

Lobar Brain Age Profiles for the Early Identification of Dementias in Aging Populations

Funded by the Brandau-Laibach-Stiftung

Project lead: Dr. Elena Doering

Co-PIs: Univ.-Prof. Dr. Thilo van Eimeren, Univ.-Prof. Dr. Alexander Drzezga

Affiliation: Clinic and Policlinic for Nuclear Medicine, University Hospital Cologne; Department of Neurology, University Hospital Cologne; Institute of Neuroscience and Medicine (INM-2), Research Center Jülich

Kurzzusammenfassung auf Deutsch

- **Datenbasis aufgebaut (WP1 abgeschlossen).** Aus 22 geprüften internationalen Datenquellen wurden sechs ausgewählt und zu einer Datenbank mit 6.465 Personen zusammengeführt. Sie enthält MRT-Aufnahmen des Gehirns von gesund Alternden sowie von Patientinnen und Patienten mit allen vier Demenzformen, die in diesem Projekt untersucht werden (Alzheimer-Demenz, frontotemporale Demenz, Lewy-Körper-Demenz, vaskuläre Demenz).
- **Bildverarbeitung standardisiert.** Eine einheitliche Aufbereitung der MRT-Bilder wurde entwickelt. Vier der sechs Datenquellen (1.902 Personen) sind bereits vollständig verarbeitet, die übrigen zwei werden aktuell finalisiert.
- **Modell zur Schätzung des Gehirnalters implementiert und trainiert.** Ein KI-Modell schätzt das Alter eines Gehirns aus dem MRT-Bild mit einer durchschnittlichen Abweichung von etwa drei Jahren, auch bei Daten aus Studien, die das Modell im Training nicht gesehen hat. Das Modell ist damit auf neue Daten übertragbar.
- **Unerwarteter Befund: Die Hirnalterung unterscheidet sich zwischen Frauen und Männern.** Herkömmliche Modelle schätzen Männergehirne systematisch älter als Frauengehirne. Unsere Analysen zeigen, dass getrennte Modelle für Frauen und Männer diese Verzerrung beheben, ohne an Genauigkeit zu verlieren. Diese Ergebnisse wurden auf zwei internationalen Fachkonferenzen vorgestellt; eine Veröffentlichung ist in Vorbereitung.
- **Erste Ergebnisse bei unterschiedlichen Demenztypen.** Gehirne von Demenzpatientinnen und -patienten erscheinen deutlich älter als die gesund Alternder, am stärksten bei frontotemporaler Demenz (rund 8 Jahre), gefolgt von Alzheimer-Demenz (rund 4 Jahre). Diese Ergebnisse sind vorläufig, da noch nicht alle Daten vollständig aufbereitet sind.

Mit den bisherigen Fortschritten sind die drei Voraussetzungen für die lobäre Auswertung geschaffen: eine große Datenbank mit allen vier Demenzformen, einheitlich aufbereitete MRT-Bilder und ein funktionierendes, geprüftes Modell für das Gesamtgehirn. Die regionalen Modelle bauen direkt auf der Pipeline dieses Gesamthirnmodells auf, und dessen Ergebnisse dienen zugleich als Vergleichsmaßstab für den Mehrwert der regionalen Auswertung. Entsprechend den Erkenntnissen zu Geschlechtsunterschieden werden auch die regionalen Modelle getrennt für Frauen und Männer trainiert.

1. Project overview

The aim of this Brandau-Laibach-funded project is to estimate global (BAG) and lobar brain age gaps (L-BAGs) from T1-weighted MRI scans and to evaluate their utility for the differential diagnosis of the most prevalent dementia subtypes: Alzheimer's disease (AD), frontotemporal dementia (FTD), dementia with Lewy bodies (DLB), and vascular dementia (VaD). The work is structured into three work packages: (WP1) curation of a multi-national imaging database, (WP2) implementation of (L-)BAG estimation models, and (WP3) analysis of subtype-specific BAG profiles.

2. WP1 – Database curation (completed)

WP1 was carried out as a research assistant project by a Master's student (Malin Gömann, Neuroscience program at University of Cologne). The dual aim of WP1 was to identify a new multi-cohort dataset comprising T1-weighted MRI scans and demographic and clinical data from patients diagnosed with AD, FTD, DLB, or VaD, as well as healthy agers, and to acquire these data in preparation for downstream brain-age modelling.

Database. A structured online search of existing imaging databases was conducted between January and March 2026, deliberately prioritizing cohorts containing VaD and DLB, as dedicated datasets for these two subtypes are the most difficult to source. The database search identified 22 candidate cohorts that met the minimum criterion of providing T1-weighted MRI for at least one dementia subgroup alongside a healthy-control group (Table 1). Access was requested to all where sufficient information was available, while two requests remained unanswered. Data were requested through the AD Workbench, the Image and Data Archive (IDA) hosted by the Laboratory of Neuro Imaging (LONI), and the Dementias Platform UK (DPUK). Six cohorts were ultimately included, namely OASIS-1, NACC, SCAN, ADNI, AIBL, and NIFD, selected for their accessibility and their potential to expand representation of the harder-to-source subtypes (VaD, DLB, and FTD). No further cohorts were added at this time, because the remaining candidate databases were deemed not to improve coverage of the non-AD dementia types. Additional reasons for database exclusion included that image data was not made available in required 3D format or collaboration requests were too time intensive. For all databases, data use agreements were summarized and stored alongside downloaded files.

