

Modality Matters: Brain Age Derived from ^{18}F -FDG-PET or MRI for Prediction of Cognitive Outcomes in the Early Alzheimer's Disease Continuum

E. Doering^{1,2}, G. Antonopoulos^{3,5}, M. C. Hömig^{1,4}, T. van Eimeren^{1,2}, M. Daamen², H. Boecker², F. Jessen^{2,6}, E. Düzel⁷, S. B. Eickhoff^{3,5}, K. R. Patil^{3,5+}, A. Drzezga^{1,2,4+} for the Alzheimer's Disease Neuroimaging Initiative*

¹University Hospital Cologne, Clinic and Policlinic for Nuclear Medicine, Cologne, Germany, ²German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany, ³Research Center Juelich, Brain and Behavior (INM-7), Juelich, Germany, ⁴Research Center Juelich, Molecular Organization of the Brain (INM-2), Juelich, Germany, ⁵Heinrich-Heine-University, Institute of Systems Neuroscience, Duesseldorf, Germany, ⁶University Hospital Cologne, Department of Psychiatry, Cologne, Germany, ⁷German Center for Neurodegenerative Diseases (DZNE), Magdeburg, Germany

Correspondence: E. Doering, Kerpener Str. 62, 50937 Köln, GERMANY, phone: (+49) 0221 478 7503, fax: (+49) 0221 748 7639, email: elena.doering@uk-koeln.de

Keywords: machine learning, cognitive impairment, neuroimaging

Brain aging is characterized by anatomical and molecular changes. Brain age can differ from chronological age ("brain age gap", *BAG*); and this difference, typically estimated from structural MRI, is associated with various intracerebral abnormalities, such as e.g. observed in Alzheimer's disease (AD). ^{18}F -Fluorodeoxyglucose PET (FDG-PET) is considered to represent an earlier indicator of neurodegeneration compared to structural MRI. Possibly, processes associated with brain aging are captured by FDG-PET with greater sensitivity compared to structural MRI. Here, we compare the accuracy of brain age estimation from FDG-PET and structural MRI, and we associate BAG derived each modality with cognitive impairment and AD biomarkers. Furthermore, we present cutoffs for the prediction of cognitive outcome after two years. Analyses were conducted in individuals without (CN), with subjective cognitive decline (SCD) and with mild cognitive impairment (MCI).

Methods: Machine learning algorithms were trained to estimate brain age from 376 matched T1-weighted MRI or FDG-PET scans of CN+SCD from the Alzheimer's Disease Neuroimaging Initiative and validated in internal and external test sets. BAG was correlated with measures of amyloid and tau pathology in CN+SCD and MCI (n=596). Finally, BAG was used to predict cognitive outcome after two years using logistic regression. Cutoffs for cognitive decline were estimated from the logistic regression output.

Results: FDG-PET (mean absolute error, $MAE=2.46$ years) and MRI ($MAE=1.96$ years) both estimated chronological age well. Both, FDG-PET- and MRI-derived BAG were correlated with amyloid load across groups and with cognitive performance in MCI. FDG-PET-derived BAG exceeding 0.85 years was indicative of pending cognitive impairment in CN+SCD, while an MRI-derived BAG beyond 2.23 years suggested development of dementia in MCI. BAG from the respective other modality was not/less indicative of cognitive outcome.

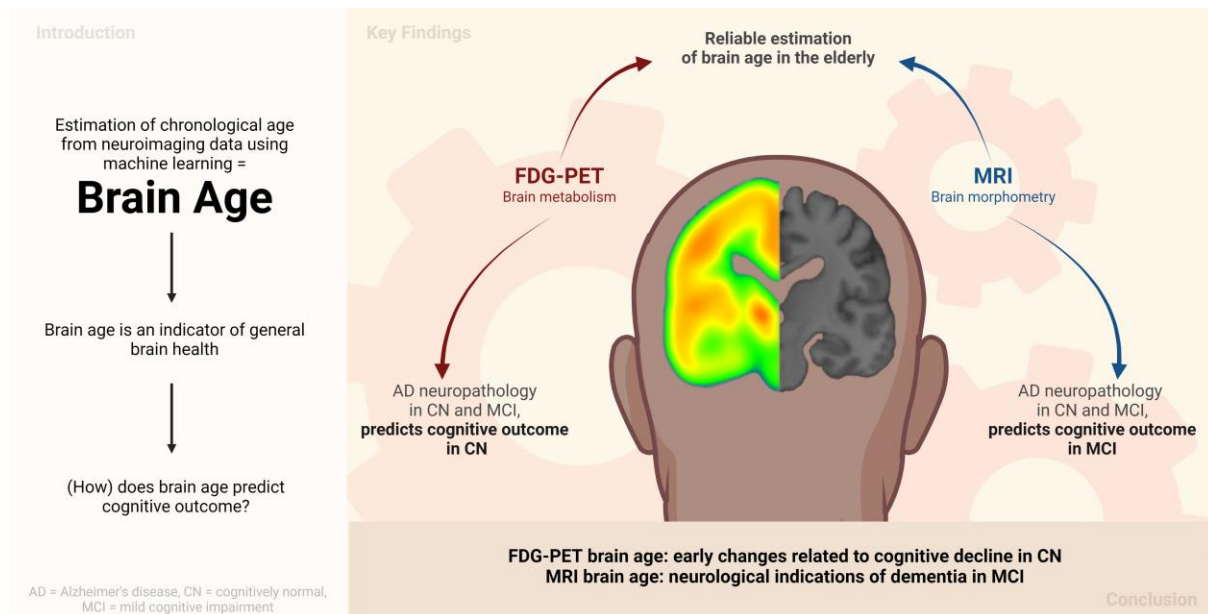
⁺ both authors contributed equally

* 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

Conclusion:

Brain age is reliably estimated from FDG-PET or MRI. FDG-PET-derived BAG is more sensitive to early intracerebral changes related to cognitive deterioration in CN+SCD, while early features indicative of impending dementia in MCI are better reflected by MRI-derived BAG.

RUNNING TITLE: FDG-PET or MRI for Brain Age Estimation



Graphical Abstract

1 Introduction

Brain aging entails changes or decline in cognitive performance, accompanied by changes in various brain functions and structural parameters of brain integrity. Brain age can be modeled using machine learning algorithms by estimating a person's chronological age from their neuroimaging data. Deviations of brain age from chronological age (the *brain age gap*, "BAG") are associated with a variety of neurological conditions, including diagnoses across the Alzheimer's disease (AD) continuum¹⁻³. A recent study¹ showed that BAG is associated with PET AD biomarkers in patients with mild cognitive impairment (MCI), and also differs significantly between individuals with impending cognitive deterioration/AD and those without. These results motivate further research into BAG as a prognostic biomarker of cognitive decline and its relation to existing biomarkers of AD, such as amyloid or tau pathology.

