

# Portable, Low-Field Magnetic Resonance Imaging Sensitive Detects and Accurately Quantifies Multiple Sclerosis Lesions

T. Campbell Arnold<sup>1,2</sup>, Danni Tu<sup>3</sup>, Serhat V. Okar<sup>4</sup>, Govind Nair<sup>4</sup>, Samantha By<sup>5</sup>, Karan Kawatra<sup>4</sup>, Timothy E. Robert-Fitzgerald<sup>3</sup>, Lisa M. Desiderio<sup>6</sup>, Matthew K. Schindler<sup>7</sup>, Russell T. Shinohara<sup>2,3</sup>, Daniel S. Reich<sup>4</sup>, Joel M. Stein<sup>2,6</sup>

1. Department of Bioengineering, School of Engineering & Applied Science, University of Pennsylvania, Philadelphia, PA, 19104, USA
2. Center for Neuroengineering and Therapeutics, University of Pennsylvania, Philadelphia, PA, 19104, USA
3. Penn Statistics in Imaging and Visualization Center and Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania, Philadelphia, PA, 19104, USA
4. National Institute of Neurological Disorders and Stroke, National Institutes of Health (NIH), Bethesda, MD, 20892, USA
5. Hyperfine, Guilford, CT, 06437, USA
6. Department of Radiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, 19104, USA
7. Department of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, 19104, USA

## Corresponding authors:

Joel M. Stein

Email address: [joel.stein@penmedicine.upenn.edu](mailto:joel.stein@penmedicine.upenn.edu)

Mailing address: Joel. M Stein

Hospital of the University of Pennsylvania

1 Silverstein Pavilion

3400 Spruce St.

Philadelphia, PA 19104-4283

Daniel S. Reich

Email address: [daniel.reich@nih.gov](mailto:daniel.reich@nih.gov)

Mailing address: Daniel S. Reich

Translational Neuroradiology Section

Division of Neuroimmunology & Neurovirology

Building 10, Room 5C103

10 Center Dr.

Bethesda, MD 20892-4128

## Word Counts:

Abstract	Article	Intro	Methods	Results	Discussion	Conclusion	Tables	Figures	References
315	3849	373	1133	1091	1170	82	2	7	44

## 1 **Abstract**

2  
3 Magnetic resonance imaging is a fundamental tool in the diagnosis and management of  
4 neurological diseases such as multiple sclerosis (MS). New portable, low-field MRI scanners  
5 could potentially lower financial and technical barriers to neuroimaging and reach underserved  
6 or disabled populations. However, the sensitivity of low-field MRI for MS lesions is unknown.  
7 We sought to determine if white matter lesions can be detected on a 64mT low-field MRI,  
8 compare automated lesion segmentations and total lesion burden between paired 3T and 64mT  
9 scans, and identify features that contribute to lesion detection accuracy. In this prospective,  
10 cross-sectional study, same-day brain MRI (FLAIR, T1, and T2) scans were collected from 36  
11 adults (32 women; mean age,  $50 \pm 14$  years) with known or suspected MS using 3T (Siemens)  
12 and 64mT (Hyperfine) scanners at two centers. Images were reviewed by neuroradiologists. MS  
13 lesions were measured manually and segmented using an automated algorithm. Statistical  
14 analyses assessed accuracy and variability of segmentations across scanners and systematic  
15 scanner biases in automated volumetric measurements. Lesions were identified on 64mT scans in  
16 94% (31/33) of patients with confirmed MS. The smallest lesions manually detected were  $5.7 \pm$   
17  $1.3$  mm in maximum diameter at 64mT vs  $2.1 \pm 0.6$  mm at 3T. Automated lesion burden  
18 estimates were highly correlated between 3T and 64mT scans ( $r = 0.89$ ,  $p < 0.001$ ). Bland-  
19 Altman analysis identified bias in 64mT segmentations (mean = 1.6 ml, standard error = 5.2 ml,  
20 limits of agreement = -19.0–15.9 ml), which over-estimated low lesion burden and under-  
21 estimated high burden ( $r = 0.74$ ,  $p < 0.001$ ). Visual inspection revealed over-segmentation was  
22 driven by flow-related hyperintensities in veins on 64mT FLAIR. Lesion size drove  
23 segmentation accuracy, with 93% of lesions  $>1.0$  ml and all lesions  $>1.5$  ml being detected.  
24 These results demonstrate that in established MS, a portable 64mT MRI scanner can identify  
25 white matter lesions, and disease burden estimates are consistent with 3T scans.

26

## 27 **Keywords**

28

29 Low-field MRI, portable MRI, point-of-care MRI, Hyperfine, multiple sclerosis, white matter  
30 lesion

## 1 **Highlights**

2

- 3 ● Paired, same-day 3T and 64mT MRI studies were collected in 36 patients
- 4 ● 64mT MRI showed 94% sensitivity for detecting any lesions in established MS cases
- 5 ● The diameter of the smallest detected lesion was larger at 64mT compared to 3T
- 6 ● Disease burden estimates were strongly correlated between 3T and 64mT scans
- 7 ● Low-field MRI can detect white matter lesions, though smaller lesions may be missed

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

# 1 **1. Introduction**

2  
3 Multiple sclerosis (MS) is a complex inflammatory and degenerative disease of the central  
4 nervous system (1). MS causes demyelinating lesions, typically assessed using magnetic  
5 resonance imaging (MRI). Imaging features related to white matter lesions (WML), such as  
6 number, volume, and dissemination in space and time, are key diagnostic criteria of MS (2) and  
7 determine treatment courses and clinical trial eligibility (3). Early diagnosis leads to better  
8 clinical outcomes, including delayed disease progression and reduced severity (4).

9  
10 Although MS affects ~800,000 people in the United States (5) and likely >2.5 million people  
11 globally (6), the significant cost, infrastructure, and technical requirements associated with  
12 traditional high-field-strength MRI limit access to imaging worldwide (7). The scarcity is  
13 particularly felt in low-resource, sparsely populated, and rural areas (8). As the lack of diagnostic  
14 imaging can lead to delayed diagnosis and treatment, which result in worsening health disparities  
15 (9), there is renewed interest in low-field MRI, which employs magnets with a field strength of 1  
16 tesla (T) or lower, as a lower-cost and potentially portable alternative to high-field MRI (10).

17  
18 Recent improvements in hardware as well as image reconstruction and processing algorithms  
19 (11) have made low-field MRI promising in contexts where modest resolution is sufficient for  
20 diagnostic purposes (12). The clinical utility of portable low-field MRI has already been  
21 demonstrated for bedside monitoring in intensive care settings, where patients may not be stable  
22 enough to transport for traditional imaging (13–15). In the outpatient treatment of diseases such  
23 as MS, portable low-field MRI has the potential to lower barriers to accessing MRI technology  
24 and allow more frequent monitoring of disease activity (16), however its sensitivity and accuracy  
25 have not been explored. In particular, WML to background signal intensities and size thresholds  
26 for detection are unknown.

27  
28 In this study, we assessed the feasibility of portable low-field MRI for MS lesion identification  
29 and lesion burden estimation. We collected paired same-day brain MRI from adults with MS at  
30 3T and 64mT at two different institutions and compared lesion detection using both manual and  
31 automated measurements. We anticipated that MS lesions would be detectable at 64mT, though

1 sensitivity to small lesions would likely be reduced. Finally, we explored a simple approach for  
2 super-resolution imaging of small lesions based on multi-acquisition image averaging.

## 3 4 **2. Materials and Methods**

### 5 6 *2.1 Participants & Imaging*

7  
8 Among adult outpatients undergoing clinical 3T brain MRI for known or suspected MS between  
9 October 2020 and April 2021, 36 patients (Fig. 1) were recruited at site A (N=21) and site B  
10 (N=15). All patients received same-day 3T and 64mT MRI. Demographic information was  
11 collected from clinical notes and included age, sex, race, clinical phenotype, disease duration,  
12 Expanded Disability Status Scale (EDSS), and current disease modifying therapy (Table 1). This  
13 study was approved by each site's institutional review board, and patients provided written,  
14 informed consent.

15

	Total (N = 36)	Site A (N = 21)	Site B (N = 15)
Age (years)	49.6 ± 14.2	45.3 ± 13.6 *	55.7 ± 12.7 *
Sex (women/men)	32/4	19/2	13/2
Race/ethnicity (White/Black/Hispanic)	26/9/1	14/7/0	12/2/1
Disease duration (years)	13.7 ± 11.2	10.2 ± 9.6 *	18.5 ± 11.5 *
EDSS (0–10)	1.5 (IQR = 2.0)	1.5 (IQR = 2.0)	2.0 (IQR = 1.25)
Phenotype		RRMS (18), CIS (1), NMO (1), RIS (1)	RRMS (10), SPMS (2), CIS (1), ITM (1), PPMS (1)
Current disease modifying therapy		ocrelizumab (9), natalizumab (2), other (6), none (4)	dimethyl fumarate (6), ocrelizumab (3), other (3), none (6)

16

17 **Table 1. Patient Demographics.** Demographic information and clinical history for 36  
18 consecutive MS patients included in the study. An asterisk indicates a significant difference

1 between sites. Abbreviations: Expanded disability status scale (EDSS), relapsing-remitting  
 2 multiple sclerosis (RRMS), primary progressive multiple sclerosis (PPMS), and secondary  
 3 progressive multiple sclerosis (SPMS), clinically isolated syndrome (CIS), radiologically  
 4 isolated syndrome (RIS), neuromyelitis optica (NMO), idiopathic transverse myelitis (ITM).

5  
 6 High-field MRI was performed on 3T scanners (Siemens, Erlangen, Germany). Each site used a  
 7 standardized, whole-brain imaging protocol, which included 3D T1-weighted (T1w), T2-  
 8 weighted (T2w), and 3D T2-FLAIR sequences (Fig. 2A). Sequence parameters are listed in  
 9 Table 2. At site B, patients received gadolinium (gadobutrol, 0.1 mmol/L) prior to 3T scanning;  
 10 64mT scans were obtained after contrast-enhanced 3T scans with mean post-gadolinium duration  
 11 of  $58 \pm 21$  minutes.

