TITLE: Linear brain atrophy measures in multiple sclerosis and clinically isolated syndromes: A 30-year follow-up

Authors: Lukas Haider (PhD)[1,2], Karen K Chung (MD)[1], Giselle Birch (BSc)[1], Arman Eshaghi (PhD)[1,3], Stephanie Mangesius (PhD)[4,5], Ferran Prados (PhD)[1,3,6], Carmen Tur (PhD)[1], Olga Ciccarelli (PhD)[1,7], Frederik Barkhof (MD)[1,3,7,8], Declan T Chard (PhD)[1,7]

Affiliations:

1. NMR Research Unit, Queen Square Multiple Sclerosis Centre, University College London Institute of Neurology, UK
2. Department of Biomedical Imaging and Image Guided Therapy, Medical University of Vienna, AUT
3. Centre for Medical Image Computing (CMIC), Department of Medical Physics and Biomedical Engineering, University College London, UK
4. Department of Neuroradiology, Medical University of Innsbruck, AUT
5. Neuroimaging Core Facility, Medical University of Innsbruck, AUT
6. Universitat Oberta de Catalunya, Barcelona, Spain
7. National Institute for Health Research (NIHR) University College London Hospitals (UCLH) Biomedical Research Centre, UK
8. Department of Radiology and Nuclear Medicine, VU University Medical Centre, Amsterdam, NL

Corresponding Author: Karen K Chung

Email: k.chung@ucl.ac.uk

Address: NMR Research Unit, Queen Square Multiple Sclerosis Centre, University College London Institute of Neurology, 10-12 Russell Square, London WC1B 5EH, United Kingdom

Manuscript word count: 4140

Date of re-submission: 15th January 2021

ABSTRACT

Objective: To determine 30-year brain atrophy rates following clinically isolated syndromes and the relationship of atrophy in the first five years and clinical outcomes 25 years later.

Methods: A cohort of 132 people who presented with a clinically isolated syndrome suggestive of multiple sclerosis (MS) were recruited between 1984-87. Clinical and MRI data were collected prospectively over 30 years. Widths of the third ventricle and the medulla oblongata were used as linear atrophy measures.

Results: At 30 years, 27 participants remained classified as having had a clinically isolated syndrome, 34 converted to relapsing remitting MS, 26 to secondary progressive MS, and 16 had died due to MS. The mean age at baseline was 31.7 years (SD 7.5) and the mean disease duration was 30.8 years (SD 0.9). Change in medullary and third ventricular width within the first five years, allowing for white matter lesion accrual and Expanded Disability Status Scale increases over the same period, predicted clinical outcome measures at 30 years. 1 mm of medullary atrophy within the first five years increased the risk for secondary progressive multiple sclerosis or multiple sclerosis related death by 30 years by 583% (OR: 5.83, 95% CI: 1.74 – 19.61, $p<0.005$), using logistic regression.

Conclusions: Our findings show that brain regional atrophy within 5 years of a clinically isolated syndrome predicts progressive multiple sclerosis or a related death, and disability 25 years later.

INTRODUCTION

Multiple sclerosis (MS) is a chronic inflammatory, demyelinating and neurodegenerative disease of the central nervous system [1]. While about 85% of people with MS initially have relapsing-remitting (RR) MS, associated with the accrual of inflammatory demyelinating lesions, in the longer-term the majority develop secondary progressive (SP) MS [2]. Neurodegeneration is considered the main cause of irreversible disability [3] and the main correlate of brain atrophy measured *in vivo* by MRI [4]. While brain atrophy occurs very early in the course of MS,[5,6] and appears to be more rapid in people with SPMS compared with RRMS,[7] it remains unknown when it becomes clinically relevant. In particular, it is unclear if early brain atrophy is significantly linked with long-term clinical outcomes and could accordingly represent an early treatment target.

We recently completed a 30-year longitudinal, clinical and MRI follow-up study of a cohort recruited soon after a clinically isolated syndrome (CIS) suggestive of MS. In the 120 people with a known outcome (of the 132 originally recruited), a third remained classified as having had a CIS, a third developed SPMS or MS contributed to their death, while the rest had RRMS, of whom ~90% were able to walk without limitation due to MS [8]. Early accrual of white matter lesions (WML), particularly infratentorial and deep white matter lesions within the first year, and disability within the first five years, were significantly associated with the development of SPMS or death due to MS by 30 years.

In the present study we investigated the relationship between regional brain atrophy and 30-year clinical outcomes. MRI scans from early time-points could not be analysed using current volumetric atrophy measurement methods, we therefore assessed rates of brain atrophy over 30 years using two linear atrophy measures.

MATERIAL AND METHODS

Participants

The clinical characteristics of this cohort have been described in detail previously [8]. Briefly, 132 people with a CIS suggestive of MS were prospectively recruited between 1984 and 1987, and underwent clinical assessment and MRI brain scan at baseline, one, five, ten, 14, 20 and 30 years. In total, clinical outcomes at 30 years were known in 120 individuals. Deceased participants who had not developed MS (10), or who had MS but died of unrelated or unknown causes (3), were excluded (Table 1), the former because we could not be sure that those who died without a diagnosis of MS would not have developed MS by 30 years, and the latter because we could not predict the eventual effects MS would have had on them. One person (who remained classified as having a CIS at 30-years) had idiopathic Parkinson's disease, otherwise no other participants had a known neurodegenerative condition. Of the remaining 107, all had at least one MRI follow-up scan, however in four participants (one with RRMS, the rest CIS) linear brain atrophy measures were not possible due to missing or poor image quality at baseline, one- or five-year follow-up, and were therefore excluded. Of the remaining 103 participants, 27 (26%) were classified at 30 years as having had a CIS (f:m = 18:9, age: 31.7 (8)), 34 (33%) developed RRMS (f:m = 22:12, age: 29.7 (6.2)) using the 2010 McDonald criteria,[9] 26 (25%) progressed to SPMS (f:m = 17:9, age: 32.1 (7.1)), and 16 (16%) had an MS related death (MSRD), (f:m = 10:6, age:35.7 (8.6)). As recruitment predated the disease modifying treatment era, the cohort was largely untreated (14% have received a disease modifying therapy (DMT) at any point)[8].

