**MANUSCRIPT DRAFT**

**COMPARISON OF STRUCTURAL AND METABOLIC BIOMARKERS OF NEURODEGENERATION FOR BRAIN AGE PREDICTION USING MACHINE LEARNING**

E. Doering1,2, G. Antonopoulos3, M. Hönig1,4, T. van Eimeren1,2, S. Eickhoff3,5, K. Patil5, A. Drzezga1,2,4

1University Hospital Cologne, Clinic and Policlinic for Nuclear Medicine, Köln, Germany, 2German Center for Neurodegenerative Diseases, Positron Emission Tomography, Bonn, Germany, 3Forschungszentrum Jülich, Brain and Behavior (INM-7), Jülich, Germany, 4Forschungszentrum Jülich, Molecular Organization of the Brain (INM-2), Jülich, Germany, 5Heinrich-Heine-University, Institute of Systems Neuroscience, Düsseldorf, Germany

Keywords: biological age, machine learning, Alzheimer’s disease

**Journal: Nature Aging**

**Abstract**

XXX (will do this at the end)

Main text – up to 4,500 words, excluding abstract, Methods, references and figure legends.

Abstract – up to 150 words, unreferenced.

Display items – up to 8 items (figures and/or tables).

Article should be divided as follows:

Introduction (without heading)

Results

Discussion

Online Methods. ​

Results and online Methods should be divided by topical subheadings.

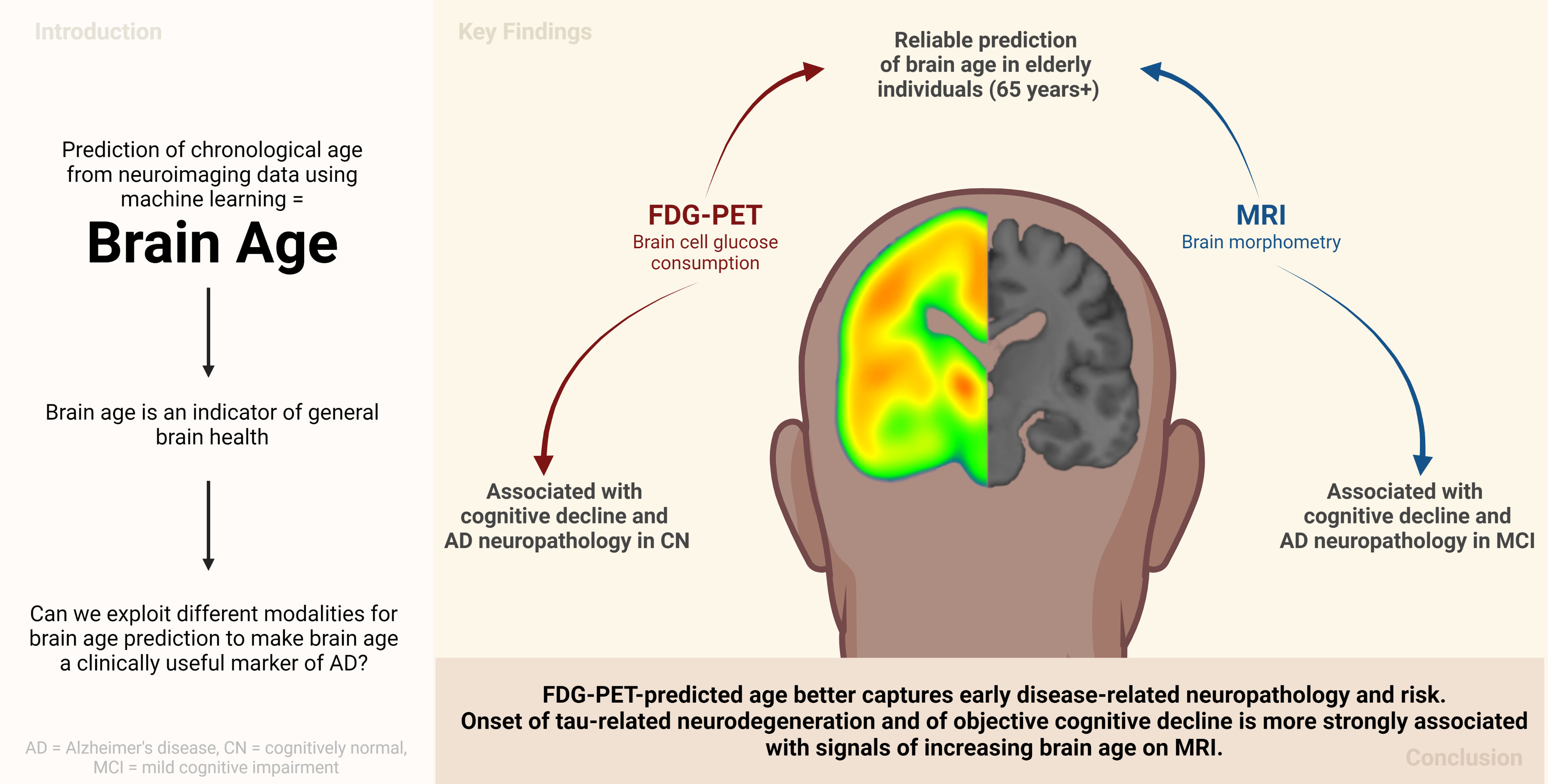
References – as a guideline, we typically recommend up to 60.

Articles include received/accepted dates.

Articles may be accompanied by supplementary information.

Articles are peer reviewed.

Clinical and public health research Articles may have longer abstracts to accommodate statistical information and must include a paragraph on limitations in the Discussion section.



**1 Introduction**

Biological aging entails the change or decline of various physiological functions. Human biological aging, as opposed to chronological aging, differs depending on the tissue under investigation – and advanced biological age of specific organs is associated with diseases of the same1. For example, advanced brain age, i.e., a positive deviation of brain age from chronological age (*brain-predicted age difference, “BPAD”*) has been repeatedly associated with presence, or future development of dementia due to Alzheimer’s disease (AD)2–4. AD is a neurodegenerative disease with an age of onset at 65 years, which is characterized by severe cognitive impairment. Biologically, AD has two neuropathological hallmarks, namely the accumulation of amyloid plaques and tau tangles. Abnormal amyloid deposition precedes symptom onset by decades5, thus providing an early target for newly emerging anti-amyloid drugs6,7 and rendering early detection of pathological abnormality pivotal. While BPAD has been associated with general diagnoses across the AD continuum2 and progression from mild cognitive impairment (MCI) to AD4,8,9, less is known about the extent to which BPAD can be used as a marker of AD neuropathology and a predictive marker of cognitive decline, especially for individuals without initial cognitive impairment. Notably, to our knowledge, no cut-off value of BPAD for elevated risk of cognitive decline, in potential dependence on other risk factors, has yet been published.

Brain age can be modeled via machine learning algorithms, by predicting a person’s chronological age from their neuroimaging data. This is most commonly achieved using structural magnetic resonance imaging data (MRI). MRI depicts anatomical changes of the brain, such as atrophy, and it represents one of two imaging biomarkers of neurodegeneration in AD. 18F-Fluorodeoxyglucose-PET (FDG-PET) is the second imaging biomarker of neurodegeneration, and it unravels the molecular changes in cerebral cell metabolism. In comparison, changes in cerebral metabolism have been shown to precede volumetric loss in the course of AD, thus suggesting different timelines of the two biomarkers10. One recent study has shown slightly better performance of brain age prediction in a cognitively normal (CN) population using FDG-PET, compared to MRI4. Therefore, FDG-PET-derived BPAD possibly yields previously unexplored associations of advanced brain aging and pathological/clinical disease severity especially at earlier stages of AD. A stage-dependent choice of modality for brain age prediction has been completely neglected in the existing literature.

Here, we tackled the above mentioned gaps by investigating FDG-PET and MRI-based brain age prediction using a cohort of cognitively normal individuals, and MCI patients. First, we compared the precision of brain-predicted age using FDG-PET or MRI in CN, and subsequently manifested the better modality to depict cognitive performance or AD neuropathology in CN and mild cognitive impairment (MCI). Finally, we applied machine learning classification to predict cognitive decline (CD) from BPAD and known risk factors in CN and MCI, and subsequently calculated a cut-off value of BPAD for elevated risk of cognitive decline.