12 Same-day, low-field MRI was performed on portable 64mT Swoop MRI systems (Hyperfine,  
 13 Guilford, CT). Whole-brain, 3D T1w, T2w, and T2-FLAIR scans, analogous to those acquired at  
 14 3T, were collected (Fig. 2B). Including localizer and pre-scan calibration, total scan time was  
 15 26:06 minutes.

16

Sequence	Site	Field Strength (T)	TE (ms)	TR (s)	TI (s)	Resolution (mm)	Scan-time (min:sec)	Averages
T1w-MPRAGE	A	3	2.48	1.9	0.9	1.0x1.0x1.0	4:18	1
T1w-MP2RAGE	B	3	2.92	5	0.7, 2.5	1.0x1.0x1.0	8:30	1
T1w	Both	0.064	6.26	1.5	0.3	1.5x1.5x5	4:52	1
T2-FLAIR 3D	A	3	398	5	1.6	1.0x1.0x1.0	5:02	1
T2-FLAIR 3D	B	3	352	4.8	1.8	1.0x1.0x1.0	7:15	1
T2-FLAIR	Both	0.064	200	4	1.4	1.6x1.6x5	9:29	1
T2w	A	3	103	5.5	N/A	0.5x0.5x5.2	1:46	2
T2w	B	3	82	5	N/A	0.34x0.34x3.0	4:30	1
T2w	Both	0.064	209	2	N/A	1.5x1.5x5	7:01	1

17

1 **Table 2. Sequence parameters for study scans.** Abbreviations: Tesla (T), T1-weighted (T1w),  
2 T2-weighted (T2w), Fluid-attenuated inversion recovery (FLAIR), echo time (TE), repetition  
3 time (TR), inversion time (TI).

#### 4 5 *2.2 Manual Review and Lesion Measurement*

6  
7 MRI scans were reviewed for WML by two neuroradiologists (DSR and JMS, 19 and 8 years  
8 experience). Maximum diameters (Dmax) of the smallest and largest WML visually detectable at  
9 each field strength were manually measured by a neuroradiologist (JMS) and a neurologist  
10 (SVO) with MS MRI expertise (3 years experience) using ITK-SNAP (17). All measurements  
11 were made on T2-FLAIR scans. Lesions were assessed in axial planes as well as sagittal and  
12 coronal reformations, and Dmax measurements were made on the plane with the largest lesion  
13 diameter. In confluent periventricular lesions, Dmax was measured perpendicular to the  
14 ventricle. Low-field imaging was evaluated prior to 3T scans to avoid interpretation bias, and  
15 image sets were reviewed separately. Inter-rater reliability was assessed using two-way random,  
16 single-measure intraclass correlation coefficients (ICC) with 95% confidence intervals (CI)  
17 reported. Patients that did not meet the McDonald criteria for dissemination of lesions in space  
18 were excluded from subsequent analyses (2).

#### 19 20 *2.3 Automated Lesion Segmentation*

21  
22 The same segmentation pipeline was applied to 3T and 64mT images. Images were preprocessed  
23 using N4 bias correction (18), and each T2-FLAIR volume was rigidly registered to the  
24 corresponding T1w volume using Advanced Normalization Tools (ANTs) (19,20). A brain mask  
25 was obtained using Multi-Atlas Skull-Stripping (MASS) (21). Finally, to enable comparisons  
26 across patients, image intensities were normalized using White Stripe (22) within each sequence.  
27 Lesion segmentation was performed using the Method for Inter-Modal Segmentation Analysis  
28 (MIMoSA) (23,24), an automated pipeline that leverages shared information (coupling) between  
29 modalities to produce probability maps of WML (Fig. S1). All probability maps were  
30 thresholded at 0.2 to generate binary lesion masks, manually selected based on prior empirical  
31 evidence.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29

#### 2.4 Automated Segmentation Evaluation

Estimation of total lesion burden was the primary performance measure compared between 3T and 64mT. Two lesion burden estimates were obtained for each patient by calculating lesion segmentation volumes for the respective scanners. The relationship between volume estimates was assessed using Pearson’s correlation. Bland-Altman plots were used to determine agreement and assess for systematic scanner biases.

Similarity between segmentation masks was assessed using the Dice-Sørensen coefficient (Dice), which measures the overlap between two images ( $X$  and  $Y$ ):

$$Dice = \frac{2|X \cap Y|}{|X| + |Y|} \tag{1}$$

Dice scores range from 0–1, with 1 indicating perfect segmentation overlap. Prior to Dice calculation, 3T and 64mT images were coregistered using ANTs (20). While Dice score may not reflect segmentation quality when the number of target objects is not known *a priori*, this measure has been included as it is widely used and allows for comparisons across studies (25). All 3T and 64mT segmentations were manually reviewed to verify overlapping regions were WMLs rather than false positive detections.

#### 2.5 Size and Intensity Analysis

Connected-components analysis was used to identify individual lesions in automated 3T and 64mT segmentations (26). Sensitivity to individual lesions at low-field MRI was assessed using the true-positive rate (TPR), or the proportion of lesions correctly identified:

$$TPR = \frac{TP}{TP + FN} \tag{2}$$



1 where true positives ( $TP$ ) are defined as lesions where 64mT and 3T segmentations overlap and  
2 false negatives ( $FN$ ) are defined as lesions with 3T segmentation but no 64mT segmentation  
3 overlap. The false-discovery rate (FDR) was assessed as:

$$4 \quad FDR = \frac{FP}{FP + TP}, \quad (3)$$

6 where false positives ( $FP$ ) are defined as lesions with 64mT segmentation but no 3T  
7 segmentation overlap. Lesion overlap was defined as at least one shared voxel between the 3T  
8 and 64mT lesion segmentations. To understand the impact of lesion features on detection rates,  
9 TPR and FDR were plotted as a function of lesion size and normalized lesion intensity (22).

## 11 12 *2.6 Super Resolution Imaging*

13  
14 Low-field MRI necessitates a trade-off between signal-to-noise ratio and image resolution, which  
15 limits the minimum detectable lesion size. However, image quality can be increased by taking  
16 advantage of partial volume effects in multiple scans (27). In one patient, multi-acquisition  
17 volume averaging was explored. Eight 64mT T2-FLAIR acquisitions (TE = 1.8 s, TR = 4 s, TI =  
18 1.4 s, averages = 80, scan time = 6:03 min, resolution = 1.8x1.8x5 mm) were collected with head  
19 repositioning between scans (total scan time: 48:24 minutes). Images were resliced to 1 mm  
20 isotropic resolution with linear interpolation, underwent affine registration to the initial  
21 acquisition, and were averaged to create a single high-resolution image. Super resolution images  
22 were iteratively generated for each additional acquisition, for a total of eight images.

23  
24 To quantify lesion conspicuity, we manually segmented the white matter lesion and a similarly  
25 sized region in normal appearing ipsilateral white matter on 3T imaging using ITK-SNAP (17).  
26 The 3T image and lesion segmentations were registered to the initial 64mT acquisition. We  
27 calculated lesion conspicuity as the ratio of the difference and the sum of mean intensity in the  
28 two segmentations:

$$29 \quad Conspicuity = \frac{X-Y}{X+Y}, \quad (4)$$

31

1 where  $X$  is the mean intensity in the lesion segmentation and  $Y$  is the mean intensity of the  
2 normal appearing white matter segmentation. The segmentations were applied to each super  
3 resolution iteration to calculate conspicuity for that iteration.

#### 4 5 *2.7 Statistics and Data/Code Availability*

6  
7 All code related to this study is publicly available. The MIMoSA algorithm is available in R as a  
8 Neuroconductor package (<https://neuroconductor.org/package/details/mimosa>) and on GitHub  
9 (<https://github.com/avalcarcel9/mimosa/>). T-tests, Pearson's correlation, and summary statistics  
10 were calculated using scipy (v1.5.2) and numpy (v1.19.2) in Python (v3.8.5). Bland-Altman  
11 plots were visualized using pyCompare (v1.5.1) while boxplots and correlations utilized seaborn  
12 (v0.11.0). Inter-rater reliability was calculated using irr (v0.84.1) in R (v4.0.3). A manuscript  
13 companion containing all analyses is available on GitHub ([https://github.com/penn-](https://github.com/penn-cnt/Arnold_LF-MRI_MS)  
14 [cnt/Arnold\\_LF-MRI\\_MS](https://github.com/penn-cnt/Arnold_LF-MRI_MS)). The data generated in this study can be made available, with  
15 protected health information removed, upon reasonable request to the corresponding author and  
16 with a data sharing agreement between institutions in place.

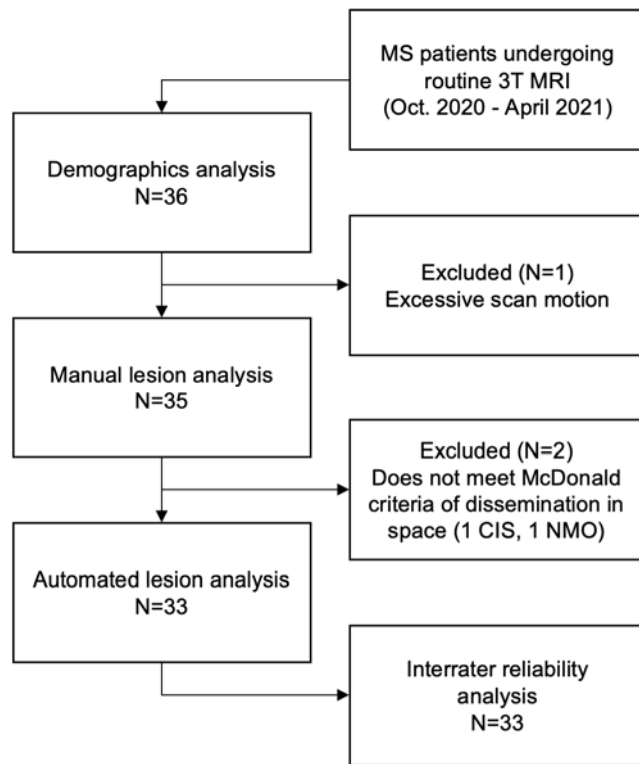
### 17 18 **3. Results**

#### 19 20 *3.1 Patient Demographics*

21  
22 We collected data from 36 adults with known or suspected MS. The patient population had a  
23 mean age of 49.6 (SD: 14.2) years and was composed of 32 women and 4 men (Table 1). The  
24 mean duration of disease was 13.7 years (SD: 11.2), and patients had a median EDSS of 1.5  
25 (interquartile range = 2). Patients from site B were significantly older than those from site A  
26 (two-sample t-test,  $t=2.2$ ,  $p = 0.03$ , site A: 45.3 years old, site B: 55.7 years old) and had a  
27 correspondingly longer duration of disease (two-sample t-test,  $t=2.3$ ,  $p = 0.03$ , site A: 10.2 years,  
28 site B: 18.5 years). Additional demographic information is provided in Table 1. After initial scan  
29 review, three patients were excluded from further analysis: One patient had excessive motion in  
30 the scanner and two patients did not meet the MS diagnostic criteria of having lesion

1 dissemination in space (DIS) (2). All subsequent analyses are based on the remaining 33 patients  
2 (Fig. 1).