This study was approved by our institutional ethics committee and the National Research Ethics Service (15/LO/0650). Participants gave informed consent, written if they attended in person, or verbal if they provided information by telephone only. For the deceased members of the cohort, death certificates were obtained where possible.

Table 1: Clinical demographics of members of the cohort with a known 30-year outcome

30-year Classification + MRI		**CIS**	**RRMS**	**SPMS**	**MS related death**	**All**
Number		27	34	26	16	103*
Age at CIS [years]		31.7 (8)	29.7 (6.2)	32.1 (7.1)	35.7 (8.6)	31.7 (7.5)
Years CIS to RRMS		NA	6.3 (6.4)	4.5 (5)	2.8 (2.2)	5 (5.4)
Years CIS to SPMS		NA	NA	16.8 (6.6)	12 (4.6)	15.3 (6.4)
Years CIS to death due to MS		NA	NA	NA	21.9 (6.6)	21.9 (6.6)
Age at death [years]		NA	NA	NA	57.6 (11.3)	NA
CIS type	Cord (n)	13	17	16	10	56
	Optic Neuritis (n)	12	12	8	2	34
	Brainstem (n)	2	5	2	4	13
Gender	Female (n)	18	22	17	10	67
	Male (n)	9	12	9	6	36
EDSS at 30 years		1 (0 - 2)	1,5 (1 - 2)	6 (6 - 6,5)	NA	3 (1.5-7)

Legend: Data are provided as counts (n), mean (standard deviation) or as median (25%-75% range) as applicable. *In the total cohort, as published by Chung et al., clinical outcomes at 30 years and at least one MRI at one follow-up were available in 107 individuals, less 4 whose MRIs could not be processed for linear atrophy measures in the present study.

Abbreviations: CIS= Clinically Isolated Syndrome; RRMS= Relapsing Remitting Multiple Sclerosis, SPMS= Secondary Progressive Multiple Sclerosis

Clinical assessment

Expanded Disability Status Scale (EDSS) scores were available from baseline (by retrospective review of notes), five, 10, 14 and 20 years. At 30 years EDSS was obtained by clinical examination or by telephone [10]. In those who attended in person (62 of the 87 living participants), the following data were also obtained: timed 25-foot walk (T25FW), 9-hole peg test (9HPT), paced auditory serial addition test (PASAT), which assesses information processing speed, and brief international cognitive assessment for MS (BICAMS) scores,[11,12] which includes the brief visuospatial memory test (BVMTR), the symbol digit modalities test (SDMT), and the California verbal learning test (CVLT). BICAMS scores were obtained from 60 people, and BICAMS z-scores (adjusted for age, sex and years of education), were available in 31 participants who were ≤65 years of age (standardized scores are currently not available for people older than 65 years)[12].

Image acquisition

A Picker 0.5T system (Marconi Medical Systems, Cleveland, OH) was used at baseline, 1, and 5 years. At 10, 14, and 20 years, a 1.5T General Electric Signa (GE Healthcare, Chicago, IL) was applied and at 30 years a 3T Philips Achieva (Philips Healthcare, Netherlands). Proton-density and/or T2-weighted scans were obtained at all time points. Examples of image quality over time are provided in Figure 1 and main imaging parameters are summarised in Supplementary Table 1.

At baseline, one- and five-year follow-up, film prints were available and were re-digitised using a Vidar Diagnostic Pro Advantage film digitiser. At 10, 14, 20 and 30 years native digital copies were available.

Image analysis

Current volumetric MRI analysis methods could not be applied to baseline, one, five and 10 year data, so linear measures were used.

Third ventricular width (TVW) is an established technique,[13] and correlates with long-term disability progression in MS [14]. TVW was measured as the width of the third ventricle at the midpoint of a line running parallel to the long axis of the ventricle on axially acquired PD/T2-weighted MR scans at baseline, one, five, 10, 14 and 20 years, and T2 weighted images at 30 years.

We developed a second linear measure, medullary width (MEDW), in light of previous work showing that brainstem measures are sensitive to atrophy in MS [15] and may act as a surrogate for cervical spinal cord volume. MEDW was measured as the dorso-ventral diameter of the medulla on a mid-sagittal image normal to the craniocaudal cord orientation. The level of medullary measurement was determined by the craniocaudal pontine length mirrored caudally from the inferior pontine notch (Supplementary Figure 1). Medullary width (MEDW) measurements were performed on sagittal scout images (originally made to determine the examination range) at baseline, one- and five-year follow-up. At 10 years, sagittal reconstructions of axially acquired scans were used. Volumetric T1-weighted, cervical spinal cord MRIs were available at 14 and 20 years, and volumetric T1-weighted brain images with cervical spinal cord coverage at 30 years (Supplementary Table 1).