**2 Results**

**2.1 Participants**

This study included 879 FDG-PET and MRI scans (respectively) from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (“CN” and “MCI” samples, adni.loni.usc.edu) and 59 from the Open Acces of Imaging Studies-3 database (OASIS-3). Scans from the ADNI were selected such that FDG-PET and MRI scans from the same individual were not more than one year apart. Data was split into three samples: the “main” ADNI sample of CN (“CN”, n = 367) was used to train models and yield predictions for later association of BPAD with cognitive performance, neuropathology and cognitive decline in CN. The small sample of CN derived from the OASIS-3 (“CN\_validation”, n = 59) was used to validate prediction accuracy (mean absolute error, MAE) in an external dataset. Finally, predictions for the ADNI sample of 513 MCI patients (“MCI”) were used to associate BPAD with cognitive performance, neuropathology and cognitive decline in MCI. To be included, individuals included had to be 65 years or older, while OASIS participants were significantly younger than ADNI participants (p < .01), especially in the MRI cohort. Compared to CN, participants in the MCI sample had a significantly lower percentage of females (χ2 = 1.5, p < .01), lower MMSE (t(424)=5.38, p < .001) and a higher percentage of amyloid positivity (χ2 = 43.7, p < .001).

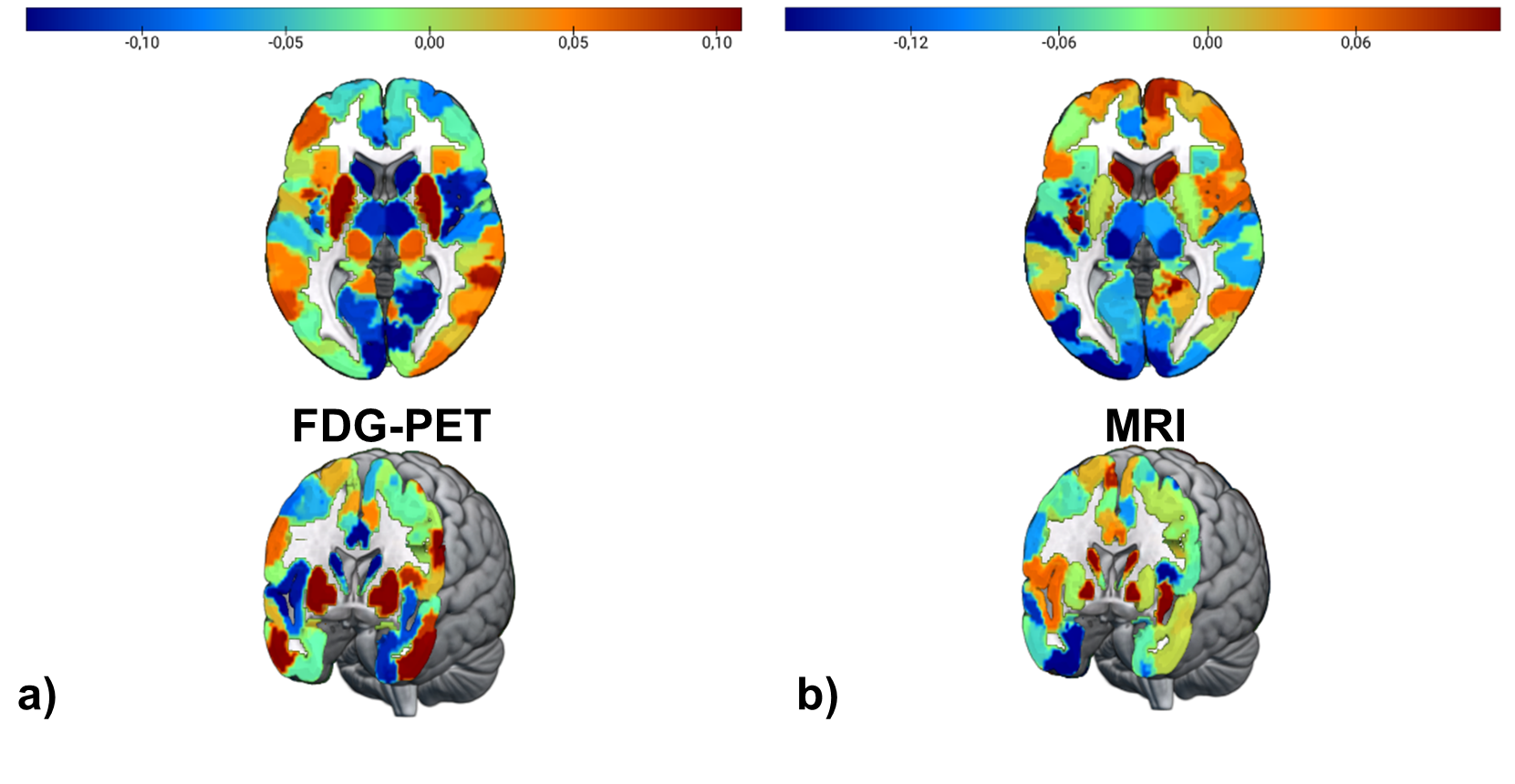
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 1.** Overview of samples | | | | |
|  | CN | CN\_validation | MCI |
| *n* total | 367 | 59 | 513 |
| Age [avg. years (SD)] | 74.2 (5.68) (PET)  74.2 (5.67) (MRI) | 71.7 (4.15) (PET)/  70.4 (4.17) (MRI) | 74.9 (5.77) |
| Sex [%female] | 51 | 59 | 40 |
| CSFAβ1-42 Status  (-/+/NA) | 171/111/85 | NA | 121/270/122 |
| MMSE [avg. score] | 29 (1.24) | 29 (.85) | 28 (1.77) |
| Education [avg. years (SD)] | 16 (2.72) | 16 (2.70) | 16 (2.70) |

**2.2 Precision of brain-predicted age**

To compare the potential of FDG-PET and MRI to predict chronological age, we used a nested five-fold cross-validation approach, yielding one test prediction for (almost) every subject in the ADNI CN sample, and five test predictions for each subject in the CN\_validation and MCI sample. Two types of models previously recommended for small sample sizes11 were implemented for brain age prediction: support vector regression (SVR) and relevance vector regression (RVR). SVR models outperformed RVR models in each fold of the outer-loop cross-validation in both modalities. Table 2 presents the results of our modality comparison. Regional FDG-PET and MRI predicted chronological age comparably well, i.e., with low mean absolute errors (MAE = |BPA – CA|). In the ADNI-derived CN test sets, individuals’ BPA as assessed with FDG-PET and MRI on average very close to their actual age, as inferred from the very low mean difference (MD = BPA – CA), thus demonstrating high potential to capture brain aging in a CN cohort. The OASIS-derived CN\_validation sample was used to validate our findings in an external dataset and showed similar MAEs as the ADNI sample, although chronological age was slightly better predicted from FDG-PET as compared to MRI, reflected in a lower MAE across the five models. A probable explanation for the higher MAE in the CN\_validation MRI sample is that the age distribution in this sample deviated more from the age distribution in the ADNI train sets compared to CN\_validation PET samples. In the MCI test sets, individuals’ brain-predicted age as assessed with FDG-PET and MRI was on average .77 and 1.57 years older than their chronological age, respectively, thus reflecting the expected advancement of brain age in this population. The bias elimination procedure successfully eliminated the correlation between chronological age and BPAD in the CN test sets, while some correlations remained marginally significant (*pFDG-PET* = [0.08,0.69], *pMRI* = [0.27, 0.86]). In the MCI sample, bias was not eliminated (*pFDG-PET* = [1.15 x 10-7, 0.70], *pMRI* = [0.001, 0.06].Therefore, age was entered as a co-variate in subsequent analyses.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Table 2.** Precision of predicting chronological age from FDG-PET and MRI scans. For CN\_validation and MCI, results of the first model and metrics over all five models are shown. | | | | | | | |
|  | CN | | CN\_validation | | MCI | | |
|  | FDG | MRI | FDG | MRI | FDG | MRI | |
| *n* total | 345⁺ | 345⁺ | [52,54]\*⁺ | [52,54]\*⁺ | 513 | | 513 | |
| MAE | 1.99 | 1.89 | 1.83 | 2.43 | 1.96 | | 2.68 | |
| MAE before bias correction | 4.04 | 3.97 |  |  |  | |  | |
| Mean (SD) over 5 models | - | - | 2.04 (.30) | 2.45 (.19) | 2.18 (.43) | | 2.50 (.12) | |
| Mean difference | -.10 | -.05 | -.80 | -.80 | .78 | | 1.75 | |
| Mean (SD) over 5 models | - | - | -.66 (.41) | -.92 (.16) | .77 (.26) | | 1.57 (.16) | |
| *Notes.* +After outlier exclusion using CN train set (IQR > 6), \*range across predictions of five final models | | | | | | | |

Three out of five, and five out of five optimal models were support vector machines with linear kernels. To assess brain regions important for the prediction of chronological age, we extracted the mean weight coefficients of these models. For non-linear kernels, weight coefficients are not available. Regional weight coefficients were strongly correlated within modalities (FDG-PET: *r* = [0.80, 0.83], MRI: *r* = [0.87, 0.90]), but average weight coefficients were not correlated between the two modalities (*r*(214) = 0.045, *p* = 0.483), i.e. the regions used for brain age prediction in the two modalities were substantially different (see **Fig 1**).