3



4

5 **Figure 1. Flow chart of study participants.** Abbreviations: multiple sclerosis (MS), clinically  
6 isolated syndrome (CIS), neuromyelitis optica (NMO).

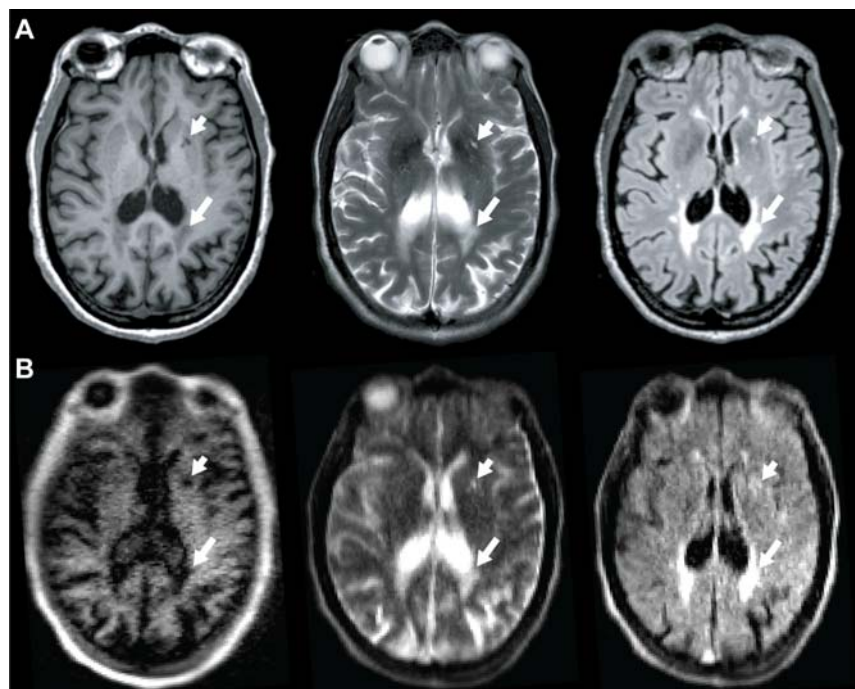
7

### 8 *3.2 Manual Measurements*

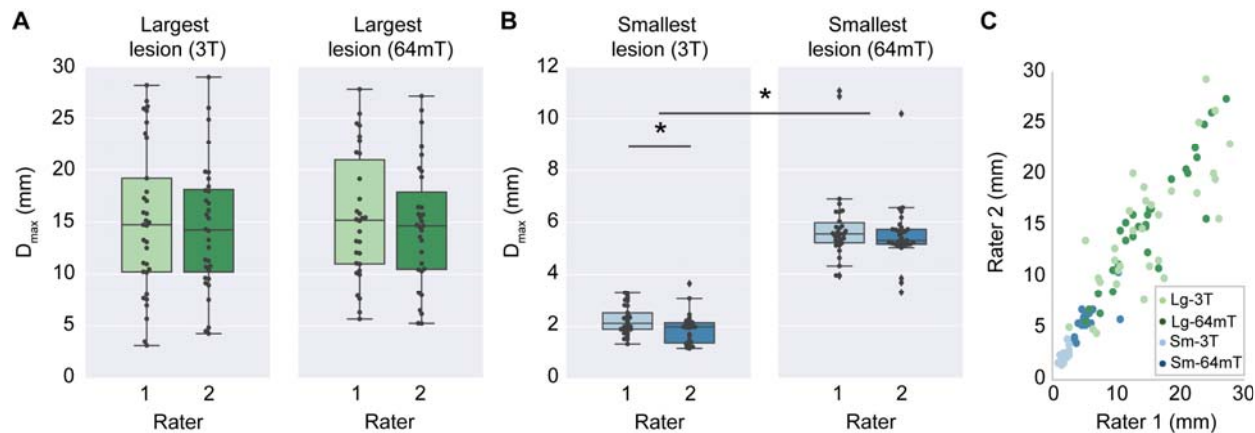
9

10 MS lesions on 64mT are characterized by T1w hypointensity and T2w/T2-FLAIR  
11 hyperintensity, similar to 3T imaging (Fig. 2). At 64mT, lesions were identified by at least one  
12 rater in 94% (31/33) of patients with confirmed lesions on 3T imaging. In one patient, only one  
13 rater identified lesions at 64mT; all other low-field ratings were concordant. The largest and  
14 smallest lesions in each image were identified, and the Dmax was recorded. The 64mT scanner  
15 showed 100% sensitivity for detecting WML when there was at least one lesion with Dmax >5  
16 mm (31/33 patients, 94%). Across patients, there was no significant difference in Dmax for the  
17 largest lesions measured at 64mT ( $15.1 \pm 5.9$  mm) and 3T ( $14.8 \pm 6.6$  mm) (Fig. 3A). However,  
18 the mean Dmax for the smallest WML detectable was significantly larger (paired t-test,  $t=19.6$ , p

1 < 0.001) on 64mT ( $5.7 \pm 1.3$  mm) compared to 3T ( $2.1 \pm 0.6$  mm) (Fig. 3B). There was no effect  
2 of scan site on Dmax measurements, however there was a significant difference between rater 1  
3 ( $2.3 \pm 0.5$  mm) and rater 2 ( $1.9 \pm 0.6$  mm) for the smallest lesion detected at 3T (paired t-test,  
4  $t=4.8$ ,  $p < 0.001$ ). No gadolinium enhancing lesions were seen on 3T or 64mT imaging.  
5



6  
7 **Figure 2. MS lesions appear similar on 3T and 64mT pulse sequences.** Paired 3T (A) and  
8 64mT (B) images from a 66-70 year-old female with stable RRMS showing deep gray matter  
9 lesions (short arrow) and periventricular white matter lesions (long arrow) on both scanners.  
10 Sequences include T1-weighted (left), T2-weighted (center), and T2-FLAIR (right).  
11  
12  
13



1  
2 **Figure 3. Manual lesion size measurements and interrater reliability.** Raters from each site  
3 independently measured the maximum diameter (D<sub>max</sub>) of the smallest lesion (Sm) and largest  
4 lesion (Lg) in 3T and 64mT imaging for all patients. (A) For the largest lesion measurements,  
5 there were no significant differences between raters at 3T ( $t = 1.3$ ,  $p = 0.19$ ) or 64mT ( $t = 1.2$ ,  $p$   
6  $= 0.23$ ); additionally, there was no difference between 3T and 64mT measurements ( $t = 0.04$ ,  $p =$   
7  $0.97$ ). (B) For the smallest lesion measurements, there was a significant difference between raters  
8 for 3T measurements ( $t = 4.83$ ,  $p < 0.001$ ) although 64mT measurements were not significantly  
9 different ( $t = 1.67$ ,  $p = 0.11$ ); additionally, the diameter of the smallest lesion was significantly  
10 lower ( $t = 19.6$ ,  $p < 0.001$ ) when measured on 3T (mean 2.1 mm) compared to 64mT (mean 5.7  
11 mm). (C) Across all lesions there was a strong correlation ( $r = 0.90$ ,  $p < 0.001$ ) between raters.  
12 There was significant intraclass correlation for the largest lesion at 3T (ICC = 0.77, CI = [0.58-  
13 0.88]), largest lesion at 64mT (ICC = 0.91, CI = [0.83-0.96]), smallest lesion at 3T (ICC = 0.62,  
14 CI = [0.12-0.83]), and smallest lesion at 64mT (ICC = 0.66, CI = [0.4-0.82]), indicating a high  
15 degree of agreement between rater measurements for both 3T and 64mT.

### 16 17 3.3 Interrater Reliability

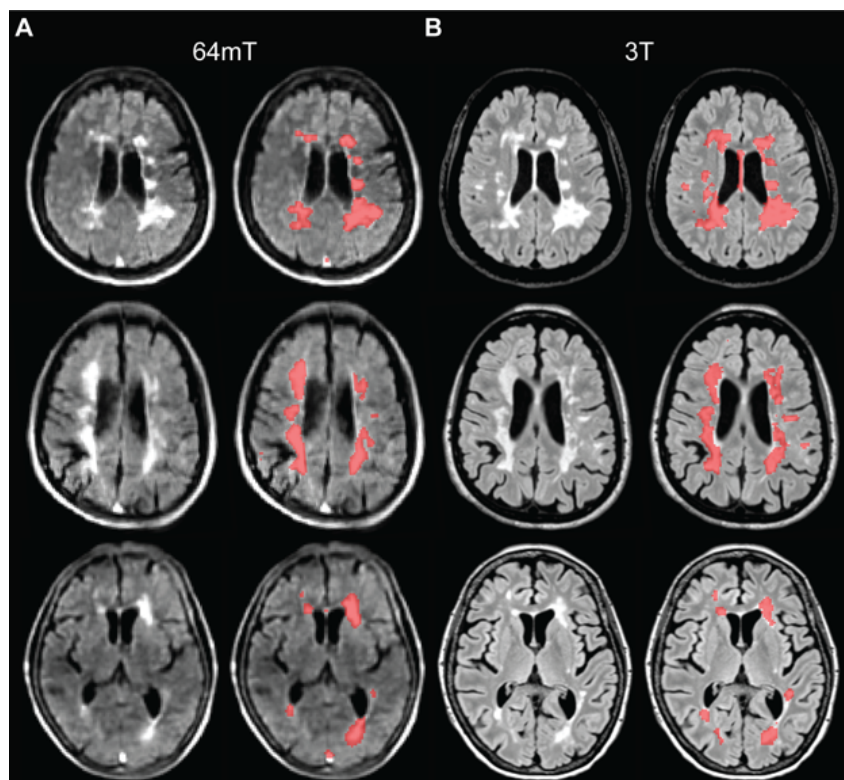
18  
19 The smallest and largest lesion in each image were independently measured by two raters to  
20 assess interrater reliability (Fig. 3). The ICC for each patient's largest lesion measured at 3T and  
21 64mT was 0.77 (CI = [0.58-0.88]) and 0.91 (CI = [0.83-0.96]) respectively, indicating high  
22 interrater reliability for large lesions on both scanners (Fig. 3A). Similarly, when measuring each  
23 patient's smallest lesion there was a significant relationship between raters at both 3T (ICC =  
24 0.62, CI = [0.12-0.83]) and 64mT (ICC = 0.66, CI = [0.4-0.82]) (Fig. 3B). This indicates that

1 measurements made by raters had a similar degree of reliability at 3T and 64mT. Of note, the  
2 average smallest lesions detected (3T:  $2.1 \pm 0.6$  mm, 64mT:  $5.7 \pm 1.3$  mm) approached the  
3 respective resolution limits for each system (3T: 1 mm, 64mT: 5 mm).

