The Use of Noncontrast Quantitative MRI to Detect Gadolinium-Enhancing Multiple Sclerosis Brain Lesions: A Systematic Review and Meta-Analysis

A. Gupta, K. Al-Dasuqi, F. Xia, G. Askin, Y. Zhao, D. Delgado and Y. Wang

AJNR Am J Neuroradiol 2017, 38 (7) 1317-1322
doi: https://doi.org/10.3174/ajnr.A5209
http://www.ajnr.org/content/38/7/1317
The Use of Noncontrast Quantitative MRI to Detect Gadolinium-Enhancing Multiple Sclerosis Brain Lesions: A Systematic Review and Meta-Analysis

A. Gupta, K. Al-Dasuqi, F. Xia, G. Askin, Y. Zhao, D. Delgado, and Y. Wang

ABSTRACT

BACKGROUND: Concerns have arisen about the long-term health effects of repeat gadolinium injections in patients with multiple sclerosis and the incomplete characterization of MS lesion pathophysiology that results from relying on enhancement characteristics alone.

PURPOSE: Our aim was to perform a systematic review and meta-analysis analyzing whether noncontrast MR imaging biomarkers can distinguish enhancing and nonenhancing brain MS lesions.

DATA SOURCES: Our sources were Ovid MEDLINE, Ovid Embase, and the Cochrane data base from inception to August 2016.

STUDY SELECTION: We included 37 journal articles on 985 patients with MS who had MR imaging in which T1-weighted postcontrast sequences were compared with noncontrast sequences obtained during the same MR imaging examination by using ROI analysis of individual MS lesions.

DATA ANALYSIS: We performed random-effects meta-analyses comparing the standard mean difference of each MR imaging metric taken from enhancing-versus-nonenhancing lesions.

DATA SYNTHESIS: DTI-based fractional anisotropy values are significantly different between enhancing and nonenhancing lesions (P = .02), with enhancing lesions showing decreased fractional anisotropy compared with nonenhancing lesions. Of the other most frequently studied MR imaging biomarkers (mean diffusivity, magnetization transfer ratio, or ADC), none were significantly different (P values of 0.30, 0.47, and 0.19, respectively) between enhancing and nonenhancing lesions. Of the limited studies providing diagnostic accuracy measures, gradient-echo-based quantitative susceptibility mapping had the best performance in discriminating enhancing and nonenhancing MS lesions.

LIMITATIONS: MR imaging techniques and patient characteristics were variable across studies. Most studies did not provide diagnostic accuracy measures. All imaging metrics were not studied in all 37 studies.

CONCLUSIONS: Noncontrast MR imaging techniques, such as DTI-based FA, can assess MS lesion acuity without gadolinium.

ABBREVIATIONS: FA = fractional anisotropy; GRE = gradient recalled-echo; MD = mean diffusivity; MTR = magnetization transfer ratio; MWF = myelin water fraction; QSM = quantitative susceptibility mapping; SMD = standardized mean difference

Multiple sclerosis represents a specific disease process for which recent investigations have shown evidence of gadolinium deposition in the brain after multiple contrast-enhanced brain MRIs. Patients with MS are likely to undergo repeat contrast-enhanced MR imaging to monitor disease status. As a result, patients with MS are at high risk for cumulative deposition of gadolinium given the relatively early disease onset, which can potentially necessitate many years of disease surveillance imaging.

Received January 4, 2017; accepted after revision February 22.

From the Department of Radiology (A.G., K.A.-D., F.X., Y.W.); Clinical and Translational Neuroscience Unit (A.G.), Feil Family Brain and Mind Research Institute; Department of Healthcare Policy and Research (G.A., Y.Z.); and Samuel J. Wood Library and C.V. Starr Biomedical Information Center (D.D.), Weill Cornell Medicine, New York, New York; and Department of Biomedical Engineering (F.X., Y.W.), Cornell University, Ithaca, New York.

This work was supported National Institutes of Health grants R01NS090464 (Y.W., A.G.) and T35EB006732 (F.X.).