Age-related changes in the brain are most evident in the brain's anatomy, such as loss of brain volume (atrophy), and metabolism (neuronal dysfunction). Brain atrophy and metabolism can be quantified by T1-weighted magnetic resonance imaging (MRI) and ¹⁸F-Fluorodeoxyglucose-PET (FDG-PET), respectively. FDG-PET is considered to be an earlier indicator of neurodegeneration compared to structural MRI, as neuronal dysfunction precedes atrophy (i.e., neuronal loss)⁴. It can therefore be assumed that some processes associated with brain aging are captured by FDG-PET with a

greater sensitivity than by structural imaging. To date, however, brain age estimation, in the vast majority of cases, is performed using MRI rather than FDG-PET. Only one recent study compared the two modalities and showed slightly better performance when using FDG-PET¹. However, in this study, FDG-PET was not investigated independently of MRI, as FDG-PET was preprocessed using partial volume correction. This argues for further exploration of FDG-PET-derived BAG, and its potentially superior performance in delineating the earliest deviations from normal aging when cognitive impairment is not yet evident.

Here, we investigated the potential of FDG-PET and MRI separately as input for brain age estimation, with a particular focus on the early stages of the AD continuum. First, we estimated brain age in cohorts of individuals who were cognitively normal (CN)/had subjective cognitive decline (SCD) and patients with mild cognitive impairment (MCI). Second, we calculated BAG and compared associations of FDG-PET- and MRI-derived BAG with cognitive performance/AD neuropathology in these cohorts. Finally, we applied machine learning classification to predict cognitive outcome from BAG, and subsequently calculated a BAG cutoff value for the prognosis of cognitive impairment which was validated in an independent sample.

2 Methods

2.1 Participants

Baseline T1-weighted MRI and FDG-PET scans of 376 CN and SCD ($CN+SCD_{ADNI}$) and 596 individuals with MCI (MCI_{ADNI}) from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu) were used. The primary goal of the ADNI study has been to test whether biological markers and clinical and neuropsychological assessments can be combined to measure the progression of MCI and dementia. For an unbiased comparison of the modalities, individuals whose MRI and FDG-PET scans were less than a year apart were selected ($CN+SCD_{ADNI}$: mean = 31 days, SD = 29 days; MCI_{ADNI} : mean = 29 days, SD = 25 days). An additional 59 MRI and FDG-PET scans of CN were acquired from the Open Access of Imaging Studies-3 database⁵ (OASIS-3, <https://www.oasis-brains.org/>, CN_{OASIS}) to validate the models in an external dataset. Finally, 80 MRI scans of MCI ($MCI_{DELCODE}$) and 88 FDG-PET scans of SCD ($SCD_{DELCODE}$) from the DZNE-Longitudinal Cognitive Impairment and Dementia Study⁶ (DELCODE) were used to validate our cognitive outcome prediction

models. To be included, participants in all samples had to be older than 60 years at the time of their scan. CN, SCD and MCI diagnoses from ADNI, OASIS, and DELCODE followed the current recommendations for the respective groups^{7,8} (details provided in the Supplementary Materials (SM) section 1a).

2.2 Acquisition & Preprocessing of Biomarkers of Neurodegeneration

FDG-PET scans in ADNI and OASIS were acquired dynamically 30-60 minutes (6x5 min frames) after injection with an average dose of 185 MBq (5 mCi). The DELCODE FDG-PET data were acquired 40-60 minutes (4x5 min frames) after injection with an average dose of 170-180 MBq. T1-weighted MRI scans were acquired according to previously published MRI acquisition protocols^{5,6,9} and were preprocessed with the CAT toolbox (version 12.5) in SPM12 based on MATLAB (r2019b). After preprocessing, both FDG-PET and MRI images were in the standard MNI152 space (see details in the SM section 1b).

2.3 Estimation of brain age

To estimate brain age, we implemented a pipeline (Fehler! Verweisquelle konnte nicht gefunden werden.) in Python 3.8.5 using the Julearn library (<https://juaml.github.io/julearn/main/index.html>), which is based on scikit-learn¹⁰. The same pipeline (**FIGURE 1**) was run independently for MRI and FDG-PET. First, regional averages of the signal of interest were extracted for the respective modality (MRI: gray matter volume, FDG-PET: SUVR) using a composite atlas containing 200 cortical¹¹ and 16 subcortical regions¹². Next, we applied a nested cross-validation (CV) approach: We repeatedly (five times) split the CN+SCD_{ADNI} sample into different training and test sets, stratified by age bin (<74, 75-84 and >85 years), such that each individual occurred in a test set exactly once. Each outer CV loop consisted of outlier exclusion and an inner CV yielding a final model, estimated parameters for bias correction¹³ and estimated, bias-corrected brain age in the test sets. Information on outlier exclusion and bias correction can be found in SM sections 1c and 1d.

An inner CV was performed to tune hyperparameters and select final models (for a list of hyperparameters see SM section 1e). Support vector regression (SVR) and relevance vector regression (RVR) were used as they are recommended for brain age estimation with small sample sizes¹⁴. The algorithms were trained using four parts of the scaled training data (*training folds*) and evaluated on the