#### 4 5 *3.4 Total Lesion Burden Estimates*

6  
7 To obtain more objective measures, 3T and 64mT image sets were processed with the same  
8 automated lesion segmentation algorithm. Initial qualitative review of segmentation overlays  
9 revealed similar patterns of lesion segmentation, particularly with respect to large periventricular  
10 lesions (Fig. 4). Quantitative comparisons indicated that estimates of total lesion burden were  
11 highly correlated ( $r = 0.89$ ,  $p < 0.001$ ) (Fig. 5A). Mean lesion burden estimates were not  
12 significantly different (paired-t-test,  $t = 1.0$ ,  $p = 0.32$ ) between 3T ( $11.9 \pm 16.5$  ml) and 64mT  
13 ( $13.5 \pm 10.2$  ml) images.

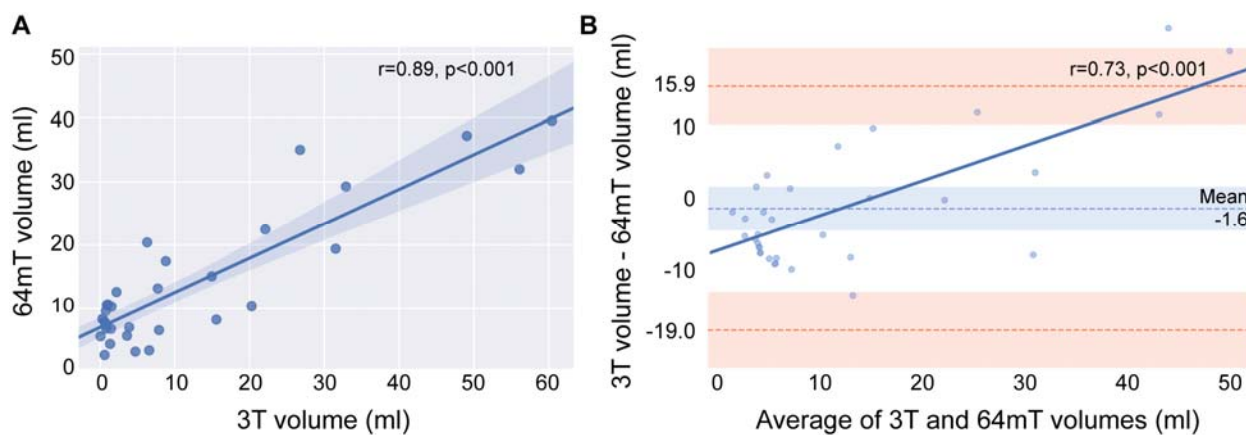
14



15  
16 **Figure 4. Automated lesion segmentations at 3T and 64mT overlap.** (A) 64mT FLAIR  
17 images for three cases (left) with automated lesion segmentations generated from the 64mT  
18 images using MIMoSA overlaid (right). (B) Corresponding 3T FLAIR images for the same three

1 cases (left) with 3T based segmentations (right). Patients from top to bottom are a 51-55 year-old  
2 female with RRMS, 41-45 year-old female with RRMS, and 71-75 year-old female with RRMS.  
3 All images were coregistered to 64mT T1-weighted images for comparison. Segmentations  
4 generated from 64mT and 3T scanners show similar patterns.  
5

6 A Bland-Altman plot for agreement between 3T and 64mT lesion burden estimates is presented  
7 in Figure 5B. The mean difference was 1.6 ml with a 5.2 ml standard error of measurement, and  
8 the 95% limits of agreement were -19.0 to 15.9 ml. There was a significant correlation ( $r = 0.74$ ,  
9  $p < 0.001$ ) between pairwise differences and averages, indicating that compared to 3T, the 64mT  
10 segmentations overestimate low lesion burdens and underestimate high lesion burdens. Visual  
11 inspection revealed that false positives contributing to over-segmentation were predominantly  
12 due to flow-related high signal intensity in veins, hyperintensity in non-lesion structures (such as  
13 the pineal gland), and areas of artifactual peripheral high signal in cortical/subcortical tissue on  
14 64mT FLAIR sequences (Fig. S2).  
15



17 **Figure 5. Total lesion volume measured at 3T and 64mT shows agreement.** (A) Pearson's  
18 correlation ( $r = 0.89$ ,  $p < 0.001$ ) and (B) Bland-Altman plot showing agreement between 3T and  
19 64mT segmentation volumes (bias -1.6 ml, standard error of measurement = 5.2 ml, 95% limit of  
20 agreement -19.0 to 15.9 ml). Pearson's correlation ( $r = 0.74$ ,  $p < 0.001$ ) in dark blue indicates  
21 over-segmentation at 64mT when lesion burden is low and under-segmentation when burden is  
22 high.  
23

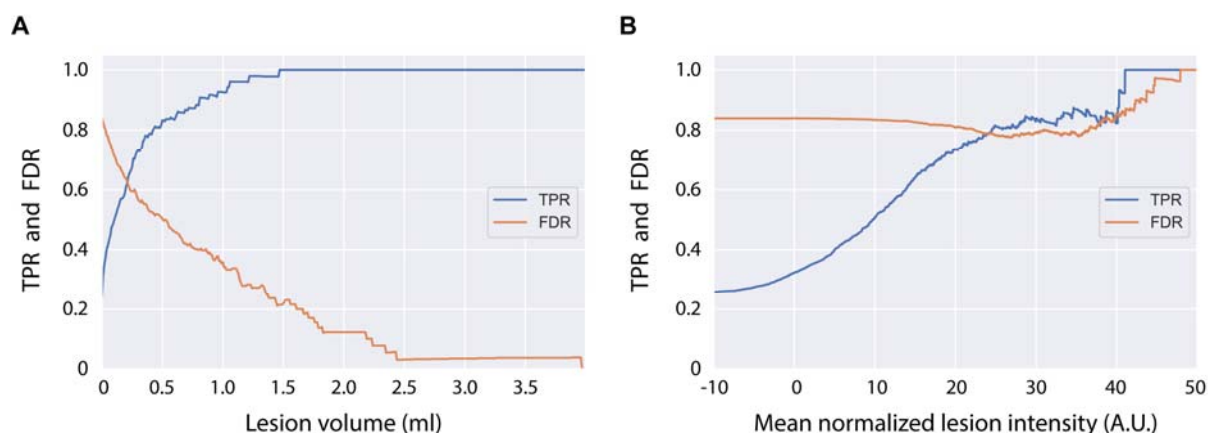
### 24 3.5 Automated Segmentation Overlap

1  
2 Across patients, there was a wide range in overlap between 3T and 64mT segmentation pairs  
3 (Dice: mean = 0.23, standard deviation = 0.21, max = 0.65, min = 0), with automated  
4 segmentations overlapping in 91% (30/33) of patients. Two patients had no segmentation overlap  
5 and one patient was excluded because the overlapping region was a hyperintense pineal gland,  
6 not a WML. All three patients without lesion overlap were in the bottom 12% of total lesion  
7 burden. To characterize the full range of segmentation quality across the dataset, Figure S3  
8 illustrates segmentations from each quartile of the Dice distribution. Larger lesion size is  
9 frequently associated with higher Dice (25). We found in our dataset that total lesion burden at  
10 3T was highly correlated with Dice ( $r = 0.81$ ,  $p < 0.001$ ) (Fig. S4).

### 11 12 *3.6 Lesion Sensitivity and False Discovery*

13  
14 In each segmentation, individual lesions were identified using connected-components analysis  
15 (26). For each lesion, volume and mean intensity were quantified. The true-positive rate (TPR)  
16 and false-discovery rate (FDR) were calculated across a range of lesion size and intensity  
17 thresholds (Fig. 6). The TPR increases dramatically with lesion size, reaching 93% for lesions >1  
18 ml and 100% for lesions >1.5 ml. The FDR decreases with lesion size, reaching 36% for lesions  
19 >1 ml, 22% for lesions >1.5ml, and 3% for lesions >2.5 ml. TPR also increases with mean lesion  
20 intensity, indicating that lesion intensity influences sensitivity; however, FDR remains high  
21 (>75%) regardless of the sensitivity, indicating a large number of false positive detections across  
22 intensity thresholds. Examples of false positive detections can be seen in Figure S2.

