NeuroImage: Clinical 12 (2016) 293–299

Contents lists available at ScienceDirect

NeuroImage: Clinical

journal homepage: www.elsevier.com/locate/ynicl

PREVAIL: Predicting Recovery through Estimation and Visualization of Active and Incident Lesions

Jordan D. Dworkin[a,*], Elizabeth M. Sweeney[b], Matthew K. Schindler[c], Salim Chahin[d], Daniel S. Reich[b,c], Russell T. Shinohara[a]

[a]Department of Biostatistics and Epidemiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, United States
[b]Department of Biostatistics, The Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD 21205, United States
[c]Translational Neuroradiology Unit, Division of Neuroimmunology and Neurovirology, National Institute of Neurological Disease and Stroke, National Institute of Health, Bethesda, MD 20892, United States
[d]Multiple Sclerosis Division of the Department of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, United States

ARTICLE INFO

Article history:
Received 24 June 2016
Received in revised form 21 July 2016
Accepted 30 July 2016
Available online 2 August 2016

Keywords:
Multiple sclerosis
Prediction
Lesion
Neuroimaging
MRI

ABSTRACT

Objective: The goal of this study was to develop a model that integrates imaging and clinical information observed at lesion incidence for predicting the recovery of white matter lesions in multiple sclerosis (MS) patients.
Methods: Demographic, clinical, and magnetic resonance imaging (MRI) data were obtained from 60 subjects with MS as part of a natural history study at the National Institute of Neurological Disorders and Stroke. A total of 401 lesions met the inclusion criteria and were used in the study. Imaging features were extracted from the intensity-normalized T_1-weighted (T_1w) and T_2-weighted sequences as well as magnetization transfer ratio (MTR) sequence acquired at lesion incidence. T_1w and MTR signatures were also extracted from images acquired one-year post-incidence. Imaging features were integrated with clinical and demographic data observed at lesion incidence to create statistical prediction models for long-term damage within the lesion.
Validation: The performance of the T_1w and MTR predictions was assessed in two ways: first, the predictive accuracy was measured quantitatively using leave-one-lesion-out cross-validated (CV) mean-squared predictive error. Then, to assess the prediction performance from the perspective of expert clinicians, three board-certified MS clinicians were asked to individually score how similar the CV model-predicted one-year appearance was to the true one-year appearance for a random sample of 100 lesions.
Results: The cross-validated root-mean-square predictive error was 0.95 for normalized T_1w and 0.064 for MTR, compared to the estimated measurement errors of 0.48 and 0.078 respectively. The three expert raters agreed that T_1w and MTR predictions closely resembled the true one-year follow-up appearance of the lesions in both degree and pattern of recovery within lesions.
Conclusion: This study demonstrates that by using only information from a single visit at incidence, we can predict how a new lesion will recover using relatively simple statistical techniques. The potential to visualize the likely course of recovery has implications for clinical decision-making, as well as trial enrichment.

1. Introduction

Multiple sclerosis (MS) is an inflammatory disease of the central nervous system, which is typically characterized by demyelinating lesions that occur in the brain and spinal cord. These lesions evolve dynamically from actively inflamed tissue over a period of months to more stable demyelinated regions of acute long-term axonal injury (Lassmann, 2013; Lassmann et al., 2007). A competing process of remyelination is also known to occur to varying degrees in patients, and has been documented in both relapsing-remitting and progressive cases (Patrikios et al., 2006; Bramow et al., 2010). Both the destructive and remyelinating processes are known to progress through the disease course (Frischer et al., 2015), and are associated with disability and morbidity. As therapeutics designed to promote tissue repair and remyelination are being developed, sensitive markers for in vivo assessment of these processes are increasingly important for studying therapeutic efficacy and patient management.

Magnetic resonance imaging (MRI) is a commonly used technique for identifying lesions, particularly in the white matter of the brain (Radü & Sahraian, 2008). The presence of new and active lesions is a key factor in the diagnosis and monitoring of MS, and several MRI sequences have been demonstrated to be effective in measuring the severity of these lesions (Polman et al., 2011; Sweeney et al., 2016; Sweeney et al., 2013; Pike et al., 2000). In recent years, successful attempts have been made to utilize quantitative methods in concert

* Corresponding author.
E-mail address: jdwor@mail.med.upenn.edu (J.D. Dworkin).

http://dx.doi.org/10.1016/j.nicl.2016.07.015

with MRI for the study of tissue damage in lesions. These techniques have included the use of advanced quantitative MRI sequences including T_1 mapping (Larsson et al., 1989; Vrenken et al., 2006), magnetization transfer imaging (van Waesberghe et al., 1998; van Waesberghe et al., 1999), and diffusion tensor imaging (Narayanan et al., 1997; Werring et al., 1999; Filippi et al., 2001), as well as statistical techniques for modeling tissue damage using conventional MRI (Shinohara et al., 2011; Mejia et al., 2015; Reich et al., 2015) and the development of time-series models to examine lesion activity (Sweeney et al., 2016; Meier et al., 2007; Meier & Guttmann, 2003; Meier & Guttmann, 2006).

Specifically, much research has engaged with the apparent paradox related to the lack of coherence between the presence of lesions and clinical disease measures (Barkhof, 2002). One recent study retrospectively related the longitudinal behavior of lesions, as opposed to simply their presence, to clinical covariates and treatment status (Sweeney et al., 2016). Significant relationships between treatment and longitudinal behavior indicated that receiving disease-modifying therapy or steroids was associated with a better healing trajectory within lesion tissue. These findings signify the presence of potentially important relationships between the repair processes in the brain, therapeutics, and disability.

