eAppendix 1

Search strategies

PubMed Medline search October 11, 2019.

Yield: 1787 titles

[Mesh] = Medical subject headings

[tiab] = words in title or abstract or author keywords

#1	"Multiple Sclerosis"[Mesh] OR multiple sclerosis[tiab] OR ms[tiab]
#2	(("Gray Matter"[Mesh] OR gray[tiab] OR grey[tiab] OR cortical[tiab] OR subcortical[tiab] OR brain[tiab]) AND ("Atrophy"[Mesh] OR atroph*[tiab] OR volum*[tiab])) OR cortical thickness[tiab]
#3	"White Matter"[Mesh] OR "Nerve Fibers, Myelinated"[Mesh] OR white[tiab] OR (t2[tiab] AND lesion*[tiab]) OR focal lesion*[tiab] OR lesion load[tiab] OR lesion burden[tiab]
#4	#1 AND #2 AND #3

Embase.com search October 11, 2019.

Yield: 1963 titles.

/exp = EMtree keyword with explosion
:ab,ti,kw = words in title or abstract or author keywords
NEXT/n = Requests terms that are within n words of each other in the order specified

1	('multiple sclerosis'/exp OR 'multiple sclerosis':ti,ab,kw OR ms:ti,ab,kw)
2	(('gray matter'/exp OR gray:ti,ab,kw OR grey:ti,ab,kw OR cortical:ti,ab,kw OR subcortical:ti,ab,kw OR brain:ti,ab,kw) AND ('atrophy'/de OR 'brain atrophy'/exp OR 'brain cortex atrophy'/exp OR 'adrenal cortex atrophy'/exp OR atroph*:ti,ab,kw OR volum*:ti,ab,kw) OR 'cortical thickness':ti,ab,kw)
3	'white matter'/exp OR 'white matter lesion'/exp OR 'myelinated nerve'/exp OR white:ti,ab,kw OR (t2 NEXT/3 lesion*):ti,ab,kw OR 'focal lesion*':ti,ab,kw OR 'lesion load':ti,ab,kw OR 'lesion burden':ti,ab,kw
4	1 and 2 and 3
5	4 not 'conference abstract'/it

To remove duplicates, records were imported into Endnote X9.2 (Clarivate Analytics, Philadelphia, PA).

eAppendix 2

Expanded results section

CIS

Eight cross-sectional and four longitudinal studies investigated patients diagnosed with clinically isolated syndrome (CIS). In the longitudinal studies, the follow-up period ranged from two to five and a half years.

The association of lesions with global gray matter (GM) measures were reported in five studies, while cortical and deep GM measures were each considered in seven studies. Four studies reported on regional white matter (WM) lesion measures.

Included studies are described in supplemental eTables 1, 2 and 3, available from Dryad.

Global GM in CIS

In two out of three cross-sectional CIS studies, no significant association was found between global GM volume and either T2[e3] or T1[e4] lesion volume (LV). One study did find such a relation: T2 LV and global GM volume was significantly correlated (r=-0.56, $p<0.020$)[e5]. In this study, 17 patients with CIS were included, and the association was not corrected for age and sex, which was done in the two studies with absent associations.

The longitudinal relationship between global WM lesion measures and global GM atrophy was reported in two studies, and both observed somewhat different, but significant associations. In one of the studies, change in GM fraction over three years correlated with the change in LV (T1: r=-0.307, $p=0.0426$, T2: r=-0.4280, $p=0.0032$), but not with baseline lesion measures[e6]. In the other available study, significant associations were found between global GM percentage change and baseline lesion measures ($p\leq0.004$) and total cumulative number of new/enlarging T2 lesions ($p=0.013$), but not with changes in LV during the 48 months

follow-up[e7]. In the first study, only three out of 58 included patients received disease modifying therapy (DMT) in the form of interferon-beta, while in the other study, all 210 patients received intramuscular interferon-beta once a week starting from the study baseline.

Cortical GM in CIS

In cross-sectional studies on patients with CIS, lower cortical GM (CGM) volume showed variable associations with global WM lesion measures.

Two studies observed a significant relation with the presence (t=2.48, p=0.020)[e8] or volume (r=-0.49, p=0.045)[e5] of T2 lesions, while three studies did not[e2,e3,e8]. In one of these studies lower pericalcarine cortical volumes were observed in patients with optic neuritis, in the presence of whole brain T2 lesions (t= 2.48, p=0.020) and T2 lesions in the optic radiation (OR) (t= 2.24, p=0.034). However, pericalcarine volume and thickness did not correlate with whole brain or OR T2 LV[e8]. Another study reporting on regional WM lesion measures observed no significant correlation with whole-brain average cortical thickness, but did observe such associations for regional cortical thickness measures ($p \leq 0.0466$) in vertexwise analyses[e9].

Of the two available longitudinal studies, one study found significant associations between cortical volume change and baseline WM lesion measures ($p \leq 0.004$) and the total cumulative number of new/enlarging T2 lesions (p=0.036) over the 48-month follow-up, however no such associations were observed for changes in LV[e7]. The second study assessed a novel, more unconventional lesion measure; atrophied T2 LV defined as T2-weighted lesional tissue subsequently substituted by CSF. Over a five-and-a-half-year follow-up period, no association with neither baseline cortical volume nor volume change was found[e10].

Deep GM in CIS

In the five available cross-sectional studies in CIS patients, all except one[e3] showed significant associations between global[e2,e4] and regional[e11,e12] WM LV and total[e2] and regional deep GM (DGM) volumes[e2,e4,e11,e12] (*p*-values ranging from <0.0001 to 0.05). By contrast, no associations with DGM volumes were found for global number of T2-lesions or the presence of Gd-enhancing lesions[e2]. Of the regional DGM volumes investigated, the most consistent relationships were found for the thalamus and hippocampus, where in one study, as much as 47% of the variance in thalamic volume was explained by T1 LV in thalamocortical WM ($p<0.001$)[e12]. This pattern was true considering both global[e2,e4] and regional WM LV[e11,e12]. Longitudinally, one of the available studies, in which all 210 patients were treated with first-line DMTs, found that change in thalamic volume over 48 months was related to global baseline lesion measures ($p\leq0.018$) and the total cumulative number of new/enlarging T2 lesions ($p=0.013$), but not to changes in LV[e7]. Similarly, another study considering regional WM LV, found no significant correlation between changes in ipsilateral LV and changes in hippocampal volume over a follow-up of 24 months[e11]. In this study, treatment status for the 36 included patients was not described.

Finally, atrophied T2 LV (described in the section above) was not associated with neither baseline thalamic volume nor volume change[e10].

RRMS

Overall, 37 cross-sectional and 14 longitudinal studies reported associations between WM lesion measures and GM atrophy in relapsing-remitting MS (RRMS) patients. Of these, 17 reported on the relation of WM lesions with global GM, 29 on that with cortical GM, and 25 on that with DGM measures. A total of 11 studies considered regional WM lesion measures. The follow-up period of the available longitudinal studies ranged from one to five and a half years.

Included studies are described in eTables 1, 2 and 3.

Global GM in RRMS

The great majority of available cross-sectional RRMS studies, i.e., eight out of ten studies, observed a significant association between global GM volume and global WM lesion load. Seven studies observed a significant correlation between global GM volumes and T1[e13,e14] and T2[e13-e18] LV, and abnormal WM[e20] (r-values ranging from -0.32 to -0.726, *p*-values ranging from <0.001 to 0.047). Furthermore, T2[e13,e17,e19] LV was as a significant predictor of GM volume in 3 studies (*p*-values ranging from <0.001 to 0.001). In contrast, two studies considering T2 LV[e21,e22], and another two studies considering Gd-enhancing LV[e13,e14], did not observe a significant correlation with global GM volumes. The four studies included a relative low number of patients (between 21 and 37 patients), no other apparent systematical differences were found between these studies, and studies observing significant associations between lesion measures and global GM volumes.

One cross-sectional study investigated the impact of regional LV on total GM volume, and reported a significant correlation with regional T1 and T2 LV in 3 and 4 out of 26 WM regions, respectively (r-values ranging from -0.49 to -0.50, $p<0.001$)[e23].

Of the seven longitudinal studies available, four did not find an association between global GM atrophy progression and global WM lesion measures. When considering Gd-enhancing lesion measures obtained at baseline, one study found a significant association (standardized β=-0.28, p=0.04) (follow-up time four years)[e24], while three others did not find that global GM atrophy progression related to either the presence[e25,e26], number[e26] or volume[e14] of Gd-enhancing lesions (follow-up time ranging from one to two years).

Three out of five studies with a follow-up period between one to four years, observed significant associations between longitudinal changes in T1[e27] and T2[e24,e27,e28] LV and GM atrophy progression (*p*-values ranging from 0.0004 to 0.03). Of note, in one of these studies, the association was only present in the patient group treated with fingolimod (r=-0.43, *p*=0.03), and not in the group treated with natalizumab[e28]. Lastly, two studies did not observe significant associations with global GM atrophy progression, neither for increasing T1[e14], T2[e14,e29] nor Gd-enhancing[e14] LV (follow-up time ranging from 12 to 24 months).

CGM in RRMS

A majority of cross-sectional studies (14 out of 19) considering global WM LV found significant associations. Negative correlations between T2 LV and total cortical volume were reported in five studies (r-values ranging from -0.245 to -0.48, *p*-values ranging from <0.0001 to 0.05)[e2,e15,e30-e32]. Furthermore, global cortical thickness was also found to associate with global WM LV (r-values ranging from -0.294 to -0.55, β-values ranging from -0.03 to -0.357, *p*-values ranging from <0.001 to <0.05)[e17,e30,e33-e35].

A total of six studies explored global T1[e36,e37] and T2[e16,e19,e30,e36-e38] lesions and their relationship with regional cortical volume, with the most consistent and strongest associations in areas in the frontal, temporal, cingulate and insular cortex (*p*-values ranging from <0.001 to <0.05). A similar pattern of associations was seen for cortical thickness measures, with cortical thinning in temporal, frontal, parietal and cingulate cortex (*p*-values ranging from <0.0001 to <0.05)[e30,e35,e39].

Although the majority of studies describing cortical measures found significant relationships with WM LV, five studies did not, either for cortical volume[e21,e40-e42] or cortical thickness[e39]. The four studies assessing cortical volumes[e21,e40-e42], included a relatively low number of patients (between 26 and 51), compared to the studies observing significant associations with LV.

In two of the eight cross-sectional studies considering regional distribution of WM lesions, the associations with cortical volume were weak or non-significant[e23,e43]. However, the other six publications demonstrated results suggestive of an anatomical or structural relationship between lesion location and regional cortical volume[e44,e45] and thickness[e9,e35,e46,e47].

Three out of five longitudinal studies considering global WM lesion measures and cortical atrophy found significant relationships between them. The only study analysing baseline WM lesion measures found that baseline T2 LV (annual additional volume loss 0.052% per cm^3 of T2 LV, $p<0.0001$) and Gd-enhancing lesion number (annual additional volume loss of 0.046% for each additional Gd-enhancing lesion, $p=0.0102$) were significant predictors of on-study total cortical volume loss[e49]. In the four studies assessing changes in global WM measures, two found an association between increasing lesion volumes[e27,e48] or numbers[e48], and regional volume loss ($p<0.01$)[e27] and cortical thinning ($p=0.040$)[e48] (follow-up time ranging from 1 to 2 years).

These results were in contrast to the third study considering global WM measures, in which the change in T2 LV did not associate with the change in cortical volume over the two-year follow-up[e32]. There were no apparent systematical differences in methodology or clinical characteristics between studies where associations were present and studies where associations were absent. Lastly, one study assessed atrophied T2 LV (described in previous section), and found no association with baseline cortical volume, or volume change[e10].

Two studies assessed the relationship between regional WM LV and cortical atrophy, both by visual inspection. One of them observed that the increase in T2 LV over the follow-up spatially coincided with areas of cortical decrease[e50], while the other study did not[e51].

DGM in RRMS

With the exception of one study[e52], all 17 cross-sectional publications reporting on global WM LV and DGM, found significant associations between the two. Three studies evaluated

DGM volume as a whole (β-values ranging from -0.258 to -0.49, *p*-values <0.0001 to 0.04)[e17,e49,e53], while the remaining evaluated the various structures separately. Thalamic volume correlated negatively with T1[e36,e37] and T2[e2,e30,e36-e38,e40,e41] LV in seven studies (r-values ranging from -0.407 to -0.81, *p*-values ranging from <0.05 to <0.00001), and an additional three studies reported significant associations in regression analyses (*p*-values ranging from <0.0001 to <0.05)[e16,e49,e55]. Lastly, one study found that surface displacement of the thalamus and other DGM structures associated with T2 LV (*p*=0.01)[e54]. Other DGM structures repeatedly showing significant associations with WM LV were the caudate nucleus (*p*-values ranging from <0.0001 to <0.05)[e2,e19,e36-e38,e41,e42,e55,e56], putamen (r-values ranging from -0.176 to -0.57, *p*-values ranging from <0.00001 to <0.05)[e2,e30,e36-e38,e53,e55], and globus pallidus (*p*-values ranging from <0.0001 to <0.05)[e2,e30,e38,e54,e55].

While two cross-sectional studies, one of which only conducted a qualitative evaluation by visual inspection[e44], did not find any associations between regional WM lesion and DGM measures[e23,e45], the majority of studies considering these structures did[e43,e44,e46,e54].

All four publications[e10,e48,e49,e57] that assessed longitudinal relations between total and regional DGM atrophy and global WM lesion measures, observed significant associations (*p*-values ranging from <0.0001 to 0.0372), and most consistently with thalamic atrophy.

Two studies observed that baseline T2 LV associated significantly with thalamic (r=-0.586, *p*=0.027[e57], β=-0.058, *p*<0.0001[e49]), and DGM (β=-0.053, *p*<0.0001)[e49] volume loss over 24 months. One of these studies also found baseline Gd-enhancing lesion number to be a significant predictor of DGM (β=-0.060, *p*=0.0007) and thalamic (β=-0.039, *p*=0.0372) atrophy[e49]. In the third available study, DGM atrophy rates were higher in patients with MRI activity (new/enlarging T2 lesions or new Gd-enhancing lesions) during the 24-month follow-up (mean atrophy rate of -0.4.51, *p*=0.024)[e48]. Finally, in the one study considering atrophied

T2 LV (described in previous section), a significant association was found for baseline thalamus volume (r=-0.384), and volume change (r=-0.430, both *p*-values=0.004)[e10].

