Supplement 1. Methods

Subjects, Inclusion criteria: Inclusion criteria for patients were: RR or SP disease course, age 18–65 years, and Expanded Disability Status Scale (EDSS) between 0 and 6.5. Patients with relapse or steroid treatment within 30 days before MRI were excluded. Participants were excluded if they were pregnant or had pre-existing medical conditions known to be associated with brain pathology. NCs were volunteers with a normal neurological examination and no history of neurologic disorders or psychiatric disorders.

Segmentation, volumetry: As was described previously [Hagemeier et al., 2018], anatomical DGM regions were segmented with FSL FIRST [Patenaude et al., 2011] on 3D T_1-weighted images that were first inpainted to avoid T_1-hypointensity misclassification [Gelineau-Morel et al., 2012]. DGM volumes were normalized using FMRIB's cross-sectional software tool SIENAX-derived scaling factor [Smith et al., 2002]. Longitudinal whole-brain volume changes were assessed using SIENA [Smith et al., 2001], while specific tissue volume changes were calculated with SIENAX-MTP [Dwyer et al., 2014]. T_2-lesions were identified on T_2-weighted/FLAIR images using a semi-automated edge detection contouring and thresholding technique, as described previously [Zivadinov et al., 2001].

QSM: We used the same reconstructed susceptibility maps as in our previous study [Hagemeier et al., 2018]. In brief, phase images were unwrapped with a best-path algorithm [Abdul-Rahman et al., 2007], background-field corrected with V-SHARP [Schweser et al., 2011; Wu et al., 2012], converted to magnetic susceptibility maps using the HEIDI algorithm [Schweser et al., 2012], and referenced to the whole brain.

Conversion to iron concentration: We assumed the following linear tissue model for magnetic susceptibility that was used previously for relating iron concentrations to magnetic susceptibility [Langkammer et al., 2012]:

$$\chi = \hat{\chi}_{\mathrm{Fe}} \cdot c_{\mathrm{Fe}} + \chi_{\mathrm{other}},$$

where c_{Fe} is the local tissue non-heme iron mass concentration (in mg/ml), $\hat{\chi}_{\mathrm{Fe}}$ is a proportionality constant that depends on the effective magnetic moment of the iron complexes in the tissue (in ppm·ml/mg) [Langkammer et al., 2012], and χ_{other} comprises all susceptibility contributions from the tissue matrix, including the shift-invariance of susceptibility maps obtained with QSM. Using this model, we converted magnetic susceptibility values to iron concentrations according to:

$$c_{\mathrm{Fe}} = \frac{\chi - \chi_{\mathrm{other}}}{\hat{\chi}_{\mathrm{Fe}}}.$$

The coefficients $\hat{\chi}_{\mathrm{Fe}}$ and χ_{other} were determined from linear regression of the region and age-dependent histochemical iron concentrations in the thalamus, caudate, putamen, and globus pallidus as reported by [Hallgren and Sourander, 1958] to the corresponding susceptibility values measured in our NCs (n=40). Hallgren and Sourander reported iron concentrations in mass fractions (mg per 100g tissue wet weight; mg/100g-ww). Here, we assumed a tissue density of 1g/ml for the conversion from tissue wet-weight to volume. We emphasize that the conversion factor $\hat{\chi}_{\mathrm{Fe}}$ serves only to arrive at a physically meaningful unit of the "iron content" metric (see below) but does not affect the statistical analyses and results.

Supplement 2. Cross-sectional iron content of normal controls and patients with MS.

Table 1 and Supplement Figure 1a summarize the calculated values of iron content in each DGM region studied cross-sectionally for NCs and MS patients. Table 2 and Supplement Figure 1b summarize the calculated values of iron content in each DGM region when comparing RR- and SPMS patients. Again, cross-sectional differences were found in the thalamus, with a significantly lower iron content observed in SPMS compared to RRMS patients, both at baseline (-23.7%, p=0.003) and follow-up (-20.9%, q=0.013). All other DGM structures, at either baseline or follow-up, were highly similar between RRMS and SPMS, differing less than 4.9% (Figure 1b).

