Journal of Neuroradiology 000 (2023) 1–5

Available online at ScienceDirect www.sciencedirect.com

Elsevier Masson France EM|consulte www.em-consulte.com

Technical Note

# Growing importance of brain morphometry analysis in the clinical routine: The hidden impact of MR sequence parameters

Michael Rebsamen[a,b], Milena Capiglioni[a,b], Robert Hoepner[d], Anke Salmen[d], Roland Wiest[a,c], Piotr Radojewski[1,a,c], Christian Rummel[*,1,a]

[a] Support Center for Advanced Neuroimaging (SCAN), University Institute of Diagnostic and Interventional Neuroradiology, University of Bern, Inselspital, Bern University Hospital, Bern, Switzerland
[b] Graduate School for Cellular and Biomedical Sciences, University of Bern, Bern, Switzerland
[c] Swiss Institute for Translational and Entrepreneurial Medicine, sitem-insel, Bern, Switzerland
[d] Department of Neurology, Inselspital, Bern University Hospital and University of Bern, Bern, Switzerland

ARTICLE INFO

*Article History:*
Available online xxx

*Keywords:*
MRI
Imaging protocol
Brain morphometry
Atrophy
Reliability

ABSTRACT

Volumetric assessment based on structural MRI is increasingly recognized as an auxiliary tool to visual reading, also in examinations acquired in the clinical routine. However, MRI acquisition parameters can significantly influence these measures, which must be considered when interpreting the results on an individual patient level.
This Technical Note shall demonstrate the problem. Using data from a dedicated experiment, we show the influence of two crucial sequence parameters on the GM/WM contrast and their impact on the measured volumes. A simulated contrast derived from acquisition parameters TI/TR may serve as surrogate and is highly correlated (r=0.96) with the measured contrast.



## Background and motivation

Quantitative image analysis of the brain is increasingly recognized as complementary to visual reading by neuroradiologists in the clinical routine and is usually performed on 3D T1-weighted MR images. In multiple sclerosis (MS) diagnosis, the latest *No Evidence of Disease Activity* (NEDA-4) criteria include brain atrophy measurement as imaging biomarker for disease progression.[15] Reliably assessing cortical atrophy on an individual patient level practically requires automated quantification of brain volume across multiple time points. While brain morphometry has contributed substantially to the understanding of brain structure in health and disease and is being used extensively in research, translation into the clinical routine remains challenging.[1,23] There is a long way from the carefully crafted cohort studies with data acquired in a well-controlled research setting, with analysis performed by technical experts using scientific tools, to the application on individual patient data acquired in the clinical routine. Commercial products try to bridge this gap and make brain volumetry and morphometry accessible in clinics providing certified products for off-the-shelf use, although technical validations are sparse and there is a considerable lack of clinical validation studies.[20] In MS, quantification of brain atrophy is increasingly used as secondary outcome measure in clinical trials, but it's use in clinical routine is still under debate due to various challenges.[27,32]

We fully embrace the trend to bring recent advances in neuroimaging research closer to the patients. Therefore, we would like to draw the attention to an, in our experience, often underestimated limitation. MR acquisition sequence parameters are known to influence quantitative image analysis methods to a much larger extent than expert reading,[9,11,12,16,21] a fact that is often neglected in clinical routine. This Technical Note shall demonstrate the issue, starting with an example of a patient with MS from the clinical routine and a simple visualization of the most relevant technical parameters of the T1-weighted MP-RAGE protocol.[3,17] Further, we propose a simulation of the expected image contrast and validate this concept with own experiments. The acquired data will be made publicly available for the clinical and scientific community as well as to vendors of commercial morphometry software to support validation studies in different environments using the same dataset.

**Example:** Patient from the clinical routine at our department visualizing the motivation behind our letter. The female patient with relapsing-remitting MS was referred for elective routine follow-up imaging. The EDSS at the time of referral was 3.5, the patient received no neuro-modulating treatment. A standard contrast enhanced MS protocol was acquired at 1.5 T MR. The examination showed no new

* Corresponding author.
*E-mail address:* crummel@web.de (C. Rummel).
[1] Shared senior authorship

https://doi.org/10.1016/j.neurad.2023.04.003


Please cite this article as: M. Rebsamen, M. Capiglioni, R. Hoepner et al., Growing importance of brain morphometry analysis in the clinical routine: The hidden impact of MR sequence parameters, Journal of Neuroradiology (2023), https://doi.org/10.1016/j.neurad.2023.04.003

MS lesion or contrast enhancement. The longitudinal volumetry suggests approx. 10% decrease in cortical gray matter (GM) volume (Fig. 1).

