0.294 (−0.497, −0.091)	0.005	−0.284 (−0.479, −0.089)	0.004	−0.196 (−0.440, 0.048)	0.115	0.039 (0.008, 0.070)	0.012	0.034 (−0.001, 0.069)	0.057

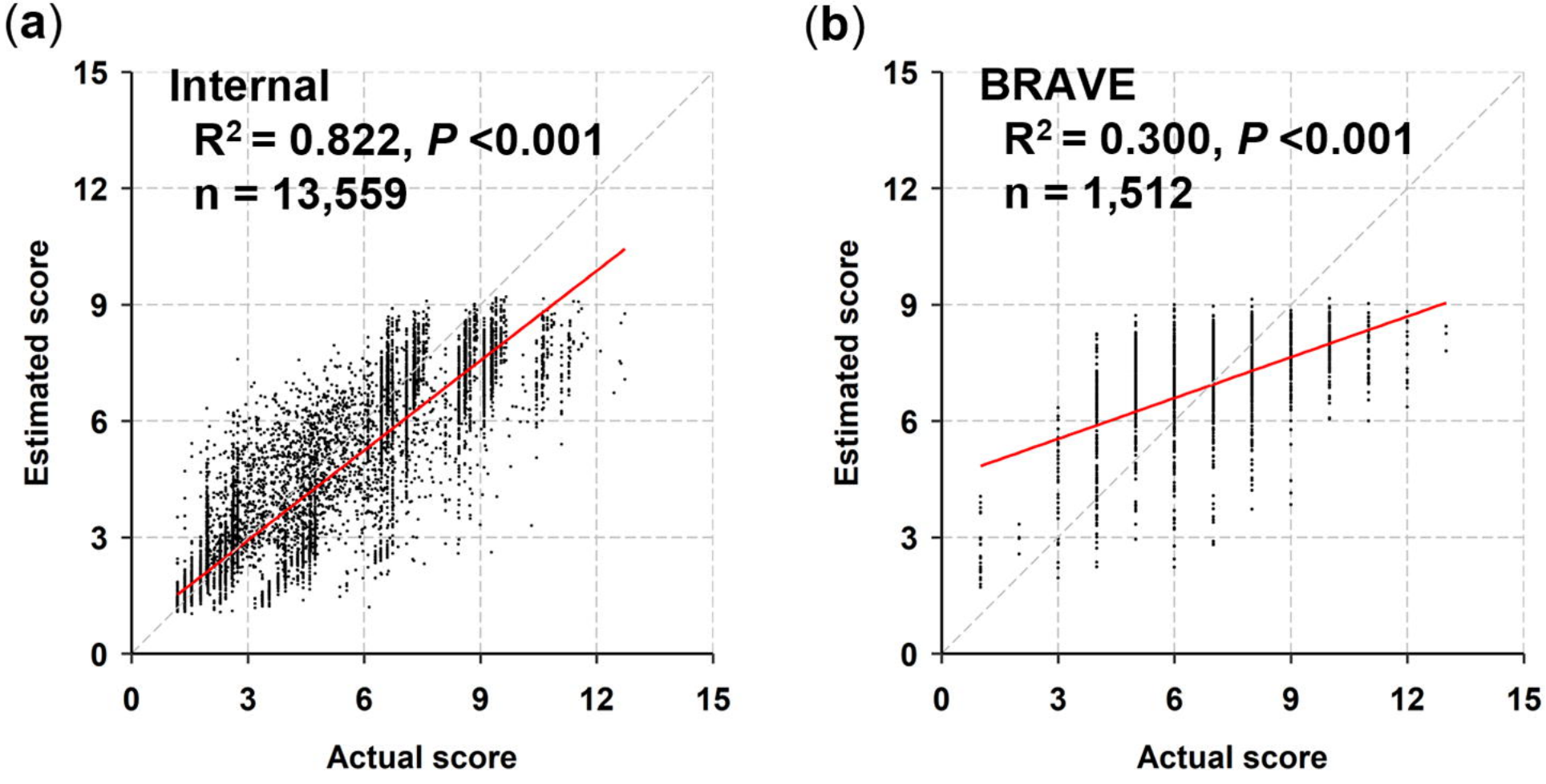
[a] Model 1 adjusted for age (years) and sex.
[b] Model 2 adjusted for age (years), sex, BMI (kg/m^2), TC (mmol/L), educational level, marriage status, drinking status, smoking status, physical inactivity,

