Journal of Neurology (2019) 266:2848–2858
https://doi.org/10.1007/s00415-019-09509-4

ORIGINAL COMMUNICATION

MRI quality control for the Italian Neuroimaging Network Initiative: moving towards big data in multiple sclerosis

Loredana Storelli[1,2] · Maria A. Rocca[1,3] · Patrizia Pantano[4,5] · Elisabetta Pagani[1] · Nicola De Stefano[6] · Gioacchino Tedeschi[7] · Paola Zaratin[8] · Massimo Filippi[1,2,3] · For the INNI Network

Received: 26 June 2019 / Revised: 12 August 2019 / Accepted: 13 August 2019 / Published online: 17 August 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

The Italian Neuroimaging Network Initiative (INNI) supports the creation of a repository, where MRI, clinical, and neuropsychological data from multiple sclerosis (MS) patients and healthy controls are collected from Italian Research Centers with internationally recognized expertise in MRI applied to MS. However, multicenter MRI data integration needs standardization and quality control (QC). This study aimed to implement quantitative measures for characterizing the standardization and quality of MRI collected within INNI. MRI scans of 423 MS patients, including 3D T_1- and T_2-weighted, were obtained from INNI repository (from Centers A, B, C, and D). QC measures were implemented to characterize: (1) head positioning relative to the magnet isocenter; (2) intensity inhomogeneity; (3) relative image contrast between brain tissues; and (4) image artefacts. Centers A and D showed the most accurate subject positioning within the MR scanner (median z-offsets = -2.6 ± 1.7 cm and -1.1 ± 2 cm). A low, but significantly different, intensity inhomogeneity on 3D T_1-weighted MRI was found between all centers ($p < 0.05$), except for Centers A and C that showed comparable image bias fields. Center D showed the highest relative contrast between gray and normal appearing white matter (NAWM) on 3D T_1-weighed MRI (0.63 ± 0.04), while Center B showed the highest relative contrast between NAWM and MS lesions on FLAIR (0.21 ± 0.06). Image artefacts were mainly due to brain movement (60%) and ghosting (35%). The implemented QC procedure ensured systematic data quality assessment within INNI, thus making available a huge amount of high-quality MRI to better investigate pathophysiological substrates and validate novel MRI biomarkers in MS.

Keywords Magnetic resonance imaging (MRI) · Multiple sclerosis (MS) · Italian Neuroimaging Network Initiative (INNI) · Big data

✉ Massimo Filippi
filippi.massimo@hsr.it

1 Neuroimaging Research Unit, Institute of Experimental Neurology, Division of Neuroscience, IRCCS San Raffaele Scientific Institute, Via Olgettina, 60, 20132 Milan, Italy

2 Vita-Salute San Raffaele University, Milan, Italy

3 Neurology Unit, IRCCS San Raffaele Scientific Institute, Milan, Italy

4 Department of Human Neurosciences, Sapienza University of Rome, Rome, Italy

5 IRCCS NEUROMED, Pozzilli, Italy

6 Department of Medicine, Surgery and Neuroscience, University of Siena, Siena, Italy

7 Department of Medical, Surgical, Neurological Metabolic and Aging Sciences and MRI-Center "SUN-FISM", University of Campania "Luigi Vanvitelli", Naples, Italy

8 Italian Multiple Sclerosis Foundation, Genoa, Italy

Introduction

Multiple sclerosis (MS) is a chronic inflammatory demyelinating and neurodegenerative disease affecting the central nervous system, which is characterized by heterogeneous clinical manifestations, course, and progression of disability over time [12]. The diagnosis of MS is based on clinical and paraclinical data, among which magnetic resonance imaging (MRI) has gained a fundamental role, thanks to its sensitivity in detecting MS-related focal abnormalities [15, 24, 39]. In addition to its role in the diagnostic workup, advanced MRI techniques (such as brain atrophy quantification, diffusion-weighted imaging and functional MRI methods) are improving the understanding of the mechanisms underlying MS heterogeneity [13, 31, 32]. However, small patient sample sizes, the inclusion of patients with specific disease clinical phenotypes and the lack of longitudinal assessments are

among the major limitations that have significantly affected the robustness and reproducibility of the findings obtained so far in the field [50]. Studies on larger cohorts of patients, including the main clinical phenotypes, are likely to improve the analysis of associations between MRI and clinical measures. Of note, during the past few years, huge amounts of MRI data from MS patients have been produced worldwide, especially in the context of research studies [22, 43, 48]. Due to the logistical demands and cost associated with MRI scanning and subject enrollment, and the effort required by the participants themselves, pooling neuroimaging data is a rewarding strategy for performing studies on larger populations [28, 48]. Such an approach is expected to provide superior statistical power and could reduce susceptibility to spurious effects [28]. During the last decade, major advances have been made in sharing neuroimaging data, giving access to thousands of MRI scans on the web [11, 22]. Several initiatives have been proposed with different aims for data sharing: some resources provide already-processed data, such as statistical maps [7, 19] or results from neuroimaging meta-analyses [29], while other contain raw data from individual subjects [25, 42, 49]. For instance, the Alzheimer's disease neuroimaging initiative (ADNI) and the UK biobank were launched to prospectively collect standardized MRI data from patients and healthy subjects [26, 38].

The Italian Neuroimaging Network Initiative (INNI) supports the creation of a repository, where MRI, clinical, and neuropsychological data from MS patients and healthy controls are collected from Italian research centers, supported by internationally recognized expertise in MS field, with the main goals of finding novel MRI biomarkers and improving the application of MRI for investigating MS [16]. Furthermore, the INNI will promote the standardized use of MRI in MS at a national level, supporting projects that integrate MRI across different centers. Specifically, the INNI started from four Italian research centers with recognized expertise in the study of MS that were pioneering in recommending guidelines for MRI acquisition protocols for the diagnosis and monitoring of MS at a national level [14]. The MRI data collected within the INNI included high-resolution T_1-weighted, T_2-weighted, diffusion-weighted, and resting-state fMRI acquisitions. However, the collection of multicenter MRI data into large-scale repositories raises the important issues of standardization and quality control (QC) [2]. Many medical image analysis methods, such as segmentation and registration, are highly sensitive to the presence of image artefacts and to spurious variations of image intensities [6, 45]. For example, large motion artefacts have been shown to affect gray-matter (GM) segmentation and parcellation techniques [30, 40]. Moreover, the heterogeneity of the MR scanners and/or acquisition protocols affects the tissue contrast in the images which, in turn, influences consistency in brain atrophy measurements [17, 31]. Thus, a high level of standardization among MRI data collected from different centers is important for obtaining reliable and optimal brain volume measures [51].

