Neuroradiology
https://doi.org/10.1007/s00234-026-04060-2

RESEARCH

An automated quantitative report for multiple sclerosis using only 3D T2-fluid-attenuated inversion recovery MRI

Zoe Mendelsohn[1,2] · Ferran Prados[2,3] · Giuseppe Pontillo[2,4] · Olivia Goodkin[2,5] · Bas Jasperse[4] · Alex Rovira[6] · Matthew Grech-Sollars[7,8] · Jordan Colman[2,9] · James Moggridge[2,7] · Weaam Hamed[2,7] · Ayodeji Ijishakin[10] · Michael A. Foster[2] · Nima Mahmoudi[1] · Mike P. Wattjes[1] · Arturo Brunetti[11] · Gabriel Gonzalez-Escamilla[12] · Sergiu Groppa[12] · Einar August Høgestøl[13,14] · Lars T. Westlye[14,15] · Jacqueline Palace[16] · Silvia Messina[16] · Rosa Cortese[17] · Nicola De Stefano[17] · Jaume Sastre-Garriga[6] · Stefan Ropele[18] · Christian Enzinger[18] · Maria A. Rocca[19] · Massimo Filippi[20] · Jeremy Chataway[2,21] · Charles Wade[2] · Ahmed Toosy[2] · Olga Ciccarelli[2,21] · Hugh Pemberton[2] · James H. Cole[10] · Michael Scheel[1] · Jawed Nawabi[1] · Tarek Yousry[2,7] · John Thornton[2,7] · Frederik Barkhof[2,4]

Received: 20 February 2026 / Accepted: 27 May 2026

Abstract

Purpose Automated tools quantifying multiple sclerosis (MS) imaging biomarkers often require non-routine MRI sequences and lack MS reference data. We developed an open-source quantitative report (QReport) that integrates validated 3D T2-FLAIR quantification methods with multi-centre MS and healthy reference models, and presents outputs in a structured graphical report to support contextualised interpretation of clinically relevant biomarkers.

Methods 2516 cross-sectional 3D T2-FLAIR scans from people with MS (pwMS) and healthy controls (HC) were retrospectively collected from 14 centres within Magnetic Resonance Imaging in MS (MAGNIMS) and affiliated sites, as well as open-source datasets. Validated T2-FLAIR-based algorithms quantified total and regional lesion count (LC), lesion volume (LV), brain volume (BV), and brain age gap estimation (BrainAGE). Distributions in pwMS and HC were estimated using quantile regression. A QReport was designed to present biomarkers and reference models in graphical formats. Four neuroradiologists assessed agreement between QReport outputs and their visual assessment, and evaluated its usefulness, in 22 cases.

Results We analysed scans from 1723 HC (age, mean ± SD: 54.5 ± 16.0; range: 18–75; F/M: 949/774) and 793 pwMS (age, mean ± SD: 43.0 ± 11.1; range: 18–75; F/M: 538/255) across 14 centres. The QReport presents single-subject measures contextualised against the 95th, 50th, and 5th percentile distributions in pwMS and HC, and includes BrainAGE. In 94% of evaluations, QReport outputs demonstrated Moderate-to-Complete agreement with visual assessment and were rated as useful in 82%.

Conclusion We developed an MS QReport requiring only 3D T2-FLAIR, integrating validated quantification algorithms and incorporating BrainAGE within a clinically interpretable framework.

Keywords Multiple sclerosis · MRI · Fluid-attenuated inversion recovery · Quantitative report · QReport

Abbreviations

MS	Multiple sclerosis	pwMS	People with MS
MRI	Magnetic resonance imaging	Healthy control	HC
FLAIR	Fluid–attenuated inversion recovery	Automated quantitative report	QReport
CNS	Central nervous system		
BrainAGE	Brain age gap estimation	T1w	T1–weighted

Extended author information available on the last page of the article

Published online: 06 June 2026

T2w	T2–weighted
LV	Lesion volume
BV	Brain volume
CE	Conformité Européenne
FDA	U.S. Food and Drug Administration
QNI	Quantitative Neuroradiology Initiative
MAGNIMS	Magnetic Resonance Imaging in MS
NHS	National Health Service
NHNN	National Hospital for Neurology and Neurosurgery
UCLH	University College London Hospital
GIF	Geodesic Information Flows
DKT	Desikan–Killiany–Tourville
GM	Grey matter
WM	White matter
CGM	Cortical grey matter
DWM	Deep white matter
DGM	Deep grey matter
BPF	Brain Parenchymal Fraction
TIV	Total Intracranial Volume
CSF	Cerebrospinal fluid
CNN	Convolutional neural network
MAE	Mean Absolute Error
NIFD	Frontotemporal Lobar Degeneration Neuroimaging Initiative/Neuroimaging in Frontotemporal Dementia
MPI	Max Planck Institute
DMT	Disease modifying therapy
GCA	Global Cortical Atrophy scale
TLV	Total lesion volume
ADNI	Alzheimer's Disease Neuroimaging Initiative
EPAD	European Prevention of Alzheimer's Dementia
AC_1	Gwet's Agreement Coefficient for binary variables
AC_2	Gwet's Agreement Coefficient for ordinal variables
RRMS	Relapsing remitting multiple sclerosis
PPMS	Primary progressive multiple sclerosis
SPMS	Secondary progressive multiple sclerosis
EDSS	Expanded Disability Severity Scale

Introduction

Automated radiological reporting tools have been developed to standardise and streamline assessment of quantitative imaging biomarkers (QIB) [1–4]. Several such tools produce regulatory-approved quantitative reports (QReports) for use in multiple sclerosis (MS) [5], which have been shown to improve inter-rater agreement and efficiency of reporting [6–8].

However, implementation of MS QReports in routine clinical practice is not yet widespread [5]. A barrier to translation is the discrepancy between imaging guidelines and the technical requirements of existing tools [5]. MS QReports require both non-contrast 3D T1-weighted (T1-w) and T2-FLAIR sequences, yet non-contrast T1-w imaging is not recommended in clinical routine [5, 9, 10]. In contrast, 3D T2-FLAIR is strongly recommended for lesion detection due to its superior sensitivity [9, 10]. Implementing MS QReports as currently designed could therefore require modifications to clinical protocols and additional scan time, which may be costly and inefficient. Several research-based algorithms for MS lesion and brain segmentation and volumetry using only 3D T2-FLAIR have been developed and validated in MS [11–19]. Leveraging these algorithms to develop clinically applicable QReports could improve compatibility with standard workflows.

The Quantitative Neuroradiology Initiative (QNI) provides an established framework for the development of QReports, emphasising the importance of aligning tools with clinical needs [20]. A recent study systematically defined end-user requirements for QReports in MS [21]. 3D FLAIR–only tools were recommended for clinical implementation, and key QIB together with their intended roles in routine practice were specified [21]. To facilitate interpretation of QIB, end users recommended the inclusion of MS-specific reference data alongside healthy control (HC) cohorts [21]. However, existing QReports lack MS reference models [5], despite evidence that MS-specific models outperform HC models in explaining disease severity and predicting disease progression [22].

In this work, we developed an MS QReport that adheres to both imaging guidelines and user-requirements, following the QNI framework [4, 9, 10, 20, 21]. We integrated 3D T2-FLAIR quantification algorithms that have been validated in MS to generate biomarkers that were systematically prioritised as clinically relevant in the user-requirement analysis [16, 18, 19, 23–27]. An MS-validated FLAIR BrainAGE algorithm was included to provide a supportive marker of global brain health [18, 19]. We compiled large, multi-centre MS and healthy reference data to contextualise QIB and designed a graphical report to facilitate clinical interpretation. Finally, following the QNI clinical validation pathway, we completed the credibility stage through multi-centre, independent neuroradiologist review of the QReport. At this stage, individual QReport components were systematically evaluated for concordance with expert routine visual interpretation in real-world data [20].

Methods

All statistical analyses were conducted in RStudio (V4.4.0).

Ethics

The project was carried out in accordance with the Declaration of Helsinki. The requirement for informed consent was waived due to the retrospective nature of the study. All scans from proprietary datasets had been performed as part of routine clinical care. A data transfer agreement (DTA) was signed by each centre that provided data. The final protocol was approved by the local ethics committee (project ID number 19143/001) and the MAGNIMS (Magnetic Resonance Imaging in MS) Study Group Steering Committee.

Participants

To build an MS and HC cohort to contextualise MS QIB, we used a retrospective, cross-sectional approach, collecting 3D T2-FLAIR scans with clinical and demographic data from healthy participants and people with MS (pwMS) from 14 sites. These included Magnetic Resonance Imaging in MS (MAGNIMS) and MAGNIMS-affiliated centres and the following open-source datasets: Frontotemporal Lobar Degeneration Neuroimaging Initiative (NIFD) [28], Max Planck Institute (MPI) [29, 30], Alzheimer's Disease Neuroimaging Initiative (ADNI) [31], and European Prevention of Alzheimer's Dementia (EPAD) [32, 33]. The data sources and respective acquisition parameters are described in Supplementary Table 1. HC were classified as healthy by the respective contributing studies or centres. Participants in both cohorts were required to be aged between 18 and 75 years; individuals older than 75 were excluded due to sparse data in this age range. PwMS met the 2010 or 2017 McDonald diagnostic criteria [34].

QReport pipeline

The fully-automated quantification pipeline for 3D T2-FLAIR scans is open source [35] and outlined below.

