The quest for measuring myelin with MRI – An interactive meta-analysis of quantitative comparisons with histology

Matteo Mancini[1,2,3], Agah Karakuzu[2], Thomas Nichols[4,5], Julien Cohen-Adad[2,6], Mara Cercignani[1,7], Nikola Stikov[2,8]

1 – Department of Neuroscience, Brighton and Sussex Medical School, University of Sussex, Brighton, United Kingdom.
2 – NeuroPoly Lab, Polytechnique Montreal, Montreal, Quebec, Canada.
3 – CUBRIC, Cardiff University, Cardiff, United Kingdom.
4 – Wellcome Centre for Integrative Neuroimaging (WIN FMRIB), University of Oxford, Oxford, United Kingdom.
5 – Big Data Institute, University of Oxford, Oxford, United Kingdom.
6 – Functional Neuroimaging Unit, CRIUGM, Université de Montréal, Montreal, Quebec, Canada
7 – Neuroimaging Laboratory, Fondazione Santa Lucia, Rome, Italy.
8 - Montreal Heart Institute, Université de Montréal, Montreal, Quebec, Canada.

Abstract

Several MRI measures have been proposed as in vivo biomarkers of myelin content, each with a concrete application ranging from plasticity to pathology. Despite the broad availability of these myelin-sensitive MRI modalities, specificity and sensitivity have been a matter of discussion. Debate about which MRI measure is the most suitable one for quantifying myelin is still ongoing.

In this study, we performed a systematic review of published quantitative validation studies, and used meta-analysis tools to clarify how different these measures are when compared to the underlying histology, controlling for the study sample size and using interactive visualization tools (https://neurolibre.github.io/myelin-meta-analysis). A first qualitative selection of 58 studies proposed 35 different measures to characterize myelin content. However, a quantitative analysis showed that most of these measures have a limited coefficient of determination and provide little information to inform future studies, because of the large prediction intervals and high heterogeneity. These results indicate that most measures are statistically equivalent regarding their relationship with histology and that future work should take inter-study variability into consideration.

Abbreviations and mathematical symbols

AD – axial diffusivity

AK – axial kurtosis

AWF – axonal water fraction

FA – fraction anisotropy

ihMTR – inhomogeneous magnetization transfer ratio

k_fm – free water-macromolecular exchange rate

k_mf – macromolecular-free water exchange rate

M0m – macromolecular pool magnetization fraction

MD – mean diffusivity

MK – mean kurtosis

MPF – macromolecular pool fraction

MT – magnetization transfer

MTR – magnetization transfer ratio

MTR-UTE – magnetization transfer ratio (using ultra-short echo time)

MTV – macromolecular tissue volume

MVF-MT – myelin volume fraction (estimated from MT)

MVF-T2 – myelin volume fraction (estimated from T2)

MWF – myelin water fraction

PD – proton density

PN – peripheral nerve

PRISMA – Preferred Reporting Items for Systematic Reviews and Meta-Analyses

QSM – quantitative susceptibility mapping

R1f – free water pool longitudinal relaxation rate

R2* – apparent transverse relaxation rate

RAFF – relaxation along a fictitious field

RD – radial diffusivity

RD-DBSI – radial diffusivity (from diffusion basis spectrum imaging)

RDe – extra-cellular compartment radial diffusivity

RK – radial kurtosis

rSPF – relative semi-solid proton fraction

SC – spinal cord

T1 – longitudinal relaxation time

T1p – adiabatic longitudinal relaxation time

T1sat – longitudinal relaxation time under magnetization transfer irradiation

T2 – transverse relaxation time

T2f – free water pool transverse relaxation time

T2int – transverse relaxation intermediate component

T2m – macromolecular pool transverse relaxation rate

T2p – adiabatic transverse relaxation time

Introduction

Myelin is a key component of the central nervous system. The myelin sheaths insulate axons with a triple effect: allowing fast electrical conduction, protecting the axon, and providing trophic support (Nave & Werner, 2014). The conduction velocity regulation has become an important research topic, with evidence of activity-dependent myelination as an additional mechanism of plasticity (Sampaio-Baptista & Johansen-Berg, 2017). Myelin is also relevant from a clinical perspective, given that demyelination is often observed in several neurological diseases such as multiple sclerosis (Y. Wang et al., 2015).

Given this important role in pathology and plasticity, measuring myelin in vivo has been an ambitious goal for magnetic resonance imaging (MRI) for almost two decades (Petiet et al., 2019). Even though the thickness of the myelin sheath is in the order of micrometres, well beyond the MRI spatial resolution, its presence influences several physical properties that can be probed with MRI, from longitudinal and transversal relaxation phenomena to water molecule diffusion processes.

However, being sensitive to myelin is not enough: to study how and why myelin content changes, it is necessary to define a specific biomarker. Interestingly, the quest for measuring myelin has evolved in parallel with an important paradigm shift in MRI research, where MRI data are no longer treated as just “pictures”, but as actual 3D distributions of quantitative measures. This perspective has led to a new field of research, quantitative MRI (qMRI), that encompasses the study of how to measure the relevant electromagnetic properties that influence magnetic resonance phenomena in biological tissues (Cercignani, Dowell, & Tofts, 2018; Cohen-Adad & Wheeler-Kingshott, 2014). From the very definition of qMRI, it is clear that its framework applies to any approach for non-invasive myelin quantification.

Similarly to other qMRI biomarkers, MRI-based myelin measurements are indirect, and might be affected by other microstructural features, making the relationship between these indices and myelination noisy. Assessing the accuracy of such measurements, and their sensitivity to change, is essential for their translation into clinical applications. Validation is therefore a fundamental aspect of their development (Cohen-Adad, 2018). The most common approach is based on acquiring MR data from in vivo or ex vivo tissue and then comparing those data with the related samples

analysed using histological techniques. Despite being the most realistic approach, this comparison involves several methodological choices, from the specific technique used as a reference to the quantitative measure used to describe the relationship between MRI and histology. So far, a long list of studies have looked at MRI-histology comparisons (Cohen-Adad, 2018; Laule & Moore, 2018; MacKay & Laule, 2016; Petiet et al., 2019), each of them focusing on a specific pathology and a few MRI measures. Despite these numerous studies, there is still an ongoing debate on what MRI measure should be used to quantify myelin and as a consequence there is a constant methodological effort to propose new measures. This debate would benefit from a quantitative analysis of all the findings published so far, specifically addressing inter-study variations and prospects for future studies, something that is currently missing from the literature.

In this study, we systematically reviewed quantitative MRI-histology comparisons and we used meta-analysis tools to address the following question: how different are the modalities for myelin quantification in terms of their relationship with the underlying histology?

Results

Literature survey

The screening process is summarized in the flowcharts in Figure 1 and Figure A1. The keywords as reported in the appendix returned 688 results on PubMed (last search on 03/06/2020). These results included 50 review articles. From the 50 review articles, six were selected as relevant for both the topics of myelin and related MRI-histology comparisons (Cohen-Adad, 2018; Laule & Moore, 2018; Laule et al., 2007; MacKay & Laule, 2016; Petiet et al., 2019; Turner, 2019). After the assessment, 58 original research studies were considered eligible, as shown in Table A1 (in the appendix) and Figure S2. All the data collected are available in the supplementary materials.

In terms of specific modalities, the survey shows that the most common MRI approach compared with histology was diffusion-weighted imaging (used in 28 studies), followed by magnetization transfer (MT, 27 studies), T2 relaxometry (19 studies) and T1 relaxometry (10 studies). Only 20 studies considered more than one approach: among the others, 20 focused exclusively on diffusion, 12 on MT, and 6 on T2 relaxometry.

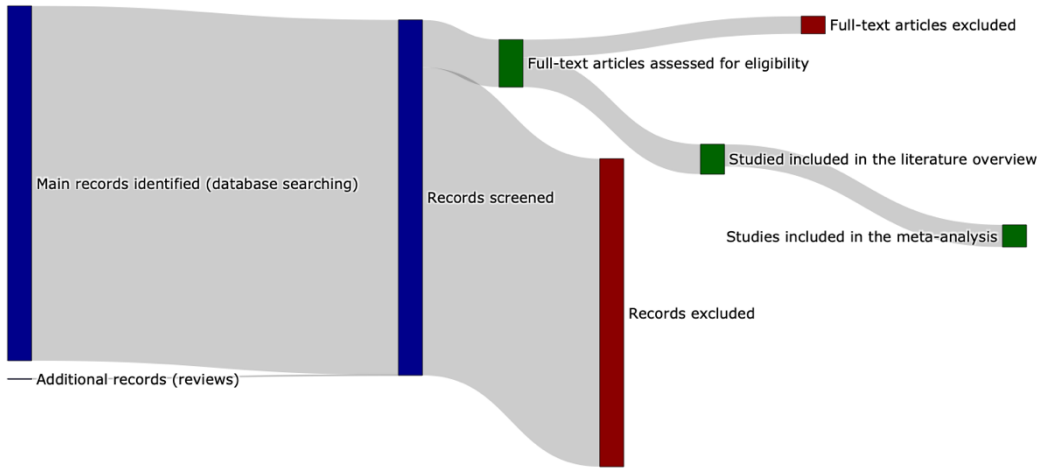

Figure 1 – Sankey diagram representing the screening procedure (PRISMA flow chart provided in the appendix). To see the interactive figure: https://neurolibre.github.io/myelin-meta-analysis/01/selection.html#Figure-1

From these 58 studies, we then focused only on brain studies and we further excluded studies not reporting either the number of subjects or the number of ROIs per subject. We also excluded one single-subject study that relied on voxels as distinct samples, whereas the other studies in this review are based on ROIs (i.e. including more than one voxel). In the end, 43 suitable studies were identified for the subsequent analyses.