Unfortunately, today there is still relatively little that can be determined in advance about the way specific lesions will recover, or the degree to which they may be responsive to treatment. The ability to visually examine the likely course of recovery for a given incident lesion would have the potential to be useful in several settings. Specifically, such visualizations could be a beneficial tool for physicians, providing them important supplemental information when making treatment decisions. Additionally, knowledge of how patients' brains are likely to recover from lesion damage could be beneficial in clinical trials, for which advanced knowledge of lesion characteristics could inform recruitment enrichment and trial design.

To build on the previous work, and to address the needs outlined above, the current study attempted to develop a statistical model that would be capable of prospectively predicting how lesions would heal over the course of a year. In this paper, we discuss the development of such prediction models for two outcome MRI modalities, we present statistical and clinical measures of validity and prediction accuracy, and we discuss the implications and potential next steps of this line of research.

2. Methods

2.1. Image acquisition and preprocessing

Details of the image acquisition and preprocessing have been previously published (Sweeney et al., 2016) and are summarized in this section. Whole-brain two-dimensional T2-weighted FLAIR, PD, T2, and three-dimensional T1-weighted volumes were acquired in a 1.5 tesla (T) MRI scanner (Signa Excite HDxt; GE Healthcare, Milwaukee, Wisconsin) using the body coil for transmission. The 2D FLAIR, PD, and T2 volumes were acquired using fast-spin-echo sequences, and the 3D T_1 volume was acquired using a gradient-echo sequence. All scanning parameters were clinically optimized for each acquired image.

For image preprocessing, we used Medical Image Processing Analysis and Visualization (http://mipav.cit.nih.gov) and the Java Image Science Toolkit (http://www.nitrc.org/projects/jist) (Lucas et al., 2010). All images for each subject at each visit were interpolated to a voxel size of 1 mm^3 and rigidly co-registered longitudinally and across sequences to a template space (Fonov et al., 2011). To coregister the T_1 images across study visits, a two-step procedure was applied: first, subject-specific templates were generated by averaging after rigid alignment of the T_1 images to the MNI template. Second, all T_1 images were then realigned to the subject-specific templates. Finally, the additional MRI sequences were aligned to the T_1 images within each study visit and this transformation was composed with the T_1-based transformation to the subject-specific template.

Extracerebral voxels were removed using a skull-stripping procedure (Carass et al., 2007) and the brain was automatically segmented using the T_1 and FLAIR images (Shiee et al., 2010) to produce a mask of normal-appearing white matter (NAWM), or white matter excluding lesions. Intensity normalization was then conducted using z-scoring based on the mean and variance of the variability in the NAWM (Shinohara et al., 2011; Shinohara et al., 2014). After preprocessing, studies were manually quality controlled by a researcher with over five years' experience with structural MRI (EMS) and studies with motion or other artifacts were removed.

2.2. Patient demographics

For this study, 60 subjects diagnosed with MS were scanned between 2000 and 2008 on a monthly basis over a period of up to 5.5 years ($mean = 2.2$ years, $sd = 1.2$) as part of a natural history study at the National Institute of Neurological Disorders and Stroke in Bethesda, Maryland. To be included in the analysis, subjects were required to meet certain pre-specified inclusion criteria. Specifically, only subjects with at least one new lesion during the observation period were included, and these subjects were required to have been rescanned at least twice 360 days after lesion incidence. 32 subjects met these criteria and were included in the analyses. The 32 subjects ranged from 18 to 60 years of age, with a mean age of 37 years ($sd = 9$). Of the 32 subjects, 11 were male and 21 were female. The majority of the subjects ($n = 27$) were diagnosed with relapsing-remitting MS, and the remaining five were characterized as secondary-progressive. Subjects were either untreated or treated with a variety of disease-modifying therapies during the observation period, including both FDA-approved therapies (Avonex, Betaseron, Daclizumab, and Rebif) and experimental therapies.

2.3. Prediction model

2.3.1. Outcomes

The outcomes of interest in this study were 1) normalized T_1-weighted voxel intensity (nT_1w) (Shinohara et al., 2014) and 2) MTR voxel intensity approximately one-year post-incidence, and is denoted by $Y_{post,i}(v)$ for subject i in voxel v. Due to the noise inherent in both sequences, outcome variables were created by averaging the intensity of each voxel at the visit immediately following the 360-day cutoff (referred to as the one-year visit), the visit prior to the one-year visit ($mean = 10.6$ months from incidence, $sd = 1.3$ months), and the visit following the one-year visit. Because no change is expected in the lesion after that length of time, this average only reduced variability due to measurement error (Meier et al., 2007). Thus, the average score represents a more precise estimate of true voxel intensity than the one-year visit intensity alone.

