Accelerated frontal grey matter atrophy in relapsing remitting multiple sclerosis

Max Korbmacher[1-3]*, Ingrid Anne Lie[4], Kristin Wesnes[5], Eric Westman[6,7], Thomas Espeseth[8,9], Ole Andreas Andreassen[10,11], Hanne Harbo[4,12], Gro Owren Nygaard[4,12], Lars Tjelta Westlye[8,10,11], Stig Wergeland[1,13,14], Kjell-Morten Myhr[1,13], Einar August Høgestøl[4,8,12,†], Øivind Torkildsen[1,13,†]
for the Alzheimer's Disease Neuroimaging Initiative[+]

[†]These authors contributed equally to this work.

[+]Part of the 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

Author affiliations:
[1]Neuro-SysMed, Department of Neurology, Haukeland University Hospital, Norway
[2]Mohn Medical Imaging and Visualisation centre, Department of Radiology, Haukeland University Hospital, Norway
[3]Department for Radiography, Western Norway University of Applied Sciences, Norway
[4]Department of Neurology, Oslo University Hospital, Norway
[5]Department of Neurology, St. Olav's Hospital, Trondheim University Hospital, Norway
[6]Division of Clinical Geriatrics, Department of Neurobiology, Care Sciences, and Society, Karolinska Institutet, Sweden
[7]Ageing Epidemiology Research Unit, School of Public Health, Imperial College London, UK
[8]Department of Psychology University of Oslo, Norway
[9]Department of Psychology, Oslo New University College, Norway
[10]Center for Precision Psychiatry, University of Oslo and Oslo University Hospital, Norway
[11]K.G. Jebsen-Centre for Neurodevelopmental disorders, University of Oslo, Norway
[12]Institute for Clinical Medicine, University of Oslo, Norway
[13]Department of Clinical Medicine, University of Bergen, Norway
[14]Norwegian MS-Registry and Biobank, Helse Bergen, Haukeland University Hospital, Norway

Correspondence to: Max Korbmacher
Jonas Lies vei 71, 5053 Bergen, Norway
max.korbmacher@helse-bergen.no

Abstract

Background. Grey matter (GM) atrophy has been suggested as the most accurate marker of neurodegeneration in multiple sclerosis (MS) progression. At the same time, few studies have examined regional atrophy longitudinally and how it correlates to different clinical domains.

Methods. We used two relapsing remitting MS (RRMS) cohorts (N=250, T_1w-scans=940) sampled for up to 12 years to map grey matter atrophy localisation, compare the examined patterns with six age-matched healthy control (HC) cohorts (N=31,427, T_1w-scans=32,793), and associate the observed atrophy pattern with disability, cognitive decline and fatigue, measured with the Expanded Disability Status Scale (EDSS), Paced Auditory Serial Addition Test (PASAT), and Fatigue Severity Scale (FSS).

Results. The strongest, replicable significant brain atrophy patterns in RRMS were found in the frontal lobes, specifically, in the superior frontal cortices ($\beta_{age} \leq -0.27$), pars orbitalis ($\beta_{age} \leq -0.25$), and thalami ($\beta_{age} \leq -0.20$). Across samples, such atrophy patterns were more pronounced than in a sample of more 20-year older HCs in the right superior frontal cortex ($Z>2.41$, $p<0.009$), right caudal middle frontal cortex ($Z>2.08$, $p<0.019$), right caudate ($Z>2.09$, $p<0.019$), and the left frontal pole ($Z>2.42$, $p<0.008$). The replicability of associations between volumetrics and clinical outcomes was limited to EDSS and the left superior frontal cortex ($\beta_{EDSS} \leq -0.09$) and the pars orbitalis ($\beta_{EDSS} \leq -0.09$).

Conclusion. Our findings indicate accelerated, early GM atrophy in people with RRMS, specific to the frontal lobes, which is more pronounced than in older HCs and uniquely related to disability-progression. These results suggest frontal lobe atrophy as an imaging biomarker in future MS-trials.

Key messages

What is already known on this topic – Grey matter atrophy is present in multiple sclerosis. The replicability of region-specific atrophy patterns has however not been systematically assessed, neither their relationship to different clinical outcomes, nor their comparison to healthy controls.

What this study adds – We present atrophy patterns with a dominant role of the frontal lobes, also presenting associations with disability progression. Additionally, we show that grey matter degenerates faster in people with multiple sclerosis than in more than 20 years older healthy controls.

How this study might affect research, practice or policy – Our findings highlight the presence of strong grey matter atrophy patterns in multiple sclerosis. Consequentially, more focus could be cast on cortical areas, and specifically the frontal lobes, when assessing neurodegeneration and disease progression in multiple sclerosis.

Introduction

Multiple sclerosis (MS) is a chronic immune-mediated disease of the central nervous system affecting more than 2.8 million people worldwide.[1] Despite significant technological and pharmacological advancements, further research is needed to achieve a deeper understanding of the disease mechanisms and development of reliable biomarkers.[2] A proposed early prognostic biomarker is grey matter (GM) loss.[3]

While numerous studies have investigated whole brain GM volume loss in MS, few have examined region-level changes over time.[4–6] Emerging evidence suggests distinct region-level volumetric differences between people with MS (pwMS) and healthy controls (HC), particularly reductions in subcortical volumes, such as the thalamus, and superior cortical regions such as the cingulate gyrus.[6,7] Moreover, regions like the thalamus and frontal lobes, have been strongly linked to disability progression, cognitive decline, psychiatric and sexual dysfunctions in pwMS.[4,8–10]

At the same time, significant variability in atrophy patterns is observed among pwMS.[7,9,11,12] Additionally, previous studies often lack the statistical power to observe small effects, typically do not utilize longitudinal designs to better approximate causality, and rarely include longitudinal comparisons between pwMS and HC. Finally, replications of specific findings are the exception rather than the norm.