1. *Image processing*

   i) Lesion Segmentation

   After voxel resampling to 1 mm^3, we applied an in-house version of nicMSlesions that uses only 3D T2-FLAIR input to perform skull-stripping and lesion segmentation [12, 14, 36]. The algorithm has been applied in MS populations [26, 27] and compared to manual segmentation (see Supplementary methods and Supplementary Tables 1 and 2).

   ii) Brain Segmentation

   N4 bias field correction [37] and lesion filling – performed using a modality-agnostic patch-based method [38] – were followed by brain tissue segmentation with T2-FLAIR-only GIF (v2), which has been validated in an MS population [11, 15, 35]. GIF is an atlas-based approach that has been shown to outperform multi-atlas methods—which rely on pairwise propagation of labels directly from each atlas to the target—particularly in settings with morphological differences between training and testing populations, such as the presence of MS lesions [11]. Probabilistic tissue maps were generated from the 3D FLAIR images followed by parcellation into regions defined by the Desikan-Killiany-Tourville (DKT) atlas [39].

   iii) Volumetry

   Total intracranial volume (TIV) was calculated as the sum of white matter (WM), grey matter (GM), and cerebrospinal fluid (CSF) volume. BV estimates derived from the segmentation stage were divided by TIV to account for head size. Brain parenchymal fraction (BPF) was calculated by dividing brain parenchymal volume (excluding CSF) by TIV.

   iv) Lesion Classification

   WM and GM were split into concentric bands to define lesion location as cortical GM (CGM), juxtacortical, deep WM (DWM), periventricular, cerebellum, brainstem, deep GM (DGM), leukocortical (largest proportion in CGM with overlap in WM), and cross-regional (largest proportion in DWM with overlap in DGM) [17]. This method has been validated against visual rating scales of regional distribution of white matter hyperintensities (WMH) and applied in MS [17, 23–25].

2. *Normative modelling*

i) Cross–centre harmonisation

As HC and MS data were collected from multiple scanners, BPF, total LV (TLV), and regional LV and BV were harmonised using ComBat (empirical Bayes) [40]. This method removed scanner-specific effects, preserving variance associated with age and sex [41]. The distribution of age, assessed prior to harmonisation, overlapped across scanners, reducing the risk of confounding (Supplementary Fig. 1).

ii) Reference curves of TLV, BPF and regional BV

We utilised linear quantile regression to estimate conditional quantiles (5th, 50th, 95th percentiles) across TLV and BPF distributions in HC and MS [42, 43]. Unlike methods that assume normality or constant variance, quantile regression is robust to data variability and effectively captures distributional shifts beyond the mean, making it well-suited for distinguishing MS from HC populations [43]. Age and sex were included as covariates. Model calibration was assessed by comparing predicted and observed proportions below each quantile (5%, 50%, 95%).
Regional BV was contextualised using percentiles derived from a Gaussian approximation of HC GM volumes, adjusted for age and sex (see Supplementary List 1).

iii) BrainAGE

BrainAGE was quantified using a T2-FLAIR-based model (Mean Absolute Error (MAE) = 3.08 years) that has been validated in MS and outperformed state-of-the-art architecture [18, 19].

Quality control measures

All MRI data processed by the pipeline are required to meet predefined acquisition and metadata standards, including T2-FLAIR imaging with an approximate voxel size of $1 \times 1 \times 1$ mm and essential demographic information (age and sex). Prior to processing, scans and metadata underwent visual review by an experienced reader to ensure overall image quality and protocol consistency. The pipeline operates fully automatically and incorporates internal consistency checks that halt processing if required information is missing or if any step fails, thereby preventing the generation of invalid results.

QNI credibility study

The QNI framework requires an initial clinical evaluation of a QReport—termed a "credibility study"—to establish real-world concordance between the report and routine expert assessment prior to large-scale accuracy studies. To fulfil this step, we assessed agreement between QReport outputs (including segmentation outputs, volumetric measures, reference percentile contextualisation, and graphical presentation) and expert visual interpretation of corresponding real-world 3D T2-FLAIR scans.

Four independent neuroradiologists from four centres (10–40 years of experience) participated. Clinicians were blinded to all clinical and demographic information except age and sex, reflecting routine reporting conditions. A dedicated web platform (http://qni.cs.ucl.ac.uk/) was designed to display the scans in three orthogonal planes alongside the QReport in a randomised order, together with assessment questions (Fig. 1).

For each QReport component (lesion, CGM, and DGM masks; TLV; BPF; BrainAGE; and regional LV and BV quantifications), clinicians graded agreement with their visual assessment using a three-point Likert scale (1 = Limited, 2 = Moderate, 3 = Complete-almost complete). To evaluate interpretability and design, raters also judged the usefulness of QReport results per case as either "useful" or "not useful".

Participants

HC and MS participants were retrospectively selected and excluded from the data collected for the pipeline development (Supplementary Table 3). To ensure that results were generalisable, we included data from multiple centres, sex balance, a broad age range, different scanner types and field strengths, and an even distribution of degrees atrophy and lesion burden. A neuroradiologist (G.P.) scored atrophy using the Global Cortical Atrophy (GCA) scale [44] and manually counted lesions classifying them as None, Minimal (1–10), Moderate (10–30), and Extensive (>30). HCs were included that had no lesions or atrophy.

Statistical analysis

The 88 case-by-rater credibility study evaluations were analysed descriptively. Inter-rater reliability was assessed using Gwet's agreement coefficient (AC) [45]. Unlike Cohen's kappa, Gwet's AC is particularly suited for imbalanced data with few observations at one level of the ordinal scale, e.g. 1 = Limited [46].

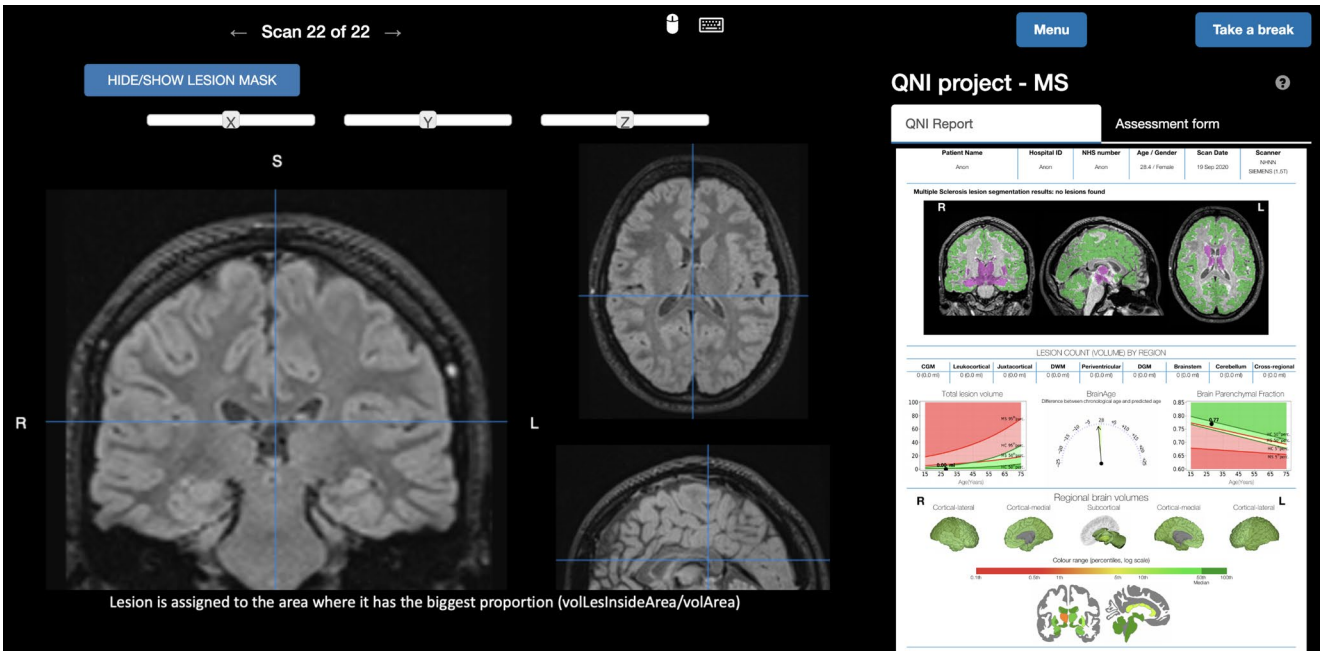

Fig. 1 Screenshot of the credibility study website portraying a scrollable MRI scan in three planes (left) and the Qreport (right)

Table 1 Demographic, clinical, and imaging characteristics, and sources of healthy control cohort

Participants, *N* cases	Source	Age, Mean ± SD (years)	Age, range (years)	Female sex, %
1723	ALL	54.5 ± 16.0	18–75	55
43	[a]ADNI	68.9 ± 4.4	58–75	63
46	Barcelona [b](MAGNIMS)	44.5 ± 9.7	23–60	72
794	[c]EPAD	63.6 ± 6.3	50–75	58
33	Graz [b](MAGNIMS)	30.6 ± 8.9	22–52	61
14	Hannover	52.1 ± 8.6	41–71	29
46	Mainz	29.9 ± 9.6	20–57	54
151	Milan [b](MAGNIMS)	33.2 ± 11.6	18–72	50
110	Max Planck Institute	35.9 ± 18.3	22–73	20
61	Naples	43 ± 11.2	24–64	50
110	[d]NIFD	62.0 ± 6.9	36–75	55
277	Oslo	56.2 ± 15.0	20–75	63
11	Oxford [b](MAGNIMS)	35.9 ± 13.6	24–58	46
27	Siena [b](MAGNIMS)	36.1 ± 8.4	25–56	44

[a]*ADNI* = Alzheimer's Disease Neuroimaging Initiative, [b]*MAGNIMS* = Magnetic Resonance Imaging in MS, [c]*EPAD* = European Prevention of Alzheimer's Dementia, [d]*NIFD* = Neuroimaging in Frontotemporal Dementia

Results

HC cohort

For the HC reference model, 1723 participants were included (mean age: 54.5 ± 16.0 years; range: 18–75; female/male: 949/774). Scans were performed between 2013 and 2022. Demographic characteristics are summarised in Table 1, while site-specific information, including imaging acquisition protocols, is presented in Supplementary Table 3.

