NeuroImage: Clinical 31 (2021) 102705

Contents lists available at ScienceDirect

NeuroImage: Clinical

journal homepage: www.elsevier.com/locate/ynicl

Quantifying disease pathology and predicting disease progression in multiple sclerosis with only clinical routine T2-FLAIR MRI

Tom A. Fuchs [a,b], Michael G. Dwyer [a,c], Dejan Jakimovski [a], Niels Bergsland [a,d], Deepa P. Ramasamy [a], Bianca Weinstock-Guttman [b], Ralph HB Benedict [b], Robert Zivadinov [a,d,*]

[a] Buffalo Neuroimaging Analysis Center, Department of Neurology, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, State University of New York, Buffalo, NY, USA
[b] Jacobs Multiple Sclerosis Center, Department of Neurology, School of Medicine and Biomedical Sciences, University at Buffalo, State University of New York, Buffalo, NY, USA
[c] Center for Biomedical Imaging at Clinical Translational Science Institute, University at Buffalo, State University of New York, Buffalo, NY, USA
[d] IRCCS, Fondazione Don Carlo Gnocchi ONLUS, Milan, Italy

ARTICLE INFO

Keywords:
Multiple sclerosis
Quantifying pathology
Disability progression
Predictivity
T2-FLAIR
MRI

ABSTRACT

Background: Although quantitative measures from research-quality MRI provide a means to study multiple sclerosis (MS) pathology in vivo, these metrics are often unavailable in legacy clinical datasets.
Objective: To determine how well an automatically-generated quantitative snapshot of brain pathology, measured only on clinical routine T2-FLAIR MRI, can substitute for more conventional measures on research MRI in terms of capturing multi-factorial disease pathology and providing similar clinical relevance.
Methods: MRI with both research-quality sequences and conventional clinical T2-FLAIR was acquired for 172 MS patients at baseline, and neurologic disability was assessed at baseline and five-years later. Five measures (thalamus volume, lateral ventricle volume, medulla oblongata volume, lesion volume, and network efficiency) for quantifying disparate aspects of neuropathology from low-resolution T2-FLAIR were applied to predict standard research-quality MRI measures. They were compared in regard to association with future neurologic disability and disease progression over five years.
Results: The combination of the five T2-FLAIR measures explained most of the variance in standard research-quality MRI. T2-FLAIR measures were associated with neurologic disability and cognitive function five-years later ($R^2 = 0.279$, $p < 0.001$; $R^2 = 0.382$, $p < 0.001$), similar to standard research-quality MRI ($R^2 = 0.279$, $p < 0.001$; $R^2 = 0.366$, $p < 0.001$). They also similarly predicted disability progression over five years (%-correctly-classified = 69.8, $p = 0.034$), compared to standard research-quality MRI (%-correctly-classified = 72.4%, $p = 0.022$) in relapsing-remitting MS.
Conclusion: A set of five T2-FLAIR-only measures can substitute for standard research-quality MRI, especially in relapsing-remitting MS. When only clinical T2-FLAIR is available, it can be used to obtain substantially more quantitative information about brain pathology and disability than is currently standard practice.

1. Introduction

Multiple sclerosis (MS) is a disease of the central nervous system (CNS) associated with progressive neurologic symptoms (Scalfari et al., 2010). Although its hallmark is the presence of inflammatory demyelinating plaques in the white matter (WM), it is a highly multi-factorial disease including diffuse tissue abnormalities, pathological processes occurring directly in gray matter (GM), atrophy of both WM and GM, and spinal cord involvement (Zivadinov et al., 2016). The relationship between MS pathology and clinical outcomes is further complicated by the variable location of focal pathology within eloquent brain systems and tracts and by the potentially different staging of individual region involvement (Eshaghi et al., 2018).

Over the last decade or so, a variety of sophisticated neuroimaging acquisition protocols and analysis tools have been developed to quantify and study these various aspects of MS pathology and their clinical impacts. These include: quantitative lesion volumetry on T2-FLAIR images, often with reference to proton density, T2-weighted, and T1-weighted

* Corresponding author at: Buffalo Neuroimaging Analysis Center, 100 High Street, Buffalo, NY 14203, USA.
E-mail address: rzivadinov@bnac.net (R. Zivadinov).

https://doi.org/10.1016/j.nicl.2021.102705
Received 7 April 2021; Received in revised form 12 May 2021; Accepted 19 May 2021
Available online 24 May 2021

images, to elucidate focal WM pathology; diffusion tractography on multi-directional DWI imaging protocols to characterize structural connectivity; brain volumetry measures like SIENA on high-resolution 3D T1-weighted images to evaluate brain atrophy (Smith et al., 2002) tissue specific volumetry tools like SIENAX to measure WM and GM involvement separately on high-resolution 3D T1-weighted images (Smith et al., 2002); structure-specific volumetry tools like FIRST and FreeSurfer on high-resolution 3D T1-weighted images to identify specific regional involvement (Fischl et al., 2004); and multiparametric spinal cord analysis like the Spinal Cord Toolbox applied on cervical MRI to quantify spinal cord atrophy and tissue degeneration (Moccia et al., 2019). Unfortunately, though, these measures depend on a number of different MRI contrasts (notably including high-resolution 3D T1-weighted imaging). While these are often acquired in a research setting, they are frequently omitted in routine clinical research. A recent large, retrospective, multi-center study in a clinical routine found that the only common denominator was T2-FLAIR (Zivadinov et al., 2018). While some centers include more sequences, the situation is unlikely to change dramatically in the near future since recent consensus guidelines have made high-resolution 3D T1 optional (Sastre-Garriga et al., 2020) to balance time and expense constraints. Similarly, spinal cord imaging also remains less common, and diffusion imaging is rarely acquired (Rovira et al., 2015).

Given this background, there exists a very large amount of clinical routine imaging acquired in MS that does not currently contribute to basic or clinical research because it does not fit the tools available. This is unfortunate, because the large number of scans and subjects in these latent datasets could theoretically provide power to more broadly investigate disease heterogeneity, individual prognostic factors, real-world treatment response, and many other aspects of the disease. However, we hypothesize that in practice a substantial amount of the information that is traditionally obtained from a battery of research MRI images (high-resolution 3D T1, PD/T2, DWI, spinal cord imaging) can be inferred from a carefully selected set of measures on single T2-FLAIR brain images using modified techniques. While we do not expect these to be a replacement for those higher-fidelity measures, we do expect that with adequate tools a much better compromise can be reached than not using this widespread source of data at all for quantitative MS research.

Based on this background, we synthesized five methods for automatically and reliably processing T2-FLAIR MRI, selected based on their ability to provide complementary data reflecting the various known axes of MS pathology. They include NeuroSTREAM-based lateral ventricular volume (LVV) (Dwyer et al., 2017) as a proxy for global atrophy, DeepGRAI-based thalamus volume (Dwyer et al., 2021) as a proxy for regional GM atrophy, T2-FLAIR based medulla oblongata volume (MOV) as a proxy for spinal cord atrophy, salient central lesion volume (SCLV) (Dwyer et al., 2019) as a proxy for overall WM lesion burden, and NeMo-based network efficiency (Kuceyeski, Maruta, Relkin, & Raj, 2013; Rubinov & Sporns, 2010b) as a proxy for diffusion-based structural (dis)connectivity.

We expected that together these measures would capture a substantial amount of the pathophysiological variance usually identified by standard research-quality MRI metrics, would correlate strongly with neurologic disability, and would predict future disability progression well relative to standard research-quality MRI. To test these hypotheses, we applied these measures and the related standard measures to a large retrospective dataset with multimodal research-quality MRI images, clinical routine T2-FLAIR MRI images, and longitudinal clinical and cognitive data. Because the association between MRI and neurologic outcomes can vary by disease phenotype (Filippi et al., 2019), we also carried out our analyses using both the full cohort and stratified by disease sub-types (i.e. relapsing-remitting and progressive).

