REAL-WORLD EVALUATION OF A COMMERCIAL MACHINE LEARNING-BASED TOOL FOR THE LONGITUDINAL ANALYSIS OF MULTIPLE SCLEROSIS PLAQUES ON BRAIN MRI

S Nesar, J L Chen, M A Esmeraldo, F Zaccagna, T Das

Cambridge University Hospitals NHS Foundation Trust, United Kingdom

BACKGROUND

There has been rapid development of artificial intelligence (AI) tools within radiology, including those aiding detection of demyelinating lesions in multiple sclerosis (MS).[1]

It is critical that these tools are evaluated appropriately. However, there is currently a relative paucity of real-world clinical evaluation data.

The aim of this study was to evaluate the accuracy of a recently deployed MDR Class IIa tool (**Pixyl.Neuro.MS, Pixyl,** France) in a standard clinical setting.[2,3]

METHODS

3-month retrospective review of MRI brain scans that were performed for the longitudinal follow-up of patients with MS and were processed by Pixyl.Neuro.MS, with the results available only to the reporting radiologist.

We assessed performance of Pixyl.Neuro.MS at both the **lesion level** and **case level** in the detection, annotation, classification, and quantification of new demyelinating lesions on 3D T2 FLAIR sequences, with the presence of any new demyelinating lesions taken to indicate active disease.[4]

Where the Pixyl.Neuro.MS and consultant neuroradiologist reports were concordant, these were accepted as the ground truth. Where discordant, ground truth was established by researcher review (consisting of four radiologists, including a consultant neuroradiologist).

RESULTS

43 case-pairs were evaluated (patient age range 17–72, mean age 42; male : female ratio = 1 : 1.39; mean scan interval 16.5 months).

Pixyl.Neuro.MS identified a total of 2697 stable lesions across all patients and 75 new demyelinating lesions across 18 patients (41.9% of cohort). 46 (61.3%) of these were correctly identified; the remaining 29 (38.7%), across 10 patients (23.3% of cohort), represented misidentification of pre-existing lesions or other brain structures. In 3 patients (7.0%), this would have resulted in erroneous classification as active disease.

Pixyl.Neuro.MS failed to identify 1 new brainstem lesion in each of 2 patients (4.7%), which in 1 patient (2.3%) would have resulted in erroneous classification as inactive disease. However, Pixyl.Neuro.MS identified 1 new lesion in a patient, which was not reported by the neuroradiologist.