To address these issues, this study aimed at proposing and implementing quantitative measures that are able to characterize the quality and the standardization level of the large amounts of brain MRI data collected retrospectively within the INNI network. The measures proposed (described in "Methods") will also be systematically applied to all MRI uploaded to the INNI repository, to maintain a quality-assessed database. From this perspective, the INNI should ultimately promote the standard use of MRI for the study of MS at both multicenter national and international levels.

Materials and methods

The INNI project is promoted by the Neuroimaging Study Group of the Italian Society of Neurology and is financially supported by a research Grant from the Fondazione Italiana Sclerosi Multipla (FISM). FISM is the owner of the database, according to Italian copyright law. The INNI currently involves four MS centers in Italy (Milan, Neuroimaging Research Unit, San Raffaele Scientific Institute; Rome, Department of Human Neuroscience, Sapienza University; Naples, Department of Neurological Sciences, Second University of Naples/Neurological Institute for Diagnosis and Care ''Hermitage Capodimonte''; Siena, Department of Medicine, Surgery and Neuroscience, University of Siena).

Data set

For the current analysis, MRI scans of 423 MS patients with the main clinical phenotypes (relapsing–remitting [RR], primary progressive [SP], secondary progressive [SP], and clinically isolated syndrome [CIS]), including isotropic 3D T_1- and T_2-weighted scans from the four participating centers (labeled here A, B, C, and D in no particular order) were retrieved from the INNI online repository (https://database.inni-ms.org). FLAIR images were available only for Center B. All MRI scans were acquired using 3.0 T scanners: Intera and Achieva, for Centers A and D, respectively (Philips Medical Systems, Best, The Netherlands); Signa HDxt, for Center B (GE Healthcare, Milwaukee, USA); Magnetom Verio, for Center C (Siemens, Erlangen, Germany). Demographic information for this dataset and MRI acquisition parameters for each center is summarized in Tables 1 and 2.

Patients' positioning was performed according to the usual internal procedures for each center. MR acquisitions from Centers A and B were performed using 8-channel head coils, while Centers C and D used 12-channel and 32-channel head coils, respectively.

Table 1 Demographic and clinical data from patients included in the current analysis, collected from the INNI database

	Patients $n=423$
Age, years (SD)	39.53 (10.94)
Females/males	284/139
Clinical Phenotype: RRMS/PPMS/SPMS/CIS	321/24/63/15
Disease duration, years (SD)	20.33 (6.6)
EDSS, median (IQR)	2.0 (1.5, 4.0)

Data are presented as mean (SD), median (IQR), or absolute number when not specified

MS multiple sclerosis, *SD* standard deviation, *RR* relapsing–remitting, *PP* primary progressive, *SP* secondary progressive, *CIS* clinically isolated syndrome, *EDSS* expanded disability status scale, *IQR* interquartile range

Ethics committee approval

Ethical approval was received from the local ethical standards committee of each participating center, and written informed consent was obtained from all participants at the time of data acquisition.

Methods

A QC procedure was implemented to characterize and monitor the INNI database. We introduced several quantitative and qualitative measures to inspect the quality of MRI data uploaded to the INNI repository for future multicenter MS studies. The factors included in the QC procedure were: (1) subject head positioning within the MR scanner; (2) the degree of intensity inhomogeneity on 3D T_1-weighted images; (3) the relative image contrast between brain tissues; and (4) the presence of image artefacts.

1. The metric for head positioning quality was estimated on the 3D T_1-weighted scans as the distance between the magnet isocenter and the centroid of the subject's brain, quantified as the z-offset along the axis of the scanner bed and also the Euclidean distance (ED), as

$$ED = \sqrt{(x_I - x_C)^2 + (y_I - y_C)^2 + (z_I - z_C)^2}, \quad (1)$$

where x_I, y_I, z_I are the coordinates of the magnet isocenter and x_C, y_C, z_C are the coordinates of the centroid of the subject's brain, estimated after brain tissue extraction on 3D T_1-weighted scans. The coordinates of the isocenter were obtained from the image header for each patient. The z-offset can be positive or negative, and is indicative of systematic positioning errors, whereas the ED is indicative of the absolute magnitude of the positioning error.

2. The intensity inhomogeneity (bias field) in MRI is mainly caused by the spatially non-uniform transmission and reception properties of the radiofrequency coils, which results in a smooth intensity modulation [45]. The well-established N4 method from ANTs toolbox [1, 41] was applied to estimate and correct the bias field using the 3D T_1-weighted scans. This iterative method is based on image intensity histogram and seeks the smooth multiplicative field that maximizes

Table 2 MRI acquisition parameters for each research center participating in the INNI

	Center A		Center B		Center C		Center D	
MR scanner	Philips medical system—intera		GE medical system signa HDxt		Siemens Verio		Philips medical system—achieva	
MR imaging sequence	Dual-echo	T_1-weighted TFE	T_2-weighted FLAIR	T_1-weighted IR–FSPGR	Dual-echo	T_1-weighted MPRAGE	Dual-echo	T_1-weighted FFE
Coil	8-channel head coil		8-channel head coil		12-channel head coil		32-channel head coil	
Imaging plane	Axial	Axial	Axial	Sagittal	Axial	Sagittal	Axial	Axial
Acquisition voxel [mm^3]	1×1×3	1×1×1	1×1×3	1×1×1.2	1×1×4	1×1×1	1×1×3	1×1×1
FOV [mm^2]	243×243	230×230	256×256	256×256	220×220	256×256	240×240	256×256
TR [ms]	2910	25	9002	6.988	5310	1900	4000	10
TE [ms]	16–80	4.6	120	2.85	10–103	2.9	15–100	3.9
TI [ms]	–	–	2500	650	–	900	–	900
FA [°]	90	30	90	8	150	9	90	8
ETL	6	1	1	1	6	1	4	128
TA [min]	2.43	2.75	2.02	1.9	2.92	3.07	8.6	8.2

FOV field of view, *TR* repetition time, *TE* echo time, *TI* inversion time, *FA* flip angle, *ETL* echo train length, *TA* acquisition time, *TFE* turbo field echo, *IR* inversion recovery, *FSPGR* fast spoiled gradient echo, *MPRAGE* magnetization prepared gradient echo, *FFE* fast field echo

the high-frequency content of the distribution of tissue intensity. Being a multiplicative field, a value of 1 of the estimated bias field means no intensity modulation. For modulation values > 1, the image intensity is enhanced, while for modulation values < 1, intensity is reduced. The obtained bias field is a measure of the extent of the intensity inhomogeneity estimated by the algorithm between the original and the corrected images. As a consequence, this measure is related to the degree of homogeneity found on images. The bias field was assessed on four image sub-volumes as well as on the entire brain volume for each patient acquired at each center. The image sub-volumes were obtained by dividing the entire image brain volume into superior-right (sub-volume 1), superior-left (sub-volume 2), inferior-right (sub-volume 3), and inferior-left (sub-volume 4). The N4-corrected MRI are also provided to all centers, after the QC procedure.