Table 1. Requested and acquired datasets for database

Dataset	Dementia types	Platform	Acquired	Included
OASIS I	AD	Website	Yes	Yes
OASIS II	AD	Website	Yes	
OASIS III	AD	Website	No response	
OASIS IV	AD	Website	No response	
MIRIAD	AD	Website	03.02.2026	
NACC	AD/FTLD/VaD/DLB	Website	09.02.2026	Yes
SCAN	AD/FTLD/VaD/DLB	LONI	12.02.2026	Yes
A4/LEARN Study	AD	LONI	05.02.2026	
ADNI	AD	LONI	Yes	Yes
AIBL	AD	LONI	Yes	Yes

Table 1. Requested and acquired datasets for database

Dataset	Dementia types	Platform	Acquired	Included
GS1	CU subjects at risk of dementia	LONI	*	
HABS-HD	AD	LONI	09.02.2026	
MCSA	unspecified	LONI	10.02.2026	
POINTER	CU subjects at risk of dementia	LONI	09.02.2026	
WRAP	AD	LONI	09.02.2026	
NIFD	FTLD	LONI	19.02.2026	Yes
ADDIA	AD	AD Workbench	05.02.2026	
AMPLE	AD/DLB	DPUK	*	
Swedish BioFINDER study	AD/FTLD/VaD/DLB	DPUK	*	
MEMENTO	unspecified	DPUK	*	
NIMROD	AD/FTLD/VaD/DLB	DPUK	*	
QMIN MC	unspecified	DPUK	*	
European DLB Consortium	DLB	Insufficient information		
Mayo Clinic DLB	DLB	Insufficient information		

* Application efforts were discontinued; AD = Alzheimer's disease; CU = cognitively unimpaired; DLB = dementia with Lewy bodies; FTLD = frontotemporal lobar degeneration; VaD = vascular dementia

To be included into our multi-cohort dataset, participants were required to be at least 50 years of age, and both a clinical and MRI assessment needed to be available no more than six months apart. If multiple scans were available, the most recent scan was retained to obtain the most pronounced cases of neurodegeneration for the current analyses. Healthy agers were required to have a Clinical Dementia Rating (CDR) of 0 and have been otherwise found cognitively normal by study administrators of each database. A diagnosis of dementia provided by each database was accepted if patients had a CDR of at least 0.5. In datasets where VaD or DLB cases were grouped within broader vascular or Lewy body disease labels, an additional Mini-Mental State Examination (MMSE) threshold of 26 or lower was applied to reduce the likelihood of including individuals without clinically relevant cognitive impairment.

The resulting database comprises 6,465 participants, all with a T1-weighted MRI scan, and it uniquely combines all four targeted dementia subtypes with a large healthy-control reference within a single cross-sectional resource (Table 2). Expectedly, FTD and DLB subjects were younger compared to the other dementia types, as these dementias tend to occur earlier in life. Given the strong under-representation of women in the DLB cohort, we will revisit acquisition of DLB data in parallel to downstream analyses, wherein we will primarily focus on obtaining information and data from the Mayo Clinic DLB cohort and the European DLB consortium.

Table 2. Demographic and clinical overview of the curated database.						
	Healthy agers	Alzheimer's disease (AD)	Frontotemporal dementia (FTD)	Lewy body dementia (DLB)	Vascular dementia (VaD)	Total
N	3,596	2,363	212	160	134	6,465
N Preprocessing finished	1,418	341	114	12	17	1,902
% Female	64.9	51.1	39.2	18.8	54.5	57.6
Age (M $\pm$ SD)	75.5 $\pm$ 8.4	75.6 $\pm$ 8.0	66.3 $\pm$ 6.1	71.0 $\pm$ 6.5	76.1 $\pm$ 7.4	—
MMSE (M $\pm$ SD)	29.1 $\pm$ 1.3	22.3 $\pm$ 5.3	21.4 $\pm$ 7.1	22.8 $\pm$ 6.5	25.8 $\pm$ 3.4	—
CDR (median [IQR])	0 [0–0]	1 [1–2]	1 [0.5–1]	1 [0.5–1]	0.5 [0.5– 0.5]	—

All demographic and clinical variables (age, sex, MMSE, CDR, and years of education) were harmonized across cohorts. Education data were available for five of the six datasets. In one dataset (NACC), education was recorded categorically (12 years = high-school completion, 16 = bachelor's, 18 = master's, 20 or more = doctoral), therefore, education information in the remaining four datasets was adapted to the same scheme. MMSE scores were reported for five datasets. In the SCAN cohort, where MMSE was not reported, MMSE scores were estimated from CDR Sum-of-Boxes using a validated mapping function (Brogaard et al., 2025). Geographically, 95.6% of participants were scanned in North America (predominantly the United States) and the remainder in Australia.

Harmonization. As a second part of WP1, a literature review on harmonization strategies for multi-cohort neuroimaging data was performed. Broadly, two categories of image harmonization techniques have emerged, namely statistics- and deep learning-based ones, with no current gold standard. A statistics-based method called “ComBat” (Fortin et al., 2017) is currently the most widely used, which harmonizes scans across acquisition sites, while preserving variance due to covariates like age and diagnosis. However, classical statistical harmonization methods such as ComBat require prior knowledge of all scanner sites to estimate and remove site-related variance. This dependence can introduce cross-cohort data leakage and therefore limits the applicability of the pipeline to independent external datasets. In machine learning applications such as brain age estimation, strict independence between training and test data is essential to demonstrate that the model has learned generalizable patterns rather than an overfitted input-output mapping. Several deep learning strategies have recently been proposed that could be applied to independent data, as well as capture more complex batch effects compared to statistical methods. However, there is no most frequently used deep learning harmonization method thus far, the available methods have primarily been demonstrated on raw rather than preprocessed data, and there is currently no open-source provision of a fully trained model. Given the lack of a universally beneficial method, we decided to treat image harmonization as an optional, additive component rather than a standard part of the project, to be activated only if the project aims cannot be achieved without it.