**Fig 1** **Feature importance for brain age prediction.** a) Average weights of support vector regression across three linear SVR for brain age prediction using FDG-PET. Weights were highly correlated across models (r > 0.8). b) Average weights of support vector regression across five linear SVR for brain age prediction using MRI. Weights were highly correlated across models (r > 0.8).

**2.3 BPAD and Cognitive Performance**

Partial spearman correlations between cross-sectional BPAD and memory (ADNI-MEM) and executive function scores (ADNI-EF) were calculated to evaluate whether BPAD is associated with cognitive performance in the two modalities. Age and sex were used as covariates. To adjust for multiple comparison, threshold levels of significance were adjusted by Bonferroni correction (p = .025). In CN (n = 345), there was no significant partial correlation between BPAD and ADNI-MEM in either modality when controlling for age and sex. A weak, negative, partial correlation was detected between MRI-BPAD and ADNI-EF (*ρ*(341)=-.150, *p* = .005).

In MCI (n = 511), significant, negative partial correlations between BPAD and ADNI-MEM, as well as between BPAD and ADNI-EF existed with BPAD derived from either modality and from each of the five models (Table 3). Across models, median correlation coefficients were significantly stronger between MRI-BPAD and ADNI-MEM (*z* = 3.56, *p* < .001) compared to FDG-BPAD.

**2.4 BPAD and AD Neuropathology**

Partial spearman correlations were calculated between cross-sectional BPAD and PET amyloid status (global AV45), CSF β-amyloid1–42 (CSF Aβ1-42), CSF total-tau (CSF Tau) and CSF phospho-tau181 (CSF pTau181) to evaluate whether BPAD is associated with AD neuropathology in the two modalities. Age and sex were used as covariates. To adjust for multiple comparison, threshold levels of significance were adjusted by Bonferroni correction (p = .0125). In CN (n = 266), a weak, negative, partial correlations existed between FDG-BPAD and CSF Aβ1-42 (*ρ*(262)= -.160, *p* = .009). MRI-BPAD was also partially correlated with CSF Aβ1-42 (*ρ(262)* = -.126, *p* = .040), however this correlation did not withstand Bonferroni correction. Other neuropathological measures were not associated with BPAD in CN.

In MCI (n = 392 for CSF; n = 345 for AV45), partial correlations between BPAD and AD neuropathology revealed that FDG-BPAD was only marginally correlated with CSF Aβ42 across models, not always passing correction multiple comparison (CSF Aβ1-42: p < .05). MRI-BPAD was significantly correlated with measures of amyloid across models. Moreover, partial correlations were observed between MRI-BPAD and (p-)tau, which, however, did not withstand multiple comparison adjustment in the predictions of two (CSF total tau) and one (CSF pTau181) model(s) (CSF Tau: *p* = .006 - .046; pTau: *p* = .004 – .025).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Table 3. Correlation strength between BPAD and neuropathology/cognitive function in MCI patients across five different models. | | | | |
|  | FDG-PET | | MRI | |
|  | Zero-order | Partial | Zero-order | Partial |
| CSF Aβ42 | -.184  [-.215, -.150] | -.174  [-.216, -.120] | -.290  [-.294, .283] | -.262  [-.264, -.258] |
| Global AV45 | ns | ns | .204  [.189, .225] | .196  [.183, .205] |
| CSF Tau | ns | ns | ns | ns |
| CSF pTau | ns | ns | ns | ns |
| ADNI-MEM | -.236 [-.243; -.207] | -.208  [-.224; -.196] | -.437  [-.442; -.422] | -.409  [-.419; -.392] |
| ADNI-EF | -.237 [-.272; -.231] | -.224  [-.246; -.203] | -.300  [-.339; -.307] | -.290  [-.301; -.286] |
| *Notes.* Median [range] of Spearman correlation coefficients are displayed when significant correlation existed in brain-predicted age according to all five models (p < 0.0125 for neuropathology, p < 0.025 for cognitive performance). ns = not significant in all five models. | | | | |

**2.5 BPAD and Cognitive Decline**

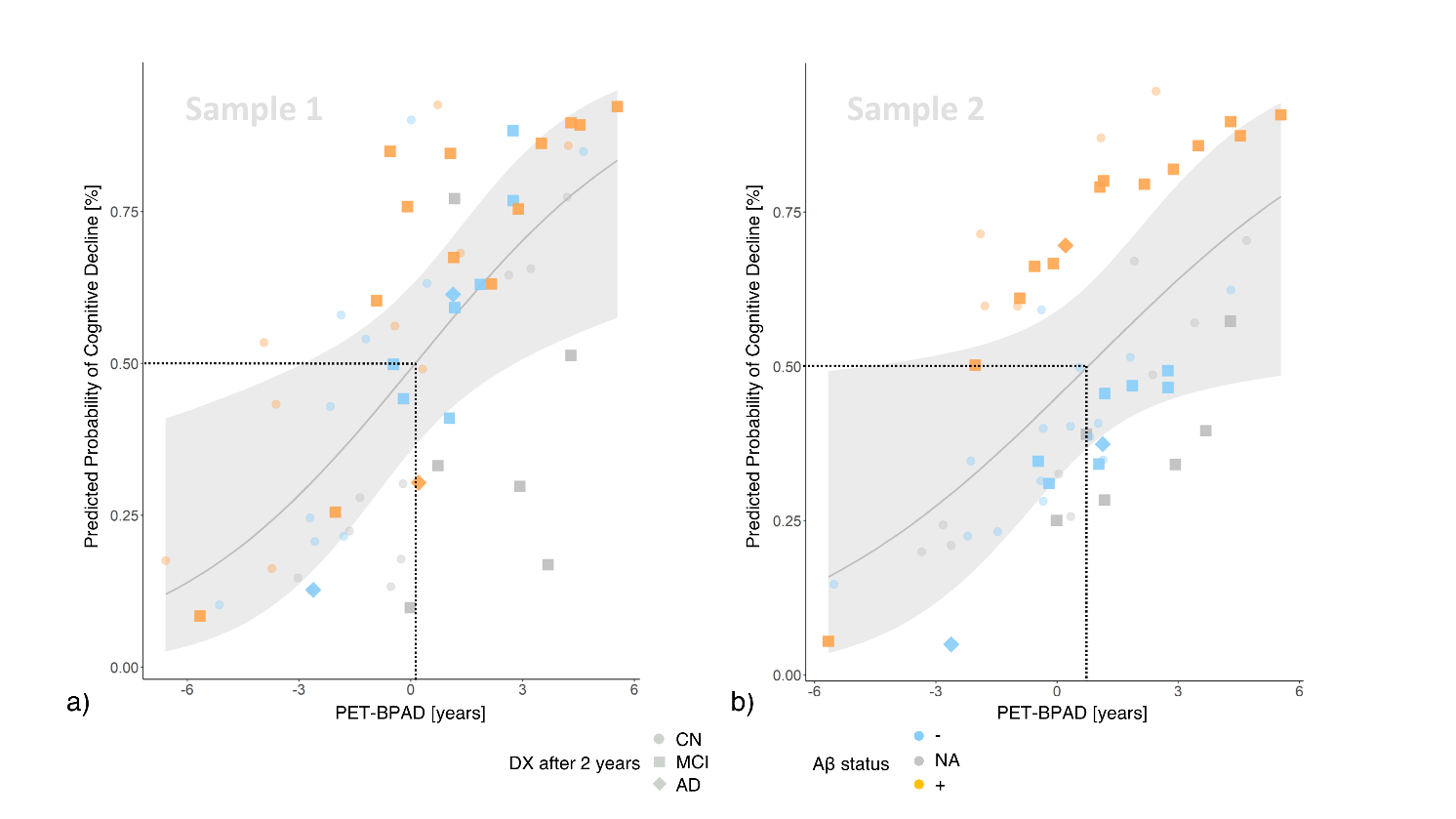
To assess the potential of BPAD in the two modalities to serve as an indicator of CD, and to calculate cut-off values for elevated risk of cognitive decline in CN and MCI, individuals’ diagnosis at year two was predicted from PET-BPAD, MRI-BPAD, APOE-e4 carriership, amyloid status and years of education. Here, we applied 10-fold cross-validated logistic regression in two subsamples per group (CN/MCI), each containing all individuals who show cognitive decline within two years, and an exclusive matched sample of non-decliners (matched in number by age and sex). As amyloid status was not available for all individuals, analyses were conducted two-fold: once including individuals with missing amyloid information (NA values coded as separate category and amyloid negativity coded as reference; “whole samples”), and once excluding these individuals (“complete samples”, results in Supplementary Materials).