2.3.2. Predictors

A dataset made up of scan data and relevant demographic variables was created to predict the one-year post-incidence voxel intensities. For each voxel, this included the MTR as well as the nFLAIR, nPD, nT_2w, and pre- and post-contrast nT_1w intensities at incidence, denoted by $\boldsymbol{Y_{inc,i}}(v)$. After applying a 3D Gaussian smoother with variance parameter 3 mm and width 5 mm, each voxel's blurred intensities on the five scan modalities, $G\boldsymbol{Y_{inc,i}}(v)$, were also included, as well as the distance, in number of voxels, from the voxel to the nearest boundary of the lesion, $d_i(v)$, and the size, in number of voxels, of the lesion, $s_i(v)$. Additional predictors $\boldsymbol{X_i}$ included were the patient's age, sex, disease subtype, expanded disability status score (EDSS; (Kurtzke, 1983)), disease-modifying treatment status (treated versus untreated, with use of one or more therapies counting as treated), and steroid status (receiving steroids versus not on steroids) at the time of lesion incidence.

2.3.3. Model creation and validation

All statistical modeling was conducted in the R statistical environment (R: a Language and Environment for Statistical Computing, 2015). Separate linear regression models for nT_1w intensity one-year post-incidence and MTR intensity one-year post-incidence were created using all of the variables in the dataset, as well as the interaction between voxel distance to lesion boundary and lesion size:

$$Y_{post,i}(v) = \alpha + \boldsymbol{X_i'}\boldsymbol{\beta} + \boldsymbol{Y_{inc,i}}(v)'\boldsymbol{\gamma} + G\boldsymbol{Y_{inc,i}}(v)'\boldsymbol{\gamma} + d_i(\boldsymbol{v})\delta + d_i(\boldsymbol{v})s_i(\boldsymbol{v})\delta_{int} + \epsilon_i(v)$$

Predictions were obtained using leave-one-lesion-out cross-validation. In this cross-validation technique, the prediction for each voxel in a given lesion is made using a model trained on all of the data except those from that lesion. As a result, 397 models were developed, each excluding one of the 397 lesions in the dataset. This method ensures that the prediction for each voxel is not influenced by the true outcome of that voxel. Secondary cross-validation was also performed using a leave-one-subject-out technique, but due to the small sample size and large number of variables in the model, performance was assessed on the predictions obtained by the leave-one-lesion-out procedure.

2.4. Performance assessment

The performance of the model was assessed in two ways. First, prediction accuracy was measured quantitatively by calculating the root mean square error (RMSE) of both the nT_1w and MTR predicted intensities for estimating the average one-year intensity outcomes. This measure gives an estimate of the average difference between the predicted intensity and observed intensity of the voxels in the dataset. Because the RMSE is dependent on the scale of the outcome, the measurement error of voxel intensity was estimated for comparison. The measurement errors for T_1w and MTR were estimated by calculating the RMSE of the voxel intensities at the pre-one-year visit for estimating the average one-year intensity outcomes. This directly compares the accuracy of the model's prediction using information at incidence to the accuracy of a scan taken at approximately 11 months (*mean* = 10.6, *sd* = 1.3 months), for predicting the average intensity of a voxel at one year.

The second method for assessing accuracy was a rater study. Three board certified MS clinicians (two neurologists and one neuroradiologist) with between 5 and 12 years of research experience in MS participated in this validation. To assess the prediction performance visually, the raters were asked to individually score how similar the model-predicted one-year appearance was to the observed one-year appearance for a random sample of 100 lesions. Raters viewed images of the lesion voxels' intensities at incidence, their predicted intensities one-year post-incidence, and their observed intensities one-year post-incidence.

For each lesion in the sample, raters were asked to determine "overall how well the prediction reflects the appearance of the lesion after one year," "how well the degree of recovery in the prediction reflects the degree of recovery after one year," and "how well the pattern of recovery in the prediction reflects the pattern of recovery after one year."