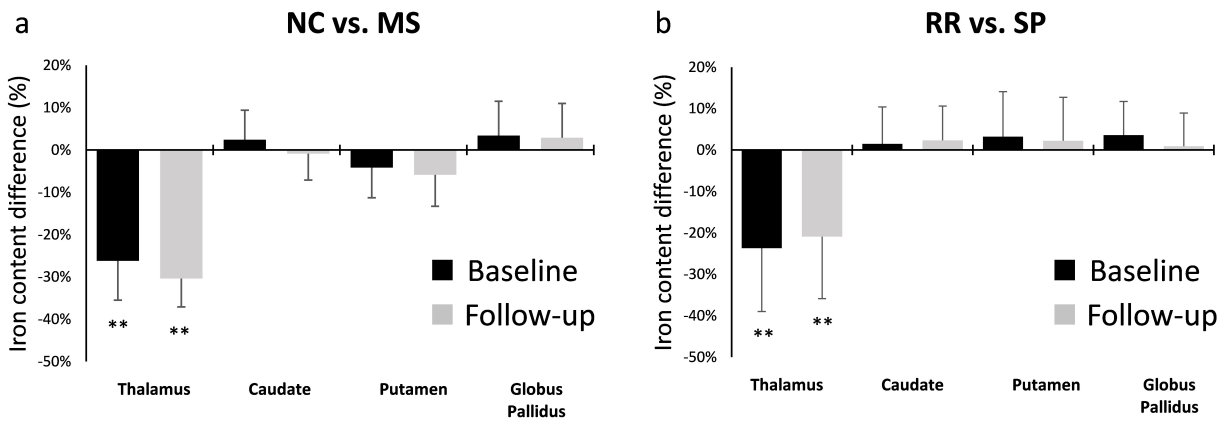
In both healthy individuals and patients with MS, the highest iron *concentration* is typically found in the globus pallidus, with other regions showing substantially lower concentrations (putamen: 50% lower; caudate: 65% lower) [Hagemeier et al., 2018]. In contrast, iron *content* was the highest in the putamen, and all other regions showed substantially lower iron content. This observation is a result of the different volumes of the structures.

Comparing the *cross-sectional* iron concentration results to iron content (Supplement Figure 2) reveals that DGM iron concentrations depend on disease status, with higher values in more advanced disease stages (MS > NC, SP > RR) in the caudate, putamen and globus pallidus. In these structures, volume reductions (atrophy) followed a similar trajectory [Hagemeier et al., 2018]. When investigating iron content, however, a disease progression dependency could not be observed in caudate, putamen, and globus pallidus (Fig. 1). In the thalamus, MS patients had both lower iron *concentration and content* (Supplement Table 2), cementing the thalamus as a structure with substantial decreases in overall iron, as previously suggested [Bergsland et al., 2018; Burgetova et al., 2017; Hagemeier et al., 2018; Schweser et al., 2018; Zivadinov et al., 2018]. Furthermore, Hernández-Torres et al. recently studied *iron content* in MS patients with a similar

methodology as employed in the present work (but based on R_2* instead of QSM and using a cross-sectional study design) and also reported lower iron content in the thalamus [Hernández-Torres et al., 2019]. In terms of iron concentration, several other studies did not find differences in the thalamus [Al-Radaideh et al., 2013; Langkammer et al., 2013], or found higher thalamic magnetic susceptibility in MS [Rudko et al., 2014]. Discrepant results may be due to technical differences in the employed QSM techniques or due to demographic and clinical differences of the cohorts (as discussed in [Schweser et al., 2018]). In particular, the latter may lead to different findings because thalamic iron follows a peculiar non-linear aging trajectory *in controls* [Hallgren and Sourander, 1958] with a reduction in iron *concentration* starting around 30-40 years of age.