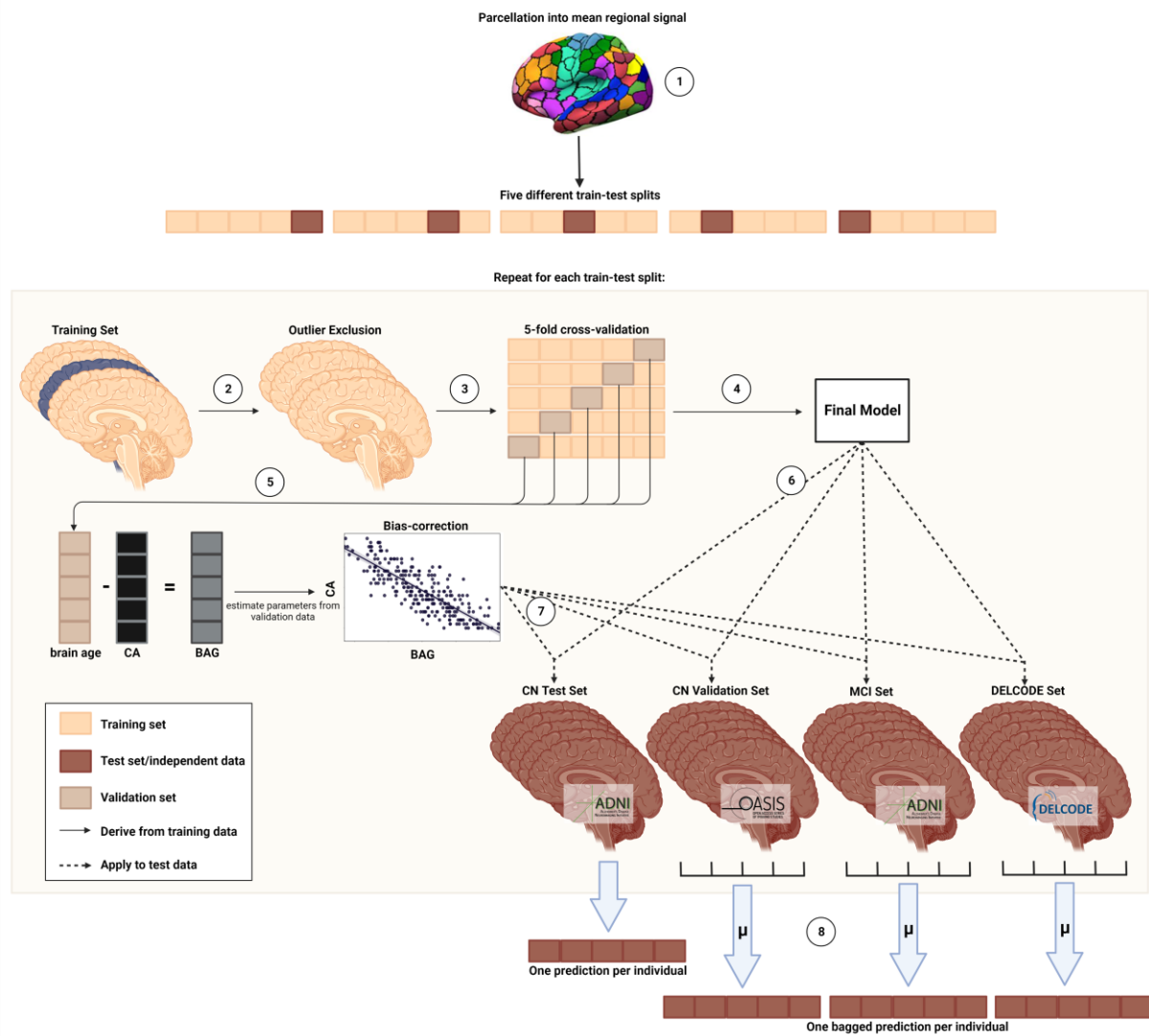


FIGURE 1. Nested cross-validation approach for brain age prediction. Five different train-test splits were used to train and test the models. (1) Region-of-interest parcellation. (2) Outlier exclusion. (3) Five-fold CV. (4) Selection of final model. (5) Bias correction. (6) Estimation of brain age in test sets. (7) Bias correction in test sets. (8) Bagging. BAG = brain age gap; CA = chronological age; CV = cross-validation. Created with BioRender.

fifth *validation fold* (data scaled with “robust scaler” from scikit-learn, scaling was estimated from training folds). The SVR and RVR models with the smallest average mean absolute error (MAE) between chronological age and brain age across the validation sets were selected as optimal models. The final model was then selected as the optimal model with the smallest average validation MAE.

As a result of the nested CV approach, we obtained five final models per modality. Thus, per modality, we obtained one brain age estimate per (non-outlier) subject in the CN+SCD_{ADNI} sample, and five estimates per (non-outlier) subject in the CN_{OASIS}, SCD_{DELCODE}, MCI_{ADNI}, and MCI_{DELCODE} samples. In the CN_{OASIS}, SCD_{DELCODE}, MCI_{ADNI} and MCI_{DELCODE} samples, the average of the five estimates was treated as the final brain age estimate (‘bagging’). The feature importance (δ) of each brain region was

assessed by considering the learned weights of linear kernel models. ‘Most important’ regions were such where δ was at least 2 standard deviations larger or smaller than the average δ . For non-linear kernels, weight coefficients cannot be directly inferred. Finally, BAG was calculated for each individual as the difference between brain age and chronological age, such that higher bag reflected more advanced brain age.

2.4 Associations of brain age gap with cognitive performance and Alzheimer’s disease neuropathology

To assess whether BAG is associated with cognitive performance, we calculated partial correlations between BAG and composite scores of memory (ADNI-MEM¹⁵) and executive function (ADNI-EF¹⁶) while correcting for age, sex, years of education and APOE- ϵ 4 carriership.

To assess whether BAG is associated with AD neuropathology, we calculated partial correlations between BAG and PET amyloid load (AV45-PET), as well as markers of beta-amyloid₁₋₄₂ (CSF A β ₁₋₄₂), total tau (CSF Tau) and phosphor-tau_{181p} (CSF p-Tau₁₈₁) in cerebrospinal fluid (CSF), while correcting for age, sex, years of education and APOE- ϵ 4 carriership. Correlations with cognitive performance and AD neuropathology were calculated in the ADNI samples and Bonferroni-corrected. Descriptions of the variables assessed are provided in SM sections 1f and 1g.

2.5 Prediction of cognitive outcome using the brain age gap

Finally, we trained a logistic regression classifier to predict cognitive outcome within two years from baseline (where BAG was assessed) in the CN+SCD_{ADNI} and MCI_{ADNI} samples. Cognitive outcome was a binary variable (“stable” vs. “decline”), based on the final diagnosis at the *two-year follow-up*. CN who received an MCI or AD diagnosis *within two years* were labeled as cognitive *decliners*, while CN who maintained CN status *until 24 months after BAG assessment* were labeled as *stables*. For MCI, decliners were those individuals who progressed to dementia within two years, while individuals who maintained the MCI diagnosis until 24 months after BAG assessment were considered stable. In the ADNI and DELCODE study, a diagnosis of dementia at follow-up entailed fulfillment of NINCDS/ADRDA criteria for probable (ADNI) and possible or probable AD (DELCODE), respectively. Both, in CN+SCD_{ADNI} and MCI_{ADNI}, we extracted an equally sized subsample of stables, matched by age and sex to the cohort of *all* decliners. MRI- and FDG-derived BAG in these samples,