Please address correspondence to Ajay Gupta, MD, Weill Cornell Medicine, 525 East 68th St, Starr 8A, Box 141, New York, NY 10065; e-mail: ag9004@med.cornell.edu

Indicates open access to non-subscribers at www.ajnr.org
Indicates article with supplemental on-line appendix and tables.
Indicates article with supplemental on-line photo.

http://dx.doi.org/10.3174/ajnr.A5209

Such repeat doses of gadolinium administration in patients with MS have been correlated with increased T1-weighted signal intensity in the dentate nucleus in the cerebellum, a finding that has been associated with secondary-progressive MS and increased clinical disability. Although the precise causal role, if any, that gadolinium injections play in MS pathogenesis remains unknown, in 2016, officials from the National Institutes of Health issued guidelines recommending that the necessity of gadolinium administration in specific clinical indications should be carefully re-evaluated, given the uncertain long-term public health impact of the deposition of gadolinium in the brain.

Gadolinium enhancement of MS lesions is a well-established method to evaluate disease status, with contrast-enhancing lesions reflective of the blood-brain barrier breakdown and inflammatory response known to occur in acutely demyelinating lesions. Several quantitative MR imaging biomarkers may provide diagnostic information capable of discriminating between enhancing and nonenhancing MS lesions. Such MR imaging markers may also have the added benefit of providing insight into MS pathobiology not possible by the simple dichotomous characterization of the presence or absence of contrast enhancement. Identifying the most promising MR imaging biomarkers of MS lesion inflammatory activity is important to inform future MS imaging research in which such biomarkers could be further standardized and evaluated. We therefore performed a systematic review and meta-analysis of the existing scientific literature to evaluate whether quantitative noncontrast MR imaging metrics are able to accurately discriminate between enhancing and nonenhancing MS brain lesions.

**MATERIALS AND METHODS**

We followed both the Meta-Analysis of Observational Studies in Epidemiology group guidelines and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement in our study protocol.

**Data Sources and Searches**

An experienced medical librarian performed comprehensive searches of on-line data bases including Ovid MEDLINE, Ovid Embase, and the Cochrane Library from data base inception to August 15, 2016. Key words included “MR imaging,” “multiple sclerosis,” “gadolinium,” “enhancement,” “susceptibility,” “diffusion,” and “perfusion.” We searched first in Ovid MEDLINE and then adapted headings and keywords for other data bases and identified additional records by using the “Cited by” and “View references” features in Scopus (see the On-line Appendix for search methodology details).

**Study Selection**

We included peer-reviewed journal articles focused on patients with MS who had MR imaging brain studies in which T1-weighted postcontrast sequences were compared with noncontrast sequences obtained during the same MR imaging examination. Specific study criteria were as follows: 1) patients with MS who underwent MR imaging of the brain with and without the administration of a gadolinium-based intravenous contrast agent; 2) evaluation of MS lesion gadolinium enhancement on T1-weighted postcontrast images; and 3) ROI comparison of MR imaging biomarker values obtained in the location of gadolinium-enhancing lesions versus values obtained in T2 hyperintense, nonenhancing lesions. We included studies evaluating any of the following noncontrast quantitative MR imaging measures: gradient recalled-echo (GRE) derived quantitative susceptibility mapping (QSM); DWI-derived ADC; DTI-derived fractional anisotropy (FA) or mean diffusivity (MD); PWI-derived assessment of CBF; myelin water imaging-derived assessment of myelin water fraction (MWF); or quantitative MR imaging techniques allowing the absolute quantification of longitudinal relaxation, transverse relaxation, proton density, and magnetization transfer ratio; and 4) studies including ≥5 subjects to avoid the inclusion of case reports or very small case series. We chose to include only peer-reviewed journal articles rather than conference abstracts so that the studies included would have sufficient detail for systematic review and meta-analysis. If data from a single patient cohort were published more than once, the single article with the largest sample size was included to minimize analysis of duplicate study samples. We contacted via e-mail the corresponding author for additional details to clarify our data extraction as needed.