All measurements were performed blind to clinical status by at least two trained assessors: MEDW by LH and GB, and TVW by LH, KC and GB. FIJI (ImageJ) [16] was used by LH and GB, Jim (version 6.0) (Xinapse Systems, http://www.xinapse.com) by KC. Inter- and intra-rater statistics are provided in Supplementary tables 2 and 3, and Supplementary Figure 2.

Lesion counts were available from our earlier analysis of the MRI data [8].

Statistical analysis

Statistical analyses were undertaken with R studio [17] and Stata [18]. P-values less than 0.05 (two-tailed) were considered statistically significant. As total intracranial volume (TIV) was only available for those with MRI follow-up at 30 years, which would potentially bias towards a more favourable disease course, we did not correct for TIV. Stepwise selection models were used for model selection to predict 30-year clinical outcome scores and disease phenotype based on atrophy within the first five years. Measures that significantly improve the model fit, as measured by AIC, do not necessarily exceed a conventional level of statistical significance ($p<0.05$). We report all measure that contribute to the final models, along with their respective individual p-values. Beta correlation coefficients are provided with 95% CI and p-values, model fits with AIC and R^2 and the Hosmer-Lemeshow 'goodness of fit' for logistic regression. We did not adjust for multiple comparison.

Rates of change

Nested mixed effect models were built in Stata and R to calculate annualized atrophy rates and the differences between CIS, RRMS, SPMS, MSRD outcomes at 30 years. Subject, as a random effect, was nested in time, thus accounting for differences in intervals, and the models were adjusted for age and gender.

Early predictors of 30-year clinical scores

Linear regression models corrected for potential age and gender effects were computed to predict clinical outcome measures at 30-years based on change in disability (EDSS), WML and linear atrophy measures between baseline and five-years. As models predicting BICAM z-scores are already adjusted for age, gender and years of education, age and gender were not additionally included in this comparison. To facilitate the comparison of the regression coefficients, z-scores were computed for each of the MRI measures. For an individual to be included in the prediction of a given clinical outcome score, all three MRI measures, TVW, MEDW and WML, need to be available (no missing values).

Five-year predictors of 30-year clinical scores

Logistic regression models were built to predict the 30-year outcome: CIS or RRMS compared with SPMS or MSRD based on disability (EDSS), atrophy and WML accrual within the first five years, correcting for age and gender. Intra- and inter-rater correlations are provided in the supplementary materials. For an individual to be included in the model seeking to predict clinical outcome scores, a clinical outcome score at 30-years and all three baseline to 5-year MRI measures (TVW, MEDW and WML) were required.

RESULTS

Rate of change

Figure 2 shows the 30-year time course of EDSS, MEDW and TVW of each participant, grouped by their respective 30-year outcome (CIS, RRMS, SPMS and MS related death = MSRD).

Over 30 years, MEDW decreased in the CIS group at 0.02 mm/year (95% CI: 0.003 - 0.028 mm/year, p< 0.019), in RRMS 0.02 mm/year (95% CI: 0.013 – 0.034 mm/year, p<0.001), in SPMS 0.03 mm/year (95% CI: 0.019 – 0.044 mm/year, p< 0.001) and in those with MSRD during follow-up 0.09 mm/year (95% CI: 0.067 - 0.122 mm/year, p< 0.001). The differences in the rate of change of the MEDW over time was not significant comparing CIS and RRMS (p<0.353) and CIS and SPMS (p= 0.079) but was significant between CIS and MSRD (p< 0.001) (Figure 2B, Supplementary Table 4). Over the first 5 years following a CIS, the difference in the rate of change of the MEDW over time was not significant comparing CIS and RRMS (p<0.0605), but reached the level of statistical significance for the comparison of CIS and SPMS (p= 0.0004) and CIS against MSRD (p= 0.0275) (see Supplementary Table 5 for details).

Over 30 years, TVW increased in CIS at 0.09 mm/year (95% CI: 0.056 – 0.129 mm/year , p<0.001), in RRMS 0.11 mm/year (95% CI: 0.075 – 0.137 mm/year , p<0.001), in SPMS 0.19 mm/year (95% CI: 0.154 – 0.230 mm/year , p<0.001) and in MSRD at 0.31 mm/year (95%CI: 0.245 – 0.380mm/year , p<0.001) (Figure 2C, Supplementary Table 4). The rate change of the TVW within the first five years was not significantly predictive for CIS or RRMS vs SPMS or MSRD, as was gender and the EDSS accrual within the first five years. The rate of change of the TVW over time was significant between CIS and SPMS (p<0.001) and CIS and MSRD (p<0.001). Over the first 5 years following a CIS, the differences in the rate of change of the TVW did not reach the level of statistical significance (see Supplementary Table 5 for details).

A brainstem CIS occurred in 27 participants and a cord syndrome in 36. Neither brainstem CIS (beta= -0.11; 95%CI: -0.31 – 0.08; p=0.242) nor cord CIS (beta: -0.12; 95%CI: -0.27 -

0.03; p=0.1235) were significantly associated with 30-year MEDW atrophy rates. Similarly no effect of the CIS type was measurable on the 5 year MEDW atrophy rates (data not shown). All models (rate changes and differences between rate changes for MEDW, TVW) were adjusted for age and gender, neither of which were statistically significant.

Five-year predictors of 30-year clinical scores

In combined models, MEDW change in the first five years significantly contributed to the prediction of all clinical scores, except PASAT, at 30 years, and TVW to EDSS, T25FWT and 9HPT (Table 2). WML lesion count changes, in these combined models, did not significantly improve prediction.

