Original research

Brain age gap in neuromyelitis optica spectrum disorders and multiple sclerosis

Ren Wei,[1] Xiaolu Xu ,[1] Yunyun Duan ,[1] Ningnannan Zhang,[2] Jie Sun,[2] Haiqing Li,[3] Yuxin Li,[3] Yongmei Li,[4] Chun Zeng,[4] Xuemei Han,[5] Fuqing Zhou,[6] Muhua Huang,[6] Runzhi Li,[7] Zhizheng Zhuo,[1] Frederik Barkhof,[8,9] James H Cole ,[9,10] Yaou Liu [1]

Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/jnnp-2022-329680).

For numbered affiliations see end of article.

Correspondence to
Professor Yaou Liu, Department of Radiology, Beijing Tiantan Hospital, Beijing, China; yaouliu80@163.com

RW and XX contributed equally.

Received 27 June 2022
Accepted 12 September 2022

To cite: Wei R, Xu X, Duan Y, *et al*. *J Neurol Neurosurg Psychiatry* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/jnnp-2022-329680

In this study, we used a novel deep-learning brain age model to investigate the utility of BAG as a neuroimaging biomarker to predict EDSS worsening in NMOSD and RRMS in a large multicentre dataset.

METHODS

Participants

Data from patients with NMOSD and RRMS were retrospectively collected from six tertiary neurological centres in China covering the period between November 2009 and April 2018. Patients who fulfilled the following criteria were included: (a) confirmed diagnosis of NMOSD according to 2015 revised diagnostic criteria[8] or RRMS according to 2017 McDonald criteria[9]; (b) complete demographic and clinical information, including baseline EDSS score and disease duration and (c) good quality baseline three-dimensional (3D) T1-weighted structural images (T1WI). Clinical evaluation, diagnosis, treatment and follow-up assessments of the participants were conducted at each centre by local neurologists with expertise in demyelinating diseases. EDSS worsening was defined as an increase in EDSS score ≥1.0 for baseline EDSS ≤5.5 or an increase in EDSS score ≥0.5 for baseline EDSS >5.5, as previously published.[10]

Data for deep learning model training

Training data for our deep learning-derived brain age included MRI scans from healthy controls (HCs, n=9794) from publicly available datasets, including Alzheimer's Disease Neuroimaging Initiative (ADNI), The Australian Imaging, Biomarkers and Lifestyle (AIBL),[11] Brain Genomics Superstruct Project (GSP)[12] and Southwest University Longitudinal Imaging Multimodal (SLIM),[13] as well as a group of healthy people scanned at Beijing Tiantan Hospital from January to December 2019 (online supplemental table 1, online supplemental figure 1). After training, the model was tested on two further independent datasets. Internal validation data comprised another group of healthy participants (n=462) scanned at Beijing Tiantan Hospital from January to April 2020 on two different scanners (see online supplemental table 1). The external validation dataset included HCs from the multicentre NMOSD and MS cohorts (n=267).

Image acquisition and data preprocessing

All the MRI scans of participants as well as the validation dataset were acquired using 3.0 T scanners at 1.0 mm isotropic resolution using Magnetization Prepared-RApidGradient Echo imaging (MP-RAGE) or similar sequences. Non-contrast 3D T1-weighted scans were affinely registered to Montreal Neurological Institute (MNI) space. Skull stripping was performed by HD-BET on the registered scans.[14] The signal intensity of the resulting images was normalised by dividing by the mean intensity within the cerebral mask. Scans were then resampled to 1 mm isotropic resolution using linear interpolation and served as the input of the proposed convolutional neural network (CNN).

Age at each scan was determined by either of two methods: (1) the demographic metadata (in years) provided by owners of the dataset; (2) calculated from the difference between date of birth and image acquisition date recorded in DICOM metadata, which was done in days and converted to years. Inconsistent data were omitted from the study.

Brain volume measurement

Brain volume segmentation was performed using the automated recon-all procedure in FreeSurfer package (V.6.0.0) as described by Fischl .[15] The total brain volume was calculated and normalised by dividing by the estimated total intracranial volume.[16]

Model construction, training and prediction

We built a 3D CNN called the 3D Simple Fully Convolutional Neural Network (SFCN) network as per the work of Peng .[17] We modified the output structure so that the network could predict age across a larger range of 6–90 years. Model training and mathematical details are described in the online supplementary material.

