Brain age as a surrogate marker for information processing speed in multiple sclerosis

Denissen S1,2,3, Engemann DA3,4, De Cock A1, Costers L1,2, Baijot J1, Laton J1,5, Penner IK5,7, Grothe M6, Kirsch M1, D’hooghe MB10,11, D’Haeseleer M10, Dive D12, De Mey J1,3, Van Schependom J1,14, Sima DM12, and Nagels G1,2,15

1AIMS lab, Center for Neurosciences, UZ Brussels, Vrije Universiteit Brussel, Brussels, Belgium
2icometrix, Leuven, Belgium
3Université Paris-Saclay, Inria, CEA, Palaiseau, France
4Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
5Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, United Kingdom
6Cogito Centre for Applied Neurocognition and Neuropsychological Research, Düsseldorf, Germany
7Department of Neurology, Medical Faculty, Heinrich Heine University Düsseldorf, Düsseldorf, Germany
8Department of Neurology, University Medicine Greifswald, Ferdinand-Sauerbruchstraße, 17475, Greifswald, Germany
9Institute for Diagnostic Radiology and Neuroradiology, University Medicine of Greifswald, Germany
10National Multiple Sclerosis Center Melsbroek, Melsbroek, Belgium
11Center for Neurosciences, Vrije Universiteit Brussel, Brussels, Belgium
12Department of Neurology, University Hospital of Liege, Esneux, Belgium
13Department of Radiology, UZ Brussels, Brussels, Belgium
14Department of Electronics and Informatics (ETRO), Vrije Universiteit Brussel, Brussels, Belgium
15St Edmund Hall, University of Oxford, Oxford, UK

Background Data from neuro-imaging techniques allow us to estimate a brain’s age. Brain age is easily interpretable as “how old the brain looks”, and could therefore be an attractive communication tool for brain health in clinical practice. This study aimed to investigate its clinical utility.

Objectives To investigate the relationship between brain age and information processing speed in MS.

Methods A ridge-regression model was trained to predict age from brain MRI volumetric features and sex in a healthy control dataset (HC_train, n=1690). This model was used to predict brain age in two test sets: HC_test (n=50) and MS_test (n=201). Brain-Predicted Age Difference (BPAD) was calculated as BPAD=brain age minus chronological age. Information processing speed was assessed with the Symbol Digit Modalities Test (SDMT).

Results Brain age was significantly related to SDMT scores in the MS_test dataset (r=-0.44, p<0.001), and contributed uniquely to variance in SDMT beyond chronological age, reflected by a significant correlation between BPAD and SDMT (r=-0.21, p=0.003) and a significant weight (-0.21, p=0.011) in a multivariate regression equation with age.

Conclusions Brain age is a candidate biomarker for information processing speed in MS and an easy to grasp metric for brain health.

Introduction

About half of the people with multiple sclerosis (MS) experience cognitive impairment(1), aggravating the impact of MS on their daily life and that of their caregivers(2). The most prominent of these cognitive difficulties is a slowing of information processing abilities(3), which literature suggests to be the key driver for deficits on other cognitive domains in MS(4). As difficulties with information processing speed (IPS) can induce problems such as falling, reduced quality of life and employment issues(4), and might negatively affect mental health in caregivers(5), timely identification is necessary. This requires regular and consistent follow-up in standard clinical care. Currently, neuropsychological testing remains the gold standard to detect cognitive problems(6), the most popular test in clinical practice for IPS being the Symbol Digit Modalities Test (SDMT)(7). Although the SDMT is a brief screening test(8), it is prone to practice effects(9).

An objective biomarker to diagnose IPS deficits might circumvent the aforementioned problem. Currently, predominantly structural brain characteristics, extracted from brain imaging techniques such as MRI, were found to be related to IPS(10). Yet more information can be extracted from an MRI than meets the eye. By using large datasets of brain images of healthy individuals, a machine learning model can be trained to estimate the age of a given brain. For a new brain image, the model will output the best guess of the age of that person’s brain, i.e. the “brain age”, which can look older or younger than the actual, chronological age of that person. The elegance of brain age is in its interpretable nature; it is easily graspable as how old a brain appears to be. In several brain disorders(11), among which MS(11–13), brains typically look older than those of their healthy peers.