Table 4 presents an overview of logistic regression estimates and p-values on the whole samples. Two hundred ninety eight individuals from the baseline CN sample received either a CN diagnosis at year two (“stables”; n = 269), or a diagnosis of cognitive impairment (MCI or dementia) six months to two years after baseline (“decliners”; n = 29). PET- and MRI-BPAD were not significantly correlated in the two samples. In sample 1, PET-BPAD significantly predicted CD with an odds ratio (OR) of 1.404 (95% CI [1.113, 1.874]). In sample 2, PET-BPAD (OR = 1.298, 96% CI [1.013, 1.734]) and amyloid status (OR = 5.011, 95% CI [1.197, 25.363]) marginally to significantly predicted CD. In both samples, we then fit a new logistic regression model using only (marginally) significant variables to determine a cut-off for BPAD, that is, PET-BPAD for sample 1 and amyloid status and PET-BPAD for sample 2. These models had an AUC of 71% and 73%, respectively. 50% disease probability (our criterion for the cut-off) was reached at 0.2 and 0.4 years PET-BPAD in samples 1 and 2, respectively (Fig. 3), indicating that any positive deviation of BPA from CA yields an elevated risk for CD. Stratified by amyloid status (only significant in sample 2), we observed that that the cut-off for CD in amyloid positive CN was considerably lowe (-3.1 years PET-BPAD) compared to amyloid negative CN (2.1 years PET-BPAD; Fig. 3b). This underlines that amyloid pathology increases the risk for CD in the metabolically normal, or even younger brain.

After removing those individuals who did not have information on amyloid status available, a complete sample of 23 decliners remained, thus constituting a sample size of 46. Results from the complete samples were largely consistent with results obtained from the whole samples and can be found in the Supplementary Materials, section “Prediction of Cognitive Decline”.

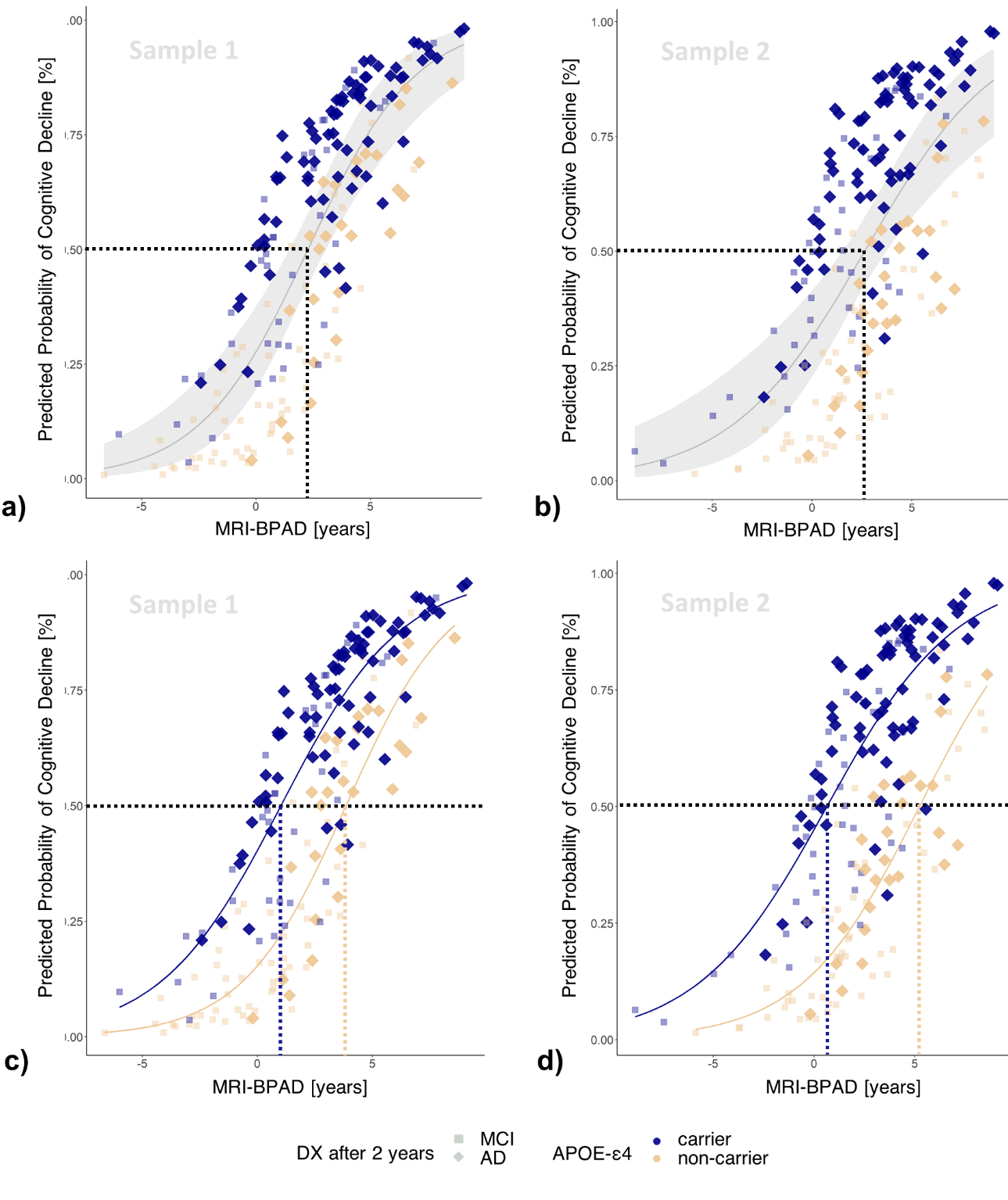
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Table 4 Estimates (p-values) of logistic regression for prediction of cognitive decline. | | | | |
|  | CN | | MCI | |
|  | Sample  (n = 58) | Sample 2  (n = 58) | Sample 1  (n = 200) | Sample 2  (n = 200) |
| PET-BPAD [Years] | .340 (.008) | .261 (.052) | .132 (.122) | .172 (.040) |
| MRI-BPAD [Years] | .037 (.784) | .032 (.769) | .362 (<.0001) | .320 (<.0001) |
| Aβ+ | .026 (.973) | 1.612 (.036) | 1.407 (.009) | .603 (.292) |
| APOE-ε4+ | 1.464 (.104) | -.066 (.923) | .775 (.046) | 1.466 (<.001) |
| Education [Years] | -.136 (.221) | -.060 (.571) | -.031 (.646) | -.087 (.213) |