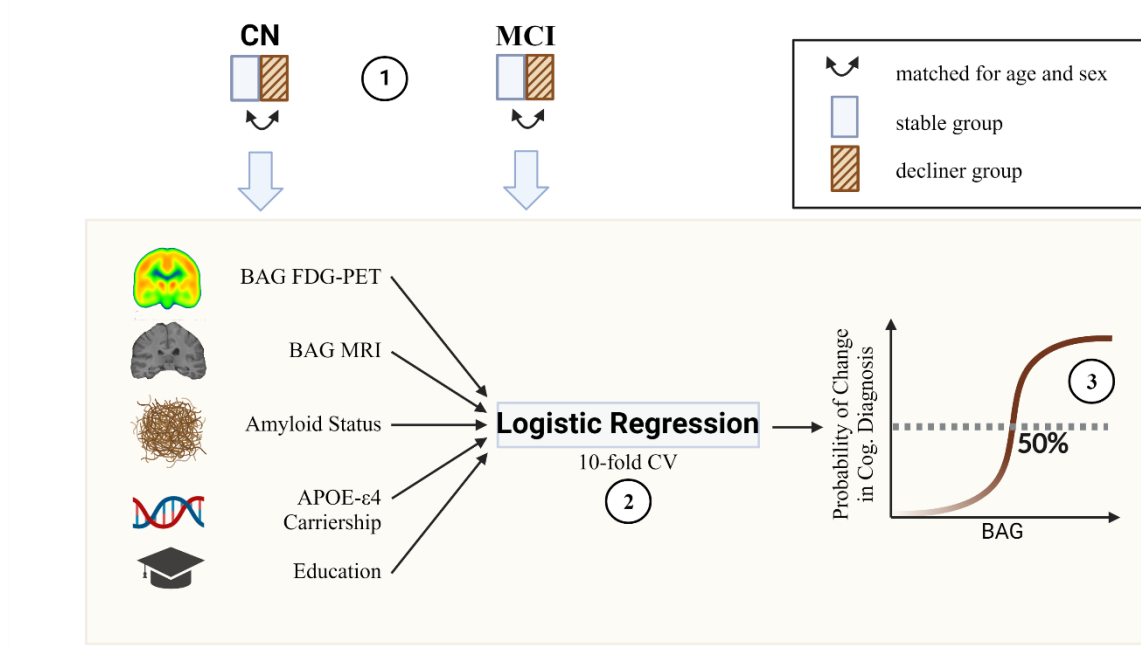


FIGURE 2. Estimation of a brain age gap cutoff for cognitive decline. (1) Matching. (2) 10-fold CV. (3) Derivation of cutoff. Created with BioRender.

together with amyloid status (CSF $A\beta_{1-42} \leq 1100$ pg/ml¹⁷), APOE-ε4 carriership and years of education, were used as input to predict cognitive outcome using a 10-fold cross-validated logistic regression classifier (**FIGURE 2**). As amyloid status was not available for all individuals, analyses were conducted in two ways: once including individuals with missing amyloid information (NA values coded as separate category 0, amyloid negativity coded as reference (-1), amyloid positivity coded as 1; “sample 1”) and once excluding individuals with missing amyloid information (“sample 2”). Significant predictors ($\alpha=.05$) of cognitive outcome were recorded. To derive a BAG cutoff for elevated risk of cognitive deterioration, a logistic regression was fitted to model the relationship between BAG from the significant imaging modality and cognitive outcome. The intercept of this curve at 50% probability was identified as a cutoff and validated in the current (ADNI), as well as the corresponding DELCODE sample.

3 Results

3.1 Participants

An overview of participant characteristics is shown in **Table 1**. In the cognitively unimpaired cohorts, CN_{OASIS} and $SCD_{DELCODE}$ were significantly younger than $CN+SCD_{ADNI}$ ($t_{OASIS} = 3.44$, $p_{OASIS} < .001$; $t_{DELCODE} = 4.45$, $p_{DELCODE} < .0001$), and the MMSE of $SCD_{DELCODE}$ was higher compared to $CN+SCD_{ADNI}$

($t_{\text{DEL CODE}} = -2.30, p = .03$). Among MCI samples, $\text{MCI}_{\text{DEL CODE}}$ had significantly fewer years of education ($t = 6.01, p < .001$).

3.2 Accuracy and demographic profile of estimated brain age

MRI- and FDG-PET-derived brain age both yielded a small MAE and therefore estimated brain age well, while the MAE of MRI-derived brain age was significantly lower compared to FDG-PET (paired t-test, $t = -6.69, p = .026$; **Table 2**). In the external CN_{OASIS} sample, bagged estimations of brain age were comparably accurate, thus showing high generalization of the models. MCI_{ADNI} individuals' brains were estimated to be significantly older compared to $\text{CN} + \text{SCD}_{\text{ADNI}}$ ($t_{\text{MCI}} = 7.89, p < .001$; $t_{\text{FDG-PET}} = 5.11, p < .001$). Both in $\text{CN} + \text{SCD}_{\text{ADNI}}$, and in MCI_{ADNI} , MRI- and FDG-PET-derived BAG were correlated across modality ($\text{CN} + \text{SCD}_{\text{ADNI}}$: normally distributed: $r = .288, p < 0.001$; MCI_{ADNI} : not normally distributed; $\rho = .428, p < .001$).

Model selection returned linear SVRs five out of five and four out of five times for MRI and FDG-PET, respectively (see Table SM1). Regional weight coefficients were strongly correlated across models of the same modality (MRI: $r = [.79, 0.89]$, FDG-PET: $r = [.74, .79]$), but average weight coefficients were not correlated between the two modalities ($r = .048, p = .48$), i.e., the regions used for brain age estimation in the two modalities were substantially different (**FIGURE 3**). For FDG-PET, important regions (very low or very high weight coefficient) included parts of the temporal and pre-frontal cortex, as well as sub-cortical regions (globus pallidus, nucleus accumbens, and caudate nucleus). Notably, SUVR in the most important regions in FDG-PET was pre-dominantly right hemispheric, and consistently negatively correlated with chronological age after Bonferroni correction ($\alpha = .05/9$). For MRI, important regions included parts of the parietal, pre-frontal and occipital cortex and sub-cortical regions (e.g., hippocampus, nucleus accumbens, globus pallidus, and caudate nucleus). The majority of most important regions for brain age estimation from MRI were negatively correlated with age after Bonferroni correction ($\alpha = .05/13$) and spread across hemispheres. A full list of most important regions with correlation results is presented in SM section 2b.