Meta-analysis

To compare the studies of interest, we first organized them according to the terms of MRI measure used. Figure 2 and Figure 3 (and also Figure S3-S4) show the R^2 values for the selected studies across measures: the highest values (R^2>0.8) are obtained mostly from MT measures, but they are associated with small sample sizes (with an average of 32 sample points). The studies with largest sample sizes are associated with R^2 values between 0.6 and 0.8 for MT and T2 relaxometry, but with lower values for T1 relaxometry and other approaches.

To combine the results for each measure, we then used a mixed-effect model: in this way we were able to express the overall effect size in terms of a range of R^2 values within a confidence interval, but also to assess prediction intervals and inter-study differences. The results are shown in Figure 4 (and also Figure S5).

Apart from MPF and MWF, all the measures showed R^2 overall estimates in the range 0.21-0.53. When considering the prediction intervals, for most measures the interval spanned from 0.1 to 0.9. This implies that future studies relying on such measures can expect, on the basis of these studies, to obtain any value in this broad interval. The only exceptions were MPF (0.49-1) and MWF (0.45-0.95), whose intervals were narrower than the alternatives. Finally, I^2 (a measure of how much of the variability in a typical study is due to heterogeneity in the experimental design) was generally quite high (Table 1). MWF showed the lowest I^2 across measures, but considering that it was based on only 4 studies while the other measures included around 10 studies, it was relatively high. Excluding MWF, MPF also showed a relatively low I^2. Qualitative comparisons across experimental conditions and methodological choices highlighted differences across pathology models, targeted tissue types and reference techniques (Figure 5 and Figure S6). Other factors such as magnetic field, co-registration, specific tissue and the related conditions (Figure S7) showed comparable distributions.

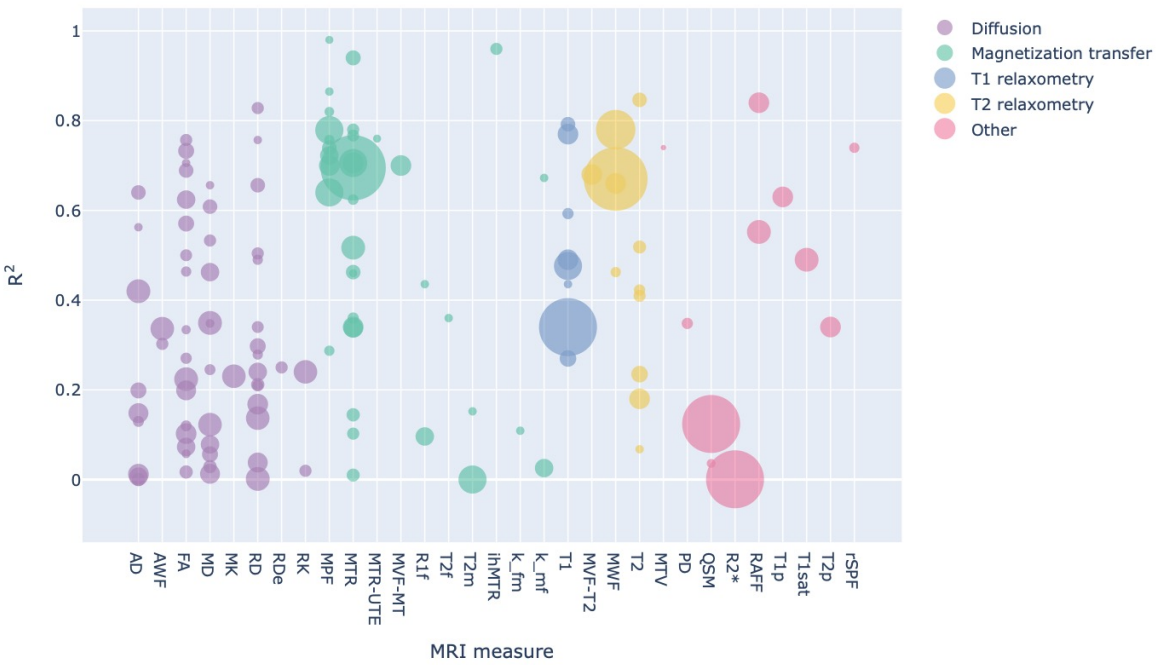

Figure 2 – Bubble chart of R^2 values between a given MRI measure and histology for each study across MRI measures, with the area proportional to the number of samples. To see the interactive figure: https://neurolibre.github.io/myelin-meta-analysis/02/closer_look.html#Figure-3

Discussion

Indirect measures are the most popular (for better or worse)

The literature survey offers an interesting perspective on popular research trends (Figure S2). The first consideration one can make is that the most common families of approaches, DWI and MT, both provide indirect measures of myelin. The MT effect is driven by saturation pulses interacting with myelin macromolecules that transfer their magnetization to water, whereas in diffusion experiments myelin is just not part of the picture. Diffusion acquisitions are blind to direct myelin measurement (Campbell et al., 2018) because the TEs used are too long (~100ms) to be influenced by the actual macromolecules (with T2 of ~10us) or even the water molecules trapped in the myelin sheath (with T2 of ~30ms). To infer myelin content, one needs to rely on the interaction between intra-cellular and extra-cellular water compartments. The majority of diffusion studies included in this analysis used tensor-based measures (with fractional anisotropy being the most common), but some also used kurtosis-based analysis. The main issue with this approach is that other factors affect those measures, making it difficult to specifically relate changes in the water compartments to changes in myelin.

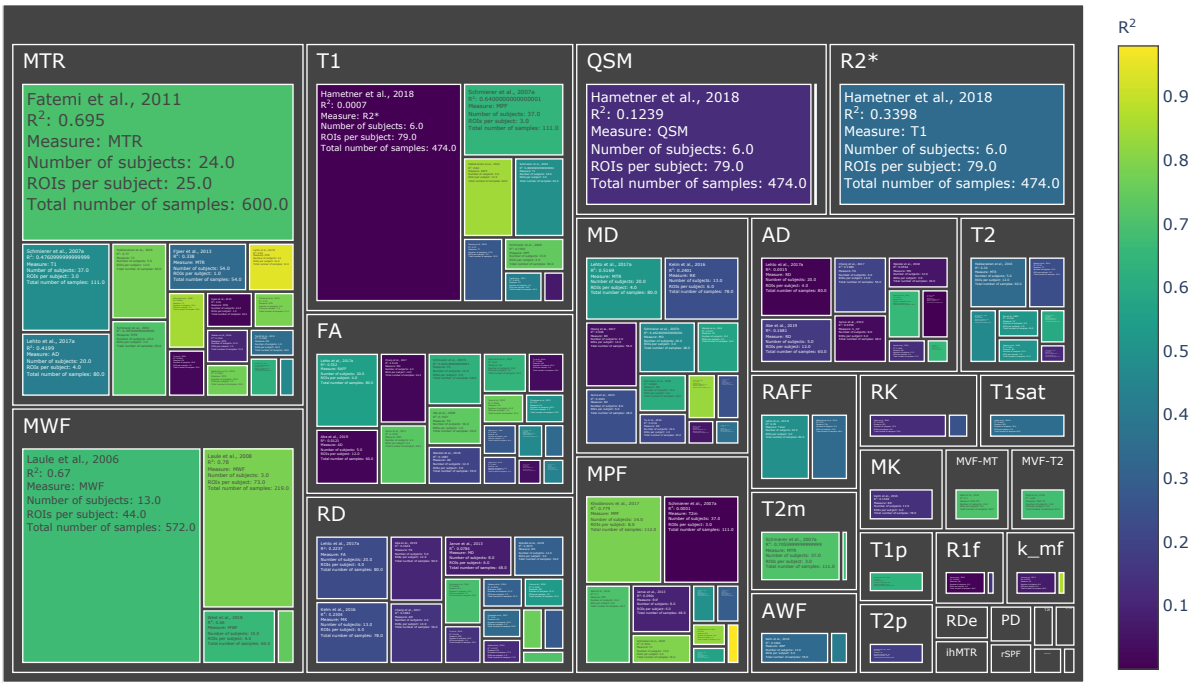

Figure 3 – Treemap chart of the studies considered for the meta-analysis, organized by MRI measure. The color of each box represents the reported R^2 value while the size box is proportional to the sample size. To see the interactive figure: https://neurolibre.github.io/myelin-meta-analysis/02/closer_look.html#Figure-4