In the whole samples of MCI, 393 individuals either maintained an MCI diagnosis until year two (“stables”; n = 293) or received a clinical diagnosis of dementia six months to two years after baseline (“decliners”; n = 100). Here, we present the logistic regression results using brain age predictions from the first of five models. Results from models two to five can be found in the Supplementary materials (Table S1) and are highly concordant with the results presented here. Across the two matched sub-samples, PET- and MRI-BPAD were moderately strongly correlated (r­sample1 = .439; psample1 < .0001; rsample2 = .372; psample2 < .0001). In both samples, higher MRI-BPAD very significantly predicted CD (sample 1: OR = 1.436, 95% CI [1.241, 1.688]; sample 2: OR = 1.377, 95% CI [1.202, 1.599]) together with APOE-ε4 carriership. Notably, MRI-BPAD showed considerably higher significance compared to other risk factors (see Table 3). Odds ratios and confidence intervals of both samples are available in Supplementary Fig. S1. Logistic regression models with only the significant variables (sample 1: MRI-BPAD, CSF Aβ1-42 and APOE-ε4 carriership; sample 2: PET-BPAD, MRI-BPAD and APOE-ε4 carriership) yielded AUCs of 80% and 79% in samples 1 and 2, respectively, with a cut-off for CD at 2.3 and 2.5 years MRI-BPAD. Stratified by APOE-ε4 carriership, we observed that a lower cut-off was required for APOE-ε4 carriers (sample 1: 1.1 years MRI-BPAD, sample 2: 0.7 years MRI-BPAD) compared to non-carriers (sample 1: 3.8 years MRI-BPAD, sample 2: 5.3 years MRI-BPAD, see Fig. XX). Stratified by amyloid status (only significant in sample 1), a similar picture emerged, where amyloid positive MCI patients had a lower cut-off for CD (2.0 years MRI-BPAD) compared to amyloid positive MCI patients (3.5 years MRI-BPAD). While the MCI brain shows advanced aging patterns on MRI, a BPAD of over 2 years thus depicts increased risk for cognitive decline, while individuals with genetic risk for AD show an even lower threshold.

**

**Fig. 3 Cross-validated probability of CD within two years after a baseline diagnosis of CN by PET-BPAD.** PET-BPAD predicted CD within two years with variable control groups in the whole samples (displayed here) and in one reduced sample. Stable individuals made transparent for visibility. a) Higher PET-BPAD was the only predictor of CD in sample one. The PET-BPAD-derived threshold for CD in sample 1 (50% probability of CD; dotted line) was .2 years. b) Higher PET-BPAD and amyloid positivity predicted CD in sample two. The PET-BPAD-derived decision boundary in sample 2 was .9 years. Shaded area represents standard error. CD = cognitive decline; DX = diagnosis.

86 MCI patients had full information on all considered variables, thus constituting the decliner group of the complete samples. Results from the complete samples were largely consistent with results obtained from the whole samples and can be found in the Supplementary Materials, section “Prediction of Cognitive Decline”. Finally, due to the correlation observed between PET- and MRI-BPAD in the MCI sample, we additionally assessed logistic regression models with unimodal BPAD12. Considered in separate models, both MRI- and PET-BPAD very significantly predicted CD (see Supplementary Tables S2 and S3 for estimates of logistic regression in whole samples) together with APOE-e4 carriership (and amyloid status). However, MRI-BPAD continued to show higher significance compared to PET-BPAD.

**Fig. 4 Cross-validated probability of cognitive decline within two years after a baseline diagnosis of MCI by MRI-BPAD.** Higher MRI-BPAD and a positive amyloid status, together with higher XX PET-BPAD and a positive APOE-ε4 carriership predicted CD within two years with variable control groups. Stable individuals made transparent for visibility. a) The MRI-BPAD-derived threshold for CD in sample 1 (50% probability of disease progression; dotted line) was 2.3 years. b) The MRI-BPAD-derived decision boundary in sample 2 was 2.7 years. Shaded area represents standard error. c) and d) 50% probability of cognitive decline requires a lower MRI-BPAD in APOE-e4 carriers compared to carriers. CD = cognitive decline; DX = diagnosis.

**Discussion**

The existing literature on brain age mainly defined brain age as chronological age predicted from MRI scans. Here, we have shown that FDG-PET predicts brain age as well as MRI, and FDG-BPAD can serve as a marker of cognitive decline where MRI-BPAD falls short. Precisely, we demonstrated that FDG-BPAD is superior in representing AD neuropathology and risk of cognitive decline in CN. MRI-BPAD, on the other hand, was more closely associated with a decline of executive function within the range of a CN diagnosis, as well as memory function, AD neuropathology and cognitive decline in MCI.

Our findings are in favor of a time line in which FDG-PET brain age better captures early disease-related neuropathology and risk, while the later onset of tau-related neurodegeneration and of objective cognitive decline are more strongly associated with signals of increasing brain age on MRI. This is in line previous work delineating that FDG-PET shows greater and more consistent changes early in the AD continuum, whereas MRI is superior in delineating AD-related changes with an AD diagnosis10. The current results expand on this finding by showing that MRI superiority is already present at the MCI stage. In clinical practice, the general brain health of CN could be assessed with FDG-PET-derived measures of brain age. This may be especially relevant for CN experiencing subjective cognitive impairment (SCI). Persons with SCI recognize cognitive deficits before they become objectively measurable. These individuals were shown to be more likely to develop MCI or AD compared to CN without SCI13. Differences on MRI brain age between CN and SCI have previously been shown, as SCI demonstrated a brain age advanced by 1.1 years3. The same model predicted the brain age of MCI to be advanced by 1.6 years, which is very close to our current results on MRI. To establish an optimal modality for brain age assessment in SCI will be an interesting topic for future research. For MCI patients, MRI scans are already routinely done, thus the assessment of brain health in these individuals and especially their risk of cognitive decline could be assessed without further strain.

Early detection of pathological abnormality is among the most crucial concepts in preventing AD. According to the *amyloid cascade hypothesis*, amyloid deposition is the causative agent of AD, causing a downstream effect of tau deposition, neurodegeneration and dementia14. Several promising anti-amyloid therapies are currently under assessment or have recently been approved for the treatment of MCI and early AD. The inclusion of BPAD into clinical trials of AD could have several advantages. First, we have shown that BPAD serves as a biomarker of cognitive decline in CN and especially MCI. Since cognitive decline is often an outcome factor of these trials, the notion of BPAD could help to identify those individuals who are most at risk of cognitive decline, thereby potentially strongly reducing the number of participants and thus cost associated with drug discovery. The usability of BPAD for cognitive decline prediction in CN will have to await confirmation by future studies including a larger sample of decliners. Secondly, BPAD is an established summary marker of brain health15. Brain age prediction algorithms are trained on a cognitively normal cohort and we have shown that BPAD reliably detects current and pending deviations from normal cognitive performance. Thus, it also appears possible to consider BPAD itself as an outcome variable of neuroscientific clinical trials, potentially reflecting drug action on the whole brain above and beyond variables of interest. Importantly, given that it is hardly possible to restore brain structures lost to neurodegeneration, MRI brain age will mostly not be expected to decrease, but only to decelerate. On the other hand, it appears possible that a decreased metabolism can be strengthened and increased again by certain interventions. This would have to be considered in choosing the most appropriate modality for BPAD as an outcome measure.

PET-BPAD or amyloid status (CN), and MRI-BPAD and APOE-e4 carriership (MCI) were identified as predictors of cognitive decline. For MCI to AD conversion, our out-of-sample prediction results compare well to a previous in-sample prediction of cognitive decline from MRI-brain age, which achieved an AUC of .83 for prediction of conversion to AD in 12 months, and an AUC for conversion in 36 months9. The combined observation of biomarkers of neurodegeneration and APOE-e4 carriership has previously been found beneficial for AD risk assessment (CITE GATEKEEPING PAPER). Other factors, such as hormones16 or lifestyle17 have a documented influence on brain age, a profound understanding the additional consideration of lifestyle factors may significantly improve prediction of cognitive decline

* BPAD = 0 enough for APOE positive in MCI 🡪 no severe neurodegeneration in aging relevant areas 🡪 does APOE speed up pending neurodegeneration?

APOE “doomed” for cognitive decline? Non-APOE != AD?