BAG was calculated by subtracting chronological age from predicted brain age, with a positive BAG indicating an older-looking brain. To investigate the possible influence of brain lesions on age prediction, we performed a correlation analysis between raw and lesion-filled 3D T1WI images. Lesion filling was performed by default pipeline of Lesion Segmentation Tool (V.3.0.0, https://www.applied-statistics.de/lst.html).

Statistical analysis

Statistical analyses were conducted using R (V.3.6.3). Graphs were plotted with ggplot2 package. Intergroup comparison was conducted using the χ^2 test (for categorical variables), Wilcoxon signed-rank test (for EDSS) and Student's t-test or analysis of variance with Tukey's range test as post hoc analysis (for continuous variable). Survival analysis with Kaplan-Meier curve and Cox proportional hazards model were used to analyse time-to-progression data. Other details are described in the online supplementary material. All statistical tests were two-sided, and $p<0.05$ was considered statistically significant.

RESULTS

Participants

In total, 199 patients with NMOSD, 200 patients with RRMS and 269 age-matched and sex-matched HC subjects were included (table 1). Patients with NMOSD were older at baseline (41.0±13.0 years vs 37.1±11.4 years, p=0.005), had a longer disease duration (4.5±5.1 years vs 3.2±4.4 years, p=0.006) and had less severe disability measured by EDSS at baseline (2.0 vs 3.5, p<0.001) than patients with RRMS. Of the patients with NMOSD included, 52 (26.1%) patients received disease-modifying therapy (DMT), others received immunosuppressants including cyclophosphamide and azathioprine. In the RRMS group, 86 (43.0%) patients received an MS-specific DMT, others received the above other treatment.

Follow-up data were available for 85 patients with NMOSD and 124 patients with RRMS (median follow-up duration: 5.8±1.9 years and 5.2±1.7 years, respectively). During follow-up, 31 patients with NMOSD and 42 patients with RRMS experienced EDSS worsening.

Brain morphometry of the participants

Both the NMOSD and RRMS groups had lower brain volumes than the HCs (1080.1±121.5 mL and 1058.9±94.4 mL vs 1154.6±98.5 mL, both p<0.001). While the NMOSD and RRMS groups were not significantly different in raw brain volume (p=0.108), normalised brain volumes revealed less pronounced atrophy in patients with NMOSD (0.750±0.038 vs 0.731±0.045, p<0.001). The lesion load in the NMOSD group was lower than that in the RRMS group (4.9±8.1 mL vs 12.7±17.9 mL, p<0.001) (table 1).

Table 1 Demographic characteristics, baseline status and deep learning-derived brain age of participants

	NMOSD	RRMS	HCs	P value
Baseline				
N	199	200	269	
Age at baseline, year (min–max)	41.0±13.0 (16.9–66.0)	37.1±11.4 (16.6–66.9)	38.5±12.7 (17.0–69.0)	NMOSD versus HC 0.071 RRMS versus HC 0.468 NMOSD versus RRMS 0.005
Female, n (%)	176/199 (88.4)	128/200 (64.0)	152/269 (56.5)	<0.001
Seropositive for AQP4-IgG, n (%)	84/132 (63.6)	–	–	–
First onset to diagnosis, year (min–max)	4.5±5.1 (0.0–35.0)	3.2±4.4 (0.0–21.0)	–	0.006
Baseline use of DMT, n (%)	52 (26.1%)	86 (43.0%)	–	–
EDSS at baseline, median (IQR) (min–max)	2.0 (2.0) (0.0–9.0)	3.5 (3.0) (0.0–9.0)	–	<0.001
Brain segmentation volume without ventricles, mL (min–max)	1058.9±94.4 (798.7–1390.1)	1080.1±121.5 (742.6–1484.5)	1154.6±98.5 (910.7–1434.0)	NMOSD versus HC <0.001 MS versus HC <0.001 NMOSD versus RRMS 0.108
Normalised brain volume (min–max)	0.750±0.038 (0.647–0.891)	0.731±0.045 (0.590–0.858)	0.765±0.030 (0.700–0.894)	<0.001*
Total volume of lesion, mL (min–max)	4.9±8.1 (0.0–43.9)	12.7±17.9 (0.0–134.0)	–	<0.001*
Deep learning-derived brain age				
Predicted brain age, year (min–max)	46.4±16.0 (18.8–77.5)	49.8±17.5 (19.5–77.8)	39.3±13.7 (14.8–73.8)	<0.001*
Brain age gap, year (95% CI)	5.4±8.2 (4.3 to 6.5)	13.0±14.7 (10.9 to 15.0)	0.8±6.2 (0.1 to 1.6)	<0.001*
Predicted brain age SD, year (95% CI)	6.0±3.0 (5.6 to 6.5)	7.2±4.2 (6.6 to 7.8)	4.8±1.1 (4.7 to 4.9)	<0.001*
Follow-up				
N with follow-up data, n (%)	85 (42.7)	124 (62.0)	–	
Mean follow-up time, year (min–max)	5.8±1.9 (1.9–9.9)	5.2±1.7 (1.5–9.2)	–	0.020
EDSS worsening, n (%)	31 (36.5)	42 (33.9)	–	0.764