## Materials and methods

### *Simulated MR contrast*

Cortical gray matter segmentation is ultimately dependent on the gray (GM)/white (WM) contrast,[16,34] i.e., the MR signal difference of these adjacent tissue classes.[18] For the sake of the argument, we restrict to the generic inversion recovery (IR) pulse sequence [13] here and discuss possible extensions to other sequences in the Discussion section. Based on fundamental MR physics, we calculate a *"simulated GM/WM contrast"* defined by the parameters predominantly influencing the signal: inversion time (TI) and repetition time (TR). The signal intensity $S_x$ of a voxel in tissue class $x$ is given by

$$S_x = K \cdot \rho_x \cdot (1 - 2 \cdot e^{-TI/T1_x} + e^{-TR/T1_x}) \quad (1)$$

where $K$ is a constant irrelevant for our present purposes, $\rho_x$ is the proton density and $T1_x$ is the spin-lattice (longitudinal) relaxation time specific to a tissue class $x$. TI and TR can be easily extracted from the standard DICOM header of the chosen MRI protocol. The contrast between GM and WM signals is given by

$$Contrast = \left| \frac{S_{gm} - S_{wm}}{S_{gm} + S_{wm}} \right| \quad (2)$$

Assuming constant tissue properties across the whole brain of $\rho$=85/70 [13] and T1=1820/1084 ms [30] for GM/WM results in the contrast depending on the two sequence parameters TI and TR. We suggest depicting the sequence parameters on a heatmap of the simulated contrast to visualize changes in technical acquisition parameters that might impact morphometric measures (cf. Fig. 1).

### *Validation with experiments on the "Phantom of Bern"*

In an experiment, we acquired nine same-session scans of two healthy volunteers with eight different combinations of the sequence parameters TI and TR [5] and compared the suggested simulated contrast to the GM/WM contrast [26] measured with the freely available software package FreeSurfer 6.0 [6,7] and its impact on cortical GM volume by calculating Pearson correlation coefficients ($r$). Baseline was the ADNI-3 protocol [8,33] from which we modified TR/TI combinations while the remaining parameters remained unchanged. The results were also replicated with SIENA [28] using the longitudinal pipeline to quantify relative brain volume change, with DL+DiReCT [22] to estimate cortical thickness, and with FSL-FAST [29,36] for total brain volume.

## Results

The simulated contrast showed a very high correlation to the measured global mean GM/WM contrast of FreeSurfer ($r = 0.96$). The repeated acquisitions with the MR sequence according to the original ADNI-3 protocol, which had the highest (simulated and measured) contrast were separated by only $\sim 5\%$ of the total variability of the measured contrast between 18% and 30%, showing very high reproducibility between start and end of the scan session of about 45 minutes duration (Fig. 2). Moreover, cortical GM volumes were substantially negatively correlated ($r = -0.54$, $r = -0.64$) with the simulated contrast. In our experiments we have observed up to 2.5% variation of the estimated GM volume from same-session acquisitions, which can only be attributed to changes in TI and TR, since all other parameters remained unchanged. A qualitative example comparing the images with the lowest to the highest contrast is shown in Fig. 3.

The repeated scans with identical scanning-parameters in both subjects revealed an uncertainty of about 2 ml for measurement of total GM volume (about 0.5%). The variability due to different contrast and sequence parameters is much larger than these values ($\sim$ 9.5 to 12.5 ml, i.e. $\sim$2.0 to 2.5% of the total GM volume).

Similar effects were observed for relative brain volume using SIENA (Supplementary Section S1), for global mean cortical thickness using DL+DiReCT (Supplementary Section S2), and for total brain volume using FSL-FAST (Supplementary Section S3).