MS cohort

For the MS reference model, 793 participants were included (mean age: 43.0 ± 11.1 years; range: 18–75; female/male: 538/255) across all three MS phenotypes — relapsing-remitting (RRMS), primary progressive (PPMS), and secondary progressive MS (SPMS), scanned between 2017 and 2021. Demographic and clinical data are described in Table 2. Site-specific information, including imaging acquisition parameters, is detailed in Supplementary Tables 3 and

Table 2 Demographic, clinical, and imaging characteristics of MS cohort sourced from the National Hospital for Neurology and Neurosurgery in London, United Kingdom

Participants, *N* cases	Age, mean ± SD (years)	Age, range (years)	Female sex, %	[a]EDSS, Mean ± SD	Disease duration, Mean ± SD (years)	Phenotype at scan ([b]RRMS/[c]PPMS/[d]SPMS)	T2–FLAIR lesion volume, Mean ± SD (mL)
793	43.0 ± 11.1	18–75	66	3.7 ± 2.3	13.3 ± 8.2	683/45/65	13.0 ± 11.4

[a]*EDSS* = Expanded Disability Status Scale, [b]*RRMS* = relapsing remitting MS, [c]*PPMS* = primary progressive MS, [d]*SPMS* = secondary progressive MS

disease-modifying therapies (DMT) in use at the time of scanning are detailed in Supplementary Table 4.

The MS QReport

The QReport displays segmentation masks, total and regional LC and LV, BPF, regional BV, and BrainAGE (Fig. 2).

At the top of the report, an anatomical slice with peak lesion burden is automatically selected and overlaid with segmentation masks, ensuring that the most affected region is visualised. LV per slice is plotted in each plane to retain 3D information.

The TLV and BPF graphs show sex-adjusted distributions from HC and MS reference cohorts. Across all ages, LV at the 50th percentile was consistently higher in pwMS than in HC. From age 45 onwards, the 95th percentile in HC surpassed the MS 50th percentile, possibly due to age-related accumulation of lesions in the highest 5% of the HC population. The MS reference curves of BPF were consistently lower than the HC curves.

Regional BV is visualised using animated renderings with colour-coded percentile estimates relative to HC reference data, sex- and age-adjusted [47]. A logarithmic colour scale is used to highlight low percentiles.

BrainAGE is displayed on a dial depicting the difference between predicted and chronological age. If predicted age is higher than chronological age, the segment on the dial is red. For a lower predicted age, the segment is green.

The QReports in Fig. 3 highlight the importance of interpreting quantification results within the context of reference data. In the 50-year-old participant with MS (Fig. 3a), the TLV exceeds both the 50th percentile for pwMS and the 95th percentile for HC. In contrast, a similar lesion load in an older individual could reflect age-related changes, provided MS diagnostic criteria are not met. The 75-year-old healthy participant (Fig. 3b) has a comparable lesion load, but it falls between the 50th and 95th percentiles of the HC distribution. The 50-year-old also shows a BrainAGE of +25 years, indicating that the predicted age is substantially greater than the chronological age. By contrast, the 75-year-old HC has a BrainAGE near zero, suggesting no deviation from actual age. Additionally, for the 75-year-old, both BPF and regional BV are within normal limits, approximately at or above the 50th percentile for HC. In contrast, Fig. 3a shows pathological deviation, with BPF at the 5th percentile for MS and thalamic volumes in the lowest HC percentiles.

QReport credibility study

The credibility study included 17 MS and 5 HC cases. Clinical, demographic, and imaging characteristics are summarised in Table 3 and described per case, along with the data sources, in Supplementary Table 5.

In line with the QNI framework, credibility was evaluated as the agreement between multi-centre expert assessment and QReport-derived outputs in real-world cases. Figure 4 summarises agreement across predefined levels (1 = Limited, 2 = Moderate, and 3 = Complete-almost complete). Overall, Moderate-to-Complete agreement was observed in 94% of evaluations, with Complete agreement in 60%.

Global measures demonstrated consistently high agreement. TLV, BPF, CGM masks, and BrainAGE achieved the highest agreement level in over 80% of evaluations, with up to 100% rated as Level 2 or 3. Lesion masks showed Moderate-to-Complete agreement in 93% of evaluations, including 61% at the highest level of agreement.

QReport regional measures were most frequently rated as showing moderate agreement with visual assessment: 81% of evaluations for regional LV, 94% for deep GM masks, and 90% for regional BV. The highest agreement level was achieved in 20%, 24%, and 45% of evaluations, respectively.

Inter-rater reliability for agreement between visual assessment and QReport outputs ranged from $AC_2 = 0.52$ ("Moderate") to 0.86 ("Almost perfect") across components ($p < 0.001$ for all; Supplementary Table 6). Highest reliability was observed for evaluation of BPF (AC_2=0.86, 95% CI: 0.78–0.95) and TLV (AC_2=0.80, 95% CI: 0.71–0.90). Agreement with regional BV and LV showed the greatest variability (AC_2=0.52 and 0.57, respectively).

Overall, the QReport was considered useful in 82% of evaluations across MS and HC cases. Component-specific usefulness is shown in Fig. 5. TLV estimation and visualisation were deemed useful in 97% of evaluations, including 100% of MS cases. BPF and BrainAGE outputs were considered useful in 96% of evaluations.

Regional LV and BV estimations and visualisations were rated as useful in approximately 50% of evaluations. Notably, for regional LV, the proportion of "not useful" evaluations remained at 50% even at the highest agreement level, and for regional BV at 25%.

Discussion

We developed an open-source MS QReport that integrates validated 3D T2-FLAIR quantification algorithms and systematically defined clinical requirements into a structured, reference-contextualised reporting framework. The report includes total and regional lesion volume and count, brain volume measures contextualised against MS and HC distributions, and BrainAGE. All underlying quantification algorithms have been previously technically validated in MS [16–19, 23–27]. Through collaboration and data sharing

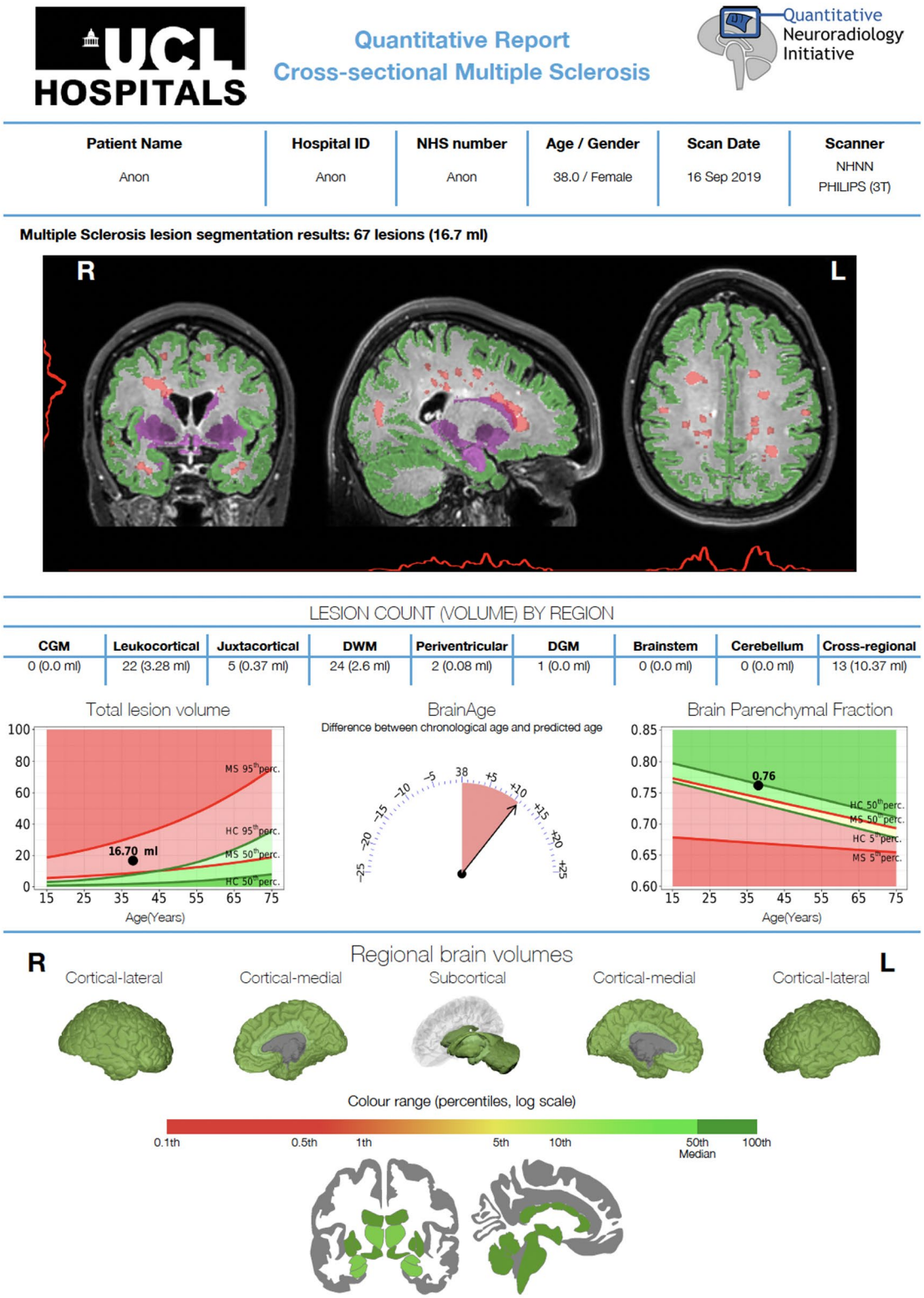

Fig. 2 Quantitative report (QReport) of a 3D T2-FLAIR MRI scan from a 38-year-old female with MS. The QReport includes key demographic information in the header. The total lesion count and total lesion volume (TLV) were 67 and 16.7 mL, respectively. An anatomical slice in axial, sagittal, and coronal planes are overlaid with segmentation masks (red = lesions; green = cortical grey matter (CGM); purple = deep grey matter (DGM)). Regional lesion volume and count are displayed in a table (leukocortical = lesions with the largest proportion in CGM but overlapping deep white matter (DWM); cross-regional = largest proportion in DWM with overlap in GM). TLV and brain parenchymal fraction (BPF) are presented graphically and contextualised with healthy control (HC) and MS reference distributions. Brain age gap estimation (BrainAGE) is visualised using a dial (arrow at + 10 years) with the participant's chronological age (38 years) at the centre. Perc. = percentile

