DR. ANNE CHRISTIN RAHN (Orcid ID : 0000-0002-9051-3621)

Article type : Review Article

Magnetic resonance imaging as a prognostic disability marker in clinically isolated syndrome: A systematic review

Rahn, Anne Christin[1*] PhD; Köpke, Sascha[2] PhD; Stellmann, Jan-Patrick[1] MD; Schiffmann, Insa[1] MD; Lukas, Carsten[3] MD; Chard, Declan[4,5] PhD FRCP, Heesen, Christoph[1,6] MD

[1]Institute of Neuroimmunology and Multiple Sclerosis, University Medical Center Hamburg-Eppendorf, Germany

[2]Nursing Research Unit, University of Lübeck, Germany

[3]Department of Radiology, St. Josef Hospital Bochum, Ruhr University, Germany.

[4]NMR Research Unit, Queen Square Multiple Sclerosis Centre, University College London (UCL) Institute of Neurology, UK

[5]National Institute for Health Research (NIHR) University College London Hospitals (UCLH) Biomedical Research Centre, UK

[6]Department of Neurology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/ane.13010

*Corresponding author:

Anne Christin Rahn

Institute of Neuroimmunology and Multiple Sclerosis

University Medical Center Hamburg-Eppendorf

Martinistraße 52

20246 Hamburg

Phone: +49(40) 7410-53215

Fax: +49(40) 7410 - 56973

Email: a.rahn@uke.de

Short title: MRI as a prognostic marker in CIS

Keywords

Clinically isolated syndrome, disability, multiple sclerosis, MRI, prognosis, systematic review

Abstract

Magnetic resonance imaging (MRI) is the key prognostic tool in people with a clinically isolated syndrome (CIS). There is increasing interest in treating people following a CIS in the hope that conversion to multiple sclerosis (MS) will be prevented and future disability reduced. So far, the prognostic value of MRI for disability following a CIS has not been evaluated systematically.

We systematically searched MEDLINE and EMBASE. Cohort studies were selected if they reported associations of MRI and disability following a CIS, included at least 50 people with a CIS at baseline, had at least 5 years of follow-up and obtained at least one structural MRI measurement (T1 lesions, T2 lesions, T1 contrast enhancing lesions or brain atrophy). We assessed the studies for quality and rated the completeness of MRI reporting.

In total, 13 studies were identified reporting on the following: T2 lesion number and volume, T2 infratentorial lesion number and volume, T1 contrast enhancing lesions and grey matter fraction. T2 brain lesion number determined soon after the occurrence of a CIS was associated with disability progression after 5 - 7 years, with an increased risk when 10 or more lesions were present. Infratentorial lesions were also associated with a higher risk of subsequent disability.

The number and distribution of MRI-visible lesions soon after a CIS is associated with disability later on, and may offer additional useful information when making treatment decisions in people with early MS. Further work is required to determine if other measures have a higher predictive potential.

Introduction

About 85% of people with multiple sclerosis (MS) present with a clinically isolated syndrome (CIS) commonly affecting optic nerves, the spinal cord, the cerebellum or the brainstem (1).

Magnetic resonance imaging (MRI) has an established role in the diagnosis of MS, estimating the risk of developing MS following a CIS (2), and assessing MS activity when making treatment decisions (3, 4). There is also evidence that early MRI features at least partly predict disability progression in the short (<2 - 3 years) and longer term (≥5 years) (5, 6).

Disability prognosis is of great importance to people with a CIS, helping them to make informed treatment decisions (7, 8). Recent work on patients' attitudes towards MRI has shown marked interest in the information provided by results from MRI scans, although the use of this information in MS management varies considerably at a national and international level (7). Beliefs, accessibility and costs seem to determine MRI usage more so than a clear evidence base.

There are literature reviews on MRI factors and their association with the risk of developing MS following a CIS (9, 10), however most do not fulfil systematic review criteria (11, 12). To the best of our knowledge, no systematic review has assessed the prognostic value of MRI for disability following a CIS. The present systematic review aimed to fill this gap.

Methods

To maximise the amount of information we could gather about outcomes following a CIS, and noting that studies on MS also include information about CIS (i.e. those who convert to MS); we performed a literature search to identify studies assessing the prognostic role of MRI in people who had had a CIS or MS. For completeness, the literature search was performed for all disease courses of MS, but only findings related to early relapse-onset MS (CIS) were included.

Criteria for considering studies for this systematic review

We included retrospective and prospective longitudinal studies that reported on associations of brain and/ or spinal MRI and disability following a CIS. As disability is usually low following recovery from a CIS and the predictive value of the short term disability evolution within the first years is low (13), only studies with a follow-up ≥5

years were included to assess disease progression. Placebo arms of randomised controlled MS long-term immunotherapy treatment studies were not considered, because the placebo arms usually switch to treatment after 2 years (14, 15).

Cohort size

Studies were eligible for inclusion when they included ≥50 people with a CIS at baseline. As we did not expect sample size calculations for the cohort studies, but rather convenient samples, we chose to only include studies with at least 50 people to receive meaningful results. While we included studies with at least 5 years of follow up, we expected a considerable loss to follow-up increasing the risk of substantial selection bias in very small cohorts. Cohorts only consisting of people under 18 years old were not included.

MRI prognostic factors

At least one of the following MRI measurements had to be assessed at baseline: number or volume of T1 lesions, number or volume of T2 lesions, T1 contrast enhancing lesions or brain volume. Where available, the relative prognostic value of regional lesion accrual was also considered, e.g. supra- and infratentorial location of lesions. Studies, which reported on new MRI techniques, such as spectroscopy, or prognostic models only were not included.

Primary outcome measures

The primary outcome was disability progression assessed by the expanded disability status scale (EDSS) (16) in association with the prognostic factor (MRI). Studies that assessed mean EDSS and time to reach EDSS 3 or 6 were included. Studies were only included in case they reported EDSS as an outcome.

Secondary outcome measures

We also assessed transition from a CIS to clinically definite MS (CDMS), the development of secondary progressive MS (SPMS) and mortality.

Search methods

Electronic searches were carried out in MEDLINE and EMBASE (via OVID). We performed a search on all disease courses and combined the term “multiple sclerosis” and variations with the term “MRI” and variations using the Boolean operator AND. Search filters on prognosis were applied (17, 18). The MEDLINE search strategy (supplemental material 1) was adapted for EMBASE. We included published studies until May 2013 (EMBASE) and accordingly September 2017 (MEDLINE). The full-text had to be available in English or German language. A hand-search of reference lists of the included studies was performed to detect further eligible studies. EndNote X5-X7 was used for bibliographic management including duplicate removal.

Data collection and analysis

To minimise selection bias, paired reviewers (JB/SK, CH/AR) independently screened the search findings (title/abstract). The reviewers were over-inclusive. Full-texts were examined independently by paired reviewers (JB/SK, CH/AR) according to the predefined inclusion criteria. Non-agreement of including or excluding a study was resolved by discussion.

Paired researchers (AMK/AR, CH/AR) independently extracted demographic, disease specific and outcome data. Results were compared by the researchers. All numbers of included studies assessing disability, CDMS or SPMS in relation to baseline MRI data were extracted. Additional results, e.g. from regression analyses, were considered.

Paired researchers (AMK/AR and CH/AR) independently assessed the quality of included studies. Different results were resolved by discussion.

We used the 'Quality In Prognosis Studies' (QUIPS) tool (19, 20) to assess the validity of the included studies. With this tool, markers on disease activity and outcomes (e.g. disability progression) can be assessed separately for potential bias. It comprises six domains: participation, attrition, prognostic factor measurement (in our case MRI measurements), outcome measurement, confounding and analysis. A judgment concerning the related risk of bias will be assigned for each domain (yes, partly, no risk). Further, an afore applied checklist was used to assess the quality and completeness of MRI reporting (21).

Heterogeneity in study designs and cohorts made it impossible to conduct robust meta-analyses. A scheme reaching from "strong evidence of effect" to "no evidence", proposed in a protocol for an exemplary prognostic review by the Cochrane group (22), was used to illustrate the consistency and the strength of the results for the primary endpoint (table 3).

Results

Results of the search

A total of 3708 titles and abstracts were screened for inclusion (figure 1). In that stage, 3639 studies were excluded. In total, 69 studies, which potentially met the inclusion criteria, were assessed in full text. Of those, 13 studies (5, 6, 23-33) fulfilled

the eligibility criteria. Further, 18 studies reported on the topic, but focussed on other disease courses, which are not subject of this publication.