For clinical translatability, longitudinal case-control comparisons are essential to map regional brain atrophy patterns. Such data provide insight into trajectories rather than mere correlations, allowing for analyses on temporal dynamics. The predictive value of regionally accelerated atrophy for clinical outcomes could pave the way for the development of new image-derived biomarkers. Furthermore, comparing GM atrophy rates between pwMS and HC, can help determine the region-specific pattern of accelerated brain ageing in MS. However, comparing ageing processes poses challenges, as brain volumes remain relatively stable from the twenties to forties.[13] Detecting subtle changes is particularly challenging over short scanning intervals, where known measurement variability can exceed the actual rate of brain change.[14] Since the rate of brain atrophy accelerates with age,[15–19] using samples from older HC provides a more robust comparison when examining longitudinal atrophy processes with younger pwMS.

In this study, we therefore investigated whether generalisable patterns of brain GM volume changes can be identified by replicating our findings in two independent cohorts of pwMS. Moreover, we studied age-related associations and atrophy patterns by utilizing a large cross-sectional and longitudinal magnetic resonance imaging (MRI) dataset of HC spanning a large age range (see Figure 1). This allowed us to compare atrophy patterns across samples to determine whether RRMS is associated with accelerated atrophy in different brain regions. Finally, we investigated whether these patterns were associated with disability progression, fatigue and cognitive decline.

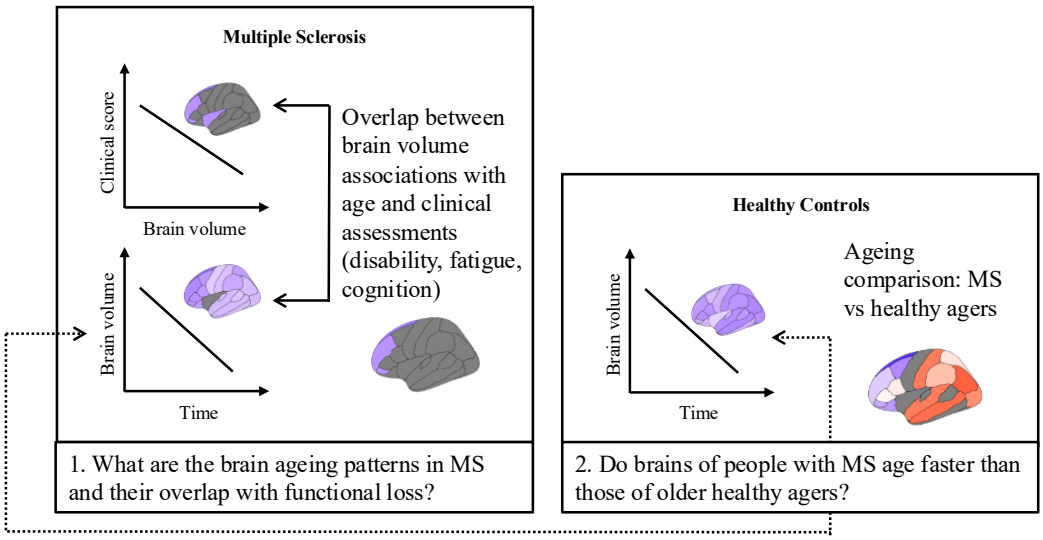

Figure 1. Key research questions and analysis workflow.

Methods

Participants

Multiple RRMS and HC datasets were used, which were sample across time (see Figure 2).

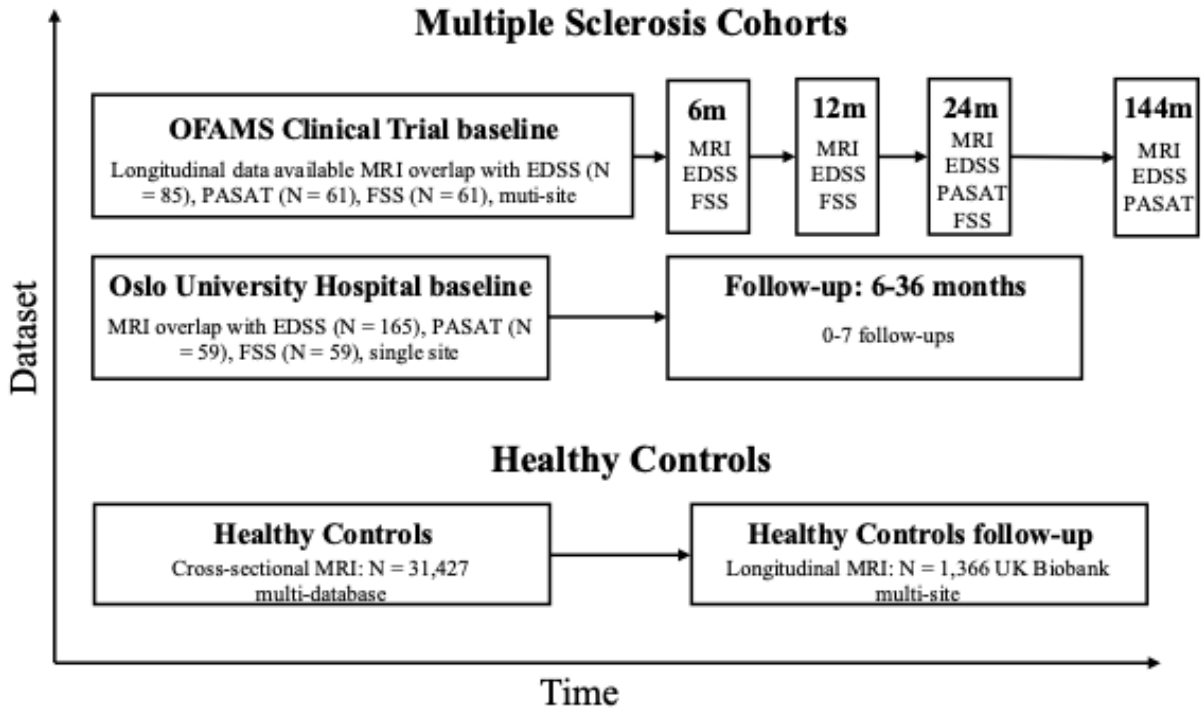

Figure 2. Overview of datasets, sampling and data availability. m = months. *Note*: Data availability indicated in the baseline condition describes the number of participants across time.