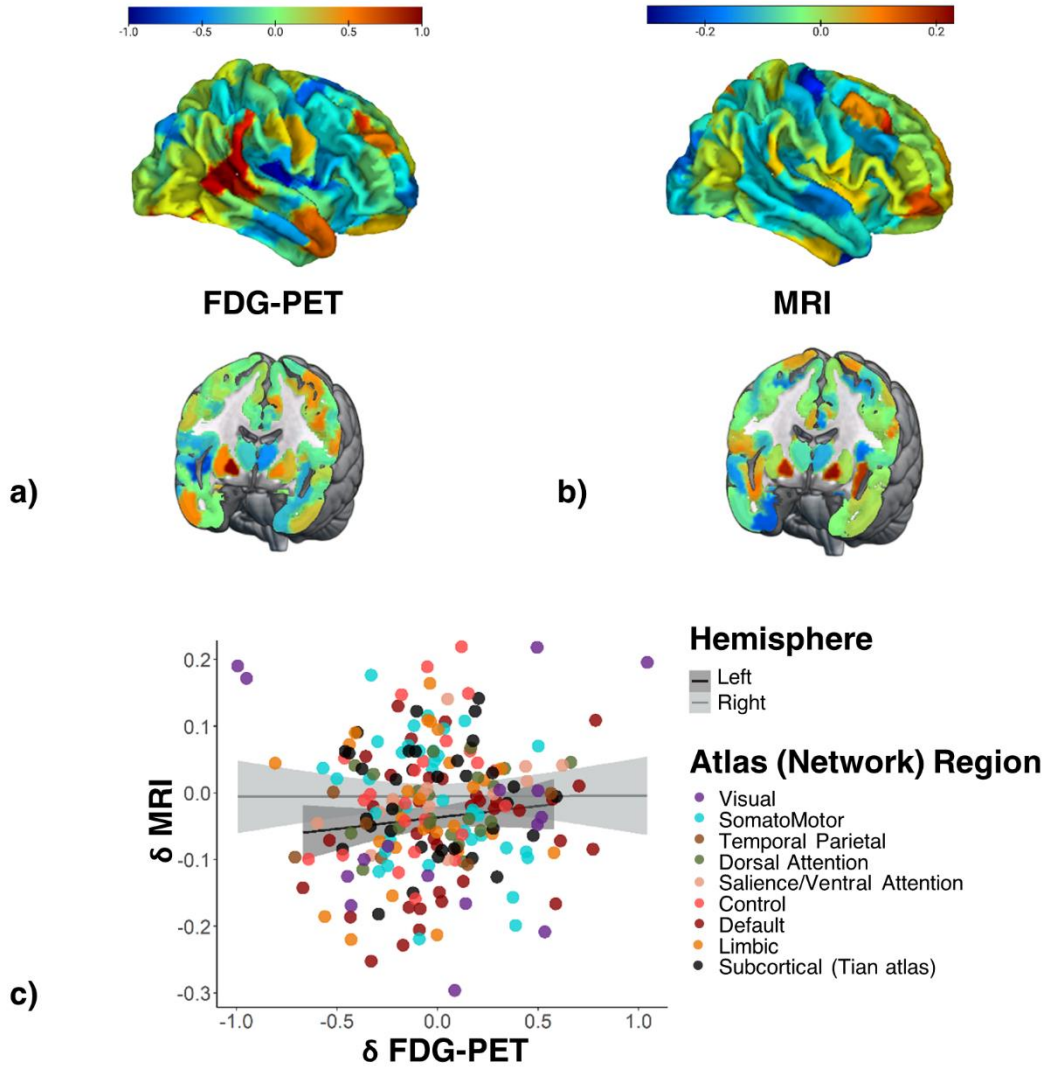


FIGURE 3 Feature importance for brain age prediction. a) Average weights for brain age prediction using FDG-PET. b) Average weights for brain age prediction using MRI. More relevant weights are depicted in blue and red. c) Scatter plot of average feature importance in FDG-PET and MRI and regression lines (per hemisphere).

3.3 BAG and cognitive performance

In CN+SCD_{ADNI}, MRI-derived BAG was associated with lower ADNI-EF scores (normally distributed; $r = -.137$, $p = .01$). A marginally significant correlation was observed between FDG-PET-derived BAG and ADNI-MEM scores (normally distributed; $r = -.100$, $p = .06$).

In MCI_{ADNI}, both, MRI- and FDG-PET-derived BAG were negatively correlated with ADNI-MEM (not normally distributed; $\rho_{\text{MRI}} = -.379$, $p_{\text{MRI}} < .0001$; $\rho_{\text{FDG-PET}} = -.237$, $p_{\text{FDG-PET}} < .0001$) and ADNI-EF (not normally distributed; $\rho_{\text{MRI}} = -.272$, $p_{\text{MRI}} < .0001$; $\rho_{\text{FDG-PET}} = -.243$, $p_{\text{FDG-PET}} < .0001$).

3.4 BAG and AD neuropathology

In CN+SCD_{ADNI}, higher MRI- and FDG-PET-derived BAG were both correlated with a decrease in CSF A β_{1-42} (not normally distributed; $\rho_{\text{MRI}} = -.152$, $p_{\text{MRI}} = .012$; $\rho_{\text{FDG-PET}} = -.152$, $p_{\text{FDG-PET}} = .012$), but none of the other neuropathological variables. In MCI_{ADNI}, higher MRI- and FDG-PET-derived BAG were associated with enhanced amyloid accumulation, as reflected by significant negative correlations with CSF A β_{1-42} (not normally distributed; $\rho_{\text{MRI}} = -.232$, $p_{\text{MRI}} < .0001$; $\rho_{\text{FDG-PET}} = -.144$, $p_{\text{FDG-PET}} = .002$) and (marginal) positive correlations with global AV45 (not normally distributed; $\rho_{\text{MRI}} = .144$, $p_{\text{MRI}} = .004$; $\rho_{\text{FDG-PET}} = .110$, $p_{\text{FDG-PET}} = .026$). Moreover, higher MRI-derived BAG was marginally associated with higher levels of CSF pTau₁₈₁ (not normally distributed; $\rho_{\text{MRI}} = .093$, $p_{\text{MRI}} = .048$).

3.5 BAG and Cognitive Outcome

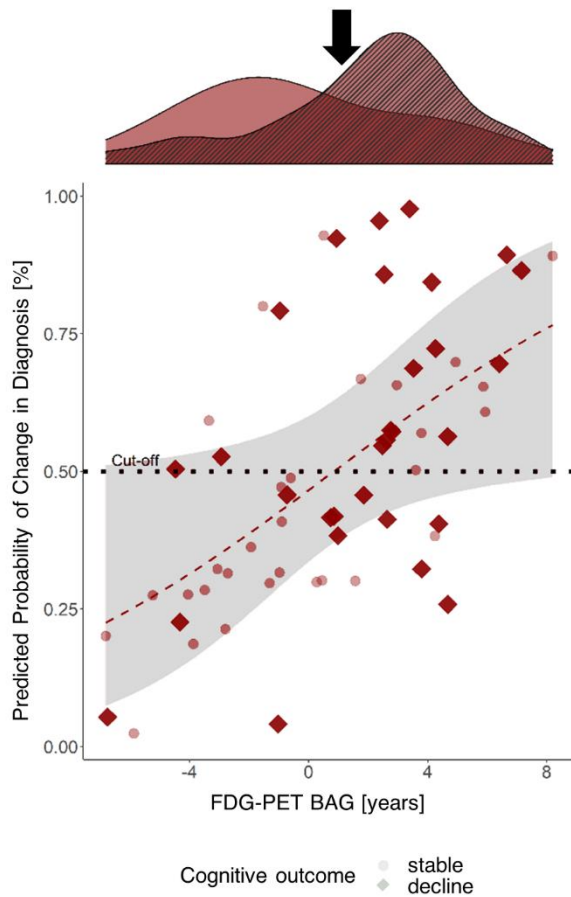
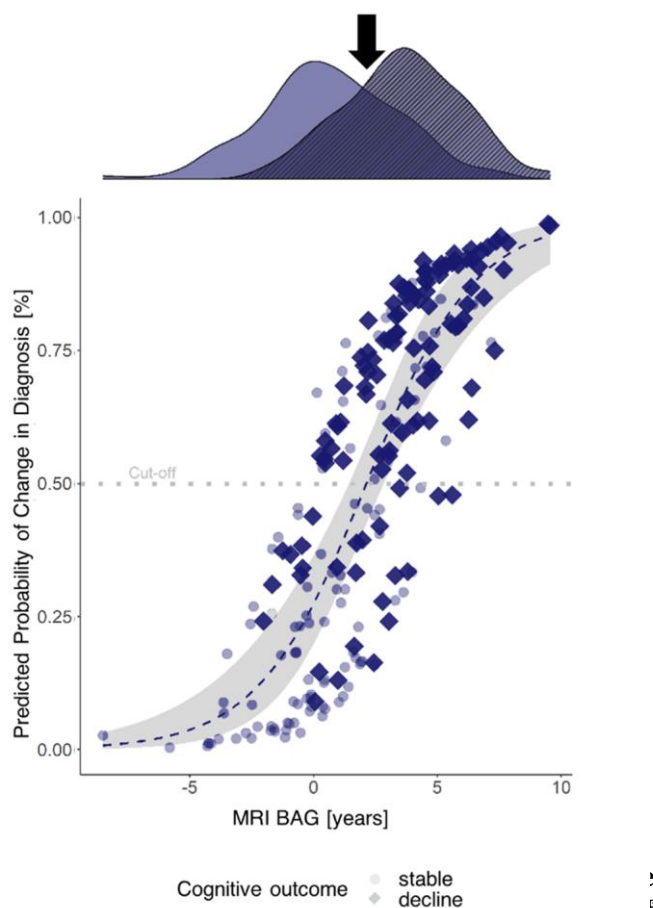


