

Disagreement Among Neuroradiologists Detecting Change in Multiple Sclerosis Disease Activity

L. Bajus¹, K. M. Leyden¹, M. Iv², E. Tong², N. Siebenborn³, C. Granziera³

¹Octave, Menlo Park, CA, USA, ²Stanford University, Palo Alto, CA, USA, ³University Hospital Basel, Basel, Switzerland

P176



Introduction

Multiple Sclerosis (MS) Neurologists' clinical decisions rely on accurate assessment of whether a patient is demonstrating change in MS disease activity (DA) over time. Variability of Neuroradiologist (NRs) impression of change on magnetic resonance imaging (MRI) can significantly impact clinical decisions.

Objectives

Evaluate the consistency of NRs impression of a patient's DA status within standard of care structured brain MRI reports.

Methods

- 90 de-identified MS subjects with 2 (median 1 year apart) MRI exams from 2012 to 2019 were retrospectively enrolled from the University Hospital Basel. Enrolled subjects were randomly selected to represent varying levels of disease activity.
- Each MRI exam consisted of 3D pre- and post-gadolinium T1 and 3D T2 FLAIR sequences.
- Three fellowship-trained NRs (NR 1, NR 2, and NR 3) with up to 9 years of post-fellowship clinical experience visually interpreted the images for each subject and assessed report metrics shown in Table 2. The NRs were instructed to provide their interpretation in a standardized structured report template.
- Descriptive statistics were reported as frequency and percentage for categorical variables.
- Inter-rater reliability analysis with Fleiss' Kappa statistics was performed to evaluate the degree of consistency between NRs in their assessment of specific metrics. Percentage agreement in each metric were also reported.
- The Fleiss' Kappa value was interpreted as:
 - ≤ 0 = no agreement
 - 0.01–0.20 = none-slight
 - 0.21–0.40 = fair
 - 0.41–0.60 = moderate
 - 0.61–0.80 = substantial
 - 0.81–1.00 = almost perfect agreement

Results

- Of the 90 MS subjects, 64.6% were female and the mean age was 51.7 ± 10.4 years old, with an average EDSS score of 3.3 ± 1.7 .
- Table 1 describes the distribution of disease activity metrics in all 270 MRI reports, then by the number and percentage of 90 individual subjects.
- Table 2 describes the inter-rater reliability analysis of disease activity metrics between the 3 NRs.
- Change in Disease Activity**
 - NRs reported evidence of change in 81 of 270 MRI reports (30%)
 - NRs had 100% agreement for 47 patients (10 change, 37 stable).
 - 53 patients were noted as demonstrating a change by at least 1 NR, while 80 patients were noted as stable by at least 1 NR.
 - Inter-rater analysis of change status showed fair agreement (FK= 0.24, 52% agreement).
- Presence of Enhancing Lesions**
 - 3 patients (3%) had discrepant ratings for the presence of new lesions.
 - Inter-rater analysis of the presence of new lesions showed substantial agreement among NRs (FK =0.62, 95% agreement).
- Presence of New Lesions**
 - 39 patients (53%) had discrepant ratings for the presence of new lesions.
 - 83 patients had at least 1 NR report no new lesions, while 46 patients had at least 1 NR report new lesions.
 - Inter-rater analysis of the presence of new lesions showed fair agreement among NRs (FK =0.25, 68% agreement).
- Presence of Enlarging lesions**
 - 27 of 90 patients had discrepancy for the presence of enlarging lesions (%).
 - 89 patients had at least 1 NR report no enlarging lesions, while 28 patients had at least 1 NR note the presence of enlarging lesions.
 - Inter-rater analysis of the presence of enlarging lesions showed none to slight agreement (FK=0.03, 50% agreement).

Table 1. Neuroradiologist responses across individual reports (N=270) and for each Subject (N=90).