The first RRMS dataset included 88 pwMS with RRMS who previously participated in a multicentre clinical trial on omega-3 fatty acid in MS (OFAMS)[20]. The trial entailed data collection every 6 months over 2 years, followed by a single follow-up visit 10 years after the original trial concluded, with N=85 subjects participating in the 10-year follow-up visit (T_1w-scans=250). Participants were aged 38.9±8.3 years at baseline and 49.6±8.6 at the final follow-up, with 65.9% females at the first visit and 70.0% females at the last visit. The study was approved by the Norwegian Regional Committees for Medical and Health Research Ethics (REK, 814351 and 2016/1906) and registered as clinical trial (clinicaltrials.gov identifier: NCT00360906).

The second RRMS dataset was collected at the Oslo university hospital (OUH)[21,22] using all available data. Altogether 165 pwMS with RRMS participated (T_1w-scans=690), aged 38.1±10.0 (range 18-67) years at baseline and 40±10 at the last follow-up, with 74.5% being

females at the first visit, and 72.8% at the last visit. Ethics approval was obtained from REK (2011/1846).
Both RRMS cohorts included data on disability (Expanded Disability Status Scale, EDSS), cognitive function (3 second Paced Auditory Serial Addition Test, PASAT), and fatigue (Fatigue Severity Scale, FSS). EDSS scores were available at all MRI timepoints, however, PASAT scores were only available at baseline, month 24 and at 10-year follow-up for all OFAMS participants (N=85, overlap with T_1w-scans=282) and across MRI sessions in OUH data (N=53, overlap with T_1w-scans=82). FSS were obtained at baseline, month 6, 12 and 24 in OFAMS (N=85, overlap with T_1w-scans=127), and across MRI sessions in OUH data (N=82, T_1w-scans=125).

Additionally, we used several databases and samples to assemble large cross-sectional and longitudinal HC cohorts (N=31,427, T_1w-scans=32,793, of which T_1w-follow-up-scans=1,366), which matched the age-range of the two RRMS datasets (N=250, T_1w-scans=940, age range: 18-67). Longitudinal data was available for UK Biobank data (N=1,586) who were on average 59.1±5.2 years old at baseline and 61±5 at the single follow-up, with 55.8% being females. The cross-sectional sample (N=31,427) was, on average, aged 57.56±9.24 years, with 55.2% being female. For a full overview of the HC samples, including ethics approvals, see Supplementary Table 1. All participants gave their written informed consent. All methods were performed in accordance with the relevant guidelines and regulations (Declaration of Helsinki).

Magnetic resonance imaging (MRI) and data processing

T_1-weighted MRI data were obtained using various protocols, scanners, sites, and field strengths (1.5T or 3T). An overview of the acquisition protocols can be found in the original studies and Supplement (HCs: Supplemental Table 1, OFAMS: Supplemental Table 4 and original study[20], OUH: Supplemental Note 4 and original studies[21,23]). Volumetrics were extracted using the longitudinal pipeline of FreeSurfer (v7.1.1 for OFAMS and v6.0.0 for OUH data) and averaged across the 34 cortical from each hemisphere specified in the Desikan-Killany atlas. Additionally, we used the 7 subcortical brain regions provided by the standard FreeSurfer subcortical parcellation, including the thalamus pallidum, amygdala, hippocampus, putamen, caudate, in both hemispheres and the corpus callosum. Note that version differences were addressed in the next step, during harmonisation.

MRI Data harmonisation

We applied Combat, a data harmonisation pipeline, originally developed for batch effects in laboratory samples,[24] with recent extensions and applications to neuroimaging data.[25] The data were harmonised across scanners within both RRMS datasets to allow for replications of the findings from one dataset to another. HC data were harmonised independent from the MS data to avoid spillover harmonisation effects between RRMS and HC datasets. A longitudinal pipeline was applied to longitudinal data and a cross-sectional pipeline for cross sectional data.[26] As a quality control, we also report results without using Combat and instead the inclusion of the scanner as a random intercept in the Supplement.

Statistical analysis

First, we assessed the rate of volume change in the longitudinal samples using linear mixed models with a random intercept at the level of the individual, associating age with regional brain volume change, controlling for sex and the estimated total intracranial volume. The same association were examined without the random intercept, in the cross-sectional HC data, using simple linear models. We tested for 82 brain regions, leading to an adjusted $\alpha = \frac{0.05}{82}$, which was applied to each of the tests which attempted to replicate the findings across samples.

Second, we extracted the selected brain regions with significant volume change in step 1, across both RRMS datasets. As previously described,[27] these coefficients were statistically compared against the coefficients observed in the longitudinal UK Biobank sample.

$$Z = \frac{\alpha - \beta}{\sqrt{(SE\alpha)^2 + (SE\beta)^2}}$$

α and β are the respective standardized coefficients and $SE\alpha$ and $SE\beta$ representing their standard errors. For display, we present all significant Z-values in Figure 4, and Bonferroni adjusted significant values in-text (marked $p_{Bonferroni}$) for data combining both cohorts with $\alpha = \frac{0.05}{Nr+Nz}$, where Nr=82 is the number of brain regions, and Nz the number of Z-comparisons, equal to the number of significant replicating age-related volume changes across RRMS samples.

Third, to highlight the clinical relevance of brain volumes, we assessed longitudinal associations between the volume changes and EDSS scores for each dataset individually:

OUH and OFAMS data individually, as well as a combined dataset to maximise power, using random intercept models of the same structure as reported in step one but adding EDSS (over time) as explanatory variable of brain volume change across time. Again, we present uncorrected alpha and an adjusted $\alpha = \frac{0.05}{Nr+Nz}$. In case of no effect surviving the multiple comparison correction, we present uncorrected p-values, which are not marked as $p_{Bonferroni}$. The full analysis workflow is presented in Figure 1.

