The clinico-radiological paradox of cognitive function and MRI burden of white matter lesions in people with multiple sclerosis: a systematic review and meta-analysis.

PLOS ONE

RESEARCH ARTICLE

The clinico-radiological paradox of cognitive function and MRI burden of white matter lesions in people with multiple sclerosis: A systematic review and meta-analysis

Daisy Mollison[1], Robin Sellar[1], Mark Bastin[1], Denis Mollison[2], Siddharthan Chandran[1], Joanna Wardlaw[1], Peter Connick[1]*

1 Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, United Kingdom, **2** Department of Actuarial Mathematics and Statistics, Heriot-Watt University, Edinburgh, United Kingdom

* pconnick@exseed.ed.ac.uk

OPEN ACCESS

Citation: Mollison D, Sellar R, Bastin M, Mollison D, Chandran S, Wardlaw J, et al. (2017) The clinico-radiological paradox of cognitive function and MRI burden of white matter lesions in people with multiple sclerosis: A systematic review and meta-analysis. PLoS ONE 12(5): e0177727. https://doi.org/10.1371/journal.pone.0177727

Editor: Orhan Aktas, Heinrich-Heine-Universitat Dusseldorf, GERMANY

Received: January 26, 2017

Accepted: May 2, 2017

Published: May 15, 2017

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: Daisy Mollison is funded by a Rowling Scholars Clinical Academic Fellowship, http://annerowlingclinic.com/training.html. PC is funded by the Wellcome Trust, https://wellcome.ac.uk/funding, award number: 106716/Z/14/Z. The funders had no role in study design, data collection

Abstract

Background

Moderate correlation exists between the imaging quantification of brain white matter lesions and cognitive performance in people with multiple sclerosis (MS). This may reflect the greater importance of other features, including subvisible pathology, or methodological limitations of the primary literature.

Objectives

To summarise the cognitive clinico-radiological paradox and explore the potential methodological factors that could influence the assessment of this relationship.

Methods

Systematic review and meta-analysis of primary research relating cognitive function to white matter lesion burden.

Results

Fifty papers met eligibility criteria for review, and meta-analysis of overall results was possible in thirty-two (2050 participants). Aggregate correlation between cognition and T2 lesion burden was $r = -0.30$ (95% confidence interval: -0.34, -0.26). Wide methodological variability was seen, particularly related to key factors in the cognitive data capture and image analysis techniques.

Conclusions

Resolving the persistent clinico-radiological paradox will likely require simultaneous evaluation of multiple components of the complex pathology using optimum measurement techniques for both cognitive and MRI feature quantification. We recommend a consensus

and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

initiative to support common standards for image analysis in MS, enabling benchmarking while also supporting ongoing innovation.

Introduction

Cognitive impairment is seen in 43–70% of people with multiple sclerosis (MS), exhibiting a variable pattern of deficits between individuals [1]. The most frequently detected deficits include a reduction in information processing speed, executive functions, attention, and long-term memory. Impairment of information processing may represent the core cognitive deficit [2], consistent with the model of a disconnection syndrome [3]. The underlying pathology is complex, including both focal and diffuse abnormalities of the central nervous system that affect both white and grey matter structures [4]. Components of this pathology have been increasingly amenable to in vivo quantification through magnetic resonance imaging (MRI) and associated image analysis techniques [5]. The impact of pathology on phenotype is also influenced by lifetime intellectual enrichment ('cognitive reserve') [6], lifestyle variability (cognitive leisure), as well as comorbidities, ageing, and medications.

Although aspects of the 'global' pathological burden affecting the brains of people with MS can be readily estimated by abnormalities such as T2 hyperintense lesions that are visible on structural MRI, limited correlation exists between these measures and the clinical phenotype. This has been termed the 'clinico-radiological paradox' (CRP) and is well described for both physical and cognitive impairments [7]. The CRP presents a fundamental challenge with respect to mechanistic understanding of the relationship between pathology and phenotype in MS, and to the use of MRI metrics in clinical decision-making at the individual-subject level. Several explanations have therefore been proposed in order to resolve the CRP, including recognition that summation of whole-brain metrics fails to account for variability between subjects in the spatial patterning of multifocal pathology [8]. Although such consideration has a clear corollary through the fundamental principles of localisation in clinical neurology, it provides a less satisfactory explanation with respect to cognitive impairments, particularly for cognitive functions such as information processing speed where the functional neuroanatomy involves widespread connectivity between brain regions [9, 10].

Additional potential contributors to the 'cognitive CRP' include fundamental issues with the evaluation of cognition such as whether any existing test can isolate and quantify a neuroanatomically distributed cognitive function. Or, whether multidimensional cognitive assessment through 'cognitive batteries' provides a valid quantitative assessment of 'global' cognitive performance. Such 'global' evaluation may be contingent on pre-requisite and relatively localised functions such as sustained attention that are subject to the potential confound of variable spatial patterning. Although mitigated by the use of batteries with normative data simultaneously developed for component tests, judgement with respect to the latter remains a fundamental principle applied in clinical neuropsychology [11]. Separately, critical issues also arise in the quantification of total pathological burden by structural brain MR imaging. These include the known insensitivity of existing metrics for potentially critical aspects of disease pathology such as grey matter lesions [12] and a failure to adequately quantify neuroaxonal loss or underlying subvisible/diffuse pathology. Critically, these aspects of the complex pathology may be independent at the individual-subject level [13], rendering assessment of T2 hyperintense lesion burden as an inadequate account of the total (multifaceted) pathological burden.

In addition, modest correlation of total MRI-visible white matter lesion burden to cognitive status may reflect attenuation due to psychometric limitations in the methodologies for quantification of both cognitive and MRI features [14], as well as aspects of study design including participant selection. We therefore performed a systematic review and meta-analysis of the published literature describing the relationship between cognitive function and the total burden of white matter pathology detected by standard structural brain MRI. Our aim was first to confirm the modest correlations that have been previously described [15], and second to explore the potential methodological issues that may affect the observed relationship.

Methods

Design of the systematic review, meta-analysis, and manuscript was based on PRISMA ('Preferred Reporting Items for Systematic Reviews and Meta-Analyses') guidelines [16].

Protocol, information sources and search strategy

The study protocol was documented in advance. Medline, Embase, and Web of Science databases were searched for English language papers on 1st July 2015, with no date restrictions (S1 Appendix). Review articles were excluded, but relevant reviews published in the last 10 years were screened for references. Archives of the journals Neurology, Multiple Sclerosis and the American Journal of Neuroradiology were also hand-searched for relevant articles published in the previous ten years. Search terms were: 'magnetic resonance imaging', 'multiple sclerosis', 'cognitive', 'cognition', related terms and abbreviations of these.