SPMS

Eleven cross-sectional and two longitudinal studies reported on patients with secondary progressive MS (SPMS), five of which explored associations between WM lesions and global GM volume, while eight and ten studies focused on cortical and DGM measures, respectively. Four studies considered regional WM lesion measures.

Included studies are described in eTables 1 and 2.

Global GM in SPMS

Three out of four cross-sectional studies showed significant negative associations between WM LV and global GM volume (r-values ranging from -0.46 to -0.72, all *p*-values <0.001)[e16,e58]. In one paper, T2 LV did correlate with normalized GM volume (r=-0.36, *p*<0.01), but did not remain as a significant predictor in the final regression model[e17]. The one study considering regional T1 and T2 LV and global GM volume, found no significant associations[e23].

Longitudinally, neither baseline, nor on-study changes in WM lesion measures predicted changes in GM fraction over the subsequent four-year follow-up of 19 patients with SPMS[e24].

CGM in SPMS

The observed relationship between cortical volume or thickness and global WM lesion load in patients with SPMS was not consistent in the available four studies. Two studies found significant associations with lower cortical volume, mainly in frontal, temporal, cingulate and

cerebellar regions (*p*-values <0.001[e40] and <0.05[e16]). Furthermore, cortical thickness was evaluated in another two studies based on one and the same study population: neither study found any significant association between T2 LV and global mean cortical thickness[e17,e34].

In the four studies investigating regional WM lesions, results were also variable. Measured qualitatively by visual inspection, one study found that the distribution of T2 lesions were spatially close to regions with lower GM volume[e45], while another two studies found weak[e46] or non-significant[e23] associations. Lastly, one study observed relatively strong correlations between lower cortical volume in a given lobe, and T2 LV in the same or adjacent lobes (r-values ranging from -0.67 to -0.79, $p<0.001$)[e40].

Longitudinally, the only available study assessed atrophied T2 LV (described in previous section), and combined patients with SPMS and primary progressive MS (PPMS) into one progressive MS group (PMS) of 42 patients: no relation to either baseline cortical volume or volume changes were found (follow-up time five and a half years)[e10].

DGM in SPMS

In the six cross-sectional publications that considered global WM lesions, results were somewhat conflicting. Two studies found no associations with lower DGM volume[e16,e55], in the other four however, T1[e53] or T2[e17] LV associated significantly with both total DGM volume (r=-0.65[e17], β=-0.385[e17] and -1.11[e53], *p*-values ranging from <0.001 to 0.04), and separate DGM structures like the hippocampus[e52], thalamus and caudate nucleus[e40] (r-values ranging from -0.69 to -0.88, *p*-values ranging from<0.001 to 0.018). Two out of four studies that included results for regional WM LV or distribution, and the relationship with DGM volume, showed significant associations (r-values ranging from -0.64 to -0.87, $p<0.001$[e40], average standardized β=-0.264, p<0.05[e46]). Two studies did not find such associations[e23,e45].

There were no apparent systematic methodological or clinical differences in the discrepant studies, neither for those assessing global nor regional WM lesions.

As described in the above section, the only identified longitudinal publication studied atrophied T2 LV in PMS, and found no relation to baseline thalamus volume or volume change[e10].

PPMS

The relationship between WM lesions and various GM measures in patients with PPMS was assessed in 11 cross-sectional studies, and longitudinally in two studies. Three studies performed analyses involving global GM volume, while cortical and DGM measures were each considered nine studies. Three studies considered regional WM LV or distribution.

Included studies are described in eTables 1 and 2.

Global GM in PPMS

The cross-sectional associations between WM lesion load and global GM volume in patients with PPMS were variable. One study reported a significant correlation (r= -0.68, $p<0.001$) with T2 LV, but in a multiple regression model the association with global GM volume was no longer significant[e17]. In the other available study, no significant associations were found between GM volume and either T2, T1 or gadolinium-enhancing lesion volumes or numbers[e59].

In the available longitudinal study, no association was found between baseline WM lesion measures, and GM volume change over 12 months[e17].

CGM in PPMS

Cross-sectional results for global WM lesion measures and cortical GM volume or thickness in patients with PPMS were divided. Three studies found associations between T1[e61] and T2[e34,e40] LV and total (r=-0.508, $p<0.05$)[e61] and regional (r-values ranging from -0.605 to -0.85, *p*-values ranging from <0.001 to <0.01)[e40] cortical volume and total cortical thickness (standardized β=-0.425, $p<0.05$)[e34]. In the three studies with non-significant results, no associations were found for either cortical volume[e31,e62] or thickness[e17].

Out of three publications assessing regional WM lesion measures, one found a significant association between cortical volume and regional WM lesion measures in anatomically connected areas (r-values ranging from -0.83 to -0.91, $p<0.001$)[e40], while in the other two the associations with cortical thickness or volume were weak[e46] or absent[e45]. Again, there were no apparent systematic methodological or clinical differences in the studies observing significant and non-significant observations, and in all studies the number of included patients were relatively low (between 18 and 31).

Only one longitudinal study was identified (described in previous section), finding no associations between atrophied T2 LV and baseline cortical volume, or volume change[e10].

DGM in PPMS

All but one[e55] of the six cross-sectional studies reporting on the relationship between global WM lesions and DGM volume observed significant associations. In PPMS patients, correlations were significant for both DGM volume as a whole (r-values ranging from -0.651 to -0.71, *p*-values ranging from <0.001 to <0.01)[e17,e61], and for the separate structures. The most consistent association with global WM LV was seen for the thalamus[e62], for both T2 (r-values ranging from -0.48 to -0.94, *p*-values ranging from <0.001 to <0.05)[e40,e61,e63] and T1 LV (r-values ranging from -0.44 to -0.554, *p*-values ranging from 0.002 to <0.05)[e61,e63]. Of note, in the only study with absent associations[e55], 25 patients with both SPMS and PPMS were pooled in a combined PMS group.

Three cross-sectional publications analysed regional WM lesions and the relationship with DGM volume. Lower thalamus volume were related to T2 LV in all lobes analysed in one study (r-values ranging from -0.85 to -0.93, $p<0.001$)[e40], and another study found that regional T2 LV was a significant factor in multiple regression models for the nucleus accumbens, hippocampus and globus pallidus, but not for amygdala, putamen or thalamus (average standardized $\beta=-0.438$, $p<0.05$)[e46]. The third study conducted a qualitative analysis by visual inspection, and found no spatial correspondence between T2 lesions and lower DGM volume[e45].

Lastly, only one longitudinal study was available (described in previous section), and for both baseline thalamic volume and volume change, no relation to atrophied T2 LV was found[e10].

Results for mixed MS groups

In a mixed MS patient group, i.e., comprising different disease types but studying the entire patient group as a whole, 32 publications assessed the cross-sectional, and 9 the longitudinal association between WM lesion measures and GM volume. In this section, studies including patients with CIS in the total patient population are also considered. The follow-up time in the longitudinal studies ranged from two to five and a half years.

Global GM volume and its relation to WM lesions was described in 13 studies, CGM volume or thickness was considered in 17, and DGM volume in 23 studies. Associations with regional WM lesions were reported in eight studies in total.

Included studies are described in eTables 1, 2 and 3.

Global GM in mixed MS groups

In all but one of the nine available cross-sectional studies, global WM lesion measures were consistently associated with global GM volume in patients with MS. Significant correlations were found in seven studies (r-values ranging from -0.43 to -0.63, *p*-values ranging from <0.001 to 0.016)[e16,e17,e65-e69], and two studies observed WM LV as a significant predictor of global GM volume (r=-0.52[e17], β=-0.226[e17] and -0.27[e1], *p*-values ranging from <0.01 to <0.001). In all of the above studies, patients with RRMS made up the majority of the mixed MS groups. Patients with or without Gd-enhancing lesions however, did not exhibit different GM volumes[e64].

One cross-sectional study reported results on regional WM LV, and observed significant correlations between total GM fraction and regional T2 and T1 LV in 9 and 5 out of 26 regions, respectively (r-values ranging from -0.24 to -0.45, $p<0.001$)[e23].

The four available longitudinal studies considering global GM atrophy and global WM lesion measures load reported mainly absent associations. In the three studies assessing baseline WM lesion measures, changes in GM volume were not associated with the presence of Gd-enhancing lesions[e64], T1 LV[e88] or abnormal WM fraction[e89] (follow-up period ranging from two to five and a half years). However, in one of these studies, baseline global GM volume correlated significantly with variation in abnormal WM fraction over the 2-year follow-up (r=-0.180, $p<0.001$)[e89]. Lastly, in the one study assessing T2 lesion shrinking over a three-year follow-up, this did not relate to changes in global GM volume[e87].

CGM in mixed MS groups

The majority of studies, i.e., ten out of twelve studies considering cortical volume or thickness and global WM lesion measures in MS, found significant associations between the two. T2 LV was associated with lower GM volume in the cerebellar cortex, temporal lobe and the pre- and postcentral gyrus in the frontal and parietal lobe in two studies (both *p*-values <0.05)[e16,e73]. One study using estimated regional cortical, DGM and brainstem volumes in an

event-based model to determine the sequential occurrence of atrophy, observed that the event-based model stage at baseline was related to the T2 LV ($p<0.001$)[e71]. Five studies found a significant association between global WM LV and both global (*p*-values ranging from <0.001 to <0.05)[e17,e34,e70,e75] and regional cortical thickness (*p*-values ranging from 0.007 to <0.05)[e74,e75,e77]. Lastly, one study found that global T2 LV correlated significantly with cortical surface area ($r= -0.62$, $p<0.05$)[e72].

In two publications, one of which all included patients were treated with natalizumab, cortical thickness and/ or volume did not associate with either global T2 LV[e76] or the presence of Gd-enhancing lesions[e64].

Associations between regional WM lesion measures and cortical atrophy were reported in five available studies. One study found that regional T2 LV in the left and right frontal, temporal, parietal, and occipital areas were associated with lower cortical thickness in widespread bilateral cortical regions ($p<0.05$)[e77], and another study observed that LV in connected WM tracts was a significant explanatory variable of cortical thickness (average standardized $\beta=-0.116$, $p<0.05$) in 28 out of 34 areas[e46]. In the three remaining publications, associations between regional WM lesion measures and cortical volume or thickness were weak ($r=-0.267$, $p\leq0.01$)[e23] or non-significant[e9,e78].

Four publications explored the longitudinal relationship between global WM lesion measures and cortical atrophy, and none of them observed significant associations. No relations were found between the presence of Gd-enhancing lesions at baseline[e64] or changes in T2 LV[e77], and changes in cortical volume[e64] or thickness (either global or regional)[e64,e77] (follow-up time ranging from 3 to 4 years). In one study using an event-based model (described in the above section); the rate of increase in T2 LV over the mean follow-up of 2.4 years did not associate with the rate of change in the event-based model stage[e71]. Assessing atrophied T2 LV

(described in previous section), associations with baseline cortical volume or volume change were not found[e10].

One of the four longitudinal publications also considered regional T2 LV, and found that changes in left temporal and occipital T2 LV associated negatively with cortical thinning in the left temporal, parietal and occipital areas (p<0.05)[e77].

DGM in mixed MS groups

Seventeen studies considered the cross-sectional relationship between global WM lesion and DGM measures in patients with MS; as many as 14 observed significant associations. Global T1[e53] and T2[e17] LV was significantly associated with total DGM volume (r=-0.70[e17], standardized β= -0.407[e17] and -0.41[e53], p-values ranging from <0.001 to 0.02) as well as with the separate structures. A total of ten publications reported significant associations between global WM LV and thalamic (r-values ranging from -0.36 to -0.77, p-values ranging from <0.0001 to <0.05)[e16,e53,e55,e76,e80,e81,e83-e85] and pulvinar nucleus (p<0.05) volume[e73]. Associations were also seen for caudate nucleus (p-values ranging from <0.01 to <0.05)[e16,e55,e73,e76], putamen (p-values ranging from 0.003 to <0.05)[e53,e55,e69,e73] and hippocampal volume (r-values ranging from -0.46 to -0.56, all p-values=0.008)[e82]. Furthermore, global T2 LV was associated with a more advanced event-based model stage, as referred to in the above section regarding cortical GM atrophy in MS patients (p<0.001)[e71]. As opposed to the results listed above, two studies did not observe significant associations between WM LV and lower GM volume either in the caudate nucleus[e79] or in any other DGM structure[e74]. Furthermore, the presence of Gd-enhancing lesions did not correlate with total or regional DGM volume[e64].

Of the available studies analysing the cross-sectional relationship between regional WM LV and DGM volume, three found significant (*p*-values ranging from 0.02 to <0.05)[e46,e73,e78], or near significant associations[e23,e86].

Only one out of three studies[e10,e64,e71] reporting on the longitudinal relationship between DGM atrophy and global WM lesion measures in MS patients found a significant association: in this study, atrophied T2 LV correlated significantly with baseline thalamic volume (r=-0.620, p=0.003) and thalamic volume change (r=-0.672, p=0.003)[e10]. The presence of Gd-enhancing lesions at baseline did not associate with changes in total or regional DGM volume over a 3-year follow up, in patients treated with natalizumab[e64]. Furthermore, the rate of change in the event-based model stage was not related to the rate of increase in T2 LV[e71].

Two longitudinal studies considering regional WM lesion measures and DGM atrophy were available, and both found significant associations. One study found that over a 12-month period, new, chronic enlarging and chronic shrinking T1 lesion number along the optic radiation was related to reduced ipsilateral lateral geniculate nucleus volume (*p*-values ranging from 0.0001 to 0.0056)[e90]. The second study assessed regional DGM atrophy over five years, and its association to percentage lesion disruption, i.e., the percentage of normative data-base derived connected tract streamlines that passes through a given lesion mask and are considered disrupted. When controlling for T2 LV, tract disruption could predict DGM volume atrophy rate in 5 of 14 regions, but only 1 after correction for multiple comparisons (p<0.001)[e86].

Comparisons between disease phenotypes

Four studies described global GM atrophy, while eleven studies reported results on CGM and DGM atrophy, each. Five studies considered regional WM lesion measures and their possible association with GM atrophy.

The included studies are described in eTables 1 and 2.