For comparability, we report standardized regression coefficients. The significance-level was set at a conventional α=0.05, and Bonferroni-adjusted for multiple comparisons. To reduce the convoluting effect of age differences between HC and RRMS samples, the older, longitudinal UK Biobank sample was restricted to participants not exceeding the maximum age of the RRMS samples. Similarly, the cross-sectional lifespan sample was restricted to the age range of the RRMS samples. For statistical analyses, R version 4.5.0 was used, for spatial representations of the examined brain areas, we used the ggseg R package,[28] and for power simulations the simr package, reported in Supplementary Note 3.[29]

Results

When examining the effect of age on brain GM volume change in pwMS in the two RRMS cohorts, we found replicable patterns of grey matter loss in multiple cortical and subcortical regions (Figure 3, Supplemental Table 2). The strongest replicable effect of age on grey matter volume were found in the *superior frontal cortex* (OUH: β_{age}<-0.27, $p_{Bonferroni}$<0.001; OFAMS: β_{age}<-0.51, $p_{Bonferroni}$<0.001), *pars orbitalis* (OUH: β_{age}<-0.25, $p_{Bonferroni}$<0.001; OFAMS: β_{age}<-0.41, $p_{Bonferroni}$<0.001), and *thalami* (OUH: β_{age}<-0.20, $p_{Bonferroni}$<0.001; OFAMS: β_{age}<-0.52, $p_{Bonferroni}$<0.001) in each hemisphere. The overall effect on whole brain GM atrophy was weaker than the observed regional effects in the OUH data (β_{age}=-0.25 [-0.15, -0.35], $p_{Bonferroni}$<0.001), but stronger in OFAMs data (β_{age}=-0.52 [-0.40, -0.64], $p_{Bonferroni}$<0.001).

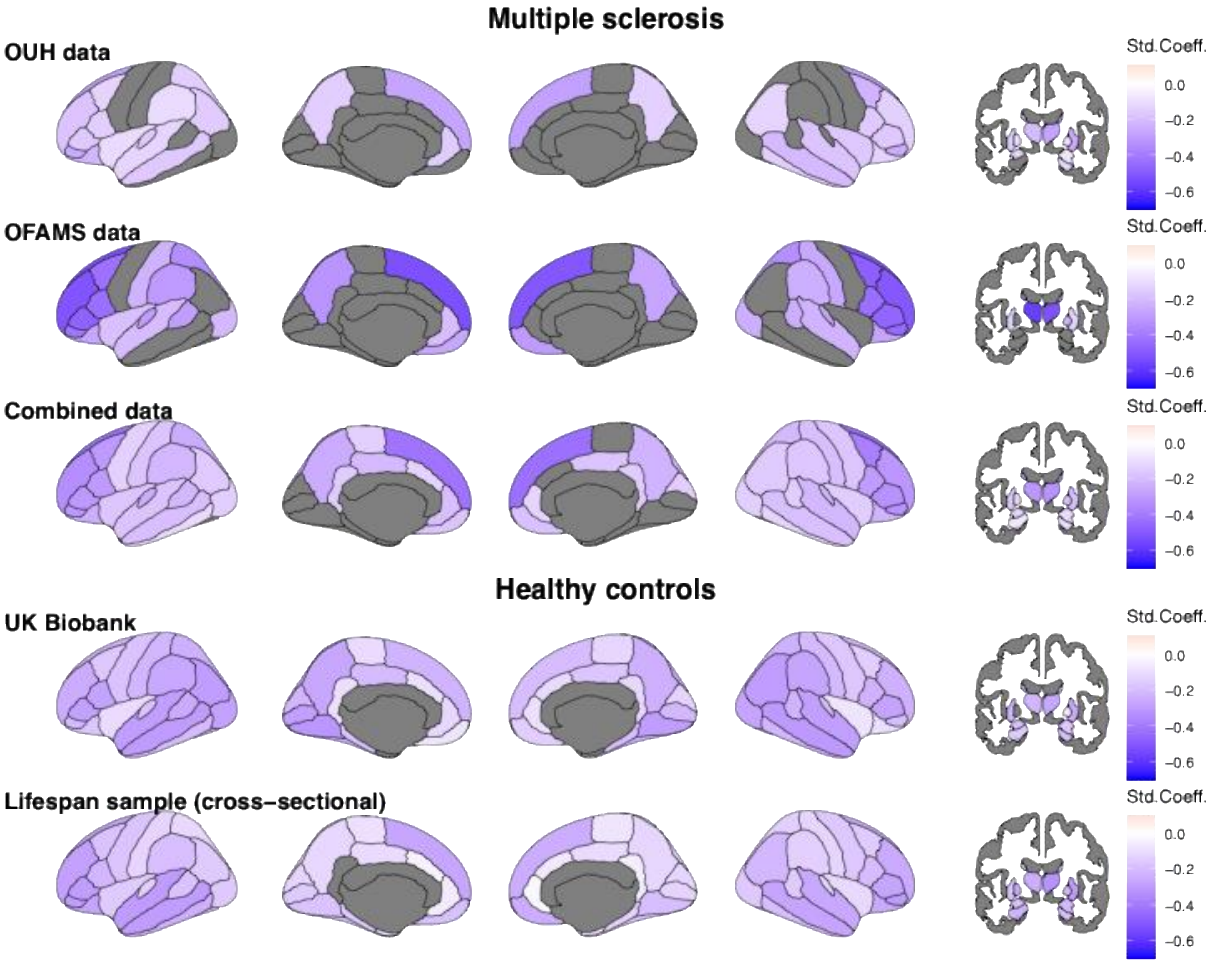

Figure 3. Volume change in RRMS and the UK Biobank, and age-associations in a large HC cohort. Top row: ageing effects in the OUH sample (baseline N=165). Second row: ageing effects in the OFAMS sample (baseline N=85). Third row: ageing effects in both RRMS cohorts combined. Fourth row: ageing effects in the longitudinal UK Biobank sample of healthy controls (N=1,586). Fifth row: age-associations in a cross-section sample of HC

(N=31,427). The age range across time points for the longitudinal MS dataset was 18.46-67.64 years, for the longitudinal healthy control sample 46.83-67.63, and for the cross-sectional sample 18.52 to 67.63 years. Grey regions indicate non-significant age effects (Bonferroni-corrected p≥0.05), blue areas indicate significant age associations after Bonferroni correction, with darker colours indicating stronger relationship indicated here by standardized coefficients (Std.Coeff.). For comparison, the effect of age on regional brain volumes in older HCs in the longitudinal sample (UK Biobank), and the cross-sectional lifespan sample, was smaller than in pwMS (Supplemental Note 5).