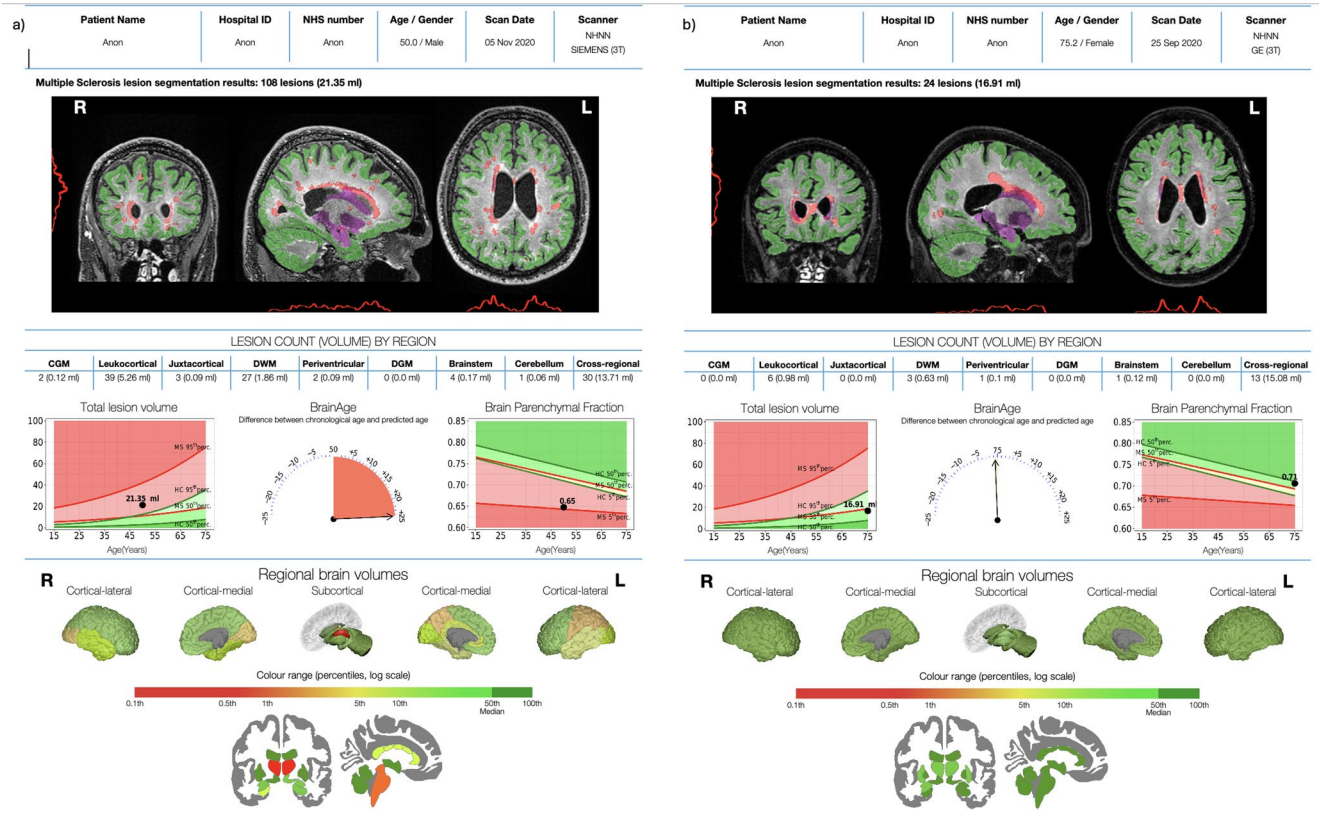

Fig. 3 Quantitative reports (QReports) from (**a**) a 50-year-old male with MS and (**b**) a 75-year-old healthy female participant with similar lesion volumes (total lesion volume (TLV): 21 mL and 17 mL, respectively). In (**a**), TLV is above the 95th percentile of the healthy control (HC) distribution and above the 50th percentile of the MS distribution. In (**b**), TLV falls between the 50th and 95th percentiles of the HC distribution and below the 50th percentile of the MS distribution. Brain parenchymal fraction (BPF) is below the 5th MS percentile in (**a**), and near the 50th HC percentile in (**b**). The dial titled BrainAge depicts predicted age as 25 years greater than chronological age in (**a**), while predicted age is slightly lower than chronological age in (**b**). Regional brain volumes (BV) in (**a**) show lower percentiles (e.g., thalami), colour-coded in red; in (**b**), regional BV fall within higher HC percentiles, coded in green *CGM = cortical grey matter; DGM = deep grey matter; BrainAGE = brain age gap estimation*

Table 3 Clinical, demographic, and imaging characteristics of credibility study testing data

	Healthy controls	MS
Participants, *N* cases	5	17
Field strength, Tesla	1.5/3	1.5/3
Scanner Manufacturer	GE/Philips/Siemens	GE/Philips/Siemens
Age, Mean ± SD (years)	38.3 ± 12.7	45.8 ± 11.1
Age, range (years)	22–63	28–75
Female sex, %	40	80
[a]GCA scale score, N cases (0/1/2/3)	5/0/0/0	6/4/4/3
Lesion burden, N cases (none/minimal/moderate/severe)	5/0/0/0	0/7/6/5
[b]EDSS, Mean ± SD	–	3.7 ± 2.2
Disease duration, Mean ± SD (years)	–	12.1 ± 11.0
Phenotype at scan, N cases ([c]RRMS/[d]PPMS/[e]SPMS)	–	13/2/3

[a]*GCA* = Global Cortical Atrophy, [b]*EDSS* = Expanded Disability Status Scale, [c]*RRMS* = relapsing remitting MS, [d]*PPMS* = primary progressive MS, [e]*SPMS* = secondary progressive MS

between 14 centres and open sources, we assembled demographically diverse HC and clinically heterogeneous MS datasets, acquired on scanners from three manufacturers and two field strengths. Using these data, we built generalisable normative reference curves to contextualise QIB. We conducted a QNI-defined credibility study as the initial clinical evaluation step within the framework, involving multi-centre, independent expert neuroradiologist review of QReport outputs alongside routine visual assessment [20]. Moderate-to-Complete agreement between visual assessment and QReport outputs was observed in 94% of evaluations, with the highest level of agreement in approximately two-thirds. The tool was considered useful in 82% of evaluations.

Despite evidence that MS QReports can improve inter-rater agreement and reduce reporting time, their uptake in routine clinical practice remains limited [5, 21]. A recent systematic review found that MS QReports require both 3D T1 and T2-FLAIR sequences as input [5]. This requirement poses a major barrier to implementation, as current imaging guidelines for clinical routine do not recommend

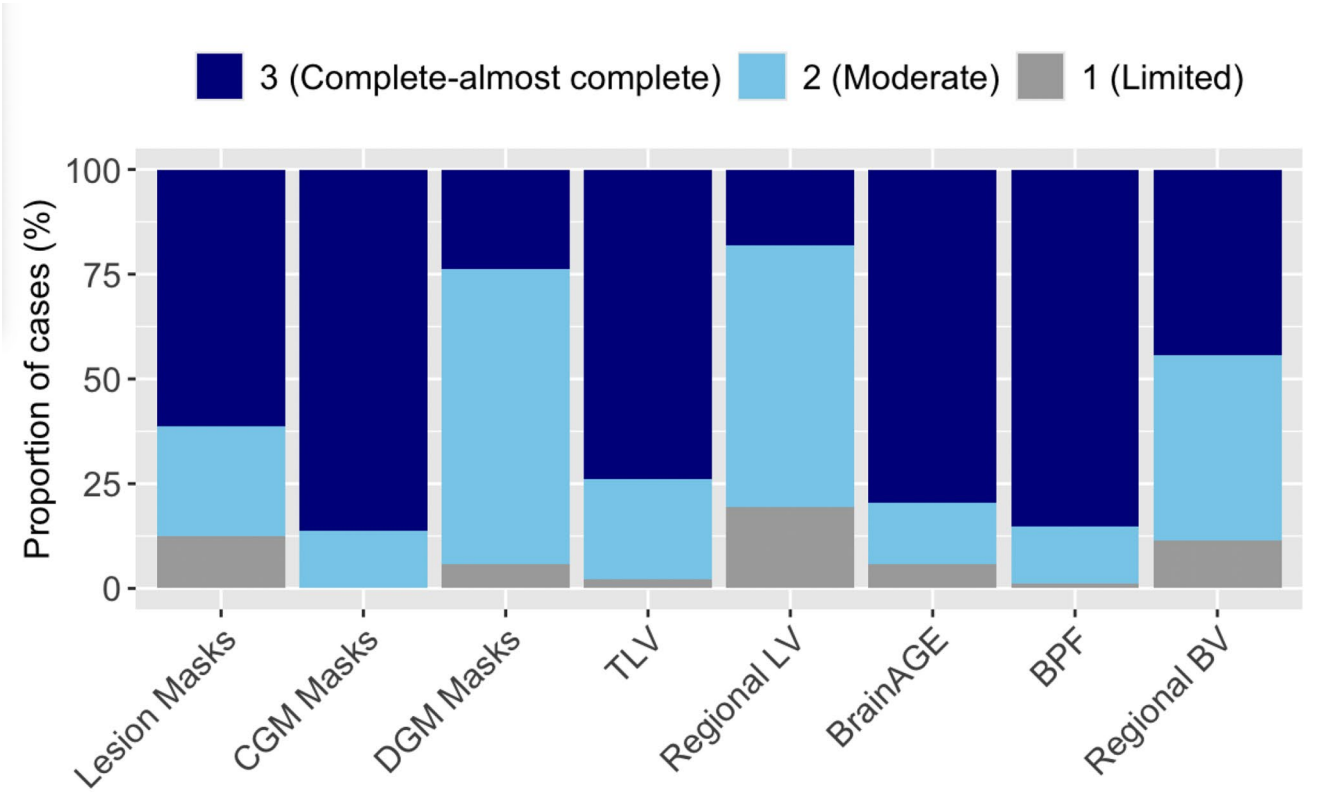

Fig. 4 Bar chart of the proportion of evaluations (%) at each level of agreement between visual assessment and QReport results in the credibility study (1 = Limited; 2 = Moderate; 3 = Complete-almost complete). *CGM = cortical grey matter*, *DGM = deep grey matter*, *TLV = total lesion volume*, *LV = lesion volume*, *BrainAGE = brain age gap estimation*, *BPF = brain parenchymal fraction*, *BV = brain volume*