	Disease Activity (N, %)	Enhancing Lesions (N, %)	New Lesions (N, %)	Enlarging Lesions (N, %)
# of Reports with 'Yes'	81 (30%)	10 (3%)	78 (29%)	31 (11%)
# of Subjects with ≥ 1 NR Responding 'Yes'	53 (59%)	5 (6%)	46 (51%)	28 (31%)
# of Subjects with ≥ 1 NR Responding 'No'	80 (89%)	88 (98%)	83 (92%)	89 (99%)
# of Subjects with 100% NR Agreement	47 (52%)	87 (97%)	51 (57%)	63 (70%)

Table 2. Results of inter-rater reliability analysis across report metrics

	Disease Activity	Enhancing Lesions	New Lesions	Enlarging Lesions
Percentage of Agreement	52%	96%	64%	46%
Fleiss Kappa Coefficient	0.24	0.62	0.31	0.04
Coefficient Interpretation	Fair	Substantial	Fair	None-slight

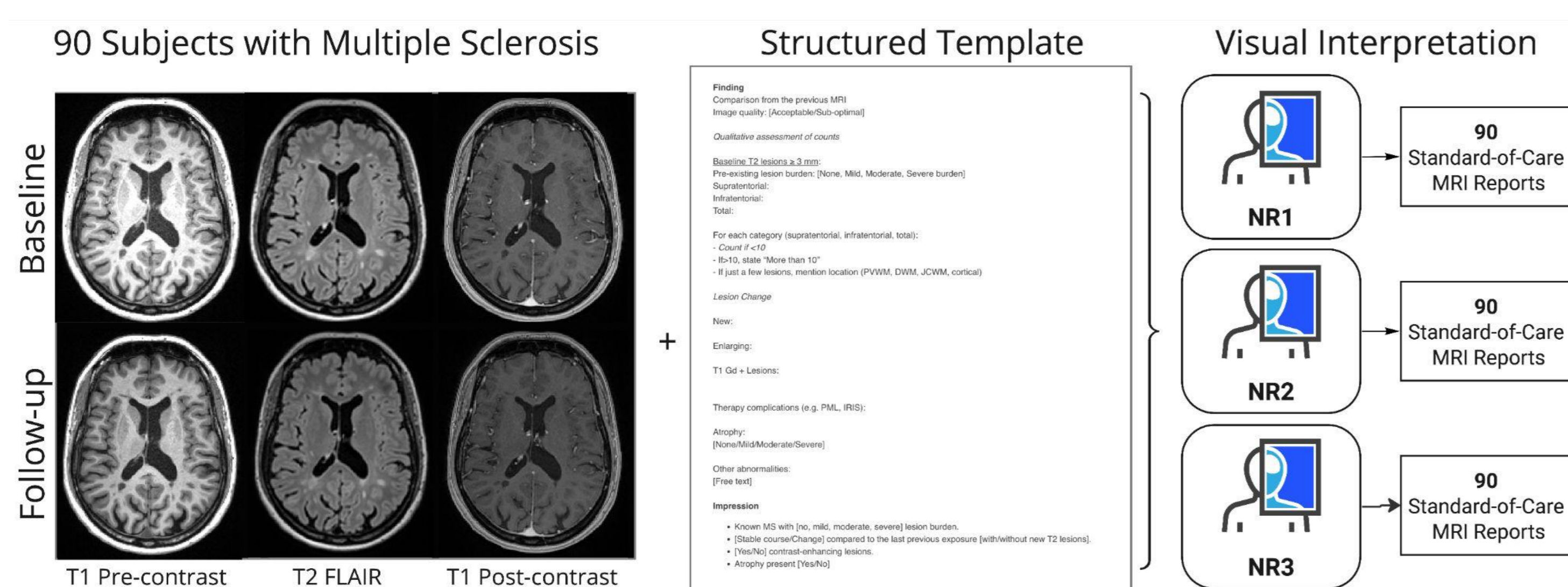


Figure 1: Study design workflow with example MRI slices and structured report template.

Conclusions

The results demonstrate significant variability across Neuroradiologist's impressions of whether the subjects demonstrated evidence of disease activity. While Neuroradiologists were most consistent when reporting the presence of enhancing lesions, all other metrics of disease activity such as new and enlarging lesions varied from none-slight to fair agreement.

Accurate measures of longitudinal brain changes & new and enlarging lesion counts within an MRI report are critical for MS Neurologists who are responsible for clinical treatment change decision making- directly impacting the decision to change treatment, schedule follow-up MRIs or order tests. The demonstrated variability in impressions of change across NRs may significantly impact patient outcomes and their cost of care.