Study selection and eligibility criteria

Initial screening of abstracts was performed by a single author (DM). Full articles were then retrieved and eligibility assessment performed in a standardized manner, with a final decision over study inclusion taken in consensus with a second reviewer (PC). Eligibility criteria were: English language and peer-reviewed publications reporting data from adults with clinically definite MS as primary research with a primary aim of relating cognition to T1w, T2w, FLAIR or PD metrics of total brain white matter lesion burden. Imaging outcomes for total lesion volume or area, and lesion counts or scores, were all accepted as valid measures of whole brain lesion burden. Similarly, any measure of cognitive function with face-validity was accepted. Studies were excluded if reporting exploratory or secondary analysis, or if lesion burden was only related to longitudinal change in cognitive function. Where studies examined both cross-sectional and longitudinal outcomes, the baseline cross-sectional analyses were used. When overlap of reported cohorts was identified and clarification from the original investigators was not possible, a conservative approach was adopted with inclusion of only the earliest dated relevant article. Studies within the systematic review were suitable for meta-analysis if they reported an overall effect for the relationship of imaging metrics to a single measure of cognition defined by either a single cognitive test, or a summary result from a cognitive battery.

Data collection

Data was extracted by a single author (DM) using a standardized form that captured (1) characteristics of the participants, including age, sex and disease phenotype; (2) cognitive testing methods including blinding and identity of the tester; (3) image acquisition methods; (4) image analysis methods including training and blinding of investigators, software tools used,

whether measures of intra- and inter-rater reliability were provided; and (5) statistical analysis methods including controlling for potential confounding factors. A study quality assessment tool (S2 Appendix) was also developed based on STROBE ('Strengthening the Reporting of Observational studies in Epidemiology') guidelines [17] to evaluate the risk of bias in individual studies. The authors for one paper were contacted for further information and numerical data was provided.

Summary measures and synthesis of results

Summary measures were recorded if relating MRI metrics to an overall measure of cognitive function or to a single cognitive test. Where summary measures were provided both unadjusted and adjusted for potentially confounding clinical covariates, adjusted results were used. Correlation coefficients or the difference in lesion burden between groups defined by cognitive status were accepted as summary measures, with preference given to correlations if both were available [18]. All reported summary measures were converted into effect sizes and inverted as necessary so that negative values always indicated an association of lower cognitive scores to higher lesion burdens. Standardized mean differences were calculated from studies reporting group comparisons, prior to conversion to equivalent correlations (r). An approximation to the standard deviation was estimated as necessary based on available measures of dispersion (e.g. interquartile range or range). In studies with two impaired groups defined by specific cognitive deficits, these groups were combined before calculation of a standardized difference from a non-impaired group. The Fisher's z transformation was used prior to calculation of an aggregate summary effect, with conversion back to r for reporting of overall meta-analysis findings and confidence intervals.

An aggregate summary effect was calculated using maximum likelihood estimation [19], taking into account the size of the various studies; this method allows incorporation of those studies reporting non-significant results without providing their estimate. Separate analyses were carried out for studies measuring hyperintense lesion burden on T2w, FLAIR and/or PD sequences, and for the subgroup of studies evaluating T1w hypointense lesion volume. Heterogeneity was assessed by Cochran's Q and the I^2 statistic [20], based on the studies providing specific estimates of the effect size. All analyses were performed using the statistical software *R*, version 3.2.4.

Risk of bias across studies

Our eligibility criteria required a stated primary aim to evaluate the relationship between cognitive status and brain imaging metrics so that we might minimize the influence of reporting bias from *post hoc* analyses. Within the included studies, we recorded analyses that were described without results being provided. A funnel plot was also evaluated visually and tested formally using Egger's regression test for asymmetry.

Additional analyses

An alternative aggregate effect size was calculated using quality scores as an additional scaling factor. A sensitivity analysis examining the effect of using an alternative random-effects model, with DerSimonian and Laird methodology, was carried out for all studies providing data compatible with precise estimates of the effect size. Subgroup analyses of studies using the Paced Auditory Serial Addition Test (PASAT) and Symbol Digit Modalities Test (SDMT) were also pre-specified to investigate whether focusing on distributed cognitive function would improve correlations with overall lesion burden and replicate previous findings [15].

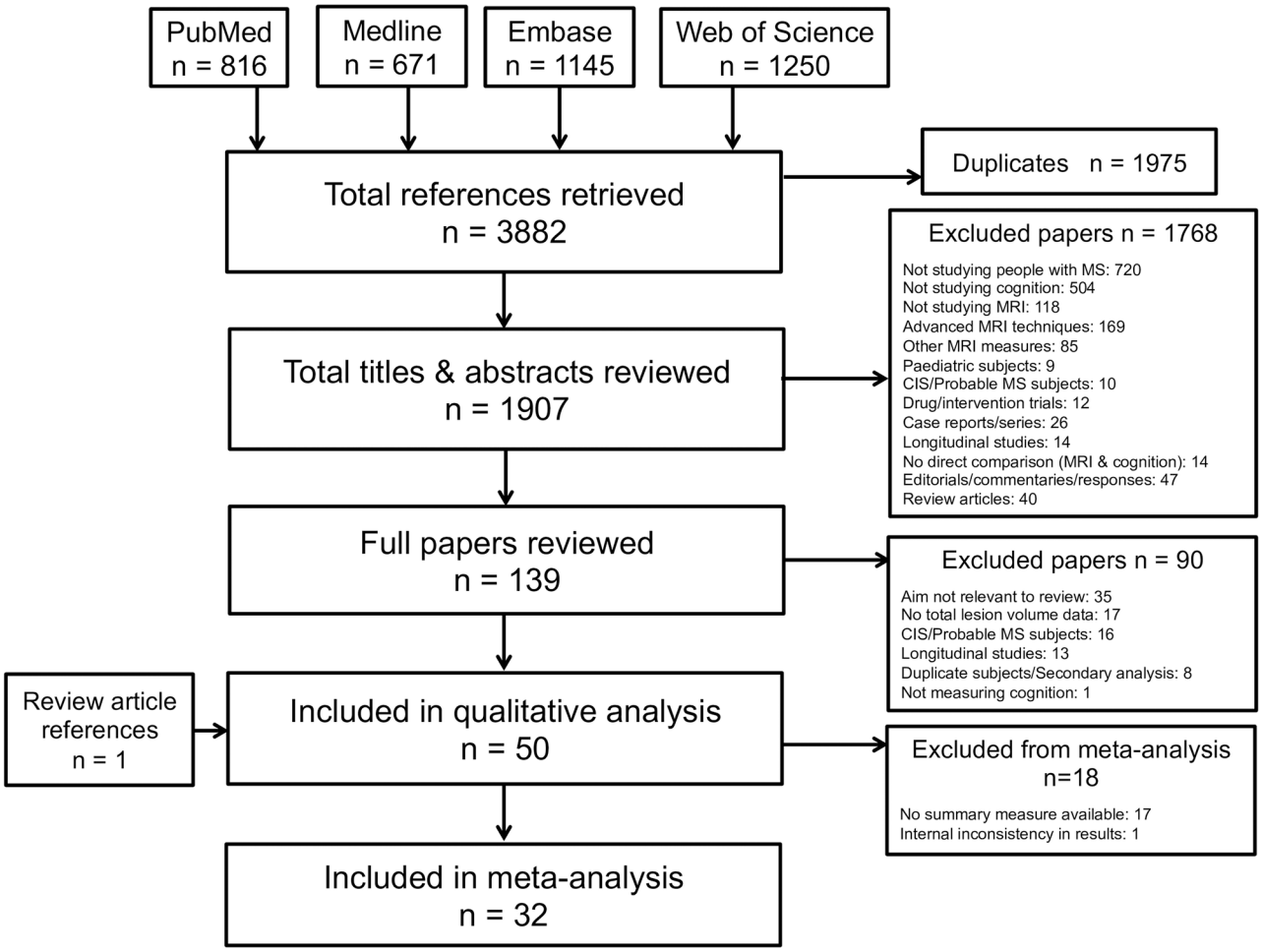


Fig 1. Flowchart showing articles retrieved and considered at each stage of the review process.