## Discussion

Motivated by the growing interest to apply quantitative assessment of structural MRI in the clinical routine, it is essential to highlight the impact of MRI acquisition parameters on volumetry, which users like neuroradiologists or neurologists need to consider when

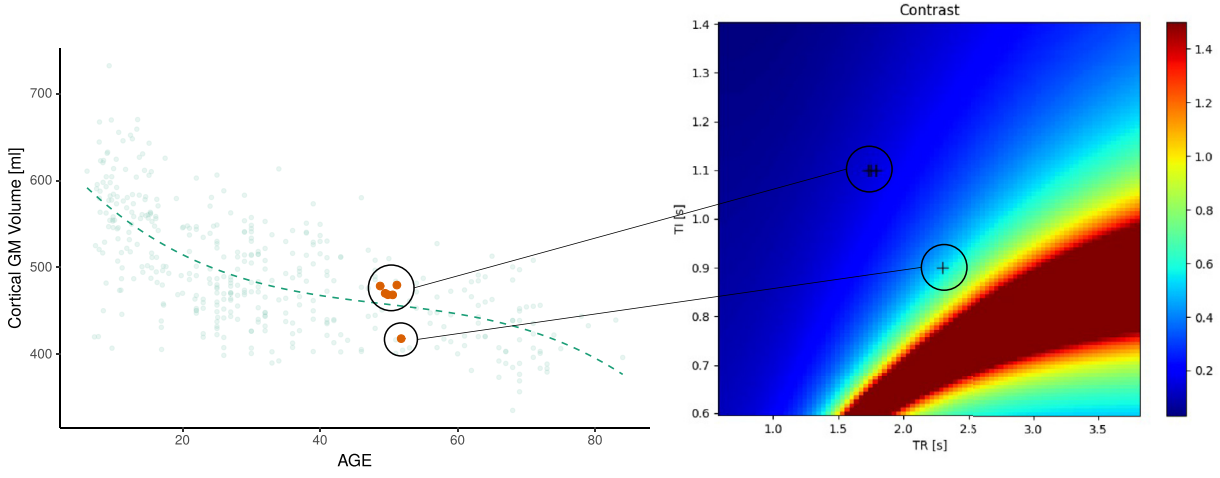


**Fig. 1.** *swimming in cold waters*: Longitudinal assessment of cortical gray matter volume in an MS patient contrasted to healthy controls (left panel, follow-up examinations of the patient acquired at the same scanner highlighted in red and normative dataset in pale green; green dashed line represents a polynomial fit to the normative measures of the healthy controls). The latest examination suggests a significant drop in GM volume suspected as an artifact resulting from a change in MR sequence parameters in the clinical routine (right, for the description of the contrast map see Materials and Methods section). Corresponding results using SIENA can be found in Supplementary Figure S1. Image parameters of the lastest scan: TR 2300 ms, TI 900 ms, TE 2.57 ms, voxel size 1 mm. The previous scans were acquired with TR 1790 ms, TI 1100 ms, TE 2.58 ms, voxel size 1 mm. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