23



24

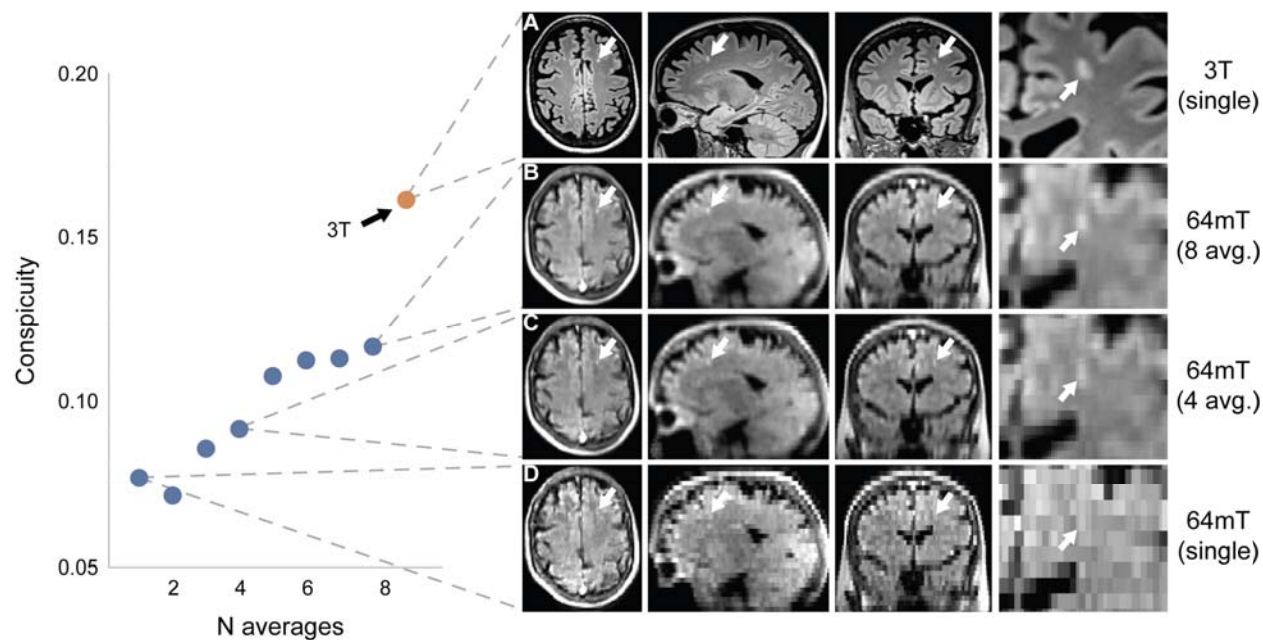


1 **Figure 6. Lesion size and intensity influence detection rate.** (A) The detection rate, or true  
2 positive rate (TPR), steadily increases with lesion size, with 93% detected at >1 ml, and all  
3 lesions greater than 1.5 ml being detected. The false discovery rate (FDR) decreases with lesion  
4 size, with 36% false discovery rate at >1 ml, 22% at >1.5 ml, and 3% at >2.5 ml. Though the x-  
5 axis was limited to 4 ml for illustrative purposes, lesions >20 ml were found in the dataset. (B)  
6 To analyze the relationship between lesion intensity and detection rate, image intensity values  
7 were first normalized using White Stripe (Shinohara et al. 2014). While detection rate increases  
8 as mean lesion intensity increases, the FDR remains high (>75%) across lesion intensities. The  
9 high number of false positive detections was driven by hyperintense veins and peripheral signal  
10 artifacts, as seen in figure E3.

### 11 12 *3.7 Super Resolution Imaging*

13  
14 In one patient, a 3x4x5 mm (0.06 ml) subcortical lesion was evident near the left middle frontal  
15 gyrus on 3T (Fig. 7A) but not in any single low-field acquisition (Fig. 7D). After multi-  
16 acquisition volume averaging of 3 to 8 acquisitions, the lesion became detectable on the low-  
17 field system, and lesion intensity relative to ipsilateral white matter steadily increased with  
18 additional acquisitions. With 8 volume averages, there was a 53% increase in lesion conspicuity,  
19 which was equivalent to 72% of conspicuity at 3T (Fig. 7B). With multi-acquisition image  
20 averaging, lesions as small as 0.06 ml were discernible on 64mT scans.

21



1  
2 **Figure 7. Multi acquisition image averaging increases resolution.** This figure depicts a  
3 3x4x5mm (0.06 ml) subcortical left frontal white matter lesion in a 51-55 year-old female with  
4 stable RRMS and compares 64mT FLAIR images generated from multi-acquisition image  
5 averaging to 3T imaging. The lesion is readily apparent on 3T imaging (A); however, it could not  
6 be discerned in a single 64mT acquisition (D). Volume averaging of multiple acquisitions with  
7 repositioning between scans did reveal the lesion on the low-field system (B & C). The lesion  
8 was discernible for  $N \geq 3$  multi acquisition averages. The lesion was manually segmented on 3T,  
9 and the ratio of mean lesion intensity to ipsilateral adjacent white matter (WM) is given as an  
10 estimate of lesion conspicuity (red dot). In 64mT imaging, the ratio steadily increases with  
11 additional acquisition averages (blue dots). With 8 volume averages, there was a 53% increase in  
12 lesion conspicuity.

## 13 14 **4. Discussion**

15  
16 In this study, we compared manual and automated lesion detection in paired 3T and portable  
17 64mT brain MRI from patients with known or suspected MS at two sites. On visual inspection of  
18 64mT images, neuroradiologists were able to detect white matter lesions in 94% (31/33) of  
19 patients with discernable 3T lesions. An automated lesion segmentation algorithm detected  
20 overlapping lesions in 91% (30/33) of patients, and estimates of total lesion burden were highly

1 correlated between 3T and 64mT scans ( $r = 0.89$ ,  $p < 0.001$ ). Together, these results suggest that  
2 portable 64mT imaging could have diagnostic utility in the context of MS.

3  
4 Our investigation is motivated by recent advances in hardware development and reconstruction  
5 software that address the reduced signal-to-noise and resolution associated with low-field MRI.  
6 Commercial MRI systems have predominantly trended towards higher field strengths, with low-  
7 field systems relegated to niche applications (7,11,28). Currently, there is renewed commercial  
8 interest in developing low-field MRI systems (29), including Hyperfine's portable 64mT Swoop  
9 system, Synaptive Medical's 0.5T Evry system, and Siemens's 0.55T Magnetom Free.Max  
10 system, all of which have received FDA clearance since 2020. As evidence indicates that low-  
11 field systems can detect relatively subtle pathologies, such as CNS demyelination (30–32), we  
12 were motivated to investigate the sensitivity of the newly available portable 64mT MRI for MS  
13 lesions.

14  
15 Our study found that lesions could be identified on 64mT scans in 94% of patients with  
16 discernible lesions at 3T. Additionally, we found a strong correlation between total lesion burden  
17 estimates on 3T and 64mT scanners. However, the smallest detected lesion size was significantly  
18 larger at 64mT ( $5.7 \pm 1.3$  mm) compared to 3T ( $2.1 \pm 0.6$  mm), indicating that smaller lesions  
19 are missed at low field. Taken together, these findings suggest that the 64mT device may be  
20 useful for tracking disease severity over time, although the device may be less suitable for  
21 making an initial diagnosis when high-field scanners are available. However, while high-field  
22 MRI will likely remain the diagnostic tool of choice in population centers of developed  
23 countries, the lower costs and infrastructure requirements of portable low-field MRI could  
24 expand clinical options for patients in low and middle income countries and rural areas (9).  
25 Within the United States, 60% of rural hospitals lack an on-site MRI (33). Additionally, most  
26 MS patients will experience mobility impairment, which can impact quality of care (34). Mobile  
27 MRI units could bring imaging to patients, providing otherwise unavailable service to sparsely  
28 populated areas and individuals who cannot travel (35,36).

29  
30 The lower cost and ease of use associated with low-field MRI could also impact how early and  
31 frequently patients are imaged (37). Rovira et al. were able to predict which individuals would

1 develop clinically definite MS using an MRI collected less than 3 months after the patient's  
2 initial symptom onset (38). Increased scan frequency could also potentially permit earlier  
3 assessment of therapy response or detection of treatment complications, such as progressive  
4 multifocal leukoencephalopathy (16). Further studies should assess low-field MRI sensitivity for  
5 new or growing lesions over time in each clinical scenario. Our work indicates an ability to  
6 detect individual lesions at least as small as 0.06 ml using super-resolution imaging.

7  
8 Low-field MRI also offers the potential to conduct large-scale studies or screening of high-risk  
9 individuals at lower cost. In MS, high-risk asymptomatic family members have an increased  
10 incidence of neurological dysfunction and neuroimaging findings associated with MS (39).  
11 Additionally, patients with radiologically isolated syndrome (i.e., patients who meet MS criteria  
12 radiologically but are clinically asymptomatic) are known to be at high risk for development of  
13 clinical MS (40). However, studies of asymptomatic individuals require large sample sizes,  
14 which cause recruitment and cost restraints. The reduced cost of low-field MRI could  
15 significantly impact the type of population based and longitudinal studies available to researchers  
16 (37).

17  
18 While machine learning methods for MS lesion segmentation have yet to consistently outperform  
19 manual segmentation, they reduce the cost, time, and subjectivity associated with manual  
20 labeling (23,24). Combining low-field MRI with automated techniques can further address  
21 barriers to MRI access and image interpretation. In our work, the average Dice overlap between  
22 automated 3T and 64mT segmentations was only 0.23, with three subjects having no overlap.  
23 The low overlap was driven in part by hyperintense venous structures and peripheral artifacts on  
24 64mT FLAIR imaging, which also resulted in a higher number of false positives (22% for lesions  
25 >1.5ml) despite comparatively high lesion sensitivity (100% for lesion >1.5 ml). Our work  
26 illustrates the importance of reassessing algorithm performance for low-field MRI sequences and  
27 indicates that retraining or tuning may be necessary to address differences in image quality and  
28 tissue contrast. In practice, an automated segmentation could serve as a biomarker for  
29 determining eligibility or endpoints in clinical trials or as a starting point for further manual  
30 refinements.

31

1 Whether gadolinium can be used to assess contrast-enhancing lesions on the 64mT device  
2 remains unknown. In our study, we saw no contrast-enhancing lesions on 3T or 64mT. Patients  
3 at site B received a standard dose of 0.1 mmol/L of a contrast agent (gadobutrol) optimized for  
4 high-field MRI; however, since 3T imaging was performed first, there was an average delay of  
5  $58 \pm 21$  minutes between contrast administration and 64mT imaging. At lower field strengths,  
6 the benefit from T1 shortening contrast agents is significantly reduced, and a higher gadolinium  
7 dosage (41), or potentially an alternative high-relaxivity agent (41,42), could be useful for low-  
8 field MRI. Additionally, the 58-minute delay approaches gadolinium's half-life and this washout  
9 period can further attenuate the signal. Studies with minimal post-contrast delay and potentially  
10 higher doses of gadolinium could be conducted to better assess the potential for lesion  
11 enhancement at low field. However, even at higher doses, low-field devices may have reduced  
12 sensitivity to contrast-enhancing lesions (43).