3. The relative image contrast (RC) between normal appearing white matter (NAWM) and GM tissue, and between NAWM and MS lesions were estimated on 3D T_1-weighted and T_2-weighted scans, respectively, using the following equation [34, 36]:

$$\mathrm{RC} = \frac{S_{\mathrm{tissue1}} - S_{\mathrm{tissue2}}}{S_{\mathrm{tissue2}}}, \quad (2)$$

where $S_{\mathrm{tissue(x)}}$ is the average signal amplitude within the tissue of interest. The average signal within GM and NAWM was estimated after segmentation into GM, WM, and cerebrospinal fluid (CSF) using the 3D T_1-weighted images (SPM12—MATLAB; http://www.fil.ion.ucl.ac.uk/spm/software/spm12). MS lesions were semi-automatically segmented by an expert physician at each of the participating centers;

4. Image artefacts on MRI were categorized into several main classes according to the source of image corruption [35, 44]: (1) movements artefacts; (2) distortion artefacts; (3) MR acquisition technique-related artefacts; (4) B_0 inhomogeneity artefacts; (5) radio-frequency (RF) artefacts; (6) Noise; (7) Ghosting. For each of the artefact classes, subcategories were identified (see Table 3 for details). The presence of image artefacts was visually checked on the T_1-weighted and T_2-weighted images for each subject of the study and classified by an expert physician (trained in MRI) in collaboration with physicists.

Table 3 Main classifications of the MR image artefacts implemented for the INNI dataset

1. Movement artefacts	Motion
	Flow:
	From cerebral arteries
	From sagittal sinus
	From cerebrospinal fluid
	Physiological movements:
	Swallowing
	Eye
2. Distortion artefacts	Fold over
	Cross talk
	Partial volume
	Gradients non-linearity
3. MR acquisition technique-related artefacts	Truncation
	Chemical shift
	Parallel imaging
4. B0 artefacts	Susceptibility artefacts
	Eddy currents
	Metallic artefacts
	Imperfect fat saturation
5. RF artefacts	RF interference
	Spike of noise in the raw data
	Dielectric effect
	Stimulated echoes
	RF Coil artefacts
6. Noise	
7. Ghosting	
8. Other signal artefacts not ascribable to the previous categories	

B_0 static magnetic field, *RF* radio frequency

The relevance of the artefact (or artefacts) identified on each image was also visually inspected and related to the future quantitative analyses that could be performed (whole-brain atrophy quantification, subcortical atrophy, voxel-based morphometry, and lesion load quantification). Thus, based on the extent of the image artefacts on MRI, a feasibility opinion for future quantitative analyses is also provided as part of the QC procedure.

The four measures above have been included in a standard and semi-automatic QC procedure that can easily be applied to each MRI study uploaded to the INNI repository. The results are stored in a database that allows users of the INNI platform to inspect the QC measures, so that data for multicenter studies in MS can be reliably selected.

Statistical analysis

Statistical analyses were performed using the R software package, version 3.1.1, New Jersey. Kruskal–Wallis tests due to non-normal distribution of data (identified using the Kolgoroff–Smirnoff test) were used to evaluate (*a*) differences between head positioning z-offsets, (*b*) differences in intensity inhomogeneity, and (*c*) differences in relative tissue contrast among the different centers. Paired *t* tests were used to evaluate differences in intensity inhomogeneity among image sub-volumes for the same Center. Coefficients of variation (CoVs) were used to assess the variability of ED values for head positioning within each center and the intensity inhomogeneity among sub-volumes of the same image. Bonferroni correction was applied to adjust significance levels for multiple comparisons. $p < 0.05$ was considered to indicate a statistically significant result.

Data availability statement

Anonymized data and codes will be shared upon reasonable request by contacting the corresponding author.

Results

Figure 1 illustrates the z-offsets for the patient head positioning in the MR scanner for each center. Even though the positioning procedures for each center are focused on placing the region of the highest magnetic field homogeneity approximately around the center of the brain (the structure of interest), there were significant differences in positioning between the centers ($\chi^2 = 121.68$, $p < 0.001$), with the highest displacement of the brain center with respect to the magnet isocenter (median z-offset: $+4.8 \pm 4.5$ cm) and variability among subjects for Center C (CoV = 0.67), while the lowest displacements (medians z-offsets: -2.6 ± 1.7 cm and -1.1 ± 2 cm) and variability were found for Centers A and D ($CoV_A = 0.37$; $CoV_D = 0.49$). The 3D T_1-weighted MRI showed an increasing visible geometrical distortion at the maximum distance from the isocenter (in cases with high displacement values) for Centers C and D, as shown in the example of Fig. 2. Geometrical distortions are usually corrected by scanner manufacturers' image reconstruction software across all acquisition directions. However, some manufacturers' software does not apply this image distortion correction automatically for all the acquisition directions, and this option must be manually activated by the user. In this case, by obtaining the displacement field in each gradient direction from the manufacturer, which results from the gradient non-linearity of the MR system, it is possible to correct a posteriori the distorted MR image volume. This is something we are working on incorporating in the future [21].

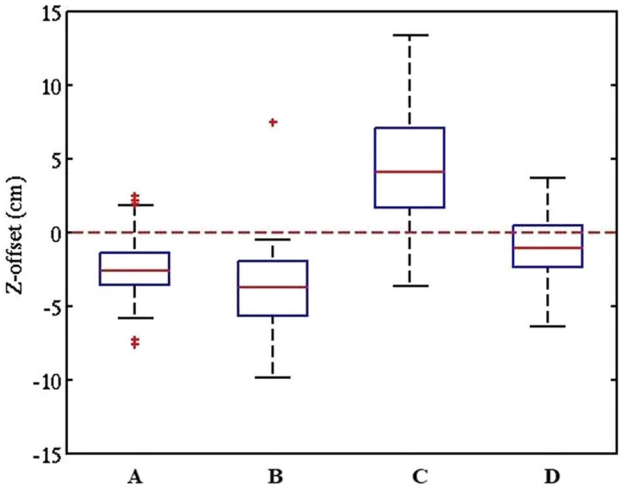

Fig. 1 Box and whiskers plots representing the distance (on the *z*-axis) between the isocenter of the static magnetic field and the center of the brain (red-dashed line) for each 3D T_1-weighted MRI for the different centers