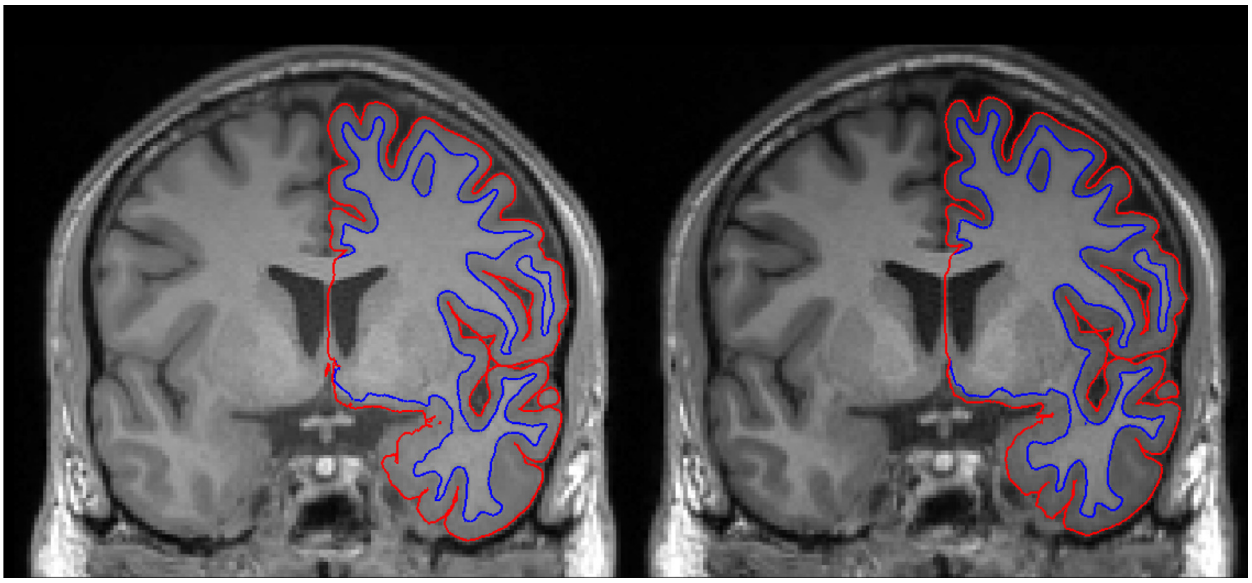

**Fig. 3.** Qualitative comparison of the MRIs with lowest (left) and highest GM/WM contrast (right) in subject *POBHC0002 of the "Phantom of Bern"*. To allow better judgment of the contrast between tissue classes, FreeSurfer's pial surface (red) and GM/WM boundary (blue) are shown in the left hemisphere only (displayed on the right side of the image in radiological orientation). Quantitative analysis reveals that the remarkably small differences in placement of the boundaries detectable by visual inspection accumulate systematically with all used software packages (see Fig. 2, Supplementary Figures S2,S4,S6). Left MRI: TI=1.1 s, TR=1.84 s, TE=2.96 ms, cortical GM volume=511 ml, mean GM/WM contrast=19.0%. Right MRI: TI=0.9 s, TR=2.3 s, TE=2.96 ms, cortical GM volume=489 ml, mean GM/WM contrast=29.8%. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

interpreting data from individual patients. We have designed a dedicated experiment with two healthy volunteers to depict the issue. Our results confirm and visualize that acquisition parameters like TI and TR can introduce a systematic bias in volumetric and morphometric measures.[9]

Our observed within-session reliability of repeated acquisitions with identical parameters (0.5%) is consistent with previous studies,[16] which roughly corresponds to the threshold 0.4% for annual brain volume loss required to diagnose *no evidence of disease activity* according to NEDA-4. Deriving annual brain atrophy rate across more than two time points for longitudinal assessments might partially alleviate this problem. However, a volume variability of up to 2.5%, which can be attributed to different sequence parameters, significantly exceeds this threshold, clearly demonstrating the importance of considering this confounding influence. Consequently, we agree with the general recommendation to keep acquisition parameters fixed,[11,35] but stress that this is often only possible in well-controlled research settings. For clinical applications, morphometry tools should clearly indicate such protocol deviations, prompting users like neurologists and neuroradiologists to critically assess the generated output.

Calculating a simulated contrast based on two MR parameters yielded a highly accurate (r=0.96) prediction of the true (measured) GM/WM contrast. This metric might be used to scrutinize the comparability of different acquisitions. A current limitation is that our argument is based on a generic IR sequence, whereas in practice other T1-weighted sequences might be used, like the MP-RAGE protocol [3,17] in our own experiment. However, based on our experience, the general effects are very similar for the MP-RAGE sequence while the formula for the IR sequence depends only on parameters readily available from the standard DICOM tags which simplifies applicability.

Future methodological improvements might mitigate part of the problem, e.g., generalizing the formula for the synthetic contrast to other sequences, or statistical models correcting for the influence of acquisition parameters [24] for which the simulated contrast might serve as confounding factor. Post-processing techniques, such as AI-based solutions for de-noising [14] or harmonizing MRI contrast [4] might be worthwhile investigating. Exploring learning-based