EDSS change over the first 5 years in these models was predictive for EDSS and PASAT at 30 years, but was not predictive of 30-year outcomes as measured by T25FWT, 9HPT, SDMT-z and BVMTR-z (Table 2).

Factors with contribution to the overall model fit (AIC) but p-values above 0.05 are reported in light grey in Table 2.

Table 2: 30-year clinical outcome prediction based on initial five-year changes

			Change within *first five years* after CIS in:					
			EDSS	**TVW**	**MEDW**	**WML**	**Gender (f=0, m=1)**	**Age**
prediction of *30-year* outcome as measured by:	**EDSS**	Estimate (β)	0.30	0.20	0.60	-	-	-
		(95%CI)	(0.108 - 0.501)	(0.013 - 0.384)	(0.372 - 0.825)	-	-	-
		p-value=	0.004	0.042	0.0001	-	-	-
		Overall	R2: 0.60; AIC.:98; p-value< 0.0001					
	T25FWT	Estimate (β)	-	0.62	0.57	-	-	-
		(95%CI)	-	(0.157 - 1.082)	(0.122 - 1.021)	-	-	-
		p-value=	-	0.015	0.020	-	-	-
		Overall	R2: 0.33; AIC.:84; p-value= 0.0108					
	9HPT	Estimate (β)	-	0.61	0.58	-	0.46	-
		(95%CI)	-	(0.207 - 1.013)	(0.186 - 0.977)	-	(0.081 - 0.841)	-
		p-value=	-	0.007	0.009	-	0.026	-
		Overall	R2: 0.48; AIC.:77; p-value= 0.002					
	SDMT-z	Estimate (β)	-	-	-0.46	-0.90	-	-
		(95%CI)	-	-	(-0.889 - -0.037)	(-1.852 - 0.05)	-	-
		p-value=	-	-	0.048	0.081	-	-
		Overall	R2: 0.37; AIC.:57; p-value= 0.019					
	CVLT-z	Estimate (β)	0.44	-	-0.49	-	-	-
		(95%CI)	(-0.094 - 0.968)	-	(-0.942 - -0.041)	-	-	-
		p-value=	0.125	-	0.047	-	-	-
		Overall	R2: 0.29; AIC.:60; p-value= 0.056					
	BVMTR-z	Estimate (β)	-	-	-0.70	-	-	-
		(95%CI)	-	-	(-1.159 - -0.239)	-	-	-
		p-value=	-	-	0.008	-	-	-
		Overall	R2: 0.34; AIC.:55; p-value= 0.008					
	PASAT	Estimate (β)	-0.62	-	-	-	-	-
		(95%CI)	(-1.061 - -0.185)	-	-	-	-	-
		p-value=	0.010	-	-	-	-	-
		Overall	R2: 0.25; AIC.:68; p-value= 0.010					

Legend: Changes in EDSS, TVW, MEDW and WML, within the first five years, allowing for age and gender, were used as predictors of 30-year outcome measures. The number of

available participants for the prediction of the outcome scores were: EDSS n=43; T25FWT and 9HPT n=26; SDMT-z, CVLT-z and BVMTR-z n=20; and PASAT n=25.

Factors contributing significantly to the overall model fit (as measured by Akaike Information Criterion [AIC]) need not individually exceed the conventional threshold ($p<0.05$) for statistical significance, and these are reported in light gray along with their individual p-values. Factors that did not significantly increase the AIC were not retained in the model, and so coefficients are not available for them.

Abbrev.: AIC= Akaike information criterion; EDSS= Expanded Disability Status Scale; TVW= Third Ventricular Width; MEDW= Medullary Width; T25FW= Timed 25-foot Walk; 9HPT= 9-hole Peg Test; BVMTR-z= Brief Visuospatial Memory Test; CVLT-z= California Verbal Learning Test; SDMT-z= Symbol Digit Modalities Test; PASAT= Paced Auditory Serial Addition Test.

Predictors of 30-year disease phenotype after 5 years

For each new WML that occurred within the first five years the risk for SPMS or MSRD compared to CIS or RRMS increased by 7% (OR: 1.07 95% CI: 1.01-1.14, p= 0.027) and for each mm reduction in MEDW by 583% (OR: 5.43, 95% CI: 1.74 – 19.61, $p<0.005$). In the first five years individuals with SPMS or MSRD outcome developed on average 32 WML (sd 42) and those who remained CIS or RRMS, 6 (sd 11). The medullary width reduced on average by 1.3 mm (sd 0.6) in individuals with SPMS or MSRD compared to 0.3 mm (sd 0.8 mm) in those who remained CIS or RRMS (Supplementary Figure 5).

Age at CIS did not reach the threshold for statistical significance (OR: 1.10, 95% CI: 0.98 – 1.23, p= 0.100). The rate change of TVW and EDSS within the first five years, as well as gender, were not retained in the model. Hosmer-Lemeshow goodness of fit testing found no evidence for poor model fitting (p=0.47).

Juxtacortical, deep white matter, periventricular and infratentorial white matter lesion counts did not outperform global white matter lesion counts in the predictive models (data not shown).

DISCUSSION

In this study, which to the best of our knowledge is the first to assess brain atrophy in MS over 30 years, we found that atrophy was faster over the duration of the study in those who developed SPMS (on average 16.8 years after CIS) or who died due to MS (on average 21.9 years after CIS) when compared with those who remained classified as CIS (Figure 2, Tables 1 and 2). Brain atrophy within 5 years of symptom onset significantly predicted neurological and cognitive function at 30-year follow-up.