3. WP2 – Preprocessing and (global) BAG estimation (in progress)

3.1 Preprocessing pipeline

A dedicated MRI preprocessing pipeline was established as part of this project that aimed to reduce signal bias and noise, while maintaining information content. The pipeline comprises linear co-registration to MNI standard space, N4 bias-field correction, SANLM denoising and brain masking. Of the six databases, MRI and clinical data of four (ADNI, AIBL, SCAN, NIFD) have been fully preprocessed and quality checked so far; the remaining two (NACC, OASIS-1) are currently being processed with the same pipeline. Exemplary images before and after pre-processing are shown in Fig. 1.

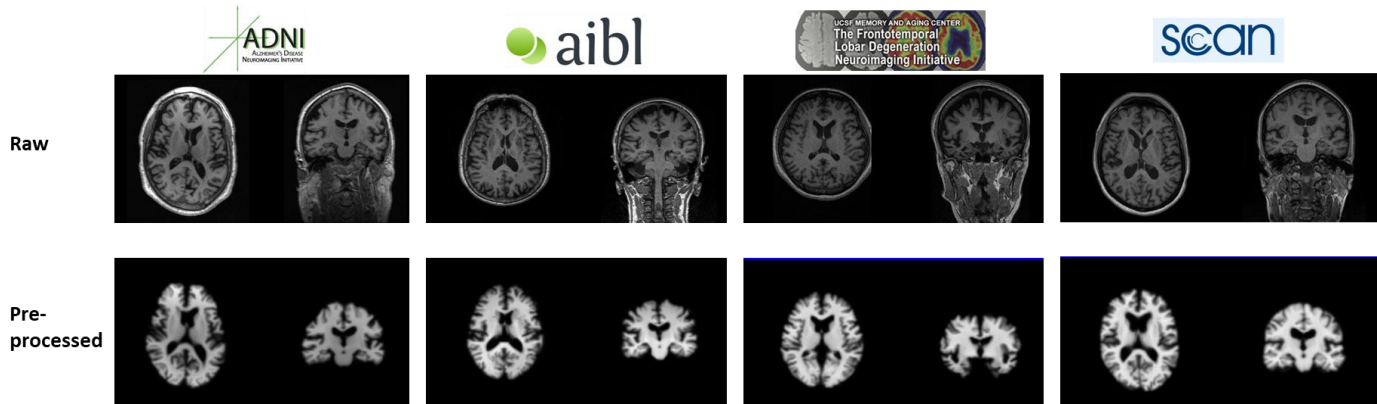


Figure 1 Raw and preprocessed data from four cohorts.

3.2 Global brain-age model

Next, a convolutional neural network (CNN) was implemented and trained to estimate age from preprocessed MRI scans of the ADNI healthy-ager cohort (N = 720, 50% women). ADNI was chosen as it represented the largest control sample. Training was repeated using five different train-test splits (i.e., performed using 5-fold cross-validation), with the test folds and the three remaining, already preprocessed cohorts (AIBL, SCAN, and NIFD) serving as test cohorts. This process yielded five trained brain age models, and final brain ages were obtained as the average across them. Brain age accuracy was assessed using the mean absolute error (MAE) and the coefficient of determination (r^2) between predicted and chronological age in healthy agers. BAGs were computed as brain age – chronological age, with higher BAGs representing older-appearing brains. Regression-to-the-mean bias, wherein younger subjects are typically overpredicted and vice versa, was corrected using a linear model fit on a portion of the training data (de Lange & Cole, 2020).

3.3 An unexpected, valuable detour: Quantification of sex differences in brain aging

An unexpected but important early outcome of this WP was that BAGs differed significantly between men and women when using conventional CNN training (Fig. 2B). This bias is reproduced, yet largely ignored, in the current literature, where consistently, women's brains tend to be estimated younger than men's (Doering et al., 2024; Lee et al., 2022). Left unaddressed, this bias potentially produces systematic bias in the interpretation of BAG, and thus needs careful consideration before brain age can be used as a marker in clinical routine. Hence, we allocated some time to investigating how sex needs to be accounted for in the context of BAG based on the ADNI cohort.

One possibility would be that the higher BAGs in men simply necessitate a sex-specific interpretation in clinical contexts, such as, e.g., sex-specific cut-off values for establishing an “elevated” BAG. This would mean that while women and men share common aging trajectories, men’s brains compared to women’s are truly phenotypically older than women’s. A further possibility is that brain aging itself differs between women and men, meaning that the phenotypical acceleration in men is a computational artifact and separate brain-age models need to be trained for women and men. To shed further light on this area, we trained the CNN using three different training strategies, namely sex-agnostic training (women and men pooled), sex-aware training (women and men pooled, but sex information is provided to the CNN), or sex-specific training (women only or men only used for training). We then compared 1) accuracy, 2) brain regions reflective of aging as per the different models, and 3) how BAG relates to stages and domains of cognitive impairment across models and sexes.