non-contrast 3D T1-weighted imaging [4, 9, 10]. FLAIR-only input may therefore enhance clinical feasibility of MS QReports [21]. Another limitation of existing tools is the absence of MS-specific reference models [5]. Evidence suggests that MS data can better explain disease severity and predict disease progression than HC data [22]. Furthermore, a recent publication systematically defined clinical user requirements for MS QReports, highlighting a preference for 3D T2-FLAIR-only tools and specifying clinically relevant QIB. The MS QReport presented here aligns technical feasibility with systematically defined user requirements by integrating clinically prioritised, 3D T2-FLAIR-derived QIBs within a structured framework contextualised against large multi-centre MS and HC reference datasets.

Contextualised global markers, including TLV, BPF, and BrainAGE, most frequently achieved Complete-to-almost complete agreement with neuroradiologist visual assessment. We found that global markers were considered useful in 97% of evaluations. To our knowledge BrainAGE has not yet been included on an MS QReport, despite evidence that it correlates with disease severity and progression [48–53]. Therefore, this represents the first evaluation of end-user perception of BrainAGE on an MS QReport.

Lower usefulness and agreement scores for regional quantification, compared with global markers, may reflect known algorithmic limitations of regional measurements that have been reported in MS populations, including systematic over- or underestimation of regional BV relative to reference methods [54, 55]. These effects are mitigated by contextualising outputs against reference datasets processed with the same pipeline, thereby establishing normative bounds across the lifespan and reducing the risk of misinterpretation. In addition, reduced regional BV and LV agreement likely arises from methodological differences between automated segmentation—which generates deterministic 3D regional boundaries—and clinicians' 2D, slice-by-slice visual assessments, underscoring the importance of structured training for QReport interpretation.

This work had several limitations. Studies analysing the participant base of the open sources included in this work indicate that non-Caucasian ethnicities were underrepresented [31]. Future reference models will include data from ADNI-4, which specifically aims to recruit a more ethnically diverse cohort [31, 56]. Clinicians participating in the credibility study had comparable MS experience; however, familiarity with interpreting QReports may have varied.

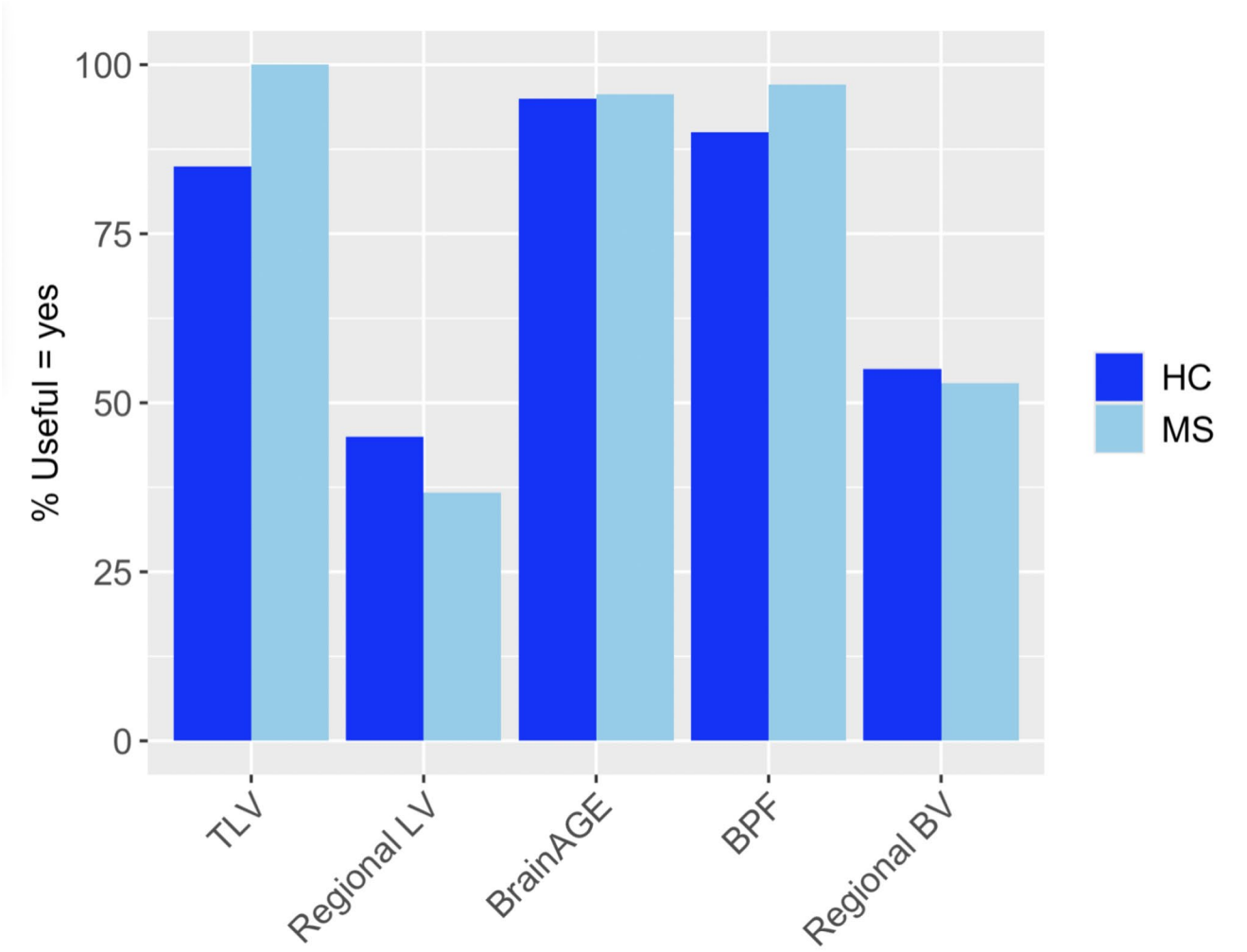

Fig. 5 Bar chart of proportion of evaluations (%) in which QReport components were considered useful, stratified by healthy control (HC) and MS cohorts. *TLV = total lesion volume*, *LV = lesion volume*, *BrainAGE = brain age gap estimation*, *BPF = brain parenchymal fraction*, *BV = brain volume*

Future evaluations will incorporate self-reported QReport experience to account for this variability. Although the MS reference cohort was derived from a single institution, acquisitions spanned six scanners with heterogeneous protocols and field strengths, reflecting real-world variability. Scanner-related effects were harmonised using the same method applied to the multi-centre HC data, thereby mitigating technical bias and supporting the robustness of the reference models. While the credibility study comprised 22 cases, each was independently assessed by four neuroradiologists, yielding 88 independent case-by-rater assessments and repeated evaluations of each QReport component. This multi-rater design enabled proof-of-concept assessment of individual components of the QReport within the credibility phase of the QNI framework.

In accordance with the QNI framework, the credibility study should be followed by a larger accuracy study evaluating impact on inter-rater agreement, reporting accuracy, and efficiency [20]. Another key next step will be the development of a longitudinal report designed to quantify within-subject changes over time, which could support prognostication and monitoring.

To conclude, we developed an open-source tool for quantification of clinically relevant 3D T2-FLAIR biomarkers in MS. The QReport addresses both technical barriers to implementation and systematically defined clinical user requirements by integrating validated FLAIR-based algorithms within a structured, reference-contextualised graphical framework. We provided generalisable MS- and HC-specific reference models and communicated results in a user-friendly format. Finally, in accordance with the QNI clinical validation pathway, we conducted a credibility study, in which independent, multi-centre expert neuroradiologists evaluated QReport outputs against routine visual

assessment of real-world cases, demonstrating high agreement and usefulness.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00234-026-04060-2.

Acknowledgements The authors are members of the MAGNIMS study group (Magnetic ResonanceImaging in Multiple Sclerosis; https://www.magnims.eu/), a collaborative network ofEuropean ce ntres conducting research projects using MRI in multiple sclerosis. Thesteering committee includes Frederik Barkhof (Amsterdam), Nicola de Stefano(Siena), Jaume Sastre–Garriga (Barcelona, Co–Chair), Olga Ciccarelli (London), Christian Enzinger (Graz), Massimo Filippi (Milan), Claudio Gasperini (Rome), Ludwig Kappos (Basel), Jacqueline Palace (Oxford), Hugo Vrenken (Amsterdam), Alex Rovira (Barcelona), Maria Assunta Rocca (Milan, Co–Chair), and Tarek Yousry(London). Z. Mendelsohn, F. Prados, J. Thornton, and F. Barkhof receive funding from National Institute for Health Research (NIHR), Biomedical Research Centre initiative at University College London Hospitals (UCLH).

Author contributions Z.M., F.P., G.P., O.G., M.G.S., T.Y., and J.T., and F.B. contributed to the conception and design of the study, data acquisition, and data analysis. Z.M. conducted data interpretation and wrote the main manuscript. A.R., B.J., W.H., and A.I. contributed to data analysis. All other authors contributed to data acquisition. All authors critically reviewed and approved the final manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data availability The software is open-source and available at: https://niftyweb.cs.ucl.ac.uk/index.php. The publicly accessible datasets are available at: ADNI - https://ida.loni.usc.edu/login.jspEPAD - https://ep-ad.org/index.php/open-source-data/MPI - https://openneuro.org/datasets/ds000221/versions/00002NIFD - https://ida.loni.usc.edu/collaboration/access/appLicense.jsp.