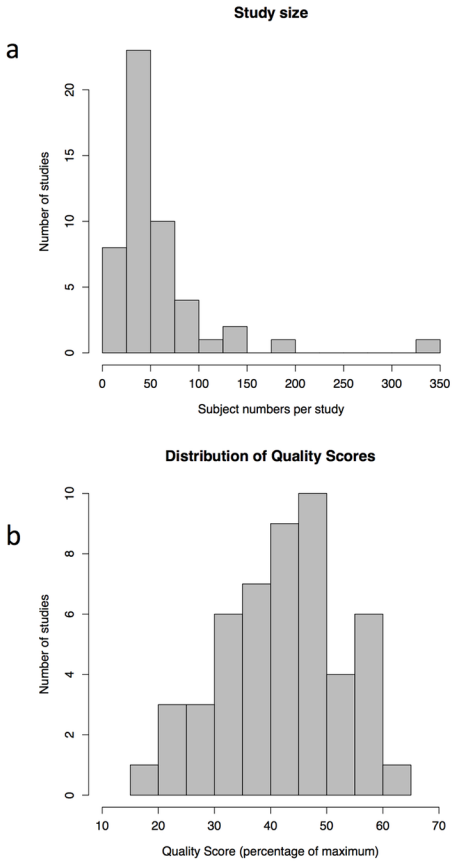
https://doi.org/10.1371/journal.pone.0177727.g001

Results

Study selection

A total of 3882 studies were identified by the initial literature search, 1975 of which were duplicates (Fig 1). Year-on-year increases were seen in the publication rate identified through the initial search (S1 Fig). No additional studies were included following hand searching of journal archives. After review of abstracts, 139 manuscripts were retrieved. Ninety were subsequently excluded, most frequently (35/90 = 39%) because the study aim was not relevant. A total of fifty papers met all inclusion criteria [21–70] spanning the period 1987–2015.

Thirty studies provided usable summary measures relating hyperintense T2w/FLAIR/PD lesion burden to cognitive function. Two studies reported a 'non-significant' result and one study was excluded from meta-analysis as the reported summary measure was internally inconsistent with other reported results and significance levels. The remaining seventeen studies did not provide results suitable for use in meta-analysis, reporting only individual results for each cognitive subtest (n = 12) or multiple regression modeling with simultaneous assessment of several brain imaging metrics (n = 5). Thirteen studies related cognition to T1



c

Information reported	Studies gaining each mark (%)		
	1	0.5	0
Eligibility criteria	60	22	18
Individual outcome variables results	52	40	8
Overall outcome results with precision	46	40	14
Quantitative variable handling	42	48	10
Recruitment pattern	40	-	60
Participant characteristics	38	48	14
Statistical methodology	32	50	18
Blinding of assessors	18	50	32
Participant dropout	14	-	86
Objective clearly stated	8	50	42
Cognitive testing & imaging delay	6	58	36
Study design specified	6	10	84
Clearly defined outcomes	4	88	8
Potential confounding factors	0	64	36
Measurement methodology	0	56	44
Study size rationale	0	-	100

Fig 2. Showing factors relevant to study quality including histograms of a) numbers of participants with MS in individual studies and b) overall quality scores, and c) the reporting of individual factors contributing to the overall quality score.

https://doi.org/10.1371/journal.pone.0177727.g002

hypointense lesion burden, of which eleven provided usable summary measures and two reported 'non-significant' results.

Participant characteristics

The total number of subjects from all included studies was 2891. Individual study size ranged from 17 to 327 participants (mean 58, median 45; Fig 2a). Forty-four studies specified the sex ratio, all but one having a female majority. The range of mean participant age (provided in 47/50 studies) was 31–55 years. No study used age of disease onset in its eligibility criteria. Twenty-six studies included participants with a mixture of disease courses; thirteen studies recruited exclusively relapsing-remitting disease, six studies progressive disease, two 'benign', and three did not specify the participants' disease course.

MR brain imaging acquisition

The majority (29/50 studies) used 1.5T scanners. Ten studies used scanners with below 1.5T magnets for some or all participants' imaging, seven used 3T scanners, one used both 1.5 and 3T scanners and three did not specify the scanner field strength. Details of the imaging protocol were given in all but seven studies.

Image analysis

The sequence(s) used to measure lesion volume was specified in forty-three studies. Twenty-six specified the number of people involved in the lesion analysis, a single observer in fourteen. The anatomical boundaries of evaluation were explicitly defined in two studies and a sample image was provided by five studies. A wide variety of approaches were used for the quantification of lesion burden. These included lesion counts (two studies) or weighted lesion scores (six studies), manual lesion outlining either on hard copies (two studies) or within viewing software (six studies), and the use of semi-automated methods (thirty-one studies). Of the six studies using lesion scores, five different scoring systems were used. One study used both manual and semi-automated measurements (for different sequences), one used manual lesion outlining and an absolute lesion count, and in one study the methodology was unclear. In the thirty-two studies using semi-automated measurement tools, the software used was specified or references provided in 25 studies (78%), covering fourteen different software packages. In eighteen of these studies the named software was publically available (eleven different softwares). The remaining studies did not specify their software. A manual editing stage for software-generated lesion masks was specified in five studies (16%) and the person performing this was described in two studies. In the ten studies using fully manual lesion outlining, the person performing this was described in six. Only two studies provided an indication of inter-observer agreement and one study intra-observer reproducibility. Seven studies mentioned previous measures of reproducibility or results on training data sets. Only five percent of studies calculating a lesion volume or area (2/42) normalized to intracranial volume.

Cognitive testing

The cognitive assessor and their training was unclear in thirty-eight studies. Of defined batteries, the most commonly used was Rao's Brief Repeatable Battery (12/50), followed by the Minimal Assessment of Cognitive Function in MS (5/50), used with modifications or additional tests in eight (67%) and two (40%) studies respectively. Unique collections of tests were found in twenty-seven studies. The SDMT or PASAT were used either exclusively or as part of a wider battery in thirty studies. Substantial variability was seen in how raw cognitive scores were processed prior to their use in the evaluation of a possible relationship with imaging metrics. Methods included use of unadjusted scores, standardization, and the deployment of group classifiers. Standardization was performed using either historic- (published or unpublished) or contemporary- (matched or unmatched for participant characteristics) control data. Group classifiers were either based on internal (patient) or external (normative) reference cohorts. The specific thresholds used to define impairment on individual tests were also variable, including 1, 1.5, and 2 standard deviations from the reference mean, and those based on centiles. Moreover, the number of failed tests used to define overall cognitive impairment was also variable (S1 Table). Consideration of the effect of potential confounders also varied between studies, both in the recording of relevant data and whether it was adjusted for in the analysis. Some studies adjusted for age (n = 18), sex (n = 12), education level (n = 13) and/or affective disorders (n = 15). Drug treatments and premorbid IQ were both adjusted for in three studies. Cognitive leisure activities were neither measured nor adjusted for in any study.