Continuous variables other than EDSS are reported as the mean±SD. EDSS is reported as the median (IQR).
*For all pairwise comparisons, that is, for NMOSD versus HC, RRMS versus HC and NMOSD versus RRMS.
DMT, disease-modifying therapy; EDSS, Expanded Disability Status Scale; HC, healthy control; MS, multiple sclerosis; NMOSD, neuromyelitis optica spectrum disorder; RRMS, relapsing-remitting multiple sclerosis.

Performance of the brain age prediction model

Model training (using 9794 HCs) was terminated at epoch 108. The mean absolute error (MAE) before inverse linear bias correction was 2.63 years in the developmental validation set, and this model was used as the final model for further analysis.

The model was then tested using 462 images for internal (across-scanner) validation and 267 images for external validation (across-centre). In the internal validation dataset, the MAE was 2.9±3.1 years, with no significant difference across scanner types (p=0.581, n=2). The Pearson's correlation coefficient (r) between age and brain age was 0.957. In the external validation set, the MAE was 4.5±3.9 years, and the Pearson's r was 0.890. The MAE was not significantly different across different centres (p=0.660, n=5; online supplemental table 2).

Increased BAG in NMOSD and RRMS compared with healthy controls

The difference in BAG among patients with NMOSD, patients with RRMS and HCs was relatively consistent across baseline chronological ages (figure 1A). At baseline, patients with NMOSD had a significantly higher BAG than HCs (NMOSD−HC=4.6 years, 95% CI 2.4 to 6.9, p<0.001), but patients with RRMS had a markedly higher BAG than HCs (MS−HC=12.1 years, 95% CI 9.9 to 14.3, p<0.001). BAG was lower in NMOSD than in RRMS (NMOSD−RRMS=−7.5 years, 95% CI 5.2 to 9.9, p<0.001) (table 1, figure 1B).

Furthermore, we performed subgroup analyses of BAG in AQP4 seropositive versus seronegative patients with NMOSD, as well as in patients with NMOSD with versus without brain lesions. We observed that there was no significant difference in BAG between the AQP4 seropositive and seronegative subgroups (5.8±8.8 vs 4.2±6.9 years, p=0.256). However, the BAG in patients with brain lesions was significantly higher than those without (7.1±8.5 vs 3.4±7.2 years, p=0.001) (online supplemental table 5).

A significant difference in BAG across centres (p<0.001) was noted, although post hoc analysis revealed consistent trends in disease effects on BAG in all six centres (figure 1C). Sample images and the corresponding output from both the NMOSD and RRMS groups were provided for better understanding (figure 1D–G).

The correlation between raw and lesion-filled 3D T1WI images was very high (R^2=0.984, p<0.001, online supplemental figure 3A). A Bland-Altman plot showed that the mean difference between raw and lesion-filled brain age was 0.28±2.11 years with no apparent systematic bias (online supplemental figure 3B), indicating that the lesion filling process did not have a particular impact on the model.