FIGURE 4 Cross-validated probability of a change in diagnosis from CN+SCD to MCI/dementia within two years by FDG-PET BAG. The red line shows the logistic regression on cognitive outcome by FDG-PET BAG, with the shaded area representing standard error. The density plot above shows FDG-PET BAG distribution of stables (clear) and decliners (striped) in the subsample and the black error points to the cutoff.

CN+SCD_{ADNI} sample 1 consisted of 30 decliners and 30 stables. Holding all other predictor variables constant, FDG-PET-derived BAG and APOE- $\epsilon 4$ carriership significantly predicted cognitive outcome at year two follow-up. The odds of a cognitive impairment diagnosis within two years were increased by 29% (95% CI [1.079, 1.604], $p = .010$) for every FDG-PET-derived BAG year and by eightfold with a positive APOE- $\epsilon 4$ carriership status (95% CI [1.496, 71.814], $p = .028$). To obtain a cutoff for prognoses of cognitive impairment, we fitted a logistic regression model on cognitive outcome by BAG on FDG-PET. The intersection of the curve with a 50% probability of receiving such a diagnosis was at 0.85 years FDG-PET BAG (**FIGURE 4**), i.e., cognitively unimpaired individuals with a brain age

advanced by 0.85 years had an elevated risk of converting to cognitive impairment. Stratification by this cutoff in the current CN+SCD_{ADNI} sample 1 yielded a sensitivity of 70% and a specificity of 67% (positive predictive value $PPV = 68\%$, negative predictive value $NPV = 69\%$). We additionally applied the cutoff to the SCD_{DELCODE} cohort ($n_{\text{decliners}} = 8$, $n_{\text{stables}} = 80$), where we obtained a sensitivity of 88% and a specificity of 34% ($PPV = 13\%$, $NPV = 96\%$).

MCI_{ADNI} sample 1 consisted of 113 decliners and 113 stables. Holding all other predictor variables constant, elevated MRI-derived BAG (OR=1.52, 95% CI [1.304, 1.788]), a positive amyloid status in CSF (OR=4.11, 95% CI [1.632, 11.255]) and APOE- $\epsilon 4$ carriership (OR=3.03, 95% CI [1.495, 6.254]) significantly predicted a dementia diagnosis at year-two follow up. The intersection of the curve with a 50% probability of receiving a diagnosis of dementia within two years was at 2.14 years of MRI BAG (**FIGURE 5**). Stratification by this cutoff in the current MCI_{ADNI} subsample yielded a sensitivity of 73% and a specificity of 72% ($PPV = 72\%$, $NPV = 73\%$). In the MCI_{DELCODE} sample ($n_{\text{stables}} = 41$, $n_{\text{decliners}} = 28$), the cutoff had a sensitivity of 68% and a specificity of 73% ($PPV = 63\%$, $NPV = 79\%$).



Analyses excluding subjects without information on amyloid status (sample 2) in both, CN+SCD_{ADNI} ($n=48$) and MCI_{ADNI} ($n=186$) yielded results consistent with those obtained with sample 1 (see SM section 2c).

MCI by MRI-BAG. The blue line shows the logistic regression on cognitive outcome by MRI BAG, with the shaded area representing standard error. The blue line intersects 50% at 2.14 years BAG on MRI. The density plot above shows MRI BAG distribution of stables (clear) and decliners (striped) in the subsample and the black error points to the cutoff.

Finally, given the correlations observed between FDG-PET- and MRI BAG in CN+SCD_{ADNI} and MCI_{ADNI} (see Section 3.2), we assessed logistic regression models with unimodal BAG¹⁸. Considered in

separate models, both MRI- and FDG-PET-BAG significantly predicted cognitive outcome in MCI_{ADNI}, while only FDG-PET predicted cognitive outcome in CN+SCD_{ADNI} (see Tables SM2-SM5 for estimates of logistic regression in sample 1 analyses using only one imaging modality).

4 Discussion

Previous studies have mainly used MRI to estimate brain age. FDG-PET is an early indicator of neurodegeneration-related cerebral changes and a recent study showed for the first time that it could also be used successfully to estimate brain age¹. Here, we compared the accuracy of FDG-PET and MRI-estimated brain age and provided a comprehensive overview of the cognitive and neuropathological profile of FDG-PET and MRI-derived BAG in different cognitive groups. We showed that 1) MRI and FDG-PET both accurately estimated brain age; 2) MRI and FDG-PET-derived BAG both reflect neuropathological abnormalities in CN/SCD and MCI, as well as cognitive dysfunction in MCI, and 3) BAG derived from FDG-PET, but not MRI, yields prognostic value for cognitive outcome in CN/SCD, while MRI yields better prognostic value in MCI. We also calculated and validated cutoffs for predicting cognitive outcome. Importantly, this study is the first to present modality-dependent BAG cutoffs, which can aid in the prognosis of cognitive outcome not only in MCI patients, but even patients without objective cognitive impairment.

Our findings suggest that BAG derived from FDG-PET captures greater and more consistent changes associated with early and subtle neurodegeneration as observed in cognitively unimpaired individuals. On the other hand, MRI-derived BAG was superior in delineating dementia-related changes in MCI⁴. While Lee and colleagues showed that both, MRI and FDG-PET BAG are significantly increased in CN converting to MCI or AD at baseline¹, we demonstrated that prognostic value exists only for FDG-PET BAG. Among the CN population, our results are likely most relevant for individuals with SCD. Individuals with SCD are 1) assumed to recognize cognitive deficits before they become clinically measurable⁷, 2) more likely to develop MCI or AD compared to CN¹⁹, and 3) likely to be seen by a physician given their subjective symptoms. Prediction of cognitive outcome in our cohorts based on FDG-PET BAG was moderately to highly sensitive with moderately to very high NPV. Together, these findings provide strong evidence that FDG-PET BAG could complement the identification of at-

risk individuals, as individuals with a BAG below our proposed cutoff are unlikely to develop cognitive impairment within two years.