**Data Extraction and Quality Assessment**

After the titles and abstracts of preliminary articles were read, potentially eligible articles were shortlisted on the basis of the information contained with the abstract. Two readers (K.A.-D. and F.X., research fellows each with ~2 years of experience in brain imaging research) read the shortlisted articles in their entirety to determine eligibility, with disagreements resolved by consensus. We extracted data in duplicate by using a data-collection template, with a third tie-breaking reader (A.G., faculty neuroradiologist) resolving disagreements in data extraction. We extracted the following study characteristics: first author, study design (prospective versus retrospective), major study inclusion criteria, total number of subjects, basic study demographics, median MS disease duration, median Expanded Disability Status Scale, and specific MR imaging techniques used. We performed our analysis at the level of each MS lesion, and divided regions of the brain into 2 categories: enhancing lesions and nonenhancing, T2 hyperintense lesions. When they were available, we collected receiver operating characteristic and diagnostic accuracy measures based on proposed threshold values to describe the ability of the MR imaging biomarkers to discriminate between enhancing and nonenhancing lesions. We also evaluated the risk of bias in each study by generating specific questions to evaluate potential selection, detection, reporting, and confounding bias. Two study investigators (K.A.-D. and F.X.) assessed risks of bias with disagreements in assessment resolved by a third tie-breaking evaluator (A.G.). We did not use a threshold to exclude studies based on a bias scoring system.

**Data Synthesis and Analysis**

We calculated a standardized mean difference (SMD), comparing the difference of values of each MR imaging metric taken from enhancing-versus-nonenhancing lesions. If at least 5 studies were present that evaluated a given imaging biomarker, we statistically pooled data and performed meta-analyses of the individual quan-
titative MR imaging SMDs. We used a random-effects rather than a fixed-effects model because we adopted the conservative assumption that individual studies did not have the same effect size, given the likelihood of between-study heterogeneity in terms of subject characteristics and MR imaging methods. We performed the I² statistical heterogeneity test to evaluate the combinability of each study. The presence of publication bias was evaluated through the Begg-Mazumdar rank correlation test. We considered all \( P \) values < .05 statistically significant.

RESULTS
Study Selection and Characteristics
From a total of 6129 articles produced by our data base searching, we identified a total of 37 articles\(^8\)–\(^44\) that met inclusion for our systematic review (On-line Figure). In total, the studies included 985 patients comprising 1495 enhancing and 8529 nonenhancing MS lesions from which we were able to extract data (On-line Table 1). Most studies were prospective, and each included an average of 25 patients (range, 5–89 patients). The mean age range for patients with MS in the included studies was between 27.5 and 50 years, with all studies containing cohorts of patients who were predominantly female. Four MR imaging metrics studied in 27 unique articles had an adequate number of published studies (\( > 4 \)) whose raw data were presented in a manner amenable for meta-analysis: magnetization transfer ratio (MTR) (\( n = 11 \)),\(^{12,17–20,27,28,30,31,36,39} \) ADC (\( n = 7 \)),\(^{23,25,29,32,33,35,44} \) FA (\( n = 6 \)),\(^{8,13,24,34,35,40} \) and MD (\( n = 6 \)).\(^{8,13,14,24,34,40} \) There were 4 or fewer studies focused on specific quantitative MR imaging tissue parameters such as GRE-QSM or T1 or T2 relaxation times. We found no studies comparing noncontrast PWI metrics with enhancing-versus-nonenhancing MS lesions. Only 3 studies\(^9,23,45\) provided optimal threshold values to distinguish enhancing from nonenhancing lesions with area under the curve measurements, with 1 study\(^45\) showing an area under the curve of 0.95 for QSM to distinguish enhancing-versus-nonenhancing lesions. The remainder of the studies (\( n = 34 \)) reported differences in summary biomarker values in enhancing-versus-nonenhancing lesions without diagnostic accuracy measures. The specific MR imaging protocols used by each study and the detailed breakdown of MR imaging test values are provided in On-line Tables 2 and 3, respectively.

Meta-Analysis Results
Magnetization Transfer Ratio. The SMD (Fig 1A) between measurements of enhancing and nonenhancing lesions for MTR was...
0.009 (95% CI, −0.032−0.015; P = .47). The publication bias test (Egger) was not statistically significant for publication bias (P = .74). The heterogeneity statistic revealed I² = 94.4%, consistent with strong heterogeneity.

**Apparent Diffusion Coefficient.** The SMD (Fig 1B) between measurements of enhancing and nonenhancing lesions for ADC was −0.00008 (95% CI, −0.00020−0.00004; P = .19). There was no evidence of publication bias (P = .85), though there was significant between-study heterogeneity, with I² = 92.5%.

**Fractional Anisotropy.** The SMD (Fig 1C) between measurements of enhancing and nonenhancing lesions for FA was −0.052 (95% CI, −0.097 to −0.007; P = .02). There was no evidence of publication bias (P = .92). Finally, there was evidence of strong heterogeneity, with I² = 91.4%.