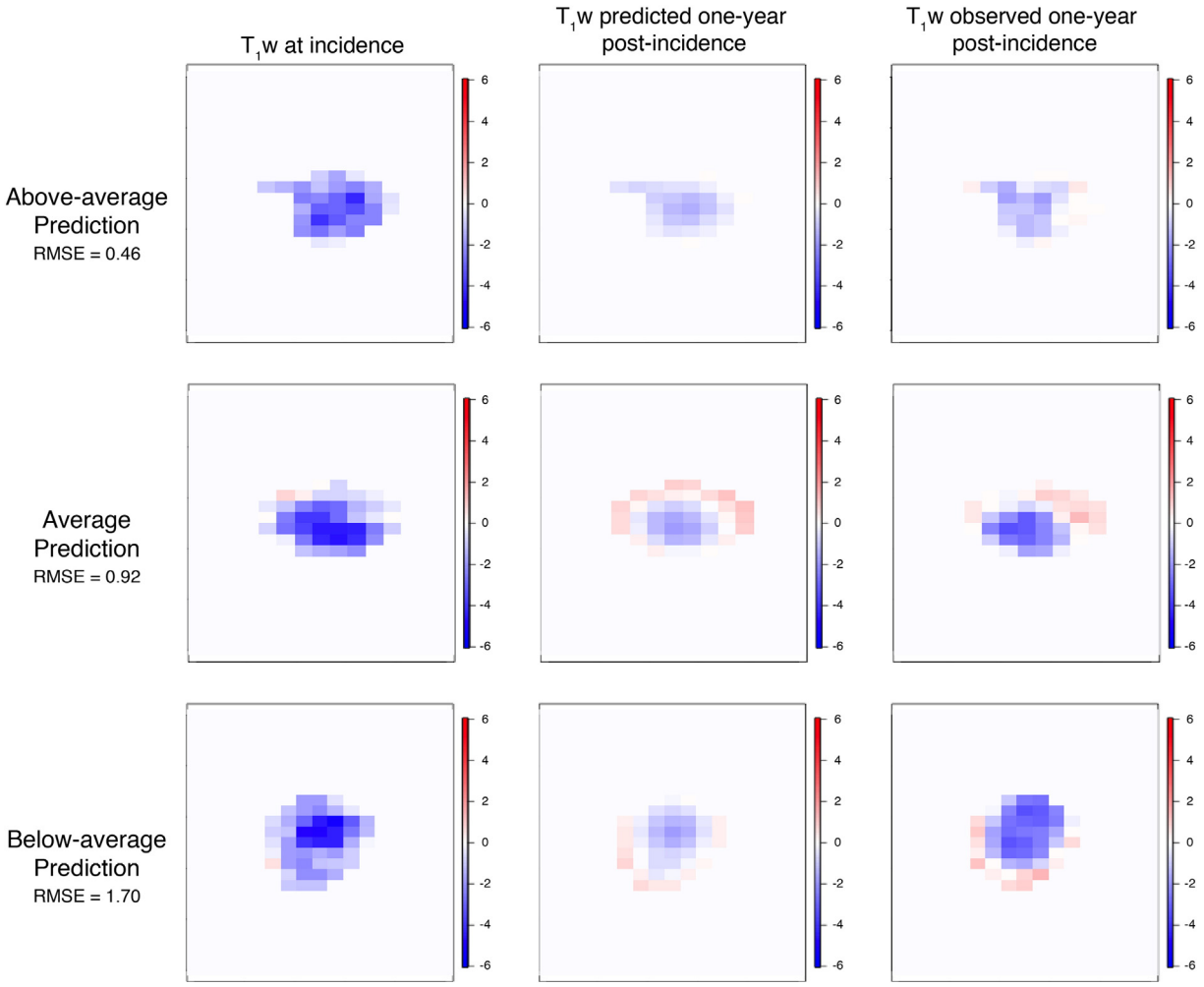

Fig. 1. Representative predictions for the nT_1w model. Images are axial slices of lesions, with rows representing three example lesions with varying levels of predictive accuracy. For nT_1w intensities, red areas represent hyperintensity and blue areas represent hypointensity, with 0 (white) representing the intensity of normal-appearing white matter. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Each question was asked for both the nT_1w and MTR predictions, resulting in six ratings per lesion. Ratings were given on a 1-to-4 scale, with labels of "1 - Failed miserably," "2 - Some redeeming features," "3 - Passed with minor errors," and "4 - Passed." Raters were broadly instructed that "Failed miserably" indicated no correspondence whatsoever between prediction and observed images, "Some redeeming features" indicated some correspondence, "Passed with minor errors" indicated correspondence, and "Passed" indicated excellent correspondence. Images were observed privately and ratings were given independently, with no discussion by raters occurring during the rating process.

3. Results

3.1. RMSE

Using this measure, both cross-validated models performed well. The overall RMSE of the T_1w prediction was 0.95 (*mean of lesion RMSEs* = 0.89, *sd* = 0.35), as compared to a measurement error in T_1w scans of 0.48. This indicates that the correspondence between the predicted intensity and the observed intensity of each voxel was only slightly (approximately one-half standard deviation unit) lower than the correspondence between the voxel intensity of a scan taken at approximately 11-months and the true 12-month observed intensity. The RMSE of the MTR prediction was 0.064 (*mean of lesion RMSEs* = 0.057, *sd* = 0.023), as compared to a measurement error in MTR scans of 0.078, where the average MTR outcome value was.36. This demonstrates that the correspondence between the model-predicted intensity and the observed intensity of each voxel was better than the correspondence between the voxel intensities approximately one month apart. This is consistent with the literature, as MTR has been demonstrated to be more noisy than T_1w (Reich et al., 2015). Prediction images demonstrating above-average, average, and below-average performance based on the RMSE measure are presented in Figs. 1 and 2.

3.2. Rater study

3.2.1. Consistency between raters

Of the 100 lesions included in the rater study, 4 were excluded for segmentation errors. Six ratings were obtained for each image, resulting in 576 distinct scores for each rater. In 17% (n = 98) of these cases, all three raters assigned the same score. In another 54% (n = 309), two raters assigned the same score and the third gave a score either one unit above or one unit below the score given by the other two. Thus, in 71% of the ratings, the raters all assigned scores within one unit of each other.

3.2.2. Ratings

Scores were averaged across the three raters to obtain mean ratings for each image. Using these mean ratings, both the nT_1w and MTR prediction models performed well. For both models, the median score for the overall similarity between the predicted image and the true lesion image was 3.0, corresponding to a rating of "Passed with minor errors." For the specific rating of similarity of degree of recovery between the predicted and observed images, the median for both nT_1w and MTR