We found significant differences between the centers for the average estimated bias fields on the whole 3D T_1-weighted scan ($p < 0.05$), with the exception of the comparison between Centers A and C, which were similar ($p = 0.59$). In Fig. 3a, bias field distributions for each center are shown. Considering sub-volumes of the 3D T_1-weighted image (Fig. 3b), none of the centers showed significant spatial differences of intensity inhomogeneity within the same image ($p > 0.05$). Thus, there was no particular spatial pattern of intensity inhomogeneity within the same image, possibly due to the geometric symmetry of RF coil design. Although all centers showed, on average, low intensity inhomogeneity, (median modulation values ≈ 1), the highest estimated inhomogeneity was for Center B (median root mean square error [RMSE] of the

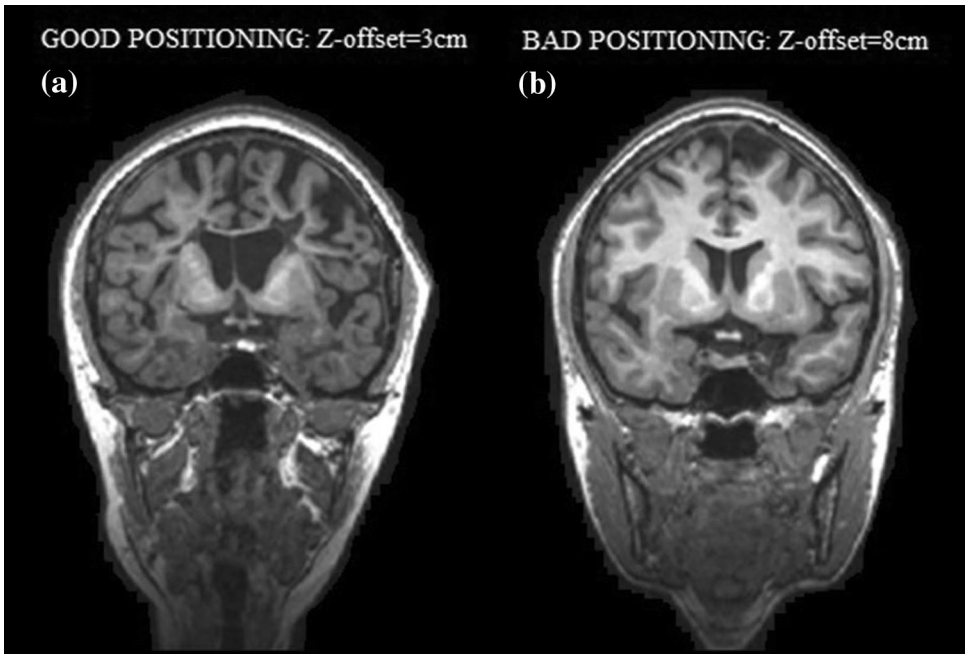

Fig. 2 Example of 3D T_1-weighted MRI acquired with a low (**a**) and high (**b**) z-offset from center B. In (**b**) a pronounced geometrical distortion in the superior part of the head is visible, caused by the poor positioning of the patient within the MR scanner

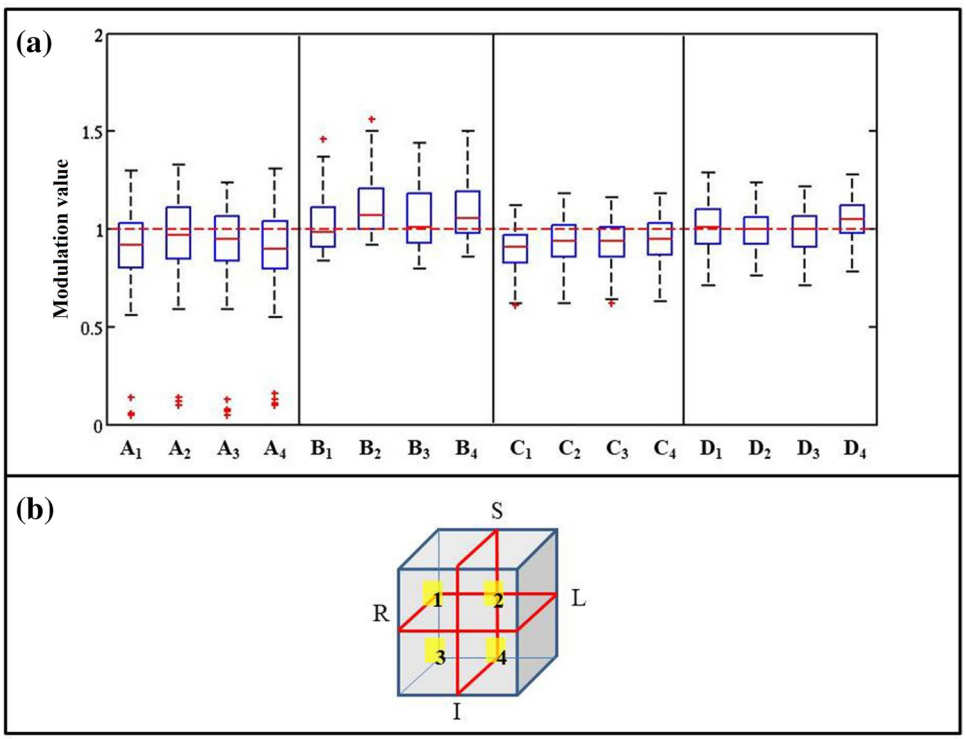

Fig. 3 Box and whiskers plots showing bias field modulation values estimated using the N4 correction method on each sub-volume (1, 2, 3, 4) of the 3D T_1-weighted acquired at each Center (A, B, C, D) are shown in (**a**). In (**b**) the partition of the image sub-volumes is represented. *R* right, *L* left, *S* superior, *I* inferior

bias field = 0.13 ± 0.13), while the lowest was for Center D (median RMSE = 0.08 ± 0.07).

Significant differences in relative tissue contrast between NAWM and GM on the 3D T_1-weighted scans were found among the four centers ($\chi^2 = 171.05$, $p < 0.001$), except for Centers B and C which showed similar T_1-weighted tissue contrasts ($p > 0.05$). In particular, Center D showed the highest relative contrast between NAWM and GM (0.63 ± 0.04) on 3D T_1-weighted images (Fig. 4). Due to a longer echo time, FLAIR MRI showed higher values of relative tissue contrast between NAWM and MS lesions (0.21 ± 0.06) in comparison with the T_2-weighted images without the inversion pulse (Fig. 5) from the other centers. For a fair assessment, we performed statistical