Zusammenhang amyloid cascade & APOE, APOE & metabolismus (🡪

NEURODEGENERATION = FAST AGING?;

* Feature-importance shows a8ging and disease related structures, only cognitively normal individuals included 🡪 neurodegeneration = faster aging?
  + Amyloid negatives + positives in sample

Brain age predicts conversion in MCI:

, ε4 carriers showed increased acceleration of individual brain aging as compared to non-carriers in pMCI and AD patients. 18 This is in line with recent studies suggesting that APOE ε4 carriers are suffering from faster pathologic processes than non-carriers, higher accuracy in APOE-e4 carriers than non-carriers

Brain age predicts conversion above and beyond established biomarkers (which might be more easily accessible) 9

* COMBINATION OF FACTORS:
* Limitations:
  + NO AMYLOID STATUS FOR CN\_VALIDATION 🡪 difference OASIS: different amyloid status, age distribution more closely matched
  + Small sample of converters with baseline CN diagnosis resulted in large standard error 🡪 pending confirmation by larger sample sizes, possibly outside of ADNI
  + FDG-PET quite expensive, will not be done in CN population without reason
  + Amyloid not predictive of conversion to AD in MCI sample 2 🡪 likely due to high collinearity of amyloid status and APOE
  + “In cases where the number of features for each data point exceeds the number of training data samples, the SVM will underperform.” 🡪 5-fold cv on 367 samples with 216 features: not enough samples 🡪 next approaches could investigate influence of feature reduction techniques on brain age prediction
  + Nfl easier accessed, but just one-dimensional 🡪 Likely, one-dimensional feature spaces such as a single plasma or cerebrospinal fluid biomarker will be insufficient
  + Amyloid positive individuals in training set 🡪 “cognitively normal” rather than “neuroscientifically healthy”, but latter definition hard to achieve due to multitude of variables to control for: decided to take representative sample of old individuals without cognitive impairment as reference group

analyses in the MCI sample revealed that both PET- and MRI-BPAD indicate conversion to AD. Notably, an MRI-BPAD greater than 2.3 years was identified as a specific and sensitive marker of CD, which outperformed other risk factors of CD in MCI in sample one in terms of significance and specificity. However, sensitivity and specificity scores are not high enough for MRI-BPAD to serve as a stand-alone biomarker of CD. Potentially, MRI-BPAD can serve as a reliable biomarker of CD together with APOE-e4 carriership, as Figure XX suggests different BPAD thresholds may apply as a function of the latter. XX

**Method**

**Participants**

Baseline T1-weighted MRI and FDG-PET scans of 367 CN and 513 individuals with MCI used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database ([adni.loni.usc.edu](https://ida.loni.usc.edu/collaboration/access/adni.loni.usc.edu)). The primary goal of ADNI has been to test whether biological markers and clinical and neuropsychological assessment can be combined to measure the progression of MCI and AD. To be included, time passed between the FDG-PET and MRI scan of the same individual could not exceed one year, and at the earlier scan, all individuals had to be at least 65 years of age. The age restriction was due to the fact that age at onset of AD is 65 years, and CN data below this age is rare in ADNI, thus potentially creating algorithms not suited for age prediction below 65 years.

To test our algorithms in an external dataset, we additionally considered 59 CN elderly  participants from the Open Access of Imaging Studies-3 (OASIS-3) database (https://www.oasis-brains.org/) 19. Given the small sample size of participants who received both an sMRI and 18F-FDG-PET within 12 months, we eliminated this time constraint for the OASIS test set, while all individuals were still necessitated to be 65 years at acquisition of the earlier scan.

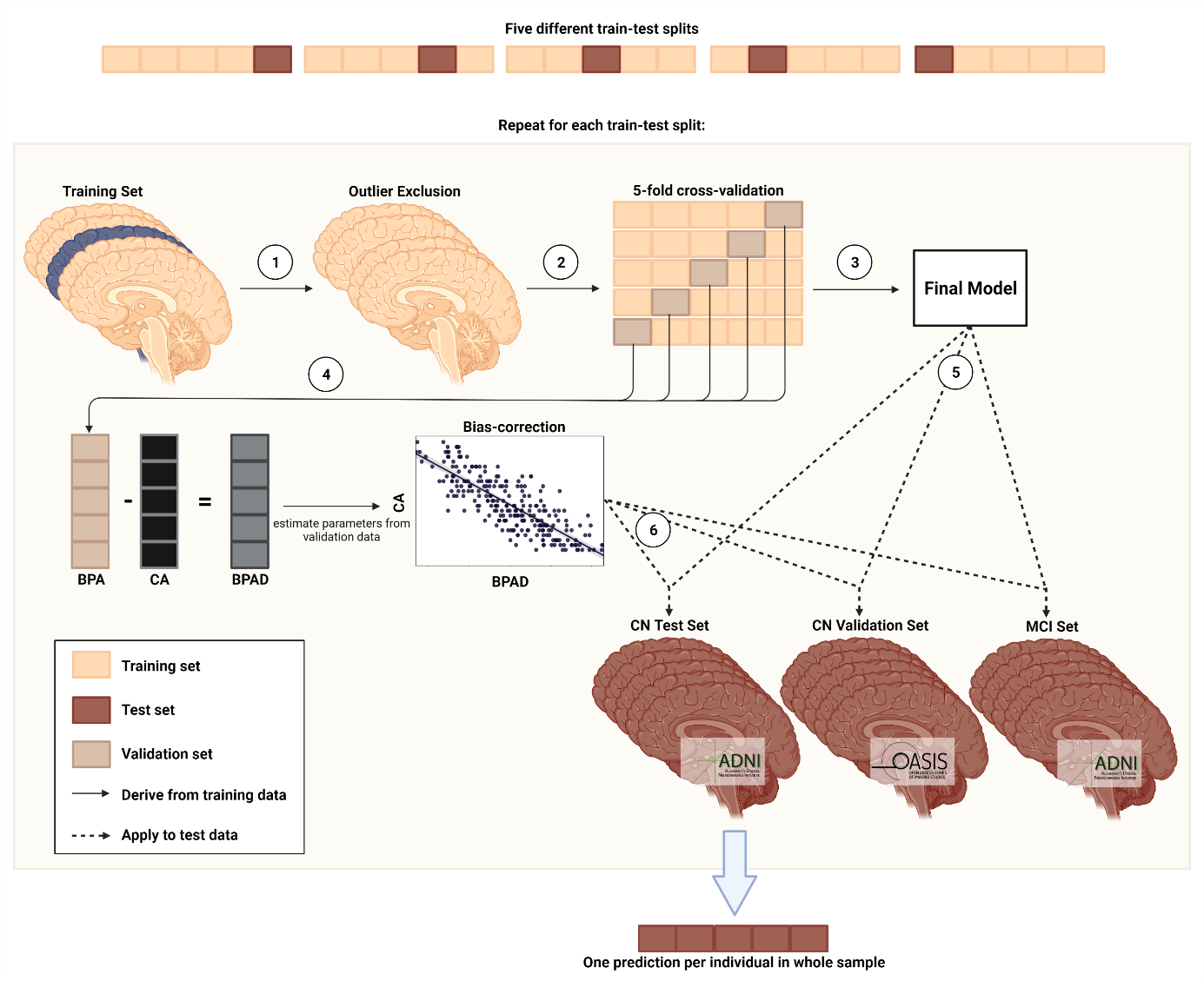
**Acquisition & Preprocessing of Biomarkers of Neurodegeneration**

18F-FDG-PET scans in both samples were acquired dynamically 30-60 minutes (6x5min frames) after injection with an average dose of 185 MBq (5mCi) and downloaded with minimal pre-processing (“Co-registered, averaged”-format). Pre-processing was performed using the Statistical Parametric Mapping 12 toolbox (SPM12; [www.fil.ion.ucl.ac.uk](http://www.fil.ion.ucl.ac.uk)): All 18F-FDG-PET scans were aligned to the anterior commissure/posterior commissure, and subsequently co-registered and normalized to a template in standard space. Lastly, standardized uptake value ratios (SUVr) were calculated (reference: pons20).

T1-weighted MRI scans were acquired on XX-T scanners according to the ADNI MRI acquisition protocol21. First, scans were pre-processed using denoising (spatial-adaptive Non-Local Means), spatial registration, bias-correction and skull-striping. Then the images are segmented by an adaptive maximum a posteriori approach (Rajapakse et al. 1997) with partial volume model (Tohka et al. 2004). For non-linear transformation, the Geodesic Shooting Algorithm (Ashburner & Friston 2011) was used based on SPM12 and compiled for containerization in Singularity (2.6.1).