Figure 1: PRISMA flow diagram (12)

Descriptions of included cohorts and publications

Marked heterogeneity was found in study duration ranging from 5 - 20 years and MRI assessments (suppl. material 2). Cohort study reports from 2 centres in London (2 cohorts) and Barcelona (1 cohort) and 3 publications on the long-term follow-up of a RCT on treatment of optic neuritis (ONTT) were included (see suppl. material 2 and tables 1 and 2 for details). Therefore, most studies report a follow-up study or reanalysis of the same patient cohort. For example, 6 studies (6, 23, 27, 28, 33, 34) reported on one of the London cohorts, which included participants from 1984-1987.

The Optic Neuritis Treatment Trial (ONTT) (29-31)

The randomised Optic Neuritis Treatment Trial (ONTT) was conducted in 15 centres in the United States to assess corticosteroids in the treatment of acute optic neuritis (ON). Further, a long-term follow-up aiming to investigate the relationship between ON and the development of MS was conducted. In total, 457 patients were enrolled between 1988 and 1991 and randomised to one of 2 different corticosteroid regimen or placebo. MRI scans were performed at baseline following a standardised protocol (30). Further, the EDSS was assessed at each follow-up visit (after 6 months, yearly until 1997, 2001-2002, 2006) and diagnosis of CDMS was based on Poser criteria (30, 31). All study arms were included in the analysis, because patients received only a short-term treatment.

London (The National Hospital, Queen Square; Moorfields Eye Hospital) (6, 23, 24, 27, 28, 32, 33)

Two different cohorts were identified: 1984-1987 (London I) and 1995-2004 (London II). In London I (National Hospitals for Neurology and Neurosurgery and Moorfields Eye Hospital), 140 patients were recruited between 1984 and 1987 (6, 23, 27, 28). MRI scans were performed at baseline (6). Follow-up assessment, including EDSS, was conducted after 1, 5 (23), 10 (24), 14 (28) and 20 years (6). CDMS was diagnosed according to the Poser criteria (6).

Filippi (33) performed a reanalysis of the 5-year-cohort data (23) and Sailer (27) of 58 patients of the 5- and 10-year-cohort data. Both used semi-automated techniques to analyse the MRI scans.

In London II, 143 patients with ON were included from 1995 to 2004 at the Moorfields Eye Hospital and National Hospitals for Neurology and Neurosurgery (32) and EDSS was assessed at 5 years. The ON cohort was part of a follow-up study. CDMS was diagnosed according to the Poser criteria (32).

Barcelona (Vall d'Hebron University Hospital) (5, 25, 26)

The Barcelona centre-based cohort was started in 1995 and is ongoing. People with CIS are included consecutively and evaluated regularly regarding EDSS and occurrence of relapses (every 3 - 6 months (5, 25, 26) or annually (5)). Baseline MRI scans were performed at the centre 3 - 5 months after a CIS diagnosis, after 12 months and then every 5 years (5). The EDSS was assessed at each visit and only considered for analysis during stable periods as indicated by the authors (5, 25, 26). CDMS was diagnosed according to the Poser criteria. In the first study (25), 175 patients were followed over 7 years. Another study (26) focused on the prognostic

role of baseline infratentorial lesions by analysing a subgroup of 77 patients with infratentorial lesions out of a cohort including 246 patients retrospectively. The latest publication from 2015 reports on 1058 patients (5).

Quality assessment

Heterogeneity of studies was high due to different study designs (randomised controlled trials (RCTs) and centre-based cohorts), inclusion criteria (ON only, mixed onset, age) and time periods for inclusion (e.g. 1984-1987, 1995-ongoing).

Quality assessment for potential bias (20) showed mixed results. Especially for "prognostic factor" (MRI assessment), "outcome measurement" (e.g. EDSS assessment) and "confounding measurement and account" (e.g. immunotherapy) detailed information was missing in most studies. In most studies, sufficient information concerning the "analysis" was provided (table 1).

Table 1: Quality assessment

Potential bias: No = ■, Partly = ■, Yes = □

Different MRI sequence techniques, field strengths of MRI and strategies of rater blinding increased the heterogeneity of studies (suppl. material 2). The synthesis of information on the MRI showed that information on field strength and scan resolution was provided by most studies. Information on "rater blinded to clinical details" is missing for 8, and "T2 lesion sequence" (e.g. T2/proton density or fluid-attenuated inversion recovery (FLAIR)) as well as "number of raters" for 5 of 13 studies (suppl. material 2).

Study results

Mean patient age at baseline ranged from 29 to 32 years across the cohorts, with more women (approx. 67% overall). Results are presented separately for CIS and ON only. As we found very limited information on transition to SPMS, data are presented in the "comments and additional" results section of table 2 and in the text. No information about mortality following a CIS and its relationship with MRI was reported. For comparison with the MRI prognostic features predicting disability following a CIS, we also included MRI features predicting conversion to CDMS (Table 3), but as these have already been considered previously (4, 9, 10), we do not further discuss these.

Reported outcomes

In table 2 it is listed on which different prognostic factors and outcomes were reported by the included studies, besides disability and conversion to CDMS. For more details see table 3 and the reported results section on disability progression.

Table 2: Prognostic factors and outcome reporting

CIS = clinically isolated syndrome, ON = optic neuritis

Table 3: Included studies: Magnetic resonance imaging as a prognostic factor for people with clinically isolated syndrome

Abbreviations: BL = baseline, CP = Cumulative probability, DO = dropout, FU = follow-up, Gd = Gadolinium, HR = hazard ratio, les = lesions, LV = lesion volume, MD = missing data, RRMS = relapsing-remitting multiple sclerosis, SPMS = secondary-progressive multiple sclerosis, r = Spearman's rank correlation, T2LV = T2 lesion volume
*T2 lesions unless otherwise stated (in outcomes), **no information on kind of average measurement, ***EDSS is reported at first and afterwards the MRI measure (e.g. EDSS 0: 0 (0-4) T2 lesions), ⦸ = no information provided

Disability progression

For baseline MRI measures of lesions, we present the associations with outcomes according to years of follow-up. Details about the strength of associations between MRI outcomes and disability are provided in table 4. We present the numbers of people with disability progression according to MRI results (normal or abnormal) in table 5.

CIS

Lesions seen on baseline MRI

5 to 7 years

Morrisey 1993 (23) (London I) showed that the number of baseline T2 MRI lesions correlated moderately with disability scores (EDSS) after 5 years (r (Spearman's rank correlation)=0.45, p=0.003), which was also shown by Tintore 2006 (25) (Barcelona) after 5 years (r=0.43, p<0.001).

Filippi (1994) (33), who reanalysed data from London I (23), reported a correlation between total baseline T2 lesion volume and disability scores at follow-up (r=0.62, p<0.0001). In addition, Filippi (33) (London I) showed that next to the total lesion volume also the infratentorial lesion volume at baseline appeared higher in patients with an EDSS ≥3 than in those with an EDSS score <3 (p<0.0001) after 5 years.

Tintore 2015 (5) (Barcelona, median follow-up: 6.8 years) found that the risk of an increased EDSS was higher in patients with 10 or more brain lesions (adjusted hazard ratio 2.9, 95% CI 1.4–6.0). Further, 11% of CIS patients with 10 or more lesions reached an EDSS score ≥3 during the first 5 years of follow-up based on Kaplan-Meier curves estimates. The authors performed a subgroup analysis of patients with spinal cord MRI and found an association between ≥1 spinal cord lesion and disability progression.

8 to 14 years

Tintore 2010 (Barcelona cohort) (26) reported on the prognostic value of baseline infratentorial lesions after a follow-up period of 7.7 years. Patients with infratentorial lesions were associated with a higher risk of reaching EDSS 3 (32.5% vs 12.4%, HR 2.4, 95% CI 1.3–4.3, p=0.003). However, the risk of mild disability was associated with the presence of at least 1 brainstem lesion (HR 2.5, 95% CI 1.1–5.4, p=0.026), but not with cerebellar lesions.

After a mean follow-up of 10 years, O'Riordan (24) (London I) showed that the number of T2 lesions at baseline was related to the EDSS after 10 years (r=0.45, p<0.001).