13  
14 The current study has several limitations. Our findings suggest that portable 64 mT FLAIR scans  
15 will be sensitive for white matter lesions in MS and more generally, but we focused on patients  
16 with established MS and did not assess the specificity of MS lesion detection relative to other  
17 disease processes or normal controls. In addition, sensitivity at the patient or lesion level will  
18 depend on the lesion burden and size distribution (44). We used automated 3T segmentation as  
19 ground truth, though complete labeling accuracy is challenging even at high field, and we  
20 considered lesion overlap and volume rather than lesion counts. We only evaluated a single  
21 lesion segmentation method; results may not generalize to other algorithms. Indeed, our findings  
22 indicate that both image acquisition strategies and segmentation methods can be further  
23 optimized to increase the sensitivity and accuracy of low field lesion detection for larger  
24 prospective studies. We did not assess longitudinal imaging or the ability to detect new or active  
25 lesions. Gadolinium was only administered at one of the two sites and was not administered  
26 directly for 64mT imaging. As discussed above, given that none of the patients in our cohort had  
27 gadolinium enhancing lesions on their high-field scans and the post-contrast delay before each  
28 64mT scan, we cannot assess whether contrast enhancing lesions can be seen at 64mT.

29

## 30 **5. Conclusion**

31

1 In conclusion, increased imaging capabilities of low-field MRI systems warrants their re-  
2 evaluation across a range of pathologies and indications. We found that a portable 64mT scanner  
3 was sensitive to WML and that an automated algorithm designed for 3T image segmentation can  
4 be applied to 64mT data. Although additional work will be needed to evaluate portable low-field  
5 MRI systems and their capacity to carry out specific clinical functions, our findings suggest  
6 promising avenues to more accessible imaging technologies for MS around the world.

## 7 8 **Acknowledgments**

9  
10 We thank the team at Hyperfine, Inc. (Guilford, CT), particularly Jonathan Rothberg, PhD, and  
11 Edward B. Welch, PhD, for technical assistance and the use of Hyperfine low-field MRI  
12 scanners. We thank the Penn Neuroradiology Research Core for assistance with patient  
13 recruitment and scanning. We also acknowledge the staff of the NINDS Neuroimmunology  
14 Clinic; Yeajin Song, MPS; Rose Cuento, CRNP; and the staff of the NIH NMR Center.

## 15 16 **Funding**

17  
18 T. Campbell Arnold was funded in part by the HHMI-NIBIB Interfaces Initiative  
19 (5T32EB009384-10). This study received support from a research services agreement between  
20 Hyperfine, Inc. and the Trustees of the University of Pennsylvania (JMS - principal investigator).  
21 The study was partially funded by the Intramural Research Program of NINDS/NIH.

## 22 23 **Disclosures**

24  
25 Samantha By is a former employee of Hyperfine and works for Bristol Myers Squibb. Russell T.  
26 Shinohara receives consulting income from Octave Bioscience, compensation for reviewing  
27 scientific articles from the American Medical Association and for reviewing grants for the  
28 Emerson Collective, National Institutes of Health, and the Department of Defense. Daniel S.  
29 Reich is supported by the Intramural Research Program of NINDS and additional research  
30 support from Abata Therapeutics, Sanofi-Genzyme, and Vertex Pharmaceuticals. Joel M. Stein

1 has two sponsored research agreements with Hyperfine and receives consulting income from  
2 Centaur Diagnostics, Inc.

3

4

## 1   **References**

- 2
- 3   1.    Reich DS, Lucchinetti CF, Calabresi PA. Multiple Sclerosis. Longo DL, editor. *N Engl J*
- 4        *Med.* *N Engl J Med*; 2018;378(2):169–180. doi: 10.1056/NEJMra1401483.
- 5   2.    Thompson AJ, Banwell BL, Barkhof F, et al. Diagnosis of multiple sclerosis: 2017
- 6        revisions of the McDonald criteria. *Lancet Neurol.* Lancet Publishing Group; 2018. p.
- 7        162–173. doi: 10.1016/S1474-4422(17)30470-2.
- 8   3.    Filippi M, Preziosa P, Banwell BL, et al. Assessment of lesions on magnetic resonance
- 9        imaging in multiple sclerosis: practical guidelines. *Brain.* Oxford University Press; 2019.
- 10       p. 1858–1875. doi: 10.1093/brain/awz144.
- 11   4.    Noyes K, Weinstock-Guttman B. Impact of diagnosis and early treatment on the course of
- 12        multiple sclerosis. *Am. J. Manag. Care.* Ascend Media; 2013. p. s321-31.
- 13        <https://europepmc.org/article/med/24494633>. Accessed October 18, 2021.
- 14   5.    Wallin MT, Culpepper WJ, Campbell JD, et al. The prevalence of MS in the United
- 15        States: A population-based estimate using health claims data. *Neurology.* Lippincott
- 16        Williams and Wilkins; 2019;92(10):E1029–E1040. doi:
- 17        10.1212/WNL.00000000000007035.
- 18   6.    Tullman MJ. Overview of the epidemiology, diagnosis, and disease Progression
- 19        Associated With multiple Sclerosis. *Am J Manag Care.* 2013;19:S15–S20.
- 20   7.    Marques JP, Simonis FFJ, Webb AG. Low field MRI: An MR physics perspective. *J*
- 21        *Magn Reson Imaging.* John Wiley and Sons Inc.; 2019;49(6):1528–1542. doi:
- 22        10.1002/jmri.26637.
- 23   8.    Ogbole GI, Adeyomoye AO, Badu-Pepurah A, Mensah Y, Nzeh DA. Survey of magnetic
- 24        resonance imaging availability in West Africa. *Pan Afr Med J. African Field*
- 25        *Epidemiology Network*; 2018;30. doi: 10.11604/pamj.2018.30.240.14000.
- 26   9.    Maru DS-R, Schwarz R, Andrews J, Basu S, Sharma A, Moore C. Turning a blind eye: the
- 27        mobilization of radiology services in resource-poor regions. *Global Health.* 2010;6(1):18.
- 28        doi: 10.1186/1744-8603-6-18.
- 29   10.   Wald LL, McDaniel PC, Witzel T, Stockmann JP, Cooley CZ. Low-cost and portable
- 30        MRI. *J Magn Reson Imaging.* John Wiley and Sons Inc.; 2020;52(3):686–696. doi:
- 31        10.1002/jmri.26942.



- 1 11. Campbell-Washburn AE, Ramasawmy R, Restivo MC, et al. Opportunities in  
2 Interventional and Diagnostic Imaging by Using High-Performance Low-Field-Strength  
3 MRI. *Radiology*. Radiological Society of North America Inc.; 2019;293(2):384–393. doi:  
4 10.1148/radiol.2019190452.
- 5 12. Danni TU, Goyal MS, Dworkin JD, et al. Multi-modal biomarkers of cerebral edema in  
6 low resolution MRI. *bioRxiv*. Cold Spring Harbor Laboratory; 2020. p.  
7 2020.12.23.424020. doi: 10.1101/2020.12.23.424020.
- 8 13. Sheth KN, Mazurek MH, Yuen MM, et al. Assessment of brain injury using portable, low-  
9 field magnetic resonance imaging at the bedside of critically ill patients. *JAMA Neurol*.  
10 American Medical Association; 2020;E1–E7. doi: 10.1001/jamaneurol.2020.3263.
- 11 14. Mazurek MH, Cahn BA, Yuen MM, et al. Portable, bedside, low-field magnetic resonance  
12 imaging for evaluation of intracerebral hemorrhage. *Nat Commun*. Nature Publishing  
13 Group; 2021;12(1):5119. doi: 10.1038/s41467-021-25441-6.
- 14 15. Turpin J, Unadkat P, Thomas J, et al. Portable Magnetic Resonance Imaging for ICU  
15 Patients. *Crit Care Explor*. Ovid Technologies (Wolters Kluwer Health);  
16 2020;2(12):e0306. doi: 10.1097/ccx.0000000000000306.
- 17 16. Scarpazza C, Signori A, Cosottini M, Sormani MP, Gerevini S, Capra R. Should frequent  
18 MRI monitoring be performed in natalizumab-treated MS patients? A contribution to a  
19 recent debate. *Mult Scler J*. SAGE Publications Ltd; 2020;26(10):1227–1236. doi:  
20 10.1177/1352458519854162.
- 21 17. Yushkevich PA, Piven J, Hazlett HC, et al. User-guided 3D active contour segmentation  
22 of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage*.  
23 Academic Press; 2006;31(3):1116–1128. doi: 10.1016/J.NEUROIMAGE.2006.01.015.
- 24 18. Tustison NJ, Avants BB, Cook PA, et al. N4ITK: Improved N3 Bias Correction. *IEEE*  
25 *Trans Med Imaging*. 2010;29(6):1310–1320. doi: 10.1109/TMI.2010.2046908.
- 26 19. Tustison NJ, Cook PA, Holbrook AJ, et al. The ANTsX ecosystem for quantitative  
27 biological and medical imaging. *Sci Rep*. NLM (Medline); 2021;11(1):9068. doi:  
28 10.1038/s41598-021-87564-6.
- 29 20. Avants, Brian B., Tustison, Nick, Song G. Advanced normalization tools (ANTs). *Insight*  
30 *j* 2. 2009;1–35.
- 31 21. Doshi J, Erus G, Ou Y, Gaonkar B, Davatzikos C. Multi-Atlas Skull-Stripping. *Acad*