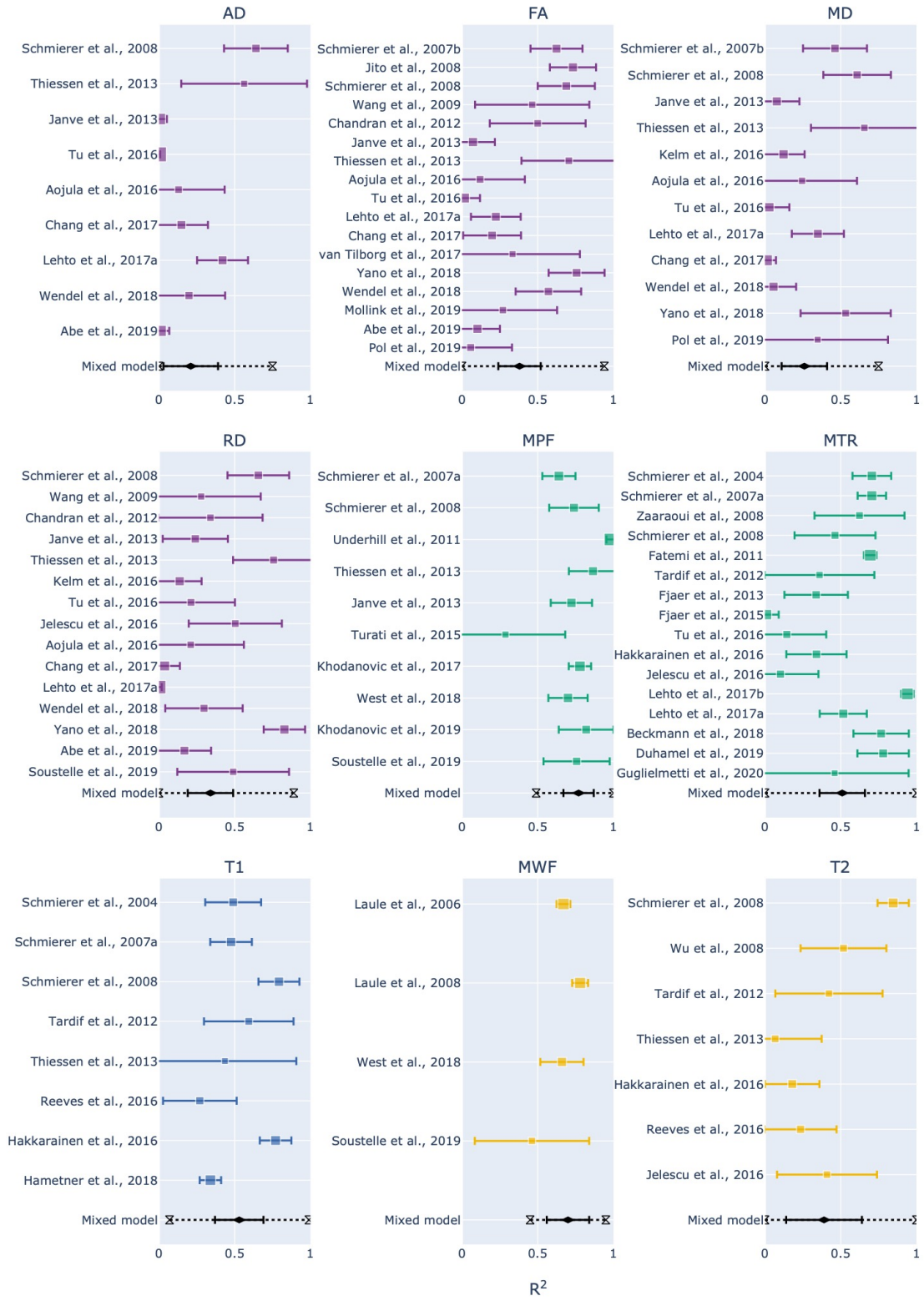

Figure 4 – Forest plots showing the R^2 values reported by the studies and estimated from the mixed-effect model for each measure. The hourglasses and the dotted lines in the mixed-effect model outcomes represent the prediction intervals.
To see the interactive figure:
https://neurolibre.github.io/myelin-meta-analysis/03/meta_analysis.html#Figure-5

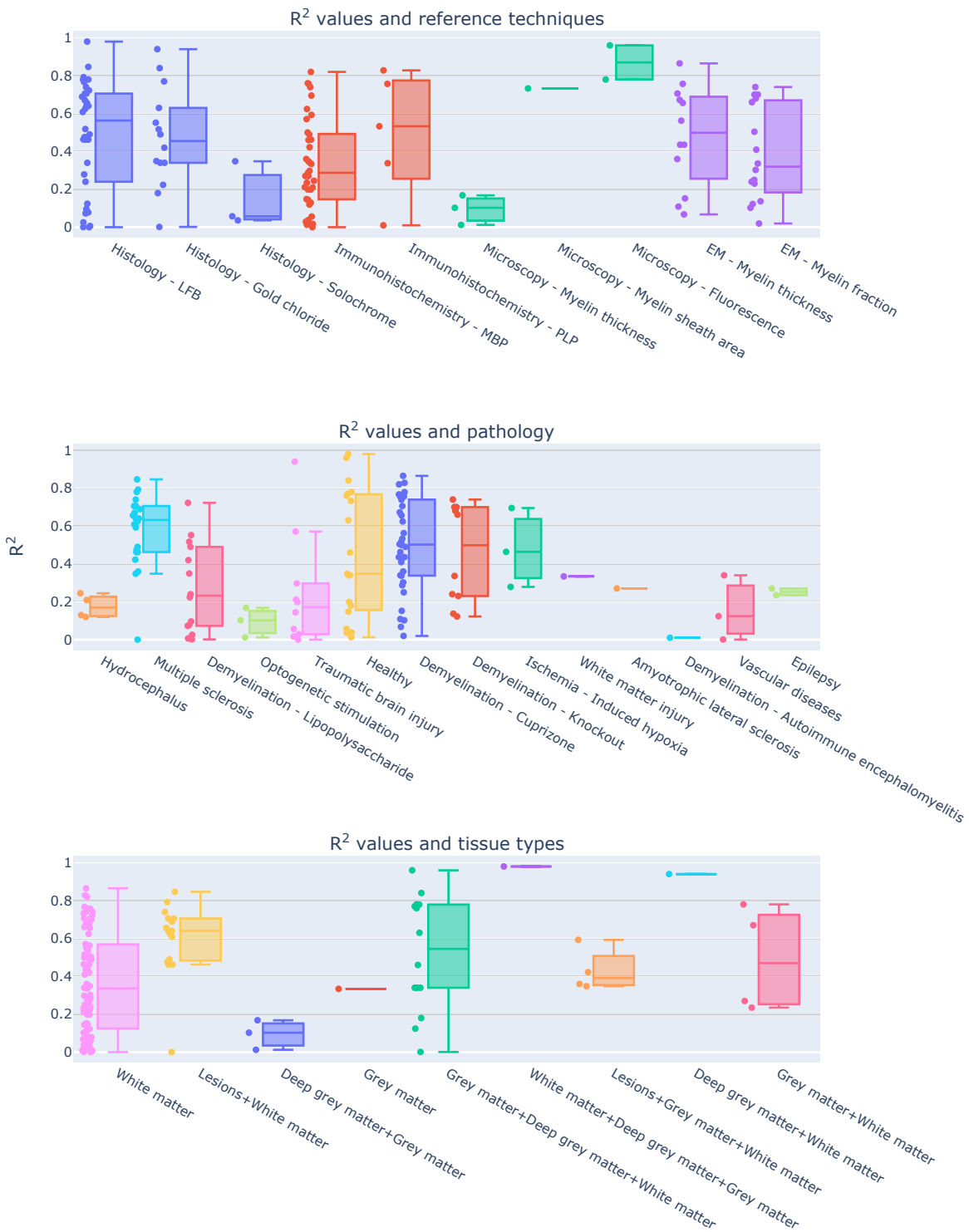

Figure 5 – Experimental conditions and methodological choices influencing the R^2 values (top: reference techniques; middle: pathology model; bottom: tissue types). To see the interactive figure: https://neurolibre.github.io/myelin-meta-analysis/03/meta_analysis.html#Figure-6

MEASURE	NUMBER OF STUDIES	ESTIMATE	STANDARD ERROR	TAU²	I²
MTR	16	0.508	0.0691	0.07	96.03%
MPF	10	0.7657	0.0455	0.0128	83.18%
FA	17	0.3766	0.0663	0.0652	87.49%
RD	15	0.3364	0.0679	0.0615	92.30%
MD	12	0.2639	0.0679	0.044	87.35%
T1	8	0.5321	0.0692	0.0328	86.51%
AD	9	0.2095	0.0802	0.048	97.69%
T2	7	0.3938	0.1023	0.0651	84.49%
MWF	4	0.6997	0.0432	0.0041	73.19%

Table 1 Results from the mixed-effect models: for each measure we reported the number of studies, the estimate and standard error of the overall R^2 distribution, the tau^2 and the I^2.

Despite this issue, the use of diffusion as a proxy for myelin is quite widespread, specifically outside the field of quantitative MRI. This is probably a consequence of how popular DWI has become and how widely available are the related acquisition sequences. MT, the second most popular technique for quantifying myelin, estimates myelin by acquiring data with and without saturating the macromolecular proton pool. The simplest MT measure, MT ratio (MTR), incorporates non-myelin contributions in the final measurement. Recent acquisition variations include computing MTR from acquisitions with ultra-short echo times (Guglielmetti et al., 2020) or relying on inhomogeneous MT (Duhamel et al., 2019). More complex experiments, e.g. quantitative MT, are based on fitting two compartments to the data, the free water and the macromolecular compartments, or pools. In this way, one is able to assess myelin through MPF with higher specificity, although still potentially including contributions from other macromolecules. Additional measures have also been considered (including the T2 of each pool, the exchange rate between the pools). The drawback of qMT is the requirement for a longer and more complex acquisition. Recently there have been alternative techniques to estimate only MPF, resulting in faster acquisitions with similar results (M. Khodanovich et al., 2019; M. Y. Khodanovich et al., 2017).

Following diffusion and MT, the most popular approach is T2 relaxometry. Unlike diffusion and MT, in T2 relaxometry experiments one can directly observe the contribution from the water trapped between the myelin bilayers, and can therefore

estimate the myelin water fraction. A simpler but less specific approach consists in estimating the transverse relaxation time considering the decay to be mono-exponential. The main drawback of these approaches is that they require longer acquisitions.

Finally, other studies used a diverse collection of other measures, including T1 relaxometry, apparent transversal relaxation rate (R2*), proton density (PD), macromolecular tissue volume (MTV), relaxation along a fictitious field (RAFF), and quantitative susceptibility mapping (QSM).