2. Methods

2.1. Participants

This retrospective study of prospectively collected data in the Cardiovascular Environmental and Genetics (CEG) (Kappus et al., 2015) study was approved by the Institutional Review Board (IRB) of the University at Buffalo. All study participants provided written informed consent. Included in this study were MRI data and clinical data at baseline and clinical data five years later from 172 MS patients who met inclusion and exclusion criteria. These criteria were: (a) diagnosis of clinically definite MS, relapsing-remitting (RRMS) or progressive (PMS), according to the McDonald criteria (Polman et al., 2011), or clinically isolated syn drome (CIS) (b) age 18 + years, (c) English fluency, (d) able to provide informed consent to procedures, (e) no neurologic disorder other than MS, (f) no psychiatric disorder other than mood or behavior change following onset of MS, (g) no history y of developmental disorder, (h) no history of substance abuse, past or present, (i) no motor or sensory defect that might interfere with cognitive test performance, (j) T2-FLAIR and high spatial resolution T1w imaging data available/passing quality inspection, and (k) neurologic examination completed at time of MRI and again five years later. If clinical relapse was detected during expected evaluation times, clinical and MRI examinations were re-scheduled to a time when participants were clinically stable.

2.2. Clinical progression

Neurologic disability was measured with the Expanded Disability Status Scale (EDSS) (Kurtzke, 1983), Nine-Hole-Peg Test (NHPT) (Feys et al., 2017), and the Timed 25 Foot Walk (T25W) (Goldman et al., 2013). The change in EDSS from baseline to follow-up was calculated and used to determine whether each person with MS exhibited disability progression (DP) over the five-year window. The DP converters were classified as those with EDSS change ≥ 1.5 if baseline EDSS was < 1.0, EDSS change of ≥ 1.0 if baseline EDSS was 1.0–5.5, and EDSS change of ≥ 0.5 if baseline EDSS ≥ 5.5. We also measured cognitive processing speed using the Symbol Digit Modalities Test (SDMT) (Smith, 1982). The NHPT, T25FW, and SDMT were available f or all study participants at the five-year follow-up, and for a smaller subset ($n = 45$) at baseline.

2.3. Image acquisition

MRI data were c ollected using a 3 T GE scanner. Because our aim was to determine how well a quantitative snapshot of brain pathology at one time can predict future disability progression, compared to standard research-quality MRI, we used only baseline MRI data. Imaging included a high spatial resolution 3D T1-weighted, 2D T2-FLAIR, and diffusion tensor imaging, 2D echo-planar imaging (EPI). For details, see Supplemental Materials.

2.4. Image processing

Five measures (thalamus volume, LVV, MOV, SCLV and network efficiency) were applied to autom atically quantify disparate aspects of brain pathology in MS from clinical quality T2-FLAIR MRI (Fig. 1). The measures from standard research-quality MRI for comparison included thalamus volume, brain volume, MOV, T2 lesion volu me (T2LV), T1 lesion volume (T1LV), diffusion-weighted network efficiency, and cortical GM volume (CGMV). For all T2-FLAIR processing protocols, see the Supp lemental materials for additional details.

2.5. Global and cortical bra in atrophy

Global/GM brain atrophy is well-established for its clinical relevance in MS (Bermel & Bakshi, 2006; Steenwijk et al., 2016). We therefore elected to capture this phenomenon on clinical T2-FLAIR and standard

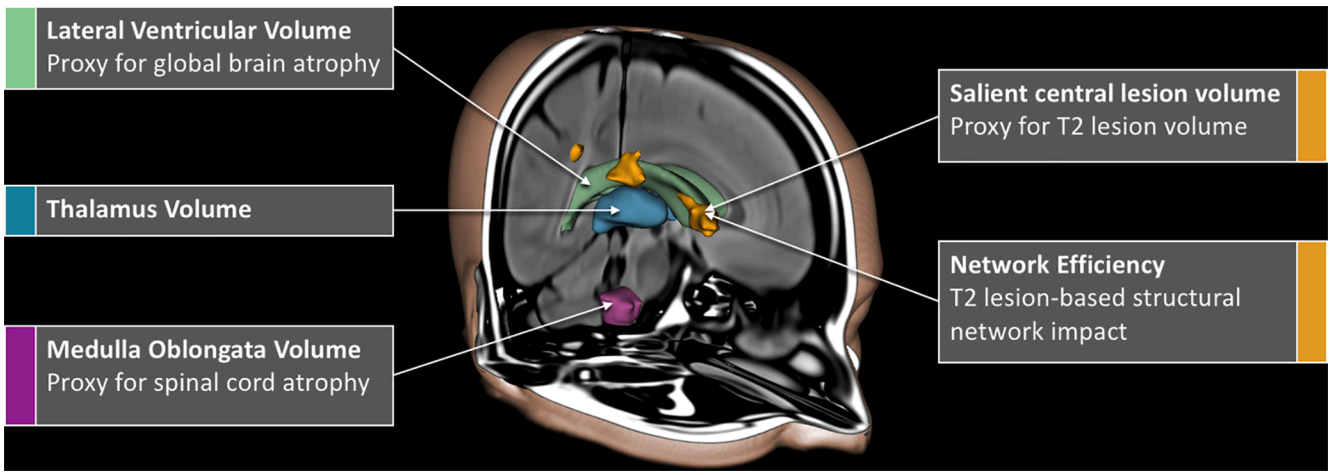

Fig. 1. T2-FLAIR image processing. Automated methods of measuring LVV, SCLV, MOV, thalamus volume, and network efficiency from clinical quality T2-FLAIR images were applied. Abbreviations: T2-FLAIR, T2-weighted fluid attenuated inversion recovery; LVV, lateral ventricular volume; SCLV, salient central lesion volume; MOV, medulla oblongata volume.

research-quality MRI. The standard research-quality MRI measure of global brain atrophy was the SIENAX (Smith et al., 2002) brain parenchymal volume (BPV), calculated using the high-resolution T1w image and normalized head size. The SIENAX measure for CGMV was also collected to assess cortical tissue loss.

The T2-FLAIR proxy measure for global brain atrophy, LVV, was acquired using the previously described Neurological Software Tool for Reliable Atrophy Measurement (NeuroSTREAM) (Dwyer et al., 2017). This tool performs automated processing including basic pre-processing, multi-atlas template based segmentation, level-set refinement, and partial volume estimation of the LVV. No proxy of CGMV was available on T2-FLAIR.

2.6. Regional subcortical gray matter atrophy

In addition to global brain atrophy, regional subcortical GM atrophy, thalamic atrophy specifically, has been established as particularly relevant to cognition and important for distinguishing prognosis and disease severity in MS (Azevedo et al., 2018; Houtchens et al., 2007; Zivadinov et al., 2013). We therefore elected to measure regional GM atrophy via a measure of thalamic volume. The standard research-quality MRI measure of thalamus volume was generated using FSL FIRST (Patenaude, Smith, Kennedy, & Jenkinson, 2011) on high-resolution T1w images.

The T2-FLAIR measure of thalamus volume was acquired using the previously described, Deep Gray Rating via Artificial Intelligence (DeepGRAI) method (Dwyer et al., 2021). This tool uses a semantic segmentation convolutional neural network architecture to segment the thalamus and measure thalamic partial volumes.

2.7. Indirect assessment of spinal cord atrophy

In addition to atrophy measures within the brain, we aimed to also capture proxies for atrophy in the spinal cord. This is because spinal cord atrophy progresses independently, occurs preferentially in PMS subtypes, and correlates well with physical disability (Song et al., 2020; Tsagkas et al., 2019; Zeydan et al., 2018). The standard research-quality MRI metric as a proxy for spinal cord atrophy was MOV, calculated using the FreeSurfer brainstem segmentation tool (Iglesias et al., 2015; Sander et al., 2019) on T1w images. FreeSurfer data was reviewed for quality and manual corrections (e.g. control points) were applied as needed.

The T2-FLAIR MOV was acquired as follows. T2-FLAIR images were non-linearly registered to a probabilistic medulla oblongata (MO) atlas. They were then deskulled and processed to obtain tissue and cerebrospinal fluid partial volume estimates (PVE). The tissue PVE, thresholded at 0.6, was used to mask the MO atlas, and then the absolute MOV was calculated. Adjusted volumes, used for the final analyses, were derived by accounting for the proportion of the MO atlas present after warping the T2-FLAIR to the template.