On top of being an interpretable metric, recent evidence indicates that brain age could explain clinical symptomatology in MS. More specifically, increased brain age is associated to physical disability as quantified by the Expanded Disability Status Scale (EDSS)(11, 13) and nine hole peg test (9HPT)(12). Beyond physical disability, recent findings by Kaufmann et al 2019 in dementia suggest that brain age could explain cognitive disability as well, namely by finding...
that increased brain age was associated with lower scores on the Mini Mental State Examination (MMSE), independent of chronological age(11).

In summary, brain age is an interpretable imaging-derived metric that is sensitive to MS-related pathology. However, it is currently unknown how brain aging is related to MS-specific dysfunction in processing speed. So far, efforts have mostly uncovered anatomical correlates by directly linking brain volumetry to information processing speed. Yet, people might be more easily capable of imagining “how old a brain looks” compared to “how voluminous a brain is”, posing an opportunity for a new communication tool in clinical practice that avoids medical jargon, which answers the desire of patients to be informed in plain language(14). In this study, we explore brain age as a tool for studying dysfunction in cognitive processing speed in MS on an international, multi-center dataset.

Methods

A. Data description. Data are described by subdividing them in HC_train, HC_test and MS_test, used to train and test the brain age decoding model as shown in figure 1. HC_train was constructed from a large sample of 1690 healthy control (HC) subjects from online open-source repositories. We refer to supplementary table S1 and figure S1 for a more detailed description. HC_test was collected as part of a study(15, 16) on understanding the neural origins of cognitive disturbances in MS, and for which both HC (HC_test, n = 50) and MS (n = 97) subjects were assessed. The MS subjects of this study were recruited at the National Multiple Sclerosis Center (NMSC) of Melsbroek, and were included in the MS_test dataset, which was further enriched with data from the Universitätsmedizin Greifswald (n = 104). For all data, T1-weighted Magnetic Resonance Imaging (MRI), sex and age at image acquisition were available. For the test datasets, Fluid Attenuated Inversion Recovery (FLAIR) MRI and results from the Symbol Digit Modalities Test (SDMT)(7) were available. The SDMT is a brief test, designed to measure information processing speed, that is attractive for its psychometric properties(17), quick administration(8) and its capability of predicting scores on other cognitive tests(18). Finally, MS_test data also contained EDSS, disease duration and type of MS. A summary of all data is available in table 1.

B. Ethics. Participants of the BRUMEG study, Brussels, all provided their written informed consent prior to MRI assessment. The study protocol (B.U.N. 143201423263) was approved by the ethical committee of the UZ Brussels (Commissie Medische Ethiek (O.G. 016), Reflectiegroep Biomedische Ethiek) on 2015/02/25. For the patients from Greifswald, Germany, the study was approved by the ethics committee of the Medical Faculty of the University of Greifswald (BB159/18), and all participants gave their written informed consent. HC_train data consists of publicly available data originating from other projects, listed in supplementary table S1. Ethical approval was received by each project separately.

C. MRI preprocessing and brain age pipeline. Several preparatory steps were involved in the construction of our brain age model. They are summarized below.

1. Brain MRI segmentation. From the HC_train dataset, we extracted 3D T1-weighted MRI, sex and chronological age. Next, the T1-weighted MR images were evaluated by the FDA cleared icobrain software (version 4.4) of ico metamrix NV (Leuven, Belgium). This is an end-to-end automatic software that segments and subsequently quantifies distinct regions of the brain. For a comprehensive description of the pipeline, originally published in Jain et al. 2015(19), we refer to Golan et al 2020(20). The pipeline relies on T1-weighted MR images, and when available, also uses FLAIR images to segment white matter lesions in the brain. These lesions are filled in the T1 image with white matter intensities, and the T1 image is subsequently segmented by fitting a probabilistic model for grey matter, white matter and cerebrospinal fluid image intensities. Ultimately, the pipeline generates volumes of the segmented regions, yielding a set of features, normalised for head size, that describe the brain’s morphology. These can be subdivided into general volumes (whole brain, grey matter, white matter, cortical grey matter, lateral ventricles), lobe-specific cortical grey matter (frontal, temporal, parietal, occipital) and subcortical volumes (e.g., hippocampus and thalamus (total, left and right)). Together with the subjects’ sex, this forms the total set of features used for the brain age pipeline. HC_test and MS_test were also segmented with icobrain, yielding the aforementioned set of features. The sole difference with HC_train is the additional availability of FLAIR images to perform lesion filling in the T1 image.