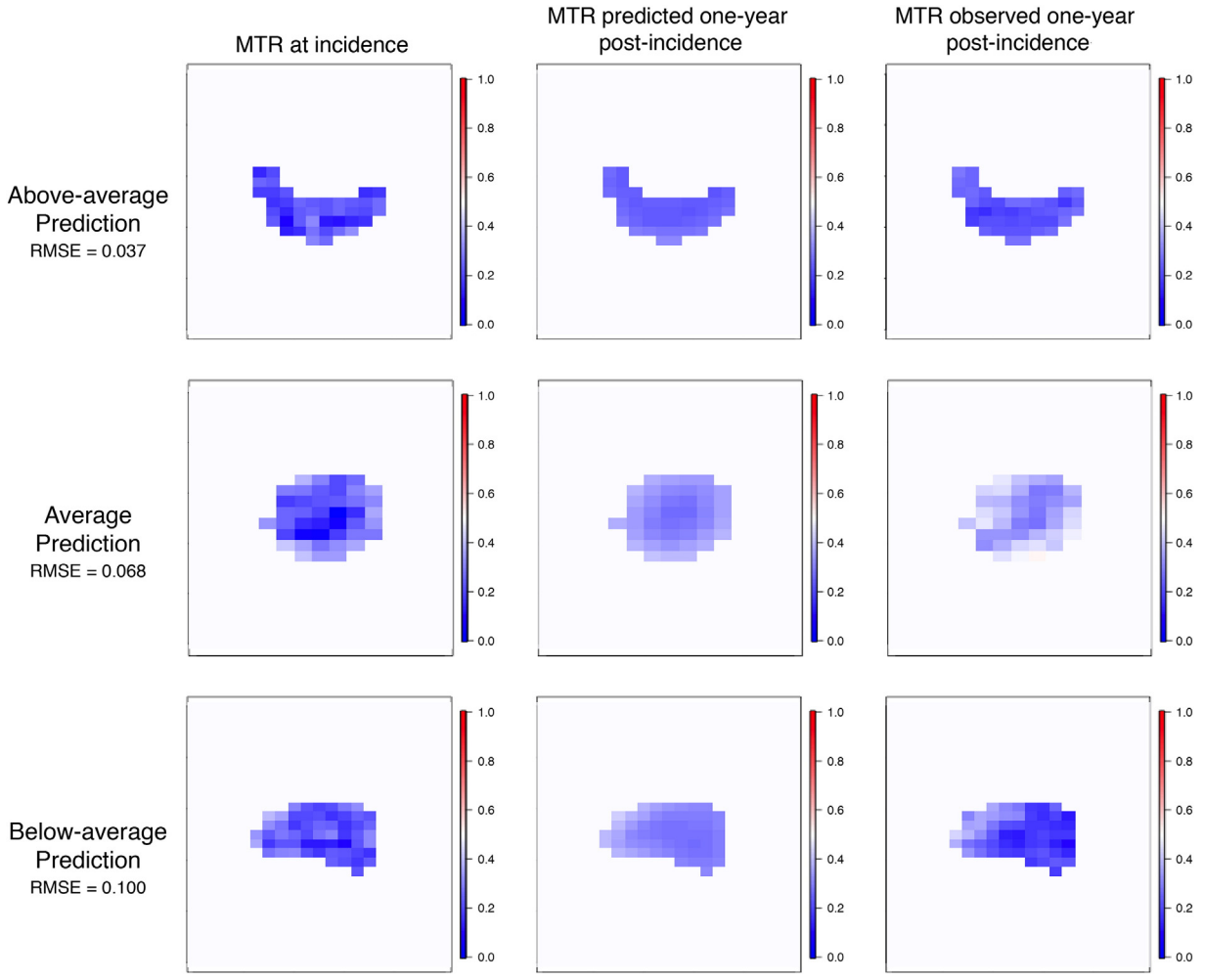

Fig. 2. Representative predictions for the MTR model. Images are axial slices of lesions, with rows representing three example lesions with varying levels of predictive accuracy. For MTR, intensities range from 0 to 1, with a mean of approximately 0.5 for normal-appearing white matter.

was 2.6. This score was partway between the rating "Some redeeming features" and the rating "Passed with minor errors." For the similarity of "pattern of recovery" between the predicted and observed images, the median for both nT_1w and MTR was 3.00 (Fig. 3). Overall ratings of accuracy were significantly ($p < 0.001$) negatively associated with a lesion's RMSE ($r_{T1} = -0.60$, $r_{MTR} = -0.56$). Additionally, ratings of degree were more associated with RMSE ($r_{T1} = -0.71$, $r_{MTR} = -0.68$) than ratings of pattern ($r_{T1} = -0.52$, $r_{MTR} = -0.53$), but all associations were statistically significant ($p < 0.001$).

4. Discussion

In this paper, we developed models to estimate the appearance of white matter MS lesions one year after their incidence. Both the model for normalized T_1w voxel intensity and the model for MTR voxel intensity produced accurate, cross-validated predictions of intensity at one-year post-incidence. This accuracy was measured quantitatively using RMSE for estimating the true one-year intensity with the statistically predicted one-year intensity. RMSE for the predicted intensity was compared to the measurement error of voxel intensity, operationalized as the RMSE for estimating intensity at a one-year post-incidence visit with the intensity at the visit immediately preceding it. In the model for nT_1w intensity, the prediction accuracy was approximately one-half standard deviation unit larger than the value of measurement error, indicating that the nT_1w model was able to predict one-year intensities almost as accurately as these can be assessed directly by re-acquiring the image. For the model of MTR intensity, the prediction accuracy surpassed that of measurement error. The accuracy of the model was also confirmed by three board-certified MS clinicians, who rated the correspondence between images produced by the model and images of true one-year intensities.