2.8. White matter lesion burden

As WM lesions are a hallmark of MS-related pathology (Lassmann, 2018), we aimed to capture the effects of WM damage from multiple perspectives. The standard research-quality MRI metric for WM lesion burden was T2 lesion volume (T2LV), measured from the T2-FLAIR images using a semi-automated edge-detection contouring/thresholding technique, with reference to PD/T2 images (Zivadinov et al., 2012). T2LV masks were reviewed for quality by a trained neurologist, and corrected as needed. A similar approach was also used on 2D T1w images to measure T1 black-hole lesion volume (T1LV).

The T2-FLAIR proxy measure for T2LV, SCLV, was acquired as previously described (Dwyer et al., 2019). No proxy of T1LV was available on T2-FLAIR. For the SCLV approach, images were bias-field corrected, and a fully-automated random-forest based lesion classifier was run to produce a lesion probability map. Voxels were retained if they were within 20 mm of the lateral ventricles (centrality) and were and brighter than central normal appearing brain tissue (salience).

2.9. White matter network disruption

We also set out to consider WM pathology from a network perspective, one that considers that manner WM damage alters network efficiency (Rubinov & Sporns, 2010a). The standard research-quality MRI metric used to investigate network disruption was global efficiency, derived from diffusion-weighted imaging (only available for 98 participants). For this measure, tract disruption was measured as previously described (Ashton et al., 2019) according to the intensity/severity of microstructural (fractional anisotropy) abnormality. This data was processed using the Network Modification tool (Kuceyeski et al., 2013) to derive 86x86 connectivity/disconnectivity matrices, for all GM regions in the 86-region FreeSurfer atlas. Then the Brain Connectivity Toolbox (Rubinov & Sporns, 2010b) was used to calculate global efficiency, weighted for the strength of connectivity/disconnectivity between GM region-pairs.

The T2-FLAIR measure of WM tract disruption network impact, global efficiency, was also calculated. For this, WM lesion masks were generated using the lesion segmentation tool (LST) and processed as in previous publications (Fuchs et al., 2018). Lesion-based tract disruption

between pairs of connected GM regions was determined using the network modification tool (Kuceyeski et al., 2013), resulting in the production of 86x86 connectivity/disconnectivity matrices, as above, and global network efficiency was calculated.

2.10. Statistical analysis

To evaluate the ability of the proposed five measure panel to replicate data that would normally be obtained from multimodal imaging, multiple regression was applied with each standard research-quality MRI measure as a dependent variable and the five automatically-generated T2-FLAIR measures as the independent variables. To understand the potential common underlying axes of pathology that the five separate measures might be reflecting, we further explored the statistical variance within/between the five T2-FLAIR measures using principal component analysis (R stats package) with automatic dimensionality reduction. This was repeated for the standard research-quality MRI measures.

To determine the extent to which the five T2-FLAIR measures relate to clinical outcomes, they were applied in regressions, predicting baseline and follow-up (5 years later) EDSS, and performance on the SDMT, T25FW, and NHPT. They were then applied in a binary logistic regression to predict DP over the five-year window. This was repeated with the standard research-quality MRI measures for comparison and was performed on the whole study sample, as well as in subsets for those with RRMS, PMS, and CIS. Results were not considered for samples of 20 or fewer subjects. For a direct comparison of the predictive ability of models, no false discovery rate correction was applied. All beta values reported were standardized and all R^2 values were adjusted for the number of factors in each model. P values < 0.05 were considered statistically significant.

2.11. Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

3. Results

In the sample (n = 172, mean age = 47.2, mean disease duration = 13.7), the mean time between initial data collection and follow-up was 5.4 ± 0.6 years. For demographic and clinical data, see Table 1.

In the regressions predicting each of the standard measures from the five T2-FLAIR measures (Table 2, Fig. 2), the combination of T2-FLAIR measures best predicted T2LV ($R^2 = 0.97$, $p < 0.001$) and thalamus volume ($R^2 = 0.90$, $p < 0.001$), with the lowest prediction for MOV ($R^2 = 0.476$, $p < 0.001$). See Table 2 for more details.

Table 1
Demographics for MS participants (N = 172) at baseline. All values represent mean ± SD unless otherwise specified. Abbreviations: SD, standard deviation, EDSS, Expanded Disability Status Scale; CIS, Clinically Isolated Syndrome; NHPT, nine-hole peg test; T25FW, Timed 25 Foot Walk; SDMT, Symbol Digit Modalities Test.

Age (years)	47.2 ± 11.1
Disease Duration (years)	13.7 ± 9.8
Female (n, %)	128, 74.4%
Disease Course (n, %)	
Relapsing-Remitting	100, 58.1%
Progressive	50, 29.1%
CIS	22, 12.8%
EDSS (Median, IQR)	2.5, 1.5–5.5
NHPT (s)	26.5 ± 14.8
T25FW (s)	9.5 ± 13.1
SDMT (raw score)	50.0 ± 14.8

Table 2
Predicting variance in standard research-quality MRI measures using all five T2-FLAIR MRI measures (β) and each individually (r). All β values reported are standardized and R^2 values are adjusted for the number of factors in each model. Abbreviations: T2-FLAIR, T2-weighted fluid attenuated inversion recovery; LVV, lateral ventricular volume; MOV, medulla oblongata volume; T2LV, T2-weighted lesion volume; SCLV, salient central lesion volume; T1LV, T1-weighted hypointense lesion volume; CGMV, neocortical gray matter volume; WM, white matter.

		Combined Model		Independent Correlations	
		β	p	r	P
Brain Volume	**LVV**	**−0.372**	**<0.001**	**−0.601**	**<0.001**
Combined Model	**Thalamus Volume**	**0.325**	**<0.001**	**0.659**	**<0.001**
$R^2 = 0.546$					
$p < 0.001$	**MOV**	0.111	0.104	**0.315**	**<0.001**
	SCLV	**−0.215**	**0.011**	**−0.557**	**<0.001**
	Network Efficiency	0.081	0.343	**−0.483**	**<0.001**
Thalamus Volume	**LVV**	−0.019	0.513	**−0.457**	**<0.001**
Combined Model	**Thalamus Volume**	**0.827**	**<0.001**	**0.946**	**<0.001**
$R^2 = 0.902$					
$p < 0.001$	**MOV**	**0.141**	**<0.001**	**0.419**	**<0.001**
	SCLV	0.059	0.121	**−0.574**	**<0.001**
	Network Efficiency	**−0.099**	**0.012**	**−0.419**	**<0.001**
MOV	**LVV**	0.043	0.529	0.005	0.938
Combined Model	**Thalamus Volume**	−0.020	0.835	**0.353**	**<0.001**
$R^2 = 0.476$					
$p < 0.001$	**MOV**	**0.722**	**<0.001**	**0.678**	**<0.001**
	SCLV	−0.018	0.841	*−0.144*	*0.059*
	Network Efficiency	0.071	0.440	−0.060	0.433
T2LV	**LVV**	0.017	0.296	**0.457**	**<0.001**
Combined Model	**Thalamus Volume**	−0.001	0.982	**−0.623**	**<0.001**
$R^2 = 0.970$					
$p < 0.001$	**MOV**	−0.001	0.939	**−0.269**	**<0.001**
	SCLV	**0.893**	**<0.001**	**0.982**	**<0.001**
	Network Efficiency	**0.107**	**<0.001**	**0.796**	**<0.001**
Network Efficiency *	**LVV**	0.009	0.918	**0.427**	**<0.001**
Combined Model	**Thalamus Volume**	**−0.449**	**<0.001**	**−0.708**	**<0.001**
$R^2 = 0.550$					
$p < 0.001$	**MOV**	−0.007	0.941	**−0.367**	**<0.001**
	SCLV	0.013	0.908	**0.624**	**<0.001**
	Network Efficiency	**0.333**	**0.001**	**0.684**	**<0.001**
T1LV	**LVV**	−0.097	0.142	**0.287**	**0.001**
Combined Model	**Thalamus Volume**	−0.121	0.197	**−0.485**	**<0.001**
$R^2 = 0.506$					
$p < 0.001$	**MOV**	0.033	0.642	*−0.137*	*0.071*
	SCLV	0.090	0.305	**0.586**	**<0.001**
	Network Efficiency	**0.634**	**<0.001**	**0.714**	**<0.001**
CGMV	**LVV**	**0.156**	**0.011**	**−0.202**	**0.008**
Combined Model	**Thalamus Volume**	**0.739**	**<0.001**	**0.727**	**<0.001**
$R^2 = 0.565$					
$p < 0.001$	**MOV**	**0.147**	**0.027**	**0.334**	**<0.001**
	SCLV	0.085	0.291	**−0.371**	**<0.001**
	Network Efficiency	−0.046	0.570	**−0.379**	**<0.001**

*DWI data were only available for 98 study participants.