Statistical analysis

Summary measures were provided through univariate correlations (n = 37) and/ or group comparisons based on cognitive status (n = 24). Four studies divided participants into groups dependent on radiological features. Fourteen studies constructed statistical models predicting

cognitive performance based on imaging and other laboratory, demographic, or clinical markers.

Reporting quality and risk of bias within studies

A range of study-specific quality scores was seen (mean 42%, SD 11%; Fig 2b). Among individual elements of the composite quality score, complete reporting was provided most frequently for eligibility criteria and outcome measures (Fig 2c). In contrast, no study provided complete reporting of potential confounding factors, measurement methodology, or study size rationale.

Results of individual studies

Studies directly reporting correlation coefficients relating cognitive performance to T2 hyperintense lesion burden ranged from -0.6 to -0.23. Standardised mean differences ranged from -2.70 to 0.23, equivalent to correlations of -0.80 to 0.11.

Synthesis of results

The aggregate effect size relating cognitive performance to T2 hyperintense lesion burden was r = -0.30 (95% confidence interval: -0.34 to -0.26; Fig 3). There was evidence of possible heterogeneity (Q = 43.62, df = 29, p = 0.04; I^2 = 33.5%). The aggregate effect size relating cognitive performance to T1 hyperintense lesion burden was r = -0.26 (95% CI: -0.32, -0.20; Q = 20.4, df = 10, p = 0.025, I^2 = 51.0%, see S3 Appendix for further details).

Risk of bias across studies

Funnel plot inspection (S2 Fig) and Egger's test of asymmetry gave an equivocal result (p = 0.05). We therefore explored possible underlying sources of heterogeneity [71]. Reporting biases could not be evaluated as study protocols were not published prospectively. Despite methodological heterogeneity apparent from our quality scoring, no correlation was seen between overall quality score and effect size (r = -0.18, p = 0.34). Exploratory meta-analysis using quality scores as an additional weighting factor returned an effect size similar to that of our primary analysis (r = -0.30; 95% CI: -0.36, -0.24). In order to explore the possibility of 'true heterogeneity' between study effect sizes, we performed a sensitivity meta-analysis using a random effects model, giving an overall effect size similar to that of our primary analysis (r = -0.33; 95% CI: -0.38, -0.27). Further sensitivity analyses, comparing scanner field strength and type of lesion quantification method did not demonstrate a measurable subgroup difference in heterogeneity from the small number of studies using high (3T) or low (below 1T) field scanners, or from those using lesion counts or scores.

Additional analyses

Alternative cognitive endpoints. Exploratory meta-analyses were performed on two widely used measures of information processing speed (IPS), the SDMT and PASAT (S4 & S5 Appendices). Our *a priori* hypothesis was that total lesion burden would have a stronger correlation with these tests of distributed cognition function (IPS) compared to the mixture of distributed and localized functions in our primary analysis. The summary effect size for SDMT was r = - 0.37 (95% CI: -0.43, -0.31; n = 13 studies) and for PASAT was r = - 0.28 (95% CI: -0.34, -0.22; n = 15 studies).

Study	Size	Quality score
Huber 1987	32	3
Franklin 1988	60	4
Pozzilli 1991	17	6
Ron 1991	58	5
Patti 1995	26	5.5
Rovaris 1998	30	9
Camp 1999	63	6
Comi 1999	22	8.5
Snyder 2001	41	7.5
Kalkers 2001	134	7.5
Nocentini 2001	44	4.5
Zivadinov 2001	63	9.5
Bermel 2002	23	7
Christodoulou 2003	37	7
Lazeron 2005	82	6
Benedict 2006	82	6.5
Amato 2008	47	8.5
Lin 2008	36	8
Karlinska 2008	60	4.5
Rovaris 2008	62	9.5
Sanchez 2008	52	8.5
Krause 2009	22	6
Patti 2009	327	9.5
Heesen 2010	50	7
Akbar 2010	62	9.5
Lund 2012	50	4.5
Mesaros 2012	82	7.5
Rossi 2012	142	8
Francis 2013	45	7.5
Laffon 2014	75	5
Yildiz 2014	78	7.5
Sacco 2015	46	6.5
Aggregate effect size		

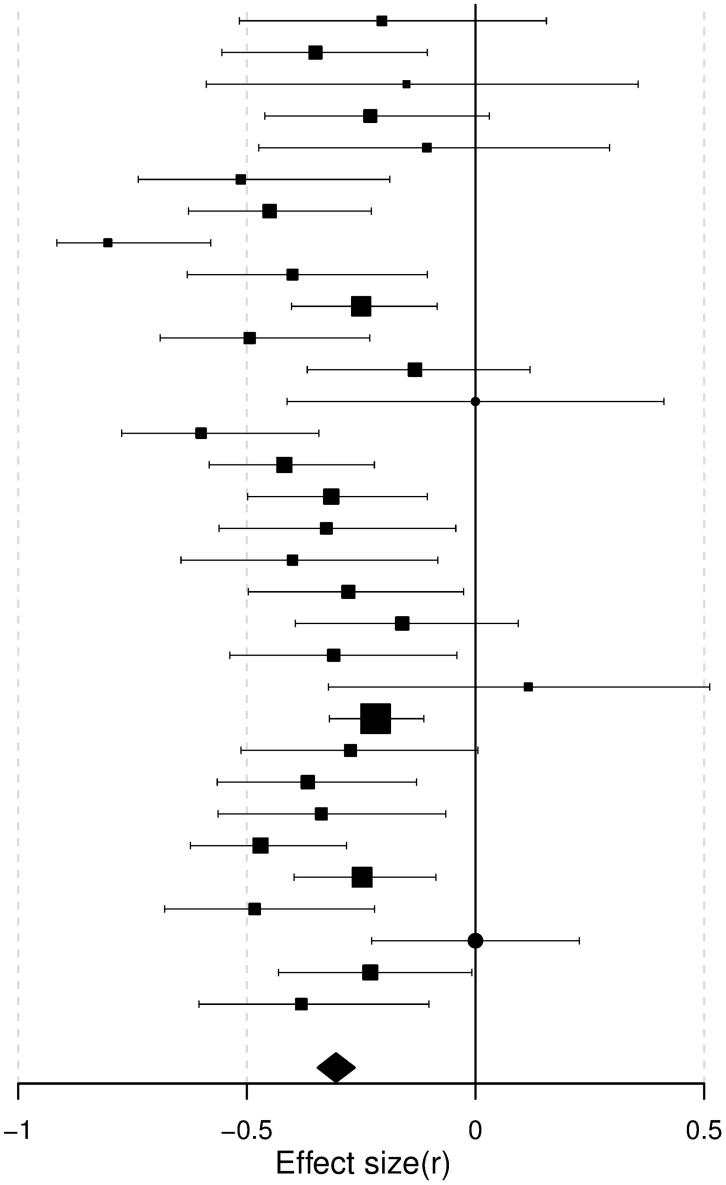


Fig 3. Forest plot of the individual studies showing their effect sizes as correlation coefficients. Box sizes are inversely proportional to study variance. Manuscripts reporting "non-significant" results without a point estimate are represented by circles. Aggregate effect size: r = -0.30; 95% confidence interval: -0.34, -0.26.

https://doi.org/10.1371/journal.pone.0177727.g003

Discussion

Our results confirm a modest correlation (r = - 0.30) between MRI measures of total brain white matter lesions and cognitive function in people with MS. Although some variability was observed between studies in the magnitude of the reported relationship, no large (>100 participants) single study demonstrated a strong correlation. We therefore sought to explore whether technical and methodological factors may have been important in attenuating the reported correlation.