Insightful findings emerged. First, although sex-specific training makes use of only half the data compared to sex-agnostic or -aware models, age estimation was equally accurate compared to sex-agnostic or sex-aware training (Fig. 2A and Table 3). This could mean that reducing the training data to only one sex eliminates a major source of variance that sex-agnostic and -aware models otherwise have to disentangle. Interestingly, estimation error was consistently higher in women and seemed to slightly decrease in women through sex-specific modeling, whereas it remained roughly stable in men. Sex-specific modeling of BAGs also mitigated the sex difference in BAG, wherein men’s brains typically look older compared to women’s (Fig. 2B). Second, the regions deemed important for age estimation were different across women and men, providing further evidence that there are relevant differences in brain aging across sexes (Fig. 3). In line with the observation that estimation error changed more in women than in men through sex-specific modeling, visual assessment of regions important to the individual models suggested a male-dominated pattern even in pooled (women+men) training strategies. Consistently, applying the female model to MRI scans of men yielded significantly higher BAGs, and vice versa, especially in women (Table 3 “cross-sex” row). A possible interpretation of these findings in healthy agers is that women’s brain aging is more heterogeneous compared to men’s, and that pooled training strategies underrepresent this heterogeneity because the overall estimation error can be relatively minimized without accounting for heterogeneity in only one half of the cohort.

Regarding the third aim, BAGs increased with stage of cognitive impairment (ANCOVA, $F>32$ and $P<0.001$ across all modeling strategies, Fig. 2B). In amyloid-positive subjects, regression analysis revealed that higher BAGs were significantly associated with more tau pathology (linear model $BAG \sim \text{tau SUVR} + \text{education} + \text{APOE}$; women: $\beta=0.31$, $P=0.025$; men: $\beta=0.53$, $P<0.001$ in sex-specific model), suggesting that BAGs carry biological staging potential for AD across sexes. Higher sex-specific BAGs were also associated with worse cognitive performance across domains and sexes (Fig. 4).

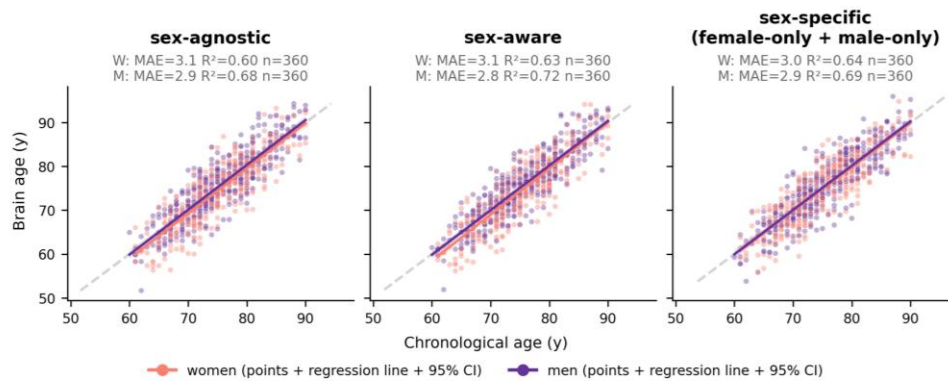
Together, these results provide compelling evidence that brain aging differs sufficiently between women and men to warrant both sex-specific training and interpretation of brain age. Based on these results, we will proceed with sex-specific modeling of BAGs. Although sex-specific modelling was not part of the original Brandau-Laibach proposal, it strengthens the robustness of the project's main analyses and aligns this project with the need to understand sex differences in neurodegeneration. Initial results of this subproject were presented at two international conferences (see section 6). A manuscript based on these findings is currently in preparation and aimed for submission in September 2026.

Table 3. Accuracy and average brain age gap (BAG) of different training strategies and diagnostic groups in ADNI cohort.

	Healthy agers					MCI			Dementia		
	MAE Women	MAE Men	Women	Men	W vs M	Women	Men	W vs M	Women	Men	W vs M
Sex-agnostic	3.15	2.95	-0.60	0.24	0.008	0.84	2.91	0.007	4.93	6.16	0.03
Sex-aware	3.06	2.76	-0.73	0.13	0.003	0.66	1.64	0.009	4.82	5.67	0.08
Sex-specific	3.00	2.91	-0.14	0.13	0.37	1.29	1.52	0.72	4.45	4.97	0.24
Cross-sex	NA	NA	-1.84	0.59	<0.001	NA	NA	NA	NA	NA	NA

Notes. MAE column indicates average absolute deviation of brain age from chronological age in years („mean absolute error“). „W vs M“ column indicates whether there was a significant difference between women and men. In the cross-sex row, the women-only model was applied to MRI scans of men and vice versa.