When statistically comparing the standardized longitudinal effects of age on brain volumes between younger pwMS and a 20 year older cohort of HCs (Figure 4), we found replicable faster GM atrophy in pwMS in the right superior frontal cortex (OUH: Z=-2.41, p=0.008; OFAMS: Z=-6.49, p<0.001), the right caudal middle frontal cortex (OUH: Z=-2.09, p=0.018; OFAMS: Z=-6.38, p<0.001), right caudate (OUH: Z=-2.46. p=0.007; OFAMS: Z=-2.10, p=0.018), and left frontal pole (OUH: Z=-2.43. p=0.007; OFAMS: Z=-2.99, p=0.001). For a full overview of differences in effects see Supplemental Table 3.

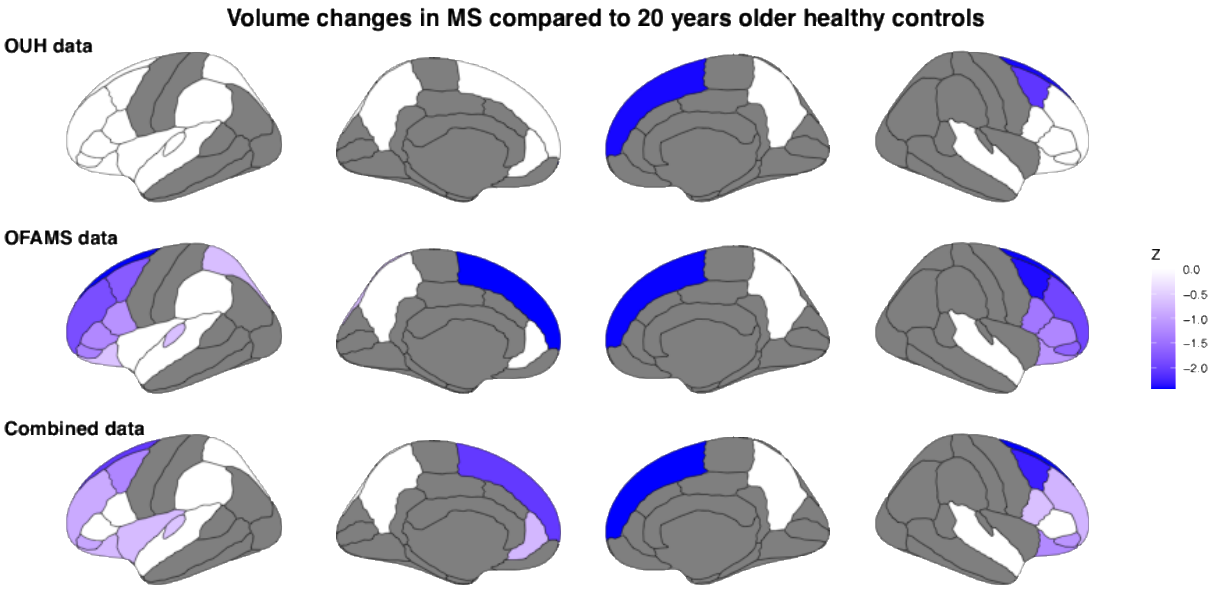

Figure 4. Volume changes in younger pwMS compared to 20 years older healthy controls. The regions considered for this analysis were selected based on cross-sample replicated significant brain changes/atrophy. Blue colours indicate faster volume change in RRMS, white colour indicates no difference (p>0.05). The top two rows indicate comparisons between standardised regression coefficients of single RRMS datasets and UK Biobank data. The bottom row indicates such comparisons between coefficients from combined RRMS data

and UK Biobank healthy controls. The presented significance level is not adjusted for multiple comparisons. Only cortical areas are displayed due to the lack of significant subcortical areas, except the caudate and thalamus, which are highlighted in-text.

None of the brain regions with significant GM volume loss over time (Figure 3) were significantly associated to disability, fatigue, or cognitive function after Bonferroni corrections (requiring $p<0.00045$) *in both samples*. A more lenient $\alpha=0.005$, to minimize false negatives in the smaller OUH dataset, highlights the association of EDSS as a predictor with the left (OFAMS: $\beta_{EDSS} = -0.09$ [-0.04, -0.14], $p<0.001$, OUH: $\beta_{EDSS} = -0.17$ [-0.07, -0.28], $p=0.002$) and right (OFAMS: $\beta_{EDSS} = -0.11$ [-0.06, -0.16], $p<0.001$, OUH: $\beta_{EDSS} = -0.16$ [-0.06, -0.27], $p=0.002$) pars orbitalis, and the left superior frontal cortex (OFAMS: $\beta_{EDSS} = -0.08$ [-0.04, -0.12], $p=0.001$, OUH: $\beta_{EDSS} = -0.12$ [-0.04, -0.21], $p=0.004$). However, the same standard did not yield consistent associations for fatigue, only highlighting the right caudal middle frontal cortex (OFAMS: $\beta_{fatigue} = -0.17$ [-0.06, -0.28], $p=0.002$, OUH: $p=0.836$) and the left precuneus (OFAMS: $\beta_{fatigue} = -0.20$ [-0.07, -0.34], $p=0.003$, OUH: $p=0.510$) in OFAMS data only; or PASAT, where the right lateral orbitofrontal cortex was significant in OUH but not OFAMS data (OFAMS: $p=0.264$, OUH: $\beta_{PASAT} = 0.15$ [0.06, 0.24], $p=0.003$), and vice versa for the left thalamus (OFAMS: $\beta_{PASAT} = 0.22$ [0.08, 0.35], $p=0.002$, OUH: $p=0.099$). For an analysis approach combining the two RRMS cohorts see Supplementary Note 1.
Conversely, EDSS across time consistently predicted whole grey matter atrophy, but with small effect sizes (OFAMS: $\beta_{EDSS}=-0.10$ [-0.06, -0.14], $p<0.001$; OUH: $\beta_{EDSS}=-0.12$ [-0.03, -0.20], $p=0.006$). Such associations were not found for FSS (OFAMS: $\beta_{fatigue}=-0.13$ [-0.02, -0.24] $p=0.018$, OUH: $p=0.331$), or PASAT (OFAMS: $p=0.327$, OUH: $\beta_{PASAT}=0.08$ [0.009, 0.15], $p=0.030$).