During the development of the models, flexible spline modeling was explored to account for possible nonlinear relationships. This method did not improve the predictions, and thus the simpler linear regression

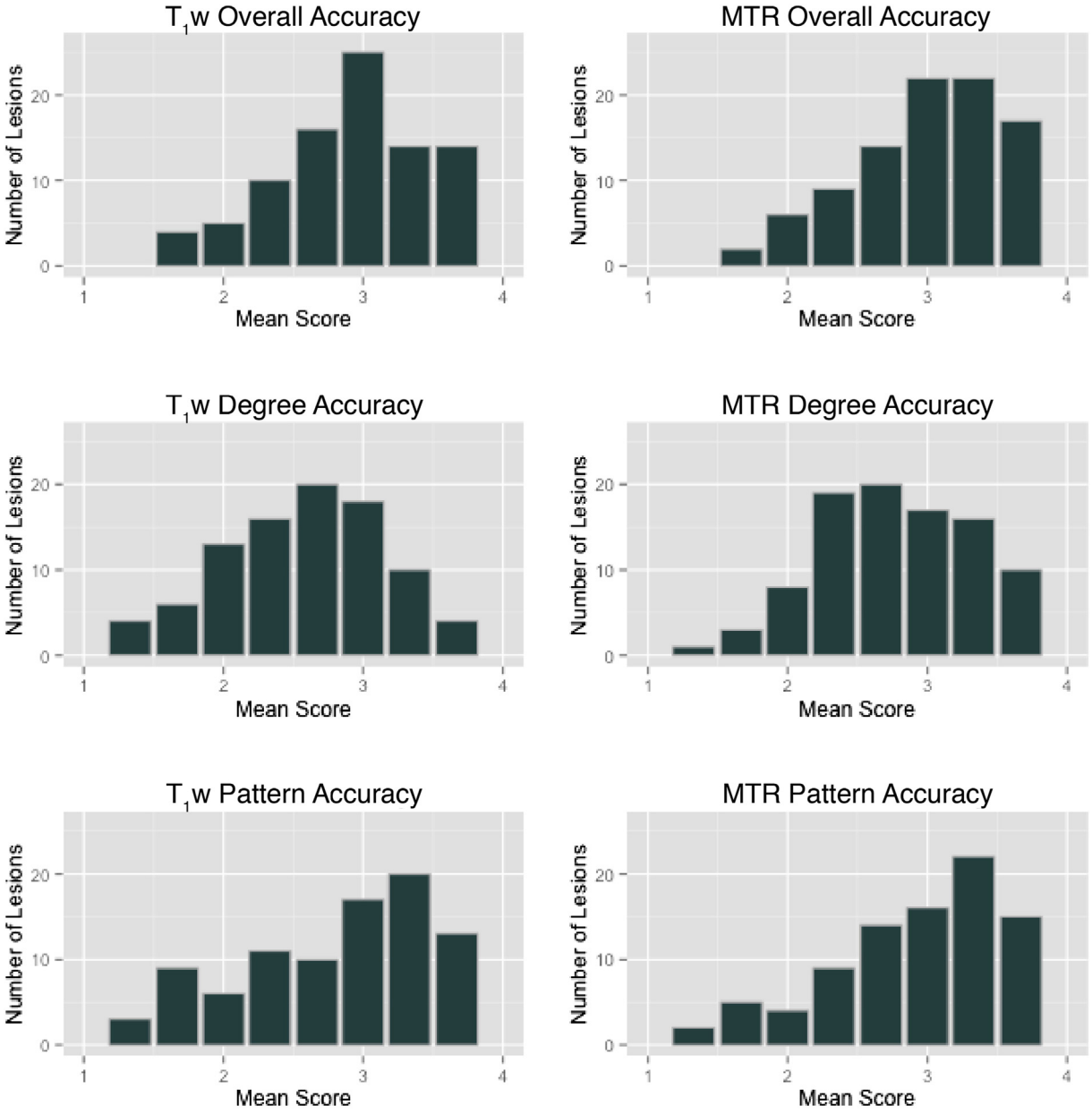

Fig. 3. Scores for the six rater study questions, averaged across the three raters. The three rows show the distributions of the ratings of overall accuracy, accuracy of the degree of healing, and accuracy of the pattern of healing, respectively. Plots in the first column are distributions of ratings of the nT_1w prediction images, and plots in the second column are distributions of ratings of the MTR prediction images.

approach was used for the final models. It is also worth noting that voxels containing vasogenic edema, as described previously (Sweeney et al., 2016), were included in the data for both model development and accuracy testing. Though edema behaves differently than structurally damaged lesion tissue, and its healing pattern is potentially easier to predict, it was determined that for the purpose of these predictions it was appropriate to leave edema voxels in the data. This is largely due to the fact that in many cases, differentiating between edema and lesion is difficult without examining how the tissue changes over the course of several weeks (Sweeney et al., 2016). Since our model was developed for the purpose of predicting healing using only information obtained at the time of lesion incidence, using data from later visits to categorize and remove edema would not accurately represent the way the model would be used clinically. However, voxels were categorized as edema or lesion after the models had been developed for performance assessment, and we found that the prediction accuracy was comparable between edema voxels and lesion voxels for both nT_1w ($RMSE_{edema} = 0.80$, $RMSE_{lesion} = 1.04$) and MTR models ($RMSE_{edema} = 0.061$, $RMSE_{lesion} = 0.066$).