Principal component analysis of the T2-FLAIR MRI measures identified five statistical components (Fig. 3), with the first two components (eigenvalues = 2.83 & 1.01) explaining 56.7% and 20.1% (76.8% together) of the total variance (Fig. 3A). The strongest contributors to the first component (dimension 1) were SCLV, network efficiency, and thalamus volume (Fig. 3B), largely representing WM and network disruption pathology. In contrast, the strongest contributors to the second component (dimension 2) were LVV and MOV (Fig. 3C), largely reflecting global CNS atrophy. The distribution of individuals across

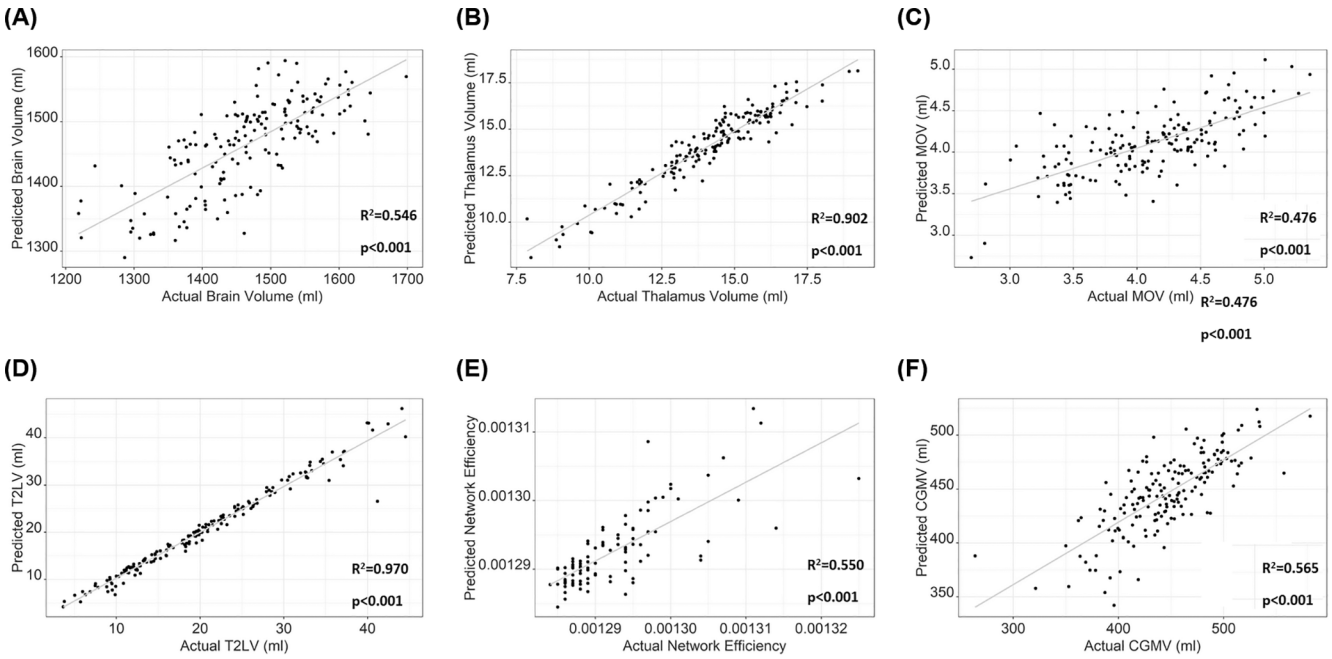

Fig. 2. Regression-based predictions of standard research-quality MRI metrics from the five automatically-generated T2-FLAIR processing measures. Abbreviations: T2-FLAIR, T2-weighted fluid attenuated inversion recovery; MOV, medulla oblongata volume; T2LV, T2 lesion volume *DWI data were only available for 98 study participants.

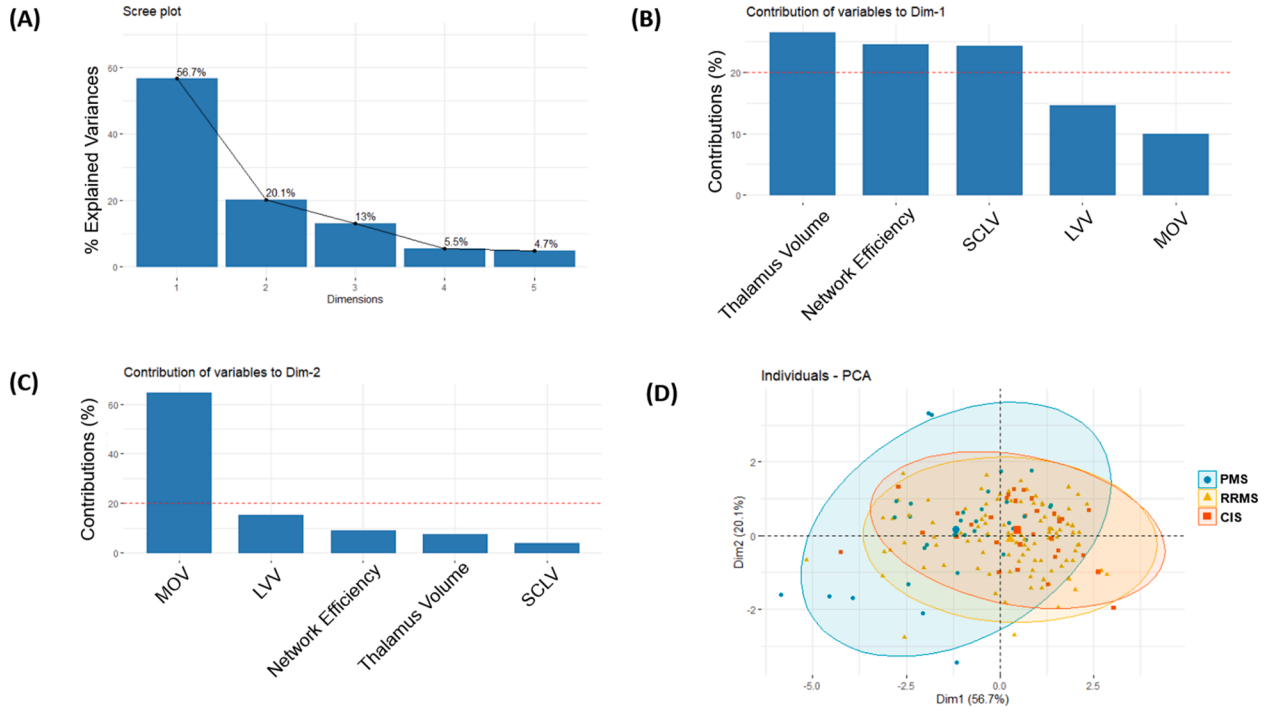

Fig. 3. Principal component analysis of the five automated T2-FLAIR processing measures (A-D). Two components explained most of the overall variance in the T2-FLAIR measures (A). The first component was comprised largely of measures relating to white matter damage and thalamus volume (B). The second component was comprised largely of measures relating to central nervous system atrophy (C). Subjects with progressive MS exhibited a wider degree of variability across the second component (D). Abbreviations: T2-FLAIR, T2-weighted fluid attenuated inversion recovery; Dim, dimension; SCLV, salient central lesion volume; LVV, lateral ventricular volume; MOV, medulla oblongata volume; RRMS, relapsing-remitting MS; PMS, progressive MS; CIS, clinically isolated syndrome.

these two major components varied by disease-course (Fig. 3D) such that individuals with PMS exhibited greater variability over dimension 2 (global CNS atrophy).