Substantial variability was seen with respect to study design, including the approaches used to quantify both cognitive function (see review by Fischer *et al* [72]) and lesion burden, and the adjustment for other variables that influence cognition (e.g. education, premorbid IQ and drugs). For cognitive assessment, this may represent a largely historic issue as a global movement is now established to harmonise evaluation and scoring through the Brief International Cognitive Assessment for MS (BICAMS) initiative [73]. In contrast, the optimum method to generate quantifiable measures of lesion burden from brain imaging data lacks emergent consensus. Recent initiatives to harmonise MR acquisition protocols are welcome [74, 75], however no similar initiative exists for image analysis techniques. Semi-automated approaches were the most frequently used (62%) and therefore merit particular consideration. While effective manual editing is clearly dependent on adequate training of the operator, the automated (software) component is more challenging to benchmark. We recommend that authors should routinely report the software used. Separately, the field risks delaying progress and reducing the potential for collaboration due to the many differing software packages used. Of the twenty-four studies naming software, ten different publicly available (commercial or open source) packages were used, and a further three packages that were developed 'in house'. To our knowledge, no comparative study has been performed on a common dataset to evaluate agreement between these varied approaches. It is therefore our view that a new consensus initiative is required to support an image analysis framework in MS that enables benchmarking while also supporting ongoing innovation.

Despite our finding of substantial methodological variability between studies, formal testing for heterogeneity in our primary meta-analysis returned an equivocal result. This indicates that methodological variability between studies cannot provide a sufficient explanation for the cognitive-CRP. However, measurement errors within all published studies may have attenuated observed correlations in the face of a higher 'true' correlation [14]. Greater recognition and transparency around measurement error for both cognitive and lesion-burden quantification would therefore be beneficial to the field.

As previously noted, resolving the cognitive-CRP may require consideration of the spatial patterning of lesions with simultaneous evaluation of other aspects of MS pathology that may be both phenotypically relevant and independent from the burden of white matter hyperintensities [76, 77]. However, a further potential contributor is the pathological variability of white matter T2 hyperintense lesions. Conventional MRI is unable to distinguish the extent of intralesional inflammatory infiltrate, demyelination, remyelination, axonal damage, or gliosis [78]. If cognitive impairment reflects only some of these pathological features, then the remainder will contribute only measurement error; improved MR-based quantification of individual lesion characteristics may therefore be critical.

Our findings may have been limited by an overly inclusive approach to both the evaluation of cognition and white matter lesion burden. With respect to the former, we saw a higher aggregate correlation between white matter lesion burden and cognition measured by the SDMT–a measure of information processing speed, understood to reflect widely distributed brain connectivity–than was seen for cognition as defined in the primary analysis. Notably, relatively few studies in our review used >1.5T field-strength scanners, in part reflecting the recent shift away from exploring the relationship between phenotype and T2 hyperintense lesion burden, focusing instead on the possible relevance of other MR metrics. We therefore interpret our sensitivity analysis for the effect of magnet field strength to lack sufficient data for a definitive conclusion. We would encourage re-evaluation of this relationship as the literature evolves with respect to 3T acquisition and/or in cohorts scanned on both low and high magnetic field scanners. Finally, a substantial body of potentially relevant data was excluded from our review as the primary aim of the study was unclear or reported findings were

secondary/exploratory analyses. Finally, despite our best efforts to apply a systematic approach, all reviews are conducted by researchers who bring unconscious bias [79].

Conclusions

We replicate the finding that modest correlation (r = - 0.30) exists between MRI measures of total brain white matter lesion burden and cognitive function in people with MS. However, the quantification techniques for both cognitive and MR features were highly variable, and this may have attenuated the observed strength of the association. An accurate assessment of the relationship requires optimum measurement techniques; this is a prerequisite to meaningful investigation of the clinico-radiological paradox through simultaneous evaluation of multiple components of the complex pathology. We therefore make the following recommendations:

1. A new consensus initiative is advanced to support an image analysis framework in MS that enables benchmarking while also supporting ongoing innovation.
2. Greater recognition and transparency are fostered around measurement error for both cognitive and lesion burden quantification.
3. Reporting of observational research adheres to best-practice guidance such as provided by the STROBE statement.
4. Attempts to resolve the cognitive clinico-radiological paradox should adopt a more multidimensional approach to understanding white matter lesions with simultaneous consideration to multiple elements of the quantifiable pathology, whilst also incorporating potential clinical confounders of the relationship.

Supporting information

S1 Appendix. Record of database search.
(DOCX)

S2 Appendix. Quality assessment tool for evaluation of manuscripts, based on the STROBE checklist.
(DOCX)

S3 Appendix. Sub-analysis of studies relating T1 hypointense lesion volume to overall cognitive performance.
(DOCX)

S4 Appendix. Sub-analysis of studies relating T2 hyperintense lesion volume to Symbol Digit Modalities Test (SDMT) performance.
(DOCX)

S5 Appendix. Sub-analysis of studies relating T2 hyperintense lesion volume to Paced Auditory Serial Additions Test (PASAT) performance.
(DOCX)

S6 Appendix. Database containing relevant data extracted from primary literature.
(XLSX)

S7 Appendix. PRISMA checklist.
(DOC)

S8 Appendix. PRISMA flow diagram.
(DOC)

S1 Fig. Number of results retrieved from database search by year of publication. The 2015 point is an extrapolated value from the 6-month figure.
(EPS)

S2 Fig. Funnel plot of effect sizes, on Fisher's z scale, against the inverse of their standard error (itself inversely related to study size) with asymmetry towards increased reporting of stronger correlations for smaller study sizes. The vertical dashed line indicates the summary effect on the same scale (z = -0.32).
(EPS)

S1 Table. Summarising cognitive testing and scoring protocols.
(DOCX)

Acknowledgments

Our thanks to Dr Charles Guttmann for providing additional data.

Author Contributions

Conceptualization: Daisy Mollison RS MB SC JW PC.

Data curation: Daisy Mollison.

Formal analysis: Daisy Mollison Denis Mollison JW PC.

Funding acquisition: Daisy Mollison PC.

Investigation: Daisy Mollison PC.

Methodology: Daisy Mollison RS PC.

Project administration: Daisy Mollison PC.

Supervision: RS MB SC JW PC.

Visualization: Daisy Mollison.

Writing – original draft: Daisy Mollison PC.

Writing – review & editing: Daisy Mollison Denis Mollison RS MB SC JW PC.