After this general overview, it is clear that each modality could be a suitable candidate for a quantitative myelin biomarker. To then make a choice informed by the studies here reported, it becomes necessary to consider not only effect sizes in terms of correlation, but also sample sizes and acquisition times.

<u>There is no myelin MRI measure true to histology</u>

When looking at the R^2 values across the different measures, the first detail that catches one's eye is how most measures present a broad range of values (Figure 2 and Figure 3). When taking into account the sample size, the largest studies show higher correlations for MT and T2 relaxometry studies than any other approach (Figure S3 and Figure S4). In quantitative terms, the meta-analysis corroborates this idea, showing that MPF and MWF tend to be more specific to myelin compared to the other measures (respectively with R^2=0.7657 and R^2=0.6997), in line with the underlying theory. Notably, diffusion-based measures show the lowest overall estimates (with values between R^2=0.3766 for FA and R^2=0.2095 for AD): this could be due to the fact, as already mentioned, that DWI does not specifically measure myelin properties, and despite FA and RD being influenced by the myelin content, they are also influenced by other factors that make them unsuitable as measures of myelin.

Despite these considerations on the advantages of MPF and MWF, one should refrain from concluding that they are the "true" MRI measures of myelin. The reason for this caution is given not by the overall effect sizes observed here, but by the collateral outcomes of the meta-analysis. The first one is given by the prediction intervals: most measures showed quite a large interval (Figure 4), not supporting the idea of them

being robust biomarkers. MPF and MWF seem to be again the most suitable choices for future studies, but a range between 0.5 and 1 is still quite large.

The second important aspect to consider is given by the differences across studies: the meta-analysis showed how such differences strongly limit inter-study comparisons for a given measure (Figure 5). This result should be expected, given that the studies here examined are inevitably influenced by the specific experimental constraints and methodological choices. A clear example of this aspect is the validation modality used as a reference, which will be dictated by the equipment availability and cost. However, such a choice has an impact on the actual comparison: histology and immunochemistry, despite being specific to myelin, do not offer a volumetric measure of myelin, but rather a proxy based on the transmittance of the histological sections. So far, the only modality able to give a volumetric measure would be electron microscopy, which is an expensive and resource-consuming approach. Also, electron microscopy has several limitations, including tissue shrinkage, degradation of the myelin sheath structure due to imperfect fixation, imperfect penetration of the osmium stain, polishing, keeping focus over large imaging regions. All these effects participate in the lack of precision and accuracy when quantifying myelin content with EM-based histology (Cohen-Adad, 2018). Another important observation is that none of the studies here reviewed considered histology reproducibility, which is hard to quantify as a whole given that a sample can be processed only once: collateral factors affecting tissue processing (e.g. sectioning distortions, mounting and staining issues) constitute an actual limitation for histology-based validation.

In addition to differences in experimental and methodological designs, there are also several considerations that arise out of the lack of shared practices in MRI validation studies. The first evident one is the use of correlations: despite being a simple measure that serves well the purpose of roughly characterizing a relationship, Pearson correlation is not the right tool for quantitative biomarkers, as it does not characterize the actual relationship between histology and MRI. Linear regression is a step forward but has the disadvantage of assuming a linear relationship. Despite Pearson correlation and linear regression being the most common measures used in the studies here reviewed, it is still not clear if the relationship is actually linear. Only one study among the considered ones computed both Pearson and Spearman correlation

values (Tardif, Bedell, Eskildsen, Collins, & Pike, 2012), and reported higher Spearman correlations, pointing out that non-linear relationships should actually be considered. One last consideration regarding the use of correlation measures for validating quantitative biomarkers is about the intercept in the MRI-histology relationship. Notably, only MWF is expected to assume a value equal to zero when myelin is absent (West et al., 2018). For the other measures, it would be necessary to estimate the intercept, which leads to the calibration problem in the estimate of myelin volume fraction. Notably, calculating Pearson correlation does not provide any information for such calibration.

Another arbitrary practice that would benefit from some harmonization is the choice of ROIs. The studies reported here examined a diverse list of ROIs, in most cases hand-drawn on each modality, encompassing different types of tissue, and the most common approach is to report a single, pooled correlation. This is problematic, as different types of tissue (e.g. grey matter and white matter) will show different values for MRI-based measures but also for histology-based ones, making linearity assumptions about the two modalities. However, with this approach gross differences between tissues drive the observed correlation, without actually showing if the MRI-based measure under analysis is sensitive to subtle differences and therefore a suitable quantitative biomarker for myelin.

It should be clear at this point that any debate about a universal MRI-based measure of myelin is pointless, at least at the moment, as the overall picture provided by previous studies does not point to any such ideal measure. Nevertheless, is debating about a universal measure helpful for future studies?

<u>Better biomarkers require more reproducibility studies</u>

Rather than debating about a perfect measure, we would argue that what is missing at the moment is a clear picture of what can be achieved with each specific MRI modality. The studies examined here focus on a large set of different measures, and more than half of them considered at most two measures, highlighting how the field is mostly focused on formulating new measures. While it is understood that novel measures can provide new perspectives, it is also fundamentally important to understand the concrete capabilities and limitations of current measures. From this

meta-analysis, what the literature clearly lacks is reproducibility studies, specifically answering two main questions: (1) what is the specificity of each measure? We should have a practical validation of our theoretical understanding of the relevant confounds; (2) what is the "parameter sensitivity" of each measure? Here we refer to parameter sensitivity in a broad sense, that includes also experimental conditions and methodological choices. The results here presented show how certain conditions (e.g. pathology) seem to affect the coefficient of determination more than others but given the limited number of studies for each modality, we refrained from additional analyses to avoid speculation. A warning message that is evident from these results is the inherent limitation of DWI for estimating myelin content. If estimating myelin content is relevant in a diffusion study, it is important to consider complementing the diffusion measure with one of the modalities here reviewed; in this way, it would be possible to decouple the influence of myelin content from the many other factors that come into play when considering diffusion phenomena.

Finally, an important factor to take into account when choosing a biomarker of myelin is the actual application. For animal research, long acquisitions are not a major issue. However, when considering biomarkers for potential clinical use, the acquisition time can become a relevant issue. This is particularly problematic for measures such as MWF, that could only be used for a specific slice in a hypothetical clinical scenario. Faster techniques have been proposed for estimating it with gradient- and spin-echo (GRASE) sequences (Faizy et al., 2018). Even in this case, the acquisition time still reaches 15 minutes for acquiring roughly the whole brain with an isotropic resolution of 2mm. Complex MT acquisitions such as qMT suffers from the same problem, although it is possible to use optimized and faster protocols to focus specifically on MPF (M. Khodanovich et al., 2019; M. Y. Khodanovich et al., 2017).

Conclusions

Several MRI measures are sensitive to myelin content and the current literature suggests that most of them are statistically equivalent in terms of relationship with the underlying histology. Measures highly correlated with histology are also the ones with a higher expected specificity. This suggests that future studies should try to better address how specific each measure is, for the sake of clarifying suitable applications.

Methods

Review methodology

The Medline database (https://pubmed.ncbi.nlm.nih.gov) was used to retrieve the articles. The keywords used are specified in the appendix. We followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines for record screening and study selection. The results were first screened to remove unrelated work. Specifically we discarded: work relying only on MRI; work relying only on histology or equivalent approaches; work reporting only qualitative comparisons. After this first screening, the remaining papers were assessed. At this stage, we discarded: studies using MRI-based measures in arbitrary units (e.g. T1-weighted or T2-weighted data); studies using measures of variation in myelin content (defined either as the difference between normal and abnormal myelin content) either for MRI or for histology; studies using arbitrary assessment scales; studies comparing MRI-based absolute measures of myelin with histology-based relative measures (e.g. g-ratio); studies reporting other quantitative measures than correlation or R^2 values; studies comparing histology from one dataset and MRI from a different one. As an additional source for potential candidate studies, we screened the review articles in the initial results, and we selected the relevant studies that were not already present in the studies already selected.

From the final papers, we collected first the following details: the DOI; which approach was used (diffusion, MT, T1 relaxometry, T2 relaxometry, or other); which specific MRI measures were compared to histology or equivalent techniques; the magnetic field; the technique used as a reference (histology, immunochemistry, microscopy, electron microscopy); the focus of the study in terms of brain, spinal cord or peripheral nerve; if the subjects were humans or animals, and if the latter which animal; if the tissue under exam was *in vivo*, *in situ* or *ex vivo*, and in the latter case if the tissue was fixed or not; if the tissue was healthy or pathological, and if the latter which pathology; the specific structures examined for correlation purposes; which comparison technique was used (e.g. Pearson correlation, Spearman correlation, linear regression); the number of subjects; the number of ROIs per subject; the male/female ratio; if registration procedures were performed to align MRI and histology; in case of pathological tissue, if control tissue was considered as well; other relevant notes. If

before calculating the correlations the data were averaged across subjects, the number of subjects was considered to be one. The same consideration was made for averaging across ROIs. This is because the numbers of subjects and ROIs were used to take into account how many sample points were used when computing the correlation. We set each of those numbers to 1 for all the studies where the data were averaged respectively across subjects and across ROIs. Finally, in those cases where the number of ROIs or the number of subjects were given as a range rather than specific values, we used the most conservative value and added the related details to the notes.