**Calculation of brain-predicted age**

Mean gray matter volume (GMV) and SUVr were extracted for T1w MRI and 18F-FDG-PET, respectively, using a composite atlas containing 200 cortical22 and 16 sub-cortical regions23. Calculation of BPA was achieved using the scikit-learn library24 in Python 3.8.5. We applied a nested cross-validation approach, where ADNI data of CN individuals was split in a stratified manner into five different train (70%) and a test sets (30%). Through stratification, the original proportions of young-old (65 - 74 years), middle-old (75 - 84 years) and oldest-old classes (85 years+)25 in the whole ADNI dataset were maintained in the train and test sets. Each train set was then used to derive a final model for brain age prediction using a pipeline consisting of 1) outlier exclusion, 2) cross-validated prediction, and 3) bias correction (see Fig 5).

1. Outlier exclusion was performed to ensure data quality in an automated manner. Interquartile ranges (IQR) were inferred from the test sets, and subsequently applied to cognitively normal test sets, where subjects outside 6xIQR were removed from the analyses (range outliersADNI = [3, 7], range outliersOASIS = [5, 7]). Importantly, as previous works have shown, MCI subjects show an advanced brain age, which translates to a reduced signal in age-relevant brain regions9. Thus, outlier exclusion was not applied to the MCI sample.
2. To estimate BPA using FDG-PET or MRI, we compared relevance vector regression (RVR) and support vector regression (SVR). These machine learning models are prominently used for brain age prediction and are especially suited for training on small datasets26. Optimal (hyper)parameters were determined using five-fold stratified cross-validation in scikit-learn (for 

**Fig 5. Nested cross-validation approach for brain age prediction.** Five different train-test splits were used to train and test the models. (1) Outlier exclusion ranges were inferred from the training data, and applied to both the training and test data. (2) Models were trained using five-fold cross-validation. (3) The model with the smallest MAE on the validation folds was chosen as the final model. (4) PBA and CA from the validation folds was used to derive bias correction parameters. (5) The final model was subsequently applied to the test sets. (6) Bias correction parameters were applied to predictions in the test set.

1. a list of hyperparameters, see Supplementary Materials Table 1). During each iteration of cross-validation, four parts of the training data were first scaled (by removing the median and scalithe data according to the IQR, “robust scaler” from scikit-learn library) and then used to fit the models. The respective scaling parameters were subsequently applied to the validation set (fifth part of training data). Fitted models were used to predict CA from either neuroimaging modality in the validation set and these predictions were stored for bias-correction. As a result of cross-validation, one optimal RVR and one optimal SVR was yielded, where “optimal” refers to the respective (hyper-)parameter configuration that allowed for the smallest average MAE between

CA and BPA across the validation sets, and the final model was the one with the smallest average MAE across the remaining two optimal models.

1. BPA is subject to a frequently reported bias, in which BPA of older individuals is under- and BPA of younger individuals is overestimated27, regardless of the data or method under consideration28. Several approaches have been suggested for the correction of this bias, which can be broadly summarized into *methods including CA in the correction* and *methods not including CA in the correction*29. Bias correction was inferred from validation folds, and estimated parameters were subsequently applied to the test folds. To obtain an in-depth understanding of the effect of the different methods of bias correction on the prediction of CA from MRI and FDG-PET in our data, we implemented both following previous approaches27,30 and compared them with regards to MAE, R², and the percentage of test sets where bias remained eliminated. Bias correction with CA yielded the best average MAE and R² (see Supplementary Materials for a description of bias-correction without CA and Supplementary Table XX for results of the methodological comparison) and was thus for the calculation of BA. Thus, a linear regression model was fit on BPAD versus CA. Bias-free brain age was then calculated using slope (ɑ) and an intercept (β):

As a result from the above described nested cross-validation approach, we obtained five final models per modality, yielding one prediction per (non-outlier) subject in the CN sample (n = 345), and five predictions per (non-outlier) subject in the CN\_validation (rangen = [52, 54]) and MCI sample (n = 513). For each individual, BPAD was calculated as BPA – CA. MAE and average BPAD of all samples was compared between modalities using a dependent t-test.

**Associations of BPAD with cognitive performance, AD neuropathology, and cognitive decline**

To assess the associations of BPAD with cognitive performance, BPAD was correlated with two composite scores, ADNI-MEM31 and ADNI-EF32. Age and sex were entered as covariates in all correlations and significance levels were Bonferroni-corrected. ADNI-MEM combines several scores used to evaluate individuals’ memory performance from the Rey Auditory Verbal Learning Test, Alzheimer’s Disease Assessment Scale and Mini Mental State Exam. ADNI-EF is a summary score of several executive function tasks, including: Category Fluency, Trails, Digit span backwards, Wechsler Adult Intelligence Scale-R Digit Symbol Substitution, Number Cancellation, and Clock Drawing items

To assess the associations of BPAD with AD neuropathology, BPAD was correlated with amyloid deposition in the brain (AV45-PET), as well as with amyloid, tau and phosphor-tau accumulation in cerebrospinal fluid (CSF). Age and sex were entered as covariates in all correlations and significance levels were Bonferroni-corrected. For AV45-PET, mean standardized uptake value ratios (SUVr) are publicly available from previous analyses33–36. Briefly, the scans were co-registered to corresponding MRI images in native space and SUVrs were calculated voxel-wise using the whole cerebellum as a reference region. Global SUVr was then calculated using a brain mask including frontal, anterior/posterior cingulate, lateral parietal and lateral temporal regions, with amyloid positivity being defined with a cut-off of 1.1133. CSF measures of amyloid, tau and phospho-tau were acquired via lumbar puncture and analyzed using the Roche Elecsys® beta-Amyloid(1-42), Total Tau and Phospho Tau (181p) immunoassays37. The CSF cut-off for amyloid positivity was 1100 pg/ml38. The measuring range of the beta-amyloid assay is 200 – 1700 pg/ml.

Finally, BPAD, amyloid status, APOE-ε4 carriership and years of education were used to predict cognitive decline in CN and MCI using ten-fold cross-validated logistic regression. Cognitive decline was defined as a diagnosis of (more severe) cognitive impairment within two years (inclusive) after BPAD assessment. Thus, CN who received a diagnosis of MCI or AD within two years were cognitive “decliners”, while CN who maintained the CN diagnosis until 24 months after BPAD assessment yielded the group of “stables”. For MCI, decliners were those individuals who progressed to Dementia within two years. We created subsets of data matched for age and sex, where the number of decliners and stables was balanced.

1. Nie, C. *et al.* Distinct biological ages of organs and systems identified from a multi-omics study. *Cell Rep.* **38**, (2022).

2. Beheshti, I., Mishra, S., Sone, D., Khanna, P. & Matsuda, H. T1-weighted MRI-driven brain age estimation in Alzheimer’s disease and Parkinson’s disease. *Aging Dis.* **11**, (2020).

3. Rokicki, J. *et al.* Multimodal imaging improves brain age prediction and reveals distinct abnormalities in patients with psychiatric and neurological disorders. *Hum. Brain Mapp.* **42**, (2021).

4. Lee, J. *et al.* Deep learning-based brain age prediction in normal aging and dementia. *Nat. Aging* (2022) doi:10.1038/s43587-022-00219-7.

5. Jack, C. R. & Holtzman, D. M. Biomarker modeling of alzheimer’s disease. *Neuron* vol. 80 (2013).

6. Salloway, S. *et al.* Amyloid-Related Imaging Abnormalities in 2 Phase 3 Studies Evaluating Aducanumab in Patients with Early Alzheimer Disease. *JAMA Neurol.* **79**, (2022).

7. Haeberlein, S. B. *et al.* *EMERGE and ENGAGE topline results: two phase 3 studies to evaluate aducanumab in patients with early Alzheimer’s disease.* (2019).

8. Löwe, L. C., 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* **11**, (2016).

9. 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* **8**, (2013).

10. Dukart, J. *et al.* Generative FDG-PET and MRI Model of Aging and Disease Progression in Alzheimer’s Disease. *PLoS Comput. Biol.* **9**, e1002987 (2013).

11. Beheshti, I. *et al.* Predicting brain age using machine learning algorithms: A comprehensive evaluation. *IEEE J. Biomed. Heal. Informatics* (2021) doi:10.1109/JBHI.2021.3083187.

12. Ranganathan, P., Pramesh, C. & Aggarwal, R. Common pitfalls in statistical analysis: Logistic regression. *Perspect. Clin. Res.* **8**, (2017).

13. Parfenov, V. A., Zakharov, V. V., Kabaeva, A. R. & Vakhnina, N. V. Subjective cognitive decline as a predictor of future cognitive decline a systematic review. *Dement. e Neuropsychol.* **14**, (2020).