Previous studies have shown that brain atrophy occurs early in MS but is most prominent in those with clinically progressive disease [19]. In the absence of a control cohort we could not determine how much faster brain atrophy was relative to normal ageing. However, relative to those who remained classified as having a CIS we did observe more rapid brain atrophy in people with SPMS, and within 5 years of their CIS in those with RRMS who eventually (a decade or more later) developed SPMS. Our data highlight a potentially long window of opportunity to alter the rate of brain atrophy before it clinically manifests, and conversely that treatments designed to slow or prevent progressive disease should be considered far in advance of its clinical onset. However, it is also worth recalling that a significant proportion of disability accrual remains unexplained by atrophy measures, even in the long-term, reminding us that other processes also have substantial and enduring clinical effects.

MS was diagnosed using the 2010 version of the McDonald diagnostic criteria [20]. Given this, it is possible that some of those classified as having a CIS in the present cohort could now have been diagnosed with RRMS. However, we have not observed a significant difference in regional atrophy between the CIS and RRMS groups over 30 years and, importantly, this will have made no difference to the classification of people with SPMS or those for whom MS contributed to their death, as both were based on clinical features. Overall, use of the updated criteria should not materially alter apparent differences with the more severely clinically affected groups. The occurrence of a cord or brainstem CIS was not significantly associated with subsequent MEDW atrophy rates over five- and over 30-years, suggesting that local inflammatory demyelination has a modest effect on local atrophy in the

brainstem. Previous studies have similarly found a modest association of cord atrophy and cord lesions, while other processes, such as brain lesion formation and brain atrophy, have been linked with cord atrophy,[7] and so this observation is not entirely surprising.

In our previous clinical report on the present cohort, [8] we found that despite accruing lesions and having relapses, those with RRMS at 30-year follow-up were indistinguishable from those who remain classified as CIS in terms of EDSS scores. In the present study we also found no significant differences in brain atrophy between those who remain classified as CIS and those who developed RRMS (Figure 2), albeit with the caveat that linear measures of brain atrophy may have been insensitive to some MS-related brain atrophy that could have been detected using current volumetric methods. This suggests that for some people with a clinically less aggressive disease course, clinically significant brain atrophy is not inevitable.

In our earlier study WM lesions at baseline and one year predicted outcomes at 30 years [8]. In the present study, in multivariable models over five years, brain atrophy was the stronger predictor for clinical outcome (Table 2), except for PASAT. This does not mean that WM lesions are irrelevant to longer-term outcomes. First, while trends towards differences in atrophy between 30-year outcome groups were observed at one year, they only reached significance at five years (see supplemental materials), while lesion location and accrual did so over the first year [8]. Second, we do not know the main cause of brain atrophy (and the neurodegeneration underlying it), and while cross-sectionally, neither WM nor GM lesions have been strongly associated with atrophy,[21] correlations between lesion accrual and brain atrophy appear to increase with the passage of time,[22] suggesting that it takes time for associations between new lesions and eventual atrophy to manifest. As such, lesion accrual and atrophy may both need to be treated early on.

With due caution when directly comparing the apparent TVW and MEDW regional atrophy trajectories (see limitations noted below), while at a group level there appears to be a consistent pattern of more rapid atrophy in the MS groups with increasingly disabling outcomes, the TVW and MEDW do not entirely reflect each other. In particular, while in the

models of early atrophy (within 5 years of clinical onset) and 30-year clinical outcomes, MEDW (potentially indicative for cervical spinal cord volume [15]) emerged as being more significant than TVW, over the entire 30 years, progressive differences in TVW were more apparent (Figure 2).

In part this may relate to measurement sensitivity, particularly at the earlier time-points, as linear measures rely on the accurate identification of landmarks, which would be influenced by factors such as MRI slice thickness, image contrast (Figure 1, Supplementary Table 1), and head positioning. Given this, apparent differences in the prognostic relevance of MEDW and TVW measures may actually be due to differences in measurement noise rather than a true predilection for atrophy to affect the medulla over the supratentorial compartment. However, it also reminds us that brain atrophy in MS is not uniform,[23] being faster in some regions than others. Practically, while volumetric measurements would be preferred over the linear measurements, whole brain volumetric measurements may be less sensitive to clinically significant regional atrophy, and if the present results are to be translated into clinical trials or practice, this should be considered. The new method of medullary atrophy measurement was developed to serve as additional measure of atrophy, so increasing confidence in the results overall if they were broadly consistent with the third ventricular width findings. We do not propose that our new method be used outside of research, and only then when more advanced whole brain or regional volumetric methods cannot be applied.