- 1 Radiol. Elsevier; 2013;20(12):1566–1576. doi: 10.1016/j.acra.2013.09.010.
- 2 22. Shinohara RT, Sweeney EM, Goldsmith J, et al. Statistical normalization techniques for  
3 magnetic resonance imaging. *NeuroImage Clin.* Elsevier Inc.; 2014;6:9–19. doi:  
4 10.1016/j.nicl.2014.08.008.
- 5 23. Valcarcel AM, Linn KA, Vandekar SN, et al. MIMoSA: An Automated Method for  
6 Intermodal Segmentation Analysis of Multiple Sclerosis Brain Lesions. *J Neuroimaging.*  
7 Blackwell Publishing Inc.; 2018;28(4):389–398. doi: 10.1111/jon.12506.
- 8 24. Valcarcel AM, Linn KA, Khalid F, et al. A dual modeling approach to automatic  
9 segmentation of cerebral T2 hyperintensities and T1 black holes in multiple sclerosis.  
10 *NeuroImage Clin.* Elsevier Inc.; 2018;20:1211–1221. doi: 10.1016/j.nicl.2018.10.013.
- 11 25. Oguz I, Carass A, Pham DL, et al. Dice overlap measures for objects of unknown number:  
12 Application to lesion segmentation. *Lect Notes Comput Sci (including Subser Lect Notes*  
13 *Artif Intell Lect Notes Bioinformatics).* Springer Verlag; 2018. p. 3–14. doi: 10.1007/978-  
14 3-319-75238-9\_1.
- 15 26. Boudraa AO, Dehak SMR, Zhu YM, Pachai C, Bao YG, Grimaud J. Automated  
16 segmentation of multiple sclerosis lesions in multispectral MR imaging using fuzzy  
17 clustering. *Comput Biol Med.* Elsevier Science Ltd; 2000;30(1):23–40. doi:  
18 10.1016/S0010-4825(99)00019-0.
- 19 27. Jovicich J, Czanner S, Han X, et al. MRI-derived measurements of human subcortical,  
20 ventricular and intracranial brain volumes: Reliability effects of scan sessions, acquisition  
21 sequences, data analyses, scanner upgrade, scanner vendors and field strengths.  
22 *Neuroimage.* Academic Press; 2009;46(1):177–192. doi:  
23 10.1016/j.neuroimage.2009.02.010.
- 24 28. O’Reilly T, Teeuwisse WM, Gans D, Koolstra K, Webb AG. In vivo 3D brain and  
25 extremity MRI at 50 mT using a permanent magnet Halbach array. *Magn Reson Med.*  
26 John Wiley and Sons Inc; 2021;85(1):495–505. doi: 10.1002/mrm.28396.
- 27 29. Sarracanie M, Salameh N. Low-Field MRI: How Low Can We Go? A Fresh View on an  
28 Old Debate. *Front Phys.* Frontiers Media S.A.; 2020;8:172. doi:  
29 10.3389/fphy.2020.00172.
- 30 30. Mateen FJ, Cooley CZ, Stockmann JP, Rice DR, Vogel AC, Wald LL. Low-field Portable  
31 Brain MRI in CNS Demyelinating Disease. *Mult Scler Relat Disord.* Elsevier;

- 1           2021;102903. doi: 10.1016/j.msard.2021.102903.
- 2   31.   Orrison WW, Stimac GK, Stevens EA, et al. Comparison of CT, low-field-strength MR  
3       imaging, and high-field-strength MR imaging. Work in progress. *Radiology*. Radiology;  
4       1991. p. 121–127. doi: 10.1148/radiology.181.1.1887020.
- 5   32.   Arnold TC, Baldassano SN, Litt B, Stein JM. Simulated diagnostic performance of low-  
6       field MRI: Harnessing open-access datasets to evaluate novel devices. *Magn Reson*  
7       Imaging. Elsevier; 2021; doi: 10.1016/j.mri.2021.12.007.
- 8   33.   Ginde AA, Foianini A, Renner DM, Valley M, Camargo CA. Availability and quality of  
9       computed tomography and magnetic resonance imaging equipment in U.S. emergency  
10      departments. *Acad Emerg Med*. *Acad Emerg Med*; 2008;15(8):780–783. doi:  
11      10.1111/j.1553-2712.2008.00192.x.
- 12   34.   Sutliff MH. Contribution of impaired mobility to patient burden in multiple sclerosis.  
13      *Curr. Med. Res. Opin*. *Curr Med Res Opin*; 2010. p. 109–119. doi:  
14      10.1185/03007990903433528.
- 15   35.   Emerging Ethical Issues Raised by Highly Portable MRI Research in Remote and  
16      Resource-Limited International Settings Shen FX, Wolf SM, Bhavnani S, et al. Emerging  
17      Ethical Issues Raised by Highly Portable MRI Research in Remote and Resource-Limited  
18      International Settings. *Neuroimage*. Academic Press; 2021;118210. doi:  
19      10.1016/j.neuroimage.2021.118210.
- 20   36.   Deoni SC, Deoni AT, Burton P, et al. Residential MRI: Development of A Mobile  
21      Anywhere-Everywhere MRI Lab. 2021; doi: 10.21203/rs.3.rs-1121934/v1.
- 22   37.   Deoni SCL, D’Sa V, Volpe A, et al. Remote and At-Home Data Collection:  
23      Considerations for the NIH HEALthy Brain and Cognitive Development (HBCD) Study.  
24      *Dev Cogn Neurosci*. Elsevier; 2022;101059. doi: 10.1016/j.dcn.2022.101059.
- 25   38.   Rovira À, Swanton J, Tintoré M, et al. A single, early magnetic resonance imaging study  
26      in the diagnosis of multiple sclerosis. *Arch Neurol*. *Arch Neurol*; 2009;66(5):587–592.  
27      doi: 10.1001/archneurol.2009.49.
- 28   39.   Xia Z, Steele SU, Bakshi A, et al. Assessment of early evidence of multiple sclerosis in a  
29      prospective study of asymptomatic high-risk family members. *JAMA Neurol*. American  
30      Medical Association; 2017;74(3):293–300. doi: 10.1001/jamaneurol.2016.5056.
- 31   40.   Hosseiny M, Newsome SD, Yousem DM. Radiologically isolated syndrome: A review for

- 1           neuroradiologists. *Am. J. Neuroradiol.* American Society of Neuroradiology; 2020. p.  
2           1542–1549. doi: 10.3174/ajnr.A6649.
- 3 41.   Desai NK, Runge VM. Contrast Use at Low Field: A Review. *Top Magn Reson Imaging.*  
4           2003;14(5):360–364. doi: 10.1097/00002142-200310000-00002.
- 5 42.   Bendszus M, Roberts D, Kolumban B, et al. Dose Finding Study of Gadopiclesol, a New  
6           Macrocyclic Contrast Agent, in MRI of Central Nervous System. *Invest Radiol.*  
7           Lippincott Williams and Wilkins; 2020;55(3):129–137. doi:  
8           10.1097/RLI.0000000000000624.
- 9 43.   Ertl-Wagner B, Reith W, Sartor K. Low field-low cost: Can low-field magnetic resonance  
10          systems replace high-field magnetic resonance systems in the diagnostic assessment of  
11          multiple sclerosis patients? *Eur Radiol.* Springer; 2001;11(8):1490–1494. doi:  
12          10.1007/s003300000806.
- 13 44.   Commowick O, Istace A, Kain M, et al. Objective Evaluation of Multiple Sclerosis Lesion  
14          Segmentation using a Data Management and Processing Infrastructure. *Sci Rep.* Nature  
15          Publishing Group; 2018;8(1). doi: 10.1038/s41598-018-31911-7.
- 16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31

## 1 **Supplemental Material**

2

### 3 *Study size calculation*

4

5 Based on preliminary results, the maximum diameter (Dmax) of the smallest lesion on 3T  
6 imaging was estimated to be 2 mm, while Dmax was estimated to be 6 mm on 64mT imaging.

7 The standard deviation for manual measurements was estimated to be 2 mm. Using a

8 significance level of 0.01 and a power level of 0.95, we estimated the necessary sample size to be

9 18 patients to detect a significant difference between modalities (Eng 2003). Our study is well

10 powered with over 30 patients.

11

12 1. Eng J. Sample size estimation: how many individuals should be studied? Radiology. 2003