We then proceeded to collect the quantitative results reported for each measure and for each study in the form of R^2. Given that different studies may rely on a different strategy when reporting correlations, we adopted the following reasoning to limit discrepancies across studies while still objectively representing each of them. In case of multiple correlation values reported, for our analysis we selected the ones referring to the whole dataset and the entire brain if available, and considering each ROI in a given subject as a sample if possible; if only correlation values for specific ROIs were reported, the one for the most common reported structure would be chosen. In the case of multiple subjects, if data were provided separately for each group, the correlation for the control group was used. When different comparison methods were reported (e.g. both Pearson and Spearman correlation) or if the MRI data was compared with multiple references (e.g. both histology and immunohistochemistry), the correlations used were chosen on the basis of the following priority orders (from the most preferable to the least): for multiple comparison methods, linear regression, Spearman correlation, Pearson correlation; for multiple references, electron microscopy, immunohistochemistry, histology. Finally, in any other case where more than one correlation value was available, the most conservative value was used. Any other additional value was in any case mentioned in the notes of the respective study.

Meta-analysis

For the quantitative analysis, we restricted our focus on brain studies and only on the ones providing an indication of both the number of subjects and the number of ROIs. For each study, we computed the sample size as the product between the number of

subjects and the number of ROIs per subject. In this way, we were able to compare the reported R^2 values across measures taking into account the related number of points actually used for correlation purposes. We note that correlation or regression analyses run on multiple ROIs and subjects represents a repeated measures analysis, for which the degrees of freedom computation can be complex; however, most papers neglected the repeated measures structure of the data and thus the sample size computation here represents a very approximate and optimistic view of the precision of each R^2 value.

Assuming that an underlying Pearson's correlation r has variance $(1-\rho^2)^2/N$, where ρ is the true correlation, we use a first order expansion to approximate the variance of R^2 as $4\rho^2(1-\rho^2)^2/N$, where we plug in R^2 for ρ^2. Again, we recognise that some papers computed Spearman correlation, for which this calculation is again optimistic and may underestimate the sampling variability of the squared Spearman correlation.

We then proceeded to fit a mixed-effect model to each measure where more than one study was available. This model takes into account potential differences across studies as random effects and allows us to estimate a prediction interval of effect sizes (in our case, the R^2 values) on the basis of the effects and sample sizes of each study. We reported two additional measures from the mixed-effect model, I^2 and tau^2: the former expresses as a percentage how much of variability in a typical study is due to heterogeneity (i.e. the variation in study outcomes between studies) rather than chance (Higgins & Thompson, 2002), while the latter can be used to calculate the prediction interval (Raudenbush, 2009), which gives the expected range for the measure of interest in future studies. We used forest plots to represent the outcomes, and both the mixed effects estimate of the population estimated R^2, with both a 95% confidence and a (larger) 95% prediction interval.

To compare the R^2 across measures, we used the Jupyter notebook provided in the supplementary materials. For model fitting and forest plots, we used the Metafor package, version 2.4-0 (Viechtbauer, 2010).

Acknowledgements

MM was funded by the Wellcome Trust through a Sir Henry Wellcome Postdoctoral Fellowship [213722/Z/18/Z].

Appendix

Search keywords:

(myelin[Title/Abstract] AND ((magnetic[Title/Abstract] AND resonance[Title/Abstract]) OR mr[Title/Abstract] OR mri[Title/Abstract])) AND (histology[Title/Abstract] OR histopathology[Title/Abstract] OR microscopy[Title/Abstract] OR immunohistochemistry[Title/Abstract] OR histological[Title/Abstract] OR histologically[Title/Abstract] OR histologic[Title/Abstract] OR histopathological[Title/Abstract] OR histopathologically[Title/Abstract] OR histopathologic[Title/Abstract])

Results obtained from the Medline database: 688 (03/06/2020)

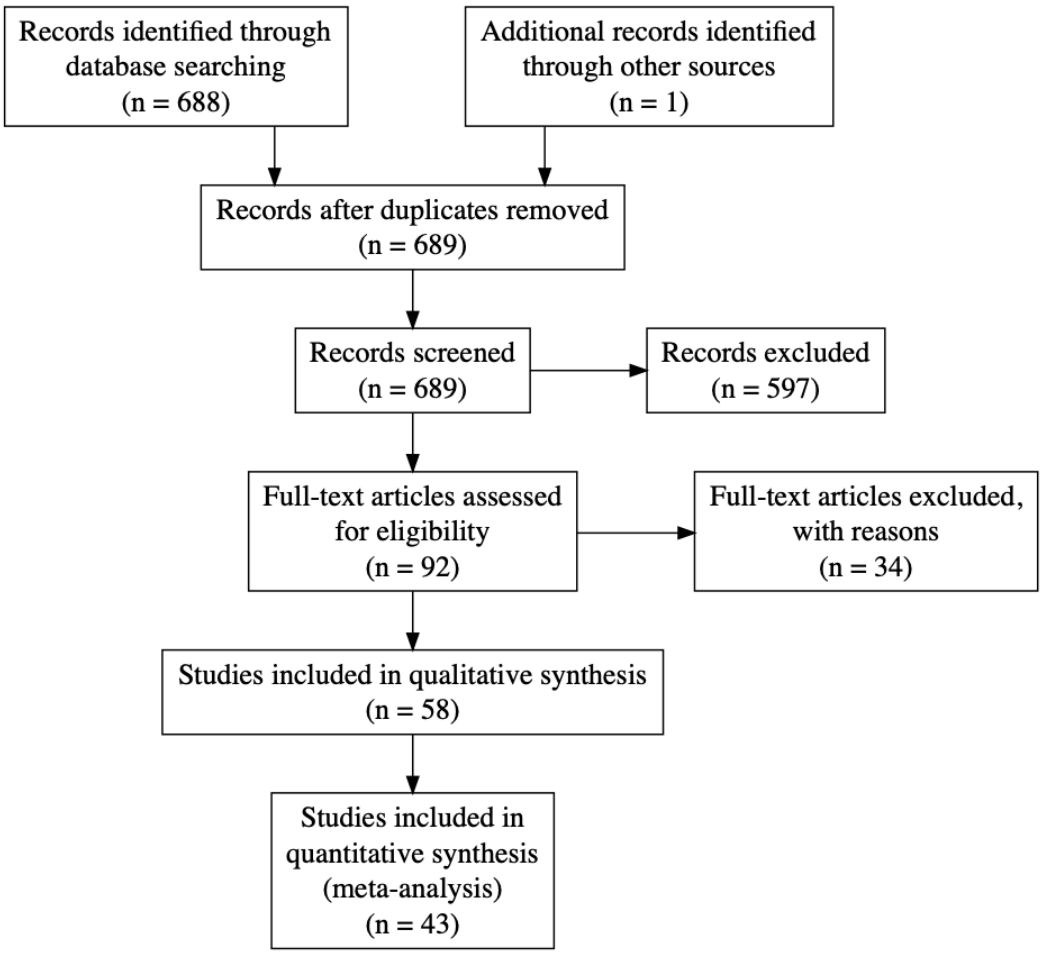

Figure A1 - PRISMA flowchart for the meta-analysis.

Table A1 - Selected studies for qualitative analysis.

Study	MRI measure(s)	Histology/microscopy measure	Tissue	Condition	Focus
(Schmierer, Scaravilli, Altmann, Barker, & Miller, 2004)	T1, MTR	Histology - LFB	Human	Multiple sclerosis	Brain
(Odrobina, Lam, Pun, Midha, & Stanisz, 2005)	T1, T2, T2int, MWF, M0m, MTR	Microscopy - Myelin fraction	Animal - Rat	Demyelination - Tellurium	PN
(Pun et al., 2005)	T1, T2int, MWF	Microscopy - Myelin fraction	Animal - Rat	Demyelination - Tellurium	PN
(Laule et al., 2006)	MWF	Histology - LFB	Human	Multiple sclerosis	Brain
(Schmierer, Tozer, et al., 2007)	T1, MTR, MPF, T2m	Histology - LFB	Human	Multiple sclerosis	Brain
(Schmierer, Wheeler-Kingshott, et al., 2007)	FA, MD	Histology - LFB	Human	Multiple sclerosis	Brain
(Jito et al., 2008)	FA	Microscopy - Myelin sheath area	Animal - Mouse	Healthy	Brain
(Kozlowski et al., 2008)	MWF, FA, AD, RD, MD	Immunohistochemistry - MBP	Animal - Rat	Injury - Dorsal columnar transection	SC
(Laule et al., 2008)	MWF	Histology - LFB	Human	Multiple sclerosis	Brain

(Schmierer et al., 2008)	T1, T2, MTR, MPF, MD, FA, AD, RD	Histology - LFB	Human	Multiple sclerosis	Brain
(Wu et al., 2008)	T2	Histology - LFB	Animal - Mouse	Demyelination - Cuprizone	Brain
(Zaaraoui et al., 2008)	MTR	Immunohistochemistry - MBP	Animal - Mouse	Demyelination - Cuprizone	Brain
(Takagi et al., 2009)	FA, AD	EM - Myelin thickness	Animal - Rat	Degeneration - Contusive injury	PN
(S. Wang et al., 2009)	FA, RD	Histology - LFB	Animal - Rat	Ischemia - Induced hypoxia	Brain
(Zhang et al., 2009)	RD	Histology - LFB	Animal - Rat	Injury - Dorsal columnar transection	SC
(Schmierer et al., 2010)	MTR, T2	Histology - LFB	Human	Multiple sclerosis	Brain
(Fatemi et al., 2011)	MTR	Immunohistochemistry - MBP	Animal - Mouse	Ischemia - Induced hypoxia	Brain
(Laule et al., 2011)	MWF	Immunohistochemistry - MBP	Human	Multiple sclerosis	Brain
(Underhill, Rostomily, Mikheev, Yuan, & Yarnykh, 2011)	MPF	Histology - LFB	Animal - Mouse	Healthy	Brain
(Chandran et al., 2012)	FA, RD	Immunohistochemistry - MBP	Animal - Mouse	Demyelination - Cuprizone	Brain