A Accuracy in sex-agnostic, sex-aware and sex-specific models



B Women's and men's brain age gaps (BAGs) across cognitive groups

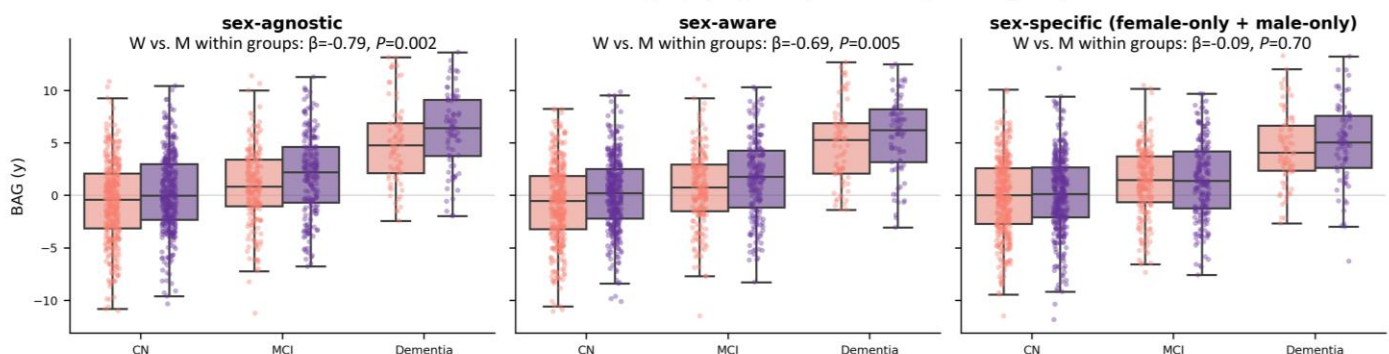


Figure 2 Accuracy and BAGs of three different training strategies in ADNI. A: Scatter and regression plots of Brain age by chronological age show high accuracy of the implemented models (sex-agnostic, sex-aware, sex-specific). All three panes show the same 360 healthy agers. Dashed line shows identity. B: Boxplots of women's and men's BAGs show a step-wise increase of BAG with increasing clinical impairment (CN = cognitively normal [healthy agers], MCI = mild cognitive impairment and dementia). Linear models ($BAG \sim Sex + DIAGNOSIS + PTEDUCAT + APOE4$, reference = Male) show that the significant ($P < 0.05$) over-prediction of men's BAGs within each diagnostic group is mitigated for sex-specific modeling.

Interim Progress Report

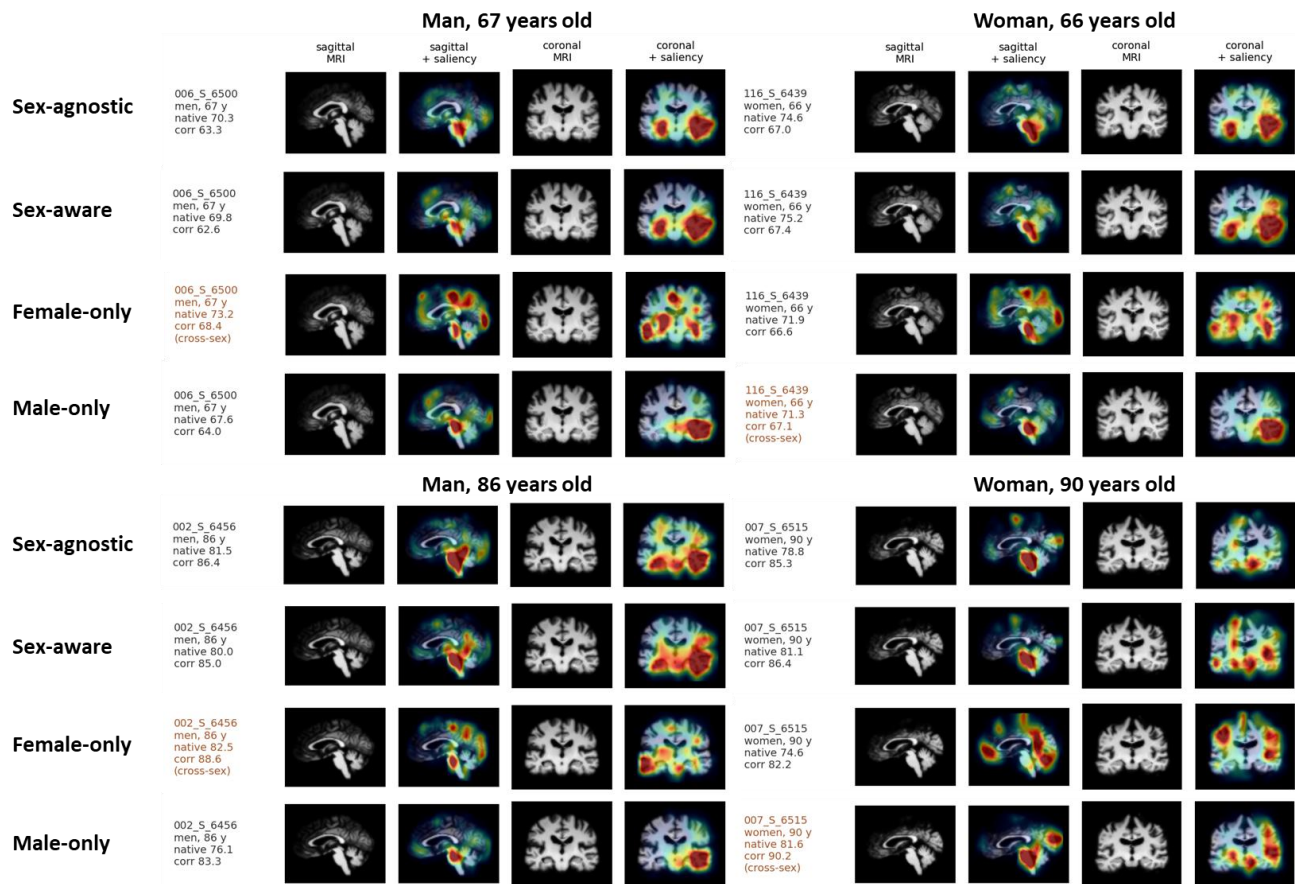


Figure 3 Exemplary saliency maps depicting brain regions relevant to age estimation by model type. For men, saliency was more right-hemispheric and focal, whereas for women, models tended to learn a more global assessment of aging patterns. Models trained using a pooled approach (sex-agnostic and sex-aware) appear to have learned a combined pattern that was visually dominated by patterns associated with aging in men.