**Mean Diffusivity.** The SMD (Fig 1D) between measurements of enhancing and nonenhancing lesions for MD was −0.0001 (95% CI, −0.0003−0.0001; P = .30). No significant publication bias was present (P = .03), while I² = 92.3% was suggestive of strong heterogeneity.

**Assessment of the Quality of Included Studies**
Most studies provided detailed imaging methodology and patient inclusion/exclusion criteria that would allow the study to be reproduced (On-line Tables 4–8). We found that most included studies were prospective and all involved some form of blinding to ensure that the presence of gadolinium enhancement occurred without knowledge of the results of the MR imaging quantitative biomarker being studied. Most studies did not use >1 investigator to analyze quantitative MR imaging biomarker imaging data; therefore, most included studies did not report measures of inter-rater reproducibility.

**DISCUSSION**
The presence of lesion gadolinium contrast enhancement is a well-established marker of MS lesion inflammation and relative acuity and has figured prominently in the diagnosis and monitoring of patients with MS. Given the emergence of promising imaging biomarkers in MS and mounting concerns over the unknown long-term health consequences of repeat gadolinium administration to patients, we sought to critically analyze the body of literature evaluating differences in quantitative MR imaging markers between enhancing and nonenhancing lesions. In this systematic review and meta-analysis, we found that DTI-based FA values are significantly different between enhancing and nonenhancing lesions, with enhancing lesions showing decreased FA compared with nonenhancing lesions. Concerning the other MR imaging biomarkers we identified in the literature that most frequently addressed this question (MD, MTR, or ADC), when pooling data across studies, no other imaging markers were significantly different between enhancing and nonenhancing lesions.

In our study, we found that only 3 studies provided an area under the curve showing the performance of the proposed imaging biomarker to differentiate enhancing-versus-nonenhancing lesions. It is possible that most studies did not report diagnostic accuracy measures because the investigators did not explicitly have the goal of determining whether a given MR imaging biomarker could function as a replacement for gadolinium enhancement as a diagnostic test of lesion acuity. Indeed, most of the literature we studied was published before the recent reports that initially surfaced in 2015 over the deposition of gadolinium in the brain, and our pooled analyses of this existing body of literature suggest that re-evaluating whether these noncontrast quantitative MR imaging techniques are a viable alternative to gadolinium-based techniques in MS may now be warranted.

The differences between enhancing and nonenhancing lesions in terms of FA was a robust finding in our analysis. White matter fiber tracts are composed of aligned myelinated axons that limit the free diffusivity of water along the axis of the axonal fibers. The lower mean FA value found in enhancing MS lesions is consistent with the pathobiology of acute MS lesions in which inflammation-mediated disruption of myelin sheath integrity and an increase in local tissue vasogenic edema would together be expected to cause increased 3D diffusivity of water (lower FA).13,24 The relatively higher FA found in nonenhancing MS lesions raises the possibility that as lesions transition from enhancing to nonenhancing, some degree of axonal remyelination and/or a decrease in vasogenic edema may underlie increasing FA values following blood-brain barrier closure.21 Most interesting, we found no significant difference in ADC between enhancing and nonenhancing lesions, perhaps reflecting immune cell concentration in enhancing lesions being balanced by FA-associated myelin content immediately following blood-brain barrier closure. Because both cellular and myelin content affect ADC, the multidirectional nature of DTI may allow a more sensitive differentiation of enhancing from nonenhancing lesions compared with an ADC map alone.

Furthermore, although only 1 included study45 focused on the relatively new technique of QSM, investigators noted high diagnostic accuracy measures using specific relative susceptibility values to distinguish enhancing from nonenhancing lesions. The ability of QSM to detect acute enhancing lesions is a natural extension of the recent observational data showing that the magnetic susceptibility of an MS lesion increases as it changes from enhancing to nonenhancing. Another study included in our systematic review that focused on MR frequency shifts derived from gradient-echo data also found reduced quantitative susceptibility in enhancing-versus-nonenhancing lesions. The usefulness of tissue susceptibility in monitoring MS disease status may reflect the role played by activated macrophages and microglia in removing diamagnetic myelin fragments and depositing iron within MS lesions, which results in enhancing lesions demonstrating lower susceptibility values than nonenhancing lesions.