Declarations

Competing interests F. Prados received a Guarantors of Brain fellowship 2017-2020. G. Pontillo received research grants from ECTRIMS-MAGNIMS and ESNR. O. Goodkin is an employee of Bayer Plc. A. Rovira serves/ed on scientific advisory boards for Novartis, Sanofi-Genzyme, SyntheticMR, TensorMedical, Roche, Biogen, and OLEA Medical, and has received speaker honoraria from Bayer, Sanofi- Genzyme, Merck-Serono, Teva Pharmaceutical Industries Ltd, Novartis, Roche, Bristol-Myers and Biogen. Ayodeji Ijishakin is a founder of Mecha Health. M.A. Foster has received speaker honoraria from Merck. J. Palace has received support for scientific meetings and honorariums for advisory work from Merck Serono, Novartis, Chugai, Alexion, Roche, Medimmune, Argenx, Vitaccess, UCB, Mitsubishi, Amplo, Janssen. Grants from Alexion, Argenx, Roche, Medimmune, Amplo biotechnology. Patent ref P37347WO and licence agreement Numares multimarker MS diagnostics Shares in AstraZeneca. Her group has been awarded an ECTRIMS fellowship and a Sumaira Foundation grant to start later this year. A Charcot fellow worked in Oxford 2019-2021. She acknowledges partial funding to the trust by Highly specialised services NHS England. She is on the medical advisory boards of the Sumaira Foundation and MOG project charities, is a member of the Guthy Jackon Foundation Charity and is on the Board of the European Charcot Foundation and the steering committee of MAGNIMS and the UK NHSE IVIG Committee and chairman of the NHSE neuroimmunology patient pathway and ECTRIMS Council member on the educational committee since June 2023. On the ABN advisory groups for MS and neuroinflammation. S. Messina received honoraria for lecturing and advisory board activity from UCB and Biogen, and travel grant from Roche and Merck. R. Cortese received speaker honoraria and travel support from Roche, Merck, Sanofi-Genzyme, Novartis and Janssen. She was awarded a MAGNIMS-ECTRIMS fellowship in 2019. EAH received honoraria for advisory board activity from Sanofi-Genzyme, and his department has received honoraria for lecturing from Biogen and Merck. L.T. Westlye is a shareholder and advisor of baba.vision. M.A. Rocca received consulting fees from Biogen, Bristol Myers Squibb, Eli Lilly, Janssen, Roche; and speaker honoraria from AstraZeneca, Biogen, Bristol Myers Squibb, Bromatech, Celgene, Genzyme, Horizon Therapeutics Italy, Merck Serono SpA, Novartis, Roche, Sanofi and Teva. She receives research support from the MS Society of Canada, the Italian Ministry of Health, the Italian Ministry of University and Research, and Fondazione Italiana Sclerosi Multipla. She is Associate Editor for Multiple Sclerosis and Related Disorders. M. Filippi is Editor-in-Chief of the Journal of Neurology, Associate Editor of Human Brain Mapping, Neurological Sciences, and Radiology; received compensation for consulting services from Alexion, Almirall, Biogen, Merck, Novartis, Roche, Sanofi; speaking activities from Bayer, Biogen, Celgene, Chiesi Italia SpA, Eli Lilly, Genzyme, Janssen, Merck-Serono, Neopharmed Gentili, Novartis, Novo Nordisk, Roche, Sanofi, Takeda, and TEVA; participation in Advisory Boards for Alexion, Biogen, Bristol-Myers Squibb, Merck, Novartis, Roche, Sanofi, Sanofi-Aventis, Sanofi-Genzyme, Takeda; scientific direction of educational events for Biogen, Merck, Roche, Celgene, Bristol- Myers Squibb, Lilly, Novartis, Sanofi-Genzyme; he receives research support from Biogen Idec, Merck-Serono, Novartis, Roche, the Italian Ministry of Health, the Italian Ministry of University and Research, and Fondazione Italiana Sclerosi Multipla. A.Toosy has received speaker honoraria from Merck, Biomedia, Sereno Symposia International Foundation, Bayer and At the Limits and meeting expenses from Merck, Biogen Idec and Novartis He was the UK PI for two clinical trials sponsored by MEDDAY pharmaceutical company (MD1003 in optic neuropathy [MS-ON -NCT02220244] and progressive MS [MS-SPI2 - NCT02936037]). He has been supported by recent grants from the MRC (MR/S026088/1), NIHR BRC (541/CAP/OC/818837) and RoseTrees Trust (A1332 and PGL21/10079). He is an associate editor for Frontiers in Neurology - Neuro-ophthalmology section and on the editorial board for Neurology and Multiple Sclerosis Journal. O. Ciccarelli is an NIHR Research Professor (RP-2017-08-ST2-004); acts as a consultant for Biogen, Merck, Novartis, Roche, and Teva; and has received research grant support from the MS Society of Great Britain and Northern Ireland, the NIHR UCLH Biomedical Research Centre, the Rosetree Trust, the National MS Society, and the NIHR-HTA. H. Pemberton is an employee of GE Healthcare. J.H. Cole is a scientific advisor to and shareholder in BrainKey and Claritas HealthTech PTE. F. Barkhof receives NIHR funding and is a steering committee and iDMC member for Biogen, Merck, Roche, EISAI. Consultant for Roche, Biogen, Merck, IXICO, Jansen, Combinostics. Research agreements with Novartis, Merck, Biogen, GE, Roche. Cofounder and shareholder of Queen Square Analytics LTD. The remaining authors report no competing interests.F. Prados received a Guarantors of Brain fellowship 2017-2020. G. Pontillo received research grants from ECTRIMS-MAGNIMS and ESNR. O. Goodkin is an employee of Bayer Plc. A. Rovira serves/ed on scientific advisory boards for Novartis, Sanofi-Genzyme, SyntheticMR, TensorMedical, Roche, Biogen, and OLEA Medical, and has received speaker honoraria from Bayer, Sanofi- Genzyme, Merck-Serono, Teva Pharmaceutical Industries Ltd, Novartis, Roche, Bristol-Myers and Biogen. Ayodeji Ijishakin is a founder of Mecha Health. M.A. Foster has received speaker honoraria from Merck. J. Palace has received support for scientific meetings and honorariums for advisory work from Merck Serono, No-

vartis, Chugai, Alexion, Roche, Medimmune, Argenx, Vitaccess, UCB, Mitsubishi, Amplo, Janssen. Grants from Alexion, Argenx, Roche, Medimmune, Amplo biotechnology. Patent ref P37347WO and licence agreement Numares multimarker MS diagnostics Shares in AstraZeneca. Her group has been awarded an ECTRIMS fellowship and a Sumaira Foundation grant to start later this year. A Charcot fellow worked in Oxford 2019-2021. She acknowledges partial funding to the trust by Highly specialised services NHS England. She is on the medical advisory boards of the Sumaira Foundation and MOG project charities, is a member of the Guthy Jackon Foundation Charity and is on the Board of the European Charcot Foundation and the steering committee of MAGNIMS and the UK NHSE IVIG Committee and chairman of the NHSE neuroimmunology patient pathway and ECTRIMS Council member on the educational committee since June 2023. On the ABN advisory groups for MS and neuroinflammation. S. Messina received honoraria for lecturing and advisory board activity from UCB and Biogen, and travel grant from Roche and Merck. R. Cortese received speaker honoraria and travel support from Roche, Merck, Sanofi-Genzyme, Novartis and Janssen. She was awarded a MAGNIMS-ECTRIMS fellowship in 2019. EAH received honoraria for advisory board activity from Sanofi-Genzyme, and his department has received honoraria for lecturing from Biogen and Merck. L.T. Westlye is a shareholder and advisor of baba.vision. M.A. Rocca received consulting fees from Biogen, Bristol Myers Squibb, Eli Lilly, Janssen, Roche; and speaker honoraria from AstraZeneca, Biogen, Bristol Myers Squibb, Bromatech, Celgene, Genzyme, Horizon Therapeutics Italy, Merck Serono SpA, Novartis, Roche, Sanofi and Teva. She receives research support from the MS Society of Canada, the Italian Ministry of Health, the Italian Ministry of University and Research, and Fondazione Italiana Sclerosi Multipla. She is Associate Editor for Multiple Sclerosis and Related Disorders. M. Filippi is Editor-in-Chief of the Journal of Neurology, Associate Editor of Human Brain Mapping, Neurological Sciences, and Radiology; received compensation for consulting services from Alexion, Almirall, Biogen, Merck, Novartis, Roche, Sanofi; speaking activities from Bayer, Biogen, Celgene, Chiesi Italia SpA, Eli Lilly, Genzyme, Janssen, Merck-Serono, Neopharmed Gentili, Novartis, Novo Nordisk, Roche, Sanofi, Takeda, and TEVA; participation in Advisory Boards for Alexion, Biogen, Bristol-Myers Squibb, Merck, Novartis, Roche, Sanofi, Sanofi-Aventis, Sanofi-Genzyme, Takeda; scientific direction of educational events for Biogen, Merck, Roche, Celgene, Bristol- Myers Squibb, Lilly, Novartis, Sanofi-Genzyme; he receives research support from Biogen Idec, Merck-Serono, Novartis, Roche, the Italian Ministry of Health, the Italian Ministry of University and Research, and Fondazione Italiana Sclerosi Multipla. A.Toosy has received speaker honoraria from Merck, Biomedia, Sereno Symposia International Foundation, Bayer and At the Limits and meeting expenses from Merck, Biogen Idec and Novartis He was the UK PI for two clinical trials sponsored by MEDDAY pharmaceutical company (MD1003 in optic neuropathy [MS-ON -NCT02220244] and progressive MS [MS-SPI2 - NCT02936037]). He has been supported by recent grants from the MRC (MR/S026088/1), NIHR BRC (541/CAP/OC/818837) and RoseTrees Trust (A1332 and PGL21/10079). He is an associate editor for Frontiers in Neurology - Neuro-ophthalmology section and on the editorial board for Neurology and Multiple Sclerosis Journal. O. Ciccarelli is an NIHR Research Professor (RP-2017-08-ST2-004); acts as a consultant for Biogen, Merck, Novartis, Roche, and Teva; and has received research grant support from the MS Society of Great Britain and Northern Ireland, the NIHR UCLH Biomedical Research Centre, the Rosetree Trust, the National MS Society, and the NIHR-HTA. H. Pemberton is an employee of GE Healthcare. J.H. Cole is a scientific advisor to and shareholder in BrainKey and Claritas HealthTech PTE. F. Barkhof receives NIHR funding and is a steering committee and iDMC member for Biogen, Merck, Roche, EISAI. Consultant for Roche, Biogen, Merck, IXICO, Jansen, Combinostics. Research agreements with Novartis, Merck, Biogen, GE, Roche. Co-founder and shareholder of Queen Square Analytics LTD. The remaining authors report no competing interests.