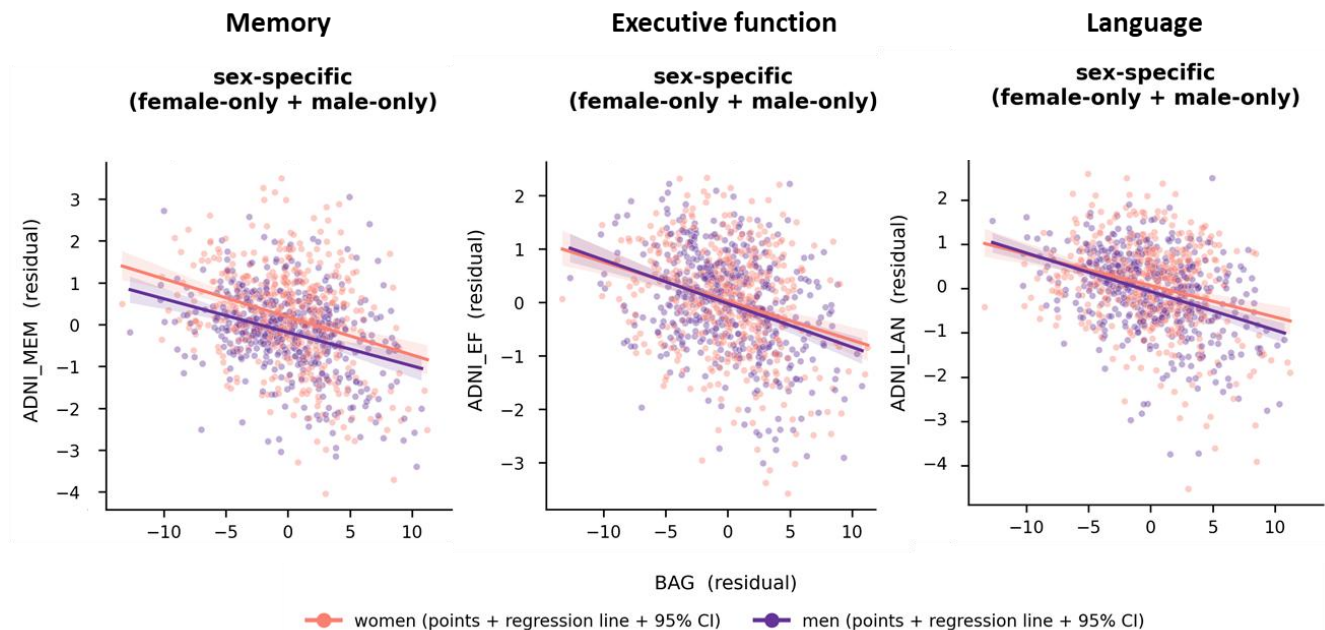


Figure 4 Association of sex-specific BAGs with performance in three cognitive domains.

3.4 Global brain age accuracy

Next, we aimed to understand how accurately the global brain age model estimated chronological age in the other, external cohorts. Towards this end, we fed the healthy ager data from the AIBL, SCAN and NIFD to the ADNI-trained models and averaged brain age estimates to obtain one brain age per subject.

Table 4 summarizes average accuracy in the already-preprocessed subset of the database as evaluated based on the healthy agers of each cohort. A total of 1902 subjects (1418 healthy agers and 484 dementia patients) were evaluated. Of note, the currently pre-processed data shows a strong under-representation of VaD and DLB subjects, with no MRI scans of women currently available for DLB. This gap is aimed to be closed after inclusion of NACC data (N>3500) and will be a major focus point.

All cohorts showed acceptable accuracy in healthy agers (MAE = 2.91-3.55 years) close to the pre-registered target of 2.5-3.0 years, while r^2 was slightly lower than anticipated (target: $r^2>0.8$). Women in the NIFD cohort, as well as both women and men in the AIBL cohort showed slight miscalibrations of BAG, shifted towards over-prediction. These minor deviations will be further investigated and optimized through additional CNN architecture considerations as the project continues, while the absence of major cohort effects is in line with our hypothesis that major harmonization efforts might not be needed.

Table 4. Healthy-aging brain-age accuracy of the sex-specific CNN trained in ADNI. ADNI scores are test set estimates not seen during training. AIBL, SCAN and NIFD are external datasets also not seen during training. MAE = mean absolute error; mean BAG near zero indicates good calibration. Total = 1,418 subjects.

Database	Role	Sex	n	Mean age (yrs)	MAE (yrs)	r^2	Mean BAG (yrs)
ADNI	Internal test (5-fold CV)	Women	360	74.9	3.00	0.64	-0.14
		Men	360	75.5	2.91	0.69	0.13
AIBL	External	Women	178	73.3	3.26	0.54	0.99
		Men	115	73.5	3.09	0.63	0.86
SCAN	External	Women	207	68.9	3.55	0.83	0.02
		Men	83	68.7	2.97	0.80	0.34
NIFD	External	Women	69	66.5	3.17	0.79	1.56
		Men	46	65.7	3.33	0.67	0.36