(Tardif et al., 2012)	T1, T2, MTR, PD	Immunohistochemistry - MBP	Human	Multiple sclerosis	Brain
(Fjaer et al., 2013)	MTR	Immunohistochemistry - PLP	Animal - Mouse	Demyelination - Cuprizone	Brain
(Harkins, Valentine, Gochberg, & Does, 2013)	MWF, MPF	Microscopy - Myelin fraction	Animal - Rat	Edema - Hexaclorophene	SC
(Janve et al., 2013)	MPF, R1a, k_ba, FA, RD, MD, AD	Histology - LFB	Animal - Rat	Demyelination - Lipopolysaccharide	Brain
(Thiessen et al., 2013)	MPF, R1f, k_fm, k_mf, T2f, T2m, MD, RD, AD, FA, T1, T2	EM - Myelin thickness	Animal - Mouse	Demyelination - Cuprizone	Brain
(Kozlowski, Rosicka, Liu, Yung, & Tetzlaff, 2014)	MWF	Immunohistochemistry - MBP	Animal - Rat	Injury - Dorsal columnar transection	SC
(X. Wang et al., 2014)	RD, RD-DBSI	Immunohistochemistry - MBP	Animal - Mouse	Demyelination - Autoimmune encephalomyelitis	SC
(Fjaer, Bo, Myhr, Torkildsen, & Wergeland, 2015)	MTR	Immunohistochemistry - PLP	Animal - Mouse	Demyelination - Autoimmune encephalomyelitis	Brain

(Seehaus et al., 2015)	FA, RD, MD	Histology - Silver	Human	Healthy	Brain
(Turati et al., 2015)	MPF	Immunohistochemistry - MBP	Animal - Mouse	Demyelination - Cuprizone	Brain
(Y. Wang et al., 2015)	RD-DBSI	Histology - LFB	Human	Multiple sclerosis	SC
(Aojula et al., 2016)	FA, AD, RD, MD	Immunohistochemistry - MBP	Animal - Rat	Hydrocephalus	Brain
(Hakkarainen et al., 2016)	T1, T2, MTR, T1p, T2p, RAFF	Histology - Gold chloride	Animal - Rat	Healthy	Brain
(Jelescu et al., 2016)	RD, RK, AWF, Rde, T2, MTR	EM - Myelin fraction	Animal - Mouse	Demyelination - Cuprizone	Brain
(Kelm et al., 2016)	MD, RD, MK, RK, AWF	EM - Myelin fraction	Animal - Mouse	Demyelination - Knockout	Brain
(Reeves et al., 2016)	T1, T2	Immunohistochemistry - MBP	Human	Epilepsy	Brain
(Tu et al., 2016)	FA, AD, RD, MD, MTR	Immunohistochemistry - MBP	Animal - Rat	Traumatic brain injury	Brain
(Chang et al., 2017)	FA, AD, RD, MD	Immunohistochemistry - MBP	Animal - Mouse	Healthy	Brain
(Chen, Holmes, Liu, Tetzlaff, & Kozlowski, 2017)	MWF	EM - Myelin fraction	Animal - Rat	Injury - Dorsal columnar transection	SC

(M. Y. Khodanovich et al., 2017)	MPF	Histology - LFB	Animal - Mouse	Demyelination - Cuprizone	Brain
(Lehto, Albors, et al., 2017)	RAFF, MTR, T1sat, FA, MD, AD, RD	Histology - Gold chloride	Animal - Rat	Demyelination - Lipopolysaccharide	Brain
(Lehto, Sierra, & Grohn, 2017)	MTR	Histology - Gold chloride	Animal - Rat	Traumatic brain injury	Brain
(van Tilborg et al., 2018)	FA	Immunohistochemistry - MBP	Animal - Rat	White matter injury	Brain
(Beckmann et al., 2018)	MTR	Histology - LFB	Animal - Mouse	Demyelination - Cuprizone	Brain
(Berman, West, Does, Yeatman, & Mezer, 2018)	MTV	EM - Myelin fraction	Animal - Mouse	Demyelination - Knockout	Brain
(Hametner et al., 2018)	R2*, T1, QSM	Histology - LFB	Human	Vascular diseases	Brain
(Praet et al., 2018)	MK, RK, AK, FA, MD, RD, AD	Immunohistochemistry - MBP	Animal - Mouse	Amyloidosis	Brain
(Wendel et al., 2018)	FA, AD, RD, MD	Immunohistochemistry - MBP	Animal - Mouse	Traumatic brain injury	Brain
(West et al., 2018)	MPF, MWF, MVF-T2, MVF-MT	EM - Myelin fraction	Animal - Mouse	Demyelination - Knockout	Brain

(Yano et al., 2018)	FA, RD, MD	Immunohistochemistry - PLP	Animal - Mouse	Demyelination - Cuprizone	Brain
(Abe et al., 2019)	FA, RD, AD	Microscopy - Myelin thickness	Animal - Mouse	Optogenetic stimulation	Brain
(Duhamel et al., 2019)	ihMTR, MTR	Microscopy - Fluorescence	Animal - Mouse	Healthy	Brain
(M. Khodanovich et al., 2019)	MPF	Immunohistochemistry - MBP	Animal - Mouse	Demyelination - Cuprizone	Brain
(Mollink et al., 2019)	FA	Immunohistochemistry - MBP	Human	Amyotrophic lateral sclerosis	Brain
(Peters et al., 2019)	FA, MD	Histology - LFB	Human	Tuberous sclerosis complex	Brain
(Pol et al., 2019)	QSM, FA, MD	Histology - Solochrome	Animal - Mouse	Healthy	Brain
(Soustelle et al., 2019)	MPF, RD, MWF, rSPF	Immunohistochemistry - MBP	Animal - Mouse	Demyelination - Cuprizone	Brain
(Guglielmetti et al., 2020)	MTR, MTR-UTE	Immunohistochemistry - MBP	Animal - Mouse	Healthy	Brain

References

Abe, Y., Komaki, Y., Seki, F., Shibata, S., Okano, H., & Tanaka, K. F. (2019). Correlative study using structural MRI and super-resolution microscopy to detect structural alterations induced by long-term optogenetic stimulation of striatal medium spiny neurons. *Neurochem Int, 125*, 163-174. doi:10.1016/j.neuint.2019.02.017

Aojula, A., Botfield, H., McAllister, J. P., 2nd, Gonzalez, A. M., Abdullah, O., Logan, A., & Sinclair, A. (2016). Diffusion tensor imaging with direct cytopathological validation: characterisation of decorin treatment in experimental juvenile communicating hydrocephalus. *Fluids Barriers CNS, 13*(1), 9. doi:10.1186/s12987-016-0033-2

Beckmann, N., Giorgetti, E., Neuhaus, A., Zurbruegg, S., Accart, N., Smith, P., . . . Shimshek, D. R. (2018). Brain region-specific enhancement of remyelination and prevention of demyelination by the CSF1R kinase inhibitor BLZ945. *Acta Neuropathol Commun, 6*(1), 9. doi:10.1186/s40478-018-0510-8

Berman, S., West, K. L., Does, M. D., Yeatman, J. D., & Mezer, A. A. (2018). Evaluating g-ratio weighted changes in the corpus callosum as a function of age and sex. *Neuroimage, 182*, 304-313. doi:10.1016/j.neuroimage.2017.06.076

Campbell, J. S. W., Leppert, I. R., Narayanan, S., Boudreau, M., Duval, T., Cohen-Adad, J., . . . Stikov, N. (2018). Promise and pitfalls of g-ratio estimation with MRI. *Neuroimage, 182*, 80-96. doi:https://doi.org/10.1016/j.neuroimage.2017.08.038

Cercignani, M., Dowell, N. G., & Tofts, P. S. (2018). *Quantitative MRI of the brain* (2nd edition ed.). Boca Raton: CRC Press.

Chandran, P., Upadhyay, J., Markosyan, S., Lisowski, A., Buck, W., Chin, C. L., . . . Day, M. (2012). Magnetic resonance imaging and histological evidence for the blockade of cuprizone-induced demyelination in C57BL/6 mice. *Neuroscience, 202*, 446-453. doi:10.1016/j.neuroscience.2011.10.051

Chang, E. H., Argyelan, M., Aggarwal, M., Chandon, T. S., Karlsgodt, K. H., Mori, S., & Malhotra, A. K. (2017). Diffusion tensor imaging measures of white matter compared to myelin basic protein immunofluorescence in tissue cleared intact brains. *Data Brief, 10*, 438-443. doi:10.1016/j.dib.2016.12.018

Chen, H. S., Holmes, N., Liu, J., Tetzlaff, W., & Kozlowski, P. (2017). Validating myelin water imaging with transmission electron microscopy in a rat spinal cord injury model. *Neuroimage, 153*, 122-130. doi:10.1016/j.neuroimage.2017.03.065

Cohen-Adad, J. (2018). Microstructural imaging in the spinal cord and validation strategies. *Neuroimage, 182*, 169-183. doi:10.1016/j.neuroimage.2018.04.009

Cohen-Adad, J., & Wheeler-Kingshott, C. A. (2014). *Quantitative MRI of the spinal cord*: Academic Press.