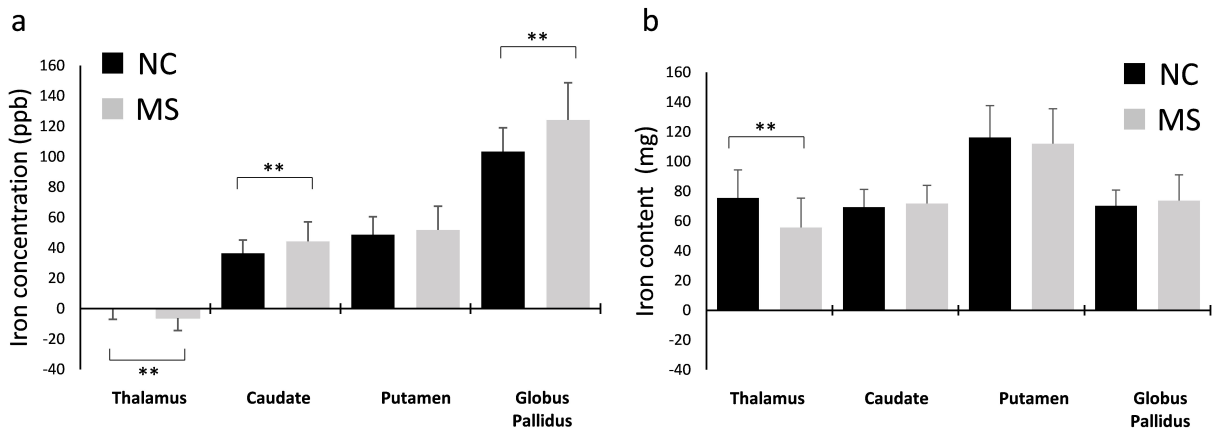
Taken together, our cross-sectional results support the theory of iron depletion in the thalamus. Considering that in MS the thalamus is one of the regions that is most affected by atrophy [Azevedo et al., 2018; Zivadinov et al., 2013], the fact that substantial *reductions* in iron content are still observed, suggests that the thalamus is especially susceptible to iron loss in middle-aged MS patients.

Supplement Figure 1.



Cross-sectional percent group difference of regional iron content at both baseline (black) and two-year follow-up (gray). (a) Difference between patients with multiple sclerosis (MS) and normal controls (NC) relative to the iron content in NCs (negative values mean less iron in patients). (b) Difference between secondary progressive (SP) MS and relapsing-remitting (RR) MS relative to RR (negative means less iron in SP). * unadjusted $p<0.05$. ** false discovery rate adjusted $q<0.05$. Error bars represent the 95% confidence interval of the change.

Supplement Figure 2.



Comparison of (a) baseline iron concentration (ppb) in multiple sclerosis patients (MS; gray) and normal controls (NC; black) see [Hagemeier et al., 2018], and (b) baseline iron content (mg) in patients (gray) and NC (black).

Supplement Table 1. Clinical, demographic, and volumetric MRI data at baseline and change over follow-up. This table is mostly identical to Table 1 in our previous work [Hagemeier et al., 2018].

	Controls	Multiple Sclerosis	Relapsing MS	Progressive MS	p-value (NC v MS)	p-value (RR v SP)
N	40	120	98	22	NA	NA
Female, n (%)	24 (60%)	81 (67.5%)	69 (70.4%)	12 (54.5%)	0.387[b]	0.151[b]
Age, years	43.7 ± 12.3	44.2 ± 10.2	42.0 ± 9.6	53.7 ± 7.2	0.819[c]	<0.001[c]
Baseline EDSS, median ± IQR	NA	2.5 ± 1.5 – 4.0	2.5 ± 1.5 – 3.5	6.0 ± 3.5 – 6.5	NA	<0.001[d]
Follow-up EDSS, median ± IQR	NA	3.0 ± 1.5 – 4.0	2.5 ± 1.5 – 3.5	6.0 ± 4.0 – 6.5	NA	<0.001[d]
Disease duration in years	NA	12.8 ± 9.4	11.2 ± 8.4	19.7 ± 10.6	NA	0.002[c]
Baseline relapses in the last year, median ± IQR	NA	0 ± 0 – 1	0 ± 0 – 1	0 ± 0 – 0	NA	0.041[d]
Follow-up relapses in the last year, median ± IQR (sum)	NA	0 ± 0 – 0 (29)	0 ± 0 – 0	0 ± 0 – 0	NA	0.310[d]
Treated with DMT, n (%)[†]	NA	103 (85.3%)			NA	
Follow-up time, years	1.9 ± 1.2	2.1 ± 1.11	2.0 ± 1.1	2.2 ± 1.1	0.516[c]	0.579[c]
Baseline T_2 lesion volume, ml	0.3 ± 0.7	13.9 ± 18.1	12.6 ± 17.9	19.4 ± 18.1	<0.001[c*]	0.157[c*]
Follow-up T_2 lesion volume, ml	0.7 ± 1.7	12.6 ± 14.8	11.6 ±14.6	17.4 ± 15.2	<0.001[c*]	0.157[c*]
Whole brain volume, ml	1,532.0 ± 76.2	1,457.8 ± 91.4	1,471.8 ± 87.6	1,395.8 ± 83.2	<0.001[c*]	<0.001[c*]
Gray matter volume, ml	768.0 ± 49.2	743.6 ± 54.6	751.4 ± 52.5	708.9 ± 51.3	0.048[c*]	<0.001[c*]
White matter volume, ml	770.1 ± 43.8	723.7 ± 52.5	720.4 ± 50.2	686.9 ± 47.8	<0.001[c*]	0.017[c*]
Lateral ventricle volume, ml	33.6 ± 16.2	48.1 ± 20.9	46.4 ± 22.1	55.7 ± 12.5	<0.001[c*]	0.145[c*]
PBVC	-0.71 (.82)	-1.12 (1.16)	-1.19 ±1.22	-0.80 ± 0.78	0.139[c*]	0.152[c*]
PGMVC	-1.29 (2.04)	-1.45 (1.66)	-1.54 ± 1.69	-1.00 ± 1.43	0.807[c*]	0.293[c*]
PWMVC	0.29 (1.54)	0.50 (2.07)	0.46 ± 2.11	0.73 ± 1.89	0.764[c*]	0.703[c*]
PVVC	3.74 (6.11)	5.12 (7.40)	1.95 ± 8.68	1.54 ± 4.48	0.535[c*]	0.784[c*]