2. Ridge regression and selecting regularisation parameter lambda. Along with chronological age, i.e. the target variable to be predicted, the total set of features (z-normalised) served as input for training a supervised machine learning model. Since the brain volumetric features in our data were mutually correlated, and ordinary linear regression could therefore result in large weights that are often indicative for overfitting(21), we used ridge regression. This method forces weights to shrink, a process that is controlled by regularisation parameter lambda. Lambda was established prior to training using a random search hyperparameter tuning.

3. Brain age correction. According to Le et al. 2018, age-predicting models are prone to “regression towards the mean”, a phenomenon that results in overestimation of the brain age of younger subjects, and underestimation of the brain age of older subjects(22). Inspired by Cole et al. 2020, we corrected for this bias by correcting brain age linearly for age(13). First, we
Table 1. Data characteristics. Abbreviations: SD = standard deviation, M = Male, F = Female, IQR = interquartile range, T = Tesla, CIS = Clinically Isolated Syndrome, RRMS = Relapsing Remitting MS, SPMS = Secondary Progressive MS, PPMS = Primary Progressive MS. *We refer to supplementary table S1 for more details.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>HC_train</th>
<th>HC_test</th>
<th>MS_test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>Open Source</td>
<td>Brussels</td>
<td>Brussels</td>
</tr>
<tr>
<td>N</td>
<td>1690</td>
<td>50</td>
<td>97</td>
</tr>
<tr>
<td>Age:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Mean ± SD</td>
<td>41.6 ± 19.6</td>
<td>47.9 ± 12.0</td>
<td>48.0 ± 9.5</td>
</tr>
<tr>
<td>- Range (min - max)</td>
<td>8 – 94</td>
<td>26 – 68</td>
<td>26 – 72</td>
</tr>
<tr>
<td>EDSS (Median; IQR)</td>
<td>/</td>
<td>/</td>
<td>3.6; 2.0</td>
</tr>
<tr>
<td>Disease duration (years)</td>
<td>/</td>
<td>/</td>
<td>15.7 ± 8.4</td>
</tr>
<tr>
<td>MS subtype</td>
<td>/</td>
<td>/</td>
<td>CIS: 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RRMS: 82</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SPMS: 6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PPMS: 7</td>
</tr>
<tr>
<td>SDMT (Mean ± SD)</td>
<td>53.8 ± 9.6</td>
<td>48.0 ± 11.4</td>
<td>51.2 ± 15.0</td>
</tr>
</tbody>
</table>

A. The brain age pipeline. Linear regression between brain age, obtained with 10-fold cross-validation on HC_train, and chronological age yielded the following correction weights: \( \beta_0 = 8.32, \beta_1 = 0.80 \). In the same dataset, the mean absolute error (MAE) between corrected brain age and chronological age was 7.84 years. The result of the brain age correction was added to the supplementary material as figure S2.

Second, we calculated the error of predicting age from brain images with the mean absolute error (MAE) between true and predicted age for every dataset. Next, to establish the association between brain age and SDMT in the MS_test data, we used Pearson correlation. To assess whether it contains unique information beyond chronological age, we used two approaches. First, we considered brain age and chronological age together in a multivariate linear regression equation: SDMT = \( \beta_0 + \beta_1 \text{BrainAge} + \beta_2 \text{ChronologicalAge} + \epsilon \). Second, previous literature used the “Brain-Predicted Age Difference” (BPAD)(13). It quantifies brain age overestimation by subtracting chronological age from brain age. It was calculated in our test sets from the corrected brain age values (figure 1). Pearson correlation was used to establish the relationship between BPAD and SDMT in the MS_test data.

To test the complexity of the brain age measure, we compared it to the first principal component (\( PC_1 \)) of the feature space in the MS_test data by means of Pearson correlation.