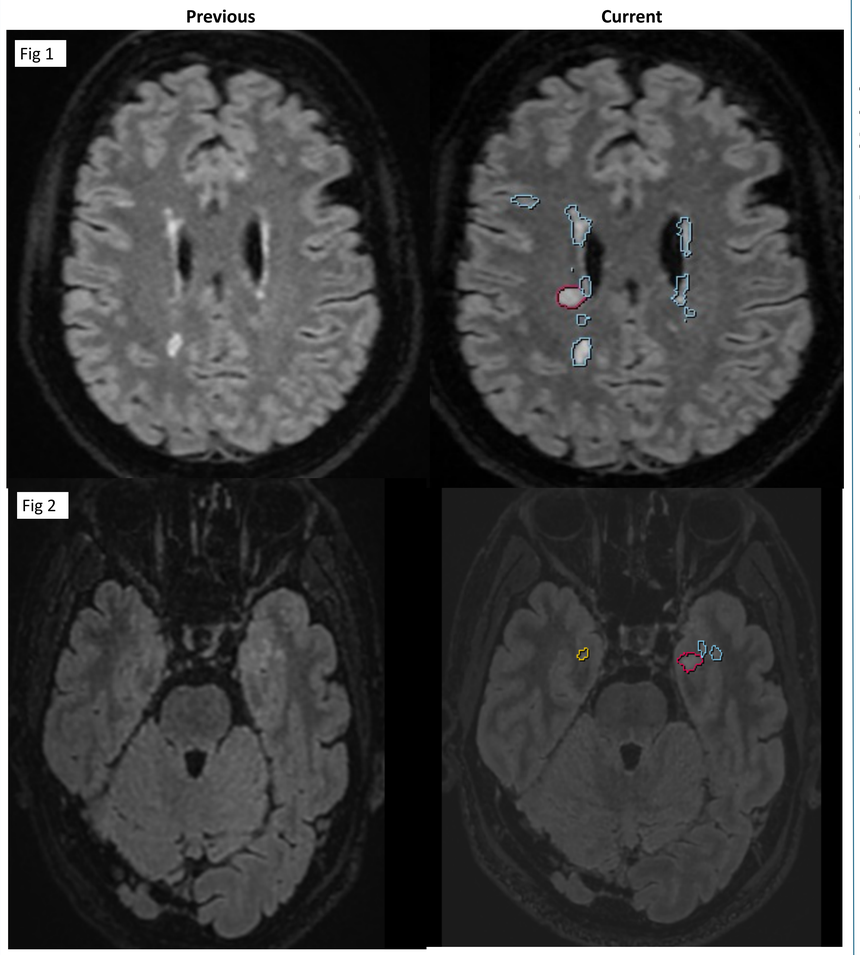
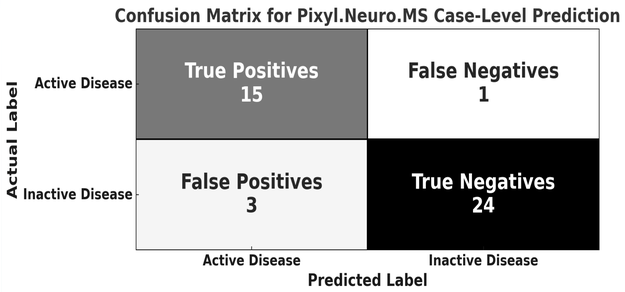


Figure 1: Pixyl.Neuro.MS correctly identifying a new lesion (outlined in red). Lesion types colour-coded: new lesions outlined in red, stable lesions in blue, and enlarging lesions in yellow.
Figure 2: Pixyl.Neuro.MS incorrectly identifying a new lesion.



Case-Level Metrics	Pixyl.Neuro.MS Results
Sensitivity	93.4%
Specificity	88.9%
Positive Predictive Value	83.3%
Negative Predictive Value	96.0%
Accuracy	90.7%

CONCLUSION

Pixyl.Neuro.MS is a **promising tool** for the longitudinal analysis of demyelinating lesions in multiple sclerosis, particularly **identification of new lesions**, signifying active disease, or confirming stable disease with no new lesions.

It demonstrated a **low false positive rate at the lesion level overall**, but a **high false positive rate when considering only new lesions**. It had **good performance at the case level** (90.7% accuracy) and could identify new lesions not identified by a neuroradiologist.

However, it did occasionally produce false positive and false negative results at the case level, and careful neuroradiologist scrutiny is still necessary.

Finally, Pixyl.Neuro.MS has significant potential to improve **reporting efficiency**, which we aim to evaluate further.

REFERENCES

1. Mendelsohn, Z. *et al.* Commercial volumetric MRI reporting tools in multiple sclerosis: a systematic review of the evidence. 2023. Neuroradiology, 65(1), 5-24.
2. Roca, P. *et al.* Artificial intelligence to predict clinical disability in patients with multiple sclerosis using FLAIR MRI. 2020. Diagnostic and Interventional Imaging, 101(12), 795-802.
3. Pixyl.Neuro.MS. https://pixyl.ai/multiple-sclerosis/
4. Giovannoni, G, *et al.* "Is it time to target no evident disease activity (NEDA) in multiple sclerosis?." Multiple Sclerosis and Related Disorders 4.4 (2015): 329-333.

Acknowledgments
- NHS Fellowship in Clinical AI
- NHS Health Education England – East of England.