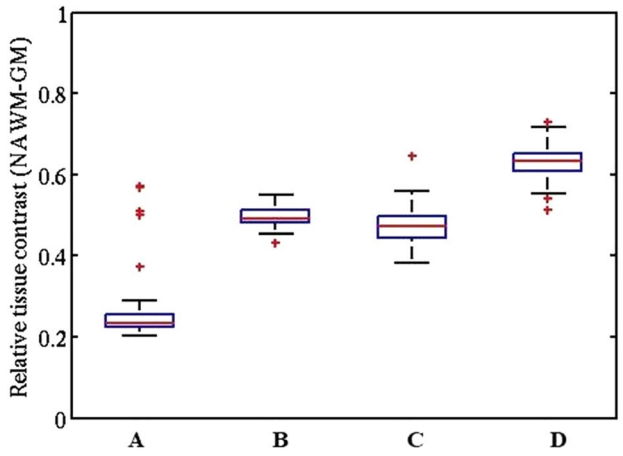

Fig. 4 Box and whiskers plots representing the estimated relative tissue contrast values between normal appearing WM (NAWM) and GM on 3D T_1-weighted scans for each Center

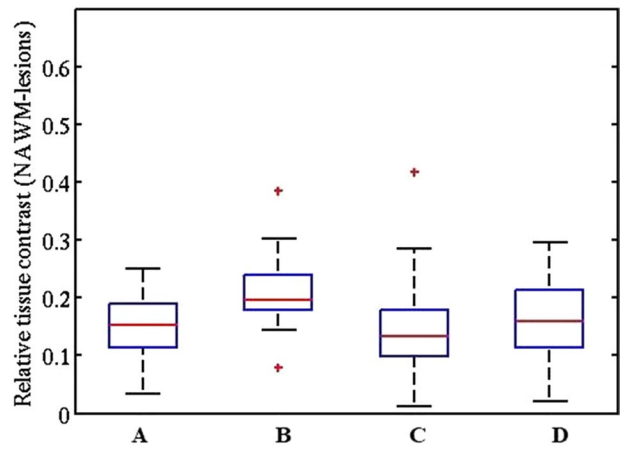

Fig. 5 Box and whiskers plots representing the relative tissue contrast values estimated between normal appearing WM (NAWM) and MS lesions on T_2-weighted scans for each center

comparisons only between the centers that acquired conventional T_2-weighted MRI (Centers A, C, and D), and we found no significant differences in relative tissue contrast between NAWM and MS lesions for these centers ($\chi^2 = 3.05$, $p > 0.05$).

According to the classification proposed (Table 3), the majority of image artefacts were due to motion, including flow, physiological motion, and head movement, representing 60% of all the artefacts found on 3D T_1-weighted and T_2-weighted images for each center. The second major source of image corruption was due to object ghosting (35% of all the other artefacts found on 3D T_1-weighted images). In particular, Center D showed a higher percentage of object ghosting image artefacts and RF interference (noise) in comparison with the other centers.

Discussion

One of the main goals of the INNI is to identify and validate novel MRI biomarkers for the study of MS. The small samples of patients enrolled and the recruitment of selective clinical phenotypes are among the common limitations of the majority of previous MRI studies in MS. Within the INNI, large-scale multicenter studies can be performed using MRI, and clinical and neuropsychological data collected from Italian research centers using high-field scanners and standardized MRI protocols within each center. In multicenter contexts, QC is a very important issue, since poor execution can compromise the reliability of a study [4]. In the literature, this topic has been widely explored, but very often, it has been mostly focused on quality assurance (i.e., avoiding the occurrence of problems by improving a process) rather than QC, mainly due to the prospective collection of standardized MRI data [18]. On the other hand, QC has mainly been performed by visual inspection at each step in the processing pipeline [23]. However, the huge amount of imaging data that will be collected in repositories, as for the INNI, makes QC at each step in a processing pipeline via visual inspection unfeasible.

Against this background, we proposed and implemented a quantitative QC procedure for all data uploaded to the INNI repository, as an initial pre-processing step. This allowed us to characterize the quality and the level of standardization of MRI acquisitions from different centers, thus guaranteeing a repository of quality-controlled MR acquisitions for subsequent multicenter analyses. Maintaining the quality and homogeneity of these data across the different centers is necessary for obtaining reliable results in large-scale studies in MS. However, this preliminary characterization of MRI currently included in the INNI database revealed that the integration of data from different centers (with similar but not standardized MRI acquisition protocols) is still challenging.

The correct positioning of the subject's brain within the MR scanner and the manufacturer's distortion correction is crucial steps for avoiding significant geometric distortions on images at greater distance from the magnet isocenter (the region with the highest field homogeneity). In this study, we found a significant variability among the centers in the positioning of the patients within the MR scanner, with the highest intra-center consistency for Centers A and D. Center C showed the highest percentage of patients with inaccurate centering of the brain (47% with ED > 5 cm). If the anatomy of interest is too far from the isocenter, parts of the image field of view experience poorer B_0 homogeneity and gradient non-linearity [47]. This results in worse geometrical distortion that would

likely affect advanced MRI measures [8]. Geometrical distortions due to gradient non-linearity may be dealt by manufacturer-supplied software. However, this correction can be optional or even not applicable for some MR scanners. Therefore, several methods have been proposed to a posteriori remove geometrical distortions, starting from the knowledge of proprietary information of the vendor describing the gradient linearity of the scanner [20, 21]. In these cases, the measure of the spatial deformation on images could be obtained. For the INNI Initiative, we apply the a posteriori correction for the geometrical distortions on images from two participating centers, since the manufacturer-supplied software performed this correction only in 2D. For the other two centers, it was already implemented by the scanner software in 3D. For all images, we inspected the presence of the distortion correction in 3D and, if not applied, we included the a posteriori method (already proposed also for the ADNI initiative) as a preprocessing step for the INNI database [21]. Correction of geometrical distortion performed by the MR scanner decreases the degree of image deformation, but this is not as effective as acquiring images closer to the scanner isocenter [37, 46]. According to the literature, it is recommended to keep the distance between the center of the brain (or of the structure of interest) and the isocenter below 5 cm to avoid significant geometrical distortions on images [3, 37].

Intensity inhomogeneity on MR images arises from the imperfections of the image acquisition process mainly due to the properties of the MRI device (for example, the uniformity of the radiofrequency field) and/or to the imaged object itself (i.e., shape, position, and orientation of the object inside the magnet). This spatial modulation of image intensity is also related to the wavelength of the RF: if the wavelength is not much longer than the size of the object, wave effects can spatially modulate the image intensity [5]. With the same design of the RF coil, at higher MR field strengths, the RF wavelength is typically shorter than the size of the object to be imaged, thus corrupting images with a spatial intensity modulation. Several MRI techniques, such as image segmentation and registration, are highly sensitive to these smooth intensity variations across the image. Therefore, the correction of the bias field is an important step for efficient segmentation and registration of brain MRI. All the 3D T_1-weighted images evaluated were affected by low average spatial signal intensity inhomogeneity, as found by the N4 method [41], although significant differences of the estimated bias field were found among the centers. Overall, due to the low bias field on 3D T_1-weighted images (modulation values ≈ 1), the intensity correction method was successfully applied on these images to adjust intensity histograms, which should improve the performance of post-processing algorithms in future studies, such as brain tissue segmentation or atrophy quantification. Centers participating in the INNI can also benefit from the systematic quantitative monitoring of intensity distortions to identify possible issues in the performance of the MR device. Future studies could further examine the relationship between intensity inhomogeneity and the accuracy of the applied analysis techniques (e.g., image segmentation and registration).