Including BAG in AD clinical trials could have several advantages. Numerous anti-amyloid therapies are currently under assessment or have recently been approved for the treatment of MCI and early AD. Since cognitive decline is an important outcome factor of these trials, BAG could support the exclusion of individuals at lower risk of cognitive decline, thereby helping to reduce the number of participants and therefore the cost and time required for treatment trials. Moreover, BAG, which is a summary measure of overall brain health²⁰ and reflects amyloid neuropathology even in cognitively unimpaired individuals, could potentially provide useful information on drug efficacy²¹.

Similar to previous studies^{1,4}, we found differences in brain regions displaying aging as observed on FDG-PET and MRI. Thus, different aging processes may be observed depending on the modality, which underlines the importance of considering the appropriate modality for a research question. The regions deemed most important by our MRI and FDG-PET models have previously been described to be substrates heavily affected by aging⁴, and almost all regions we identified as ‘most important’ for brain age estimation were correlated strongly with age in our ADNI cohort. The greater left-hemisphere involvement in brain age estimation on MRI compared to FDG-PET could explain the better association of MRI-derived BAG with AD biomarkers and cognitive outcome in MCI, as the left hemisphere is known to be affected early on in AD aetiology²². Given the overall strong association of regions deemed important for brain age estimation with AD, our work further supports the claim that AD-related neurodegeneration, at least in part, resembles a form of advanced brain aging.

Some limitations should be acknowledged. First, the sensitivity and especially specificity values for prognoses of cognitive outcome are not high enough to use these measures as standalone biomarkers of cognitive outcome. Moreover, obtaining FDG-PET scans from a cognitively unimpaired population is not straightforward, as it requires logistical availability, high cost, and the injection of a radioactive tracer in the absence of evident cognitive impairment. However, accurate prognoses, particularly for cognitively unimpaired individuals, are difficult to establish and we believe that BAG assessment with a group-dependent choice of modality can aid this process by providing a first indicator of cognitive outcome. Future work should evaluate the combined potential of FDG-PET BAG and APOE-ε4

carriership as a prognostic biomarker of cognitive outcomes. Second, the different FDG-PET scanning protocol of DELCODE (acquisition time: 40-60 min post injection) compared to ADNI and OASIS (acquisition time: 30-60 min post injection) might have influenced generalization of our models to the DELCODE cohort. Yet, we believe that the difference would not be substantial, as we averaged time frames over the entire acquisition time. Moreover, the average BAG (ME) of SCD_{DELCODE} exceeds the previously reported BAG on MRI (1.1 years²³), and the MRI and FDG-PET BAG of MCI patients in our analyses. These differences may be driven by methodological differences, including substantially younger age compared to the study by Rockiki et al., which likely lowers the risk of these individuals to be incipient AD patients, and a different choice of modality (MRI instead of FDG-PET). Whether the FDG-PET BAG is abnormally high in our analyses, or whether higher FDG-PET BAG in SCD reflects very early neurological dysfunction needs further investigation.

In summary, we have shown that MRI and FDG-PET can both be used to estimate brain age and, with some respects, show different benefits depending on the group analyzed: While MRI- and FDG-PET BAG both accurately reflect neuropathological burden across groups and cognitive performance in MCI, only FDG-PET BAG can aid in the prognosis of cognitive outcome in cognitively unimpaired individuals. On the other hand, MRI-derived BAG demonstrates a better estimate for dementia risk in MCI. Estimating cognitive outcome using our BAG cutoffs could complement the identification of patients in need of frequent monitoring at an early time point of cognitive decline and could support clinical trials, both methodologically and financially.

AUTHOR CONTRIBUTIONS

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by EID and GA, with support from KRP and MCH. KRP, TvE, SBE and AD jointly supervised this work. DELCODE data preparation was supervised by MD and HB (PET), EmD (MRI) and FJ (clinical data). The first draft of the manuscript was written by ED and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge data gathering and provision by the ADNI, OASIS and DELCODE studies for the preparation of this manuscript. The authors further thank the Interdisciplinary Program Molecular Medicine at the University of Cologne (IPMM) for their continued support.

DISCLOSURE

KRP and SBE were partly supported by the Deutsche Forschungsgemeinschaft (DFG, PA 3634/1-1 and EI 816/21-1). MCH received funding from the Alzheimer Forschung Initiative. This work was partly supported by the Helmholtz Portfolio Theme “Supercomputing and Modelling for the Human Brain”.

Data collection and sharing for this project was funded by the 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, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; 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 organization 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 Neuro Imaging at the University of Southern California.

The DELCODE study was funded by the German Center for Neurodegenerative Diseases (Deutsches Zentrum für Neurodegenerative Erkrankungen, DZNE: reference number BN012).

KEY POINTS

QUESTION: What is the neuropathological and predictive profile of brain age gaps (BAGs) derived from structural MRI or FDG-PET?

PERTINENT FINDINGS: BAG was computed from structural MRI and FDG-PET and subsequently associated with neuropathological markers of Alzheimer's disease, as well as risk of cognitive deterioration. While both, MRI- and FDG-PET-derived BAG were indicative of existing amyloid pathology already in individuals without cognitive impairment, the predictive capacity of BAG for cognitive outcome was group-dependent: FDG-PET-derived BAG predicted cognitive deterioration in cognitively unimpaired individuals and MRI-derived BAG predicted cognitive deterioration in patients with mild cognitive impairment.

IMPLICATIONS FOR PATIENT CARE: A group-dependent choice of modality for BAG assessment can complement care management plans of cognitively unimpaired and impaired individuals by providing estimates of cognitive outcomes.

References

1. Lee J, Burkett BJ, Min H-K, et al. Deep learning-based brain age prediction in normal aging and dementia. *Nat Aging*. 2022;2:412–424. doi:10.1038/s43587-022-00219-7
2. Löwe LC, Gaser C, Franke K. The effect of the APOE genotype on individual BrainAGE in normal aging, Mild cognitive impairment, and Alzheimer's Disease. *PLoS One*. 2016;11(7):e0157514. doi:10.1371/journal.pone.0157514
3. Gaser C, Franke K, Klöppel S, Koutsouleris N, Sauer H. BrainAGE in Mild Cognitive Impaired Patients: Predicting the Conversion to Alzheimer's Disease. *PLoS One*. 2013;8(6):e67346. doi:10.1371/journal.pone.0067346
4. Dukart J, Kherif F, Mueller K, et al. Generative FDG-PET and MRI Model of Aging and Disease Progression in Alzheimer's Disease. *PLoS Comput Biol*. 2013;9(4):e1002987. doi:10.1371/journal.pcbi.1002987
5. LaMontagne PJ, Benzinger TLS, Morris JC, et al. OASIS-3: Longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and Alzheimer disease. *medRxiv*. Published