**Results**

**A. The brain age pipeline.** Linear regression between brain age, obtained with 10-fold cross-validation on HC_train, and chronological age yielded the following correction weights: \( \beta_0 = 8.32, \beta_1 = 0.80 \). In the same dataset, the mean absolute error (MAE) between corrected brain age and chronological age was 7.84 years. The result of the brain age correction was added to the supplementary material as figure S2. In a next step, represented by panel A in figure 1, we fitted a final brain age model by using all available HC_train data for training. The model’s feature weights and lambda estimated \( \beta_0 = 8.32, \beta_1 = 0.80 \). In the same dataset, the mean absolute error (MAE) between corrected brain age and chronological age was 7.84 years. The result of the brain age correction was added to the supplementary material as figure S2. In a next step, represented by panel A in figure 1, we fitted a final brain age model by using all available HC_train data for training. The model’s feature weights and lambda estimated BrainAge_raw on HC_train by adopting 10-fold cross-validation. Second, we fitted a linear regression equation between the obtained raw brain ages and the respective chronological ages: \( \text{BrainAge}_{\text{raw}} = \beta_0 + \beta_1 \text{ChronologicalAge} + \epsilon \). Herein, \( \beta_0 \) and \( \beta_1 \) represent the intercept and slope of the regression line, respectively, whereas the error term \( \epsilon \) represents the residuals between data points and the regression line. \( \beta_0 \) and \( \beta_1 \) serve to correct each brain age predicted by our model using: \( \text{BrainAge} = (\text{BrainAge}_{\text{raw}} - \beta_0)/\beta_1 \). They were first used to correct the raw brain age estimates of HC_train.

Altogether, the preparatory steps yield the features and regularisation parameter lambda that are used for the training phase (figure 1, panel A) of the final brain age model, i.e. establishing the final feature weights on the full HC_train data using ridge regression. Next, we investigated whether the learned weights generalise to other datasets as well (figure 1, panel B). We first calculated raw brain age on the HC_test and MS_test dataset by calculating the weighted sum of the features (i.e. sex and brain volumes from icobrain) with their respective weight, which were then corrected by using the correction formula of preparatory step 3. In the remainder of this paper, “brain age” consistently refers to the corrected brain age.

**D. Statistical Analyses.** Statistical analyses and visualizations were performed in Python and R. Significance level alpha was set to 0.05 for all reported test results. Where Pearson correlation was used, the distributions of the variables were checked for normality with a Shapiro-Wilk test. Raincloud plots were generated with use of the PtitPrince package(23).

First, to compare brain age, BPAD and chronological age distributions between MS_test and HC_test, we used a Mann-Whitney U test. To compare BPAD of both test sets with 0, we used a one-sample Wilcoxon signed rank test.

**Results**

**A. The brain age pipeline.** Linear regression between brain age, obtained with 10-fold cross-validation on HC_train, and chronological age yielded the following correction weights: \( \beta_0 = 8.32, \beta_1 = 0.80 \). In the same dataset, the mean absolute error (MAE) between corrected brain age and chronological age was 7.84 years. The result of the brain age correction was added to the supplementary material as figure S2. In a next step, represented by panel A in figure 1, we fitted a final brain age model by using all available HC_train data for training. The model’s feature weights and lambda estimated BrainAge_raw on HC_train by adopting 10-fold cross-validation. Second, we fitted a linear regression equation between the obtained raw brain ages and the respective chronological ages: \( \text{BrainAge}_{\text{raw}} = \beta_0 + \beta_1 \text{ChronologicalAge} + \epsilon \). Herein, \( \beta_0 \) and \( \beta_1 \) represent the intercept and slope of the regression line, respectively, whereas the error term \( \epsilon \) represents the residuals between data points and the regression line. \( \beta_0 \) and \( \beta_1 \) serve to correct each brain age predicted by our model using: \( \text{BrainAge} = (\text{BrainAge}_{\text{raw}} - \beta_0)/\beta_1 \). They were first used to correct the raw brain age estimates of HC_train.

Altogether, the preparatory steps yield the features and regularisation parameter lambda that are used for the training phase (figure 1, panel A) of the final brain age model, i.e. establishing the final feature weights on the full HC_train data using ridge regression. Next, we investigated whether the learned weights generalise to other datasets as well (figure 1, panel B). We first calculated raw brain age on the HC_test and MS_test dataset by calculating the weighted sum of the features (i.e. sex and brain volumes from icobrain) with their respective weight, which were then corrected by using the correction formula of preparatory step 3. In the remainder of this paper, “brain age” consistently refers to the corrected brain age.