The presented age-effects were reproduced when using scanner as a fixed effect instead of Combat harmonisation (Supplemental Figures 1-2). However, the smaller effects of clinical assessments on brain volumes could not be reproduced in both examined RRMS cohorts without appropriate harmonisation (Supplemental Note 2).

Discussion

We used two independent RRMS cohorts to replicate regional GM atrophy patterns, examine their associations with clinical outcomes, and compare these patterns with those observed during normal aging. Consistent with previous studies,[7,9,11,12] our findings highlight the superior frontal cortex and thalamus as key regions affected by RRMS, associated with disability progression, and a less clear extent cognitive dysfunction, and fatigue. A key novel finding was that GM loss in the superior frontal lobes progressed more rapidly in RRMS than in HC who were approximately 20 years older. While cortical volume loss in HC appeared more widely distributed, atrophy in RRMS was more region-specific, with a pronounced impact on the superior frontal regions.

Mapping GM atrophy patterns in MS and distinguishing these from normal ageing is essential for understanding disease progression. Recent studies suggest that GM atrophy in MS is predictive of disability and death.[9,10] These findings suggest that GM atrophy could serve as an early marker of MS disease progression. However, more detailed mechanistic explanations, as well as the development of new prognostic and diagnostic imaging-derived markers, require spatially refined and adequately powered approaches.

Our findings align with previous cross-sectional and longitudinal studies emphasizing frontal GM atrophy,[7,9,30] which has been associated with a range of disabilities, including psychiatric signs, cognitive decline, autonomic and sexual dysfunctions.[8–10] However, there is variability across findings concerning the exact atrophy pattern, with different studies highlighting also other areas than the frontal lobes. For example, in line with our findings, one study shows the strongest atrophy pattern in the superior frontal lobe, yet also in the insula and paracentral gyrus.[30] Similarly, while other studies also report the frontal lobe as one of the most significantly changing regions, strong atrophy patterns were also found in the cingulate,[7] or parietal and occipital gyri.[9] The present study differs from previous studies by presenting several notable strengths. First, we leveraged two independent, longitudinal RRMS cohorts, allowing for replication of findings across different samples and timescales. Replications across samples provide a crucial validation step, allowing specifically to better account for measurement error than when examining a single isolated sample. Additionally, difference between samples – the extended follow-up period in the OFAMS cohort (up to 12 years) and the shorter but more recent OUH dataset – provided complementary insights into disease progression over varying temporal scales. Second, the inclusion of HC allowed us to

benchmark MS-specific atrophy patterns against the trajectories of normal aging. Third, we employed a harmonised image-processing pipeline, alongside standardized statistical models to minimize technical variability and enhance comparability of findings. Fourth, most previous studies do not report standardized effects, making comparability across studies difficult. Standardized effects, as presented here, provide a useful tool for the comparability of study results. Finally, our region-specific approach enabled us to identify key neuroanatomical targets - such as the superior frontal cortex and thalamus – that are linked to clinical outcomes, establishing a foundation for future research into MS-related neurodegeneration.

Our findings suggest that frontal lobe atrophy is implicated in multiple aspects of clinical disability, including physical disability (EDSS), cognitive function (PASAT), and fatigue (FSS). In contrast, thalamic atrophy appears to be more specifically associated with disability as measured by the EDSS. To our knowledge, there are no previous studies comparing different clinical measures to regional GM volumes. However, some studies show an association between thalamic volumes and EDSS over time.[4,31] Yet, these studies are limited by a lower field strength of 1.5T compared to the utilized 3T in the present study, entailing a lower signal-to-noise ratio and consequently poorer estimates of the true volumes. Our data suggested smaller associations between volumes and clinical outcomes. A possible explanation might be that most of the pwMS in our cohorts were only very slowly progressing. This is concordant with the progression of most pwMS on modern treatment plans and especially early diagnosis, underlining the usefulness of the presented findings.

Moreover, our findings highlight the differential vulnerability of distinct brain regions to MS-related neurodegeneration, each contributing uniquely and interactively to the overall symptomatology.[4,8–10] While regional atrophy provides valuable insights, it is the integration of these regions—representing complex network-level dysfunction—that likely underlies the heterogeneous clinical presentation of MS.[32–35] This interdependence poses a significant challenge for research, as focusing solely on region-specific effects may overlook broader patterns of distributed dysfunction. Nonetheless, the spatial detail provided by a region-based approach makes it superior to analyses of whole-brain volumes for understanding the nuanced impacts of MS.

Part of the diverse set of symptoms which characterise MS may be explained by the high level of connectivity between the frontal lobes and other brain regions. Functional

connectivity MRI studies on pwMS have consistently demonstrated the involvement of the frontal lobes (see for review[36]). Another key region identified in this study was the thalamus, which holds prognostic value due to its susceptibility to atrophy processes early in the disease course and association with physical disability.[4,9,10,12,37]

Ultimately, MS may be best understood as a network of symptoms and biomarkers, encompassing not only regional atrophy but also a constellation of image-derived markers and functional disruptions.[32–35] In this context, our study provides a foundational step by identifying generalizable patterns of brain atrophy, which can serve as a scaffold for more detailed investigations into the network-level mechanisms underlying disease progression.

There are limitations to our study, primarily related to sample-specific variability origination from multiple sources of error. While we identified distinct and replicable atrophy patterns, these sources of error may have limited the discovery rate of true effects across samples. Data from the OFAMS trial were sampled over a longer time (12 years) compared to those from the OUH sample (2 years), potentially explaining the stronger effects of age observed in the OFAMS sample. Measurement errors in the OUH sample may have been larger than the percentage change observed during the shorter sampling period.[13,14] The OUH data were collected at a single hospital, while OFAMS data were collected at multiple sites with different scanners and protocols, introducing variability. Differences in available disease-modifying treatments in the OFAMS study and the OUH data may have influenced the results, as data collection started in 2004 and 2012, respectively.
For practical reasons, we used two different versions of FreeSurfer (6.0.0 and 7.1.1) for data processing. However, the variability due to these processing choices is expected to be minor.[38,39] The Desikan-Killany atlas-based regional averages used in this study are less susceptible to spatial inaccuracies compared to atlases with smaller parcels or voxel-level analyses. Even for FreeSurfer versions further apart (5.3.0 and 7.4.1) than reported here (6.0.0 and 7.1.1), the correlations between features are shown to be acceptable (mean $r = 0.88$).[40] These parcels do, however, have reduced spatial specificity, both for detecting true underlying effects and for identifying version-induced variability in the spatial pattern.