3.5 Global brain age gaps in different types of dementia

To understand how global BAGs are associated with dementia in general, as well as individual subtypes in particular, we next obtained dementia patients' brain ages similar to the procedure in healthy agers. Group-level global BAGs are shown in Fig. 5 and Table 5. Pooled across cohorts, brains of dementia patients appeared as significantly older compared to healthy agers (ANCOVA pooled: $F=538$, $P<0.001$; women: $F=291$, $P<0.001$; men: $F=253$, $P<0.001$). BAGs were also different across dementia subtypes (pooled: $F=12$, $P<0.001$; women: $F=16.5$, $P<0.001$; men: $F=4.03$, $P=0.008$). In line with prior literature (Persson et al., 2023), FTD showed the largest acceleration of global brain aging compared to healthy agers (+8.1 years, Tukey test: $P<0.001$), AD (+4.2 years, $P<0.001$) and DLB (+3.6 years, $P=0.03$) smaller positive gaps, and VaD (+1.5 years, $P=0.66$) no significant deviation from healthy agers. These results are preliminary, however, as the sample sizes of non-AD dementia patients are currently below the anticipated sample size required to obtain robust effects (N>261). Nevertheless, these initial results provide a promising and competitive baseline for lobar BAG estimates that will follow in WP3.

Table 5. *Global brain age gaps (BAGs) estimated by the sex-specific model trained in ADNI. Total = 484 subjects.*

Group	Sex	n	Mean age	Mean BAG (yrs)	Δ vs HC (yrs)	P
Alzheimer's disease (AD)	Women	184	75.7	3.8	3.5	<0.001
	Men	157	75.2	4.8	4.5	<0.001
Frontotemporal dementia (FTD)	Women	48	67.3	8.8	8.5	<0.001
	Men	66	64.1	7.6	7.2	<0.001
Lewy body dementia (DLB)	Women	0	NA	NA	NA	NA
	Men	12	68.8	3.6	3.3	0.03
Vascular dementia (VaD)	Women	12	77.0	1.7	1.4	0.52
	Men	5	77.2	1.0	0.7	0.99

Notes. Δ vs HC = mean BAG difference in years versus the healthy-control reference; P-values are FDR corrected (Tukey vs healthy agers). Results are intermediate, as they are based on the four databases preprocessed to date and will be updated once the remaining cohorts are added and the analysis is finalized.

4. WP3 – Lobar BAG analyses (upcoming)

Following completion of WP2, in WP3, we will (i) train lobe-specific BAG models for the frontal, temporal, parietal, and occipital lobes as well as the limbic system; (ii) compare (L-)BAGs between dementia subtypes and healthy controls using ANCOVAs similar to those in WP2; and (iii) identify subtype-specific (L-)BAG profiles via random-forest classification, as specified in the proposal. If time allows, we will further extend the investigations into the potential relevance of hemispheric (left versus right) brain aging, as well as aging in disease-specific meta-regions of interest (e.g., the hippocampus and entorhinal cortex for AD).