Global GM in comparisons between phenotypes

In both available cross-sectional studies, a negative association between global GM volume and global WM LV was observed in all investigated disease types (RRMS (r=-0.334, $p<0.001$[e16], r=-0.57, $p<0.01$[e17]), SPMS (r=-0.460, $p<0.001$[e16], r=-0.36, $p<0.01$[e17]), PPMS (r=-0.68, $p<0.01$[e17])), although one of these found that in multivariable logistic regression, the association with global WM lesion volume remained only for RRMS (standardized β=-0.231, $p<0.01$), but not SPMS or PPMS[e17].

In the single study assessing regional WM LV, significant associations were only found in RRMS (r-values ranging from -0.20 to -0.50, $p\leq0.001$), but not in SPMS[e23].

Longitudinally, the presence of Gd-enhancing lesions at baseline (standardized β=-0.28, $p=0.04$) and increase in T2 LV during the four-year follow-up (standardized β=-0.46, $p=0.0004$) predicted faster GM fraction change in RRMS, but not in SPMS[e24].

In all of the above studies, the RRMS group was considerably larger than the groups consisting of patients with SPMS or PPMS.

CGM in comparisons between phenotypes

The six available cross-sectional studies found more consistent associations between global WM lesion load and cortical atrophy in RRMS than the other phenotypes, though the association was not uniformly present in all phenotypes[e2,e16,e17,e31,e34,e40]. Four studies found that global T2 LV was associated with total cortical volume or thickness in RRMS (r-values ranging from -0.420 to -0.55[e2,e17,e31], standardized β=-0.319[e34] and -0.357[e17], *p*-values ranging from <0.0001 to <0.05), while in CIS it was absent[e2], in SPMS absent[e17,e34], and in PPMS

present (r=-0.43, p<0.05[e17], standardized β=-0.425, p<0.05[e34]) or absent[e17, e31] (in one of the studies[e17] depending on the statistical analysis used); of note, two of these studies investigated the same patient group[e17, e34]. Except the one study including patients with CIS and RRMS[e2], the patient groups were again unbalanced with large RRMS groups compared to groups with progressive disease types. Regional cortical volumes associated with global T2 LV in both RRMS and SPMS in one study (all p-values <0.05)[e16], but only in SPMS (r-values ranging from -0.63 to -0.75), and PPMS (r-values ranging from -0.83 to -0.85), not RRMS, in another study (all p-values<0.001)[e40].

In the five cross-sectional studies considering regional WM LV, two assessed cortical thickness, finding the most consistent and widespread associations in RRMS patients (p-values ranging from 0.0002 to <0.05)[e9,e46]. In CIS[e9], SPMS[e46] and PPMS[e46], the association was significant only in certain cortical regions. In the three studies assessing cortical volume, associations were absent in RRMS and SPMS[e23], present in RRMS and SPMS (qualitative association by visual inspection)[e45] and present in SPMS (r-values ranging from -0.65 to -0.79, p<0.001) and PPMS (r-values ranging from -0.83 to -0.91, p<0.001)[e40].

DGM in comparisons between phenotypes

All three cross-sectional studies assessing the relationship between global WM lesion load and total DGM volume, observed a significant association in RRMS (r=-0.613[e2] and -0.74[e17], standardized β=-0.407[e17] and -0.49[e53], p-values ranging from <0.0001 to 0.04) while finding it to be present (r=-0.329, p<0.0001)[e2] in CIS, absent[e53] or present (r=-0.65[e17], standardized β=-0.385[e17] and -1.11[e53], p-values ranging from <0.001 to 0.04) in SPMS (in one of the studies[e53] depending on the lesion measure), and absent[e17] or present (r=-0.71, p<0.001)[e17] in PPMS (depending on the statistical analysis used[e17]). The two papers considering patients with RRMS and progressive disease types[e17,e53] had unbalanced patient groups with small groups

of SPMS (53[e17] and 12[e53] patients) and PPMS (25 patients[e17]) patients. In the six studies considering regional DGM volumes, global WM lesion measures were, with the exception of one study[e52], consistently related in RRMS (*p*-values ranging from <0.0001 to 0.04)[e2,e16,e40,e53,e55]. Furthermore, associations were present in CIS (r-values ranging from -0.218 to -0.328, *p*-values ranging from <0.0001 to <0.001)[e2], absent[e16,e55] or present (r-values ranging from -0.69 to -0.88, *p*-values ranging from <0.001 to 0.018)[e40,e52] in SPMS, and absent[e55] or present (r=-0.94, $p<0.001$)[e40] in PPMS.

In the four cross-sectional studies assessing the relationship between regional WM lesion measures and regional DGM volume, significant associations were absent in RRMS and SPMS[e23], present in RRMS and SPMS, but absent in PPMS (qualitative association by visual inspection)[e45], present in RRMS, SPMS and PPMS (standardized β-values ranging from -0.264 to -0.438, all *p*-values <0.05)[e46], and present in SPMS and PPMS (r-values ranging from -0.64 to -0.87, all *p*-values<0.001)[e40].

Supplemental eTable 1. Characteristics of cross-sectional studies

Cross-sectional studies					
Reference (e-ref)	*Patients (n)*	*GM atrophy measure*	*WM lesion measure*	*Results*	*Quality of evidence*
CIS					
Durhan et al., Int J Neurosci, 2016 (e3)	33	Total GM, global volume. Cortical GM, global volume. Cortical GM, regional volume. Deep GM, regional volume.	Global T2 lesion volume.	Significantly higher volume in the left cerebellar cortex, left accumbens area and right GP in smoking CIS patients than in non-smoking CIS patients, but no significant correlation between T2 LV and GMV in these areas.	Fair
Henry et al., J Neurol Neurosurg Psychiatry, 2008 (e4)	41	Total GM, global volume. Deep GM, regional volume.	Global T1 lesion volume.	T1 LV was negatively associated with bilateral thalami (R^2=0.50, p=0.020 (left) and p=0.037 (right)) and left hippocampus (R^2=0.38, p=0.037) volume[a]. No significant association between T1 LV and global GMV.	Good
Henry et al., J Neurol Sci, 2009 (e12)	24	Deep GM, regional volume (thalamus).	Regional T1 lesion volume.	In a stepwise regression model, 47% of the variance in thalamic volume was explained by T1 LV in thalamocortical WM (p<0.001)[a].	Good
Jenkins et al., J Neurol Neurosurg Psychiatry, 2011 (e8)	28	Cortical GM, regional volume and thickness (pericalcarine cortex).	Global T2 lesion volume. Regional T2 lesion volume (optic radiation). Global presence of T2 lesions Regional presence of T2 lesions (optic radiation). Regional T2 lesion length (optic nerve).	Lower pericalcarine cortical volumes were significantly associated with the presence of whole brain T2 lesions (t=2.48, p=0.020) and OR T2 lesions (t=2.24, p=0.034), in patients with optic neuritis[a]. No significant association between pericalcarine volume or thickness and optic nerve lesion length, OR T2 LV or whole brain T2 LV.	Fair
Labiano-Fontcuberta et al., Medicine, 2016 (e5)	17	Total GM, global volume. Cortical GM, global volume.	Global T2 lesion volume.	Significant negative correlation between T2 LV and GMV (r=-0.56, p<0.020), and CV (r=-0.49, p<0.045)[a].	Fair
RRMS					
Al-Radaideh et al., Clin Neuroradiol, 2019 (e30)	30	Cortical GM, global volume, thickness and surface area. Cortical GM, regional volume, thickness and surface area. Deep GM, regional volume.	Global T2 lesion volume.	Significant correlation between T2 LV and cortical surface area (r=-0.208, p<0.05), Cth (r=-0.294, p<0.05), and cortical (r=-0.245, p<0.05) putamen (r=-0.408, p<0.05), GP (r=-0.410, p<0.05) and thalamus (r=-0.407, p<0.05) volumes[a].	Fair

Supplemental eTable 1. Characteristics of cross-sectional studies

Bergsland et al., Mult Scler, 2015 (e47)	51	Cortical GM, regional thickness and surface area (primary motor cortex).	Regional T2 lesion volume (corticospinal tract).	Significant correlation between CST T2 LV and PMC thickness (r=-0.327, p=0.022)[a], but not with PMC surface area (r=-0.039, p=0.791). Significant correlation between PMC thickness and lesion probability along large parts of the CST (p-values ranging from 0.004 to 0.048)[a]. CST T2 LV not a significant predictor of PMC thickness in final multiple regression model.	
Calabrese et al., Neurology, 2010 (e33)	100	Cortical GM, global thickness.	Global T2 lesion volume.	Significant correlation between T2 LV and mean Cth (r=-0.31, p=0.004)[a].	Good
Ceccarelli et al., J Neurol, 2007 (e21)	32	Total GM, global volume. Cortical GM, global volume.	Global T2 lesion volume.	No significant correlation between total T2 LV and any GM volume measure.	Fair
Chard et al., Brain, 2002 (e13)	26	Total GM, global fraction (to ICV).	Global T1, T2 and Gd-enhancing lesion volume.	Significant correlation between GMF and T1 LV (r=-0.525, p=0.006) and T2 LV (r=-0.726, p<0.001), but not with Gd-enhancing LV (r=0.109, p=0.595)[a]. In a general linear model, GMF decreased by -0.003 for every ml T2 LV (SE=0.001, p<0.001).	Fair
Charil et al., Neuroimage, 2007 (e35)	425	Cortical GM, global thickness. Cortical GM, regional thickness.	Global WM lesion volume. Regional WM lesion volume.	Significant association between global WM LV and mean Cth (t=-15.3, p<0.0001), with an average loss of 0.03 mm in Cth per cm^3 of WM LV, accounting for 38% of the variance[a]. Regionally, global WM LV was associated with Cth in almost all vertices in the surface model, e.g., bilateral cingulate gyrus, prefrontal, temporal and parietal association cortical regions (average loss 0.03 to 0.07 mm per cm^3, t-values ranging from -11.7 to -16.2, all p<0.05)[a]. An average lesion density map showed that anterior cortical thinning correlated more with anterior lesions than posterior cortical thinning correlated with posterior lesions[a].	Fair
Datta et al., Mult Scler Relat Disord, 2015	924	Cortical GM, regional volume. Deep GM, regional volume.	Global T1, T2 and Gd-enhancing lesion volume.	Significant negative correlation between T2 and T1 LV and thalamus (T2:r=-0.492, T1:r=-0.473), putamen (T2:r=-0.193, T1:r=-0.176), CN (T2:r=-0.188, T1:r=-	Fair

Supplemental eTable 1. Characteristics of cross-sectional studies

(e36)				0.173), inferior frontal gyrus (T2:*r*=-0.120, T1:*r*=-0.101), septal nuclei (T2:*r*=-0.146, T1:*r*=-0.153), and amygdala (T2:*r*=-0.109) volumes[a]. T2 and T1 LV correlated positively with hippocampus volume (T2:*r*=+0.117, T1:*r*=+0.169)[a]. All $p<0.05$.	
Dolezal et al., Int Rev Neurobiol, 2007 (e15)	41	Total GM, global volume. Cortical GM, global volume.	Global T2 lesion volume.	Significant correlation between global T2 LV and GMV (*r*=0.4, *p*=0.003) and CV (*r*=0.48, *p*=0.001)[a]. Higher total T2 LV was associated with lower CV (R^2=0.17, *p*=0.007)[a].	Fair
Duan et al., Eur J Radiol, 2012 (e41)	26	Cortical GM, regional volume. Deep GM, regional volume.	Global T2 lesion volume.	Significant correlation between T2 LV and the right CN (*r*=−0.409, *p*=0.019), left thalamus (*r*=−0.596, *p*=0.001) and right thalamus (*r*=−0.694, *p*=0.000) volume[a].	Fair
Hasan et al., J Magn Reson Imaging, 2009 (e56)	32	Deep GM, regional volume percentage (of ICV) (caudate nucleus).	Global T2 lesion volume percentage (of ICV).	Significant correlation between global T2 LV percentage and CN volume percentage (*r*=-0.482, *p*=0.005)[a].	Fair
Hasan et al., J Neurol Sci, 2012 (e38)	54	Cortical GM, regional volume percentage (of ICV). Deep GM, regional volume percentage (of ICV).	Global T2 lesion volume.	Significant age-adjusted correlation between global T2 LV and corpus striatum (*r*=-0.519, $p<0.0001$), thalamus (*r*=-0.569, $p<0.00001$), CN (*r*=-0.46, *p*=0.001), putamen (*r*=-0.494, $p<0.00001$), GP (*r*=-0.395, *p*=0.003), accumbens area (*r*=-0.407, *p*=0.002) and temporal cortex (*r*=-0.315, *p*=0.022) volume percentage[a].	Fair
Kuceyeski et al., AJNR, 2015 (e43)	121	Cortical GM, regional volume. Deep GM, regional volume.	Change in Connectivity (ChaCo) score (the percentage of tracts connecting to a GM region that pass through the WM abnormality mask).	Out of 86 GM regions, no significant correlation was found between regional ChaCo score and lower GMV. Significant correlation between ChaCo score and bilateral thalami (*r*=0.30), putamen (*r*=0.38), CN (*r*=0.26), GP (*r*=0.26) and nucleus accumbens (*r*=0.28) volume[a]. No significant correlation between regional lower GMV and ChaCo score in temporal poles.	Fair
Magon et al., HBM, 2014 (e54)	118	Deep GM, regional shape.	Global T2 lesion volume. Regional T2 lesion volume.	Global T2 LV significantly related to surface displacement of the thalamus (F(3, 114)=2.31, *t*=3.12), striatum (F(3, 114)=1.14, *t*=2.79) and GP (F(3, 114)=0.16, *t*=3.21)[a]. Observed displacements in thalamus (F(3,114)=5.44,	Fair