There were some limitations to the predictions developed in this study. Spatial covariance was not explicitly modeled, though spatial relationships were loosely accounted for by including the voxel intensities of spatially smoothed images. Additionally, the measure of voxel location used (distance to the boundary) likely could be improved upon as well, as it is unable to account for differing lesion shapes. However, in spite of the relative simplicity of the spatial and locational variables, the models were still able to achieve good accuracy in these areas. Both T_1w and MTR predictions had a median rating of "Passed with minor errors" with respect to the pattern of healing.

An additional limitation arises in the interpretation of regression coefficients. For both T_1w and MTR models, treatment with steroids was associated with lower voxel intensity at the one-year follow-up, indicating worse healing in patients on steroids. This is most likely an effect of the observational nature of this study, which suggests that the steroids variable is capturing an aspect of disease severity or activity that was not accounted for completely by EDSS. Including other characteristics in the model could take steps to address this phenomenon, but ideally future work would test this model in the context of clinical trial data in order to obtain a clearer sense of how treatment impacts the recovery of incident lesions.

A strength of the current study is the novel integration of neuroimaging with demographic and clinical data to predict how lesion tissue will heal in MS patients. We believe that accurate predictions of this sort may have several important applications in MS treatment and research. In MS research, clinical trials may benefit from the ability to predict which patients are more likely to be responsive to treatment. Clinically, this tool could enable physicians to view a prediction of the likely course of healing of patients' incident lesions in order to make more informed and personalized treatment decisions. As such, future research will focus on the refinement of the treatment and steroid effect estimation in the model through the use of clinical trial data, with the goal of facilitating the direct comparison of predicted recovery when treated with various disease modifying therapies, steroids, or when left untreated. This would provide clinicians and researchers with previously unavailable information about the course a lesion is likely to take, and would allow for greater personalization of treatment decisions, as well as better informed and more powerful study designs.

Acknowledgments

The authors would like to thank Blake Dewey, the Neuroimmunology Branch clinical group and the technicians at the NIH who were instrumental in helping to collect and process the study data. The project described is supported in part by the NIH grants RO1 NS085211 and R21 NS093349 from the National Institute of Neurological Disorders and Stroke (NINDS). The research is also supported by the Intramural Research Program of NINDS. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