This study has several strengths, in particular participants were recruited after a CIS rather than after they developed MS; participant recruitment pre-dates the introduction of MS DMTs (and therefore provides an essentially natural history view on the clinical and radiological evolution of CIS and MS). In the present study MRI and clinical data were available over a uniquely long period, however MRI scanning has advanced substantially over the past 30 years and this is likely to have effected linear atrophy measurements. A Picker 0.5T system was used at baseline, 1 and 5 years, 1.5T General Electric Signa at 10, 14, and 20 years and 3T Philips Achieva at 30 years, and caution should be exercised comparing rates of

change calculated using each of these scanners. Further, changing scanners is likely to be a contributing factor to observed variability in individual participants' linear atrophy trajectories (Figure 2), and estimated rates of atrophy over 30 years will have been affected by this. However, despite this, our data suggests that on a group basis, changes in scanner still did not entirely obscure biological effects over a 30-year period. Current methods used to measure whole brain atrophy could not be applied to the early scans. In view of this, we used linear measures of the brain, which could be similarly applied to the archival films and modern digital MRI scans. While these correlate with brain volumetric measures (see supplemental materials), linear measures are less sensitive to atrophy than volumetric techniques [24,25]. Although this will not have led to spurious group differences being detected, it does mean that atrophy rates cannot be directly extrapolated to volumetric equivalents. Therefore, the present results cannot be used to determine prognostically significant atrophy thresholds for use with up-to-date volumetric atrophy methods. For longitudinal studies using volumetric brain atrophy measurement techniques to allow for differences in the scaling of MRI scans over time, measures are often adjusted for TIV (which is not thought to significantly change with normal adult ageing). Using the archival films it was not possible to measure intra-cranial volumes or determine bony landmarks that could be tried as a linear proxy. Again, while this circumstance will not have led to false atrophy being detected, it is likely to have reduced sensitivity. Furthermore, this study lacks a control group, and while brain atrophy rates in people with CIS and age-related healthy controls have been found to be comparable,[23] we cannot be certain that this holds true in the long-term or for the linear measures we have used. As our main conclusion is that following a CIS more rapid early brain atrophy is associated with a greater risk of disability and death in the long-term, this does not fundamentally undermine our findings. However, it does mean that we cannot determine whether or not rates of atrophy in those who had had a CIS were normal, or establish a threshold relative to normal ageing beyond which atrophy in MS is associated with poorer outcomes. Hopefully future studies will be able to address these questions, although the now widespread and early use of DMTs means that analyses

are likely to confounded by treatment effects. It should also be kept in mind that the cohort is small compared with other CIS or MS studies [26,27] and this is likely to have limited sensitivity to sub-group differences and associations with outcomes. Again, we believe that this does not negate our findings, but we may have overlooked more subtle group differences and associations with outcomes. Brain tissue volumes have been measured cross-sectionally in the present cohort at 20 years [28], and significantly smaller grey matter volumes reported in SPMS vs. RRMS, and in RRMS vs. CIS. However, from this we could not determine when accelerated atrophy was first apparent in the RRMS and CIS groups.

The present study demonstrates that accelerated atrophy with 5 years of symptom onset is associated with a progressive MS course and a broad spectrum of clinical outcome measures. Lastly, studies have shown associations between brain atrophy and subsequent disability [19], although to the best of our knowledge the longest follow-up prior to this study has been 15 years and predicted EDSS 3.0, rather than the development of progressive MS, or death due to MS [29, 30]. We are aware of no previous studies that have assessed associations between early brain atrophy and death due to MS.

It is now well recognised that co-morbidities such as vascular disease, diabetes, and smoking, are associated with more disabling outcomes in people with MS [31]. However, this awareness only developed after the inception of this cohort, and so was not systematically obtained prior to the 30-year follow-up.

In conclusion, we have shown that brain atrophy early in the course of MS independently predicts progressive disease and disability, as measured by a range of clinical and cognitive tests after 30 years. Our findings support the concept of early atrophy [32] as an important target for treatment.

ACKNOWLEDGEMENTS

We are indebted to Professor David Miller (MD, University College London) and the late Professor W Ian McDonald (PhD) for the inception and follow-up of this cohort. We thank all who participated in this study and all involved in data collection over the years. We thank Marios Yiannakas (PhD, University College London) for undertaking the MRI scanning at 30-years and Drs Jonathan O'Riordan (MD, Tayside Multiple Sclerosis Research Unit), Peter A Brex (MD, King's College Hospital National Health Service Foundation Trust) and Leonora Fisniku (PhD, Brighton and Sussex University Hospital), who undertook earlier follow-up of this cohort. No compensation was received by anyone named in this section, and written permission has been obtained (with the exception of Prof McDonald) from all persons named in this section.

COMPETING INTERESTS

L Haider has nothing to declare.

K Chung has received honoraria for speaking at meetings, advisory work or support to attend meetings from Biogen, Sanofi-Genzyme and Roche

G Birch, A Eshaghi, S Mangesius, F Prados and C Tur have nothing to declare.

O Ciccarelli serves as a consultant for Novartis, Roche, Teva and Merck, and receives personal fees from *Neurology* and *Multiple Sclerosis Journal*. Outside the submitted work, she has received research grants from Spinal Cord Research Foundation, Rosetrees trust, Progressive MS Alliance, Bioclinica & GE Neuro, and EU-H2020.

F Barkhof serves as a consultant for Bayer Schering Pharma, Sanofi-Genzyme, Biogen Idec, Teva, Merck Serono, Novartis, Roche, IXICO, GeNeuro, Apitope Ltd. and Jansen Research.

D Chard is a consultant for Biogen and Roche.

FUNDING

This study was funded by the MS Society of Great Britain and Northern Ireland [20; 984].

Lukas Haider was supported by the European Society of Neuroradiology (2018) and an ECTRIMS/MAGNIMS research fellowship (2019).

Stephanie Mangesius gratefully acknowledges research funding from ÖAW (Österreichische Akademie der Wis-senschaften; GDNG_2018-041_MedCorpInn) and the Jubliäumsfonds der Universität Innsbruck und der Medizinischen Universität Innsbruck (2019) outside of the submitted work.

Ferran Prados has received a non-clinical Guarantors of Brain fellowship.