**D. Statistical Analyses.** Statistical analyses and visualizations were performed in Python and R. Significance level alpha was set to 0.05 for all reported test results. Where Pearson correlation was used, the distributions of the variables were checked for normality with a Shapiro-Wilk test. Raincloud plots were generated with use of the PtitPrince package(23).

First, to compare brain age, BPAD and chronological age distributions between MS_test and HC_test, we used a Mann-Whitney U test. To compare BPAD of both test sets with 0, we used a one-sample Wilcoxon signed rank test.

Second, we calculated the error of predicting age from brain images with the mean absolute error (MAE) between true and predicted age for every dataset. Next, to establish the association between brain age and SDMT in the MS_test data, we used Pearson correlation. To assess whether it contains unique information beyond chronological age, we used two approaches. First, we considered brain age and chronological age together in a multivariate linear regression equation: \( \text{SDMT} = \beta_0 + \beta_1 \text{BrainAge} + \beta_2 \text{ChronologicalAge} + \epsilon \). Second, previous literature used the “Brain-Predicted Age Difference” (BPAD)(13). It quantifies brain age overestimation by subtracting chronological age from brain age. It was calculated in our test sets from the corrected brain age values (figure 1). Pearson correlation was used to establish the relationship between BPAD and SDMT in the MS_test data.

To test the complexity of the brain age measure, we compared it to the first principal component (\( PC_1 \)) of the feature space in the MS_test data by means of Pearson correlation.
value can be consulted in table 2. To test its quality, we applied it to the HC_test and MS_test set, represented in figure 1 by panel B. The corrected brain age and BPAD values are summarised in table 3, and visually displayed for HC_test and MS_test in the raincloud plots of figure 2.

B. The relation between brain age and information processing speed. Brain age was significantly correlated to SDMT (figure 3, r = -0.44, p < .001). Moreover, brain age explained unique variance in SDMT beyond chronological age, which is reflected by a significant correlation between BPAD and SDMT (figure 4, r = -0.21, p = 0.003) and the significant weight ($\beta_1 = -0.21, p = 0.011$) assigned to brain age when considering it in the multivariate regression equation $SDMT = \beta_0 + \beta_1 BrainAge + \beta_2 ChronologicalAge + \epsilon$ (figure 5). Chronological age also contributed significantly to the model ($\beta_2 = -0.35, p < .001$).

C. The relation between brain age and brain volume-try. $PC_1$ explained 85.3% of the total variance that is explained by the total set of features in the MS_test data. Figure 6 displays the relationship between $PC_1$ and brain age, revealing a strong and significant linear relationship ($r = 0.94$, $p < .001$).

Discussion

This study aimed to investigate the potential of brain age, an intuitive metric of brain health, as a biomarker for information processing speed in Multiple Sclerosis. Our results suggest that brain age could be a promising candidate; it was substantially related to information processing speed, independent of chronological age. Moreover, we showed
Table 3. Outputs from the brain age pipeline. Corrected brain age and BPAD for the HC_train, HC_test and MS_test datasets.

<table>
<thead>
<tr>
<th></th>
<th>HC_train</th>
<th>HC_test</th>
<th>MS_test</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>1690</td>
<td>50</td>
<td>97</td>
</tr>
<tr>
<td>N</td>
<td>104</td>
<td>201</td>
<td></td>
</tr>
<tr>
<td>Brain Age (M ± SD)</td>
<td>41.6 ± 22.0</td>
<td>46.8 ± 16.3</td>
<td>60.2 ± 16.8</td>
</tr>
<tr>
<td>BPAD (Mean ± SD)</td>
<td>0 ± 9.9</td>
<td>-1.1 ± 9.4</td>
<td>12.2 ± 14.2</td>
</tr>
</tbody>
</table>

Table 3. Outputs from the brain age pipeline. Corrected brain age and BPAD for the HC_train, HC_test and MS_test datasets.

