Discrepancies in the Detection Of Salient MS Lesion Counts Between Neuroradiologists and Manually Corrected Automated Lesion Segmentation

K. M. Leyden[1], F. X. Faust[1], A. Keshavan[1]
[1]Octave, Menlo Park, CA, USA

Introduction

Accurate detection and monitoring of lesions in patients with multiple sclerosis (MS) is vital for managing treatment. In particular, new and enlarging lesions are key factors for evaluating changes in treatment plans for patients with MS (pwMS). Traditional visual interpretation by neuroradiologists (NR) is subjective, leading to variability in the clinical impression of disease activity in MRI reports. Automated lesion detection algorithms are a potential tool to assist NR by providing quantitative assessments, yet their impact on clinical decision making remains underexplored.

Objectives

To quantify the discrepancy between lesion counts detected by NR to lesion counts derived from an FDA-cleared segmentation algorithm that has been manually corrected with a semi-automated lesion editing software tool.

Methods

- 3 NR were presented with 2 time points of 3D T1, 3D T2 FLAIR, and 3D T1 post-contrast scans from 90 pwMS to generate a standardized structured report based on visual interpretation.
- Structured reports included fields for the NR to count new and enlarging lesions.
  - If the pwMS had <10 lesions in the supratentorial or infratentorial compartments, the NR were asked to record the exact number of lesions. (Figure 1)
- To generate the initial lesion segmentation masks, MRI scans were processed using a commercially available, FDA-cleared software for automated lesion segmentation. (Figure 1)
- MRI QC analysts manually edited the segmentation masks to remove false positive & false negative lesions.
- Counts were generated for total, infratentorial, new & enlarging lesions to match the Neuroradiologist categories.
- Lesion counts for each category were compared between the NR reports and the manually corrected algorithm (MCA) output.
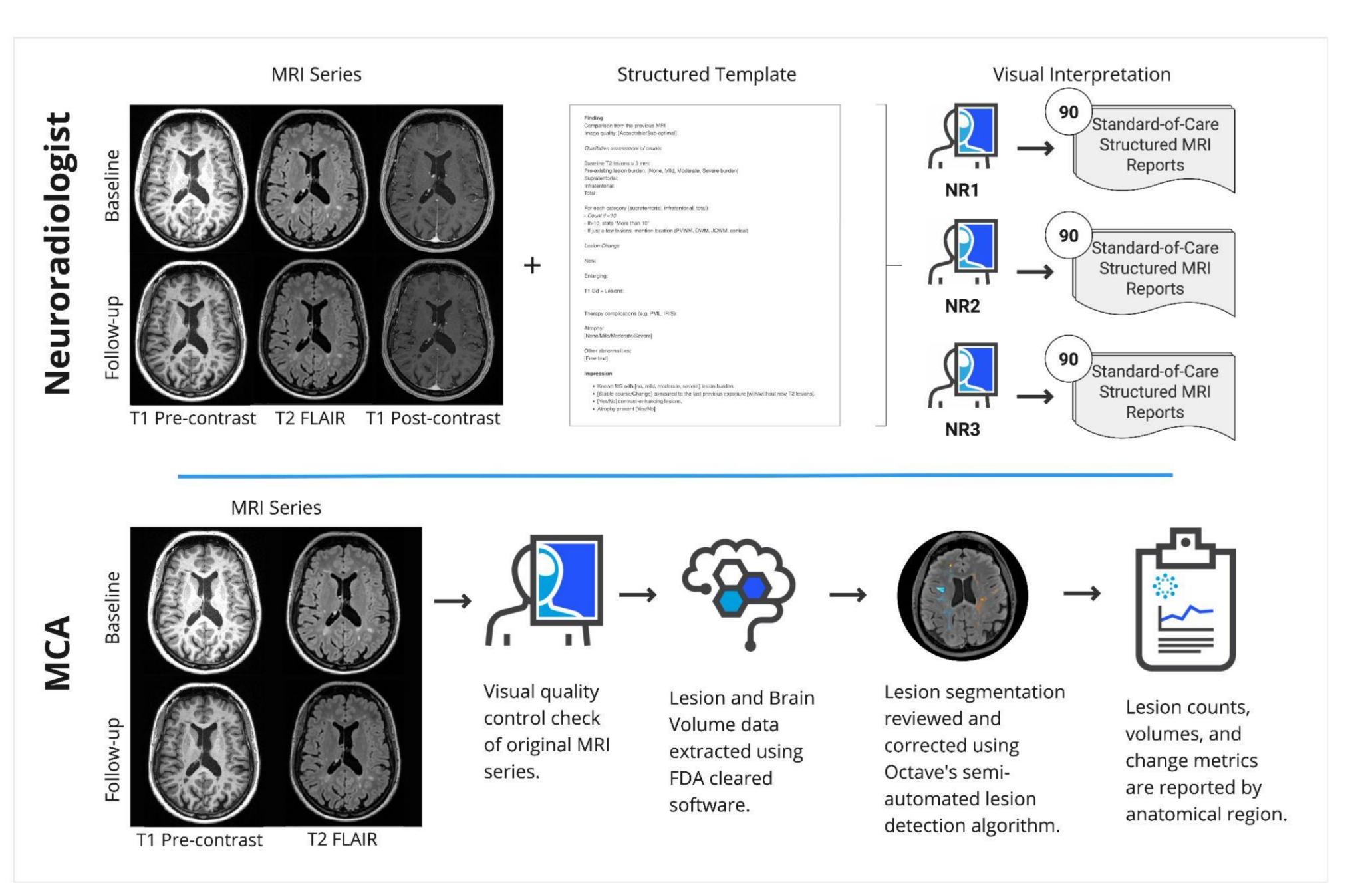