In the London cohort I (28), after a mean of 14 years, the median EDSS was 1.75 in patients with CDMS and normal baseline MRI (n=4). In comparison, the median EDSS was 3.5 in patients with CDMS and abnormal baseline MRI (n=44). There was a moderate correlation between EDSS at 14 years and number of lesions at BL (r=0.47, p<0.001), year 5 (r=0.55, p<0.001), year ten (r=0.45, p=0.001) and the number of new lesions from year 0-5 (r=0.51, p<0.001) and 10-14 (r=0.59, p<0.001). In addition, a moderate correlation between EDSS score at 14 years and BL lesion

volume (r=0.48, p<0.001), lesion volume at 5 years (r=0.60, p<0.001), lesion volume at 10 years, (r=0.48, p<0.001), change in lesion volume over the first 5 years (r=0.61, p<0.001), change in volume from year 5 to year 10 (weak correlation, r=0.29, p=0.037), and change in volume from year 10-14 (r=0.45, p=0.002) could be observed. There was a moderate correlation between change in lesion volume and change in EDSS score over 5 years (r=0.58, p<0.001) and weak correlations over the 5-10 year (r=0.41, p=0.002) and 10-14 year (r=0.35, p=0.02) periods (28). Baseline lesion volume as well as lesion number appeared larger in patients with worse clinical outcomes and those patients were associated with a larger increase in lesion volume over time (28).

15 to 20 years

In Fisniku (6) (London I), T2 lesion volume (baseline, 5, 10, 14 and 20 years) correlated moderately with the EDSS after 20 years' follow-up (r values: 0.48 to 0.67, p <0.001). Change in T2 lesion volume over the first 5 (r=0.61, p<0.001), over 5-10 (r=0.45, p=0.001), over 10-14 (r=0.40, p=0.004) and over 14-20 (r=0.49, p<0.001) years correlated weakly to moderately with change of EDSS after 20 years (6). Results for the baseline T2 lesion number and EDSS status (>3 or ≥6) after 20 years are provided in table 3.

Contrasting Barcelona and London cohort data on T2 lesions

In the Barcelona cohort (5) having 4-9 lesions (13 of 137 (10%), HR: 1.7, CI 0.8-3.8) is associated with a relevantly lower risk reaching EDSS 3 than ≥10 lesions (83 of 378 (22%), HR: 4.4, CI 2.4-8). In contrast, the descriptive London cohort data (23) show that 4 of 13 (30%) patients with 4-10 lesions reached EDSS ≥3 after 5 years. In

addition, the 20-year London data (6) indicate that having more than 1 lesion might already be associated with an increased risk (1-3 lesions: 8 of 22 (36%), 4-9 lesions: 10 of 20 (50%)). Nonetheless, the results from the London I cohort are based on a smaller sample size and it is not apparent whether findings were statistically significant. Only 4% (12 of 299) of patients with 0 lesions reached EDSS 3 during seven years of follow-up (5) in Barcelona. In contrast none of 32 patients with 0 lesions in London reached EDSS ≥3 at the 5 year follow-up (23), while 9 of 34 (26%) reached an EDSS >3 after 20 years (6).

Table 4: Evidence for an association of MRI measures with disability

+ = moderate evidence of effect, +/- = limited evidence of effect, -- = no evidence (33), n.r. = not reported (data was only considered, when p-values or confidence intervals were reported); *the number of new lesions over the first 5 years correlated with the change in EDSS over that time period

Optic neuritis only

5 to 7 years

In the ONTT, 2 of 30 CDMS patients (7%) with no MRI lesions developed EDSS ≥3 compared with 10 of 42 CDMS patients (24%) having ≥3 lesions (p=0.063) (30). Two out of 20 patients (10%) with 1 to 2 lesions developed moderate or severe disability (EDSS score ≥3).

Swanton (32) (London II) found the following markers as independent predictors of disability after 6 years: infratentorial (p=0.030), spinal cord (p=0.004), baseline Gd lesions (p=0.045) and new T2 lesions at 3-month follow-up (p=0.001). However, in people converting to CDMS, only presence of asymptomatic spinal cord lesions was associated with disability (32).

8 to 14 years

The mean EDSS score of lesion free patients (n=34) after 10-12 years was 2.4 compared to 2.6 in patients with 1 or more lesions (n=80). In the ONTT, no association between disability and presence of lesions or no lesions in baseline MRI (p=0.51) after 10-12 years was found (29). Further, the number of lesions appeared unrelated to disability (r=0.17, p=0.14) (29).

15 years

In the ONTT, there was no association between degree of disability after 15 years and number of brain lesions in the baseline MRI (r=0.07) (31).

Table 5: Main findings: MRI lesions as a prognostic factor for disability (EDSS ≥3)

Indicated are number of people (percentages) presenting with EDSS >= 3, * = only calculated in people with CDMS. °London II data were not comparable as they were subgrouped in 0-4 lesions.

Discussion

To the best of our knowledge, this is the first systematic review addressing the prognostic value of MRI for disability following a CIS.

The results illustrate that few studies assessed the prognostic MRI value beyond 5 years in people with a CIS. There is moderate evidence that the number of baseline lesions is associated with the EDSS after 5 years (23, 25). In the larger cohort from Barcelona (5), with a median follow-up of 6.8 years, it was found that the risk for an increased EDSS was higher in patients with 10 or more lesions. Further, infratentorial, especially brainstem lesions seem to be linked with greater disability.

Fisniku (6) reported that the T2 LV correlated moderately with EDSS after 20 years. It cannot be concluded based on our analysis whether lesion number or lesion volume is a better predictor for disability progression. Moreover, the study by Swanton et al (32) reporting on atrophy and gadolinium enhancing lesions, was the only study that additionally reported other MRI measurements than T2 lesions or volume.

Information is derived from only 3 different cohorts recruited in Barcelona and London as well as a trial follow-up study in ON. While a qualitative assessment of the results shows a consistent association between early brain lesion loads and subsequent disability, heterogeneity of cohorts and studies meant that meta-analysis of results was not possible. Except the ONTT (29), all cohorts were restricted to one city, which may limit generalisability to other regions. With up to 37%, there was a considerable loss to follow-up in the London cohort.

For the 5-year follow-up, 2 studies show similar results, but while recruitment for the London cohort started in 1988, recruitment in Barcelona started in 1995 and the applied MRI scanner differed considerably in terms of field strength and slice thickness. It is known that lesion detection will increase when switching from 0.5 to 1.5 and further to 3.0 Tesla, with up to ~40% more T2 lesions detectable on a 3.0 Tesla scanner compared to 1.5 Tesla (35-39), although its impact on correlations with disability has not been assessed (40).

Nonetheless, a higher detection rate seems not to lead to an earlier MS diagnosis (39-42), but this has to be confirmed for newer diagnostic criteria (2, 43). In the Barcelona cohort, McDonald 2010 diagnostic criteria but not better lesion detection technology might have led to a higher percentage of patients diagnosed with MS and short-term low disability.

The importance of infratentorial lesions was recently readdressed in the MAGNIMS consensus guidelines 2015 (4). The potential effect of infratentorial lesions on disability was indicated by both cohorts (London I and Barcelona) (26, 33) and by the study of Swanton (32) on ON (London II). However, in case of ON it was shown by Swanton et al (32) that only asymptomatic spinal cord lesions predicted disability in patients converting to CDMS during follow-up. It was also touched on by Tintore et al (5) that ≥1 spinal cord lesion was associated with disability progression at year 6.8. However, in a recent editorial based on other studies, Brownlee (44) concluded that the diagnostic and prognostic value of spinal cord lesions is still uncertain. A more complete MRI picture of the characteristics of all lesions might improve the prognostic value.

Interestingly, data from the ONTT (31) showed no association between baseline MRI and degree of disability after 15 years in people with ON, whereas development of CDMS was strongly related to the baseline MRI (31). By contrast, in the 14 year follow-up of the London I cohort (28), EDSS at year 14 correlated with T2 lesion number at baseline. The Barcelona cohort showed that patients with ON had a lower risk of reaching EDSS 3.0 and Swanton (London II) found for a follow-up <6 years that only 12% of this ON cohort had an EDSS ≥3. Together, there is consistency that disability progression in those with ON as a first symptom of CIS being is lower than those with other CIS presentations. While in the ONTT 70 of 105 patients (67%) with CDMS had an EDSS <3 after 15 years, median EDSS after 14 years in the London I cohort was 3.25 in patients with CDMS.