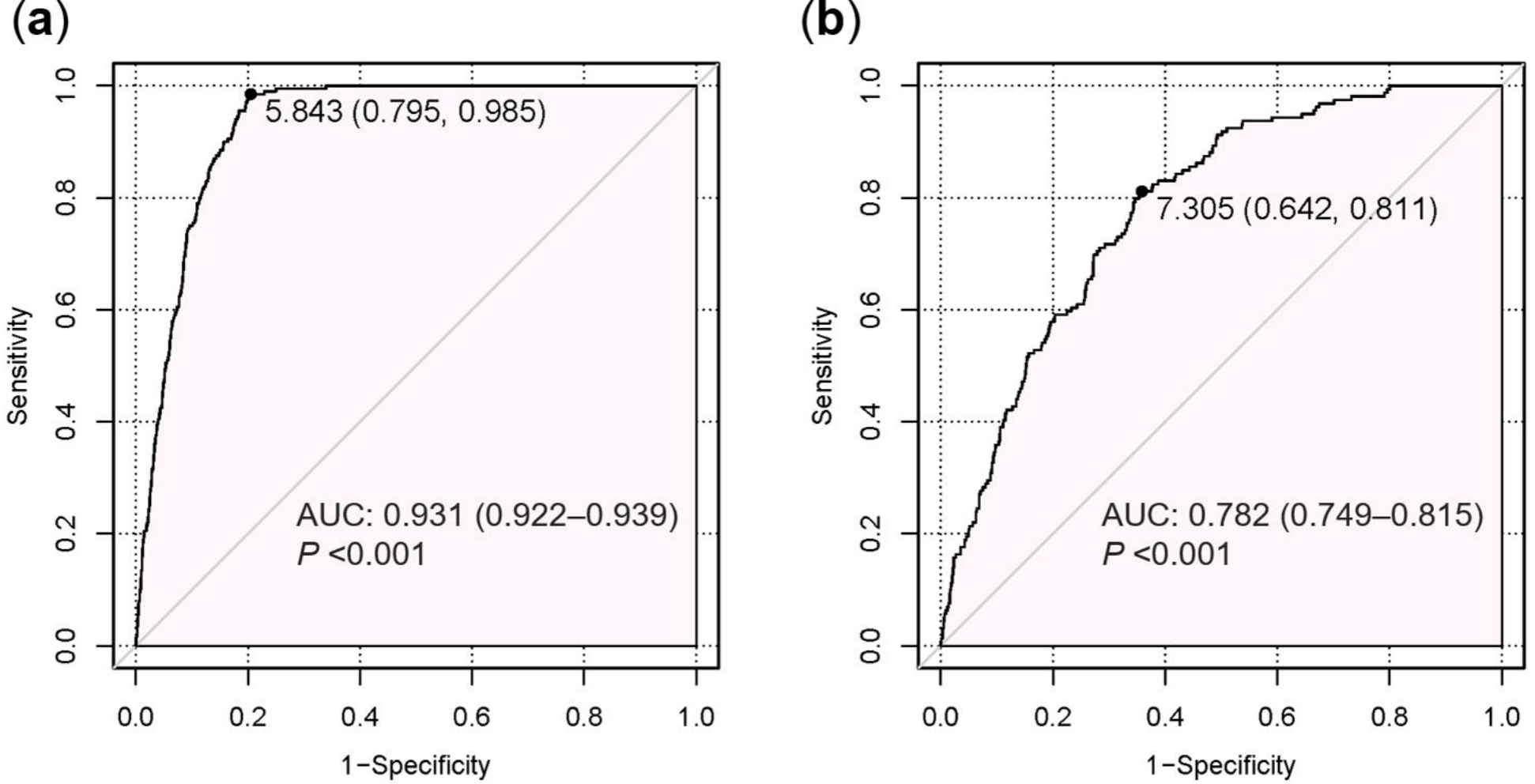
medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

depressive symptoms, *APOE* ε4 status, hypertension, diabetes, and self-reported diagnosis of coronary heart disease, stroke, cancer and chronic obstructive pulmonary disease.

Note: The lower score of MoCA, memory test and verbal fluency indicate worse performance. The longer time of TMT-A, TMT-B represent worse performance.

medRxiv preprint doi: https://doi.org/10.1101/2021.08.17.21262156; this version posted August 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.


(a)
Internal
R² = 0.822, P <0.001
n = 13,559
Estimated score
Actual score
0
3
6
9
12
15
(b)
BRAVE
R² = 0.300, P <0.001
n = 1,512
Estimated score
Actual score
0
3
6
9
12
15


(a)
5.843 (0.795, 0.985)
AUC: 0.931 (0.922–0.939)
P <0.001
Sensitivity
1−Specificity
0.0
0.2
0.4
0.6
0.8
1.0
(b)
7.305 (0.642, 0.811)
AUC: 0.782 (0.749–0.815)
P <0.001
Sensitivity
1−Specificity