Correlation of BAG with clinical variables

At baseline, univariate linear regression analysis demonstrated that BAG was positively associated with EDSS in both the NMOSD and RRMS groups (NMOSD r=0.217, β=0.86, p=0.002; RRMS r=0.268, β=2.31, p<0.001; figure 2A). Normalised brain volume was inversely associated with BAG in both NMOSD and RRMS groups (NMOSD r=−0.202, β=−48.5, p<0.001; RRMS r=−0.384, β=−126.9, p<0.001; figure 2B). Multivariable linear regression found that BAG was positively predictive of baseline EDSS independent of normalised brain volume and disease duration (NMOSD p=0.030; RRMS p=0.009; online supplemental table 3).

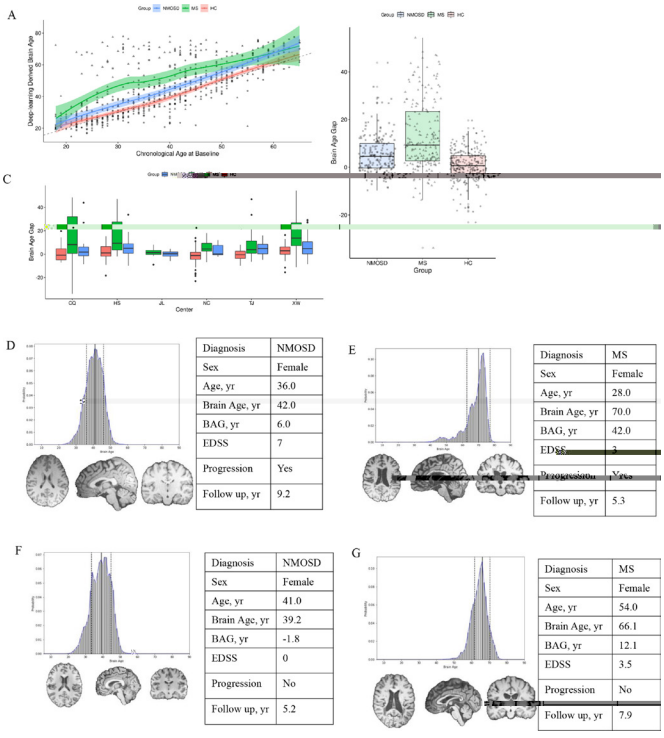

Figure 1 Deep learning-derived brain age versus chronological age in neuromyelitis optica spectrum disease (NMOSD), multiple sclerosis (MS) and healthy control (HC). (A) Deep learning-derived brain age versus chronological age in NMOSD, MS and HC groups. Predicted brain age is consistently higher in NMOSD and MS groups compared with HC group. (B) Patients with NMOSD exhibits lower brain age gap (BAG) over MS and lower BAG over HCs. (C) The difference of BAG across centres in NMOSD, MS and HC groups. The tendency that MS BAG>NMOSD BAG>HC BAG remains consistent even if there are significant differences across centres. (D, E, F, G) A sample input and prediction result of patients with NMOSD and MS. Solid line indicates brain age estimation and dashed lines indicate SD of prediction. The predicted brain age was 42.0 5.1 years for (D) and 70.0 7.2 years for (E), yielding BAG of 6.0 years and 42.0 years namely. Both (D) and (E) experienced disability progression in follow-up sessions. Predicted brain age for (F) and (G) was 39.2 5.6 years and 66.1 4.4 years yielding BAG of 1.8 years and 12.1 years namely. These patients with lower BAG did not experience disability progression within follow-up period. EDSS, Expanded Disability Status Scale.