Supplemental eTable 1. Characteristics of cross-sectional studies

				t=3.59), GP (F(3,114)=5.16, t=3.41) and striatum (F(3,114)=5.44, t=3.59) were mainly driven by frontal and parietal T2 LV (p=0.01)[a], while no surface displacements were associated with occipital T2 LV.	
Narayana et al., Neuroimage Clin, 2012 (e39)	250	Cortical GM, global thickness. Cortical GM, regional thickness.	Global T1 and T2 lesion volume.	No significant correlations between mean global Cth and T1 and T2 LV. Weak to moderate correlations between T1 and T2 LV and regional cortical areas in both 1.5 and 3T, with the strongest correlations seen for the bilateral superior frontal lobe cortex and T1 (r=-0.35 (right), r=-0,33 (left), p<0.0001) and T2 (r=-0.3 (right), r=-0.29 (left), p<0.0001) LV[a].	Fair
Prinster et al., Mult Scler, 2010 (e19)	128	Total GM, global volume. Cortical GM, regional volume. Deep GM, regional volume.	Global T2 lesion volume.	Significant linear association between T2 LV and lower global GMV (p=0.001)[a]. T2 LV significantly negatively associated with GMV in the caudate heads (t=7.20), parahippocampal gyri (t= 5.95), cingulate gyri (t=5.86), motor cortex (t=5.0), and insula (all p<0.05)[a].	Fair
Prinster et al., Neuroimage, 2006 (e42)	51	Cortical GM, regional volume. Deep GM, regional volume.	Global abnormal WM (aWM).	Significant association between aWM and lower right CN volume (p<0.05)[a]. No significant association between aWM and lower regional CV.	Fair
Quarantelli et al., Neuroimage, 2003 (e20)	50	Total GM, global fraction (to ICV).	Fractional abnormal WM (aWM/ICV = aWMf).	Lower global GMF associated significantly with aWMf (β=-0.62, R=-0.434, p<0.001)[a], with corresponding increase in CSF fraction.	Fair
Riccitelli et al., Mult Scler, 2012 (e44)	78	Cortical GM, regional volume. Deep GM, regional volume.	Regional T2 lesion distribution.	Significant correlation between T2 lesion distribution and lower GMV in the head of the right (r=-0.58) and left (r=-0.53) CN, right (r=-0.39) and left (r=-0.43) insula, left cingulate gyrus (r=-0.48), right superior frontal gyrus (r=-0.47) and left middle occipital gyrus (r=-0.54)[a].	Fair
Sbardella et al., PLoS One, 2013 (e22)	36	Total GM, global fraction (to ICV).	Global T2 lesion volume.	No significant correlation between T2 LV and GMF.	Fair
Tao et al., J Neurol Sci, 2009	88	Cortical GM, regional volume. Deep GM, regional volume.	Global T1 and T2 lesion volume.	Significant correlation between T2 LV and lower volume of the thalamus (r=-0.56), CN (r=-0.31), putamen (r=-	Fair

Supplemental eTable 1. Characteristics of cross-sectional studies

(e37)				0.5), amygdala (r=-0.45), cingulate gyrus (r=-0.47), middle temporal gyrus (r=-0.33), inferior parietal lobule (r=-0.23), entorhinal area (r=-0.34) and insular cortex (r=-0.25)[a]. Significant correlation between T1 LV and lower volume of the thalamus (r=-0.61), CN (r=-0.35), putamen (r=-0.43), pons (r=-0.21), amygdala (r=-0.28), cingulate gyrus (r=-0.43), middle temporal gyrus (r=-0.34), inferior parietal lobule (r=-0.30) and entorhinal area (r=-0.23)[a].	
Toth et al., Front Neuroanat, 2017 (e18)	52	Total GM, global volume.	Global T2 lesion volume.	Significant negative association between T2 LV and GMV (R=-0.32, p<0.021)[a].	Fair
SPMS					
Furby et al., Mult Scler, 2009 (e58)	117	Total GM, global fraction (to ICV).	Global T1 and T2 lesion volume.	Significant negative correlations between GMF and T2 (r=-0.70, p<0.001) and T1 (r=-0.72, p<0.001) LV[a]. Significant associations between GMF and T2 (r_p=-0.69, p<0.001) and T1 (r_p=-0.72, p<0.001) LV in a multiple regression model[a]. In a stepwise multiple regression analysis, T1 LV was the only significant lesion correlate of GMF (standardized β: -0.72, r_p=0.72, R^2=0.52, p<0.001)[a].	Fair
PPMS					
Galego et al., Neuroradiol J, 2015 (e61)	19	Cortical GM, global volume. Cortical GM, regional volume. Deep GM, global volume. Deep GM, regional volume.	Global T1 and T2 lesion volume.	Significant negative correlations between T1 LV and CV (r=-0.508, p<0.05), DGMV (r=-0.651, p<0.01), thalamus (r=-0.554, p<0.05), putamen (r=-0.699, p<0.01), GP (r=-0.495, p<0.05) and pre-central gyrus (r=-0.605, p<0.01) volume[a]. Significant negative correlations between T2 LV and DGMV (r=-0.630, p<0.01), thalamus (r=-0.567, p<0.05) and putamen (r=-0.696, p<0.01) volume[a].	Poor
Mesaros et al., AJNR, 2011	54	Deep GM, regional volume (thalamus).	Global T1 and T2 lesion volume.	At baseline, significant negative correlations were found between thalamus volume and T2 LV (r=-0.48, p=0.001)	Fair

Supplemental eTable 1. Characteristics of cross-sectional studies

(e63)				and T1 LV (r=-0.44, p=0.002)[a].	
Sastre-Garriga et al., Neuroimage, 2004 (e59)	43	Total GM, global fraction (to ICV).	Global T1, T2 and Gd-enhancing lesion volume. Global Gd-enhancing lesion number.	No significant correlation between GMF and T2 LV, Gd-enhancing LV or Gd-enhancing lesion number. No significant association between GMF and T2 or T1 LV in regression analyses.	Fair
Sepulcre et al., Arch Neurol, 2006 (e62)	31	Cortical GM, regional volume. Deep GM, regional volume.	Global T2 lesion volume.	Significant associations were found between higher T2 LV and lower local GMV in the anterior nuclei of both thalami (number of voxels, 893; z=4.95, p=0.003)[a], but not with any other cortical or DGM regions.	Fair
MS					
Bermel et al., Neuroreport, 2003 (e79)	24 RRMS (16) SPMS (8)	Deep GM, regional volume (caudate nucleus).	Global T1 and T2 lesion volume.	No significant correlation between CN volume and T1 (p=0.32) or T2 (p=0.23) LV.	Fair
Calabrese et al., J Neurol, 2007 (e70)	83 CIS (10) RRMS (42) SPMS (31)	Cortical GM, global thickness.	Global T2 lesion volume.	Significant correlation between Cth and T2 LV (r=-0.393, p=0.03)[a].	Fair
Deppe et al., Hum Brain Mapp, 2016 (e80)	122 CIS (12) RRMS (110)	Deep GM, regional volume percentage (of ICV) (thalamus).	Global T2 lesion volume.	There was a logarithmic dependency between RTV and T2 LV in the patient groups combined (RTV[%]=1.0028-0.1693 $\log_{10}$T2LV[mL][a], i.e., the higher the T2 LV, the lower the RTV, but the relative impact on thalamic volume is smaller with increasing LV. In a univariate general linear model, T2 LV had the second highest effect (after thalamic FA) on RTV (SS=0.0560, ms=0.0560. F=12.95, p<0.001)[a].	Fair
Fisniku et al., Ann Neurol, 2008 (e65)	73 CIS (29) RRMS (33) SPMS (11)	Total GM, global volume. Total GM, global GM fraction (to ICV).	Global T2 lesion volume.	Significant correlation between GMF and T2 LV for the whole cohort (including patients with CIS) of patients (r=-0.63, p<0.001), and for the separate MS (i.e., RRMS and SPMS) subgroup (r=-0.66, p<0.001)[a]. Significant correlation between GMV and T2 LV for the whole cohort (r=-0.57, p<0.001) and for the separate MS (i.e., RRMS and SPMS) subgroup (r=-0.67, p<0.001)[a].	Fair
Fragoso et al., J	185	Total GM, global volume.	Global T2 lesion volume.	Significant negative correlation between T2 LV and	Fair

Supplemental eTable 1. Characteristics of cross-sectional studies

Clin Neuroscience, 2017 (e66)	CIS (1) RRMS (173) PPMS (11)			GMV (p<0.001)[a].	
Hasan et al., J Neurosci, 2011 (e81)	109 CIS (9) RRMS (88) SPMS (12)	Deep GM, regional volume percentage (of ICV) (thalamus).	Global T2 lesion volume.	Significant correlation between T2 LV and percentage of total thalamus volume (r=-0.658, p<0.001)[a].	Fair
Hier et al., Neurological Research, 2007 (e72)	15 RRMS (6) SPMS (9)	Cortical GM, global surface area.	Global T2 lesion volume.	Negative correlation between T2 LV and cortical surface area (r=-0.62, p<0.05)[a].	Poor
Longoni et al., Brain Struct Funct, 2015 (e82)	103 RRMS (22) SPMS (33) PPMS (23) Benign MS (25)	Deep GM, regional volume (hippocampus).	Global T1 and T2 lesion volume.	Significant correlation between T1 LV and right (r=-0.53, p=0.008) and left (r=-0.46, p=0.008) hippocampus volume[a]. Significant correlation between T2 LV and right (r=-0.56, p=0.008) and left (r=-0.51, p=0.008) hippocampus volume[a]. Significant negative correlation between T1 and T2 LV and local volume of the lateral CA1 subfield, part of the subiculum, and the CA1 region of the hippocampal head, bilaterally (r-values ranging from -0.2 to -0.5; p<0.001; FDR rate at p<0.05 between 4 and 7%)[a].	Fair
Louapre et al., Mult Scler, 2018 (e83)	41 Early RRMS (10) RRMS (18) SPMS (13)	Deep GM, regional volume (thalamus).	Global T2* lesion volume.	Significant partial correlation between thalamic volume and T2* LV (r_p=-0.77, R^2_p=0.59 p=4×10^{-9})[a]. In a multiple regression model, lower thalamic volume associated significantly with higher T2* LV, independently of Cth, cortical LV, ICV, age, sex and disease duration. T2* LV was one of the explanatory variables retained in the model (adjusted R^2=0.68, p=0.8×10^{-5})[a].	Fair
Mehndiratta et al., Mult Scler, 2020 (e84)	90 RRMS (61) SPMS (29)	Deep GM, regional volume (thalamus).	Global T2* lesion volume.	Thalamic volume inversely correlated with T2* LV (r=-0.6, p<0.001)[a]. In multiple regression analysis, lower thalamic volume associated with T2* LV (β=-2.8×10^{-4}, 95% CI:-4.6×10^{-4},	Fair

Supplemental eTable 1. Characteristics of cross-sectional studies

				-1×10^{-4}, R^2=0.24, p=0.002), and not with thalamic or cortical LV.	
Mühlau et al., Mult Scler, 2013 (e73)	249 CIS (81) RRMS (168)	Cortical GM, regional volume. Deep GM, regional volume.	Global T2 lesion volume. Regional T2 lesion volume.	Global T2 LV associated significantly with lower regional GMV in the visual, primary auditory, motor, and somatosensory cortex, cerebellum, pulvinar nucleus, putamen and CN (p<0.05)[a]. In a linear multiple regression model, lower CN volume was associated with T2 lesion probability primarily in frontal WM. Lower pulvinar volume was associated with T2 lesion probability in a more widespread area, mainly in parietal and occipital WM (p<0.05)[a].	Good
Pareto et al., AJNR, 2015 (e74)	131 CIS (91) RRMS (40)	Cortical GM, regional thickness. Deep GM, regional volume.	Global T2 lesion volume.	On multivariate analyses, controlling for the presence and volume of juxtacortical lesions, significant associations were found between T2 LV and Cth in the left hemisphere (R^2=0.646, p=0.039), in particular the left parietal (R^2=0.649, p=0.035) and left occipital (R^2=0.639, p=0.038) lobules. For the right hemisphere, this was found for the right frontal (R^2=0.639, p=0.043), right parietal (R^2=0.696, p=0.020) and right occipital (R^2=0.753, p=0.007) lobules[a]. T2 LV did not reach significance in any DGM region.	Fair
Rocca et al., Radiology, 2010 (e85)	73 CIS (20) RRMS (34) SPMS (19)	Deep GM, regional volume fraction (to ICV) (thalamus).	Global T1 and T2 lesion volume.	Significant negative correlation between thalamic fraction and T2 LV (r=-0.75, p<0.001) and T1 LV (r=-0.60, p<0.001)[a].	Fair
Roosendaal et al., Mult Scler, 2011 (e1)	927 CIS (95) RRMS (657) SPMS (125) PPMS (50)	Total GM, global volume.	Global T1 and T2 lesion volumes (log-transformed for statistical analysis).	In separate multiple regression analyses, Log T1 and T2 LV were significant explanatory MRI-variables of GMV (both:β=-0.27, R^2=0.59 p<0.001)[a].	Fair
Sailer et al., Brain, 2003 (e75)	20 RRMS (11) SPMS (9)	Cortical GM, global thickness. Cortical GM, regional thickness.	Global T1 and T2 lesion volume.	Significant association between Cth and T1 [F(1,59)=6.17, p=0.03] and T2 [F(1,59)=18.04, p=0.001] LV[a]. Patients with a T1 LV of 3 ml or more showed significant cortical thinning when compared to healthy controls, mainly in frontal and temporal areas, which was	Fair