References

1. Chiaravalloti ND, DeLuca J. Cognitive impairment in multiple sclerosis. Lancet Neurol. 2008; 7(12): 1139–51. https://doi.org/10.1016/S1474-4422(08)70259-X PMID: 19007738
2. Denney DR, Lynch SG, Parmenter BA, Horne N. Cognitive impairment in relapsing and primary progressive multiple sclerosis: mostly a matter of speed. J Int Neuropsychol Soc. 2004; 10(7):948–56. PMID: 15803558
3. Dineen RA, Vilisaar J, Hlinka J, Bradshaw CM, Morgan PS, Constantinescu CS, et al. Disconnection as a mechanism for cognitive dysfunction in multiple sclerosis. Brain. 2009; 132(Pt 1):239–49. https://doi.org/10.1093/brain/awn275 PMID: 18953055
4. DeLuca GC, Yates RL, Beale H, Morrow SA. Cognitive impairment in multiple sclerosis: clinical, radiologic and pathologic insights. Brain Pathol. 2015; 25(1):79–98. https://doi.org/10.1111/bpa.12220 PMID: 25521179

5. Rocca MA, Amato MP, De Stefano N, Enzinger C, Geurts JJ, Penner IK, et al. Clinical and imaging assessment of cognitive dysfunction in multiple sclerosis. Lancet Neurol. 2015; 14(3):302–17. https://doi.org/10.1016/S1474-4422(14)70250-9 PMID: 25662900
6. Sumowski JF. Cognitive Reserve as a Useful Concept for Early Intervention Research in Multiple Sclerosis. Front Neurol. 2015; 6:176. https://doi.org/10.3389/fneur.2015.00176 PMID: 26347706
7. Barkhof F. The clinico-radiological paradox in multiple sclerosis revisited. Curr Opin Neurol. 2002; 15(3):239–45. PMID: 12045719
8. Hackmack K, Weygandt M, Wuerfel J, Pfueller CF, Bellmann-Strobl J, Paul F, et al. Can we overcome the 'clinico-radiological paradox' in multiple sclerosis? J Neurol. 2012; 259(10):2151–60. https://doi.org/10.1007/s00415-012-6475-9 PMID: 22446893
9. Leavitt VM, Wylie G, Genova HM, Chiaravalloti ND, DeLuca J. Altered effective connectivity during performance of an information processing speed task in multiple sclerosis. Mult Scler. 2012; 18(4):409–17. https://doi.org/10.1177/1352458511423651 PMID: 21965419
10. Costa SL, Genova HM, DeLuca J, Chiaravalloti ND. Information processing speed in multiple sclerosis: Past, present, and future. Mult Scler. 2016.
11. Harvey PD. Clinical applications of neuropsychological assessment. Dialogues Clin Neurosci. 2012; 14(1):91–9. PMID: 22577308
12. Hulst HE, Geurts JJ. Gray matter imaging in multiple sclerosis: what have we learned? BMC Neurol. 2011; 11:153. https://doi.org/10.1186/1471-2377-11-153 PMID: 22152037
13. Inglese M, Benedetti B, Filippi M. The relation between MRI measures of inflammation and neurodegeneration in multiple sclerosis. J Neurol Sci. 2005; 233(1–2):15–9. https://doi.org/10.1016/j.jns.2005.03.001 PMID: 15949493
14. Spearman C. The proof and measurement of association between two things. Am J Psychol. 1904; 15(1):72–101.
15. Rao SM, Martin AL, Huelin R, Wissinger E, Khankhel Z, Kim E, et al. Correlations between MRI and Information Processing Speed in MS: A Meta-Analysis. Mult Scler Int. 2014; 2014:975803. https://doi.org/10.1155/2014/975803 PMID: 24795824
16. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med. 2009; 6(7):e1000100. https://doi.org/10.1371/journal.pmed.1000100 PMID: 19621070
17. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol. 2008; 61(4):344–9. https://doi.org/10.1016/j.jclinepi.2007.11.008 PMID: 18313558
18. Borenstein M. Introduction to meta-analysis: Chichester, U.K.: John Wiley \& Sons; 2009.
19. Millar RB. Maximum likelihood estimation and inference: with examples in R, SAS and ADMB: John Wiley & Sons; 2011.
20. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003; 327(7414):557–60. https://doi.org/10.1136/bmj.327.7414.557 PMID: 12958120
21. Akbar N, Lobaugh NJ, O'Connor P, Moradzadeh L, Scott CJ, Feinstein A. Diffusion tensor imaging abnormalities in cognitively impaired multiple sclerosis patients. Can J Neurol Sci. 2010; 37(5):608–14. PMID: 21059506
22. Amato MP, Portaccio E, Stromillo ML, Goretti B, Zipoli V, Siracusa G, et al. Cognitive assessment and quantitative magnetic resonance metrics can help to identify benign multiple sclerosis. Neurology. 2008; 71(9):632–8. https://doi.org/10.1212/01.wnl.0000324621.58447.00 PMID: 18725589
23. Anzola GP, Bevilacqua L, Cappa SF, Capra R, Faglia L, Farina E, et al. Neuropsychological assessment in patients with relapsing-remitting multiple sclerosis and mild functional impairment: correlation with magnetic resonance imaging. J Neurol Neurosurg Psychiatry. 1990; 53(2):142–5. PMID: 2313301
24. Archibald CJ, Wei X, Scott JN, Wallace CJ, Zhang Y, Metz LM, et al. Posterior fossa lesion volume and slowed information processing in multiple sclerosis. Brain. 2004; 127(Pt 7):1526–34. https://doi.org/10.1093/brain/awh167 PMID: 15090476
25. Benedict RH, Bruce JM, Dwyer MG, Abdelrahman N, Hussein S, Weinstock-Guttman B, et al. Neocortical atrophy, third ventricular width, and cognitive dysfunction in multiple sclerosis. Arch Neurol. 2006; 63(9):1301–6. https://doi.org/10.1001/archneur.63.9.1301 PMID: 16966509
26. 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; 61(2):226–30. https://doi.org/10.1001/archneur.61.2.226 PMID: 14967771

27. Bermel RA, Bakshi R, Tjoa C, Puli SR, Jacobs L. Bicaudate ratio as a magnetic resonance imaging marker of brain atrophy in multiple sclerosis. Arch Neurol. 2002; 59(2):275–80. PMID: 11843699

28. Bomboi G, Ikonomidou VN, Pellegrini S, Stern SK, Gallo A, Auh S, et al. Quality and quantity of diffuse and focal white matter disease and cognitive disability of patients with multiple sclerosis. J Neuroimaging. 2011; 21(2):e57–63. https://doi.org/10.1111/j.1552-6569.2010.00488.x PMID: 20626570

29. Camp SJ, Stevenson VL, Thompson AJ, Miller DH, Borras C, Auriacombe S, et al. Cognitive function in primary progressive and transitional progressive multiple sclerosis: a controlled study with MRI correlates. Brain. 1999; 122 (Pt 7):1341–8.