online 2019. doi:10.1101/2019.12.13.19014902

6. Jessen F, Spottke A, Boecker H, et al. Design and first baseline data of the DZNE multicenter observational study on predementia Alzheimer's disease (DELCODE). *Alzheimer's Res Ther.* 2018;10(1):15. doi:10.1186/s13195-017-0314-2
7. Jessen F, Amariglio RE, Van Boxtel M, et al. A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. *Alzheimer's Dement.* 2014;10(6):844-852. doi:10.1016/j.jalz.2014.01.001
8. Albert MS, DeKosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dement.* 2011;7(3):270-279. doi:10.1016/j.jalz.2011.03.008
9. Jack CR, Bernstein MA, Fox NC, et al. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *J Magn Reson Imaging.* 2008;27(4):685-691. doi:10.1002/jmri.21049
10. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine learning in Python. *J Mach Learn Res.* 2011;12:2825-2830. Accessed April 20, 2021. <http://scikit-learn.sourceforge.net>.
11. Schaefer A, Kong R, Gordon EM, et al. Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cereb Cortex.* 2018;28(9):3095-3114. doi:10.1093/cercor/bhx179
12. Tian Y, Margulies DS, Breakspear M, Zalesky A. Topographic organization of the human subcortex unveiled with functional connectivity gradients. *Nat Neurosci.* 2020;23(11):1421-1432. doi:10.1038/s41593-020-00711-6
13. Beheshti I, Nugent S, Potvin O, Duchesne S. Bias-adjustment in neuroimaging-based brain age frameworks: A robust scheme. *NeuroImage Clin.* 2019;24:102063. doi:10.1016/j.nicl.2019.102063
14. Beheshti I, Ganaie MA, Paliwal V, Rastogi A, Razzak I, Tanveer M. Predicting Brain Age Using Machine Learning Algorithms: A Comprehensive Evaluation. *IEEE J Biomed Heal Informatics.* 2022;26(4):1432-1440. doi:10.1109/JBHI.2021.3083187
15. Crane PK, Carle A, Gibbons LE, et al. Development and assessment of a composite score for memory in the Alzheimer's Disease Neuroimaging Initiative (ADNI). *Brain Imaging Behav.* 2012;6(4):502-516. doi:10.1007/s11682-012-9186-z
16. Gibbons LE, Carle AC, Mackin RS, et al. A composite score for executive functioning, validated in Alzheimer's Disease Neuroimaging Initiative (ADNI) participants with baseline mild cognitive impairment. *Brain Imaging Behav.* 2012;6(4):517-527. doi:10.1007/s11682-012-9176-1
17. Hansson O, Seibyl J, Stomrud E, et al. CSF biomarkers of Alzheimer's disease concord with amyloid- β PET and predict clinical progression: A study of fully automated immunoassays in BioFINDER and ADNI cohorts. *Alzheimer's Dement.* 2018;14(11):1470-1481. doi:10.1016/j.jalz.2018.01.010
18. Ranganathan P, Pramesh C, Aggarwal R. Common pitfalls in statistical analysis: Logistic regression. *Perspect Clin Res.* 2017;8(3):148-151. doi:10.4103/picr.PICR_87_17
19. Parfenov VA, Zakharov VV, Kabaeva AR, Vakhnina NV. Subjective cognitive decline as a predictor of future cognitive decline a systematic review. *Dement e Neuropsychol.* 2020;14(3):248-257. doi:10.1590/1980-57642020dn14-030007
20. Cole JH, Marioni RE, Harris SE, Deary IJ. Brain age and other bodily 'ages': implications for neuropsychiatry. *Mol Psychiatry.* 2019;24(2):266-281. doi:10.1038/s41380-018-0098-1

21. Van Gestel H, Franke K, Petite J, et al. Brain age in bipolar disorders: Effects of lithium treatment. *Aust N Z J Psychiatry*. 2019;53(12):1179-1188. doi:10.1177/0004867419857814
22. Pfeil J, Hoenig MC, Doering E, van Eimeren T, Drzezga A, Bischof GN. Unique regional patterns of amyloid burden predict progression to prodromal and clinical stages of Alzheimer's disease. *Neurobiol Aging*. 2021;106:119-129. doi:10.1016/j.neurobiolaging.2021.06.014
23. Rokicki J, Wolfers T, Nordhøy W, et al. Multimodal imaging improves brain age prediction and reveals distinct abnormalities in patients with psychiatric and neurological disorders. *Hum Brain Mapp*. 2021;42(6):1714-1726. doi:10.1002/hbm.25323

Table 1. Overview of samples

| | CN+SCD _{ADNI} | CN _{OASIS} | SCD _{DELCODE} | MCI _{ADNI} | MCI _{DELCODE} |
|---|------------------------|---------------------------|--------------------------|---------------------|------------------------|
| <i>n</i> total | 376 | 59 | 88 | 596 | 80 |
| Age at PET scan [avg. years (SD)] | 73.9 (5.94) | 71.7 ⁻ (4.22) | 70.9 ⁻ (5.57) | 73.2 (6.93) | NA |
| Age at MRI scan [avg. years (SD)] | 73.8 (5.92) | 70.36 ⁻ (4.17) | NA | 73.2 (6.92) | 73.4 (5.87) |
| Sex [%female (<i>n</i> _{NA})] | 51 (0) | 59 (0) | 41 (0) | 42 (2) | 45 (0) |
| CSF A β ₁₋₄₂ Status [%positive (<i>n</i> _{NA})] | 39 (85) | NA | 43 (28) | 65 (139) | 52 (38) |
| MMSE [avg. score] | 29 (1.23) | 29 (1.01) | 29 ⁺ (1.03) | 28 (1.75) | 28 (1.67) |
| Education [avg. years (SD)] | 16 (2.71) | 16 (2.70) | 16 ⁻ (3.00) | 16 (2.67) | 14 ⁻ (3.06) |

Notes. Percentage of CSF A β ₁₋₄₂ status indicates percentage of amyloid positive individuals among all who received lumbar puncture (excluding NA). ⁺significantly higher than CN+SCD_{ADNI}, ⁻significantly lower than CN+SCD_{ADNI}. Comparisons within modality and group, with $\alpha = 0.05$.

Table 2. Accuracy of predicting chronological age from FDG-PET and MRI scans.

| | CN+SCD _{ADNI} | | CN _{OASIS} | | MCI _{ADNI} | | SCD _{DELCODE} | MCI _{DELCODE} |
|----------------|------------------------|------------------|---------------------|-----------------|---------------------|----------------|------------------------|------------------------|
| | MRI | FDG | MRI | FDG | MRI | FDG | FDG | MRI |
| <i>n</i> total | 357 ⁺ | 357 ⁺ | 52 ⁺ | 52 ⁺ | 596 | 596 | 88 | 80 |
| MAE | 1.96 | 2.63 | 2.23 | 2.03 | 2.62 | 2.51 | 2.64 | 2.62 |
| Range | [-6.91, 9.09] | [-8.72, 9.98] | [-5.44, 7.40] | [-6.70, 5.46] | [-9.07, 9.55] | [-10.91, 9.23] | [-5.49, 7.92] | [-4.64, 9.60] |
| ME | 0.04 | -0.02 | -0.71 | -0.04 | 1.51 | 1.07 | 2.07 | 1.42 |
| R ² | .816 | .693 | .593 | .617 | .773 | .794 | .641 | .676 |

Notes. ⁺After outlier exclusion using CN train set (IQR > 6).