Supplemental eTable 1. Characteristics of cross-sectional studies

				not seen for patients with lower T1 LV. Patients with both low and high T2 LV had increased whole brain cortical thinning when compared to controls, but motor cortex thinning was only present in patients with a T2 LV of 20 ml or more (all *p*-values<0.05)[a].	
Sanfilipo et al., Neuroimage, 2005 (e67)	41 RRMS (35) SPMS (6)	Total GM, global volume.	Global T1 and T2 lesion volume.	Significant correlation between global GMV and T1 LV (*r*=-0.46), and T2 LV (*r*=-0.43) (*p*<0.01)[a].	Fair
Sepulcre et al., Arch Neurol, 2009 (e78)	61 CIS (22) RRMS (28) SPMS (5) PPMS (6)	Cortical GM, regional volume (occipital cortex). Deep GM, regional volume (lateral geniculate nucleus).	Regional presence of T1 and Gd-enhancing lesion volume.	Significant association in the right, but not left hemisphere, between LGN volume and presence of T1 lesions in the OR (F=26.23, R^2=0.28, *p*=0.02)[a]. No other WM lesions, either within or outside the optic pathway, associated with LGN volume. No association between occipital CV and the presence of T1 lesions anywhere in the brain.	Fair
Shiee et al., PLoS One, 2012 (e76)	60 RRMS (43) SPMS (9) PPMS (8)	Cortical GM, global volume. Deep GM, regional volume.	Global T2 lesion volume.	Significant correlation between T2 LV and CN (*r*=-0.32, *p*<0.01) and thalamus (*r*=-0.36, *p*<0.005) volume[a], but not with putamen or global CV.	Fair
Tedeschi et al., Neurology, 2005 (e68)	597 RRMS (427) SPMS (140) PPMS (30)	Total GM, global fraction (to ICV).	Global abnormal WM fraction (aWMf) (to ICV).	Significant negative correlation between aWMf and GMF (*r*=-0.58, 95%CI:-0.63, -0.52), *p*<0.001).	Fair
Zimmerman et al., J Neural Transm, 2015 (e69)	37 RRMS (34) SPMS (3)	Total GM, global fraction (to TBV). Deep GM, regional fraction (to TBV) and volume (putamen).	Global T2 lesion volume.	Significant correlation between T2 LV and GMF (*r*=-0.576, *p*=0.016), and putamen fraction (*r*=-0.674, *p*=0.003)[a]. Patients with predominantly spinal lesions demonstrated significantly larger putamen volumes compared to patients with cerebral lesions and no detectable spinal lesions (*p*=0.018)[a].	Poor
Comparison between disease phenotypes					
Antulov et al., J Neuroimaging, 2011 (e23)	RRMS (67) SPMS (43)	Total GM, global volume fraction (to ICV). Cortical GM, regional volume fraction (to ICV).	Regional T1 and T2 lesion volume.	Significant correlations between regional T2 LV and total GMF (controlling for regional GMF) were found in 9 and 4 of 26 regions for all MS and RRMS patients, respectively (*r*-values ranging from -0.20 to -0.49,	Fair

Supplemental eTable 1. Characteristics of cross-sectional studies

		Deep GM, regional volume fraction (to ICV).		$p \leq 0.001$)[a]. No significant correlations were seen for SPMS patients. Significant correlations between regional T1 LV and total GMF (controlling for regional GMF) were found in 5 regions for all MS patients, and 3 regions for RRMS patients (r-values ranging from -0.33 to -0.50, $p \leq 0.001$)[a]. No correlations were significant in the SPMS group. For all MS, RR and SPMS patients, correlations between regional T2 LV and regional GMF (controlling for total GMF) were not significant for any of the 26 regions, whereas for all MS patients, regional T1 LV approached significance with GMF in the right medial orbital frontal (r=-0.267) and left posterior basal ganglia/thalamus (r=-0.276) ($p \leq 0.01$)[a].	
Bergsland et al., AJNR, 2012 (e2)	CIS (212), Early RRMS (177)	Cortical GM, global volume. Deep GM, global volume. Deep GM, regional volume.	Global T2 lesion volume and number. Global presence of Gd-enhancing lesions.	CIS patients with T2 LV above median (4.49 mL) had lower DGM (-3.97%, SED=0.51), CN (-5.73%, SED=0.13), thalamus (-4.42%, SED=0.19) (all p<0.001), GP (-2.50, SED=0.04, p<0.007), hippocampus (-3.70, SED=0.13, p<0.004), and putamen (-2.67, SED=0.14, p<0.01) volumes compared to those with T2 LV below median. CV did not differ between CIS patients with T2 LV above or below the median value. In CIS patients, no differences in any GM volume measure were found between the subgroups when divided for the number of T2 lesions and presence of Gd-enhancing lesions. In patients with CIS, correlations were found between T2 LV and total DGM, CN, thalamus (all p<0.0001), and hippocampus (p=0.001) volumes (r-values ranging from -0.218 to -0.329)[a]. In patients with early RRMS, modest to strong correlations were found between T2 LV and every GM volume measure tested: NCV, total DGMV, CN, putamen, GP, thalamus, hippocampus, nucleus	Good

Supplemental eTable 1. Characteristics of cross-sectional studies

				accumbens (all $p<0.0001$) and amygdala ($p=0.001$) volumes (r-values ranging from -0.246 to -0.619)[a].	
Ceccarelli et al., Neuroimage, 2008 (e40)	CIS (28) RRMS (26) SPMS (27) PPMS (18)	Cortical GM, regional volume. Deep GM, regional volume.	Global T2 lesion volume. Regional T2 lesion volume.	Significant correlation between T2 LV and lower GMV in the right (r=-0.70) and left (r=-0.81) thalamus in patients with RRMS[a]. Significant correlation between T2 LV and lower GMV in the thalami, CN, bilateral middle frontal gyrus, left inferior parietal lobule, bilateral parahippocampal gyrus and bilateral superior and inferior colliculus (r-values ranging from -0.63 to -0.88), in patients with SPMS[a]. Significant negative correlation between T2 LV and left thalamus (r=-0.94), left parahippocampal gyrus (r=-0.83) and right precentral gyrus (r=-0.85) volume in patients with PPMS[a]. In SPMS and PPMS, lower GMV in a given lobe was correlated to T2 LV within the same or adjacent lobes (r-values ranging from -0.65 to -0.91), while lower DGMV was associated to T2 LV in all lobes analysed (r-values ranging from -0.64 to -0.93)[a]. All $p<0.001$.	Fair
De Stefano et al., Neurology, 2003 (e31)	RRMS (65) PPMS (25)	Cortical GM, global volume.	Global T2 lesion volume.	In patients with RRMS, T2 LV correlated with CV (r=-0.47, $p<0.001$)[a]. In PPMS patients, no significant correlation was found.	Fair
Grothe et al., J Neurol, 2016 (e16)	RRMS (163) SPMS (50)	Total GM, global volume. Cortical GM, regional volume. Deep GM, regional volume. Cerebellar GM, regional volume.	Global T2 lesion volume.	Significant inverse relationship between total GMV and T2 LV in all patients (r=-0.514), patients with RRMS (r=-0.334) and SPMS (r=-0.460) (all $p<0.001$)[a]. Significant associations found between T2 LV and thalamus, various cortical and cerebellar regions in patients with RRMS (peak T-values ranging from 4.72 to 11.21), cerebellar, temporal and cingulate regions in patients with SPMS (peak T-values ranging from 5.17 to 6.17), and cortical, thalamic, cerebellar and DGM regions in the whole patient group (peak T-values	Fair

Supplemental eTable 1. Characteristics of cross-sectional studies

				ranging from 4.69 to 9.84)[a] (all $p<0.05$).	
Jehna et al., Ann Neurol, 2015 (e9)	CIS (91) RRMS (69)	Cortical GM, global and regional thickness.	Regional T2 lesion volume (periventricular and non-periventricular). Regional T2 lesion volume percentage (periventricular and non-periventricular) relative to total lesion volume.	Significant correlation between PV-LL% and mean Cth in patients with RRMS (r=-0.295, p=0.015)[a], but not in patients with CIS. In the entire cohort of patients, PV-LL% did not correlate with Cth. In patients with RRMS, PV-LL correlated significantly with Cth (r=-0.293, p=0.015)[a], but the NON-PV-LL did not. In patients with CIS, neither the PV-LL nor the NON-PV-LL correlated with Cth. Vertexwise analyses: In CIS patients, PV-LL% was associated with Cth in the caudal part of the anterior cingulate cortex (p=0.0466) and the right superior frontal cortex (p=0.003). In patients with RRMS, correlations were found for bilateral precuneus, left lingual cortex, left lateral occipital gyrus, bilateral superior parietal cortices, left middle temporal cortex, right rostral part of the middle frontal gyrus, left superior frontal cortex and right inferior parietal gyrus (p-values ranging from 0.0002 to 0.0396)[a].	Fair
Kalinin et al., AJNR, 2020 (e53)	71 RRMS (54) SPMS (12) PPMS (5)	Deep GM, global volume percentage (of ICV). Deep GM, total and regional volume percentage (of ICV).	Global T1 and T2 lesion volume.	All MS patients: DGMV was predicted by T1 LV, independent of T2 and cortical LV (R^2=0.52, standardized β=-0.41, p=0.02)[a]. Separated by disease duration: T2 LV was associated with DGM (R^2=0.34, standardized β=-0.70, SE=0.32, p=0.04), putamen (R^2=0.36, standardized β=-0.79, SE=0.31, p=0.02) and pallidum (R^2=0.32, standardized β=-0.77, SE=0.32, p=0.03) volume, in those with a disease duration <5 years. T1 LV was associated with DGM volume (R^2=0.71, standardized β=-0.62, SE=0.17, p=0.001) in those with a disease duration >5 years. Separated by EDSS:	Fair

Supplemental eTable 1. Characteristics of cross-sectional studies

				T2 LV was associated with thalamus volume (R^2=0.34, standardized β=-0.60, SE=0.22, *p*=0.01) (but none of the other volumes) in those with EDSS <4.0. No associations for T1 LV in this group. T1 LV was associated with DGM (R^2=0.77, standardized β=-0.79, SE=0.22, *p*=0.002) and thalamus (R^2=0.55, standardized β=-0.74, SE=0.31, *p*=0.03) volume in those with EDSS ≥4. No associations for T2 LV in this group. RRMS: Only T2 LV, and not T1 or cortical LV, associated with DGM (R^2=0.49, standardized β=-0.49, SE=0.22, *p*=0.04) and putamen (R^2=0.27, standardized β=-0.57, SE=0.27, *p*=0.04) volume. SPMS: T1 LV associated with DGMV (R^2=0.81, standardized β=-1.11, SE=0.44, *p*=0.04), but T2 LV did not.	
Pontillo et al., AJNR, 2019 (e55)	RRMS (52) Progressive MS (PMS) (25)	Deep GM, regional volume.	Global T2 lesion volume.	In regressions in which microstructural parameters from advanced MRI were also included, T2 LV was an independent predictor of lower volume in thalamus (standardized β=-0.347, *t*=-3.383, *p*=0.002), CN (standardized β=-0.279, *t*=-2.696, *p*=0.009), putamen (standardized β=-0.278, *t*=-2.367, *p*=0.02) and GP (standardized β=-0.481, *t*=-4.308, *p*<0.001) in all MS patients, and in thalamus (standardized β=-0.454, *t*=-3.214, *p*=0.003), CN (standardized β=-0.411, *t*=-2.938, *p*=0.006), putamen (standardized β=-0.502, *t*=-4.009, p<0.001) and GP (standardized β=-0.450, *t*=-3.407, *p*=0.001) in the RRMS patient group[a]. In the PMS group, this association was not found.	Fair
Riccitelli et al., HBM, 2011 (e45)	RRMS (22) SPMS (29) PPMS (22)	Cortical GM, regional volume. Deep GM, regional volume.	Local T2 lesion frequency map (Lesion Probability Map).	In RRMS and SPMS patients, a correspondence (by visual inspection) between the focal distribution of visible T2 lesions in WM structures and lower GMV in regions spatially close or known to be anatomically connected to these structures was seen. Such association was not found in PPMS patients.	Poor
Sicotte et al.,	RRMS (23)	Deep GM, regional volume	Global T2 lesion volume.	In RRMS patients, T2 LV was not correlated with total	Fair

Supplemental eTable 1. Characteristics of cross-sectional studies

Brain, 2008 (e52)	SPMS (11)	(hippocampus and its subregions).		or subregional hippocampal volumes. In SPMS patients, there was a significant correlation between T2 LV and CA1 atrophy (r=-0.69, p=0.018)[a].	
Steenwijk et al., Brain, 2016 (e34)	RRMS (130) SPMS (53) PPMS (25) Same patient group as in studies listed below (Steenwijk et al., HBM, 2015 and Steenwijk et al., Radiology, 2014).	Cortical GM, global thickness.	Global T2 lesion volume.	In a stepwise multiple regression analysis corrected for age and sex, T2 LV contributed to the model (with a negative association) for global Cth in all MS (standardized β=-0.172, p<0.001), RRMS (standardized β=-0.319, p<0.01) and PPMS (standardized β=-0.425, p<0.05)[a], but not in SPMS. For all MS and RRMS, FA in NAWM was also in the model, while for SPMS, there were no imaging metrics in the model at all.	Fair
Steenwijk et al., HBM, 2015 (e46)	RRMS (130) SPMS (53) PPMS (25) Same patient group as in studies listed over and below (Steenwijk et al., Brain, 2016 and Steenwijk et al., Radiology, 2014).	Cortical GM, regional thickness. Deep GM, regional volume.	Regional T2 lesion volume.	For the whole MS group, regional linear regression analyses showed that T2 LV in connected WM tracts explained lower DGMV in 7 out of 7 regions (avg. standardized β=-0.332), while lower Cth was explained by T2 LV in connected tracts in 28 out of 34 regions (avg. standardized β=-0.116)[a]. All regression analyses included mean tract NAWM FA as predictor. In RRMS patients, 26 of 34 models for regional Cth were significant and LV was a significant predictor in most models (avg. standardized β=-0.187). In SPMS 8 models were significant, with LV as a significant predictor in only 1 model (standardized β=-0.021). In PPMS 7 models were significant, with LV as a significant predictor in 2-3 models (avg. standardized β=-0.201)[a]. All 7 models for DGMV were significant in RRMS (LV significant predictor in 6/7, avg. standardized β=-0.349) and SPMS (LV significant predictor in 5/7, avg.	Fair

Supplemental eTable 1. Characteristics of cross-sectional studies

				standardized β=-0.264). In PPMS, all models except the models for amygdala, putamen and thalamus were significant, and LV was a significant predictor in all 4 models (avg. standardized β=-0.438)[a]. All $p<0.05$.	
Steenwijk, Radiology et al., 2014 (e17)	RRMS (130) SPMS (53) PPMS (25) Same patient group as in studies listed above (Steenwijk et al., HBM, 2015 and Steenwijk et al., Radiology, 2016).	Total GM, global volume. Cortical GM, global thickness. Deep GM, global volume.	Global T2 lesion volume.	Partial correlations with T2 LV corrected for age and sex were significant in all groups for GMV (all MS: r=-0.52, RRMS: r=-0.57, SPMS: r=-0.36, PPMS: r=-0.68, all $p<0.01$), DGMV (all MS: r=-0.70, RRMS: r=-0.74, SPMS: r=-0.65, PPMS: r=-0.71, all $p<0.001$) and Cth (all MS: r=-0.43, RRMS: r=-0.55, PPMS: r=-0.43, all $p<0.05$)[a] with as the sole exception the correlation between Cth and T2 LV in SPMS. In multiple regression analyses, T2 LV remained a significant predictor for GMV, DGMV and Cth in the total patient group (GMV: standardized β=-0.226, DGMV: standardized β=-0.407, Cth: standardized β=-0.216, all $p<0.01$) and patients with RRMS (GMV: standardized β=-0.231, DGMV: standardized β=-0.407, Cth: standardized β=-0.357, all $p<0.01$)[a]. For patients with SPMS, T2 LV remained as a significant predictor only for DGMV (standardized β=-0.385, $p<0.001$)[a] (absent in models predicting GMV and Cth). In patients with PPMS, T2 LV was not included in any of the models (never significant).	Fair

Abbreviations: WM, white matter; GM, gray matter; DGM, deep gray matter; CIS, clinically isolated syndrome; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis; PPMS, primary progressive multiple sclerosis; PMS, progressive multiple

Supplemental eTable 1. Characteristics of cross-sectional studies

sclerosis; Gd, gadolinium; LV, lesion volume; PV-LL, periventricular lesion load; GMV, gray matter volume; DGMV, deep gray matter volume; CV, cortical volume; NCV, normalized cortical volume; ICV, intracranial volume; TBV, total brain volume; Cth, cortical thickness; CSF, cerebrospinal fluid; GMF, gray matter fraction; aWMf, abnormal white matter fraction; NAWM, normal appearing white matter; FA, fractional anisotropy; GP, globus pallidus; CN, caudate nucleus; LGN, lateral geniculate nucleus; RTV, relative thalamic volume; CA1, cornu ammonis 1; OR, optic radiation; CST, corticospinal tract; PMC, primary motor cortex; RD, radial distance; EDSS, expanded disability status scale; FDR, false discovery rate; SS, sums of deviation squares; ms, mean square; SE, standard error; SED, standard error of difference; MRI, magnetic resonance imaging; T, tesla.