Although similar MR acquisition protocols were implemented by the different centers, the relative tissue contrast on 3D T_1-weighted brain images was significantly different between them. In particular, Center A showed a lower relative tissue contrast than the other centers, likely due to the absence of the magnetization preparation pulse on 3D T_1-weighted scans. Brain tissue and MS lesion segmentation algorithms are intrinsically dependent on the tissue contrast in an image, which in turn is dependent on the acquisition parameters. Accordingly, the appreciable differences in brain tissue contrast found on MRI acquired at the different centers could yield inconsistent segmentation results in multicenter studies. Attempting to overcome these issues, several techniques incorporate knowledge regarding the anatomical structure or other a priori information, such as atlas-guided approaches [9, 27]. In future work, it would be desirable to properly investigate the effect of these inconsistencies on a large-scale multicenter data. Automatic MS lesion segmentation methods may also benefit from better contrast between NAWM and lesions, increasing accuracy. Due to the longer echo time, we found a higher relative tissue contrast between NAWM and MS lesions for FLAIR MRI in comparison with T_2-weighted images without the inversion pulse. It has been demonstrated that, by providing better discrimination between lesions and normal appearing tissue, FLAIR MRI could yield better performance of fully automatic MS lesion segmentation methods [33]. Dual-echo T_2-weighted MRI is traditionally acquired at Italian research centers for manual MS lesion segmentation, but is increasingly being replaced by FLAIR.

In our assessment, we decided not to include estimation of the signal-to-noise ratio (SNR), since the use of a phased-array surface coils and parallel imaging influences the statistical and spatial distribution of noise. In this case, the SNR measured by evaluating a foreground and a background region would not provide an accurate assessment of SNR, with deviations that exceed 10% [10]. In addition, due to the multichannel acquisition techniques implemented at each center, the mean intensity value of the background region for the collected images is 0, making this method unfeasible for this study. Other methods for estimating SNR are based on the acquisition of at least two images with identical acquisition parameters, thus increasing the total imaging time [10].

Identifying image artefacts and understanding their causes are important for improving scanning protocols, pulse sequence designs, and data processing algorithms. The

performance of the analysis techniques could be differently affected by the presence and severity of image artefacts. As an example, even minimal motion artefacts decrease GM volumes by over 4%, potentially hiding between-group differences [6]. The classification of the sources of image corruption proposed in this study for the INNI database should be very useful for decision-making about inclusion or exclusion of MRI data for future large-scale multicenter studies in MS, according to the type of analysis (voxel-based analysis, global atrophy, etc.). Future developments could also include the quantification of the severity of the image artefacts with a view to implementing correction methods before the application of post-processing techniques.

Image QC is only a single aspect of the complex process to ensure a good quality and standardization level on MR acquisitions for future quantitative analyses, especially for multicenter studies. MR scanner maintenance is the other important aspect to avoid relevant discrepancies in the quality of MR images across time. For the INNI Initiative, to reduce variability, the procedure of image QC and pre-processing is centralized and performed at a single participating Center, while the maintenance and QC of the MR scanner are performed independently by each Center. The participating Centers have already implemented periodic maintenance procedures on their MR scanner to ensure the original and a constant quality of the MR acquisitions across time. Any major MR hardware or software change would be notified by the Center and recorded on the INNI repository for longitudinal studies. With the creation of a centralized repository containing MRI, clinical and neuropsychological data collected at different Research Centers with internationally recognized expertise, the INNI is promoting the use of these data to perform large-scale studies in MS. However, the integration of MRI data collected from different Research Centers is still challenging. Although an acceptable degree of standardization has been reached among the Centers involved in the INNI, more can be done in the MRI acquisition phase to reduce geometric distortions and improve the reliability of quantitative MRI measures.

In conclusion, the QC performed in this study will ensure a systematic quality assessment of the data collected within the INNI repository for large-scale studies in MS, with the ultimate goal of more standardized use of MRI in MS.

Acknowledgements This project has been supported by a research grant from the Fondazione Italiana Sclerosi Multipla (FISM2018/S/3), and financed or co-financed with the '5 per mille' public funding.

INNI network: Milan: Paola Valsasina, Mauro Sibilia, Paolo Preziosa; Naples: Antonio Gallo, Alvino Bisecco, Renato Docimo; Rome: Nikolaos Petsas, Serena Ruggieri, Silvia Tommasin; Siena: Maria Laura Stromillo, Riccardo Tappa Brocci.

Author contribution All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Loredana Storelli, Maria A. Rocca, Elisabetta Pagani, and Massimo Filippi. The first draft of the manuscript was written by Loredana Storelli, Maria A. Rocca, Elisabetta Pagani, and Massimo Filippi, and all authors commented on the previous versions of the manuscript. All authors read and approved the final manuscript.

Compliance with ethical standards

Conflicts of interest L Storelli, P. Pantano, E. Pagani, G. Tedeschi, and P. Zaratin have nothing to disclose. M.A. Rocca received speaker's honoraria from Biogen Idec, Novartis, Genzyme, Sanofi-Aventis, Teva, Merck Serono, and Roche and receives research support from the Italian Ministry of Health and Fondazione Italiana Sclerosi Multipla. N. De Stefano has received honoraria from Schering, Biogen Idec, Teva, Novartis, Genzyme, and Merck Serono S.A. for consulting services, and speaking and travel support. He serves on advisory boards for Biogen-Idec Merck Serono S.A. and Novartis. M. Filippi is Editor-in-Chief of the Journal of Neurology; received compensation for consulting services and/or speaking activities from Biogen Idec, Merck Serono, Novartis, Teva Pharmaceutical Industries; and receives research support from Biogen Idec, Merck Serono, Novartis, Teva Pharmaceutical Industries, Roche, Italian Ministry of Health, Fondazione Italiana Sclerosi Multipla, and ARiSLA (Fondazione Italiana di Ricerca per la SLA).

Ethical Standards Ethical approval was received from the local ethical standards committee of each participating Center, and written informed consent was obtained from all participants at the time of data acquisition.