The principal component analysis of the standard research-quality MRI measures resulted in the creation of seven statistical components (Figure e-1), with the first two components (eigenvalues = 4.17 & 1.12) explaining 59.7% and 26.0% (75.7% together) of the total variance (Figure e-1A). The strongest contributors to the first component (dimension 1) were thalamus volume, BPV, and CGMV (Figure e-1B), largely representing brain atrophy. In contrast, the strongest

contributors to the second component (dimension 2) were T1LV, T2LV, and MOV (Figure e-1C), largely representing WM damage and spinal cord atrophy.

In the full study sample, the five baseline T2-FLAIR processing measures were significantly associated with EDSS ($R^2 = 0.229$, $p < 0.001$) and SDMT ($R^2 = 0.441$, $p < 0.001$) at baseline (table e-1). The five baseline T2-FLAIR measures were also significantly associated with EDSS ($R^2 = 0.279$, $p < 0.001$) and SDMT ($R^2 = 0.382$, $p < 0.001$) five years later (Table 3). These results were similar for the standard research-quality MRI at baseline ($R^2 = 0.239$, $p < 0.001$ for EDSS; $R^2 = 0.369$, $p < 0.001$ for SDMT, table e-2), as well as five years later ($R^2 = 0.279$, $p < 0.001$ for SDMT; $R^2 = 0.366$, $p < 0.001$, table e-3).

Out of the full study sample, 33.7% (n = 58) experienced DP over the five-year window. The binary logistic regression predicting DP from baseline T2-FLAIR MRI measures correctly classified 69.8% of the RRMS study sample (n = 99, Nagelkerke $R^2 = 0.145$, $p = 0.034$, Table 4). The accuracy of this predictive model was comparable to the standard research-quality MRI measures (%correctly-classified = 72.4%, Nagelkerke-$R^2 = 0.196$, $p = 0.022$).

4. Discussion

In the present study, we confirmed the utility of five automated measures applicable to clinical routine T2-FLAIR image alone for capturing a substantial portion of the information that would otherwise normally be acquired via the kind of research-quality multi-modal imaging that is frequently unavailable in clinical routine. Based on the idea that these complementary measures, when synthesized together, can extract key information about the underlying axes of MS disease pathophysiology, we evaluated their ability to predict conventional measures and to relate to clinical outcomes. Together, the five measures (LVV, thalamus volume, MOV, SCLV and network efficiency) predicted a majority of the variance in the standard research-quality MRI measures. They also correlated with neurologic disability and cognitive function five years later and predicted disability progression over five-years, performing well relative to standard research-quality MRI measures (69% accuracy versus 72% in RRMS). The five T2-FLAIR measures at baseline related well to SDMT performance five years later in both the RRMS and PMS groups, showing that the measures investigated are robust to differences in MS sub-types when addressing cognitive function. In comparison, these associations were only significant for RRMS when addressing EDSS five years later and predictions of DP. This may be in part due to the lower PMS sample size and may also suggest that the measures addressed in this study show stronger associations with neurologic disability in RRMS patients relative to PMS.

The associations of the T2-FLAIR measures with future neurologic disability, cognition, and DP over five years, were comparable to standard research-quality MRI investigated in previous studies (Bakshi et al., 2008; Calabrese et al., 2009; Dekker et al., 2019; Eijlers, Meijer, Van Geest, Geurts, & Schoonheim, 2018; Fuchs et al., 2019; Jakimovski et al., 2018; Uher et al., 2017). With this, and because T2-FLAIR imaging is wide-spread in clinical evaluation, these metrics are a strong candidate for quantitative evaluation of brain pathology in future clinical science (Dwyer et al., 2019, 2017; R. Zivadinov et al., 2018, 2016). Because T2-FLAIR imaging has been used in clinical trials in

MS for over 20 years, these methods provide valuable ways of exploring legacy data from historical datasets.

In addition to these key findings, principal component analysis of the five T2-FLAIR measures revealed two major factors explaining most of the variance among them. The first component was largely composed of variance from measures of WM damage and thalamic atrophy (SCLV, network efficiency, thalamus volume), whereas the second component was more heavily related to global CNS atrophy (MOV and LVV). These results suggest there are two major axes of disease progression captured by our combined T2-FLAIR measures, WM pathology and CNS atrophy. This is further validated by the selective associations between each of the T2-FLAIR measures and distinct neurological outcomes, such as between thalamus volume and cognitive processing speed (SDMT) and between MOV and physical disability (EDSS). These selective correlations also have strong face validity, as thalamic integrity has long been strongly associated with cognition (Houtchens et al., 2007) and spinal cord integrity is strongly associated with physical disability in MS (Kearney et al., 2015). Interestingly, the strongest statistical contributors to each principle component differed for the T2-FLAIR measures relative to the research quality MRI measures. As such, the patterns of variance in MS sub-types (RRMS versus PMS versus CIS) across the major components differed between the research and clinical data.

Our work does not replace MRI measures usually obtained on conventional MRI sequences by standard research-quality MRI analyses, which should be considered as more accurate. Nevertheless, it opens the

Table 3
Associations between the five T2-FLAIR MRI measures at baseline and neurologic disability five years later. All β values reported are standardized and R^2 values are adjusted for the number of factors in each model. Abbreviations: EDSS, Expanded Disability Status Scale; T2-FLAIR, T2-weighted fluid attenuated inversion recovery; NHPT, nine-hole peg test; T25FW, Timed 25-Foot Walk; SDMT, Symbol Digit Modalities Test; LVV, lateral ventricular volume; MOV, medulla oblongata volume; SCLV, salient central lesion volume; RRMS, relapsing-remitting MS; PMS, progressive MS; CIS, clinically isolated syndrome.

		All		RRMS		PMS		CIS	
		B	p	β	p	β	p	B	p
EDSS	**LVV**	0.110	0.174	0.079	0.460	0.164	0.379	0.201	0.447
All: n = 172, **$R^2 = 0.279$, $p < 0.001$**	**Thalamus Volume**	−0.018	0.875	−0.129	0.401	0.079	0.746	−0.038	0.906
RRMS: n = 100, **$R^2 = 0.164$, $p < 0.001$**	**MOV**	**−0.288**	**0.001**	−0.051	0.663	*−0.375*	*0.070*	0.067	0.818
PMS: n = 50, $R^2 = 0.063$, p = 0.177	**SCLV**	**0.214**	**0.045**	0.200	0.163	−0.154	0.478	0.460	0.446
CIS: n = 22, $R^2 < 0.001$, p = 0.770	**Network Efficiency**	0.134	0.216	0.107	0.447	0.197	0.440	−0.272	0.657
SDMT	**LVV**	−0.107	0.215	−0.037	0.748	−0.163	0.358	–	–
All: n = 123, **$R^2 = 0.382$, $p < 0.001$**	**Thalamus Volume**	**0.397**	**0.002**	0.287	0.110	**0.682**	**0.001**	–	–
RRMS: n = 72, **$R^2 = 0.227$, $p < 0.001$**	**MOV**	0.083	0.403	0.065	0.633	−0.110	0.533	–	–
PMS: n = 40, **$R^2 = 0.501$, $p < 0.001$**	**SCLV**	−0.042	0.709	−0.185	0.251	0.213	0.250	–	–
CIS: n = 11, R^2 = NA, p = NA	**Network Efficiency**	−0.170	0.149	−0.057	0.712	−0.218	0.334	–	–
NHPT	**LVV**	*0.174*	*0.062*	0.187	0.120	0.210	0.314	–	–
All: n = 116, **$R^2 = 0.202$, $p < 0.001$**	**Thalamus Volume**	−0.010	0.944	0.020	0.913	−0.103	0.691	–	–
RRMS: n = 71, **$R^2 = 0.094$, $p = 0.044$**	**MOV**	*−0.199*	*0.082*	−0.020	0.886	−0.236	0.337	–	–
PMS: n = 34, *$R^2 = 0.191$, $p = 0.064$*	**SCLV**	**0.440**	**0.002**	**0.354**	**0.036**	**1.010**	**0.007**	–	–
CIS: n = 11, R^2 = NA, p = NA	**Network Efficiency**	*−0.238*	*0.092*	−0.096	0.552	**−1.094**	**0.009**	–	–
T25FW	**LVV**	0.133	0.209	0.062	0.649	**0.534**	**0.023**	–	–
All: n = 110, **$R^2 = 0.127$, $p = 0.002$**	**Thalamus Volume**	0.097	0.535	0.193	0.331	−0.329	0.277	–	–
RRMS: n = 70, $R^2 = 0.035$, p = 0.203	**MOV**	**−0.299**	**0.020**	−0.194	0.218	−0.254	0.359	–	–
PMS: n = 28, **$R^2 = 0.294$, $p = 0.037$**	**SCLV**	0.125	0.433	0.188	0.323	0.037	0.912	–	–
CIS: n = 12, R^2 = NA, p = NA	**Network Efficiency**	0.168	0.306	0.165	0.375	−0.105	0.784	–	–