Because longer observation times are needed to detect SPMS conversion in a relevant percentage results on MRI and SPMS conversion, results on MRI and the conversion to SPMS are only presented for the London I cohort (6, 24, 27, 28).

Fisniku (6) reports that 28 (42%) patients with CDMS were diagnosed with SPMS after 20 years. There was a non-significant trend that development of SPMS was associated with a higher T2 lesion volume increase over 20 years than in patients who remained RR (p=0.07). T2 lesion volume increase in patients diagnosed with SPMS appeared higher than in those with RRMS over the first 5 years (p=0.008) (6). Results beyond 15 years of follow-up are restricted to only this study (London I). Results on MRI and mortality were not reported. Here, studies with an even longer follow-up than 20 years are needed.

There were no studies, which analysed the prognostic value of brain atrophy with our search criteria, although a combination of inflammatory and degenerative measures might lead to a better matching of MRI parameters and prognosis (44). Di Filippo evaluated brain atrophy in a cohort with n=99 CIS patients and found a higher atrophy rate in patients with a later CDMS diagnoses, but brain atrophy rate was not an independent predictor of disability at 6 years (45). This was also concluded by a recent Australian study analysing the corpus callosum area in n=143 CIS patient with a follow-up of 5 years (46). In a sample from Barcelona (n=54), a significant association between lower grey matter volume and disability after 15 years was found (47). The analysis by Brownlee (48) of n=131 non-spinal CIS patients of the London II cohort respective the relevance of spinal cord atrophy and lesions indicates that the upper cervical cord cross-sectional area is an independent disability predictor at 5 years. Work is still needed to clarify the independent and interaction impact of brain and upper cervical cord atrophy as potential prognostic markers (47).

While higher brain lesion load, and in particular infratentorial lesions, appear to be linked with longer-term disability following a CIS, correlations of early brain lesion loads and later disability still only explain a small proportion of disability. Similarly, it is difficult to define thresholds that can be used at an individual level in clinical practice, as there is no definitive cut-off above which functionally significant disability is certain. As such, in the absence of robust prognostic models of longer-term disability following a CIS, it is still important to recognise that at an individual level neither CIS nor MS outcomes can be reliably predicted.

Limitations

While reviewing the literature, several limitations became apparent. With regard to the main MRI measures, these were not always presented in the same way (e.g. T2 lesion numbers were often summed up in different categories), which made a direct comparison difficult. The main clinical disability measure for these studies was the EDSS score, and while this gives a clear assessment of walking function (and so lower limb motor function), it undervalues other causes of disability, such as cognitive impairments (49). A more complete assessment of neurological and cognitive function might increase the strength of association of lesion load with outcomes. In addition, localisation of lesions such as analysis of cortical lesions might show a closer link to disability (50), but no data from larger cohorts of 5 or more years taking cortical lesion location into account are available.

Further, a publication bias could have influenced the results as studies reporting negative results might have not been published. As we did not search systematically for grey literature, we could have missed relevant studies. However, we checked the references of all included studies to detect further relevant studies.

Conclusion

There are very few long-term CIS datasets, but from these we conclude that following a CIS, the number or volume of T2 MRI brain lesions have some prognostic value for disability at 5 to 7 years, and infratentorial lesions seem to be of particular relevance. Further work is required to determine if other MRI measures, such as brain atrophy, have a higher predictive potential.

Acknowledgment

We would like to thank Judith Brand for her contribution on the literature searches and title/abstract screening as well as Anne-Marie Köpke for the assessment of included studies and data extraction. We would like to thank Ingrid Mühlhauser for the helpful comments on the draft of this paper.

Funding

This study was funded by the German Ministry of Education and Research within the Competence Network Multiple Sclerosis (Kompetenznetz Multiple Sklerose, 01GI1206). The funding body had no influence on the design, administration, analysis, and interpretation, as well as the dissemination of results of this study.

Conflicts of interest

SK has nothing to declare. AR was funded by German Ministry of Education and Research and received sponsorship from the University of Hamburg (equalisation fund) and is supported by a research grant from the National MS Society, USA (grant no. G-1508-06034). IS was supported by a research grant from the National MS Society, USA (grant no. G-1508-06034). CH has received research grants, congress travel compensations, and salaries for talks from BiogenIdec, Genzyme, Sanofi-Aventis, Bayer Healthcare, Merck Serono, Teva Pharma, and Novartis. CL received

a research grant by the German Federal Ministry for Education and Research, BMBF, German Competence Network Multiple Sclerosis (KKNMS), grant no. 01GI1601I, has received consulting and speaker's honoraria from BiogenIdec, Bayer Schering, Novartis, Sanofi, Genzyme and TEVA, and has received research scientific grant support from Bayer Schering, TEVA and MerckSerono. He holds an endowed professorship supported by the Novartis Foundation. DC has received honoraria (paid to his employer) from Ismar Healthcare NV, Swiss MS Society, Excemed (previously Serono Symposia International Foundation), Merck, Bayer and Teva for faculty-led education work; Teva for advisory board work; meeting expenses from Merck, Teva, Novartis, the MS Trust and National MS Society; and has previously held stock in GlaxoSmithKline. JPS receives research funding from the German Research Foundation, reports grants and travel funding and/or speaker honoraria from Biogen and Genzyme outside the submitted work and was supported by a research grant from the National MS Society, USA (grant no. G-1508-06034).

Authors' contributions

CH, AR and JB developed the search strategy and AR and JB performed the search. JB, SK, AR and CH screened the articles. AR, CH and AMK assessed the included studies and extracted data. AR wrote the manuscript. CH, CL, DC, PS, IS and SK commented on and edited the manuscript. All authors read and approved the final manuscript.

References

1. Miller DH, Chard DT, Ciccarelli O. Clinically isolated syndromes. Lancet Neurol. 2012;11(2):157–69.
2. 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;69(2):292–302.
3. Rovira À, Wattjes MP, Tintoré M, Tur C, Yousry TA, Sormani MP, et al. Evidence-based guidelines: MAGNIMS consensus guidelines on the use of MRI in multiple sclerosis-clinical implementation in the diagnostic process. Nature reviews Neurology. 2015;11(8):471–82.
4. Wattjes MP, Rovira A, Miller D, Yousry TA, Sormani MP, de S, et al. Evidence-based guidelines: MAGNIMS consensus guidelines on the use of MRI in multiple sclerosis--establishing disease prognosis and monitoring patients. Nature Reviews Neurology. 2015;11(10):597–606.
5. Tintore M, Rovira A, Rio J, Otero-Romero S, Arrambide G, Tur C, et al. Defining high, medium and low impact prognostic factors for developing multiple sclerosis. BRAIN. 2015;138:1863–74.
6. Fisniku LK, Brex PA, Altmann DR, Miszkiel KA, Benton CE, Lanyon R, et al. Disability and T2 MRI lesions: a 20-year follow-up of patients with relapse onset of multiple sclerosis. Brain : a journal of neurology. 2008;131(Pt 3):808–17.
7. Brand J, Köpke S, Kasper J, Rahn A, Backhus I, Pöttgen J, et al. Magnetic Resonance Imaging in Multiple Sclerosis - Patients' Experiences, Information Interests and Responses to an Education Programme. PloS one. 2014;9(11).
8. Heesen C, Kasper J, Segal J, Köpke S, Mühlhauser I. Decisional role preferences, risk knowledge and information interests in patients with multiple sclerosis. Mult Scler. 2004;10(6):643–50.
9. Odenthal C, Coulthard A. The prognostic utility of MRI in clinically isolated syndrome: a literature review. AJNR American journal of neuroradiology. 2015;36(3):425–31.
10. Zhang W-Y, Hou Y-l. Prognostic value of magnetic resonance imaging in patients with clinically isolated syndrome conversion to multiple sclerosis: A meta-analysis. NEUROLOGY INDIA. 2013;61(3):231–8.
11. Higgins JPT, editor. Cochrane handbook for systematic reviews of interventions. Reprinted. ed. Chichester: Wiley-Blackwell; 2011.
12. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ (Clinical research ed). 2009;339:b2535.
13. Weinshenker BG, Issa M, Baskerville J. Long-term and short-term outcome of multiple sclerosis: a 3-year follow-up study. Arch Neurol. 1996;53(4):353-8.
14. Comi G, Martinelli V, Rodegher M, Moiola L, Leocani L, Bajenaru O, et al. Effects of early treatment with glatiramer acetate in patients with clinically isolated syndrome. Mult Scler. 2013;19(8):1074-83.
15. Kappos L, Edan G, Freedman MS, Montalban X, Hartung HP, Hemmer B, et al. The 11-year long-term follow-up study from the randomized BENEFIT CIS trial. Neurology. 2016;87(10):978-87.
16. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology. 1983;33(11):1444–52.