References

1. Advanced Normalization Toolbox website: http://stnava.github.io/ANTs/
2. Alfaro-Almagro F, Jenkinson M, Bangerter NK, Andersson JLR, Griffanti L, Douaud G, Sotiropoulos SN, Jbabdi S, Hernandez-Fernandez M, Vallee E, Vidaurre D, Webster M, McCarthy P, Rorden C, Daducci A, Alexander DC, Zhang H, Dragonu I, Matthews PM, Miller KL, Smith SM (2018) Image processing and quality control for the first 10,000 brain imaging datasets from UK biobank. Neuroimage 166:400–424
3. Baldwin LN, Wachowicz K, Thomas SD, Rivest R, Fallone BG (2007) Characterization, prediction, and correction of geometric distortion in 3 T MR images. Med Phys 34:388–399
4. Bennett CM, Miller MB (2010) How reliable are the results from functional magnetic resonance imaging? Ann N Y Acad Sci 1191:133–155
5. Bernstein MAHJ, Ward HA (2006) Imaging artifacts at 3.0T. J Magn Reson Imaging 24:735–746
6. Blumenthal JD, Zijdenbos A, Molloy E, Giedd JN (2002) Motion artifact in magnetic resonance imaging: implications for automated analysis. Neuroimage 16:89–92
7. Brown JA, Van Horn JD (2016) Connected brains and minds—the UMCD repository for brain connectivity matrices. Neuroimage 124:1238–1241
8. Choudhri AF, Chin EM, Klimo P, Boop FA (2014) Spatial distortion due to field inhomogeneity in 3.0 tesla intraoperative MRI. Neuroradiol J 27:387–392
9. Despotovic I, Goossens B, Philips W (2015) MRI segmentation of the human brain: challenges, methods, and applications. Comput Math Methods Med 2015:450341
10. Dietrich O, Raya JG, Reeder SB, Reiser MF, Schoenberg SO (2007) Measurement of signal-to-noise ratios in MR images:

influence of multichannel coils, parallel imaging, and reconstruction filters. J Magn Reson Imaging 26:375–385
11. Eickhoff S, Nichols TE, Van Horn JD, Turner JA (2016) Sharing the wealth: neuroimaging data repositories. Neuroimage 124:1065–1068
12. Filippi M, Bar-Or A, Piehl F, Preziosa P, Solari A, Vukusic S, Rocca MA (2018) Multiple sclerosis. Nat Rev Dis Primers 4:43
13. Filippi M, Preziosa P, Rocca MA (2017) Microstructural MR imaging techniques in multiple sclerosis. Neuroimaging Clin N Am 27:313–333
14. Filippi M, Rocca MA, Bastianello S, Comi G, Gallo P, Gallucci M, Ghezzi A, Marrosu MG, Minonzio G, Pantano P, Pozzilli C, Tedeschi G, Trojano M, Falini A, De Stefano N, Neuroimaging, Neurology MSSGotISo, Functional Neuroradiology Section of the Italian Association of N (2013) Guidelines from the Italian neurological and neuroradiological societies for the use of magnetic resonance imaging in daily life clinical practice of multiple sclerosis patients. Neurol Sci 34:2085–2093
15. Filippi M, Rocca MA, Ciccarelli O, De Stefano N, Evangelou N, Kappos L, Rovira A, Sastre-Garriga J, Tintore M, Frederiksen JL, Gasperini C, Palace J, Reich DS, Banwell B, Montalban X, Barkhof F, Group MS (2016) MRI criteria for the diagnosis of multiple sclerosis: MAGNIMS consensus guidelines. Lancet Neurol 15:292–303
16. Filippi M, Tedeschi G, Pantano P, De Stefano N, Zaratin P, Rocca MA, Network I (2017) The Italian neuroimaging network initiative (INNI): enabling the use of advanced MRI techniques in patients with MS. Neurol Sci 38:1029–1038
17. Gedamu E (2011) Guidelines for developing automated quality control procedures for brain magnetic resonance images acquired in multi-centre clinical trials. Appl Exp Qual Control 26:135–158
18. Glover GH, Mueller BA, Turner JA, van Erp TG, Liu TT, Greve DN, Voyvodic JT, Rasmussen J, Brown GG, Keator DB, Calhoun VD, Lee HJ, Ford JM, Mathalon DH, Diaz M, O'Leary DS, Gadde S, Preda A, Lim KO, Wible CG, Stern HS, Belger A, McCarthy G, Ozyurt B, Potkin SG (2012) Function biomedical informatics research network recommendations for prospective multicenter functional MRI studies. J Magn Reson Imaging 36:39–54
19. Gorgolewski KJ, Varoquaux G, Rivera G, Schwartz Y, Sochat VV, Ghosh SS, Maumet C, Nichols TE, Poline JB, Yarkoni T, Margulies DS, Poldrack RA (2016) NeuroVault.org: a repository for sharing unthresholded statistical maps, parcellations, and atlases of the human brain. Neuroimage 124:1242–1244
20. Janke A, Zhao H, Cowin GJ, Galloway GJ, Doddrell DM (2004) Use of spherical harmonic deconvolution methods to compensate for nonlinear gradient effects on MRI images. Magn Reson Med 52:115–122
21. Jovicich J, Czanner S, Greve D, Haley E, van der Kouwe A, Gollub R, Kennedy D, Schmitt F, Brown G, Macfall J, Fischl B, Dale A (2006) Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. Neuroimage 30:436–443
22. Landhuis E (2017) Neuroscience: big brain, big data. Nature 541:559–561
23. Marcus DS, Harms MP, Snyder AZ, Jenkinson M, Wilson JA, Glasser MF, Barch DM, Archie KA, Burgess GC, Ramaratnam M, Hodge M, Horton W, Herrick R, Olsen T, McKay M, House M, Hileman M, Reid E, Harwell J, Coalson T, Schindler J, Elam JS, Curtiss SW, Van Essen DC, Consortium WU-MH (2013) Human Connectome Project informatics: quality control, database services, and data visualization. Neuroimage 80:202–219
24. McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD, McFarland HF, Paty DW, Polman CH, Reingold SC, Sandberg-Wollheim M, Sibley W, Thompson A, van den Noort S, Weinshenker BY, Wolinsky JS (2001) Recommended diagnostic criteria for multiple sclerosis: guidelines from the international panel on the diagnosis of multiple sclerosis. Ann Neurol 50:121–127
25. Miller KL, Alfaro-Almagro F, Bangerter NK, Thomas DL, Yacoub E, Xu J, Bartsch AJ, Jbabdi S, Sotiropoulos SN, Andersson JL, Griffanti L, Douaud G, Okell TW, Weale P, Dragonu I, Garratt S, Hudson S, Collins R, Jenkinson M, Matthews PM, Smith SM (2016) Multimodal population brain imaging in the UK Biobank prospective epidemiological study. Nat Neurosci 19:1523–1536
26. Petersen RC, Aisen PS, Beckett LA, Donohue MC, Gamst AC, Harvey DJ, Jack CR Jr, Jagust WJ, Shaw LM, Toga AW, Trojanowski JQ, Weiner MW (2010) Alzheimer's disease neuroimaging initiative (ADNI): clinical characterization. Neurology 74:201–209
27. Pham DL, Xu C, Prince JL (2000) Current methods in medical image segmentation. Annu Rev Biomed Eng 2:315–337
28. Poldrack RA, Gorgolewski KJ (2014) Making big data open: data sharing in neuroimaging. Nat Neurosci 17:1510–1517
29. Reid AT, Bzdok D, Genon S, Langner R, Muller VI, Eickhoff CR, Hoffstaedter F, Cieslik EC, Fox PT, Laird AR, Amunts K, Caspers S, Eickhoff SB (2016) ANIMA: a data-sharing initiative for neuroimaging meta-analyses. Neuroimage 124:1245–1253
30. Reuter M, Tisdall MD, Qureshi A, Buckner RL, van der Kouwe AJW, Fischl B (2015) Head motion during MRI acquisition reduces gray matter volume and thickness estimates. Neuroimage 107:107–115
31. Rocca MA, Battaglini M, Benedict RH, De Stefano N, Geurts JJ, Henry RG, Horsfield MA, Jenkinson M, Pagani E, Filippi M (2017) Brain MRI atrophy quantification in MS: from methods to clinical application. Neurology 88:403–413
32. Rocca MA, Filippi M (2007) Functional MRI in multiple sclerosis. J Neuroimaging 17(Suppl 1):36S–41S
33. Rovira À (2012) Segmentation of multiple sclerosis lesions in brain MRI: a review of automated approaches. Inf Sci 186:164–185
34. Roy S, Carass A, Prince J (2011) A compressed sensing approach for MR tissue contrast synthesis. Inf Process Med Imaging 22:371–383
35. Elhdnmkha S (2004) A short overview of MRI artefacts. SA J Radiol 8:13–17
36. Saranathan M, Tourdias T, Kerr AB, Bernstein JD, Kerchner GA, Han MH, Rutt BK (2014) Optimization of magnetization-prepared 3-dimensional fluid attenuated inversion recovery imaging for lesion detection at 7 T. Invest Radiol 49:290–298
37. Seibert TM, White NS, Kim GY, Moiseenko V, McDonald CR, Farid N, Bartsch H, Kuperman J, Karunamuni R, Marshall D, Holland D, Sanghvi P, Simpson DR, Mundt AJ, Dale AM, Hattangadi-Gluth JA (2016) Distortion inherent to magnetic resonance imaging can lead to geometric miss in radiosurgery planning. Pract Radiat Oncol 6:e319–e328
38. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, Downey P, Elliott P, Green J, Landray M, Liu B, Matthews P, Ong G, Pell J, Silman A, Young A, Sprosen T, Peakman T, Collins R (2015) UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 12:e1001779
39. Thompson AJ, Banwell BL, Barkhof F, Carroll WM, Coetzee T, Comi G, Correale J, Fazekas F, Filippi M, Freedman MS, Fujihara K, Galetta SL, Hartung HP, Kappos L, Lublin FD, Marrie RA, Miller AE, Miller DH, Montalban X, Mowry EM, Sorensen PS, Tintore M, Traboulsee AL, Trojano M, Uitdehaag BMJ, Vukusic S, Waubant E, Weinshenker BG, Reingold SC, Cohen JA (2018) Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. Lancet Neurol 17:162–173
40. Tisdall MD, Reuter M, Qureshi A, Buckner RL, Fischl B, van der Kouwe AJW (2016) Prospective motion correction with