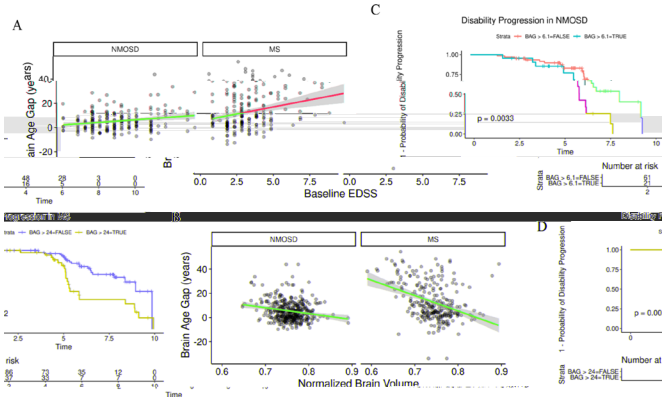

Figure 2 Correlation of brain age gap (BAG) with clinical variables and its prognostic value. (A) Increased BAG was associated with more severe baseline disability status in both neuromyelitis optica spectrum disease (NMOSD), multiple sclerosis (MS), which was more prominent in patients with MS. (B) Normalised brain volume was strongly negatively associated with BAG both in NMOSD and MS indicating possible contribution of atrophy in increased BAG. (C, D) Survival curve of BAG predicting disability progression in patients with NMOSD and MS. Cut-off point was determined by 80% specificity. Operating cut-off point for NMOSD is set to BAG >6.1 (sensitivity 38.7%, specificity 81.5%), MS is set to BAG >24.0 (sensitivity 50.0%, specificity 80.5%).

We performed 1:1 nearest neighbour propensity score matching (PSM)[18] to exclude the possible confounding influence of clinical variables on BAG. This matching yielded adequate balance for all included coefficients. The mean BAG was 5.0±7.1 years in NMOSD and 11.1±12.7 years in RRMS after adjustment for sex, age at diagnosis, baseline EDSS and normalised brain volume, with an estimated difference of −6.1 years (95% CI −8.7 to −3.4) years between NMOSD and RRMS (table 2).

The area under the curve of the receiver operating characteristic for BAG in predicting progression was 0.599 in NMOSD and 0.670 in RRMS. The optimal cut-off of BAG was 6.1 (sensitivity 38.7%, specificity 81.5%) for NMOSD and 24 (sensitivity 50.0%, specificity 80.5%) for RRMS (online supplemental figure 4). Kaplan-Meier survival analysis indicated that BAG was predictive of progression in both groups. For patients with NMOSD, the median time to progression for BAG >6.1 years was 5.79 years vs 7.99 years for BAG ≤6.1 years (p=0.003, figure 2C). The median time to progression for BAG >24.0 years was 5.36 years vs 8.95 years for BAG ≤24.0 years in patients with RRMS (p=0.002, figure 2D).

We used the Cox proportional hazards model to investigate whether BAG could be used to predict time to EDSS worsening independent of age at diagnosis, sex, disease duration, baseline EDSS and normalised brain volume. In univariate models, normalised brain volume and BAG were significantly associated with EDSS worsening in both patients with NMOSD and RRMS (table 3, univariate model). In a multivariable model, BAG was associated with EDSS worsening in patients with NMOSD (HR=1.02 (95% CI 1.00 to 1.04)). p=0.027, table 3), independent of normalised brain volume (p=0.158). However, neither normalised brain volume nor BAG was significant in the RRMS group in multivariable analysis. Interestingly, we found baseline EDSS to be negatively associated with EDSS worsening in NMOSD (multivariable model p=0.001, table 3).

Table 2 Patients with NMOSD exhibit lower brain age gap over RRMS adjusted for sex, age at diagnosis, baseline EDSS and normalised brain volume with propensity score matching

	NMOSD	RRMS	P value
N	119	119	–
Age at diagnosis, years	39.6±13.2	39.9±11.7	0.855
Female, n (%)	96 (80.7)	97 (81.5)	1.000

Analysis of predicted SD in brain age prediction

The predicted SD was positively associated with BAG in all three groups (linear model $p<0.001$ in HC and NMOSD, $p=0.011$ in RRMS, online supplemental figure 5A). The mean SD in NMOSD was higher than HC and lower than RRMS (online supplemental figure 5B), which was consistent with the trend seen in BAG, indicating a higher model uncertainty in those images with greater discrepancy between apparent and chronological age. We examined scans with high model uncertainty and found that some of them could be attributed to low image quality or incomplete anatomical coverage (online supplemental figure 5C), while others were not visually distinguishable from those with lower model uncertainty (online supplemental figure 5D). To analyse whether the difference in BAG was driven by the difference in predicted SD, we performed PSM with predicted SD added as a covariate. The difference in BAG between NMOSD and RRMS, as well as NMOSD and RRMS versus HC, remained statistically significant after PSM adjusted for age, sex, duration to diagnosis, baseline EDSS, normalised brain volume and predicted SD ($p<0.001$, online supplemental table 4).