30. Christodoulou C, Krupp LB, Liang Z, Huang W, Melville P, Roque C, et al. Cognitive performance and MR markers of cerebral injury in cognitively impaired MS patients. Neurology. 2003; 60(11):1793–8. PMID: 12796533

31. Comi G, Filippi M, Martinelli V, Campi A, Rodegher M, Alberoni M, et al. Brain MRI correlates of cognitive impairment in primary and secondary progressive multiple sclerosis. J Neurol Sci. 1995; 132(2): 222–7. PMID: 8543952

32. Comi G, Rovaris M, Falautano M, Santuccio G, Martinelli V, Rocca MA, et al. A multiparametric MRI study of frontal lobe dementia in multiple sclerosis. J Neurol Sci. 1999; 171(2):135–44. PMID: 10581380

33. Deloire MSA, Salort E, Bonnet M, Arimone Y, Boudineau M, Amieva H, et al. Cognitive impairment as marker of diffuse brain abnormalities in early relapsing remitting multiple sclerosis. Journal of Neurology, Neurosurgery and Psychiatry. 2005; 76(4):519–26. https://doi.org/10.1136/jnnp.2004.045872 PMID: 15774439

34. Francis PL, Jakubovic R, O'Connor P, Zhang L, Eilaghi A, Lee L, et al. Robust perfusion deficits in cognitively impaired patients with secondary-progressive multiple sclerosis. AJNR Am J Neuroradiol. 2013; 34(1):62–7. https://doi.org/10.3174/ajnr.A3148 PMID: 22700746

35. Franklin GM, Heaton RK, Nelson LM, Filley CM, Seibert C. Correlation of neuropsychological and MRI findings in chronic/progressive multiple sclerosis. Neurology. 1988; 38(12):1826–9. PMID: 3194059

36. Heesen C, Schulz KH, Fiehler J, Von der Mark U, Otte C, Jung R, et al. Correlates of cognitive dysfunction in multiple sclerosis. Brain Behav Immun. 2010; 24(7):1148–55. https://doi.org/10.1016/j.bbi.2010.05.006 PMID: 20621641

37. Hohol MJ, Guttmann CR, Orav J, Mackin GA, Kikinis R, Khoury SJ, et al. Serial neuropsychological assessment and magnetic resonance imaging analysis in multiple sclerosis. Arch Neurol. 1997; 54(8): 1018–25. PMID: 9267977

38. Houtchens MK, Benedict RH, Killiany R, Sharma J, Jaisani Z, Singh B, et al. Thalamic atrophy and cognition in multiple sclerosis. Neurology. 2007; 69(12):1213–23. https://doi.org/10.1212/01.wnl.0000276992.17011.b5 PMID: 17875909

39. Huber SJ, Paulson GW, Shuttleworth EC. Magnetic resonance imaging correlates of dementia in multiple sclerosis. Archives of Neurology. 1987; 44(7):732–6. PMID: 3593063

40. Izquierdo G, Campoy F Jr., Mir J, Gonzalez M, Martinez-Parra C. Memory and learning disturbances in multiple sclerosis. MRI lesions and neuropsychological correlation. Eur J Radiol. 1991; 13(3):220–4. PMID: 1756751

41. Kalkers NF, Bergers L, de Groot V, Lazeron RH, van Walderveen MA, Uitdehaag BM, et al. Concurrent validity of the MS Functional Composite using MRI as a biological disease marker. Neurology. 2001; 56(2):215–9. PMID: 11160958

42. Karlinska I, Siger M, Lewanska M, Selmaj K. Cognitive impairment in patients with relapsing-remitting multiple sclerosis. The correlation with MRI lesion volume. Neurol Neurochir Pol. 2008; 42(5):416–23. PMID: 19105110

43. Krause M, Wendt J, Dressel A, Berneiser J, Kessler C, Hamm AO, et al. Prefrontal function associated with impaired emotion recognition in patients with multiple sclerosis. Behavioural Brain Research. 2009; 205(1):280–5. https://doi.org/10.1016/j.bbr.2009.08.009 PMID: 19686782

44. Laffon M, Malandain G, Joly H, Cohen M, Lebrun C. The HV3 Score: A New Simple Tool to Suspect Cognitive Impairment in Multiple Sclerosis in Clinical Practice. Neurol Ther. 2014; 3(2):113–22. https://doi.org/10.1007/s40120-014-0021-x PMID: 26000227

45. Lazeron RH, Boringa JB, Schouten M, Uitdehaag BM, Bergers E, Lindeboom J, et al. Brain atrophy and lesion load as explaining parameters for cognitive impairment in multiple sclerosis. Mult Scler. 2005; 11(5):524–31. https://doi.org/10.1191/1352458505ms1201oa PMID: 16193889

46. Lazeron RH, de Sonneville LM, Scheltens P, Polman CH, Barkhof F. Cognitive slowing in multiple sclerosis is strongly associated with brain volume reduction. Mult Scler. 2006; 12(6):760–8. https://doi.org/10.1177/1352458506070924 PMID: 17263004

47. Lin X, Tench CR, Morgan PS, Constantinescu CS. Use of combined conventional and quantitative MRI to quantify pathology related to cognitive impairment in multiple sclerosis. J Neurol Neurosurg Psychiatry. 2008; 79(4):437–41. https://doi.org/10.1136/jnnp.2006.112177 PMID: 17673493

48. Lund H, Jonsson A, Andresen J, Rostrup E, Paulson OB, Sorensen PS. Cognitive deficits in multiple sclerosis: correlations with T2 changes in normal appearing brain tissue. Acta Neurol Scand. 2012; 125(5):338–44. https://doi.org/10.1111/j.1600-0404.2011.01574.x PMID: 21793807

49. Mesaros S, Rocca MA, Kacar K, Kostic J, Copetti M, Stosic-Opincal T, et al. Diffusion tensor MRI tractography and cognitive impairment in multiple sclerosis. Neurology. 2012; 78(13):969–75. https://doi.org/10.1212/WNL.0b013e31824d5859 PMID: 22377806

50. Mike A, Glanz BI, Hildenbrand P, Meier D, Bolden K, Liguori M, et al. Identification and clinical impact of multiple sclerosis cortical lesions as assessed by routine 3T MR imaging. AJNR Am J Neuroradiol. 2011; 32(3):515–21. https://doi.org/10.3174/ajnr.A2340 PMID: 21310857

51. Mike A, Strammer E, Aradi M, Orsi G, Perlaki G, Hajnal A, et al. Disconnection mechanism and regional cortical atrophy contribute to impaired processing of facial expressions and theory of mind in multiple sclerosis: a structural MRI study. PLoS One. 2013; 8(12):e82422. https://doi.org/10.1371/journal.pone.0082422 PMID: 24349280

52. Moller A, Wiedemann G, Rohde U, Backmund H, Sonntag A. Correlates of cognitive impairment and depressive mood disorder in multiple sclerosis. Acta Psychiatr Scand. 1994; 89(2):117–21. PMID: 8178661

53. Niino M, Mifune N, Kohriyama T, Mori M, Ohashi T, Kawachi I, et al. Association of cognitive impairment with magnetic resonance imaging findings and social activities in patients with multiple sclerosis. Clinical and Experimental Neuroimmunology. 2014; 5(3):328–35.