17. Wilczynski NL, Haynes RB. Optimal search strategies for detecting clinically sound prognostic studies in EMBASE: an analytic survey. Journal of the American Medical Informatics Association : JAMIA. 2005;12(4):481–5.
18. Wilczynski NL, Haynes RB. Developing optimal search strategies for detecting clinically sound prognostic studies in MEDLINE: an analytic survey. BMC medicine. 2004;2:23.
19. Hayden JA, Côté P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. Annals of internal medicine. 2006;144(6):427–37.
20. Hayden JA, van der Windt DA, Cartwright JL, Cote P, Bombardier C. Assessing bias in studies of prognostic factors. Annals of internal medicine. 2013;158(4):280–6.
21. Stellmann J-P, Stürner KH, Young KL, Siemonsen S, Friede T, Heesen C. Regression to the mean and predictors of MRI disease activity in RRMS placebo cohorts--is there a place for baseline-to-treatment studies in MS? PloS one. 2015;10(2):e0116559.
22. Hayden JA, Tougas ME, Riley R, Iles R, Pincus T. Individual recovery expectations and prognosis of outcomes in non- specific low back pain: prognostic factor exemplar review. Cochrane Database of Systematic Reviews. 2014(2014 Issue 9):CD011284.
23. MORRISSEY SP, Miller DH, KENDALL BE, Kingsley DP, KELLY MA, Francis DA, et al. The significance of brain magnetic resonance imaging abnormalities at presentation with clinically isolated syndromes suggestive of multiple sclerosis. A 5-year follow-up study. BRAIN. 1993;116 (Pt 1):135–46.
24. O'Riordan JI, Thompson AJ, KINGSLEY DPE, MACMANUS DG, KENDALL BE, Rudge P, et al. The prognostic value of brain MRI in clinically isolated syndromes of the CNS - A 10-year follow-up. BRAIN. 1998;121:495–503.
25. Tintore M, Rovira A, Rio J, Nos C, Grive E, Tellez N, et al. Baseline MRI predicts future attacks and disability in clinically isolated syndromes. Neurology. 2006;67(6):968–72.
26. Tintore M, Rovira A, Arrambide G, Mitjana R, Rio J, Auger C, et al. Brainstem lesions in clinically isolated syndromes. Neurology. 2010;75(21):1933–8.
27. Sailer M, O'Riordan JI, Thompson AJ, Kingsley DP, MACMANUS DG, MCDONALD WI, et al. Quantitative MRI in patients with clinically isolated syndromes suggestive of demyelination. Neurology. 1999;52(3):599–606.
28. Brex PA, Ciccarelli O, O'Riordan JI, Sailer M, Thompson AJ, Miller DH. A longitudinal study of abnormalities on MRI and disability from multiple sclerosis. New England Journal of Medicine. 2002;346(3):158–64.
29. Beck RW, Smith CH, Gal RL, Xing D, Bhatti MT, Brodsky MC, et al. Neurologic impairment 10 years after optic neuritis. Archives of neurology. 2004;61(9):1386–9.
30. Optic NSG. The 5-year risk of MS after optic neuritis. Experience of the opticneuritis treatment trial. Neurology. 1997;49(5):1404–13.
31. Optic NSG. Multiple sclerosis risk after optic neuritis: final optic neuritis treatment trial follow-up. Archives of neurology. 2008;65(6):727–32.
32. Swanton JK, Fernando KT, Dalton CM, Miszkiel KA, Altmann DR, Plant GT, et al. Early MRI in optic neuritis: the risk for disability. Neurology. 2009;72(6):542–50.
33. Filippi M, Horsfield MA, MORRISSEY SP, MACMANUS DG, Rudge P, MCDONALD WI, et al. Quantitative brain MRI lesion load predicts the course of

clinically isolated syndromes suggestive of multiple sclerosis. Neurology. 1994;44(4):635–41.
34. O'Riordan JI, Thompson AJ, Kingsley DP, MacManus DG, Kendall BE, Rudge P, et al. The prognostic value of brain MRI in clinically isolated syndromes of the CNS. A 10-year follow-up. Brain. 1998;121(Pt 3):495-503.
35. Bachmann R, Reilmann R, Schwindt W, Kugel H, Heindel W, Kramer S. FLAIR imaging for multiple sclerosis: A comparative MR study at 1.5 and 3.0 Tesla. European radiology. 2006;16(4):915–21.
36. Sicotte NL, Voskuhl RR, Bouvier S, Klutch R, Cohen MS, Mazziotta JC. Comparison of multiple sclerosis lesions at 1.5 and 3.0 Tesla. INVESTIGATIVE RADIOLOGY. 2003;38(7):423–7.
37. Nielsen K, Rostrup E, Frederiksen JL, Knudsen S, Mathiesen HK, Hanson LG, et al. Magnetic resonance imaging at 3.0 tesla detects more lesions in acute optic neuritis than at 1.5 tesla. INVESTIGATIVE RADIOLOGY. 2006;41(2):76–82.
38. Wattjes MP, Lutterbey GG, Harzheim M, Gieseke J, Traber F, Klotz L, et al. Higher sensitivity in the detection of inflammatory brain lesions in patients with clinically isolated syndromes suggestive of multiple sclerosis using high field MRI: an intraindividual comparison of 1.5 T with 3.0 T. European radiology. 2006;16(9):2067–73.
39. Wattjes MP, Harzheim M, Kuhl CK, Gieseke J, Schmidt S, Klotz L, et al. Does high-field MR imaging have an influence on the classification of patients with clinically isolated syndromes according to current diagnostic mr imaging criteria for multiple sclerosis? AJNR American journal of neuroradiology. 2006;27(8):1794–8.
40. Hagens MH, Burggraaff J, Kilsdonk ID, Ruggieri S, Collorone S, Cortese R, et al. Impact of 3 Tesla MRI on interobserver agreement in clinically isolated syndrome: A MAGNIMS multicentre study. Mult Scler. 2018:1352458517751647.
41. Swanton JK, Fernando K, Dalton CM, Miszkiel KA, Thompson AJ, Plant GT, et al. Modification of MRI criteria for multiple sclerosis in patients with clinically isolated syndromes. Journal of neurology, neurosurgery, and psychiatry. 2006;77(7):830–3.
42. Swanton JK, Rovira A, Tintore M, Altmann DR, Barkhof F, Filippi M, et al. MRI criteria for multiple sclerosis in patients presenting with clinically isolated syndromes: a multicentre retrospective study. Lancet Neurol. 2007;6(8):677–86.
43. 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;17(2):162-73.
44. Gumberz Jv, Mahmoudi M, Young K, Schippling S, Martin R, Heesen C, et al. Short-term MRI measurements as predictors of EDSS progression in relapsing-remitting multiple sclerosis: grey matter atrophy but not lesions are predictive in a real-life setting. PeerJ. 2016;4:e2442.
45. Di Filippo M, Anderson VM, Altmann DR, Swanton JK, Plant GT, Thompson AJ, et al. Brain atrophy and lesion load measures over 1 year relate to clinical status after 6 years in patients with clinically isolated syndromes. JOURNAL OF NEUROLOGY NEUROSURGERY AND PSYCHIATRY. 2010;81(2):204–8.
46. Odenthal C, Simpson S, Oughton J, van d, Rose S, Fripp J, et al. Midsagittal corpus callosum area and conversion to multiple sclerosis after clinically

isolated syndrome: A multicentre Australian cohort study. Journal of Medical Imaging & Radiation Oncology. 2017;61(4):453–60.
47. 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. Multiple Sclerosis. 2017;017:1352458517707070.
48. Brownlee WJ, Altmann DR, Alves D, Swanton JK, Miszkiel KA, Wheeler-Kingshott CG, et al. Association of asymptomatic spinal cord lesions and atrophy with disability 5 years after a clinically isolated syndrome. Multiple Sclerosis. 2017;23(5):665–74.
49. Patti F, De S, Lavorgna L, Messina S, Chisari CG, Ippolito D, et al. Lesion load may predict long-term cognitive dysfunction in multiple sclerosis patients. PLoS ONE [Electronic Resource]. 2015;10(3):e0120754.
50. Calabrese M, De Stefano N, Atzori M, Bernardi V, Mattisi I, Barachino L, et al. Detection of cortical inflammatory lesions by double inversion recovery magnetic resonance imaging in patients with multiple sclerosis. Archives of Neurology. 2007;64(10):1416-22.