DISCUSSION

In this study, we developed a deep learning model to accurately predict age from 3D structural MRI scans and demonstrated its robustness in the context of multiple centres and MRI scanners. Using this model, the BAG was estimated to be approximately +5 years in NMOSD and +13 years in RRMS. Baseline BAG was independently predictive of EDSS worsening in both NMOSD and RRMS, suggesting its additional clinical value as a non-invasive biomarker for early triage, stratified follow-up management and clinical trial enrolment.

Previous non-deep learning studies on age prediction tasks reported 2.9-year to 5.0-year MAEs on their validation sets[7 19 20] (some of which included multimodality-derived features, including functional MRI and diffusion tensor imaging), while deep learning studies reported validation MAEs as low as 2.14 years, such as in the original SFCN study.[17] We reached similar performance levels of MAE=2.5 years in the developmental validation set, and the performance was maintained in an internal test set, demonstrating the usefulness of our model and highlighting the versatility and potential of deep learning-based methods. We have also shown that the whole-brain CNN-based model was robust within scanners and centres, supporting the clinical use of the brain age paradigm.

BAG has been investigated extensively as a comprehensive biomarker for accelerated ageing. Increased BAG has been observed in dementia,[21] epilepsy[22] and traumatic brain injury.[23] We report for the first time the meaningfulness of BAG in NMOSD as well as the difference between NMOSD and RRMS. We found a BAG of 5.4 (95% CI 4.3 to 6.5) years in patients with NMOSD, which, although lower than RRMS, is still marked

indicates that the brains of patients with RRMS appear older than those of patients with NMOSD even at the same level of atrophy, implying that BAG can be seen as a global estimation that integrates information beyond simple brain volumetry while being more accessible and informative than tables of volumetric measurements.

The uncertainty and distributional pattern of predicted brain age is an important field of research that has attracted little attention. A recent study modelled brain age uncertainty with a single-layer neural network that addressed aleatoric uncertainty with quantile regression and epistemic uncertainty with the Monte Carlo dropout technique.[25] In contrast to other studies that use quantile regression, the novel method in our study renders aleatoric uncertainty a natural derivative since the model output itself is a distribution instead of the point estimate used in previous studies.[4] Epistemic uncertainty was not derived in this study due to computational cost. Although the uncertainty correlated positively with BAG, the PSM analysis indicated that the BAG difference between NMOSD and RRMS remained statistically significant even after adjustment for predicted SD. We observed that the predicted SD were higher in those scans without enough information for brain age inference (ie, low image quality, etc), and in those with a greater discrepancy between predicted and actual age. This observation suggests a potential use case for the predicted SD. The quantification of individual-level uncertainty in this way could provide an integrated, intuitive metric for image quality control, especially in healthy people, as well as provide a measure of 'confidence' for applications in clinical contexts.

Our study has a few limitations. First, the follow-up duration was relatively short, and the sample size of patients with follow-up was small, which may have introduced selection bias. Second, although previous studies have suggested the longitudinal utility of brain age in healthy cohorts[6] and accelerated ageing measured by BAG has been observed in MS cohorts,[7] our cohort lacked sufficient follow-up assessments for this type of analysis. Finally, the interpretability of the results needs to be further improved; specifically, the anatomical meaning of brain age remains ill-defined. Deep learning-based methods have been cast as 'black boxes'; however, tools such as class activation mapping, guided backpropagation and occlusion analysis are emerging that aim to extract mechanistic information from the network.[26] However, the translation of these methods to 3D data is complex, and they have yet to be validated for use in interpreting medical imaging data. Additionally, our study relied on 3D T1WI MRI, which is not always available in clinical contexts. Future work will take advantage of brain age models developed to work on routine clinical two-dimensional scans.[27]

In conclusion, NMOSD demonstrated a significant BAG compared with HCs, although less marked than RRMS. BAG is a predictive biomarker of EDSS worsening in both NMOSD and RRMS.