Olga Ciccarelli has received research funding from: UK and National MS Societies, and the National Institute for Health Research (NIHR) University College London Hospitals (UCLH) Biomedical Research Centre (BRC).

Frederik Barkhof is supported by the NIHR UCL BRC.

Declan Chard has received research funding from the International Progressive MS Alliance, the MS Society, and the National Institute for Health Research (NIHR) University College London Hospitals (UCLH) Biomedical Research Centre.

AUTHORS' CONTRIBUTIONS

L Haider had full access to all the data and was responsible for data and statistical analysis, interpretation of the results, prepared the first draft of the manuscript (together with KC and DC), and edited subsequent versions.

K Chung was responsible for clinical and MRI data collection at 30 years, and collecting and archiving clinical and MRI data from previous time-points. She contributed to image analyses, prepared the first draft of the manuscript (together with LH and DC), and edited subsequent versions.

G Birch contributed to image analysis.

A Eshaghi participated and supervised in statistical analysis and interpretation.

S Mangesius contributed to image analysis.

F Prados participated and supervised in image preparation, analysis and interpretation.

C Tur participated and supervised in statistical analysis and interpretation.

O Ciccarelli reviewed the manuscript and contributed to results interpretation.

F Barkhof contributed to study supervision, supervised in image analysis, reviewed the manuscript and contributed to results interpretation.

D Chard contributed to the study design, securing funding, and study supervision. He had full access to the data and was responsible for data and statistical analysis, interpretation of the results, prepared the first draft of the manuscript (together with KC and LH), and edited subsequent versions.

All co-authors have reviewed and approved the submission of this manuscript.

REFERENCES

1. Lassmann H. The changing concepts in the neuropathology of acquired demyelinating central nervous system disorders. Curr Opin Neurol. 2019;32(3):313–319.

2. Filippi M, Preziosa P, Langdon D, et al.; Identifying Progression In Multiple Sclerosis: New Perspectives; Ann Neurol. 2020, doi: 10.1002/ana.25808.

3. Trapp BD, Peterson J, Ransohoff RM, Rudick R, Mörk S, Bö L. Axonal transection in the lesions of multiple sclerosis. N Engl J Med. 1998;338(5):278–285.

4. Popescu V, Klaver R, Voorn P, Galis-de Graaf Y, Knol DL, Twisk JW, et al., What drives MRI-measured cortical atrophy in multiple sclerosis? Mult Scler. 2015 Sep;21(10):1280-90.

5. Brex, P. A., Jenkins, R., Fox, N. C., Crum, W. R., O'Riordan, J. I., Plant, G. T., et al., (2000). Detection of ventricular enlargement in patients at the earliest clinical stage of MS. Neurology, 54(8), 1689–1691.

6. Dalton CM, Brex PA, Jenkins R, Fox NC, Miszkiel KA, Crum WR, et al., Progressive ventricular enlargement in patients with clinically isolated syndromes is associated with the early development of multiple sclerosis. J Neurol Neurosurg Psychiatry. 2002 Aug;73(2):141-7.

7. Filippi M, Brück W, Chard D, Fazekas F, Geurts JJG, Enzinger C, et al., Lancet Neurol. 2019 Feb;18(2):198-210.

8. Chung KK, Altmann D, Barkhof F, Miszkiel K, Brex PA, O'Riordan J, et al., A 30-Year Clinical and Magnetic Resonance Imaging Observational Study of Multiple Sclerosis and Clinically Isolated Syndromes. Ann Neurol. 2020 Jan;87(1):63-74.

9. Polman CH, Reingold SC, Banwell B, Clanet M, Cohen JA, Filippi M, et al., Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Ann Neurol. 2011 Feb;69(2):292-302.

10. Lechner-Scott J, Kappos L, Hofman M, Polman CH, Ronner H, Montalban X, et al., Can the Expanded Disability Status Scale be assessed by telephone? Mult Scler. 2003 Mar;9(2):154-9.

11. Fischer JS, Rudick RA, Cutter GR, Reingold SC. The Multiple Sclerosis Functional Composite Measure (MSFC): an integrated approach to MS clinical outcome assessment. National MS Society Clinical Outcomes Assessment Task Force. Mult Scler. 1999;5(4):244–250.

12. Langdon DW, Amato MP, Boringa J, Brochet B, Foley F, Fredrikson S, et al., Recommendations for a Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS). Mult Scler. 2012 Jun;18(6):891-8.

13. Benedict RH, Weinstock-Guttman B, Fishman I, Sharma J, Tjoa CW, Bakshi R. Prediction of neuropsychological impairment in multiple sclerosis: comparison of conventional magnetic resonance imaging measures of atrophy and lesion burden. Arch Neurol. 2004 Feb;61(2):226-30.

14. Pontillo G, Cocozza S, Di Stasi M1, et al.; 2D linear measures of ventricular enlargement may be relevant markers of brain atrophy and long-term disability progression in multiple sclerosis; Eur Radiol. 2020 Feb 25.

15. Chivers T, Constantinescu C, Tench CR; MRI-Based Measurement of Brain Stem Cross-Sectional Area in Relapsing-Remitting Multiple Sclerosis; J Neuroimaging. 2015 Nov-Dec;25(6):1002-6.

16. Schindelin J, Arganda-Carreras I, Frise E, et al., Fiji: An Open-Source Platform for Biological-Image Analysis; Nat Methods. 2012 Jun 28;9(7):676-82.