Table 1: Quality assessment (19, 20)

Study	Potential bias					
	Study participation	Study attrition	Prognostic factor	Outcome measurement	Confounding measurement and account	Analysis
ONTT 1997 (30)						
ONTT 2004 (29)						
ONTT 2008 (31)						
Tintore 2006 (25)						
Tintore 2010 (26)						
Tintore 2015 (5)						
Morrissey 1993 (23)						
Filippi 1994 (33)						
O'Riordan 1998 (24)						
Sailer 1999 (27)						
Brex 2002 (28)						
Fisniku 2008 (6)						
Swanton 2009 (32)						

Table 2: Prognostic factors and outcome reporting

	Prognostic factor/ outcome	**Study**
CIS cohorts	T2 brain lesion number	Tintore 2006 (25), Tintore 2015 (5), Morrissey 1993 (23), O'Riordan 1998 (24), Brex 2002 (28), Fisniku 2008 (6)
	T2 infratentorial lesion number	Tintore 2010 (26)
	T2 brain lesion volume	Filippi 1994 (33), Sailer 1999 (27), Brex 2002 (28), Fisniku 2008 (6)
	T2 infratentorial lesion volume	Filippi 1994 (33), Sailer 1999 (27)
	Conversion to SPMS	O'Riordan 1998 (24), Sailer, 1999 (27), Brex 2002 (28), Fisniku 2008 (6)
ON only cohorts	T2 brain lesion number	ONTT 1997 (30), ONTT 2004 (29), ONTT 2008 (31), Swanton 2009 (32)
	T2 infratentorial lesion number	Swanton 2009 (32)
	T2 brain lesion volume	Swanton 2009 (32)
	Atrophy	Swanton 2009 (32)
	Gadolinium enhancing lesions	Swanton 2009 (32)

CIS = clinically isolated syndrome, ON = optic neuritis

Table 3: Included studies: Magnetic resonance imaging as a prognostic factor for people with clinically isolated syndrome

Study	BL information					Outcomes*	
	Centres and design	Clinical features	Study length (years)	T2 (number of lesions)	Lesion volume	EDSS*	CDMS*
ONTT 1997 (30)	15 centres RCT follow-up 1988-1991	ON 457 388 analysed DO: 15%	5**	0: 202 (52%) 1-2: 60 (16%) ≥3: 89 (23%) MD: 37 (10%)	⊘	EDSS ≥3 in pat. with CDMS: 2 of 30 (7%) with 0 les 2 of 20 (10%) with 1-2 les 10 of 42 (24%) with ≥3 les MD: 13	CP of developing CDMS: 16% with 0 les 44% with 1 les 26% with 2 les 51% with ≥3 les MD: 40
	Sixty-six of 457 patients were excluded from the analysis. MRI was performed in 351 patients. Patients received corticosteroids or placebo for 14 days. The number of lesions at BL was associated with the development of CDMS by 5 years.						
ONTT 2004 (29)	14 centres RCT follow-up 1988-1991	ON 147 (Sub-sample) 127 analysed DO: 14%	10 - 12**	0: 34 (27%) 1: 21 (17%) 2-4: 22 (17%) 5-8: 22 (17%) ≥9:15 (12%) MD: 13 (10%)	⊘	EDSS <3 in pat. with CDMS: 24 of 34 (71%) with 0 les 50 of 80 (63%) with ≥1 les EDSS ≥3 in pat. with CDMS: 10 of 34 (29%) with 0 les 30 of 80 (38%) with ≥1 les MD:13	⊘
	Subsample from the original cohort with 388 evaluated patients, 145 patients with CDMS included, 127 analysed. Data on MRI and EDSS were available for 114 patients.						
ONTT 2008 (31)	13 centres RCT follow-up 1988-1991	ON 389 analysed Number of BL patients is not provided	15**	0: 191 (49%) 1: 44 (11%) 2: 26 (7%) ≥3: 91 (23%) MD: 37 (10%)	⊘	EDSS <3 in pat. with CDMS: 17 of 28 (61%) with 0 les 18 of 21 (86%) with 1 les 35 of 56 (62%) with ≥2 les EDSS ≥3 in pat. with CDMS: 11 of 28 (39%) with 0 les 3 of 21 (14%) with 1 les 21 of 56 (38%) with ≥2 les	CP of developing CDMS: 25% with 0 les 60% with 1 les 68% with 2 les 78% with ≥3 les
	No relation between degree of disability and brain lesions on the BL MRI (r=0.07). CP of developing CDMS after 15 years was strongly related to BL MRI.						

Study	BL information					Outcomes*	
	Centres and design	Clinical features	Study length (years)	T2 (number of lesions)	Lesion volume	EDSS*	CDMS*
Tintore 2006 (25)	Barcelona centre based pro-spective cohort 1995-1998	CIS 175 156 analysed DO: 11%	7 median	⊘	⊘	EDSS ≥3 after 8 years (mean) 3 of 52 (6%) with 0 les 2 of 23 (9%) with 1-3 les 2 of 18 (11%) with 4-9 les 16 of 63 (25%) with ≥10 MD:5	4 of 52 (8%) with 0 les 7/23 (30%) with 1-3 les 9/18 (50%) with 4-9 les 46/63 (73%) ≥10 les
	EDDS at 5 years correlated with the number of lesions in the baseline MRI (r=0.43, p<0.001). See Tintore 2006 (22) for cumulative probabilities (figure E-1, E2).						
Tintore 2010 (26)	Barcelona centre based pro-spective cohort 1995 - 2001	CIS 246 77 with infra-tentorial lesions analysed	7.7 median	0: 84 (34%) 1-8: 74 (30%) ≥9: 83 (34%) MD: 5 (2%)	⊘	EDSS 3 12% with 0 infratentorial les 33% with ≥1 infratentorial les 11% with normal brain MRI 24% with abnormal brain MR 10% with 0 les 15% with 1-8 les 33% with ≥ 9 les	30% with 0 infratentorial les 71% ≥1 infratentorial les 10% with normal brain MRI 63% with abnormal brain MRI 8% with 0 les 47% with 1-8 les 75% with ≥ 9 les
	Retrospective MRI analysis. Infratentorial lesions were associated with a higher risk of conversion to CDMS (71.4% vs 29.6%, HR 3.3, 95% CI 2.2–4.8, p<0.001) and of reaching EDSS 3 (32.5% vs 12.4%, HR 2.4, 95% CI 1.3–4.3, p=0.003). See Tintore 2006(22) and 2015 (5) for further results.						
Tintore 2015 (5)	Barcelona centre based pro-spective cohort 1995-2013	CIS 1058 1015 analysed DO: 4%	6.8 mean	0: 299 (31%) 1-3: 137 (14%) 4-9: 137 (14%) ≥10: 378 (40%) MD: 64 (not in-corporated)	⊘	EDSS 3 12 of 299 (4%) with 0 les 10 of 137 (7%) with 1-3 les 13 of 137 (10%) with 4-9 les 83 of 378 (22%) with ≥10 les	21 of 299 (7%) with 0 les 56 of 137 (41%) with 1-3 les 71 of 137 (52%) with 4-9 les 240 of 378 (64%) with ≥10 les
	10 or more brain MRI lesions were associated with an increased risk of CDMS (adjusted HR 11.3 (6.7–19.3)) and disability (adjusted HR 2.9 (1.4–6.0)).						
Morrissey 1993 (23)	London centre based pro-spective cohort 1984-1987	CIS 132 89 analysed DO: 33%	5.3 mean	⊘	⊘	EDSS ≥3 0 of 32 (0%) with 0 les 0 of 6 (0%) with 1 les 3 of 18 (17%) with 2-3 les 4 of 13 (30%) with 4-10 les 11 of 20 (56%) with >10 les	2 of 32 (6%) with 0 les 1 of 6 (17%) with 1 les 12 of 18 (67%) with 2-3 les 12 of 13 (92%) with 4-10 les 16 of 20 (80%) with >10 les 1 of 32 (3%) with normal MRI 37of 57 (65%) with abnormal MRI