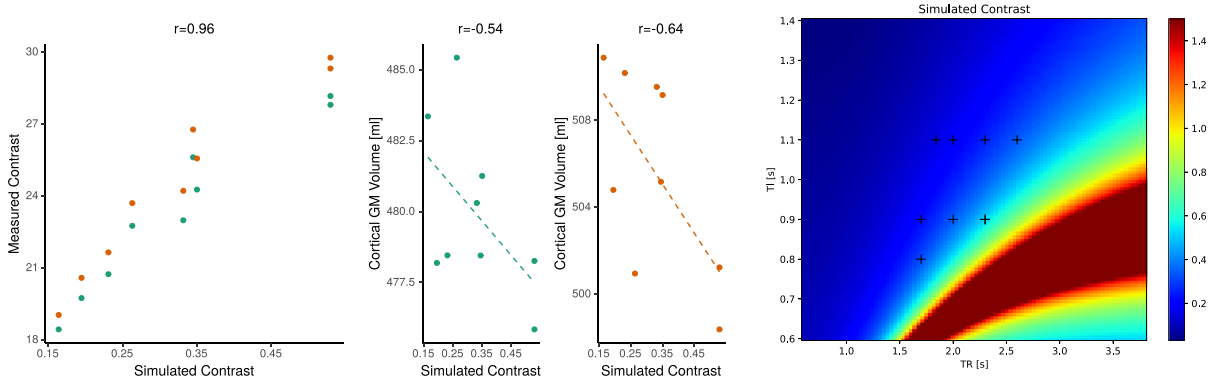


**Fig. 2.** Left: Correlation of simulated contrast with measured GM/WM contrast. Middle: Correlation with cortical GM volume for the two subjects (y-axis scales are intentionally left subject-specific to emphasize the effect). Right: Heatmap of the simulated contrast calculated from TI and TR. Black crosses show the eight different choices of combinations of acquisition parameters TI/TR in our experiment.

segmentation methods agnostic to image contrast, e.g., by leveraging large amounts of images with different contrasts as training data for semi-supervised learning,[2] might be another approach.

Quantitative assessments of MRI in clinical routine is currently performed on a regular basis only by a minority of institutes,[31] but likely to increase in the future with the adoption of NEDA-4. Data from our experiment can be used to assess the reliability of existing morphometry tools or check the robustness of new solutions against varying sequence parameters, which should be part of quality control when evaluating solutions.[10,19] Our experiments were performed using FreeSurfer, one of the most widely used and evaluated software for brain morphometry.[20,31] The findings were confirmed using SIENA, the reference method for NEDA-4,[15,27] DL+DiReCT,[22] a novel deep learning-based method that has recently proven excellent accuracy regarding atrophy detection,[25] and FSL-FAST.[29,36] We strongly assume that other tools including commercial products are likely affected to similar extent.

## Conclusions

Using the example of total cortical GM volume and the free software FreeSurfer, we demonstrated that changes in MR acquisition parameters can significantly influence estimates of quantitative measures of brain morphometry and therefore influence clinical decisions. Users interpreting such measures for diagnostic purpose should be aware of these limitations and software tools should ideally display related uncertainties. This is of utmost importance when applying brain morphometry in clinical routine with varying levels of sequence standardization and inherent variations over time. Depicting parameters of the various acquisitions on a visual heatmap of the simulated GM/WM contrast makes relevant heterogeneity immediately visible to non-technical readers, guiding them to interpret results in this context. We believe that the widespread adoption of quantitative assessments like *No Evidence of Disease Activity* (NEDA-4) in the clinical routine can only be successful if the generated morphometric data is highly reliable and reflects true biological findings.

## Data availability

The MRI dataset with repeated acquisitions of healthy subjects used for this study is available from https://openneuro.org/datasets/ds004560.

## Ethics statement

The presented MS case was enrolled in the neuroimmunological registry from Inselspital, approved by the Ethikkommision Kanton Bern (KEK-BE 2017-01369).

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgment

This work was supported by the Swiss National Science Foundation under grant numbers CRSII5_180365 (The Swiss-First Study) and 204593 (ScanOMetrics). The authors acknowledge the infrastructure and support of the Translational Imaging Center at sitem-insel and funding from the strategic research fund of the Medical Faculty of the University of Bern 2021-2023 and support by Biogen, grant number CHE-TYS-18-11316. We appreciate the support of David Romascano for the data curation of the published dataset.

## Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.neurad.2023.04.003

## References

1. Alroughani R, Deleu D, El Salem K, et al. A regional consensus recommendation on brain atrophy as an outcome measure in multiple sclerosis. *BMC Neurol*. 2016;16(1):1–9. https://doi.org/10.1186/s12883-016-0762-5.
2. Coupé P, Mansencal B, Clément M, et al. AssemblyNet: a large ensemble of CNNs for 3D whole brain MRI segmentation. *Neuroimage*. 2020;219:117026. https://doi.org/10.1016/j.neuroimage.2020.117026.
3. Deichmann R, Good CD, Josephs O, et al. Optimization of 3-D MP-RAGE sequences for structural brain imaging. *Neuroimage*. 2000;12(1):112–127. https://doi.org/10.1006/nimg.2000.0601.
4. Dewey BE, Zhao C, Reinhold JC, et al. DeepHarmony: a deep learning approach to contrast harmonization across scanner changes. *Magn Reson Imaging*. 2019;64:160–170. https://doi.org/10.1016/j.mri.2019.05.041.
5. Rebsamen M, Romascano D, Capiglioni M, et al. The Phantom of Bern: repeated scans of two volunteers with eight different combinations of MR sequence parameters. *OpenNeuro*. 2023. https://doi.org/10.18112/openneuro.ds004560.v1.0.1.
6. Fischl B, Salat DH, Busa E, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*. 2002;33(3):341–355. https://doi.org/10.1016/S0896-6273(02)00569-X.
7. Fischl B. FreeSurfer. *Neuroimage*. 2012;62(2):774–781. https://doi.org/10.1016/j.neuroimage.2012.01.021.
8. Gunter J, Thostenson K, Borowski B, et al. ADNI-3 MRI Protocol. *Alzheimers Dementia*. 2017;13(7):P104–P105. https://doi.org/10.1016/j.jalz.2017.06.2411.
9. Haller S, Falkovskiy P, Meuli R, et al. Basic MR sequence parameters systematically bias automated brain volume estimation. *Neuroradiology*. 2016;58(11):1153–1160. https://doi.org/10.1007/s00234-016-1737-3.
10. Haller S, Van Cauter S, Federau C, et al. The R-AI-DIOLOGY checklist: a practical checklist for evaluation of artificial intelligence tools in clinical neuroradiology. *Neuroradiology*. 2022:1–14. https://doi.org/10.1007/s00234-021-02890-w.
11. Han X, Jovicich J, Salat D, et al. Reliability of MRI-derived measurements of human cerebral cortical thickness: the effects of field strength, scanner upgrade and manufacturer. *Neuroimage*. 2006;32(1):180–194. https://doi.org/10.1016/j.neuroimage.2006.02.051.
12. Hedges EP, Dimitrov M, Zahid U, et al. Reliability of structural MRI measurements: the effects of scan session, head tilt, inter-scan interval, acquisition sequence, FreeSurfer version and processing stream. *Neuroimage*. 2022;246:118751. https://doi.org/10.1016/j.neuroimage.2021.118751.
13. Hornak JP. The basics of MRI. *http://www.cisritedu/htbooks/mri*. 2006.
14. Kanemaru N, Takao H, Amemiya S, et al. The effect of a post-scan processing denoising system on image quality and morphometric analysis. *J Neuroradiol*. 2022;49(2):205–212. https://doi.org/10.1016/j.neurad.2021.11.007.
15. Kappos L, De Stefano N, Freedman MS, et al. Inclusion of brain volume loss in a revised measure of no evidence of disease activity(NEDA-4) in relapsing–remitting multiple sclerosis. *Multip Sclerosis J*. 2016;22(10):1297–1305. https://doi.org/10.1177/1352458515616701.
16. Kruggel F, Turner J, Muftuler LT, et al. Impact of scanner hardware and imaging protocol on image quality and compartment volume precision in the ADNI cohort. *Neuroimage*. 2010;49(3):2123–2133. https://doi.org/10.1016/j.neuroimage.2009.11.006.
17. Mugler III JP, Brookeman JR. Three-dimensional magnetization-prepared rapid gradient-echo imaging (3D MP RAGE). *Magn Reson Med*. 1990;15(1):152–157. https://doi.org/10.1002/mrm.1910150117.
18. Nitz WR, Reimer P. Contrast mechanisms in MR imaging. *Eur Radiol*. 1999;9(6):1032–1046. https://doi.org/10.1007/s003300050789.
19. Omoumi P, Ducarouge A, Tournier A, et al. To buy or not to buyevaluating commercial AI solutions in radiology (the ECLAIR guidelines). *Eur Radiol*. 2021;31(6):3786–3796. https://doi.org/10.1007/s00330-020-07684-x.
20. Pemberton HG, Zaki LAM, Goodkin O, et al. Technical and clinical validation of commercial automated volumetric MRI tools for dementia diagnosisa systematic review. *Neuroradiology*. 2021;63(11):1773–1789. https://doi.org/10.1007/s00234-021-02818-4.
21. Potvin O, Khademi A, Chouinard I, et al. Measurement variability following MRI system upgrade. *Front Neurol*. 2019;10:726. https://doi.org/10.3389/fneur.2019.00726.
22. Rebsamen M, Rummel C, Reyes M, et al. Direct cortical thickness estimation using deep learning-based anatomy segmentation and cortex parcellation. *Hum Brain Mapp*. 2020;41(17):4804–4814. https://doi.org/10.1002/hbm.25159.
23. Rocca MA, Battaglini M, Benedict RHB, et al. Brain MRI atrophy quantification in MS: from methods to clinical application. *Neurology*. 2017;88(4):403–413. https://doi.org/10.1212/WNL.0000000000003542.
24. Rummel C, Aschwanden F, McKinley R, et al. A fully automated pipeline for normative atrophy in patients with neurodegenerative disease. *Front Neurol*. 2018;8:727. https://doi.org/10.3389/fneur.2017.00727.
25. Rusak F, Santa Cruz R, Lebrat L, et al. Quantifiable brain atrophy synthesis for benchmarking of cortical thickness estimation methods. *Med Image Anal*. 2022;82:102576. https://doi.org/10.1016/j.media.2022.102576.