[a] Full statistical details (e.g., point and/or interval estimates) not provided in published original article or supplemental material.

Supplemental eTable 2. Characteristics of longitudinal studies

Longitudinal studies					
Reference (e-ref)	*Patients (n)* *Follow-up time /(variability)*	*GM atrophy measure*	*WM lesion measure*	*Results*	*Quality of evidence*
CIS					
Dalton et al., Brain, 2004 (e6)	58 3 years (range: 2.6-6 years)	Total GM, global fraction (to ICV) percentage change.	Global T1 and T2 lesion volume and volume change. Global Gd-enhancing lesion number.	Baseline T1 and T2 LV, and number of Gd-enhancing lesions did not correlate with GMF percentage change. GMF percentage change correlated negatively with changes in T1 (r=-0.3071, p=0.0426) and T2 (r=-0.4280, p=0.0037) LV[a].	Good
Varosanec et al., AJNR, 2015 (e7)	210 4 years	Total GM, global volume percentage change. Cortical GM, global volume percentage change. Deep GM, global volume percentage change. Deep GM, regional volume percentage change.	Global T2 and Gd-enhancing lesion volume and volume percentage change. Global T2 and Gd-enhancing lesion number. Global presence of Gd-enhancing lesions.	Over the follow-up period, global GM and cortical volume change was associated with baseline T2 and Gd-enhancing LV and number, as well as the presence of Gd-enhancing lesions (all $p \leq 0.004$). DGMV change was associated with baseline Gd-enhancing lesion number (p=0.029), and thalamus volume change associated with baseline T2 LV and T2 and Gd-enhancing lesion number (all $p \leq 0.018$)[a]. Patients with the highest baseline number of T2 and Gd-enhancing lesions progressed the most in cortical (all p<0,001) and thalamic (all $p \leq 0.018$) volume change[a]. Significant associations between the total cumulative number of new/enlarging T2 lesions and global GM, cortical, DGM and thalamic volume change (p-values ranging from 0.013 to 0.036). No such associations were found for changes in T2 or Gd-enhancing LV, or cumulative Gd-enhancing lesion number[a].	Good
RRMS					
Battaglini et al., J Neurol Sci, 2009 (e50)	59	Cortical GM, regional volume change.	Regional T2 lesion volume change.	The increase in T2 LV over 3 years appeared (by visual inspection) more pronounced alongside areas with significant CV loss.	Poor

Supplemental eTable 2. Characteristics of longitudinal studies

	3 years (range: 2-4.8 years)				
Bendfeldt et al., Neuroimage, 2009 (e27)	89 1 year (SD: ±0.9 months)	Total GM, global volume change. Cortical GM, regional volume change.	Global T1 and T2 lesion volume change. Global T2 lesion number.	Patients with increasing T1 and T2 LV after 1 year follow-up (n=45) showed significant global GM atrophy, and regional GM loss in the anterior and posterior cingulate, the temporal cortex, cerebellum, frontal and parietal cortex, the uncus and precuneus (all $p<0.01$)[a]. No GMV changes were found in patients without increasing lesion burden (n=44).	Fair
Bendfeldt et al., HBM, 2010 (e51)	89 1 year (SD: ±0.9 months)	Cortical GM, regional volume change.	Regional T1 and T2 lesion volume change.	In patients with progressive T1 and T2 LV (n=45), baseline WM lesion distribution (LPMs) did not correlate with cortical GM loss. In the same patient group, areas of significant LV changes did not anatomically overlap (by visual inspection) with areas of significant GMV loss.	Poor
Damasceno et al., Mult Scler, 2016 (e48)	42 2 years	Cortical GM, global thickness atrophy rate. Deep GM, regional volume atrophy rate.	Global T2 and Gd-enhancing lesion number. Global presence of Gd-enhancing lesions.	Patients with new/enlarging T2 or Gd-enhancing lesions had increased DGM atrophy rates (mean atrophy rate of -4.51 ±2.72 mm^3, F=4.14, p=0.024), but not thalamus or cortical atrophy rates, compared to patients without new/enlarging lesions. Significant interaction effect for group by time for Cth and thalamus volume over the 24-month period, where the group with MRI activity showed more cortical thinning (F=3.36, p=0.040) and thalamus atrophy (F=2.79, p=0.067). Similar interaction was not seen for total DGM atrophy[a]. Absence of new/enlarging T2 lesions over the follow-up was the only predictor of cortical thinning (β=1.85, 95% CI:0.61-3.08, p=0.0044), DGM (β=2.09, 95% CI:0.50-3.68, p=0.011) and thalamic atrophy rate (β=2.09, 95% CI:0.33-3.84, p=0.021).	Fair
Masuda et al., J Neurol Sci, 2019 (e29)	49 1 year	Total GM, global volume atrophy rate.	Global T2 lesion volume change.	The change in T2 LV over the follow-up did not correlate significantly with the annualized atrophy rate of the total GMV.	Fair

Supplemental eTable 2. Characteristics of longitudinal studies

Preziosa et al., Neurotherapeutics, 2019 (e28)	55 2 years	Total GM, global volume percentage change.	Global T1, T2 and Gd-enhancing lesion volume change. Global T1, T2 and Gd-enhancing lesion number.	In patients treated with fingolimod (but not natalizumab), increased T2 LV correlated with percentage GMV change after two years (r=-0.43, p=0.03).	Fair
Talmage et al., PLoS One, 2017 (e57)	15 1.8 years	Deep GM, regional volume change (thalamus).	Global T2 lesion volume.	Significant correlation between baseline T2 LV and thalamic volume loss over the follow-up (r=-0.586, R^2=0.344, p=0.027)[a].	Poor
Vidal-Jordana et al., J Neuroimaging, 2016 (e26)	84 15 months	Total GM, global volume percentage change.	Global Gd-enhancing lesion number. Global presence of Gd-enhancing lesions.	GMV decreased over the follow up, regardless of the presence of Gd-enhancing lesions at baseline or at follow-up. Increased number of Gd-enhancing lesions at baseline was not predictive of GMV change.	Good
Vidal-Jordana et al., Mult Scler, 2013 (e25)	39 2 years	Total GM, global volume fraction change.	Global T2 and Gd-enhancing lesion number. Global presence of Gd-enhancing lesions.	Patients with, and without Gd-enhancing lesions at baseline, did not differ in GMF change over the follow-up period. In a linear regression analysis, no predictors were found for changes in GMF.	Fair
PPMS					
Sastre-Garriga et al., Brain, 2005 (e60)	31 1 year	Total GM, global volume fraction change.	Global T2 lesion volume. Global Gd-enhancing lesion number.	No significant associations for GMF change were found with either baseline T2 LV, or number of Gd-enhancing lesions. In stepwise regression models, baseline T2 LV and number of Gd-enhancing lesions did not predict GMF percentage change.	Fair
MS					
Pongratz et al., Brain Behav, 2019 (e87)	144 CIS (80) RRMS (64) 3 years (range: 2.5-3.5 years)	Total GM, global volume change.	Global T2 lesion volume change.	T2 lesion shrinking was not associated with changes in GMV, after about 1 year and 3 years of follow-up.	Fair

Supplemental eTable 2. Characteristics of longitudinal studies

Fox et al., Int J Mol Sci, 2016 (e90)	92 CIS (2) RRMS (69) SPMS (16) PPMS (4) RPMS (1) 1 year	Deep GM, regional volume change (lateral geniculate nucleus).	Regional T1 lesion number.	The relative frequency of patients with volume decrease in LGN over 12 months was higher in groups with higher lesion numbers along the optic radiation. In a logistic regression analysis of LGN volume reduction and ipsilateral lesion number along the optic radiation, the number of new (OR=14.01, AUC=0.65, p=0.0006), chronic enlarging (OR=1.90, AUC=0.59, p=0.0056) and chronic shrinking (OR=12.41, AUC=0.66 p=0.0001) T1 lesions were all significant predictors[a].	Poor
Lee et al., Mult Scler, 2018 (e88)	19 RRMS (12) SPMS (7) 5.5 years (range: 1.5-10.5 years)	Total GM, global volume percentage change.	Global T1 lesion volume.	Baseline T1 LV was not a significant predictor of GMV loss over the mean follow-up duration of 5.5 years after IA/aHSCT.	Fair
Tedeschi et al., Mult Scler, 2009 (e89)	267 RRMS (208) SPMS (59) 2 years (range: 24-26 months)	Total GM, global fraction (to ICV) and percentage fraction change.	Global abnormal WM lesion fraction (to ICV) and percentage fraction change.	No significant correlation between aWMf at baseline and change in GMF during the 2 year follow up. Significant correlation between baseline GMF and change in aWMf during the 2 year follow up (r=-0.180, p=0.000)[a].	Good
Comparisons between disease phenotypes					
Fisher et al., Ann Neurol, 2008 (e24)	CIS (7) RRMS (36) SPMS (27) 4 years (range: 3.4-4.8 years)	Total GM, global volume fraction change.	Global T2 lesion volume and volume change. Global presence of Gd-enhancing lesions.	RRMS: in multiple regression models, the presence of Gd-enhancing lesions at baseline (standardized β=-0.28, p=0.04) and on-study increasing T2 LV (standardized β=-0.46, p=0.0004) were associated with greater rates of GMF percentage change over 4 years[a]. SPMS: no baseline MRI-measurements or on-study changes were significant predictors of GMF change.	Good
Tavazzi et al., J Neurol, 2020 (e10)	127 CIS (20) RRMS (85) PMS (42) 5.5 years (SD: ±0.5 years)	Cortical GM, global volume and volume percentage change. Deep GM, regional volume and volume percentage change (thalamus).	Global T2 lesion volume change.	In RRMS, atrophied T2 LV was associated with baseline thalamic volume (r=-0.384, p=0.004) and thalamic volume change (r=-0.430, p=0.004)[a]. No associations were found between atrophied T2 LV and baseline CV, or CV change. In PMS and CIS, no significant associations were found.	Fair

Supplemental eTable 2. Characteristics of longitudinal studies

	In patients with disease progression over the follow-up, atrophied T2 LV correlated significantly with baseline thalamic volume (r=-0.620, p=0.003) and thalamic volume change (r=-0.672, p=0.003)[a]. No associations were found between atrophied T2 LV and baseline CV, or CV change, or in patients without disease progression.

Abbreviations: WM, white matter; GM, gray matter; DGM, deep gray matter; CIS, clinically isolated syndrome; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis; PPMS, primary progressive multiple sclerosis; PMS, progressive multiple sclerosis; RPMS, relapsing-progressive multiple sclerosis; Gd, gadolinium; LV, lesion volume; GMV, gray matter volume; DGMV, deep gray matter volume; CV, cortical volume; ICV, intracranial volume; Cth, cortical thickness; GMF, gray matter fraction; aWMf, abnormal white matter fraction; LGN, lateral geniculate nucleus; OR, odds ratio; AUC, area under the curve; CI, confidence interval; IA/aHSCT, immunoablation and autologous stem cell transplantation; MRI, magnetic resonance imaging; LPM, lesion probability map; SD, standard deviation.

[a] Full statistical details (e.g., point and/or interval estimates) not provided in published original article or supplemental material.

Supplemental eTable 3. Characteristics of cross-sectional and longitudinal studies.