Study	BL information					Outcomes*	
	Centres and design	Clinical features	Study length (years)	T2 (number of lesions)	Lesion volume	EDSS*	CDMS*
	The number of baseline MRI lesions correlated with disability at follow-up (r=0.45, p=0.003). There was a trend for an association that patients with more lesions developed MS more frequently (28 of 33 (85%) with ≥4 compared to 13 of 24 (54%) with ≤3 lesions (x^2=6.48, p<0.02).						
Filippi 1994 (33)	London Re-analyses of pros-pective cohort data (23) 1984-1987	CIS 84 analysed MD: Number of BL patients	5.3 mean	⦸	52 with abn orm al MRI: 0.83 cm^3	EDSS ≥3 11 of 21 (52%) with >1.23cm^3 T2LV 7 of 31 (23%) with <1.23cm^3 T2LV 0 of 32 (0%) with normal MRI	18 of 21 (86%) with ≥1.23cm^3 T2LV 15 of 31 (48%) with <1.23cm^3 T2LV 1 of 32 (3%) with normal MRI
	The total and infratentorial T2LV at BL were higher in patients, who developed CDMS (p<0.0001). Compared to 44 of 64 (69%) without infratentorial lesions 18 of 20 patients (90%) with infratentorial lesions developed CDMS. 14 of those 18 patients had a total T2LV >1.23cm^3. The total and infratentorial T2LV at BL was higher in patients with an EDSS ≥3 than in those with an EDSS score <3 (p<0.0001). The total BL T2LV correlated with the severity of disability at follow-up (r=0.619, p<0.0001).						
O'Riordan 1998 (24)	London centre based pro-spective cohort 1984-1987	CIS 129 81 analysed DO: 37%	9.7 mean	⦸	⦸	EDSS >3 0 of 27 (0%) with 0 les 0 of 3 (0%) with 1 les 5 of 16 (31%) with 2-3 les 4 of 15 (27%) with 4-10 les 14 of 20 (70%) with >10 les EDSS >5.5 (due to MS) 0 of 27 (0%) with 0 les 0 of 3 (0%) with 1 les 2 of 16 (13%) with 2-3 les 3 of 15 (20%) with 4-10 les 7 of 20 (35%) with >10 les	3 of 27 (11%) with 0 les 1 of 3 (33%) with 1 les 14 of 16 (87%) with 2-3 les 13 of 15 (87%) with 4-10 les 17 of 20 (85%) with >10 les 3 of 27 (11%) with normal MRI 45 of 54 (83%) with abnormal MRI
	Of those 45 with abnormal MRI and CDMS 21 (39%) had benign MS, 11 (20%) RRMS and 13 (24%) SPMS. Infratentorial lesions were not associated with development of CDMS (p=0.35). MRI lesions are described as asymptomatic lesions.						
Sailer 1999 (27)	London centre based pro-spective cohort 1984-1987	CIS 71,58 with BL, 5, 10 year MRI analysed MD: Number of BL patients	9.7 mean	2 (0-74)	0.43 cm^3 med ian	EDSS >3 9 of 34 (27%) with <3.0cm^3 T2LV 9 of 11 (82%) with >3cm^3 T2LV 5 of 26 (19%) with normal MRI Percentages were partly recalculated by the authors (table six in publication). For data on EDDS ≥6 see[33]	28 of 34 (82%) with <3.0cm^3 T2LV 11 of 11 (100%) with >3cm^3 T2LV 5 of 26 (19%) with normal MRI Percentages were partly recalculated by the authors (table six in publication).
	There was a moderate correlation between infratentorial LV at BL and the total LV at 5 (r=0.5, p=0.001) and 10-year follow-up (r=0.44, p=0.006). The total LV at BL was associated with the disease type at 5 years (CIS, CPMS, or CDMS) (p=0.0001, Kruskal-Wallis) and 10 years (CIS, CPMS, or CDMS, benign, RR, SP; p=0.0001, Kruskal-Wallis).						

Study	BL information					Outcomes*	
	Centres and design	Clinical features	Study length (years)	T2 (number of lesions)	Lesion volume	EDSS*	CDMS*
Brex 2002 (28)	London centre based pro-spective cohort 1984-1987	CIS 71 of 81 analysed MD: Number of BL patients (only for BL and 1 year FU together (109 patients)	14.1 mean	⊘	0.46 cm^3 median	EDSS >3 0 of 21 (0%) with 0 les 5 of 18 (28%) with 1-3 les 7 of 15 (47%) with 4-10 les 12 of 17 (71%) with >10 les EDSS ≥6 0 of 21 (0%) with 0 le s 2 of 18 (11%) with 1-3 les 4 of 15 (27%) with 4-10 les 9 of 17 (53%) with >10 les Median T2LV EDSS >3 0 of 21 (0%) with 0 cm^3 5 of 18 (28%) with 0.6 cm^3 7 of 15 (47%) with 0.9 cm^3 12 of 17 (71%) with 5.6 cm^3 EDSS ≥6 0 of 21 (0%) with 0 cm^3 2 of 18 (11%) with 0.6 cm^3 4 of 15 (27%) with 0.9 cm^3 9 of 17 (53%) with 5.6 cm^3	4 of 21 (19%) with 0 les 16 of 18 (89%) with 1-3 les 13 of 15 (87%) with 4-10 les 15 of 17 (88%) with >10 les 4 of 21 (19%) with normal MRI 44 of 50 (88%) with abnormal MRI
	There was a correlation between EDSS at 14 years and number of lesions at BL (r=0.47, p<0.001), year 5 (r=0.55, p<0.001), year 10 (r=0.45, p=0.001) and the number of new lesions from year 0 to 5 (r=0.51, p<0.001) and year 10 to year 14 (r=0.59, p<0.001). There was a correlation between EDSS score at 14 years and BL T2LV (r=0.48, p<0.001), the T2LV at 5 years (r=0.60, p<0.001), T2LV at 10 years, (r=0.48, p<0.001), change in T2LV over the first 5 years (r=0.61, p<0.001), change in T2LV from year 5 to year 10 (r=0.29, p=0.037); and change in T2LV from year 10 to year 14 (r=0.45, p=0.002). There was a correlation between change in T2LV and change in EDSS score over 5 years (r=0.58, p<0.001) and over the 5-to-10-year (r=0.41, p=0.002) and 10-to-14-year (r=0.35, p=0.02) periods. 17 of 44 patients with CDMS and abnormal MRI were diagnosed with SPMS.						
Fisniku 2008 (6)	London centre based pro-spective cohort 1984-1987	CIS 140 107 analysed DO: 24%	20.2 mean	⊘	0.43 cm^3 median	EDSS >3 9 of 34 (26%) with 0 les 8 of 22 (36%) with 1-3 les 10 of 20 (50%) with 4-9 les 20 of 31(65%) with ≥10 les EDSS ≥6 2 of 34 (6%) with 0 les 4 (18%) of 22 with 1-3 les 7 (35%) of 20 with 4-9les 14 (45%) of 31 with ≥10 les	7 of 34 (21%) with 0 les 18 of 22 (82%) with 1-3 les 17 of 20(85%) with 4-9 25 of 31(81%) with ≥10 les 7 of 34 (21%) with normal MRI 60 of 73 (82%) with abnormal MRI
	T2LV (BL, after 5, 10, 14 and 20 years) correlated moderately with EDSS after 20 years (r values: 0.48-0.67; p <0.001). In those patients with CDMS, 28 (42%) had developed SPMS. There was a non-significant trend for an association that patients who developed SPMS tended to have a higher T2LV growth over 20 years than those with RRMS (p=0.07). Patients with SPMS appeared to have a larger BL T2LV (p-value not provided). T2LV increase in SPMS appeared higher than in RRMS over the first 5 years (p=0.008).						