54. Nocentini U, Rossini PM, Carlesimo GA, Graceffa A, Grasso MG, Lupoi D, et al. Patterns of cognitive impairment in secondary progressive stable phase of multiple sclerosis: correlations with MRI findings. Eur Neurol. 2001; 45(1):11–8. PMID: 11150835

55. Parmenter BA, Zivadinov R, Kerenyi L, Gavett R, Weinstock-Guttman B, Dwyer MG, et al. Validity of the Wisconsin Card Sorting and Delis-Kaplan Executive Function System (DKEFS) Sorting Tests in multiple sclerosis. J Clin Exp Neuropsychol. 2007; 29(2):215–23. https://doi.org/10.1080/13803390600672163 PMID: 17365256

56. Patti F, Amato MP, Trojano M, Bastianello S, Goretti B, Caniatti L, et al. Cognitive impairment and its relation with disease measures in mildly disabled patients with relapsing-remitting multiple sclerosis: Baseline results from the Cognitive Impairment in Multiple Sclerosis (COGIMUS) study. Multiple Sclerosis. 2009; 15(7):779–88. https://doi.org/10.1177/1352458509105544 PMID: 19542262

57. Patti F, Di Stefano M, De Pascalis D, Ciancio MR, De Bernardis E, Nicoletti F, et al. May there exist specific MRI findings predictive of dementia in multiple sclerosis patients? Funct Neurol. 1995; 10(2):83–90. PMID: 7557556

58. Pozzilli C, Passafiume D, Bernardi S, Pantano P, Incoccia C, Bastianello S, et al. SPECT, MRI and cognitive functions in multiple sclerosis. J Neurol Neurosurg Psychiatry. 1991; 54(2):110–5. PMID: 2019835

59. Ron MA, Callanan MM, Warrington EK. Cognitive abnormalities in multiple sclerosis: a psychometric and MRI study. Psychol Med. 1991; 21(1):59–68. PMID: 2047506

60. Rossi F, Giorgio A, Battaglini M, Stromillo ML, Portaccio E, Goretti B, et al. Relevance of brain lesion location to cognition in relapsing multiple sclerosis. PLoS One. 2012; 7(11):e44826. https://doi.org/10.1371/journal.pone.0044826 PMID: 23144775

61. Rovaris M, Filippi M, Falautano M, Minicucci L, Rocca MA, Martinelli V, et al. Relation between MR abnormalities and patterns of cognitive impairment in multiple sclerosis. Neurology. 1998; 50(6):1601–8. PMID: 9633700

62. Rovaris M, Riccitelli G, Judica E, Possa F, Caputo D, Ghezzi A, et al. Cognitive impairment and structural brain damage in benign multiple sclerosis. Neurology. 2008; 71(19):1521–6. https://doi.org/10.1212/01.wnl.0000319694.14251.95 PMID: 18815387

63. Sacco R, Bisecco A, Corbo D, Della Corte M, d'Ambrosio A, Docimo R, et al. Cognitive impairment and memory disorders in relapsing-remitting multiple sclerosis: the role of white matter, gray matter and hippocampus. J Neurol. 2015; 262(7):1691–7. https://doi.org/10.1007/s00415-015-7763-y PMID: 25957638

64. Sanchez MP, Nieto A, Barroso J, Martin V, Hernandez MA. Brain atrophy as a marker of cognitive impairment in mildly disabling relapsing-remitting multiple sclerosis. Eur J Neurol. 2008; 15(10):1091–9. https://doi.org/10.1111/j.1468-1331.2008.02259.x PMID: 18727673

65. Sbardella E, Petsas N, Tona F, Prosperini L, Raz E, Pace G, et al. Assessing the correlation between grey and white matter damage with motor and cognitive impairment in multiple sclerosis patients. PLoS One. 2013; 8(5):e63250. https://doi.org/10.1371/journal.pone.0063250 PMID: 23696802

66. Snyder PJ, Cappelleri JC. Information processing speed deficits may be better correlated with the extent of white matter sclerotic lesions in multiple sclerosis than previously suspected. Brain Cogn. 2001; 46(1–2):279–84. PMID: 11527349

67. Sun X, Tanaka M, Kondo S, Okamoto K, Hirai S. Clinical significance of reduced cerebral metabolism in multiple sclerosis: a combined PET and MRI study. Ann Nucl Med. 1998; 12(2):89–94. PMID: 9637279

68. Swirsky-Sacchetti T, Field HL, Mitchell DR, Seward J, Lublin FD, Knobler RL, et al. The sensitivity of the Mini-Mental State Exam in the white matter dementia of multiple sclerosis. J Clin Psychol. 1992; 48(6): 779–86. PMID: 1452767

69. Yildiz M, Tettenborn B, Radue EW, Bendfeldt K, Borgwardt S. Association of cognitive impairment and lesion volumes in multiple sclerosis—a MRI study. Clin Neurol Neurosurg. 2014; 127:54–8. https://doi.org/10.1016/j.clineuro.2014.09.019 PMID: 25459243

70. Zivadinov R, De Masi R, Nasuelli D, Monti Bragadin L, Ukmar M, Pozzi-Mucelli RS, et al. MRI techniques and cognitive impairment in the early phase of relapsing-remitting multiple sclerosis. Neuroradiology. 2001; 43(4):272–8. PMID: 11338408

71. Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ. 2011; 343:d4002. https://doi.org/10.1136/bmj.d4002 PMID: 21784880

72. Fischer M, Kunkel A, Bublak P, Faiss JH, Hoffmann F, Sailer M, et al. How reliable is the classification of cognitive impairment across different criteria in early and late stages of multiple sclerosis? J Neurol Sci. 2014; 343(1–2):91–9. https://doi.org/10.1016/j.jns.2014.05.042 PMID: 24950898

73. 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; 18(6): 891–8. https://doi.org/10.1177/1352458511431076 PMID: 22190573

74. Traboulsee A, Simon JH, Stone L, Fisher E, Jones DE, Malhotra A, et al. Revised Recommendations of the Consortium of MS Centers Task Force for a Standardized MRI Protocol and Clinical Guidelines for the Diagnosis and Follow-Up of Multiple Sclerosis. AJNR Am J Neuroradiol. 2016; 37(3):394–401. https://doi.org/10.3174/ajnr.A4539 PMID: 26564433

75. Rovira A, Wattjes MP, Tintore 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. Nat Rev Neurol. 2015; 11(8):471–82. https://doi.org/10.1038/nrneurol.2015.106 PMID: 26149978

76. Steenwijk MD, Geurts JJ, Daams M, Tijms BM, Wink AM, Balk LJ, et al. Cortical atrophy patterns in multiple sclerosis are non-random and clinically relevant. Brain. 2016; 139(Pt 1):115–26. https://doi.org/10.1093/brain/awv337 PMID: 26637488

77. Meijer KA, Muhlert N, Cercignani M, Sethi V, Ron MA, Thompson AJ, et al. White matter tract abnormalities are associated with cognitive dysfunction in secondary progressive multiple sclerosis. Mult Scler. 2016; 22(11):1429–37. https://doi.org/10.1177/1352458515622694 PMID: 26733423

78. Seewann A, Kooi EJ, Roosendaal SD, Barkhof F, van der Valk P, Geurts JJ. Translating pathology in multiple sclerosis: the combination of postmortem imaging, histopathology and clinical findings. Acta Neurol Scand. 2009; 119(6):349–55. https://doi.org/10.1111/j.1600-0404.2008.01137.x PMID: 19254283

79. Ahmed I, Sutton AJ, Riley RD. Assessment of publication bias, selection bias, and unavailable data in meta-analyses using individual participant data: a database survey. BMJ. 2012; 344:d7762. https://doi.org/10.1136/bmj.d7762 PMID: 22214758