Cross-sectional and longitudinal studies					
Reference (e-ref)	*Patients (n)* *Follow-up time/ (variability)*	*GM atrophy measure*	*WM lesion measure*	*Results*	*Quality of evidence*
CIS					
Cacciaguerra et al., Mult Scler, 2019 (e11)	35 2 years	Deep GM, regional volume and volume change (hippocampus).	Regional T1, T2 and Gd-enhancing lesion volume and volume change.	Cross-sectional: significant correlation between ipsilateral T2 LV and lower hippocampal volume in the subiculum, bilaterally, at baseline and at month 3 (p-values ranging from 0.001-0.05), and in the CA1 subfield of the tail at month 12 and 24 (p-values ranging from 0.01-0.05)[a,b]. For T1 LV, similar results were found. T1 and T2 LV correlated with higher bilateral dentate gyrus volume at baseline and month 3 (p-values ranging from 0.05-0.001)[a.b]. Positive correlation between Gd-enhancing LV and left dentate gyrus volume at baseline and month 24 (p<0.05)[a.b]. Longitudinal: changes of LV did not correlate with hippocampal volume change.	Fair
RRMS					
Gaetano et al., Neurology, 2018 (e49)	2064 2 years	Cortical GM, global volume and volume change. Deep GM, global volume and volume change. Deep GM, regional volume and volume change (thalamus).	Global T2 lesion volume. Global Gd-enhancing lesion number.	Cross-sectional: at baseline, DGM volume was -0.258 (95% CI:-0.278, -0.239) and thalamus volume was -0.115 (95% CI:-0.124, -0.106) cm^3 smaller per cm^3 of T2 LV (all p<0.0001). Longitudinal: annual additional percentage volume loss per cm^2 of baseline T2 LV was 0.053% (95% CI:0.070, 0.037) for DGM, 0.058 (95% CI:0.076, 0.041) for thalamus and 0.052 (95% CI:0.068, 0.035) for cortical volume (all p-<0.0001). Annual additional percentage volume	Good

Supplemental eTable 3. Characteristics of cross-sectional and longitudinal studies.

				loss for each additional Gd-enhancing lesion at baseline was 0.060% (95% CI:0.094, 0.025, p=0.0007) for DGM, 0.039 (95% CI:0.076, 0.002, p=0.0372) for thalamus and 0.046 (95% CI:0.082, 0.011, p=0.0102) for cortical volume. Baseline T2 LV was a significant predictor of cortical and DGM atrophy, regardless of the presence of Gd-enhancing lesions at baseline (all p<0.001)[a].	
Tiberio et al., Neurology, 2005 (e14)	20 2 years (range: 23-28 months)	Total GM, global volume fraction (to ICV) and percentage volume fraction change.	Global T1, T2, and Gd-enhancing lesion volume and volume change.	Cross-sectional: significant correlation between GMF and T2 (r=-0.438, p=0.047) and T1 (r=-0.496, p=0.022) LV[a], but not with Gd-enhancing LV. Longitudinal: changes in GMF were not correlated with Gd-enhancing LV at baseline, or changes in any LV over the follow-up.	Fair
Zivadinov et al., Mult Scler Int, 2013 (e32)	136 2 years	Cortical GM, global volume and percentage volume change.	Global T2 lesion volume and volume percentage change.	Cross-sectional: Significant correlation between baseline T2 LV and baseline CV (β=-1.39, SE=0.28, p<0.0001). Longitudinal: The percentage change in T2 LV was not significantly related to the percentage change in CV, not between baseline and year 1, year 1 and year 2, or baseline and year 2.	Fair
MS					
Ciampi et al., Mult Scler, 2017 (e64)	38 RRMS (36) SPMS (2) 3 years	Total GM, global volume fraction (to ICV) and percentage fraction change. Cortical GM, global volume and thickness. Cortical GM, regional volume and thickness. Deep GM, global volume. Deep GM, regional volume.	Global presence of Gd-enhancing lesions.	Cross-sectional: patients with or without Gd-enhancing lesions at baseline did not differ in GMF. No significant correlation between the presence of Gd-enhancing lesions at baseline and total, cortical, DGM (total and regional), brainstem and cerebellar GM volume, or with Cth (total and regional). Longitudinal: presence of Gd-enhancing lesions at baseline was not associated with any of the changes in GM over the 3-year study period.	Fair

Supplemental eTable 3. Characteristics of cross-sectional and longitudinal studies.

Eshaghi et al., Brain, 2018 (e71)	1214 CIS (253) RRMS (708) SPMS (128) PPMS (125) 2.41 years (SD: ±1.97 years)	Cortical GM, regional volume and volume change. Deep GM, regional volume and volume change. (Volumes used in an event-based model, to determine sequence of occurrence of atrophy)	Global T2 lesion volume and volume change.	Cross-sectional: significant association between the event-based model stage and T2 LV at baseline (standardized β =0.11, p<0.001)[a]. Longitudinal: no significant association between the rate of change in the event-based model stage over time and the rate of increase in T2 LV.	Fair
Fuchs et al., AJNR, 2018 (e86)	176 CIS (16) RRMS (114) SPMS (39) PPMS (7) 5 years (SD: ±0.7 years)	Deep GM, regional volume and volume change.	Global T2 lesion volume and volume change. Percentage tract disruption; the percentage of normative data-base derived connected tract streamlines that pass through the given lesion mask and are considered disrupted.	Cross-sectional: T2 LV was a significant predictor for DGM volume in 12 out of 14 regions (R^2 ranging from 0.108 to 0.198, p-values ranging from 0.000 to 0.019)[a]. Adding tract disruption to the model, 13 out of 14 DGM regions could be predicted (R^2 ranging from 0.108 to 0.503, all p=0.000)[a]. Controlling for T2 LV, tract disruption could only predict DGM volume in 3 out of 14 regions (R^2 ranging from 0.008 to 0.025, p-values ranging from 0.014 to 0.049)[a], but none after correction for multiple comparisons. Longitudinal: Controlling for T2 LV, tract disruption could predict DGM atrophy rate in 5 of 14 regions (R^2 ranging from 0.018 to 0.075, p-values ranging from 0.000 to 0.034)[a], but only 1 (R^2=0.075, p=0.000)[a] after correction for multiple comparisons.	Fair
Tsagkas et al., HBM, 2020 (e77)	243 RRMS (180) SPMS (51) PPMS (12) 4.36 years (SD: ±2.03)	Cortical GM, regional cortical thickness and cortical thickness change.	Global T2 lesion volume and volume change. Regional T2 lesion volume and volume change.	Cross-sectional: Global T2 LV was negatively associated with the average Cth in regions symmetrically in nearly the whole cortex (mean t-values =-4.00±1.36 (right hemisphere) and -4.24±1.35 (left hemisphere), all p<0.05). Similarly, regional T2 LV in the left and right frontal, temporal, parietal, and occipital areas were associated with lower Cth in widespread bilateral cortical regions (all p<0.05)[a]. Longitudinal: Changes in global T2 LV did not associate with changes in Cth over time.	Good

Supplemental eTable 3. Characteristics of cross-sectional and longitudinal studies.

	Left temporal changes in T2 LV associated negatively with change in Cth in the left cuneus (mean t-value=-4.47±0.35) and precuneus (mean t-value=-4.43±0.21) (all $p<0.05$). A negative association was also shown between changes in left occipital T2 LV and Cth change in bilateral temporal, parietal and occipital regions (mean t-values ranging from -3.06 to -3.97±0.09-0.54, all $p<0.05$).

Abbreviations: WM, white matter; GM, gray matter; DGM, deep gray matter; CIS, clinically isolated syndrome; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple sclerosis; PPMS, primary progressive multiple sclerosis; Gd, gadolinium; LV, lesion volume; DGM, deep gray matter; CV, cortical volume; ICV, intracranial volume; Cth, cortical thickness; GMF, gray matter fraction; CA1, cornu ammonis 1; SD, standard deviation; SE, standard error.

[a] Full statistical details (e.g., point and/or interval estimates) not provided in published original article or supplemental material.

[b] Spearmans rho ranges indicated in color bar in Figure 4 in the original study.

Supplementary eReferences

e1. Roosendaal S.D, Bendfeldt K, Vrenken H et al. Grey matter volume in a large cohort of MS patients: relation to MRI parameters and disability. *Mult Scler*. 2011;**17**(9): p. 1098-106.

e2. Bergsland N, Horakova D, Dwyer M.G et al. Subcortical and Cortical Gray Matter Atrophy in a Large Sample of Patients with Clinically Isolated Syndrome and Early Relapsing-Remitting Multiple Sclerosis. *AJNR Am J Neuroradiol.* 2012;**33**(8): p. 1573-1578.

e3. Durhan G, Diker S, Has A.C et al. Assessment of the effect of cigarette smoking on regional brain volumes and lesion load in patients with clinically isolated syndrome. *Int J Neurosci*. 2016;**126**(9): p. 805-811.

e4. Henry R.G, Shieh M, Okuda D.T, Evangelista A, Gorno-Tempini M.L, Pelletier D. Regional grey matter atrophy in clinically isolated syndromes at presentation. *J Neurol Neurosurg Psychiatry*. 2008;**79**(11): p. 1236-44.

e5. Labiano-Fontcuberta A, Mato-Abad V, Alvarez-Linera J et al. Gray Matter Involvement in Radiologically Isolated Syndrome. *Medicine*. 2016;**95**(13): p. e3208.

e6. Dalton C.M, Chard D.T, Davies G.R et al. Early development of multiple sclerosis is associated with progressive grey matter atrophy in patients presenting with clinically isolated syndromes. *Brain*. 2004;**127**(Pt 5): p. 1101-7.

e7. Varosanec M, Uher T, Horakova D et al. Longitudinal Mixed-Effect Model Analysis of the Association between Global and Tissue-Specific Brain Atrophy and Lesion Accumulation in Patients with Clinically Isolated Syndrome. *AJNR Am J Neuroradiol*. 2015;**36**(8): p. 1457-64.

e8. Jenkins T.M, Ciccarelli O, Atzori M et al. Early pericalcarine atrophy in acute optic neuritis is associated with conversion to multiple sclerosis. *J Neurol Neurosurg Psychiatry*. 2011;**82**(9): p. 1017-21.

e9. Jehna M, Pirpamer L, Khalil M et al. Periventricular lesions correlate with cortical thinning in multiple sclerosis. *Ann Neurol*. 2015;**78**(4): p. 530-9.

e10. Tavazzi E, Bergsland N, Kuhle J et al. A multimodal approach to assess the validity of atrophied T2-lesion volume as an MRI marker of disease progression in multiple sclerosis. *J Neurol*. 2020;**267**(3): p. 802-811.

e11. Cacciaguerra L, Pagani E, Mesaros S et al. Dynamic volumetric changes of hippocampal subfields in clinically isolated syndrome patients: A 2-year MRI study. *Mult Scler*. 2019;**25**(9): p. 1232-1242.

e12. Henry R.G, Shieh M, Amirbekian B, Chung S, Okuda D.T, Pelletier D. Connecting white matter injury and thalamic atrophy in clinically isolated syndromes. *J Neurol Sci*. 2009;**282**(1-2): p. 61-6.

e13. Chard D.T, Griffin C.M, Parker G.J, Kapoor R, Thompson A.J, Miller D.H. Brain atrophy in clinically early relapsing-remitting multiple sclerosis. *Brain*. 2002;**125**(Pt 2): p. 327-37.

e14. Tiberio M, Chard D.T, Altman D.R et al. Gray and white matter volume changes in early RRMS: a 2-year longitudinal study. *Neurology*. 2005;**64**(6): p. 1001-7.

e15. Dolezal O, Dwyer M.G, Horakova D et al. Detection of cortical lesions is dependent on choice of slice thickness in patients with multiple sclerosis. *Int Rev Neurobiol*. 2007;**79**: p. 475-89.

e16. Grothe M, Lotze M, Langner S, Dressel A. The role of global and regional gray matter volume decrease in multiple sclerosis. *J Neurol*. 2016;**263**(6): p. 1137-45.

e17. Steenwijk M.D, Daams M, Pouwels P.J et al. What explains gray matter atrophy in long-standing multiple sclerosis? *Radiology*. 2014;**272**(3): p. 832-42.

e18. Toth E, Szabo N, Csete G et al. Gray Matter Atrophy Is Primarily Related to Demyelination of Lesions in Multiple Sclerosis: A Diffusion Tensor Imaging MRI Study. *Front Neuroanat*. 2017;**11**: p. 23.
e19. Prinster A, Quarantelli M, Lanzillo R et al. A voxel-based morphometry study of disease severity correlates in relapsing-- remitting multiple sclerosis. *Mult Scler*. 2010;**16**(1): p. 45-54.
e20. Quarantelli M, Ciarmiello A, Morra V.B et al. Brain tissue volume changes in relapsing-remitting multiple sclerosis: correlation with lesion load. *Neuroimage*. 2003; **18**(2): p. 360-6.
e21. Ceccarelli A, Rocca M.A, Falini A et al. Normal-appearing white and grey matter damage in MS. A volumetric and diffusion tensor MRI study at 3.0 Tesla. *J Neurol.* 2007;**254**(4): p. 513-8.
e22. Sbardella E, Petsas N, Tona F et al. Assessing the correlation between grey and white matter damage with motor and cognitive impairment in multiple sclerosis patients. *PLoS One*. 2013;**8**(5): p. e63250.
e23. Antulov R, Carone D.A, Bruce J et al. Regionally distinct white matter lesions do not contribute to regional gray matter atrophy in patients with multiple sclerosis. *J Neuroimaging*. 2011;**21**(3): p. 210-8.
e24. Fisher E, Lee J.C, Nakamura K, Rudick R.A. Gray matter atrophy in multiple sclerosis: A longitudinal study. *Ann Neurol.* 2008;**64**(3): p. 255-265.
e25. Vidal-Jordana A, Sastre-Garriga J, Perez-Miralles F et al. Early brain pseudoatrophy while on natalizumab therapy is due to white matter volume changes. *Mult Scler*. 2013;**19**(9): p. 1175-81.
e26. Vidal-Jordana A, Sastre-Garriga J, Perez-Miralles F et al. Brain Volume Loss During the First Year of Interferon-Beta Treatment in Multiple Sclerosis: Baseline Inflammation and Regional Brain Volume Dynamics. *J Neuroimaging*. 2016;**26**(5): p. 532-8.
e27. Bendfeldt K, Kuster P, Traud S et al. Association of regional gray matter volume loss and progression of white matter lesions in multiple sclerosis - A longitudinal voxel-based morphometry study. *Neuroimage*. 2009;**45**(1): p. 60-7.
e28. Preziosa P, Rocca M.A, Riccitelli G.C et al. Effects of Natalizumab and Fingolimod on Clinical, Cognitive, and Magnetic Resonance Imaging Measures in Multiple Sclerosis. *Neurotherapeutics*. 2019;**17**(1):p.208-217.
e29. Masuda H, Mori M, Hirano S et al. Relapse numbers and earlier intervention by disease modifying drugs are related with progression of less brain atrophy in patients with multiple sclerosis. *J Neurol Sci*. 2019;**403**: p. 78-84.
e30. Al-Radaideh A, Athamneh I, Alabadi H, Hbahbih M. Cortical and Subcortical Morphometric and Iron Changes in Relapsing-Remitting Multiple Sclerosis and Their Association with White Matter T2 Lesion Load : A 3-Tesla Magnetic Resonance Imaging Study. *Clin Neuroradiol*. 2019;**29**(1): p. 51-64.
e31. De Stefano N, Matthews P.M, Filippi M et al. Evidence of early cortical atrophy in MS: relevance to white matter changes and disability. *Neurology*. 2003;**60**(7): p. 1157-62.
e32. Zivadinov R, Tekwe C, Bergsland N et al. Bimonthly Evolution of Cortical Atrophy in Early Relapsing-Remitting Multiple Sclerosis over 2 Years: A Longitudinal Study. *Mult Scler Int*. 2013;**2013**: p. 231345.
e33. Calabrese M, Rinaldi F, Mattisi I et al. Widespread cortical thinning characterizes patients with MS with mild cognitive impairment. *Neurology*. 2010;**74**(4): p. 321-8.
e34. Steenwijk M.D, Geurts J.J, Daams M et al. Cortical atrophy patterns in multiple sclerosis are non-random and clinically relevant. *Brain*. 2016;**139**(Pt 1): p. 115-26.