Figure 1: Diagram of study design. The top row depicts the workflow used to generate standard of care MRI reports, visually interpreted and reported by board certified Neuroradiologists. The bottom row depicts the workflow used to process the MRIs to segment the lesions and generate the results of the MCA output.

Results

- 3 NR completed 90 standard of care structured reports for a total of 270 reports (3 per pwMS).
- MRI scans from 86 pwMS successfully passed image quality checks and volumetric segmentation to generate lesion counts, resulting in 258 MCA results that had a corresponding NR report.
  - 4 pwMS had MRI scans that were incompatible with the automated lesion segmentation due to poor image quality. This resulted in 12 reports not having a MCA comparison.
- Overall, NR detected new lesions in 56 reports, and enlarging lesions in 48 reports. The MCA detected new lesions in 30 reports and enlarging lesions in 69 reports.
- NR and MCA agreed on the presence of new lesions in 84.5% of reports (N=218) and agreed on the presence of enlarging lesions in 68.6% of the reports (N=177). NR and MCA had equal counts for new lesions in 205 reports, disagreeing in 53 reports. NR and MCA had equal counts for enlarging lesions in 181 reports, disagreeing in 77 reports. (Table 1)
  - Within the 23 instances where both NR and MCA agreed on the presence of new lesions (Table 2), 13 reports across 6 pwMS had discrepant counts of the new lesions.
  - Within the 18 instances where both NR and MCA agreed on the presence of enlarging lesions (Table 3), 7 reports across 9 pwMS had discrepant counts of the enlarging lesions.

Rates of Agreement Between NR and MCA Reports (N,%)

	Presence of New Lesions	Presence of Enlarging Lesions	New Lesion Count	Enlarging Lesion Count	Infratentorial Lesion Count	Total Lesion Count
Agree	218 (84.50%)	177 (68.60%)	205 (79.46%)	181 (70.16%)	80 (31.01%)	4 (1.55%)
Disagree	40 (15.50%)	81 (31.40%)	53 (20.54%)	77 (29.84%)	178 (68.99%)	254 (98.45%)

Table 1: The number and percentage of reports (N=258) where Neuroradiologists (NR) and the manually corrected algorithm (MCA) results agree or disagree for the presence and counts of various lesion categories.

New Lesion(s) Detected		Manually Corrected Algorithm	
		Yes (N=30)	No (N=228)
Neuroradiologists	Yes (N=56)	23	33
	No (N=202)	7	195

Table 2: The breakdown of agreement (green) or disagreement (orange) between Neuroradiologists and the manually corrected algorithm for the presence or absence of new lesions in all reports.

Enlarging Lesion(s) Detected		Manually Corrected Algorithm	
		Yes (N=69)	No (N=189)
Neuroradiologists	Yes (N=48)	18	30
	No (N=210)	51	159

Table 3: The breakdown of agreement (green) or disagreement (orange) between Neuroradiologists and the manually corrected algorithm for the presence or absence of enlarging lesions in all reports.

Conclusions

Lesion detection is a complex and nuanced process, and the methods for characterizing lesions continues to evolve. In this analysis we observed good agreement between the NRs and MCA in detecting the presence of new lesions, however, when quantifying lesions there was a marked discrepancy. The evidence of substantial variations between NRs and MCA counts in new and enlarging lesions highlight how the addition of quantitative lesion detection into clinical workflows may impact standard of care in MS.