Duhamel, G., Prevost, V. H., Cayre, M., Hertanu, A., McHinda, S., Carvalho, V. N., . . . Girard, O. M. (2019). Validating the sensitivity of inhomogeneous magnetization transfer (ihMT) MRI to myelin with fluorescence microscopy. *Neuroimage, 199*, 289-303. doi:10.1016/j.neuroimage.2019.05.061

Faizy, T. D., Kumar, D., Broocks, G., Thaler, C., Flottmann, F., Leischner, H., . . . Gellißen, S. (2018). Age-Related Measurements of the Myelin Water Fraction

derived from 3D multi-echo GRASE reflect Myelin Content of the Cerebral White Matter. *Sci Rep, 8*(1), 14991. doi:10.1038/s41598-018-33112-8

Fatemi, A., Wilson, M. A., Phillips, A. W., McMahon, M. T., Zhang, J., Smith, S. A., . . . Johnston, M. V. (2011). In vivo magnetization transfer MRI shows dysmyelination in an ischemic mouse model of periventricular leukomalacia. *J Cereb Blood Flow Metab, 31*(10), 2009-2018. doi:10.1038/jcbfm.2011.68

Fjaer, S., Bo, L., Lundervold, A., Myhr, K. M., Pavlin, T., Torkildsen, O., & Wergeland, S. (2013). Deep gray matter demyelination detected by magnetization transfer ratio in the cuprizone model. *PLoS One, 8*(12), e84162. doi:10.1371/journal.pone.0084162

Fjaer, S., Bo, L., Myhr, K. M., Torkildsen, O., & Wergeland, S. (2015). Magnetization transfer ratio does not correlate to myelin content in the brain in the MOG-EAE mouse model. *Neurochem Int, 83-84*, 28-40. doi:10.1016/j.neuint.2015.02.006

Guglielmetti, C., Boucneau, T., Cao, P., Van der Linden, A., Larson, P. E. Z., & Chaumeil, M. M. (2020). Longitudinal evaluation of demyelinated lesions in a multiple sclerosis model using ultrashort echo time magnetization transfer (UTE-MT) imaging. *Neuroimage, 208*, 116415. doi:10.1016/j.neuroimage.2019.116415

Hakkarainen, H., Sierra, A., Mangia, S., Garwood, M., Michaeli, S., Grohn, O., & Liimatainen, T. (2016). MRI relaxation in the presence of fictitious fields correlates with myelin content in normal rat brain. *Magn Reson Med, 75*(1), 161-168. doi:10.1002/mrm.25590

Hametner, S., Endmayr, V., Deistung, A., Palmrich, P., Prihoda, M., Haimburger, E., . . . Grabner, G. (2018). The influence of brain iron and myelin on magnetic susceptibility and effective transverse relaxation - A biochemical and histological validation study. *Neuroimage, 179*, 117-133. doi:10.1016/j.neuroimage.2018.06.007

Harkins, K. D., Valentine, W. M., Gochberg, D. F., & Does, M. D. (2013). In-vivo multi-exponential T2, magnetization transfer and quantitative histology in a rat model of intramyelinic edema. *Neuroimage Clin, 2*, 810-817. doi:10.1016/j.nicl.2013.06.007

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*(11), 1539-1558. doi:10.1002/sim.1186

Janve, V. A., Zu, Z., Yao, S. Y., Li, K., Zhang, F. L., Wilson, K. J., . . . Gochberg, D. F. (2013). The radial diffusivity and magnetization transfer pool size ratio are sensitive markers for demyelination in a rat model of type III multiple sclerosis (MS) lesions. *Neuroimage, 74*, 298-305. doi:10.1016/j.neuroimage.2013.02.034

Jelescu, I. O., Zurek, M., Winters, K. V., Veraart, J., Rajaratnam, A., Kim, N. S., . . . Fieremans, E. (2016). In vivo quantification of demyelination and recovery using compartment-specific diffusion MRI metrics validated by electron microscopy. *Neuroimage, 132*, 104-114. doi:10.1016/j.neuroimage.2016.02.004

Jito, J., Nakasu, S., Ito, R., Fukami, T., Morikawa, S., & Inubushi, T. (2008). Maturational changes in diffusion anisotropy in the rat corpus callosum: comparison with quantitative histological evaluation. *J Magn Reson Imaging, 28*(4), 847-854. doi:10.1002/jmri.21496

Kelm, N. D., West, K. L., Carson, R. P., Gochberg, D. F., Ess, K. C., & Does, M. D. (2016). Evaluation of diffusion kurtosis imaging in ex vivo hypomyelinated mouse brains. *Neuroimage, 124*(Pt A), 612-626. doi:10.1016/j.neuroimage.2015.09.028

Khodanovich, M., Pishchelko, A., Glazacheva, V., Pan, E., Akulov, A., Svetlik, M., . . . Yarnykh, V. (2019). Quantitative Imaging of White and Gray Matter Remyelination in the Cuprizone Demyelination Model Using the Macromolecular Proton Fraction. *Cells, 8*(10). doi:10.3390/cells8101204

Khodanovich, M. Y., Sorokina, I. V., Glazacheva, V. Y., Akulov, A. E., Nemirovich-Danchenko, N. M., Romashchenko, A. V., . . . Yarnykh, V. L. (2017). Histological validation of fast macromolecular proton fraction mapping as a quantitative myelin imaging method in the cuprizone demyelination model. *Sci Rep, 7*, 46686. doi:10.1038/srep46686

Kozlowski, P., Raj, D., Liu, J., Lam, C., Yung, A. C., & Tetzlaff, W. (2008). Characterizing white matter damage in rat spinal cord with quantitative MRI and histology. *J Neurotrauma, 25*(6), 653-676. doi:10.1089/neu.2007.0462

Kozlowski, P., Rosicka, P., Liu, J., Yung, A. C., & Tetzlaff, W. (2014). In vivo longitudinal Myelin Water Imaging in rat spinal cord following dorsal column transection injury. *Magn Reson Imaging, 32*(3), 250-258. doi:10.1016/j.mri.2013.12.006

Laule, C., Kozlowski, P., Leung, E., Li, D. K., Mackay, A. L., & Moore, G. R. (2008). Myelin water imaging of multiple sclerosis at 7 T: correlations with histopathology. *Neuroimage, 40*(4), 1575-1580. doi:10.1016/j.neuroimage.2007.12.008

Laule, C., Leung, E., Li, D. K. B., Traboulsee, A. L., Paty, D. W., MacKay, A. L., & Moore, G. R. W. (2006). Myelin water imaging in multiple sclerosis: quantitative correlations with histopathology. *Multiple Sclerosis, 12*, 747-753. doi:10.1177/1352458506070928

Laule, C., & Moore, G. R. W. (2018). Myelin water imaging to detect demyelination and remyelination and its validation in pathology. *Brain Pathol, 28*(5), 750-764. doi:10.1111/bpa.12645

Laule, C., Vavasour, I. M., Kolind, S. H., Li, D. K. B., Traboulsee, T. L., Moore, G. R. W., & MacKay, A. L. (2007). Magnetic Resonance Imaging of Myelin. *Neurotherapeutics, 4*(3), 460-484. doi:10.1016/j.nurt.2007.05.004

Laule, C., Vavasour, I. M., Leung, E., Li, D. K., Kozlowski, P., Traboulsee, A. L., . . . Moore, G. R. (2011). Pathological basis of diffusely abnormal white matter: insights from magnetic resonance imaging and histology. *Mult Scler, 17*(2), 144-150. doi:10.1177/1352458510384008

Lehto, L. J., Albors, A. A., Sierra, A., Tolppanen, L., Eberly, L. E., Mangia, S., . . . Grohn, O. (2017). Lysophosphatidyl Choline Induced Demyelination in Rat Probed by Relaxation along a Fictitious Field in High Rank Rotating Frame. *Front Neurosci, 11*, 433. doi:10.3389/fnins.2017.00433

Lehto, L. J., Sierra, A., & Grohn, O. (2017). Magnetization transfer SWIFT MRI consistently detects histologically verified myelin loss in the thalamocortical pathway after a traumatic brain injury in rat. *NMR Biomed, 30*(2). doi:10.1002/nbm.3678

MacKay, A. L., & Laule, C. (2016). Magnetic Resonance of Myelin Water: An in vivo Marker for Myelin. *Brain Plast, 2*(1), 71-91. doi:10.3233/BPL-160033

Mollink, J., Hiemstra, M., Miller, K. L., Huszar, I. N., Jenkinson, M., Raaphorst, J., . . . van Cappellen van Walsum, A. M. (2019). White matter changes in the perforant path area in patients with amyotrophic lateral sclerosis. *Neuropathol Appl Neurobiol, 45*(6), 570-585. doi:10.1111/nan.12555

Nave, K. A., & Werner, H. B. (2014). Myelination of the nervous system: mechanisms and functions. *Annu Rev Cell Dev Biol, 30*, 503-533. doi:10.1146/annurev-cellbio-100913-013101

Odrobina, E. E., Lam, T. Y., Pun, T., Midha, R., & Stanisz, G. J. (2005). MR properties of excised neural tissue following experimentally induced demyelination. *NMR Biomed, 18*(5), 277-284. doi:10.1002/nbm.951

Peters, J. M., Struyven, R. R., Prohl, A. K., Vasung, L., Stajduhar, A., Taquet, M., . . . Warfield, S. K. (2019). White matter mean diffusivity correlates with myelination in tuberous sclerosis complex. *Ann Clin Transl Neurol, 6*(7), 1178-1190. doi:10.1002/acn3.793

Petiet, A., Adanyeguh, I., Aigrot, M. S., Poirion, E., Nait-Oumesmar, B., Santin, M., & Stankoff, B. (2019). Ultrahigh field imaging of myelin disease models: Toward specific markers of myelin integrity? *J Comp Neurol, 527*(13), 2179-2189. doi:10.1002/cne.24598

Pol, S., Sveinsson, M., Sudyn, M., Babek, N., Siebert, D., Bertolino, N., . . . Zivadinov, R. (2019). Teriflunomide's Effect on Glia in Experimental Demyelinating Disease: A Neuroimaging and Histologic Study. *J Neuroimaging, 29*(1), 52-61. doi:10.1111/jon.12561

Praet, J., Manyakov, N. V., Muchene, L., Mai, Z., Terzopoulos, V., de Backer, S., . . . Verhoye, M. (2018). Diffusion kurtosis imaging allows the early detection and longitudinal follow-up of amyloid-beta-induced pathology. *Alzheimers Res Ther, 10*(1), 1. doi:10.1186/s13195-017-0329-8

Pun, T. W. C., Odrobina, E., Xu, Q. G., Lam, T. Y. J., Munro, C. A., Midha, R., & Stanisz, G. J. (2005). Histological and magnetic resonance analysis of sciatic nerves in the tellurium model of neuropathy. *Journal of the Peripheral Nervous System, 10*(1), 38-46. doi:10.1111/j.1085-9489.2005.10107.x

Raudenbush, S. W. (2009). Analyzing effect sizes: Random-effects models. In *The handbook of research synthesis and meta-analysis, 2nd ed.* (pp. 295-315). New York, NY, US: Russell Sage Foundation.