Fig. 2. Group comparison between HC_test (blue) and MS_test (orange) for brain age, BPAD and chronological age. Left: the raincloud plots show the distribution of brain age, BPAD and chronological age for MS_test and HC_test. A reference line at x = 0 was included as visual aid. Right: the scatterplot shows the relationship between brain age and chronological age for MS_test and HC_test. The dotted line was added as reference, namely where brain age = chronological age.

that brain age explained the majority of variance in brain volumetry by establishing a strong relationship with the first principal component of our total set of features.

In the past few years, most brain age-related research in MS was dedicated to establish clinical correlates of the Brain-Predicted Age Difference (BPAD)(11–13). Although BPAD might be regarded as a simplification of brain age, we in fact lose valuable information by subtracting two variables. Our results supports this statement in two ways. First, BPAD showed a weaker correlation with IPS than brain age. Second, brain age and chronological age both contributed unique information in explaining IPS in MS. Hence, in terms of clinical significance, BPAD remains at the level of proof-of-concept.

Although brain age is a fair choice for decoding processing speed in MS, we note that its performance was comparable to whole brain volume, as shown by Golan et al. 2020, reporting a correlation of r = 0.40 between whole brain volume and information processing speed(20). Nonetheless, brain age has an important advantage in contrast to any biological correlate of cognition in MS; it is easy to grasp as “how old a brain looks”, which facilitates communication with persons with MS. Constructing an uncomplicated message with minimal jargon contributes to optimally transferring medical information to patients, in turn optimising patient care(14).

However, an important hurdle in the path of brain age models to clinical practice is nicely illustrated by a statement in Ribeiro et al 2016: “if the users do not trust a model or a prediction, they will not use it”(24). First, to maximise trust of the MS clinician in our model, we used a model that is based on simple linear regression. The MS clinician has been familiarised with this method by decades of research adopting it for various purposes, e.g. studying the relation between MRI and cognitive performance in multiple sclerosis(25, 26). This contrasts with other studies on brain age in MS, mostly adopting models that are common in machine learning, but not in clinical research, such as Gaussian processes regression(13) and extreme gradient boosting(12). Second, trust in the prediction of our model, i.e. predicted brain age, was enhanced by showing that brain age explains the majority of the variance in MRI-derived volumetric features and sex, used to train our brain age model. This observation is logical in light of what is known about the aging brain, shrinking as people get older(27).

Our results imply that brain age has the potential to explain cognitive status in people with Multiple Sclerosis. Yet, the cross-sectional nature of the data limited us to investigate the potential of brain age to predict future decline in processing speed. Another limitation concerns the difference in segmentation method between the HC_train data and all test data. In HC_train, segmentation relied solely on T1-weighted images, whereas lesion filling in the T1-weighted image occurred in the test datasets with the use of FLAIR images. However, lesion filling had little impact on brain age predictions in the MS sample of Cole et al. 2020.(13) To ascertain
that our results were not driven by this methodological issue, we recalculated brain age in the MS_test dataset on the same set of features, except that lesion volume was subtracted from white matter and whole brain volume. The results of this analysis, which can be consulted in the supplementary material (figure S3 and S4), indicate that brain age predictions were hardly affected by flair hyperintensity volume, and therefore do not explain the observations in this paper. Finally, our brain age model implemented a correction step to correct for the inherent over- and underestimation of younger and older subjects respectively. Since there appears to be no firm consensus about implementing this step, we reran the entire brain age pipeline without brain age correction, which did not alter the interpretation of this paper.

In summary, the methodology of a linear brain age model can be interpreted by clinicians, and its prediction by patients. Together with its potential to explain performance in information processing speed, predicted brain age could be a valuable clinical tool to analyse and communicate results from brain imaging data in MS.

Acknowledgements
The authors wish to thank Ann Van Remoortel and Jeroen Gielen for contributing to the data acquisition in the BRUMEG study.

Funding
The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Stijn Denissen is funded by a Baekeland grant appointed by Flanders Innovation and Entrepreneurship (HBC.2019.2579, www.vlaio.be), Guy Nagels received research grants from Biogen and Genzyme, and is a senior clinical research fellow of the FWO Flanders (1805620N, www.fwo.be), and Jeroen Van Schependom is a senior research fellow of VUB.
Declaration of Conflicting Interests
Stijn Denissen is an industrial PhD candidate in collaboration with icometrix. Diana Sima is a senior researcher at icometrix. Guy Nagels is on a 10% secondment from the UZ Brussel to icometrix as medical director, and is an employee of icometrix.