Author affiliations
[1]Department of Radiology, Beijing Tiantan Hospital, Beijing, China
[2]Department of Radiology and Tianjin Key Laboratory of Functional Imaging, Tianjin Medical University General Hospital, Tianjin, China
[3]Department of Radiology, Huashan Hospital Fudan University, Shanghai, China
[4]Department of Radiology, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China
[5]Department of Neurology, China-Japan Union Hospital of Jilin University, Changchun, China
[6]Department of Radiology, The First Affiliated Hospital of Nanchang University, Nanchang, China
[7]Department of Neurology, Beijing Tiantan Hospital, Beijing, China
[8]Department of Radiology and Nuclear Medicine, Neuroscience Campus Amsterdam, VU University Medical Centre Amsterdam, Amsterdam, The Netherlands
[9]Centre for Medical Image Computing, Department of Computer Science, University College London, London, UK
[10]Dementia Research Centre, Queen Square Institute of Neurology, University College London, London, UK

Twitter James H Cole @JamesCole_Neuro

Acknowledgements Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica; Biogen; Bristol-Myers Squibb; CereSpir; Cogstate; Eisai; Elan Pharmaceuticals; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche and its affiliated company Genentech; Fujirebio; GE Healthcare; IXICO; Janssen Alzheimer Immunotherapy Research & Development; Johnson & Johnson Pharmaceutical Research & Development; Lumosity; Lundbeck; Merck & Co, Inc; Meso Scale Diagnostics; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals; Pfizer; Piramal Imaging; Servier; Takeda Pharmaceutical Company and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organisation is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuroimaging at the University of Southern California.

Collaborators Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at:http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Contributors RW and XX: conception and design of the study, acquisition and analysis of data and drafting the manuscript. Y Liu acts as the guarantor of the study and takes full responsibility for the work. YD, NZ, JS, HL, Y Li, FB, JHC: conception and design of the study, acquisition and analysis of data. Y Li, CZ, XH, FZ, MH, RL, ZZ: acquisition and analysis of data. All authors revised the manuscript and approved the final draft.

Funding This work was supported by the National Natural Science Foundation of China (No. 81870958), the Beijing Municipal Natural Science Foundation for Distinguished Young Scholars (No. JQ20035), Beijing Young Scholarship and the Capital's Funds for Health Improvement and Research (CFH2022-1-2042). FB is supported by the NIHR Biomedical Research Centre at UCLH.

Competing interests FB acts as a consultant for Combinostics, Biogen-Idec, Janssen, IXICO, Merck-Serono, Novartis and Roche. He has received grants, or grants are pending, from the Amyloid Imaging to Prevent Alzheimer's Disease (AMYPAD) initiative, the Biomedical Research Centre at University College London Hospitals, the Dutch MS Society, ECTRIMS MAGNIMS, EU-H2020, the Dutch Research Council (NWO), the UK MS Society and the National Institute for Health Research, University College London. He has received payments for the development of educational presentations from IXICO and his institution from Biogen-Idec and Merck. He is co-founder of Queen Square Analytics. He is on the editorial board of Radiology, Neuroradiology, Multiple Sclerosis Journal and Neurology. JHC is a scientific consultant to and shareholder in BrainKey and Claritas Healthcare, as has worked as a consultant to Queen Square Analytics.

Patient consent for publication Not applicable.

Ethics approval This study was approved by Ethics Committee of Beijing Tiantan Hospital, Capital Medical University (No. KY-2019-050-02). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request.

ORCID iDs
Xiaolu Xu http://orcid.org/0000-0001-9963-529X
Yunyun Duan http://orcid.org/0000-0002-8753-0260
James H Cole http://orcid.org/0000-0003-1908-5588
Yaou Liu http://orcid.org/0000-0002-9930-0331