Abbreviations: NA; not applicable; NC = normal control; RR = relapsing-remitting; SP = secondary progressive; EDSS = Expanded Disability Status Scale; IQR = inter-quartile range; DMT = disease modifying therapy; PBVC = percent brain volume change, PVVC = percent ventricle volume change; PGMVC = percent gray matter volume change; PWMVC = percent white matter volume change. Results are based on baseline and percent change data and are presented as mean ± SD unless otherwise noted. Brain volumes were corrected for head size.

† Patients were taking the following disease-modifying treatments: interferon beta: 40, natalizumab: 21, glatiramer acetate: 34, combination: 3, other: 5, no therapy: 17. 11 (9.2%) had a change in DMT from baseline to follow-up.
[a] Patients not identified as RRMS were secondary progressive (SP).
[b] Chi-squared test.
[c] Independent student *t*-test.
[d] Mann-Whitney U test
* p-values were corrected using false discovery rate (q-values).

Supplement Table 2. Comparison of significant main findings between our current work (iron content) and previous work (susceptibility and volume) [Hagemeier et al., 2018].

	Cross-sectional			Longitudinal		
	Iron Content	**Susceptibility†**	**Volume‡**	**Iron Content**	**Susceptibility†**	**Volume‡**
Thalamus	MS < NC SP < RR	MS < NC	MS < NC	MS: ↓ RR: ↓		MS: ↓
Caudate		MS > NC	MS < NC	MS: ↓ RR: ↓	NC: ↑ MS: ↑ RR: ↑	MS: ↓
Putamen		SP > RR	MS < NC	MS: ↓	RR: ↑	MS: ↓
Globus Pallidus		MS > NC SP > RR	MS < NC	MS: ↓ RR: ↓ SP: ↓		MS: ↓

† [Hagemeier et al., 2018], Table 2 and Supplement Table 2

‡ [Hagemeier et al., 2018], Supplement Table 1. Not investigated between RR and SPMS.

Supplement Table 3. Associations of iron concentration, normalized structural volume, and iron content *at baseline* with baseline EDSS, change in EDSS over two years, and disease duration, respectively, within patients with MS.

	Baseline iron concentration						**Baseline normalized volume**						**Baseline iron content**					
	Baseline EDSS		**Change in EDSS**		**Disease duration**		**Baseline EDSS**		**Change in EDSS**		**Disease duration**		**Baseline EDSS**		**Change in EDSS**		**Disease duration**	
	β	**p**	**β**	**p**	**β**	**p**	**β**	**p**	**β**	**p**	**β**	**p**	**β**	**p**	**β**	**p**	**β**	**p**
Thalamus	-0.043	0.658	**-0.257**	**0.006**	-0.131	0.237	**-0.330**	**0.001**	-0.075	0.397	**-0.379**	**0.001**	-0.175	0.064	**-0.197**	**0.025**	**-0.302**	**0.005**
Caudate	0.004	0.963	0.081	0.373	-0.161	0.142	-0.092	0.337	**-0.185**	**0.045**	-0.164	0.138	-0.103	0.287	0.022	0.802	**-0.349**	**0.002**
Putamen	-0.002	0.986	0.023	0.798	**-0.236**	**0.032**	-0.132	0.174	**-0.185**	**0.046**	-0.201	0.072	-0.098	0.321	-0.022	0.808	**-0.39**	**0.0001**
Pallidus	0.035	0.721	**0.208**	**0.027**	-0.022	0.845	**-0.203**	**0.039**	-0.108	0.263	-0.153	0.189	-0.145	0.139	0.108	0.252	-0.177	0.123