Our study has some limitations in the underlying literature that can serve to improve future research in evaluating emerging MR imaging biomarkers of MS. First, there were differences in pulse sequence parameters, postprocessing techniques, and methods of evaluating each lesion that may limit our ability to pool studies and provide absolute point estimates for specific pooled MR imaging biomarkers. For example, though most studies provided mean biomarker values by using ROI analysis in all enhancing-versus-nonenhancing lesions, some but not all studies performed subgroup analyses on the pattern of enhancement (homogeneous versus heterogeneous) and the nature of the non-
enhancing lesions ("black hole" versus no associated T1 hypointensity). Although such heterogeneity is an inherent challenge in the interpretation of quantitative MR imaging data across institutions and among patients, we attempted to minimize this impact in our meta-analysis by pooling within-subject data from each study and by reporting standardized mean differences, which should help minimize the lack of standardization that would be present if absolute biomarker quantification were used in our analysis.

Furthermore, we pooled studies using the more conservative random-effects model rather than a fixed-effects model to statistically account for between-study heterogeneity. Second, we found study design limitations in the literature, including the lack of blinding of investigators to enhancement status in a few studies and the failure to provide interreader reproducibility measures of imaging metrics in most studies. Finally, patient-level heterogeneity in terms of disease severity and duration was present in our pooled analysis. These limitations call for larger prospective observational studies with well-defined clinical cohorts who undergo MR imaging interpreted via a rigorous, multireader blinded study design.

Our study further illustrates important knowledge gaps in understanding of the role of non-gadolinium-based techniques in characterizing MS lesions. First, although the studies included in our analysis all compared the value of quantitative biomarkers in brain regions that were enhancing-versus-nonenhancing, only a small minority presented receiver operating characteristics, from which diagnostic accuracy measures, such as sensitivity and specificity, can be derived.46 In addition to the wide range of MS lesion biology that can be elucidated by quantitative MR imaging techniques, such receiver operating characteristic data are critical to evaluate the extent to which a specific noncontrast quantitative MR imaging technique is viable as an accurate replacement for contrast-enhanced imaging. Second, most of the included studies evaluated individual noncontrast MR imaging biomarkers in isolation, rather than in a combined, multiple parameter approach. Future studies evaluating the most promising quantitative imaging biomarkers, such as SWI and DTI metrics, in a composite, multiparametric approach are now warranted. For example, because low FA and low QSM values are independently associated with relatively acute lesions, it would be informative to perform receiver operating characteristic analysis and calculate the area under the curve of a composite imaging biomarker of low FA and low QSM in differentiating enhancing from nonenhancing MS lesions.

Although our focus was whether noncontrast quantitative MR imaging techniques are capable of detecting contrast-enhancing MS lesions, it will be critical for future research to also focus on the added value that these approaches provide to our understanding of MS lesion pathogenesis as a complement to the information provided by gadolinium enhancement. Gadolinium enhancement has been an attractive MR imaging biomarker in MS because it provides strong indirect evidence of blood-brain barrier breakdown in acute demyelinating lesions. However, the simple dichotomous characterization of MS lesions as enhancing or nonenhancing fails to capture the wide range of potentially clinically important pathophysiologic states of MS lesions, such as the chronic inflammation known to occur after the blood-brain barrier seals in MS lesions,47 for which non-contrast-dependent quantitative MR imaging biomarkers appear to hold significant promise.

CONCLUSIONS

Although the most widely used MS clinicoradiologic classification schemes have relied heavily on contrast enhancement as a diagnostic criterion, our systematic review and meta-analysis have revealed numerous investigations evaluating the ability of noncontrast MR imaging biomarkers to detect acute MS lesions. From this body of literature, we have found strong evidence in support of DTI-based FA as a promising approach in providing insight into lesion activity. Building on this understanding, we believe that the overall aim of future research should be to determine whether taking into account additional features of MS pathogenesis elucidated by noncontrast quantitative MR imaging techniques can increase the accuracy of MS diagnosis, improve disease prognostication, and provide a more robust marker of treatment response.

ACKNOWLEDGMENTS

The authors acknowledge Alina Jurcoane and Scott Kolbe for providing unpublished data used in this meta-analysis and to Michelle Demetres for providing peer review of the Ovid MEDLINE search strategy.


REFERENCES