Table 4
Predicting disability progression (DP) over 5 years from the baseline T2-FLAIR processing measures and standard research-quality MRI measures. Abbreviations: %CC, Percent correctly classified; RRMS, relapsing-remitting multiple sclerosis; PMS, progressive multiple sclerosis; CIS, clinically isolated syndrome; T2-FLAIR, T2-weighted fluid attenuated inversion recovery; DP, disability progression; RRMS, relapsing-remitting MS; PMS, progressive MS; CIS, clinically isolated syndrome.

		All (n = 166) %DP = 33.7		**RRMS** (n = 99) %DP = 32.0		**PMS** (n = 47) %DP = 40.8		**CIS** (n = 20) %DP = 27.3	
		β	**p**	**β**	**p**	**β**	**p**	**B**	**p**
T2-FLAIR Model	**LVV**	−0.052	0.561	−0.001	0.993	−0.132	0.491	–	–
All: %CC = 66.9, R^2 = 0.085, p = 0.072	**Thalamus Volume**	0.039	0.752	−0.048	0.760	*0.436*	*0.087*	–	–
RRMS: %CC = 69.8, R^2 = 0.145, p = 0.034	**MOV**	−0.044	0.657	−0.080	0.517	−0.177	0.405	–	–
PMS: %CC = 65.9, R^2 = 0.101, p = 0.531	**SCLV**	*−0.196*	*0.095*	−0.181	0.230	−0.151	0.504	–	–
CIS: %CC = NA, R^2 = NA, p = NA	**Network Efficiency**	−0.068	0.584	−0.122	0.411	−0.293	0.460	–	–
Standard Research-Quality MRI Model*	**BPV**	0.230	0.206	0.202	0.341	0.215	0.628	–	–
All: %CC = 69.7, R^2 = 0.100, p = 0.098	**Thalamus Volume**	0.109	0.515	0.002	0.992	0.293	0.360	–	–
RRMS: %CC = 72.4, R^2 = 0.196, p = 0.022	**MOV**	0.061	0.485	0.061	0.567	0.085	0.643	–	–
PMS: %CC = 71.1, R^2 = 0.075, p = 0.864	**T2LV**	*0.209*	*0.052*	*0.235*	*0.066*	0.119	0.632	–	–
CIS: %CC = NA, R^2 = NA, p = NA	**T1LV**	0.033	0.727	0.104	0.402	0.046	0.816	–	–
	NCV	0.092	0.494	0.273	0.084	0.253	0.535	–	–

*DWI data were only available for 98 study participants and therefore network efficiency was not used in this model.

door to better use of latent data for understanding predictors of clinical disability, especially in the real-world treatment related studies of MS patients. It also potentially contributes to foster research opportunities for underprivileged areas that did not have access to research quality MRI acquisition and analyses as part of clinical routine imaging. Therefore, the current set of carefully-chosen five T2-FLAIR MRI metrics provides a complementary approach in cases where only clinical T2-FLAIR is available, providing substantially more quantitative information about brain pathology and disability than is currently standard practice in MS.

This work has a number of limitations. First, although we investigated 172 patients with MS, larger, multicenter datasets are more suitable to fully capture the heterogeneity of MS. Second, external validation with larger and more representative samples will provide additional insights into the generalizability and reproducibility of our results. Because, T2-FLAIR image processing and evaluation was only completed for data collected from a single site, using consistent MRI scan protocols, future work is needed to fully evaluate the relevance of these measures to clinical outcomes in multi-center studies, using a variety of acquisition protocols. We further acknowledge that despite our promising results, predictivity models in patients with progressive MS would need more substantial investigation. Image processing was completed using in-house systems. Remotely accessible cloud computing would allow for a more distributable software package. In addition, our battery of automated T2-FLAIR measures did not include quantification of CGMV and T1LV. Future work will benefit from developing a more robust and reliable method for providing such measures. Future work should also consider the use of these five T2-FLAIR brain measures for evaluating longitudinal changes and whether such changes are associated with accumulation of neurologic dysfunction, as well as impact of different treatment strategies on these measures.

5. Conclusion

The present study shows that a substantial amount of the information usually obtained with multi-modal research imaging in MS can be reasonably extracted with a set of five complementary measures that can be performed fully automatically on clinical routine T2-FLAIR alone. It also confirms that this data remains clinically relevant. This panel of measures therefore provides a practical avenue for quantitative MS imaging research in settings where multi-modal imaging is not available, including large latent legacy datasets.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Tom Fuchs, Dejan Jakimovski, Niels Bergsland, and Deepa Ramasamy have nothing to declare. Bianca Weinstock-Guttman received honoraria as a speaker and/or as a consultant for Biogen Idec, Teva Pharmaceuticals, EMD Serono, Genzyme, Sanofi, Genentech, Novartis, Mallinckrodt, Celgene, Abbvie and Acorda. Dr Weinstock-Guttman received research funds from Biogen Idec, Teva Pharmaceuticals, EMD Serono, GenzymeGenentech, Sanofi, Novartis, and Acorda. Robert Zivadinov received personal compensation from EMD Serono, Sanofi, Novartis, Bristol Myers Squibb, Novartis and Key Stone Heart for speaking and consultant fees. He received financial support for research activities from Sanofi, Novartis, Bristol Myers Squibb, Mapi Pharma, Keystone Heart, V-WAVE Medical, Boston Scientific and Protembo. Michael G. Dwyer has received consultant fees from Keystone Heart and research grant support from Bristol Myers Squibb and Keystone Heart.

Acknowledgment

The method for medulla oblongata volume measurement was developed with partial funding support by Mark Diamond Research Fund, Grant ID# FA-10-12.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.nicl.2021.102705.

References

Ashton, K., Fuchs, T., Oship, D., Zivadinov, R., Weinstock-Guttman, B., Benedict, R., Dwyer, M., 2019. 5-year conversion to clinically diagnosed depression in people with multiple sclerosis is predicted by baseline tract damage in fronto-parietal network. Multiple Sclerosis J. 25, 427.

Azevedo, C.J., Cen, S.Y., Khadka, S., Liu, S., Kornak, J., Shi, Y., Zheng, L., Hauser, S.L., Pelletier, D., 2018. Thalamic atrophy in multiple sclerosis: A magnetic resonance imaging marker of neurodegeneration throughout disease. Ann. Neurol. 83 (2), 223–234. https://doi.org/10.1002/ana.25150.

Bakshi, R., Neema, M., Healy, B.C., Liptak, Z., Betensky, R.A., Buckle, G.J., Gauthier, S. A., Stankiewicz, J., Meier, D., Egorova, S., Arora, A., Guss, Z.D., Glanz, B., Khoury, S. J., Guttmann, C.R.G., Weiner, H.L., 2008. Predicting clinical progression in multiple sclerosis with the magnetic resonance disease severity scale. Arch. Neurol. 65 (11), 1449. https://doi.org/10.1001/archneur.65.11.1449.