Bibliography

9. Emilio Portaccio, Benedetta Goretta, Valentina Zipoli, Alfonso Jude, Dario Della Pina, Gian Michele Malenzi, Simona Sabatini, Pasquale Annunziata, Marco Falci, Monica Mazzone, and Maria Pia Amato. Reliability, practice effects, and change indices for Roa's brief repeatable battery. Multiple sclerosis, 16(5):611–617, may 2010. ISSN 13524585.
11. Tobias Kaufmann, Dennis van der Meer, Naih Trong Doan, Emanuel Schwarz, Martina J. Denissen et al.
Supplementary Material

Data characteristics HC_train dataset

<table>
<thead>
<tr>
<th>Source</th>
<th>N</th>
<th>Sex (M:F)</th>
<th>Scanner (Field strength (Tesla))</th>
<th>Additional info</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>studyforrest</td>
<td>18</td>
<td>10:8</td>
<td>Philips Achieva (3T)</td>
<td><a href="http://studyforrest.org/access.html">http://studyforrest.org/access.html</a></td>
<td>[1]</td>
</tr>
<tr>
<td>nki</td>
<td>352</td>
<td>121:231</td>
<td>Siemens TrioTim (3T)</td>
<td><a href="http://fcon_1000.projects.nitrc.org/indi/nki.html">http://fcon_1000.projects.nitrc.org/indi/nki.html</a></td>
<td>[3]</td>
</tr>
<tr>
<td>ds000202</td>
<td>83</td>
<td>0:83</td>
<td>Philips Achieva (3T)</td>
<td><a href="https://openneuro.org/datasets/ds000202">https://openneuro.org/datasets/ds000202</a></td>
<td>[8]</td>
</tr>
<tr>
<td>ds000102</td>
<td>24</td>
<td>14:10</td>
<td>Siemens Allegra (3T)</td>
<td><a href="https://openneuro.org/datasets/ds000102">https://openneuro.org/datasets/ds000102</a></td>
<td>[9]</td>
</tr>
<tr>
<td>brainomics</td>
<td>80</td>
<td>37:43</td>
<td>Siemens TrioTim (3T)</td>
<td><a href="https://hal-cea.archives-ouvertes.fr/cea-01213448">https://hal-cea.archives-ouvertes.fr/cea-01213448</a></td>
<td>[12]</td>
</tr>
</tbody>
</table>

Table S1. Characteristics of each data source in the HC_train dataset.

Fig. S1. Histogram of the chronological ages per data source in the HC_train dataset
Brain age correction

**Fig. S2.** Brain age and BPAD as calculated with 10-fold cross-validation on the HC_train data. Uncorrected (raw) brain age and BPAD are indicated in red (regression line is indicated in orange), whereas corrected values are shown in green (regression line is indicated in light green). Both colormaps indicate higher density when yellower. Dotted lines were added to each plot as visual reference. For the upper plots, this is the line where Brain Age = Chronological Age, whereas for the lower plots, this is the line where BPAD = 0. Before correction, the regression lines do not coincide with the dotted lines, resulting in an overestimation of brain age in young individuals and an underestimation in older subjects. This is no longer the case after correction.
Impact of lesions on brain age estimation

BrainAge_{lc} (lesion correction) and BrainAge_{n–lc} (no lesion correction) refer to the brain age that was calculated respectively with and without subtracting lesion volume from whole brain and white matter volume.

Fig. S3. Scatterplot showing the relationship between BrainAge_{lc} and BrainAge_{n–lc}.

Fig. S4. Bland-Altman plot showing the brain age difference (BrainAge_{n–lc} – BrainAge_{lc}) for every brain age mean, calculated by (BrainAge_{n–lc} + BrainAge_{lc})/2. Both the mean and difference are expressed in terms of years.
Supplementary References


Preprint template acknowledgement

This preprint was created using a \LaTeX template by Ricardo Henriques in the online \LaTeX editor Overleaf. The authors wish to express their gratitude to Ricardo Henriques for publicly sharing this preprint template.