References

1. Smith SM, Zhang Y, Jenkinson M et al (2002) Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. Neuroimage 17:479–489. https://doi.org/10.1006/nimg.2002.1040
2. Smith SM, De Stefano N, Jenkinson M, Matthews PM (2001) Normalized accurate measurement of longitudinal brain change. J Comput Assist Tomogr 25:466–475. https://doi.org/10.1097/00004728-200105000-00022
3. De Rosa AP, Benedetto M, Tagliaferri S et al (2024) Consensus of algorithms for lesion segmentation in brain MRI studies of multiple sclerosis. Sci Rep 14:21348. https://doi.org/10.1038/s41598-024-72649-9
4. Sastre-Garriga J, Pareto D, Battaglini M et al (2020) MAGNIMS consensus recommendations on the use of brain and spinal cord atrophy measures in clinical practice. Nat Rev Neurol 16:171–182. https://doi.org/10.1038/s41582-020-0314-x
5. Mendelsohn Z, Pemberton HG, Gray J et al (2023) Commercial volumetric MRI reporting tools in multiple sclerosis: a systematic review of the evidence. Neuroradiology 65:5–24. https://doi.org/10.1007/s00234-022-03074-w
6. Brune S, Høgestøl EA, Cengija V et al (2020) LesionQuant for assessment of MRI in Multiple Sclerosis—A promising supplement to the visual scan inspection. Front Neurol. https://doi.org/10.3389/fneur.2020.546744
7. Krüger J, Opfer R, Gessert N et al (2020) Fully automated longitudinal segmentation of new or enlarged multiple sclerosis lesions using 3D convolutional neural networks. Neuroimage Clin 28:102445. https://doi.org/10.1016/j.nicl.2020.102445
8. Van Hecke W, Costers L, Descamps A et al (2021) A novel digital care management platform to monitor clinical and subclinical disease activity in multiple sclerosis. Brain Sci 11:1171. https://doi.org/10.3390/brainsci11091171
9. Wattjes MP, Ciccarelli O, Reich DS et al (2021) 2021 MAGNIMS–CMSC–NAIMS consensus recommendations on the use of MRI in patients with multiple sclerosis. Lancet Neurol 20:653–670. https://doi.org/10.1016/S1474-4422(21)00095-8
10. Barkhof F, Reich DS, Oh J et al (2025) 2024 MAGNIMS–CMSC–NAIMS consensus recommendations on the use of MRI for the diagnosis of multiple sclerosis. Lancet Neurol 24:866–879. https://doi.org/10.1016/S1474-4422(25)00304-7

11. Cardoso MJ, Modat M, Wolz R et al (2015) Geodesic information flows: spatially-variant graphs and their application to segmentation and fusion. IEEE Trans Med Imaging 34:1976–1988. https://doi.org/10.1109/TMI.2015.2418298
12. Sergi Valverde (2022) nicMSlesions. In: https://github.com/sergivalverde/nicMSlesions/
13. Valverde S, Salem M, Cabezas M et al (2019) One-shot domain adaptation in multiple sclerosis lesion segmentation using convolutional neural networks. Neuroimage Clin 21:101638. https://doi.org/10.1016/j.nicl.2018.101638
14. Valverde S, Cabezas M, Roura E et al (2017) Improving automated multiple sclerosis lesion segmentation with a cascaded 3D convolutional neural network approach. Neuroimage 155:159–168. https://doi.org/10.1016/j.neuroimage.2017.04.034
15. (2025) Online brain extraction and parcelation tool based on GIF (FIFO Queue). In: http://niftyweb.cs.ucl.ac.ukprogram.php/?p=GIF
16. Goodkin O, Prados F, Vos SB et al (2021) FLAIR-only joint volumetric analysis of brain lesions and atrophy in clinically isolated syndrome (CIS) suggestive of multiple sclerosis. Neuroimage Clin 29:102542. https://doi.org/10.1016/j.nicl.2020.102542
17. Sudre CH, Gomez Anson B, Davagnanam I et al (2018) Bullseye's representation of cerebral white matter hyperintensities. J Neuroradiol 45:114–122. https://doi.org/10.1016/j.neurad.2017.10.001
18. Jordan Colman (2025) FLAIR Brain Age. In: https://github.com/jordan-colman/FLAIR-Brain-Age
19. Jordan Colman O, Goodkin M, Foster et al (2022) Adapting and applying the brain age paradigm for clinical imaging in multiple sclerosis (MS). ISMRM #1553
20. Goodkin O, Pemberton H, Vos SB et al (2019) The quantitative neuroradiology initiative framework: application to dementia. Br J Radiol. https://doi.org/10.1259/bjr.20190365
21. van Nederpelt DR, Mendelsohn ZC, Bos L et al (2025) User requirements for quantitative radiological reports in multiple sclerosis. Eur Radiol. https://doi.org/10.1007/s00330-025-11544-x
22. van Nederpelt DR, Bos L, Mattiesing RM et al (2025) Multiple Sclerosis–specific reference curves for brain volumes to explain disease severity. Neurology. https://doi.org/10.1212/WNL.0000000000213618
23. Brown JWL, Prados Carrasco F, Eshaghi A et al (2020) Periventricular magnetisation transfer ratio abnormalities in multiple sclerosis improve after alemtuzumab. Mult Scler J 26:1093–1101. https://doi.org/10.1177/1352458519852093
24. Haider L, Prados F, Chung K et al (2021) Cortical involvement determines impairment 30 years after a clinically isolated syndrome. Brain 144:1384–1395. https://doi.org/10.1093/brain/awab033
25. Pardini M, Sudre CH, Prados F et al (2016) Relationship of grey and white matter abnormalities with distance from the surface of the brain in multiple sclerosis. J Neurol Neurosurg Psychiatry 87:1212–1217. https://doi.org/10.1136/jnnp-2016-313979
26. Williams T, John N, Calvi A et al (2024) Investigating the relationship between thalamic iron concentration and disease severity in secondary progressive multiple sclerosis using quantitative susceptibility mapping: cross-sectional analysis from the MS-STAT2 randomised controlled trial. Neuroimage Rep 4:100216. https://doi.org/10.1016/j.ynirp.2024.100216
27. Williams T, John N, Calvi A et al (2023) Cardiovascular risk factors in secondary progressive multiple sclerosis: A cross-sectional analysis from the <scp>MS-STAT2 randomized controlled trial. Eur J Neurol 30:2769–2780. https://doi.org/10.1111/ene.15924
28. Boeve B, Bove J, Brannelly P et al (2020) The longitudinal evaluation of familial frontotemporal dementia subjects protocol: framework and methodology. Alzheimers Dement 16:22–36. https://doi.org/10.1016/j.jalz.2019.06.4947
29. Mendes N, Oligschläger S, Lauckner ME et al (2019) A functional connectome phenotyping dataset including cognitive state and personality measures. Sci Data 6:180307. https://doi.org/10.1038/sdata.2018.307
30. Babayan A, Erbey M, Kumral D et al (2019) A mind-brain-body dataset of MRI, EEG, cognition, emotion, and peripheral physiology in young and old adults. Sci Data 6:180308. https://doi.org/10.1038/sdata.2018.308
31. Weber CJ, Carrillo MC, Jagust W et al (2021) The worldwide Alzheimer's disease neuroimaging initiative: ADNI-3 updates and global perspectives. Alzheimers Dement Transl Res Clin Interv. https://doi.org/10.1002/trc2.12226
32. Solomon A, Kivipelto M, Molinuevo JL et al (2018) European prevention of Alzheimer's dementia longitudinal cohort study (EPAD LCS): study protocol. BMJ Open 8:e021017. https://doi.org/10.1136/bmjopen-2017-021017
33. Lorenzini L, Ingala S, Wink AM et al (2022) The Open-Access European Prevention of Alzheimer's Dementia (EPAD) MRI dataset and processing workflow. Neuroimage Clin 35:103106. https://doi.org/10.1016/j.nicl.2022.103106
34. Polman CH, Reingold SC, Banwell B et al (2011) Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Ann Neurol 69:292–302. https://doi.org/10.1002/ana.22366
35. Ferran Prados MJCNBCAW-KSO (2016) NiftyWeb: web based platform for image processing on the cloud. International Society for Magnetic Resonance in Medicine (ISMRM) 24th Scientific Meeting and Exhibition - Singapore 2016. In: International Society for Magnetic Resonance in Medicine (ISMRM) 24th Scientific Meeting and Exhibition. Singapore
36. Stutters J (2023) https://github.com/jstutters/nicpython36. In: https://github.com/jstutters/nicpython36
37. Tustison NJ, Avants BB, Cook PA et al (2010) N4ITK: Improved N3 Bias Correction. IEEE Trans Med Imaging 29:1310–1320. https://doi.org/10.1109/TMI.2010.2046908
38. Prados F, Cardoso MJ, Kanber B et al (2016) A multi-time-point modality-agnostic patch-based method for lesion filling in Multiple Sclerosis. Neuroimage 139:376–384. https://doi.org/10.1016/j.neuroimage.2016.06.053
39. Klein A, Tourville J (2012) 101 Labeled Brain Images and a Consistent Human Cortical Labeling Protocol. Front Neurosci 6. https://doi.org/10.3389/fnins.2012.00171
40. Jean-Philippe Fortin (2021) ComBat harmonization in R. In: https://github.com/Jfortin1/neuroCombat_Rpackage
41. Fortin J-P, Cullen N, Sheline YI et al (2018) Harmonization of cortical thickness measurements across scanners and sites. Neuroimage 167:104–120. https://doi.org/10.1016/j.neuroimage.2017.11.024
42. Roger Koenker (2025) quantreg: Quantile Regression. In: https://cran.r-project.org/web/packages/quantreg/index.html
43. Koenker R, Bassett G (1978) Regression Quantiles. Econometrica 46:33. https://doi.org/10.2307/1913643
44. Pasquier F, Leys D, Weerts JGE et al (1996) Inter-and intraobserver reproducibility of cerebral atrophy assessment on MRI scans with hemispheric infarcts. Eur Neurol 36:268–272. https://doi.org/10.1159/000117270
45. Kilem L, Gwet (2014) irrCAC: Computing Chance-Corrected Agreement Coefficients (CAC). In: https://CRAN.R-project.org/package=irrCAC
46. Gwet KL (2008) Computing inter-rater reliability and its variance in the presence of high agreement. Br J Math Stat Psychol 61:29–48. https://doi.org/10.1348/000711006X126600
47. Razvan V, Marinescu (2022) BrainPainter - Brain colouring software. In: https://github.com/razvanmarinescu/brain–coloring
48. Cole JH, Ritchie SJ, Bastin ME et al (2018) Brain age predicts mortality. Mol Psychiatry 23:1385–1392. https://doi.org/10.1038/mp.2017.62