Note: Regression analysis models were adjusted for age and sex. β = standardized beta. Entries with $p<0.05$ printed in boldface.

Supplement Table 4. Associations of *changes over two years* in iron concentration, normalized structural volume, and iron content with EDSS change and disease duration, respectively, within patients with MS.

	Change in iron concentration						**Change in normalized volume**						**Change in iron content**					
	Baseline EDSS		**Change in EDSS**		**Disease duration**		**Baseline EDSS**		**Change in EDSS**		**Disease duration**		**Baseline EDSS**		**Change in EDSS**		**Disease duration**	
	β	**p**	**β**	**p**	**β**	**p**	**β**	**β**	**B**	**p**	**β**	**p**	**β**	**p**	**B**	**p**	**β**	**p**
Thalamus change	-0.009	0.931	0.172	0.077	0.155	0.185	0.017	0.870	**-0.209**	**0.032**	0.105	0.369	-0.004	0.969	0.132	0.18	0.149	0.204
Caudate change	-0.074	0.462	0.038	0.695	-0.206	0.082	-0.75	0.457	0.016	0.871	0.012	0.919	-0.115	0.258	0.067	0.494	-0.19	0.11
Putamen change	-0.093	0.361	0.102	0.301	0.02	0.866	-0.073	0.462	-0.045	0.642	-0.063	0.581	-0.088	0.384	0.07	0.482	-0.043	0.718
Pallidus change	-0.027	0.794	0.027	0.786	-0.118	0.319	-0.142	0.161	0.064	0.515	0.126	0.288	-0.063	0.532	0.05	0.617	-0.031	0.794

Note: Regression analysis models were adjusted for age and sex. β = standardized beta. Entries with $p<0.05$ printed in boldface.

References

Abdul-Rahman HS, Gdeisat MA, Burton DR, Lalor MJ, Lilley F, Moore CJ (2007): Fast and robust three-dimensional best path phase unwrapping algorithm. Appl Opt 46:6623–35.

Al-Radaideh AM, Wharton SJ, Lim S-Y, Tench CR, Morgan PS, Bowtell RW, Constantinescu CS, Gowland P a (2013): Increased iron accumulation occurs in the earliest stages of demyelinating disease: an ultra-high field susceptibility mapping study in Clinically Isolated Syndrome. Mult Scler 19:896–903.

Azevedo CJ, Cen SY, Khadka S, Liu S, Kornak J, Shi Y, Zheng L, Hauser SL, Pelletier D (2018): Thalamic atrophy in multiple sclerosis: A magnetic resonance imaging marker of neurodegeneration throughout disease: Thalamic Atrophy in MS. Ann Neurol 83:223–234.

Bergsland N, Schweser F, Dwyer MG, Weinstock-Guttman B, Benedict RHB, Zivadinov R (2018): Thalamic white matter in multiple sclerosis: A combined diffusion-tensor imaging and quantitative susceptibility mapping study. Hum Brain Mapp 39:4007–4017.

Burgetova A, Dusek P, Vaneckova M, Horakova D, Langkammer C, Krasensky J, Sobisek L, Matras P, Masek M, Seidl Z (2017): Thalamic Iron Differentiates Primary-Progressive and Relapsing-Remitting Multiple Sclerosis. Am J Neuroradiol 38:1079–1086.

Dwyer MG, Bergsland N, Zivadinov R (2014): Improved longitudinal gray and white matter atrophy assessment via application of a 4-dimensional hidden Markov random field model. NeuroImage 90:207–217.

Gelineau-Morel R, Tomassini V, Jenkinson M, Johansen-Berg H, Matthews PM, Palace J (2012): The effect of hypointense white matter lesions on automated gray matter segmentation in multiple sclerosis. Hum Brain Mapp 33:2802–2814.