26. Salat DH, Lee SY, Van der Kouwe AJ, et al. Age-associated alterations in cortical gray and white matter signal intensity and gray to white matter contrast. *Neuroimage*. 2009;48(1):21–28. https://doi.org/10.1016/j.neuroimage.2009.06.074.
27. Sastre-Garriga J, Pareto D, Battaglini M, et al. MAGNIMS Consensus recommendations on the use of brain and spinal cord atrophy measures in clinical practice. *Nature Rev Neurol*. 2020;16(3):171–182. https://doi.org/10.1038/s41582-020-0314-x.
28. Smith SM, Zhang Y, Jenkinson M, et al. Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. *Neuroimage*. 2002;17(1):479–489. https://doi.org/10.1006/nimg.2002.1040.
29. Smith SM, Jenkinson M, Woolrich MW, et al. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*. 2004;23:S208–S219. https://doi.org/10.1016/j.neuroimage.2004.07.051.
30. Stanisz GJ, Odrobina EE, Pun J, et al. T1, T2 Relaxation and magnetization transfer in tissue at 3T. *Magnet Resonance Med: Off J Int Soc Magnet Resonance Med*. 2005;54 (3):507–512. https://doi.org/10.1002/mrm.20605.
31. Vernooij MW, Pizzini FB, Schmidt R, et al. Dementia imaging in clinical practice: a european-wide survey of 193 centres and conclusions by the ESNR working group. *Neuroradiology*. 2019;61(6):633–642. https://doi.org/10.1007/s00234-019-02188-y.
32. Wattjes MP, Ciccarelli O, Reich DS, et al. 2021 MAGNIMS–CMSC–NAIMS Consensus recommendations on the use of MRI in patients with multiple sclerosis. *Lancet Neurol*. 2021;20(8):653–670. https://doi.org/10.1016/S1474-4422(21)00095-8.
33. Weiner MW, Veitch DP, Aisen PS, et al. The Alzheimer's disease neuroimaging initiative 3: continued innovation for clinical trial improvement. *Alzheimer's & Dementia*. 2017;13(5):561–571. https://doi.org/10.1016/j.jalz.2016.10.006.
34. Westlye LT, Walhovd KB, Dale AM, et al. Increased sensitivity to effects of normal aging and Alzheimer's disease on cortical thickness by adjustment for local variability in gray/white contrast: a multi-sample MRI study. *Neuroimage*. 2009;47 (4):1545–1557. https://doi.org/10.1016/j.neuroimage.2009.05.084.
35. Wonderlick JS, Ziegler DA, Hosseini-Varnamkhasti P, et al. Reliability of MRI-derived cortical and subcortical morphometric measures: effects of pulse sequence, voxel geometry, and parallel imaging. *Neuroimage*. 2009;44(4):1324–1333. https://doi.org/10.1016/j.neuroimage.2008.10.037.
36. Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imag*. 2001;20(1):45–57. https://doi.org/10.1109/42.906424.