49. Cole JH, Raffel J, Friede T et al (2020) Longitudinal assessment of Multiple Sclerosis with the Brain-Age Paradigm. Ann Neurol 88:93–105. https://doi.org/10.1002/ana.25746
50. Høgestøl EA, Kaufmann T, Nygaard GO et al (2019) Cross-sectional and longitudinal MRI brain scans reveal accelerated brain aging in Multiple Sclerosis. Front Neurol. https://doi.org/10.3389/fneur.2019.00450
51. Kaufmann T, van der Meer D, Doan NT et al (2019) Common brain disorders are associated with heritable patterns of apparent aging of the brain. Nat Neurosci 22:1617–1623. https://doi.org/10.1038/s41593-019-0471-7
52. Skattebøl L, Nygaard GO, Leonardsen EH et al (2025) Brain age in multiple sclerosis: A study with deep learning and traditional machine learning. Brain Commun. https://doi.org/10.1093/braincomms/fcaf152
53. Brier MR, Li Z, Ly M et al (2023) Brain age" predicts disability accumulation in multiple sclerosis. Ann Clin Transl Neurol 10:990–1001. https://doi.org/10.1002/acn3.51782
54. Huizinga W, Poot DHJ, Vinke EJ et al (2021) Differences between MR brain region segmentation methods: Impact on single-subject analysis. Front Big Data. https://doi.org/10.3389/fdata.2021.577164
55. Weeda MM, Brouwer I, de Vos ML et al (2019) Comparing lesion segmentation methods in multiple sclerosis: input from one manually delineated subject is sufficient for accurate lesion segmentation. Neuroimage Clin 24:102074. https://doi.org/10.1016/j.nicl.2019.102074
56. Rivera Mindt M, Arentoft A, Calcetas AT et al (2024) The Alzheimer's Disease Neuroimaging Initiative-4 (ADNI-4) engagement core: a culturally informed, community-engaged research (CICER) model to advance brain health equity. Alzheimers Dement 20:8279–8293. https://doi.org/10.1002/alz.14242

Authors and Affiliations

Zoe Mendelsohn[1,2] · Ferran Prados[2,3] · Giuseppe Pontillo[2,4] · Olivia Goodkin[2,5] · Bas Jasperse[4] · Alex Rovira[6] · Matthew Grech-Sollars[7,8] · Jordan Colman[2,9] · James Moggridge[2,7] · Weaam Hamed[2,7] · Ayodeji Ijishakin[10] · Michael A. Foster[2] · Nima Mahmoudi[1] · Mike P. Wattjes[1] · Arturo Brunetti[11] · Gabriel Gonzalez-Escamilla[12] · Sergiu Groppa[12] · Einar August Høgestøl[13,14] · Lars T. Westlye[14,15] · Jacqueline Palace[16] · Silvia Messina[16] · Rosa Cortese[17] · Nicola De Stefano[17] · Jaume Sastre-Garriga[6] · Stefan Ropele[18] · Christian Enzinger[18] · Maria A. Rocca[19] · Massimo Filippi[20] · Jeremy Chataway[2,21] · Charles Wade[2] · Ahmed Toosy[2] · Olga Ciccarelli[2,21] · Hugh Pemberton[2] · James H. Cole[10] · Michael Scheel[1] · Jawed Nawabi[1] · Tarek Yousry[2,7] · John Thornton[2,7] · Frederik Barkhof[2,4]

✉ Zoe Mendelsohn
z.mendelsohn@ucl.ac.uk

Ferran Prados
f.prados@ucl.ac.uk

Giuseppe Pontillo
g.pontillo@ucl.ac.uk

Olivia Goodkin
o.goodkin@ucl.ac.uk

Bas Jasperse
m.m.s.jasperse@amsterdamumc.nl

Alex Rovira
alex.rovira.idi@gencat.cat

Matthew Grech-Sollars
m.grech-sollars@ucl.ac.uk

Jordan Colman
j.colman@nhs.net

James Moggridge
j.moggridge@nhs.net

Weaam Hamed
w.hamed@ucl.ac.uk

Ayodeji Ijishakin
a.ijishakin@ucl.ac.uk

Michael A. Foster
m.foster@ucl.ac.uk

Nima Mahmoudi
nima.mahmoudi@charite.de

Mike P. Wattjes
mike.wattjes@charite.de

Arturo Brunetti
arturo.brunetti@ibb.cnr.it

Gabriel Gonzalez-Escamilla
ggonzale@uni-mainz.de

Sergiu Groppa
segroppa@uni-mainz.de

Einar August Høgestøl
e.a.hogestol@medisin.uio.no

Lars T. Westlye
l.t.westlye@psykologi.uio.no

Jacqueline Palace
jacqueline.palace@ndcn.ox.ac.uk

Silvia Messina
silvia.messina@ndcn.ox.ac.uk

Rosa Cortese
rosa.cortese@unisi.it

Nicola De Stefano
nicola.destefano@unisi.it

Jaume Sastre-Garriga
jsastre-garriga@cem-cat.org

Stefan Ropele
stefan.ropele@medunigraz.at

Christian Enzinger
chris.enzinger@medunigraz.at

Maria A. Rocca
rocca.mara@hsr.it

Massimo Filippi
filippi.massimo@hsr.it

Jeremy Chataway
j.chataway@ucl.ac.uk

Charles Wade
c.wade@ucl.ac.uk

Ahmed Toosy
a.toosy@ucl.ac.uk

Olga Ciccarelli
o.ciccarelli@ucl.ac.uk

Hugh Pemberton
h.pemberton@ucl.ac.uk

James H. Cole
j.cole@ucl.ac.uk

Michael Scheel
michael.scheel@charite.de

Jawed Nawabi
jawed.nawabi@charite.de

Tarek Yousry
t.yousry@ucl.ac.uk

John Thornton
j.thornton@ucl.ac.uk

Frederik Barkhof
f.barkhof@ucl.ac.uk

[1] Institute of Neuroradiology, Charité - University Medicine Berlin, Berlin, Germany

[2] Queen Square Institute of Neurology, University College London, London, United Kingdom

[3] E–Health Centre, Open University of Catalonia, Barcelona, Spain

[4] MS Center Amsterdam, Radiology and Nuclear Medicine, Vrije Universiteit Amsterdam, Amsterdam, Netherlands

[5] Bayer (United Kingdom), Newbury, United Kingdom

[6] Section of Neuroradiology. Department of Radiology, Vall d'Hebron Hospital Universitari, Barcelona, Spain

[7] Lysholm Department of Neuroradiology, National Hospital for Neurology and Neurosurgery, London, United Kingdom

[8] Department of Computer Science, University College London, London, United Kingdom

[9] Department of Neuroradiology, King's College Hospital, London, United Kingdom

[10] Hawkes Institute, University College London, London, United Kingdom

[11] Department of Advanced Biomedical Sciences, University of Naples Federico II, Naples, Italy

[12] Movement Disorders, Neurostimulation and Neuroimaging, Johannes Gutenberg University Mainz, Mainz, Germany

[13] Department of Neurology, Oslo University Hospital, Oslo, Norway

[14] Department of Psychology, University of Oslo, Oslo, Norway

[15] Section for Precision Psychiatry, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway

[16] Nuffield Department of Clinical Neurosciences, Medical Sciences Division, University of Oxford, Oxford, United Kingdom

[17] Department of Medicine, Surgery and Neuroscience, University of Siena, Siena, Italy

[18] Department of Neurology, Medical University of Graz, Graz, Austria

[19] Neuroimaging Research Unit, Division of Neuroscience, Vita-Salute San Raffaele University, Milan, Italy

[20] Neurology Unit, Neurorehabilitation Unit, Neurophysiology Service, and Neuroimaging Research Unit, Division of Neuroscience, Vita-Salute San Raffaele University, Milan, Italy

[21] National Institute for Health Research, Biomedical Research Centre, University College London Hospitals NHS Foundation Trust, London, United Kingdom