Hagemeier J, Zivadinov R, Dwyer MG, Polak P, Bergsland N, Weinstock-Guttman B, Zalis J, Deistung A, Reichenbach JR, Schweser F (2018): Changes of deep gray matter magnetic susceptibility over 2 years in multiple sclerosis and healthy control brain. NeuroImage Clin 18:1007–1016.

Hallgren B, Sourander P (1958): The effect of age on the non-haemin iron in the human brain. J Neurochem 3:41–51.

Hernández-Torres E, Wiggermann V, Machan L, Sadovnick AD, Li DKB, Traboulsee A, Hametner S, Rauscher A (2019): Increased mean R2* in the deep gray matter of multiple sclerosis patients: Have we been measuring atrophy? J Magn Reson Imaging 50:201–208.

Langkammer C, Liu T, Khalil M, Enzinger C, Jehna M, Fuchs S, Fazekas F, Wang Y, Ropele S (2013): Quantitative susceptibility mapping in multiple sclerosis. Radiology 267:551–9.

Langkammer C, Schweser F, Krebs N, Deistung A, Goessler W, Scheurer E, Sommer K, Reishofer G, Yen K, Fazekas F, Ropele S, Reichenbach JR (2012): Quantitative susceptibility mapping (QSM) as a means to measure brain iron? A post mortem validation study. NeuroImage 62:1593–1599.

Patenaude B, Smith SM, Kennedy DN, Jenkinson M (2011): A Bayesian model of shape and appearance for subcortical brain segmentation. NeuroImage 56:907–22.

Rudko DA, Solovey I, Gati JS, Kremenchutzky M, Menon RS (2014): Multiple Sclerosis: Improved Identification of Disease-relevant Changes in Gray and White Matter Using Susceptibility-based MR Imaging. Radiology:132475.

Schweser F, Deistung A, Lehr BW, Reichenbach JR (2011): Quantitative imaging of intrinsic magnetic tissue properties using MRI signal phase: An approach to in vivo brain iron metabolism? NeuroImage 54:2789–2807.

Schweser F, Raffaini Duarte Martins AL, Hagemeier J, Lin F, Hanspach J, Weinstock-Guttman B, Hametner S, Bergsland NP, Dwyer MG, Zivadinov R (2018): Mapping of thalamic magnetic susceptibility in multiple sclerosis indicates decreasing iron with disease duration: A proposed mechanistic relationship between inflammation and oligodendrocyte vitality. NeuroImage 167:438–452.

Schweser F, Sommer K, Deistung A, Reichenbach JR (2012): Quantitative susceptibility mapping for investigating subtle susceptibility variations in the human brain. NeuroImage 62:2083–2100.

Smith SM, De Stefano N, Jenkinson M, Matthews PM (2001): Normalized accurate measurement of longitudinal brain change. J Comput Assist Tomogr 25:466–475.

Smith SM, Zhang Y, Jenkinson M, Chen J, Matthews PM, Federico A, De Stefano N (2002): Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. NeuroImage 17:479–89.

Wu B, Li W, Avram AV, Gho S-M, Liu C (2012): Fast and tissue-optimized mapping of magnetic susceptibility and T2* with multi-echo and multi-shot spirals. NeuroImage 59:297–305.

Zivadinov R, Rudick RA, De Masi R, Nasuelli D, Ukmar M, Pozzi-Mucelli RS, Grop A, Cazzato G, Zorzon M (2001): Effects of IV methylprednisolone on brain atrophy in relapsing-remitting MS. Neurology 57:1239–1247.

Zivadinov R, Havrdová E, Bergsland N, Tyblova M, Hagemeier J, Seidl Z, Dwyer MG, Vaneckova M, Krasensky J, Carl E, Kalincik T, Horáková D (2013): Thalamic Atrophy Is Associated with Development of Clinically Definite Multiple Sclerosis. Radiology 268:831–841.

Zivadinov R, Tavazzi E, Bergsland N, Hagemeier J, Lin F, Dwyer MG, Carl E, Kolb C, Hojnacki D, Ramasamy D, Durfee J, Weinstock-Guttman B, Schweser F (2018): Brain Iron at Quantitative MRI Is Associated with Disability in Multiple Sclerosis. Radiology 289:487–496.