13 May;227(2):309-13. Doi: 10.1148/radiol.2272012051. PMID: 12732691.

14

15

16

17

18

19

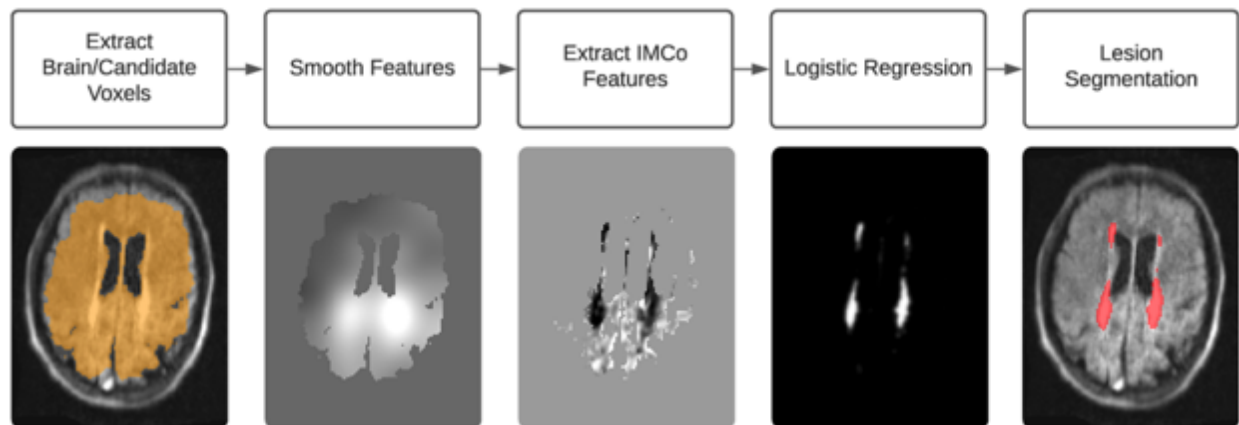
20

21

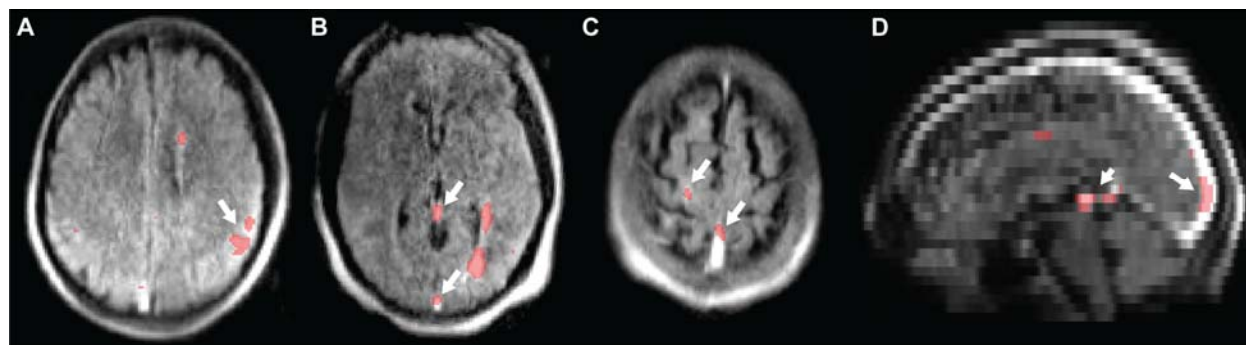
22

23

24

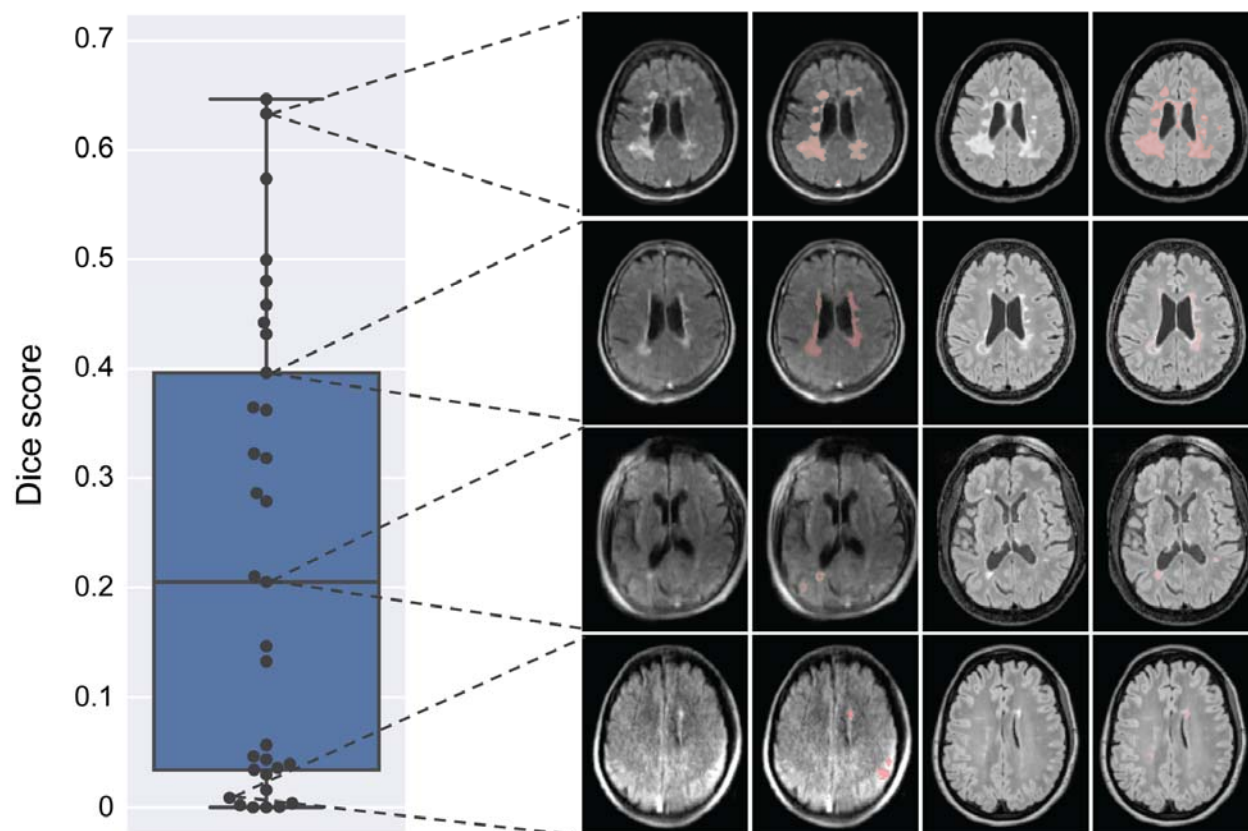


1  
2 **Figure S1. Automated segmentation algorithm.** Visual description of the steps in the  
3 MIMoSA algorithm.  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18



1  
2 **Figure S2. Examples of false positive lesion segmentations on 64mT T2-FLAIR.** (A) False  
3 positive detections (long arrow) caused by artifactual peripheral hyperintensity in a 31-35 year-  
4 old female with RRMS. False positive detections (long arrows) caused by hyperintense venous  
5 structures as seen in (B) a 31-35 year-old female with RRMS, (C) a 51-55 year-old male with  
6 RRMS, and (D) a 46-50 year-old male with clinically isolated syndrome. A hyperintense pineal  
7 cyst (short arrow) was also a source of false positive labeling in D.

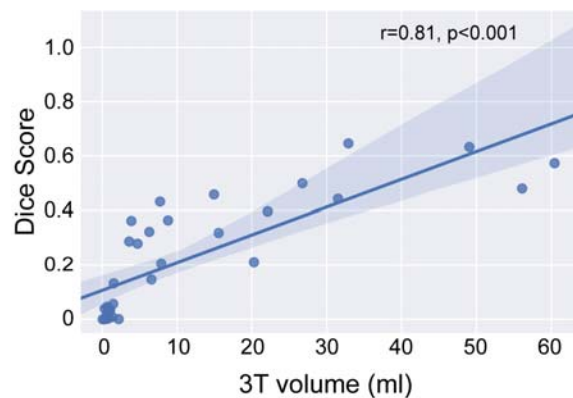
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19



1  
2 **Figure S3. Dice score distribution of lesion segmentations.** This figure highlights patients  
3 from each quartile of the Dice score distribution (mean = 0.23, standard deviation = 0.21). From  
4 left to right, the images are the original 64mT image, 64mT image with segmentation overlay,  
5 the original 3T image, and 3T image with segmentation overlay. Even patients in the lowest  
6 quartile demonstrate overlapping periventricular lesions.

7  
8  
9  
10  
11  
12  
13





1  
2 **Figure S4. Dice score increases with lesion volume.** Previous studies have found that larger  
3 lesions size are associated with higher Dice scores (25). In our study, we found a similar effect ( $r$   
4 = 0.81,  $p < 0.001$ ) such that subjects with higher lesion burden had correspondingly higher Dice  
5 scores.

6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25

1 *False positive detections*

2

3 Visual inspection of automated 64mT segmentations revealed some consistent sources of false  
4 positive lesion detection on FLAIR images, including flow-related hyperintensities in venous  
5 structures (dural venous sinuses and small cerebral or cortical veins), other hyperintense  
6 structures such as pineal cysts, and peripheral hyperintensities due to image artifacts. Examples  
7 of these artifacts can be seen in figure E3.

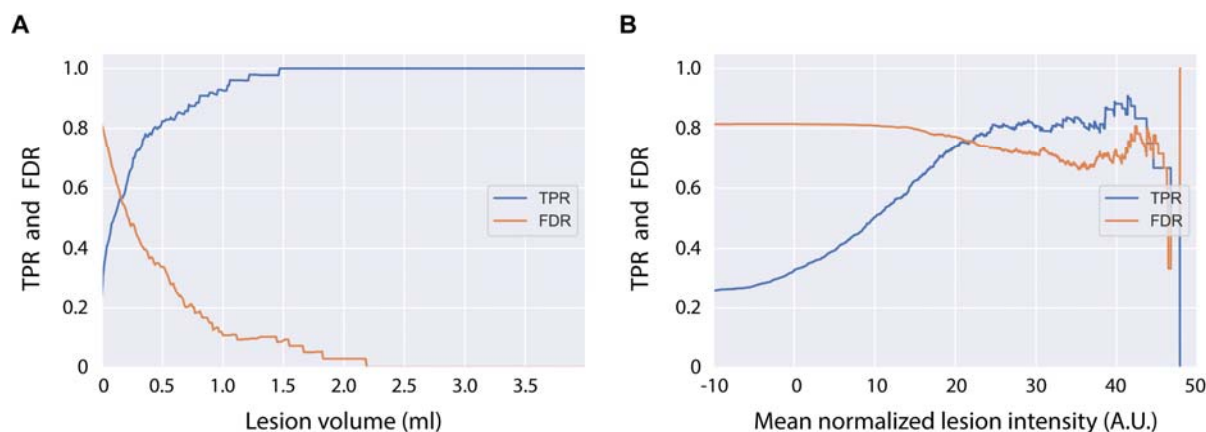
8

9 In a sub-analysis, automated segmentations were manually edited in ITK-SNAP to remove some  
10 of the common false positives. In particular, we removed false positives in major sinuses (i.e.,  
11 sagittal, straight, and transverse) as well as hyperintense pineal cysts. Peripheral hyperintensities  
12 caused by image artifacts or smaller cortical veins were not edited, as classification of these  
13 detections is more subjective. After manually editing the segmentations, the true positive rate  
14 (TPR) and false discovery rate (FDR) were recalculated.

15

16 There was a sharp decrease in the FDR as a function of lesion size after manual editing, reaching  
17 11% at >1.0ml compared to 36% in the original automated segmentation (Fig. E6A). There was a  
18 modest decrease in FDR as a function of lesion intensity; however, FDR remains >65% across  
19 sensitivities (>75% in automated segmentations) (Fig. E6B). TPR remain relatively unchanged in  
20 both plots. These results highlight how optimized FLAIR sequences without hyperintense veins  
21 could decreased the rate of false positive detections by automated algorithms.

22



23

1 **Figure S5. False discovery rate decreases after removing hyperintense veins.** After manually  
2 removing hyperintense veins and other hyperintense brain structures from segmentations, the  
3 false discovery rate (FDR) decreases. (A) The FDR decreases as a function of lesion size.  
4 Manually edited segmentations had a lower FDR relative to the unedited automated  
5 segmentations (>1 ml: unedited=36%, edited=11%, >1.5ml: unedited=22%, edited=9%, >2.5ml:  
6 unedited=3%, edited=0%). (B) When analyzing FDR as a function of mean lesion intensity, the  
7 effect was less pronounced. In unedited segmentations, FDR was high (>75%) across lesion  
8 intensities. While there was a modest reduction after manual editing, FDR remained >65%.  
9