Bermel, R.A., Bakshi, R., 2006. The measurement and clinical relevance of brain atrophy in multiple sclerosis. Lancet Neurol. 5 (2), 158–170. https://doi.org/10.1016/S1474-4422(06)70349-0.

Calabrese, M., Agosta, F., Rinaldi, F., Mattisi, I., Grossi, P., Favaretto, A., Atzori, M., Bernardi, V., Barachino, L., Rinaldi, L., Perini, P., Gallo, P., Filippi, M., 2009. Cortical lesions and atrophy associated with cognitive impairment in relapsing-remitting multiple sclerosis. Arch. Neurol. 66 (9) https://doi.org/10.1001/archneurol.2009.174.

Dekker, I., Eijlers, A.J.C., Popescu, V., Balk, L.J., Vrenken, H., Wattjes, M.P., Uitdehaag, B.M.J., Killestein, J., Geurts, J.J.G., Barkhof, F., Schoonheim, M.M.,

2019. Predicting clinical progression in multiple sclerosis after 6 and 12 years. Eur. J. Neurol. 26 (6), 893–902. https://doi.org/10.1111/ene.2019.26.issue-610.1111/ene.13904.

Dwyer, M.G., Bergsland, N., Ramasamy, D.P., Weinstock-Guttman, B., Barnett, M.H., Wang, C., Tomic, D., Silva, D., Zivadinov, R., 2019. Salient central lesion volume: a standardized novel fully automated proxy for brain FLAIR lesion volume in multiple sclerosis. J. Neuroimaging 29 (5), 615–623. https://doi.org/10.1111/jon.v29.510.1111/jon.12650.

Dwyer, M.G., Silva, D., Bergsland, N., Horakova, D., Ramasamy, D., Durfee, J., Vaneckova, M., Havrdova, E., Zivadinov, R., 2017. Neurological software tool for reliable atrophy measurement (NeuroSTREAM) of the lateral ventricles on clinical-quality T2-FLAIR MRI scans in multiple sclerosis. NeuroImage Clin. 15, 769–779. https://doi.org/10.1016/j.nicl.2017.06.022.

Dwyer, M., Lyman, C., Ferrari, H., Bergsland, N., Fuchs, T.A., Jakimovski, D., Schweser, F., Weinstock-Guttmann, B., Benedict, R.H.B., Riolo, J., Silva, D., Zivadinov, R., 2021. DeepGRAI (Deep Gray Rating via Artificial Intelligence): fast, feasible, and clinically relevant thalamic atrophy measurement on clinical quality T2-FLAIR MRI in multiple sclerosis. NeuroImage Clin. 30, 102652. https://doi.org/10.1016/j.nicl.2021.102652.

Eijlers, A.J.C., Meijer, K.A., van Geest, Q., Geurts, J.J.G., Schoonheim, M.M., 2018. Determinants of cognitive impairment in patients with multiple sclerosis with and without atrophy. Radiology 288 (2), 544–551. https://doi.org/10.1148/radiol.2018172808.

Eshaghi, A., Prados, F., Brownlee, W.J., Altmann, D.R., Tur, C., Cardoso, M.J., De Angelis, F., van de Pavert, S.H., Cawley, N., De Stefano, N., Stromillo, M.L., Battaglini, M., Ruggieri, S., Gasperini, C., Filippi, M., Rocca, M.A., Rovira, A., Sastre-Garriga, J., Vrenken, H., Leurs, C.E., Killestein, J., Pirpamer, L., Enzinger, C., Ourselin, S., Wheeler-Kingshott, C.A.M.G., Chard, D., Thompson, A.J., Alexander, D. C., Barkhof, F., Ciccarelli, O., 2018. Deep gray matter volume loss drives disability worsening in multiple sclerosis. Ann. Neurol. 83 (2), 210–222. https://doi.org/10.1002/ana.25145.

Feys, P., Lamers, I., Francis, G., Benedict, R., Phillips, G., LaRocca, N., Hudson, L.D., Rudick, R., 2017. The nine-hole peg test as a manual dexterity performance measure for multiple sclerosis. Multiple Sclerosis (Houndmills, Basingstoke, England) 23 (5), 711–720. https://doi.org/10.1177/1352458517690824.

Filippi, M., Brück, W., Chard, D., Fazekas, F., Geurts, J.J.G., Enzinger, C., Hametner, S., Kuhlmann, T., Preziosa, P., Rovira, À., Schmierer, K., Stadelmann, C., Rocca, M.A., 2019. Association between pathological and MRI findings in multiple sclerosis. Lancet Neurol. 18 (2), 198–210. https://doi.org/10.1016/S1474-4422(18)30451-4.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D.H., Dale, A. M., 2004. Automatically parcellating the human cerebral cortex. Cereb. Cortex 14 (1), 11–22. https://doi.org/10.1093/cercor/bhg087.

Fuchs, T.A., Dwyer, M.G., Kuceyeski, A., Choudhery, S., Carolus, K., Li, X., Mallory, M., Weinstock-Guttman, B., Jakimovski, D., Ramasamy, D., Zivadinov, R., Benedict, R.H. B., 2018. White matter tract network disruption explains reduced conscientiousness in multiple sclerosis. Hum. Brain Mapp. 39 (9), 3682–3690. https://doi.org/10.1002/hbm.v39.910.1002/hbm.24203.

Fuchs, T.A., Benedict, R.H.B., Bartnik, A., Choudhery, S., Li, X., Mallory, M., Oship, D., Yasin, F., Ashton, K., Jakimovski, D., Bergsland, N., Ramasamy, D.P., Weinstock-Guttman, B., Zivadinov, R., Dwyer, M.G., 2019. Preserved network functional connectivity underlies cognitive reserve in multiple sclerosis. Hum. Brain Mapp. 40 (18), 5231–5241. https://doi.org/10.1002/hbm.v40.1810.1002/hbm.24768.

Goldman, M.D., Motl, R.W., Scagnelli, J., Pula, J.H., Sosnoff, J.J., Cadavid, D., 2013. Clinically meaningful performance benchmarks in MS: Timed 25-Foot Walk and the real world. Neurology 81 (21), 1856–1863. https://doi.org/10.1212/01.wnl.0000436065.97642.d2.

Houtchens, M.K., Benedict, R.H.B., Killiany, R., Sharma, J., Jaisani, Z., Singh, B., Weinstock-Guttman, B., Guttmann, C.R.G., Bakshi, R., 2007. Thalamic atrophy and cognition in multiple sclerosis. Neurology 69 (12), 1213–1223. https://doi.org/10.1212/01.wnl.0000276992.17011.b5.

Iglesias, J.E., Van Leemput, K., Bhatt, P., Casillas, C., Dutt, S., Schuff, N., Truran-Sacrey, D., Boxer, A., Fischl, B., 2015. Bayesian segmentation of brainstem structures in MRI. NeuroImage 113, 184–195.

Jakimovski, D., Weinstock-Guttman, B., Hagemeier, J., Vaughn, C.B., Kavak, K.S., Gandhi, S., Bennett, S.E., Fuchs, T.A., Bergsland, N., Dwyer, M.G., Benedict, R.H.B., Zivadinov, R., 2018. Walking disability measures in multiple sclerosis patients: correlations with MRI-derived global and microstructural damage. J. Neurol. Sci. 393, 128–134. https://doi.org/10.1016/j.jns.2018.08.020.

Kappus, N., Weinstock-Guttman, B., Hagemeier, J., Kennedy, C., Melia, R., Carl, E., Others. (2015). Cardiovascular risk factors are associated with increased lesion burden and brain atrophy in multiple sclerosis. J Neurol Neurosurg Psychiatry, jnnp——2014.

Kearney, H., Schneider, T., Yiannakas, M.C., Altmann, D.R., Wheeler-Kingshott, C.A.M., Ciccarelli, O., Miller, D.H., 2015. Spinal cord grey matter abnormalities are associated with secondary progression and Physical disability in multiple sclerosis. J. Neurol. Neurosurg. Psychiatry 86 (6), 608–614. https://doi.org/10.1136/jnnp-2014-308241.