e35. Charil A, Dagher A, Lerch J.P, Zijdenbos A.P, Worsley K.J, Evans A.C. Focal cortical atrophy in multiple sclerosis: relation to lesion load and disability. *Neuroimage*. 2007; **34**(2): p. 509-17.
e36. Datta S, Staewen T.D, Cofield S.S et al. Regional gray matter atrophy in relapsing remitting multiple sclerosis: baseline analysis of multi-center data. *Mult Scler Relat Disord.* 2015;**4**(2): p. 124-36.
e37. Tao G, Datta S, He R, Nelson F, Wolinsky J.S, Narayana P.A. Deep gray matter atrophy in multiple sclerosis: a tensor based morphometry. *J Neurol Sci*. 2009;**282**(1-2): p. 39-46.
e38. Hasan K.M, Walimuni I.S, Abid H, Datta S, Wolinsky J.S, Narayana P.A. Human brain atlas-based multimodal MRI analysis of volumetry, diffusimetry, relaxometry and lesion distribution in multiple sclerosis patients and healthy adult controls: implications for understanding the pathogenesis of multiple sclerosis and consolidation of quantitative MRI results in MS. *J Neurol Sci*. 2012; **313**(1-2): p. 99-109.
e39. Narayana P.A, Govindarajan K.A, Goel P et al. Regional cortical thickness in relapsing remitting multiple sclerosis: A multi-center study. *Neuroimage Clin*. 2012;**2**: p. 120-31.
e40. Ceccarelli A, Rocca M.A, Pagani E et al. A voxel-based morphometry study of grey matter loss in MS patients with different clinical phenotypes. *Neuroimage.* 2008;**42**(1): p. 315-22.
e41. Duan Y, Liu Y, Liang P et al. Comparison of grey matter atrophy between patients with neuromyelitis optica and multiple sclerosis: a voxel-based morphometry study. *Eur J Radiol*. 2012;**81**(2): p. e110-4.
e42. Prinster A, Quarantelli M, Orefice G et al. Grey matter loss in relapsing-remitting multiple sclerosis: a voxel-based morphometry study. *Neuroimage*. 2006;**29**(3): p. 859-67.
e43. Kuceyeski A.F, Vargas W, Dayan M et al. Modeling the relationship among gray matter atrophy, abnormalities in connecting white matter, and cognitive performance in early multiple sclerosis. *AJNR Am J Neuroradiol.* 2015;**36**(4): p. 702-9.
e44. Riccitelli G, Rocca M.A, Pagani E et al. Mapping regional grey and white matter atrophy in relapsing-remitting multiple sclerosis. *Mult Scler.* 2012;**18**(7): p. 1027-37.
e45. Riccitelli G, Rocca M.A, Pagani E et al. Cognitive impairment in multiple sclerosis is associated to different patterns of gray matter atrophy according to clinical phenotype. *Hum Brain Mapp*. 2011;**32**(10): p. 1535-43.
e46. Steenwijk M.D, Daams M, Pouwels P.J et al. Unraveling the relationship between regional gray matter atrophy and pathology in connected white matter tracts in long-standing multiple sclerosis. *Hum Brain Mapp*. 2015;**36**(5): p. 1796-807.
e47. Bergsland N, Lagana M.M, Tavazzi E et al. Corticospinal tract integrity is related to primary motor cortex thinning in relapsing-remitting multiple sclerosis. *Mult Scler*. 2015;**21**(14): p. 1771-80.
e48. Damasceno A, Damasceno B.P, Cendes F. No evidence of disease activity in multiple sclerosis: Implications on cognition and brain atrophy. *Mult Scler*. 2016;**22**(1): p. 64-72.
e49. Gaetano L, Haring D.A, Radue E.W et al. Fingolimod effect on gray matter, thalamus, and white matter in patients with multiple sclerosis. *Neurology*. 2018;**90**(15): p. e1324-e1332.
e50. Battaglini M, Giorgio A, Stromillo M.L et al. Voxel-wise assessment of progression of regional brain atrophy in relapsing-remitting multiple sclerosis. *J Neurol Sci*. 2009;**282**(1-2): p. 55-60.

e51. Bendfeldt K, Blumhagen J.O, Egger H et al. Spatiotemporal distribution pattern of white matter lesion volumes and their association with regional grey matter volume reductions in relapsing-remitting multiple sclerosis. *Hum Brain Mapp*. 2010;**31**(10): p. 1542-55.
e52. Sicotte N.L, Kern K.C, Giesser B.S et al. Regional hippocampal atrophy in multiple sclerosis. *Brain*. 2008;**131**(Pt 4): p. 1134-41.
e53. Kalinin I, Makshakov G, Evdoshenko E. The Impact of Intracortical Lesions on Volumes of Subcortical Structures in Multiple Sclerosis. *AJNR Am J Neuroradiol*. 2020. doi: https://doi.org/10.3174/ajnr.A6513
e54. Magon S, Chakravarty M.M, Amann M et al. Label-fusion-segmentation and deformation-based shape analysis of deep gray matter in multiple sclerosis: the impact of thalamic subnuclei on disability. *Hum Brain Mapp*. 2014;**35**(8): p. 4193-203.
e55. Pontillo G, Cocozza S, Lanzillo R et al. Determinants of Deep Gray Matter Atrophy in Multiple Sclerosis: A Multimodal MRI Study. *AJNR Am J Neuroradiol*. 2019;**40**(1): p. 99-106.
e56. Hasan K.M, Halphen C, Kamali A, Nelson F.M, Wolinsky J.S, Narayana P.A. Caudate nuclei volume, diffusion tensor metrics, and T(2) relaxation in healthy adults and relapsing-remitting multiple sclerosis patients: implications for understanding gray matter degeneration. *J Magn Reson Imaging*. 2009;**29**(1): p. 70-7.
e57. Talmage G.D, Coppes O.J.M, Javed A, Bernard J. Natalizumab stabilizes physical, cognitive, MRI, and OCT markers of disease activity: A prospective, non-randomized pilot study. *PLoS One*. 2017;**12**(4): p. e0173299.
e58. Furby J, Hayton T, Altmann D et al. Different white matter lesion characteristics correlate with distinct grey matter abnormalities on magnetic resonance imaging in secondary progressive multiple sclerosis. *Mult Scler*. 2009;**15**(6): p. 687-94.
e59. Sastre-Garriga J, Ingle G.T, Chard D.T, Ramio-Torrenta L, Miller D.H, Thompson A.J. Grey and white matter atrophy in early clinical stages of primary progressive multiple sclerosis. *Neuroimage*. 2004;**22**(1): p. 353-9.
e60. Sastre-Garriga J, Ingle G.T, Chard D.T et al. Grey and white matter volume changes in early primary progressive multiple sclerosis: a longitudinal study. *Brain*. 2005;**128**(Pt 6): p. 1454-60.
e61. Galego O, Gouveia A, Batista S, Moura C, Machado E. Brain atrophy and physical disability in primary progressive multiple sclerosis: A volumetric study. *Neuroradiol J*. 2015;**28**(3): p. 354-8.
e62. Sepulcre J, Sastre-Garriga J, Cercignani M, Ingle G.T, Miller D.H, Thompson A.J. Regional gray matter atrophy in early primary progressive multiple sclerosis: a voxel-based morphometry study. *Arch Neurol*. 2006;**63**(8): p. 1175-80.
e63. Mesaros S, Rocca M.A, Pagani E et al. Thalamic damage predicts the evolution of primary-progressive multiple sclerosis at 5 years. *AJNR Am J Neuroradiol*. 2011;**32**(6): p. 1016-20.
e64. Ciampi E, Pareto D, Sastre-Garriga J et al. Grey matter atrophy is associated with disability increase in natalizumab-treated patients. *Mult Scler*. 2017;**23**(4): p. 556-566.
e65. Fisniku L.K, Chard D.T, Jackson J.S et al. Gray matter atrophy is related to long-term disability in multiple sclerosis. *Ann Neurol*. 2008;**64**(3): p. 247-54.
e66. Fragoso Y.D, Wille P.R, Abreu M et al. Correlation of clinical findings and brain volume data in multiple sclerosis. *J Clin Neurosci*. 2017;**44**: p. 155-157.
e67. Sanfilipo M.P, Benedict R.H, Sharma J, Weinstock-Guttman B, Bakshi R. The relationship between whole brain volume and disability in multiple sclerosis: a comparison of normalized gray vs. white matter with misclassification correction. *Neuroimage*. 2005;**26**(4): p. 1068-77.

e68. Tedeschi G, Lavorgna L, Russo P et al. Brain atrophy and lesion load in a large population of patients with multiple sclerosis. *Neurology*. 2005;**65**(2): p. 280-5.
e69. Zimmermann H, Rolfsnes H.O, Montag S et al. Putaminal alteration in multiple sclerosis patients with spinal cord lesions. *J Neural Transm*. 2015;**122**(10): p. 1465-73.
e70. Calabrese M, Atzori M, Bernardi V et al. Cortical atrophy is relevant in multiple sclerosis at clinical onset. *J Neurol*. 2007;**254**(9): p. 1212-20.
e71. Eshaghi A, Marinescu R.V, Young A.L et al. Progression of regional grey matter atrophy in multiple sclerosis. *Brain*. 2018;**141**(6): p. 1665-1677.
e72. Hier D.B, Wang J. Reduced cortical surface area in multiple sclerosis. *Neurol Res*. 2007;**29**(3): p. 231-2.
e73. Muhlau M, Buck D, Forschler A et al. White-matter lesions drive deep gray-matter atrophy in early multiple sclerosis: support from structural MRI. *Mult Scler*. 2013;**19**(11): p. 1485-92.
e74. Pareto D, Sastre-Garriga J, Auger C et al. Juxtacortical Lesions and Cortical Thinning in Multiple Sclerosis. *AJNR Am J Neuroradiol*. 2015;**36**(12): p. 2270-6.
e75. Sailer M, Fischl B, Salat D et al. Focal thinning of the cerebral cortex in multiple sclerosis. *Brain*. 2003;**126**(Pt 8): p. 1734-44.
e76. Shiee N, Bazin P.L, Zackowski K.M et al. Revisiting brain atrophy and its relationship to disability in multiple sclerosis. *PLoS One*. 2012;**7**(5): p. e37049.
e77. Tsagkas C, Chakravarty M.M, Gaetano L et al. Longitudinal patterns of cortical thinning in multiple sclerosis. *Hum Brain Mapp*. 2020;**41**(8): p. 2198-2215.
e78. Sepulcre J, Goni J, Masdeu J.C et al. Contribution of white matter lesions to gray matter atrophy in multiple sclerosis: evidence from voxel-based analysis of T1 lesions in the visual pathway. *Arch Neurol*. 2009;**66**(2): p. 173-9.
e79. Bermel R.A, Innus M.D, Tjoa C.W, Bakshi R. Selective caudate atrophy in multiple sclerosis: a 3D MRI parcellation study. *Neuroreport*. 2003;**14**(3): p. 335-9.
e80. Deppe M, Kramer J, Tenberge J.G et al. Early silent microstructural degeneration and atrophy of the thalamocortical network in multiple sclerosis. *Hum Brain Mapp*. 2016;**37**(5): p. 1866-79.
e81. Hasan K.M, Walimuni I.S, Abid H et al. Multimodal quantitative magnetic resonance imaging of thalamic development and aging across the human lifespan: implications to neurodegeneration in multiple sclerosis. *J Neurosci*. 2011;**31**(46): p. 16826-32.
e82. Longoni G, Rocca M.A, Pagani E et al. Deficits in memory and visuospatial learning correlate with regional hippocampal atrophy in MS. *Brain Struct Funct*. 2015;**220**(1): p. 435-44.
e83. Louapre C, Govindarajan S.T, Gianni C et al. Heterogeneous pathological processes account for thalamic degeneration in multiple sclerosis: Insights from 7 T imaging. *Mult Scler*. 2018;**24**(11): p. 1433-1444.
e84. Mehndiratta A, Treaba C.A, Barletta V et al. Characterization of thalamic lesions and their correlates in multiple sclerosis by ultra-high-field MRI. *Mult Scler*. 2021 Apr;27(5):674-683
e85. Rocca M.A, Mesaros S, Pagani E, Sormani M.P, Comi G, Filippi M. Thalamic damage and long-term progression of disability in multiple sclerosis. *Radiology*. 2010; **257**(2): p. 463-9.
e86. Fuchs T.A, Carolus K, Benedict R.H.B et al. Impact of Focal White Matter Damage on Localized Subcortical Gray Matter Atrophy in Multiple Sclerosis: A 5-Year Study. *AJNR Am J Neuroradiol*. 2018;**39**(8): p. 1480-1486.

e87. Pongratz V, Schmidt P, Bussas M et al. Prognostic value of white matter lesion shrinking in early multiple sclerosis: An intuitive or naive notion? *Brain Behav*. 2019; **9**(12): p. e01417.
e88. Lee H, Nakamura K, Narayanan S et al. Impact of immunoablation and autologous hematopoietic stem cell transplantation on gray and white matter atrophy in multiple sclerosis. *Mult Scler*. 2018;**24**(8): p. 1055-1066.
e89. Tedeschi G, Dinacci D, Comerci M et al. Brain atrophy evolution and lesion load accrual in multiple sclerosis: a 2-year follow-up study. *Mult Scler*. 2009;**15**(2): p. 204-11.
e90. Fox J, Kraemer M, Schormann T et al. Individual Assessment of Brain Tissue Changes in MS and the Effect of Focal Lesions on Short-Term Focal Atrophy Development in MS: A Voxel-Guided Morphometry Study. *Int J Mol Sci*. 2016;**17**(4): p. 489.