In multicentre studies, harmonisation is crucial to mitigate technical, non-biological effects. For this study, we applied the longCombat pipeline specifically designed for longitudinal data.[26] While the Combat approach is considered state-of-the-art and has been shown to

effectively remove scanner effects,[26] it cannot completely be excluded that the method also inadvertently alters observable biological effects. To control for multiple comparisons, we applied a conservative Bonferroni-correction for α-level adjustment. While no definite guideline exists on which correction method is optimal, this conservative approach may have led to false negatives for detecting small effects, particularly in associations between clinical outcomes and GM atrophy. Given restricted statistical power to detect subtle associations, the reported brain-behaviour finding should be investigated in larger and more densely sampled cohorts.

In conclusion, we observed accelerated and early GM degeneration the frontal lobes in RRMS. Frontal lobe atrophy progressed more rapidly than in 20-year older HC and was associated with disability-progression. Our findings suggest that frontal lobe atrophy occurs early in RRMS and is a potential imaging biomarker for future MS trials.

Open Data and Materials

Summary statistics (anonymous) and utilized code can be found in the GitHub repository: https://github.com/MaxKorbmacher/MSatrophy. Multiple of the utilized dataset are sensitive, require IRB approval for usage, and can therefore not be openly shared.

Acknowledgments

We extend our sincere gratitude to the people with MS who participated in this study, whose contributions are invaluable. This research originated from different research efforts: one national intervention study (OFAMS) and ten years follow-up examinations, as well the ongoing data collections at the MS research group at the Oslo University Hospital.
We thank the OFAMS study group at numerous hospitals across Norway, as well as the group of clinicians and researchers at the Oslo University Hospital.
Finally, we want to extend our gratitude to the thousands of scanned participants in the UK Biobank study and facilitators who made the study possible.

Funding

This project was funded by the Norwegian MS Society. Analyses on the UK Biobank were performed on the Service for Sensitive Data (TSD) platform, at the University of Oslo, operated and developed by the TSD service group at the University of Oslo IT-Department (USIT). Computations were performed using resources provided by UNINETT Sigma2

(#NS9666S) – the National Infrastructure for High Performance Computing and Data Storage in Norway, supported by the Norwegian Research Council (#223273).

Part of the data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, 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.

Another part of data collection and sharing for this project was provided by the **Human Connectome Project** (HCP; Principal Investigators: Bruce Rosen, M.D., Ph.D., Arthur W. Toga, Ph.D., Van J. Weeden, MD). HCP funding was provided by the National Institute of Dental and Craniofacial Research (NIDCR), the National Institute of Mental Health (NIMH), and the National Institute of Neurological Disorders and Stroke (NINDS). HCP data are disseminated by the Laboratory of Neuro Imaging at the University of Southern California.

Competing interests

OAA has received a speaker's honorarium from Lundbeck, Janssen, Otsuka and Lilly, and is a consultant to Coretechs.ai and Precision Health.

LTW is a minor shareholder of baba.vision.
KMM has served on scientific advisory board for Alexion, received speaker honoraria from Biogen, Lundbeck, Novartis and Roche, and has participated in clinical trials organized by Biogen, Merck, Novartis, Otivio, Roche and Sanofi.
EAH received honoraria for advisory board activity from Sanofi-Genzyme, and his department has received honoraria for lecturing from Biogen and Merck.
ØT received speaker honoraria from and served on scientific advisory boards of Biogen, Sanofi-Aventis, Merck, and Novartis, and has participated in clinical trials organized by Merck, Novartis, Roche and Sanofi.
SW received speaker honoraria from Biogen, Sanofi-Aventis, and Janssen, and has participated in commissioned research projects funded by Merck, Novartis, and EMD Serono.
MK is a postdoctoral researcher whose career progression depends in part on publishing scientific research. As such, he has a professional interest in disseminating the findings of this study.
The remaining authors declare no other competing interests.

Contributions

MK: Conceptualisation and design of the study, Drafting the manuscript, Statistical analyses. Responsible for the overall content.
All authors contributed in different capacities to data acquisition, administration, or interpretation, as well as critical revisions of the article. All authors approved the final version of the manuscript.

References

1. Walton, C. *et al.* Rising prevalence of multiple sclerosis worldwide: Insights from the Atlas of MS, third edition. *Mult. Scler. Houndmills Basingstoke Engl.* **26**, 1816–1821 (2020).
2. Kuhlmann, T. *et al.* Multiple sclerosis progression: time for a new mechanism-driven framework. *Lancet Neurol.* **22**, 78–88 (2023).
3. Stankoff, B. & Louapre, C. Can we use regional grey matter atrophy sequence to stage neurodegeneration in multiple sclerosis? *Brain* **141**, 1580–1583 (2018).
4. Cagol, A. *et al.* Association of Brain Atrophy With Disease Progression Independent of Relapse Activity in Patients With Relapsing Multiple Sclerosis. *JAMA Neurol.* **79**, 682–692 (2022).
5. Nygaard, G. O. *et al.* Cortical thickness and surface area relate to specific symptoms in early relapsing–remitting multiple sclerosis. *Mult. Scler. J.* **21**, 402–414 (2015).
6. Lansley, J., Mataix-Cols, D., Grau, M., Radua, J. & Sastre-Garriga, J. Localized grey matter atrophy in multiple sclerosis: A meta-analysis of voxel-based morphometry studies and associations with functional disability. *Neurosci. Biobehav. Rev.* **37**, 819–830 (2013).
7. Eshaghi, A. *et al.* Progression of regional grey matter atrophy in multiple sclerosis. *Brain* **141**, 1665–1677 (2018).
8. Andravizou, A. *et al.* Brain atrophy in multiple sclerosis: mechanisms, clinical relevance and treatment options. *Autoimmun. Highlights* **10**, 7 (2019).
9. Eijlers, A. J. C. *et al.* Cortical atrophy accelerates as cognitive decline worsens in multiple sclerosis. *Neurology* **93**, e1348–e1359 (2019).
10. Grey matter pathology in multiple sclerosis - The Lancet Neurology. https://www.thelancet.com/journals/lancet/article/PIIS1474-4422(08)70191-1/abstract.
11. Martinez-Heras, E. *et al.* Diffusion-based structural connectivity patterns of multiple sclerosis phenotypes. *J. Neurol. Neurosurg. Psychiatry* **94**, 916–923 (2023).
12. Meijboom, R. *et al.* Patterns of brain atrophy in recently-diagnosed relapsing-remitting multiple sclerosis. *PLOS ONE* **18**, e0288967 (2023).
13. Wang, M.-Y. *et al.* The within-subject stability of cortical thickness, surface area, and brain volumes across one year. 2024.06.01.596956 Preprint at https://doi.org/10.1101/2024.06.01.596956 (2024).
14. Vidal-Piñeiro, D. *et al.* Reliability of structural brain change in cognitively healthy adult samples. *Imaging Neurosci.* (2025) doi:10.1162/imag_a_00547.