References

Barkhof, F., 2002. The clinico-radiological paradox in multiple sclerosis revisited. Curr. Opin. Neurol. 15 (3), 239–245.
Bramow, S., Frischer, J.M., Lassmann, H., Koch-Henriksen, N., Lucchinetti, C.F., Sørensen, P.S., Laursen, H., Oct. 2010. Demyelination versus remyelination in progressive multiple sclerosis. Brain J. Neurol. 133 (10), 2983–2998.
Carass, A., Wheeler, M.B., Cuzzocreo, J., Bazin, P.-L., Bassett, S.S., Prince, J.L., 2007. A joint registration and segmentation approach to skull stripping. Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium on, pp. 656–659.
Filippi, M., Cercignani, M., Inglese, M., Horsfield, M.A., Comi, G., Feb. 2001. Diffusion tensor magnetic resonance imaging in multiple sclerosis. Neurology 56 (3), 304–311.
Fonov, V., Evans, A.C., Botteron, K., Almli, C.R., McKinstry, R.C., Collins, D.L., Brain Development Cooperative Group, Jan. 2011. Unbiased average age-appropriate atlases for pediatric studies. NeuroImage 54 (1), 313–327.
Frischer, J.M., Weigand, S.D., Guo, Y., Kale, N., Parisi, J.E., Pirko, I., Mandrekar, J., Bramow, S., Metz, I., Brück, W., Lassmann, H., Lucchinetti, C.F., Nov. 2015. Clinical and pathological insights into the dynamic nature of the white matter multiple sclerosis plaque: dynamic nature of MS plaque. Ann. Neurol. 78 (5), 710–721.
Kurtzke, J.F., Nov. 1983. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 33 (11), 1444–1452.
Larsson, H.B., Frederiksen, J., Petersen, J., Nordenbo, A., Zeeberg, I., Henriksen, O., Olesen, J., Sep. 1989. Assessment of demyelination, edema, and gliosis by in vivo determination of T1 and T2 in the brain of patients with acute attack of multiple sclerosis. Magn. Reson. Med. 11 (3), 337–348.
Lassmann, H., Oct. 2013. Pathology and disease mechanisms in different stages of multiple sclerosis. J. Neurol. Sci. 333 (1–2), 1–4.
Lassmann, H., Brück, W., Lucchinetti, C.F., Apr. 2007. The immunopathology of multiple sclerosis: an overview. Brain Pathol. 17 (2), 210–218.
Lucas, B.C., Bogovic, J.A., Carass, A., Bazin, P.-L., Prince, J.L., Pham, D.L., Landman, B.A., Mar. 2010. The Java Image Science Toolkit (JIST) for rapid prototyping and publishing of neuroimaging software. Neuroinformatics 8 (1), 5–17.
Meier, D., Guttmann, C., Oct. 2003. Time-series analysis of MRI intensity patterns in multiple sclerosis. NeuroImage 20 (2), 1193–1209.
Meier, D.S., Guttmann, C.R.G., Aug. 2006. MRI time series modeling of MS lesion development. NeuroImage 32 (2), 531–537.
Meier, D.S., Weiner, H.L., Guttmann, C.R.G., Jul. 2007. Time-series modeling of multiple sclerosis disease activity: a promising window on disease progression and repair potential? Neurother. J. Am. Soc. Exp. Neurother. 4 (3), 485–498.
Mejia, A., Sweeney, E.M., Dewey, B., Nair, G., Sati, P., Shea, C., Reich, D.S., Shinohara, R.T., Dec. 2015. Statistical estimation of T1 relaxation times using conventional magnetic resonance imaging. NeuroImage.
Narayanan, S., Fu, L., Pioro, E., De Stefano, N., Collins, D.L., Francis, G.S., Antel, J.P., Matthews, P.M., Arnold, D.L., Mar. 1997. Imaging of axonal damage in multiple sclerosis: spatial distribution of magnetic resonance imaging lesions. Ann. Neurol. 41 (3), 385–391.
Patrikios, P., Stadelmann, C., Kutzelnigg, A., Rauschka, H., Schmidbauer, M., Laursen, H., Sorensen, P.S., Brück, W., Lucchinetti, C., Lassmann, H., Dec. 2006. Remyelination is extensive in a subset of multiple sclerosis patients. Brain J. Neurol. 129 (Pt 12), 3165–3172.
Pike, G.B., De Stefano, N., Narayanan, S., Worsley, K.J., Pelletier, D., Francis, G.S., Antel, J.P., Arnold, D.L., Jun. 2000. Multiple sclerosis: magnetization transfer MR imaging of white matter before lesion appearance on T2-weighted images. Radiology 215 (3), 824–830.
Polman, C.H., Reingold, S.C., Banwell, B., Clanet, M., Cohen, J.A., Filippi, M., Fujihara, K., Havrdova, E., Hutchinson, M., Kappos, L., Lublin, F.D., Montalban, X., O'Connor, P., Sandberg-Wollheim, M., Thompson, A.J., Waubant, E., Weinshenker, B., Wolinsky, J.S., Feb. 2011. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Ann. Neurol. 69 (2), 292–302.
R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Radü, E.-W., Sahraian, M.A., 2008. MRI Atlas of MS Lesions. Springer Berlin Heidelberg, Berlin, Heidelberg.
Reich, D.S., White, R., Cortese, I.C., Vuolo, L., Shea, C.D., Collins, T.L., Petkau, J., Feb. 2015. Sample-size calculations for short-term proof-of-concept studies of tissue protection and repair in multiple sclerosis lesions via conventional clinical imaging. Mult. Scler. Houndmills Basingstoke Engl.
Shiee, N., Bazin, P.-L., Ozturk, A., Reich, D.S., Calabresi, P.A., Pham, D.L., Jan. 2010. A topology-preserving approach to the segmentation of brain images with multiple sclerosis lesions. NeuroImage 49 (2), 1524–1535.
Shinohara, R.T., Crainiceanu, C.M., Caffo, B.S., Gaitán, M.I., Reich, D.S., Aug. 2011. Population-wide principal component-based quantification of blood-brain-barrier dynamics in multiple sclerosis. NeuroImage 57 (4), 1430–1446.
Shinohara, R.T., Sweeney, E.M., Goldsmith, J., Shiee, N., Mateen, F.J., Calabresi, P.A., Jarso, S., Pham, D.L., Reich, D.S., Crainiceanu, C.M., 2014. Statistical normalization techniques for magnetic resonance imaging. NeuroImage Clin. 6, 9–19.
Sweeney, E.M., Shinohara, R.T., Shiee, N., Mateen, F.J., Chudgar, A.A., Cuzzocreo, J.L., Calabresi, P.A., Pham, D.L., Reich, D.S., Crainiceanu, C.M., 2013. OASIS is automated statistical inference for segmentation, with applications to multiple sclerosis lesion segmentation in MRI. NeuroImage Clin. 2, 402–413.

Sweeney, E.M., Shinohara, R.T., Dewey, B.E., Schindler, M.K., Muschelli, J., Reich, D.S., Crainiceanu, C.M., Eloyan, A., 2016. Relating multi-sequence longitudinal intensity profiles and clinical covariates in incident multiple sclerosis lesions. NeuroImage Clin. 10, 1–17.

van Waesberghe, J.H., van Walderveen, M.A., Castelijns, J.A., Scheltens, P., Nijeholt, G.J.L.à., Polman, C.H., Barkhof, F., Apr. 1998. Patterns of lesion development in multiple sclerosis: longitudinal observations with T1-weighted spin-echo and magnetization transfer MR. AJNR Am. J. Neuroradiol. 19 (4), 675–683.

van Waesberghe, J.H., Kamphorst, W., De Groot, C.J., van Walderveen, M.A., Castelijns, J.A., Ravid, R., Nijeholt, G.J.L.à., van der Valk, P., Polman, C.H., Thompson, A.J., Barkhof, F., Nov. 1999. Axonal loss in multiple sclerosis lesions: magnetic resonance imaging insights into substrates of disability. Ann. Neurol. 46 (5), 747–754.

Vrenken, H., Geurts, J.J.G., Knol, D.L., van Dijk, L.N., Dattola, V., Jasperse, B., van Schijndel, R.A., Polman, C.H., Castelijns, J.A., Barkhof, F., Pouwels, P.J.W., Sep. 2006. Whole-brain T1 mapping in multiple sclerosis: global changes of normal-appearing gray and white matter. Radiology 240 (3), 811–820.

Werring, D.J., Clark, C.A., Barker, G.J., Thompson, A.J., Miller, D.H., May 1999. Diffusion tensor imaging of lesions and normal-appearing white matter in multiple sclerosis. Neurology 52 (8), 1626–1632.