17. RStudio Team (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/.

18. StataCorp. 2019. Stata Statistical Software: Release 16. College Station, TX: StataCorp LLC.

19. Rocca MA, Battaglini M, Benedict RH, De Stefano N, Geurts JJ, Henry RG, et al., Brain MRI atrophy quantification in MS: From methods to clinical application. Neurology. 2017 Jan 24;88(4):403-413.

20. Thompson AJ, Banwell BL, Barkhof F, Carroll WM, Coetzee T, Comi G, et al., Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. Lancet Neurol. 2018 Feb;17(2):162-173.

21. Thompson AJ, Baranzini SE, Geurts J, Hemmer B, Ciccarelli O. Multiple sclerosis. Lancet. 2018;391(10130):1622–1636.

22. Chard DT, Brex PA, Ciccarelli O, Griffin CM, Parker GJ, Dalton C, et al., The longitudinal relation between brain lesion load and atrophy in multiple sclerosis: a 14 year follow up study. J Neurol Neurosurg Psychiatry. 2003 Nov;74(11):1551-4.

23. Eshaghi A, Marinescu RV, Young AL, Firth NC, Prados F, Jorge Cardoso M, et al., Progression of regional grey matter atrophy in multiple sclerosis. Brain. 2018 Jun;1;141(6):1665-1677.

24. Turner B, Ramli N, Blumhardt LD, Jaspan T. Ventricular enlargement in multiple sclerosis: a comparison of three-dimensional and linear MRI estimates. Neuroradiology. 2001 Aug;43(8):608-14.

25. Butzkueven H, Kolbe SC, Jolley DJ, Brown JY, Cook MJ, van der Mei IA, et.al., Validation of linear cerebral atrophy markers in multiple sclerosis. J Clin Neurosci. 2008 Feb;15(2):130-7.

26. Tintore M, Rovira À, Río J, Otero-Romero S, Arrambide G, Tur C, et al., Defining high, medium and low impact prognostic factors for developing multiple sclerosis. Brain. 2015 Jul;138(Pt 7):1863-74.

27. Cree BA, Gourraud PA, Oksenberg JR, Bevan C, Crabtree-Hartman E, Gelfand JM, et al., Long-term evolution of multiple sclerosis disability in the treatment era. Ann Neurol. 2016 Oct;80(4):499-510.

28. Fisniku L, Chard D, Jackson F, Anderson V, Altmann D, Miszkiel K., et al.; Gray matter atrophy is related to long-term disability in multiple sclerosis; Ann Neurol . 2008 Sep;64(3):247-54. doi: 10.1002/ana.21423.

29. Popescu V, Agosta F, Hulst HE, Sluimer IC, Knol DL, Sormani MP, et al.; Brain atrophy and lesion load predict long term disability in multiple sclerosis; J Neurol Neurosurg Psychiatry. 2013 Oct;84(10):1082-91.

30. Vidal-Jordana A, Sastre-Garriga J, Pareto D, Tur C, Arrambide G, Otero-Romero S, et al.; Brain atrophy 15 years after CIS: Baseline and follow-up clinico-radiological correlations; Mult Scler. 2018 May;24(6):721-727.

31. Geraldes R, Juryńczyk M, Dos Passos GR, Pichler A, Chung KK, Hagens M, et al., The role of pontine lesion location in differentiating multiple sclerosis from vascular risk factor-related small vessel disease**.** Mult Scler. 2020 Aug 6:1352458520943777. doi: 10.1177/1352458520943777. Online ahead of print.

32. Kawachi I, Lassmann H. Neurodegeneration in multiple sclerosis and neuromyelitis optica; J Neurol Neurosurg Psychiatry. 2017 Feb;88(2):137-45.

FIGURES

Figures have been uploaded as separate files as per editorial recommendation.

Figure 1: The evolution of atrophy over 30-years

Legend: The two top rows show axial images at the level of the third ventricle. Upper row: RRMS, EDSS at 30-years: 0 vs. lower row: SPMS, EDSS at 30-years: 6.5.

The two bottom rows show the corresponding sagittal images at the level of the medulla within the same two patients for each time point. The medullary images of both patients at baseline and one-year follow-up show the orientations of the axial images.

Note the high in-plane resolution of archival films (Baseline, one, and five years).

Abbrev.: EDSS= Expanded Disability Status Scale; RRMS= Relapsing Remitting Multiple Sclerosis; SPMS= Secondary Progressive Multiple Sclerosis; FU= Follow-up

Figure 2: Evolution of disability and atrophy over 30 years following a CIS

Legend:

A: expanded disability status scale changes (EDSS); B: medullary width (MEDW); and C: third ventricular width (TVW) measurements over 30 years. The participants are grouped based on their 30-year outcomes. For individual participants measurements are connected by black lines. The mean (for TVW and MEDW), as well as the median of EDSS, for each group, is shown in green for the CIS group, in blue for the relapsing remitting multiple sclerosis (RRMS) group, in orange in secondary progressive multiple sclerosis (SPMS) and in red for those with an multiple sclerosis related death (MSRD). A plot with the overlaid mean/median curves is provided at the outer right panel. The means and standard deviations for each group at each

timepoint (TVW and MEDW) as well as the medians and the 25-75% (EDSS) are available in the supplementary table 5.

Abbrev.: EDSS= Expanded Disability Status Scale; RRMS= Relapsing Remitting Multiple Sclerosis; SPMS= Secondary Progressive Multiple Sclerosis; MS= Multiple Sclerosis; FU= Follow-up