Study	BL information					Outcomes*	
	Centres and design	Clinical features	Study length (years)	T2 (number of lesions)	Lesion volume	EDSS*	CDMS*
Swanton 2009 (32)	London centre based pro-spective cohort 1995-2004	ON 143 100 of 106, who reached 5 years FU were analysed DO: 6%	6 median	5 median	T2LV: 0.46 (0.05–1.81) ml T1LV: 0 (0–0.15) ml	EDSS (5-year subgroup) Brain T2 les*** 0: 0 (0–4) 1: 3 (0-16) 1.5-2: 6 (3–16) ≥2.5: 15 (6–29) Gd les 0: 0 (0–0) 1: 0 (0–0) 1.5-2: 0 (0–1) ≥2.5: 2 (0–6) Infratentorial les 0: 0 (0-0) 1: 0 (0-0) 1.5-2: 0 (0-2) ≥2.5: 1 (0–2) T2 LV 0: 0 (0–0.34) ml 1: 0.19 (0–1.67) ml 1.5-2: 0.79 (0.23–1.81) ml ≥2.5: 2.08 (0.48–4.13) ml	
	Independent predictors linked to disability after 6 years were spinal cord (p=0.004), infratentorial (p=0.030) lesions, BL Gd lesions (p=0.045) and new T2 lesions at 3-month follow-up (p=0.001). But in patients converting to CDMS, only spinal cord lesions were associated with disability. Grey matter fraction by EDSS: 0: 0.50 (0.49–0.51); 1: 0.49 (0.47–0.51); 1.5-2: 0.48 (0.48–0.51); ≥2.5: 0.48 (0.46–0.50. White matter fraction by EDSS: 0: 0.37 (0.36–0.38); 1: 0.37 (0.36–0.39); 1.5-2: 0.37 (0.36–0.38); ≥2.5: 0.38 (0.36–0.40).						

Abbreviations: BL = baseline, CP = Cumulative probability, DO = dropout, FU = follow-up, Gd = Gadolinium, HR = hazard ratio, les = lesions, LV = lesion volume, MD = missing data, RRMS = relapsing-remitting multiple sclerosis, SPMS = secondary-progressive multiple sclerosis, r = Spearman's rank correlation, T2LV = T2 lesion volume

*T2 lesions unless otherwise stated (in outcomes), **no information on kind of average measurement, ***EDSS is reported at first and afterwards the MRI measure (e.g. EDSS 0: 0 (0-4) T2 lesions), ⦸ = no information provided

Table 4: Evidence for an association of MRI measures with disability

CIS cohorts	**5 – 7 years**	**8 – 14 years**	**15 – 20 years**
T2 lesion number	+	+/-	*n.r.*
T2 lesion volume	+/-	+/-	+/-
10 or more T2 lesions	+/-		n.r.
Change in lesion number	+/-*	+/-	n.r.
Change in lesion volume	+/-	+/-	+/-
Infratentorial lesion number	n.r.	+/-	n.r.
Infratentorial lesion volume	+/-	n.r.	n.r.
ON only cohorts	**5 – 7 years**	**8 – 14 years**	**15 years**
T2 lesion number	+/-	--	--
T2 lesion volume	n.r.	n.r.	n.r.
10 or more T2 lesions	n.r.	n.r.	n.r.
Change in lesion volume	n.r.	n.r.	n.r.
Infratentorial lesions	+/-	n.r.	n.r.
Infratentorial lesion volume	n.r.	n.r.	n.r.

+ = moderate evidence of effect, +/- = limited evidence of effect, -- = no evidence (22), n.r. = not reported (data was only considered, when p-values or confidence intervals were reported); *the number of new lesions over the first 5 years correlated with the change in EDSS over that time period

Table 5: Main findings: MRI lesions as a prognostic factor for disability (EDSS ≥3)

CIS cohorts	**Normal MRI**	**Abnormal MRI**	**Follow-up in years**
Barcelona: 2006 (25)	3 of 52 (6%)	20 of 104 (19%)	7
Barcelona: 2015 (5)	12 of 299 (4%)	106 of 652 (16%)	7
London I: 1993 (23)	0 of 32 (0%)	18 of 57 (32%)	5
London I: 1998 (24)	0 of 27 (0%)	19 of 54 (43%)	10
London I: 2002 (28)	0 of 21 (0%)	24 of 50 (48%)	14
London I: 2008 (6)	9 of 34 (26%)	38 of 73 (52%)	20
ON only cohorts°	**Normal MRI**	**Abnormal MRI**	**Follow-up in years**
ONTT: 1997* (30)	2 of 30 (7%)	12 of 62 (19%)	5
ONTT: 2004* (29)	10 of 34 (29%)	30 of 80 (38%)	10-12
ONTT: 2008* (31)	11 of 28 (39%)	24 of 77 (31%)	15

Indicated are number of people (percentages) presenting with EDSS >= 3, * = only calculated in people with CDMS. °London II data were not comparable as they were subgrouped in 0-4 lesions.

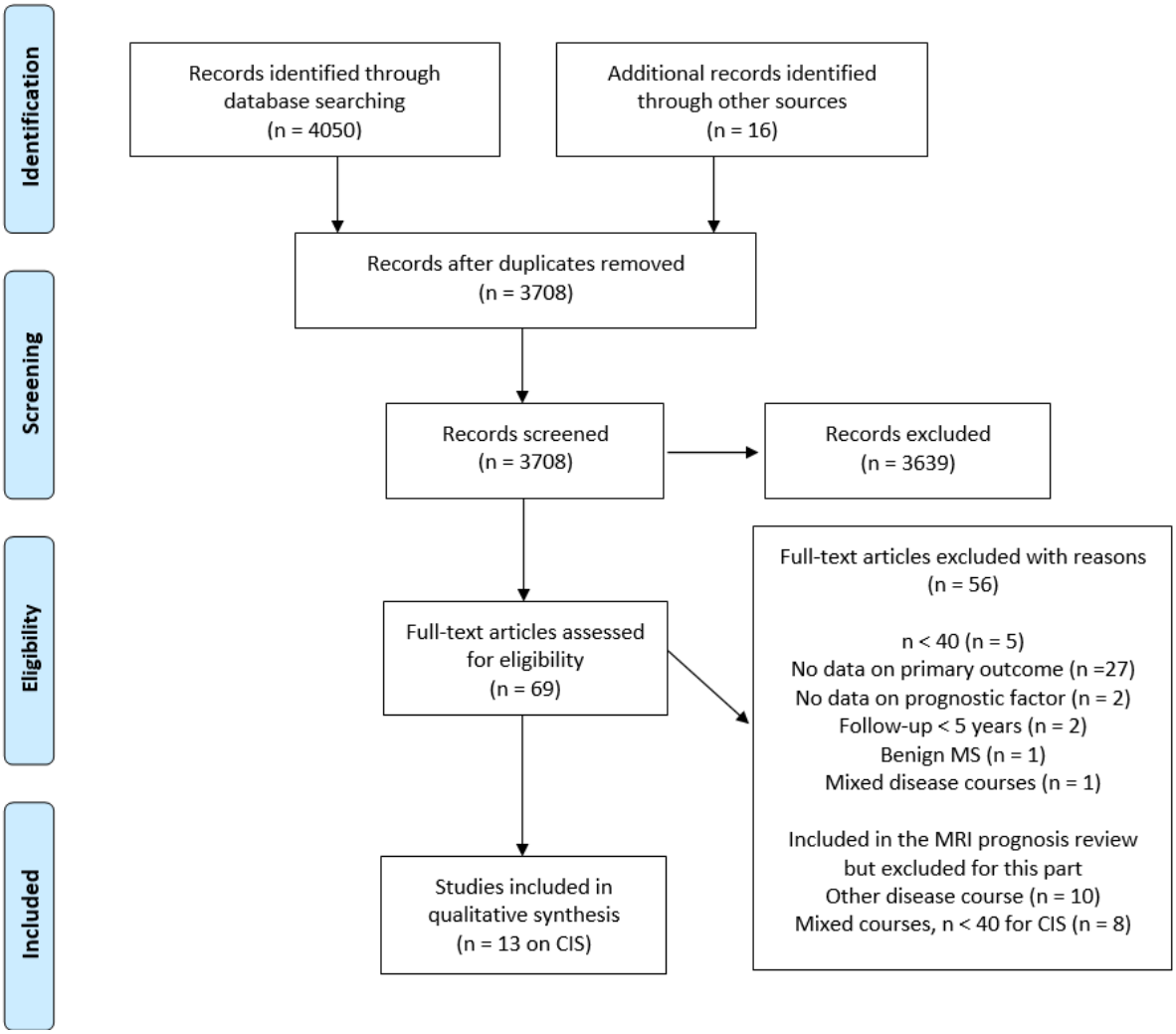

Figure 1: PRISMA flow diagram (12)