Reeves, C., Tachrount, M., Thomas, D., Michalak, Z., Liu, J., Ellis, M., . . . Thom, M. (2016). Combined Ex Vivo 9.4T MRI and Quantitative Histopathological Study in Normal and Pathological Neocortical Resections in Focal Epilepsy. *Brain Pathol, 26*(3), 319-333. doi:10.1111/bpa.12298

Sampaio-Baptista, C., & Johansen-Berg, H. (2017). White Matter Plasticity in the Adult Brain. *Neuron, 96*(6), 1239-1251. doi:10.1016/j.neuron.2017.11.026

Schmierer, K., Parkes, H. G., So, P. W., An, S. F., Brandner, S., Ordidge, R. J., . . . Miller, D. H. (2010). High field (9.4 Tesla) magnetic resonance imaging of cortical grey matter lesions in multiple sclerosis. *Brain, 133*(Pt 3), 858-867. doi:10.1093/brain/awp335

Schmierer, K., Scaravilli, F., Altmann, D. R., Barker, G. J., & Miller, D. H. (2004). Magnetization transfer ratio and myelin in postmortem multiple sclerosis brain. *Ann Neurol, 56*(3), 407-415. doi:10.1002/ana.20202

Schmierer, K., Tozer, D. J., Scaravilli, F., Altmann, D. R., Barker, G. J., Tofts, P. S., & Miller, D. H. (2007). Quantitative magnetization transfer imaging in postmortem

multiple sclerosis brain. *J Magn Reson Imaging, 26*(1), 41-51. doi:10.1002/jmri.20984

Schmierer, K., Wheeler-Kingshott, C. A., Boulby, P. A., Scaravilli, F., Altmann, D. R., Barker, G. J., . . . Miller, D. H. (2007). Diffusion tensor imaging of post mortem multiple sclerosis brain. *Neuroimage, 35*(2), 467-477. doi:10.1016/j.neuroimage.2006.12.010

Schmierer, K., Wheeler-Kingshott, C. A., Tozer, D. J., Boulby, P. A., Parkes, H. G., Yousry, T. A., . . . Miller, D. H. (2008). Quantitative magnetic resonance of postmortem multiple sclerosis brain before and after fixation. *Magn Reson Med, 59*(2), 268-277. doi:10.1002/mrm.21487

Seehaus, A., Roebroeck, A., Bastiani, M., Fonseca, L., Bratzke, H., Lori, N., . . . Galuske, R. (2015). Histological validation of high-resolution DTI in human post mortem tissue. *Front Neuroanat, 9*, 98. doi:10.3389/fnana.2015.00098

Soustelle, L., Antal, M. C., Lamy, J., Rousseau, F., Armspach, J. P., & Loureiro de Sousa, P. (2019). Correlations of quantitative MRI metrics with myelin basic protein (MBP) staining in a murine model of demyelination. *NMR Biomed, 32*(9), e4116. doi:10.1002/nbm.4116

Takagi, T., Nakamura, M., Yamada, M., Hikishima, K., Momoshima, S., Fujiyoshi, K., . . . Okano, H. (2009). Visualization of peripheral nerve degeneration and regeneration: monitoring with diffusion tensor tractography. *Neuroimage, 44*(3), 884-892. doi:10.1016/j.neuroimage.2008.09.022

Tardif, C. L., Bedell, B. J., Eskildsen, S. F., Collins, D. L., & Pike, G. B. (2012). Quantitative magnetic resonance imaging of cortical multiple sclerosis pathology. *Mult Scler Int, 2012*, 742018. doi:10.1155/2012/742018

Thiessen, J. D., Zhang, Y., Zhang, H., Wang, L., Buist, R., Del Bigio, M. R., . . . Martin, M. (2013). Quantitative MRI and ultrastructural examination of the cuprizone mouse model of demyelination. *NMR Biomed, 26*(11), 1562-1581. doi:10.1002/nbm.2992

Tu, T. W., Williams, R. A., Lescher, J. D., Jikaria, N., Turtzo, L. C., & Frank, J. A. (2016). Radiological-pathological correlation of diffusion tensor and magnetization transfer imaging in a closed head traumatic brain injury model. *Ann Neurol, 79*(6), 907-920. doi:10.1002/ana.24641

Turati, L., Moscatelli, M., Mastropietro, A., Dowell, N. G., Zucca, I., Erbetta, A., . . . Minati, L. (2015). In vivo quantitative magnetization transfer imaging correlates with histology during de- and remyelination in cuprizone-treated mice. *NMR Biomed, 28*(3), 327-337. doi:10.1002/nbm.3253

Turner, R. (2019). Myelin and Modeling: Bootstrapping Cortical Microcircuits. *Front Neural Circuits, 13*, 34. doi:10.3389/fncir.2019.00034

Underhill, H. R., Rostomily, R. C., Mikheev, A. M., Yuan, C., & Yarnykh, V. L. (2011). Fast bound pool fraction imaging of the in vivo rat brain: association with myelin content and validation in the C6 glioma model. *Neuroimage, 54*(3), 2052-2065. doi:10.1016/j.neuroimage.2010.10.065

van Tilborg, E., Achterberg, E. J. M., van Kammen, C. M., van der Toorn, A., Groenendaal, F., Dijkhuizen, R. M., . . . Nijboer, C. H. A. (2018). Combined fetal inflammation and postnatal hypoxia causes myelin deficits and autism-like behavior in a rat model of diffuse white matter injury. *Glia, 66*(1), 78-93. doi:10.1002/glia.23216

Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software, 36*(3). doi:10.18637/jss.v036.i03

Wang, S., Wu, E. X., Cai, K., Lau, H. F., Cheung, P. T., & Khong, P. L. (2009). Mild hypoxic-ischemic injury in the neonatal rat brain: longitudinal evaluation of white matter using diffusion tensor MR imaging. *AJNR Am J Neuroradiol, 30*(10), 1907-1913. doi:10.3174/ajnr.A1697

Wang, X., Cusick, M. F., Wang, Y., Sun, P., Libbey, J. E., Trinkaus, K., . . . Song, S. K. (2014). Diffusion basis spectrum imaging detects and distinguishes coexisting subclinical inflammation, demyelination and axonal injury in experimental autoimmune encephalomyelitis mice. *NMR Biomed, 27*(7), 843-852. doi:10.1002/nbm.3129

Wang, Y., Sun, P., Wang, Q., Trinkaus, K., Schmidt, R. E., Naismith, R. T., . . . Song, S. K. (2015). Differentiation and quantification of inflammation, demyelination and axon injury or loss in multiple sclerosis. *Brain, 138*(Pt 5), 1223-1238. doi:10.1093/brain/awv046

Wendel, K. M., Lee, J. B., Affeldt, B. M., Hamer, M., Harahap-Carrillo, I. S., Pardo, A. C., & Obenaus, A. (2018). Corpus Callosum Vasculature Predicts White Matter Microstructure Abnormalities after Pediatric Mild Traumatic Brain Injury. *J Neurotrauma*. doi:10.1089/neu.2018.5670

West, K. L., Kelm, N. D., Carson, R. P., Gochberg, D. F., Ess, K. C., & Does, M. D. (2018). Myelin volume fraction imaging with MRI. *Neuroimage, 182*, 511-521. doi:10.1016/j.neuroimage.2016.12.067

Wu, Q. Z., Yang, Q., Cate, H. S., Kemper, D., Binder, M., Wang, H. X., . . . Egan, G. F. (2008). MRI identification of the rostral-caudal pattern of pathology within the corpus callosum in the cuprizone mouse model. *J Magn Reson Imaging, 27*(3), 446-453. doi:10.1002/jmri.21111

Yano, R., Hata, J., Abe, Y., Seki, F., Yoshida, K., Komaki, Y., . . . Tanaka, K. F. (2018). Quantitative temporal changes in DTI values coupled with histological properties in cuprizone-induced demyelination and remyelination. *Neurochem Int, 119*, 151-158. doi:10.1016/j.neuint.2017.10.004

Zaaraoui, W., Deloire, M., Merle, M., Girard, C., Raffard, G., Biran, M., . . . Dousset, V. (2008). Monitoring demyelination and remyelination by magnetization transfer imaging in the mouse brain at 9.4 T. *MAGMA, 21*(5), 357-362. doi:10.1007/s10334-008-0141-3

Zhang, J., Jones, M., DeBoy, C. A., Reich, D. S., Farrell, J. A., Hoffman, P. N., . . . Calabresi, P. A. (2009). Diffusion tensor magnetic resonance imaging of Wallerian degeneration in rat spinal cord after dorsal root axotomy. *J Neurosci, 29*(10), 3160-3171. doi:10.1523/JNEUROSCI.3941-08.2009