volumetric navigators (vNavs) reduces the bias and variance in brain morphometry induced by subject motion. Neuroimage 127:11–22
41. Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, Gee JC (2010) N4ITK: improved N3 bias correction. IEEE Trans Med Imaging 29:1310–1320
42. Van Essen DC, Smith SM, Barch DM, Behrens TE, Yacoub E, Ugurbil K, Consortium WU-MH (2013) The WU-Minn human connectome project: an overview. Neuroimage 80:62–79
43. Van Horn JD, Toga AW (2014) Human neuroimaging as a "Big Data" science. Brain Imaging Behav 8:323–331
44. Vargas MI, Delavelle J, Kohler R, Becker CD, Lovblad K (2009) Brain and spine MRI artifacts at 3Tesla. J Neuroradiol 36:74–81
45. Vovk U, Pernus F, Likar B (2007) A review of methods for correction of intensity inhomogeneity in MRI. IEEE Trans Med Imaging 26:405–421
46. Wang D, Doddrell DM, Cowin G (2004) A novel phantom and method for comprehensive 3-dimensional measurement and correction of geometric distortion in magnetic resonance imaging. Magn Reson Imaging 22:529–542
47. Weavers PT, Tao S, Trzasko JD, Shu Y, Tryggestad EJ, Gunter JL, McGee KP, Litwiller DV, Hwang KP, Bernstein MA (2017) Image-based gradient non-linearity characterization to determine higher-order spherical harmonic coefficients for improved spatial position accuracy in magnetic resonance imaging. Magn Reson Imaging 38:54–62
48. Webb-Vargas Y, Chen S, Fisher A, Mejia A, Xu Y, Crainiceanu C, Caffo B, Lindquist MA (2017) Big Data and Neuroimaging. Stat Biosci 9:543–558
49. Weiner MW, Veitch DP, Aisen PS, Beckett LA, Cairns NJ, Green RC, Harvey D, Jack CR, Jagust W, Liu E, Morris JC, Petersen RC, Saykin AJ, Schmidt ME, Shaw L, Shen L, Siuciak JA, Soares H, Toga AW, Trojanowski JQ, Alzheimer's Disease Neuroimaging I (2013) The Alzheimer's disease neuroimaging initiative: a review of papers published since its inception. Alzheimers Dement 9:e111–194
50. Ziemssen T, Rauer S, Stadelmann C, Henze T, Koehler J, Penner IK, Lang M, Poehlau D, Baier-Ebert M, Schieb H, Meuth S (2015) Evaluation of study and patient characteristics of clinical studies in primary progressive multiple sclerosis: a systematic review. PLoS One 10:e0138243
51. Zivadinov R, Bergsland N, Korn JR, Dwyer MG, Khan N, Medin J, Price JC, Weinstock-Guttman B, Silva D, Group M-MS (2018) Feasibility of brain atrophy measurement in clinical routine without prior standardization of the MRI protocol: results from MS-MRIUS, a longitudinal observational, multicenter real-world outcome study in patients with relapsing–remitting MS. AJNR Am J Neuroradiol 39:289–295