14. Hardy, J. A. & Higgins, G. A. Alzheimer’s disease: The amyloid cascade hypothesis. *Science* (1992) doi:10.1126/science.1566067.

15. Cole, J. H., Marioni, R. E., Harris, S. E. & Deary, I. J. Brain age and other bodily ‘ages’: implications for neuropsychiatry. *Molecular Psychiatry* vol. 24 (2019).

16. Luders, E. *et al.* Potential Brain Age Reversal after Pregnancy: Younger Brains at 4–6 Weeks Postpartum. *Neuroscience* **386**, (2018).

17. Cole, J. H. Multimodality neuroimaging brain-age in UK biobank: relationship to biomedical, lifestyle, and cognitive factors. *Neurobiol. Aging* **92**, (2020).

18. Landau, S. M. *et al.* Amyloid PET imaging in Alzheimer’s disease: A comparison of three radiotracers. *Eur. J. Nucl. Med. Mol. Imaging* (2014) doi:10.1007/s00259-014-2753-3.

19. LaMontagne, P. J. *et al.* OASIS-3: Longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and Alzheimer disease. *medRxiv* (2019) doi:10.1101/2019.12.13.19014902.

20. Verger, A., Doyen, M., Campion, J. Y. & Guedj, E. The pons as reference region for intensity normalization in semi-quantitative analysis of brain 18FDG PET: application to metabolic changes related to ageing in conventional and digital control databases. *EJNMMI Res.* **11**, 1–7 (2021).

21. Jack, C. R. *et al.* The Alzheimer’s Disease Neuroimaging Initiative (ADNI): MRI methods. *Journal of Magnetic Resonance Imaging* vol. 27 (2008).

22. Schaefer, A. *et al.* Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cereb. Cortex* **28**, (2018).

23. Tian, Y., Margulies, D. S., Breakspear, M. & Zalesky, A. Hierarchical organization of the human subcortex unveiled with functional connectivity gradients. *bioRxiv* (2020) doi:10.1101/2020.01.13.903542.

24. Pedregosa, F. *et al.* *Scikit-learn: Machine Learning in Python Gaël Varoquaux Bertrand Thirion Vincent Dubourg Alexandre Passos PEDREGOSA, VAROQUAUX, GRAMFORT ET AL. Matthieu Perrot*. *Journal of Machine Learning Research* vol. 12 http://scikit-learn.sourceforge.net. (2011).

25. Suzman, R. & Riley, M. W. Introducing the ‘oldest old’. *Milbank Mem. Fund Q. Health Soc.* **63**, (1985).

26. Baecker, L. *et al.* Brain age prediction: A comparison between machine learning models using region- and voxel-based morphometric data. *Hum. Brain Mapp.* **42**, (2021).

27. Beheshti, I., Nugent, S., Potvin, O. & Duchesne, S. Bias-adjustment in neuroimaging-based brain age frameworks: A robust scheme. *NeuroImage Clin.* **24**, (2019).

28. Liang, H., Zhang, F. & Niu, X. Investigating systematic bias in brain age estimation with application to post-traumatic stress disorders. *Hum. Brain Mapp.* **40**, (2019).

29. de Lange, A. M. G. & Cole, J. H. Commentary: Correction procedures in brain-age prediction. *NeuroImage: Clinical* vol. 26 (2020).

30. Cole, J. H. *et al.* Brain age predicts mortality. *Mol. Psychiatry* **23**, (2018).

31. Crane, P. K. *et al.* Development and assessment of a composite score for memory in the Alzheimer’s Disease Neuroimaging Initiative (ADNI). *Brain Imaging Behav.* **6**, (2012).

32. Gibbons, L. E. *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.* **6**, (2012).

33. Landau, S., Koeppe, R. & Jagust, W. *Florbetaben processing and positivity threshold derivation Motivation for changing the threshold*. (2011).

34. Landau, S. & Jagust, W. *Florbetapir processing methods*. (2011).

35. Jagust, W. J. *et al.* Relationships between biomarkers in aging and dementia. *Neurology* **73**, 1193–1199 (2009).

36. Jagust, W. J. *et al.* The Alzheimer’s Disease Neuroimaging Initiative positron emission tomography core. *Alzheimer’s Dement.* **6**, (2010).

37. Blennow, K. *et al.* Predicting clinical decline and conversion to Alzheimer’s disease or dementia using novel Elecsys Aβ(1–42), pTau and tTau CSF immunoassays. *Sci. Rep.* **9**, (2019).

38. Hansson, O. *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.* **14**, (2018).

NOTES

Cole Multimodality 2020:  “Multimodality neuroimaging can improve brain-age prediction, and derived brain-PAD values are sensitive to biomedical and lifestyle factors that negatively impact brain and cognitive health.” (modeled with structural and functional MRI) 🡪 whether some single modalities are more useful for the usage of BPAD in certain tasks is unclear

Rokicki: Age prediction based on structural MRI data shows high accuracy in common brain disorders. However, brain aging is complex and heterogenous, both in terms of individual differences and the underlying biological processes.

Rokicki: MCI the difference was 1.6 years (although smaller sample size, similar finding using MRI)

“Brain-age prediction uses machine learning to estimate an individuals apparent brain aging based on structural and functional brain characteristics derived from neuroimaging, commonly magnetic resonance imaging (MRI). Subtracting chronological age from estimated brain age provides a measure of the difference between an individuals predicted and chronological age; the *brain age delta*. For instance, if a 60 year old individual exhibits a brain age delta of -5 years, their typical aging pattern resembles the brain structure of a 55 year old, i.e., their estimated brain age is younger than what is expected for their chronological age. Individual variation in delta estimations have been associated with a range of biological and cognitive variables. brain-age estimation also involves a **frequently observed bias:** brain age is overestimated in younger subjects and underestimated in older subjects, while brain age for participants with an age closer to the mean age (of the training dataset) are predicted more accurately”

Brain-predicted age (BPA)

Chronological age (CA)

Brain-predicted age difference (BPAD)

mean absolute error (MAE)

<https://pubmed.ncbi.nlm.nih.gov/11526211/>

[https://n.neurology.org/content/65/8/1227 FOR CONVERSION JACK 2005](https://n.neurology.org/content/65/8/1227%20FOR%20CONVERSION%20JACK%202005)

DUKART 2013: The model suggests greater and more consistent changes in FDG-PET compared to sMRI at earlier and the inversion of this pattern at more advanced AD stages.

* **First paragraph: aging of the brain**
  + late-life adult brain shrinks with increasing age
  + changes at all levels from metabolism to morphology
  + incidence of stroke, white matter lesions, and dementia also rise with age
  + thus, abnormal brain age could be used as a biomarker for proneness to neurodegenerative diseases, such as Alzheimer’s disease
  + however, state-of-the-art machine learning models of normal brain aging are mostly based on structural MRI, thereby restricting brain age prediction to changes in morphology
  + young-old, middle-old, oldest-old → different risk factors associated with being of high age [Suzman 1985], but inherent resilience against age-related detrimental factors may be in place which allows reaching such high age, e.g. “Overall, there is evidence that pathological substrates of cognitive deterioration in the oldest-old are different from those observed in the younger-old. Microvascular parameters such as mean capillary diameters may be key factors to consider for the prediction of cognitive decline in the oldest-old. Neuropathological particularities of the oldest-old may be related to “longevity-enabling” genes” [von Gunten 2010]
* **Second paragraph: MRI - FDG-PET comparison**
  + FDG-PET unravels the molecular changes in cell metabolism of the brain
  + structural MRI depicts anatomical changes, such as atrophy
  + FDG-PET, as compared to MRI captures first AD-related changes earlier and more accurately [1,2,3]
  + FDG-PET displays greater and more consistent changes as compared to structural MRI at early stages of AD [3]
  + therefore, FDG-PET possibly yields previously unexplored information about brain age
* **Third paragraph: What is a good brain age model**
  + brain age = neuroimaging-predicted chronological age
  + brain-predicted age difference (BPAD) = brain age - chronological age
  + ...
* **Fourth paragraph: Aim of the study**
  + 1) to find suitable bias-correction for FDG-PET and T1-weighted MRI
  + 2) to **compare the predictive value** of FDG-PET and MRI for brain age in CN
  + 3) to assess **BPAD in individuals with MCI**
  + 4) to associate **BPAD with neuropsychology and neuropathology**