Kuceyeski, A., Maruta, J., Relkin, N., Raj, A., 2013. The Network Modification (NeMo) Tool: elucidating the effect of white matter integrity changes on cortical and subcortical structural connectivity. Brain Connect. 3 (5), 451–463.

Kurtzke, J. F. (1983). Rating neurologic impairment in multiple sclerosis : An expanded disability status scale (EDSS). 0(November), 1444–1453.

Lassmann, H., 2018. Multiple sclerosis pathology. Cold Spring Harbor Perspectives in Medicine 8 (3), a028936. https://doi.org/10.1101/cshperspect.a028936.

Moccia, M., Prados, F., Filippi, M., Rocca, M.A., Valsasina, P., Brownlee, W.J., Zecca, C., Gallo, A., Rovira, A., Gass, A., Palace, J., Lukas, C., Vrenken, H., Ourselin, S., Gandini Wheeler-Kingshott, C.A.M., Ciccarelli, O., Barkhof, F., 2019. Longitudinal spinal cord atrophy in multiple sclerosis using the generalized boundary shift integral. Ann. Neurol. 86 (5), 704–713. https://doi.org/10.1002/ana.v86.510.1002/ana.25571.

Patenaude, B., Smith, S.M., Kennedy, D.N., Jenkinson, M., 2011. A Bayesian model of shape and appearance for subcortical brain segmentation. NeuroImage 56 (3), 907–922. https://doi.org/10.1016/j.neuroimage.2011.02.046.

Polman, C.H., Reingold, S.C., Banwell, B., Clanet, M., Cohen, J.A., Filippi, M., Fujihara, K., Havrdova, E., Hutchinson, M., Kappos, L., Lublin, F.D., Montalban, X., O'Connor, P., Sandberg-Wollheim, M., Thompson, A.J., Waubant, E., Weinshenker, B., Wolinsky, J.S., 2011. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Ann. Neurol. 69 (2), 292–302.

Rovira, À., Wattjes, M. P., Tintoré, M., Tur, C., Yousry, T. A., Sormani, M. P., … Montalban, X. (2015). EVIDENCE-BASED GUIDELINES MAGNIMS consensus guidelines on the use of MRI in multiple sclerosis — clinical implementation in the diagnostic process. 11(August). https://doi.org/10.1038/nrneurol.2015.106.

Rubinov, M., Sporns, O., 2010. Complex network measures of brain connectivity: uses and interpretations. NeuroImage 52 (3), 1059–1069. https://doi.org/10.1016/j.neuroimage.2009.10.003.

Sander, L., Pezold, S., Andermatt, S., Amann, M., Meier, D., Wendebourg, M.J., Sinnecker, T., Radue, E.-W., Naegelin, Y., Granziera, C., Kappos, L., Wuerfel, J., Cattin, P., Schlaeger, R., 2019. Accurate, rapid and reliable, fully automated MRI brainstem segmentation for application in multiple sclerosis and neurodegenerative diseases. Hum. Brain Mapp. 40 (14), 4091–4104. https://doi.org/10.1002/hbm.v40.1410.1002/hbm.24687.

Sastre-Garriga, J., Pareto, D., Battaglini, M., Rocca, M.A., Ciccarelli, O., Enzinger, C., Wuerfel, J., Sormani, M.P., Barkhof, F., Yousry, T.A., De Stefano, N., Tintoré, M., Filippi, M., Gasperini, C., Kappos, L., Río, J., Frederiksen, J., Palace, J., Vrenken, H., Montalban, X., Rovira, À., 2020. MAGNIMS consensus recommendations on the use of brain and spinal cord atrophy measures in clinical practice. Nat. Rev. Neurol. 16 (3), 171–182. https://doi.org/10.1038/s41582-020-0314-x.

Scalfari, A., Neuhaus, A., Degenhardt, A., Rice, G. P., Muraro, P. A., Daumer, M., & Ebers, G. C. (2010). The natural history of multiple sclerosis, a geographically based study 10: Relapses and long-term disability. Brain. https://doi.org/10.1093/brain/awq118.

Smith, A., 1982. Symbol Digit Modalities Test (SDMT) Manual (revised). Western Psychological Services, Los Angeles.

Smith, S.M., Zhang, Y., Jenkinson, M., Chen, J., Matthews, P.M., Federico, A., De Stefano, N., 2002. Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. NeuroImage 17 (1), 479–489.

Song, X., Li, D., Qiu, Z., Su, S., Wu, Y., Wang, J., Liu, Z., Dong, H., 2020. Correlation between EDSS scores and cervical spinal cord atrophy at 3T MRI in multiple sclerosis: a systematic review and meta-analysis. Multiple Sclerosis Related Disord. 37, 101426. https://doi.org/10.1016/j.msard.2019.101426.

Steenwijk, M.D., Geurts, J.J.G., Daams, M., Tijms, B.M., Wink, A.M., Balk, L.J., Tewarie, P.K., Uitdehaag, B.M.J., Barkhof, F., Vrenken, H., Pouwels, P.J.W., 2016. Cortical atrophy patterns in multiple sclerosis are non-random and clinically relevant. Brain 139 (1), 115–126. https://doi.org/10.1093/brain/awv337.

Tsagkas, C., Magon, S., Gaetano, L., Pezold, S., Naegelin, Y., Amann, M., Stippich, C., Cattin, P., Wuerfel, J., Bieri, O., Sprenger, T., Kappos, L., Parmar, K., 2019. Preferential spinal cord volume loss in primary progressive multiple sclerosis. Multiple Sclerosis J. 25 (7), 947–957. https://doi.org/10.1177/1352458518775006.

Uher, T., Vaneckova, M., Sobisek, L., Tyblova, M., Seidl, Z., Krasensky, J., Ramasamy, D., Zivadinov, R., Havrdova, E., Kalincik, T., Horakova, D., 2017. Combining clinical and magnetic resonance imaging markers enhances prediction of 12-year disability in multiple sclerosis. Multiple Sclerosis 23 (1), 51–61. https://doi.org/10.1177/1352458516642314.

Zeydan, B., Gu, X., Atkinson, E. J., Keegan, B. M., Weinshenker, B. G., Tillema, J. M., … Kantarci, O. H. (2018). Cervical spinal cord atrophy: An earlymarker of progressiveMSonset. Neurology: Neuroimmunology and NeuroInflammation. https://doi.org/10.1212/NXI.0000000000000435.

Zivadinov, R., Bergsland, N., Korn, J.R., Dwyer, M.G., Khan, N., Medin, J., Price, J.C., Weinstock-Guttman, B., Silva, D., 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. Am. J. Neuroradiol. 39 (2), 289–295. https://doi.org/10.3174/ajnr.A5442.

Zivadinov, R., Havrdová, E., Bergsland, N., Tyblova, M., Hagemeier, J., Seidl, Z., Dwyer, M.G., Vaneckova, M., Krasensky, J., Carl, E., Kalincik, T., Horáková, D., 2013. Thalamic atrophy is associated with development of clinically definite multiple sclerosis. Radiology 268 (3), 831–841. https://doi.org/10.1148/radiol.13122424.

Zivadinov, Robert, Heininen-Brown, M., Schirda, C. V., Poloni, G. U., Bergsland, N., Magnano, C. R., … Dwyer, M. G. (2012). Abnormal subcortical deep-gray matter susceptibility-weighted imaging filtered phase measurements in patients with multiple sclerosis. A case-control study. NeuroImage. https://doi.org/10.1016/j.neuroimage.2011.07.045.

Zivadinov, R., Jakimovski, D., Gandhi, S., Ahmed, R., Dwyer, M.G., Horakova, D., Weinstock-Guttman, B., Benedict, R.R.H., Vaneckova, M., Barnett, M., Bergsland, N., 2016. Clinical relevance of brain atrophy assessment in multiple sclerosis. Implications for its use in a clinical routine. Expert Rev. Neurother. 16 (7), 777–793. https://doi.org/10.1080/14737175.2016.1181543.