15. Korbmacher, M. *et al.* Distinct longitudinal brain white matter microstructure changes and associated polygenic risk of common psychiatric disorders and Alzheimer's disease in the UK Biobank. *Biol. Psychiatry Glob. Open Sci.* 100323 (2024) doi:10.1016/j.bpsgos.2024.100323.
16. Fujita, S. *et al.* Characterization of Brain Volume Changes in Aging Individuals With Normal Cognition Using Serial Magnetic Resonance Imaging. *JAMA Netw. Open* **6**, e2318153 (2023).
17. Bethlehem, R. A. *et al.* Brain charts for the human lifespan. *Nat* **604**, 525–533 (2022).
18. Korbmacher, M. *et al.* Brain-wide associations between white matter and age highlight the role of fornix microstructure in brain ageing. *HBM* **44**, (2023).
19. Korbmacher, M. *et al.* Brain asymmetries from mid-to late life and hemispheric brain age. *Nat. Commun.* **15**, 956 (2024).
20. Torkildsen, O. *et al.* ω-3 fatty acid treatment in multiple sclerosis (OFAMS Study): a randomized, double-blind, placebo-controlled trial. *Arch. Neurol.* **69**, 1044–1051 (2012).
21. Høgestøl, E. A. *et al.* Cross-Sectional and Longitudinal MRI Brain Scans Reveal Accelerated Brain Aging in Multiple Sclerosis. *Front. Neurol.* **10**, 450 (2019).
22. Sowa, P. *et al.* Restriction spectrum imaging of white matter and its relation to neurological disability in multiple sclerosis. *Mult. Scler. Houndmills Basingstoke Engl.* **25**, 687–698 (2019).
23. Sowa, P. *et al.* Restriction spectrum imaging of white matter and its relation to neurological disability in multiple sclerosis. *Mult. Scler. Houndmills Basingstoke Engl.* **25**, 687–698 (2019).
24. Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118–127 (2007).
25. Radua, J. *et al.* Increased power by harmonizing structural MRI site differences with the ComBat batch adjustment method in ENIGMA. *NeuroImage* **218**, 116956 (2020).
26. Beer, J. C. *et al.* Longitudinal ComBat: A method for harmonizing longitudinal multi-scanner imaging data. *NeuroImage* **220**, 117129 (2020).
27. Clogg, C. C., Petkova, E. & Haritou, A. Statistical methods for comparing regression coefficients between models. *Am. J. Sociol.* **100**, 1261–1293 (1995).
28. Mowinckel, A. M. & Vidal-Piñeiro, D. Visualization of Brain Statistics With R Packages ggseg and ggseg3d. *Adv. Methods Pract. Psychol. Sci.* **3**, 466–483 (2020).
29. Green, P. & MacLeod, C. J. SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods Ecol. Evol.* **7**, 493–498 (2016).

30. Steenwijk, M. D. *et al.* Cortical atrophy patterns in multiple sclerosis are non-random and clinically relevant. *Brain* **139**, 115–126 (2016).
31. Magon, S. *et al.* Volume loss in the deep gray matter and thalamic subnuclei: a longitudinal study on disability progression in multiple sclerosis. *J. Neurol.* **267**, 1536–1546 (2020).
32. Eijlers, A. J. C. *et al.* Increased default-mode network centrality in cognitively impaired multiple sclerosis patients. *Neurology* **88**, 952–960 (2017).
33. Kennedy, K. E. *et al.* Multiscale networks in multiple sclerosis. *PLOS Comput. Biol.* **20**, e1010980 (2024).
34. Fleischer, V. *et al.* Prognostic value of single-subject grey matter networks in early multiple sclerosis. *Brain* **147**, 135–146 (2024).
35. Martí-Juan, G. *et al.* Using The Virtual Brain to study the relationship between structural and functional connectivity in patients with multiple sclerosis: a multicenter study. *Cereb. Cortex* **33**, 7322–7334 (2023).
36. Rocca, M. A., Schoonheim, M. M., Valsasina, P., Geurts, J. J. G. & Filippi, M. Task- and resting-state fMRI studies in multiple sclerosis: From regions to systems and time-varying analysis. Current status and future perspective. *NeuroImage Clin.* **35**, 103076 (2022).
37. Schoonheim, M. M. *et al.* Thalamus structure and function determine severity of cognitive impairment in multiple sclerosis. *Neurology* **84**, 776–783 (2015).
38. Sokołowski, A. *et al.* The impact of FreeSurfer versions on structural neuroimaging analyses of Parkinson's disease. *bioRxiv* 2024–11 (2024).
39. Hedges, E. P. *et al.* Reliability of structural MRI measurements: The effects of scan session, head tilt, inter-scan interval, acquisition sequence, FreeSurfer version and processing stream. *NeuroImage* **246**, 118751 (2022).
40. Korbmacher, M., Westlye, L. T. & Maximov, I. I. FreeSurfer version-shuffling can enhance brain age predictions. *NeuroImage Rep.* **4**, 100214 (2024).