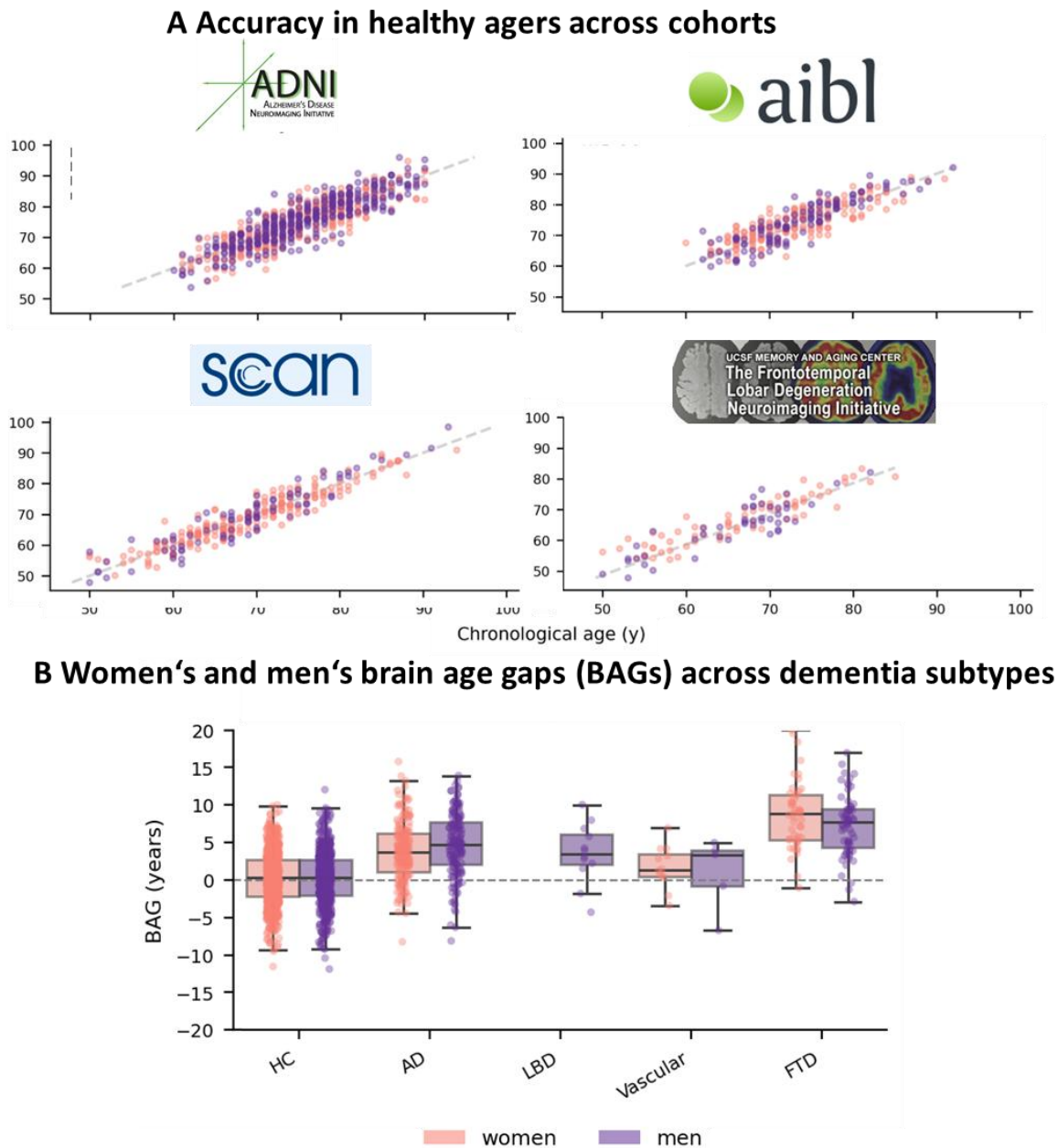


Figure 5 Sex-specific accuracy and global BAGs across healthy agers and dementia subtypes. A: Scatterplots showing high accuracy in estimating age from MRI scans across cohorts. Dashed line shows identity. B: Boxplots showing BAG estimates across healthy agers (HC) and currently available data from dementia subtypes. In this preliminary assessment, all dementia subtypes but vascular dementia show higher BAGs compared to healthy agers in both women and men.

5. Timeline and next steps

Period	Planned activities
Summer 2026	Complete preprocessing of the remaining databases (NACC, OASIS-1); re-attempt acquisition of Mayo Clinic DLB and European DLB cohorts, finalize global BAG estimation on the full sample; submission of manuscript on sex differences.
Autumn 2026	Beginning of master thesis by Malin Gömann based on WP 2 and 3; Implementation of lobar BAG models (initiation of WP3 phase).
Winter 2026/27	Subtype-specific (L-)BAG profile analyses (ANCOVA, random forest); robustness checks including sex-specific comparison.
End of funding period	Finalization of master thesis; submission of second peer-reviewed manuscript; public release of trained (L-)BAG models on GitHub.

6. Project outcomes until now

- Presentation of concept at Jülich Institutes for Neuroscience and Medicine (INM) retreat 2025 (Jülich, November 2025)
- Interim project report by Malin Gömann on data acquisition
- Journal club presentation by Malin Gömann on image harmonization strategies (Cologne, March 2026)
- Alzheimer's and Parkinson's Disease Conference (AD/PD) presentation on sex-related differences in brain age estimation by Elena Doering (Copenhagen, March 2026)
- Organization for Human Brain Mapping Conference (OHBM) poster on sex-related differences in brain age estimation by Elena Doering (Bordeaux, June 2026)

7. References

- Brogaard, N. J., Hahn-Pedersen, J. H., Gundgaard, J., Polavieja, P., Bray, B. D., Chan, M. S., Hoang, M., & Trepel, D. (2025). Extension and external validation of mapping between the Mini-Mental State Examination and the Clinical Dementia Rating scale in patients with Alzheimer's disease. *Alzheimer's and Dementia*, 21(6). <https://doi.org/10.1002/alz.70163>
- de Lange, A. M. G., & Cole, J. H. (2020). Commentary: Correction procedures in brain-age prediction. In *NeuroImage: Clinical* (Vol. 26, p. 102229). Elsevier Inc. <https://doi.org/10.1016/j.nicl.2020.102229>
- Doering, E., Antonopoulos, G., Hoenig, M., van Eimeren, T., Daamen, M., Boecker, H., Jessen, F., Düzel, E., Eickhoff, S., Patil, K., & Drzezga, A. (2024). MRI or 18 F-FDG PET for Brain Age Gap Estimation: Links to Cognition, Pathology, and Alzheimer Disease Progression. *Journal of Nuclear Medicine*, 65(1), 147–155. <https://doi.org/10.2967/jnumed.123.265931>
- Fortin, J. P., Parker, D., Tunç, B., Watanabe, T., Elliott, M. A., Ruparel, K., Roalf, D. R., Satterthwaite, T. D., Gur, R. C., Gur, R. E., Schultz, R. T., Verma, R., & Shinohara, R. T. (2017). Harmonization of multi-site diffusion tensor imaging data. *NeuroImage*, 161. <https://doi.org/10.1016/j.neuroimage.2017.08.047>

Interim Progress Report

Lee, J., Burkett, B. J., Min, H.-K., Senjem, M. L., Lundt, E. S., Botha, H., Graff-Radford, J., Barnard, L. R., Gunter, J. L., Schwarz, C. G., Kantarci, K., Knopman, D. S., Boeve, B. F., Lowe, V. J., Petersen, R. C., Jack, C. R., & Jones, D. T. (2022). Deep learning-based brain age prediction in normal aging and dementia. *Nature Aging*, 2, 412–424. <https://doi.org/10.1038/s43587-022-00219-7>

Persson, K., Leonardsen, E. H., Edwin, T. H., Knapskog, A. B., Tangen, G. G., Selbæk, G., Wolfers, T., Westlye, L. T., & Engedal, K. (2023). Diagnostic accuracy of brain age prediction in a memory clinic population and comparison with clinically available volumetric measures. *Scientific Reports*, 13(1), 14957. <https://doi.org/10.1038/s41598-023-42354-0>