REFERENCES

1 Papais-Alvarenga RM, Vasconcelos CCF, Carra A, *et al*. Central nervous system idiopathic inflammatory demyelinating disorders in South Americans: a descriptive, multicenter, cross-sectional study. *PLoS One* 2015;10:e0127757.
2 Trojano M, Avolio C, Manzari C, *et al*. Multivariate analysis of predictive factors of multiple sclerosis course with a validated method to assess clinical events. *J Neurol Neurosurg Psychiatry* 1995;58:300 6.
3 GBD 2016 Neurology Collaborators. Global, regional, and national burden of neurological disorders, 1990-2016: a systematic analysis for the global burden of disease study 2016. *Lancet Neurol* 2019;18:459 80.
4 Cole JH, Franke K. Predicting age using neuroimaging: innovative brain ageing biomarkers. *Trends Neurosci* 2017;40:681 90.
5 Zhai J, Li K. Predicting brain age based on spatial and temporal features of human brain functional networks. *Front Hum Neurosci* 2019;13:62.
6 Cole JH, Poudel RPK, Tsagkrasoulis D, *et al*. Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. *Neuroimage* 2017;163:115 24.
7 Cole JH, Raffel J, Friede T, *et al*. Longitudinal assessment of multiple sclerosis with the Brain-Age paradigm. *Ann Neurol* 2020;88:93 105.
8 Wingerchuk DM, Banwell B, Bennett JL, *et al*. International consensus diagnostic criteria for neuromyelitis optica spectrum disorders. *Neurology* 2015;85:177 89.
9 Thompson AJ, Banwell BL, Barkhof F, *et al*. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol* 2018;17:162 73.
10 Costelloe L, O'Rourke K, Kearney H, *et al*. The patient knows best: significant change in the physical component of the multiple sclerosis impact scale (MSIS-29 physical). *J Neurol Neurosurg Psychiatry* 2007;78:841 4.
11 Ellis KA, Bush AI, Darby D, *et al*. The Australian imaging, biomarkers and lifestyle (AIBL) study of aging: methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer's disease. *Int Psychogeriatr* 2009;21:672 87.
12 Buckner RL, Roffman JL, Smoller JW. Brain Genomics Superstruct Project (GSP). In: *Neuroinformatics Research Group*, 2014.
13 Liu W, Wei D, Chen Q, *et al*. Longitudinal test-retest neuroimaging data from healthy young adults in Southwest China. *Sci Data* 2017;4:170017.
14 Isensee F, Schell M, Pflueger I, *et al*. Automated brain extraction of multisequence MRI using artificial neural networks. *Hum Brain Mapp* 2019;40:4952 64.
15 Fischl B, Salat DH, Busa E, *et al*. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* 2002;33:341 55.
16 Whitwell JL, Crum WR, Watt HC, *et al*. Normalization of cerebral volumes by use of intracranial volume: implications for longitudinal quantitative MR imaging. *AJNR Am J Neuroradiol* 2001;22:1483 9.
17 Peng H, Gong W, Beckmann CF, *et al*. Accurate brain age prediction with lightweight deep neural networks. *Med Image Anal* 2021;68:101871.
18 Wang Y, Cai H, Li C, *et al*. Optimal caliper width for propensity score matching of three treatment groups: a Monte Carlo study. *PLoS One* 2013;8:e81045.
19 Elliott ML, Belsky DW, Knodt AR, *et al*. Brain-age in midlife is associated with accelerated biological aging and cognitive decline in a longitudinal birth cohort. *Mol Psychiatry* 2021;26:3829 38.
20 Smith SM, Elliott LT, Alfaro-Almagro F, *et al*. Brain aging comprises many modes of structural and functional change with distinct genetic and biophysical associations. *Elife* 2020;9:e52677.
21 Wang J, Knol MJ, Tiulpin A, *et al*. Gray matter age prediction as a biomarker for risk of dementia. *Proc Natl Acad Sci U S A* 2019;116:21213 8.
22 Pardoe HR, Cole JH, Blackmon K, *et al*. Structural brain changes in medically refractory focal epilepsy resemble premature brain aging. *Epilepsy Res* 2017;133:28 32.
23 Cole JH, Leech R, Sharp DJ, *et al*. Prediction of brain age suggests accelerated atrophy after traumatic brain injury. *Ann Neurol* 2015;77:571 81.
24 Liu Y, Wang J, Daams M, *et al*. Differential patterns of spinal cord and brain atrophy in NMO and MS. *Neurology* 2015;84:1465 72.
25 Hahn T, Ernsting J, Winter NR, *et al*. An uncertainty-aware, shareable, and transparent neural network architecture for brain-age modeling. *Sci Adv* 2022;8:eabg9471.
26 Reyes M, Meier R, Pereira S, *et al*. On the interpretability of artificial intelligence in radiology: challenges and opportunities. *Radiol Artif Intell* 2020;2:e190043.
27 Wood DA, Kafiabadi S, Busaidi AA, *et al*. Accurate brain-age models for routine clinical MRI examinations. *Neuroimage* 2022;249:118871.