

Implementation of an AI model to triage paediatric brain magnetic resonance imaging orders

Dear Editor,

Artificial intelligence (AI) is viewed as the most important recent advancement in radiology with the potential to achieve Singapore's objective of delivering value-based patient-centric care.¹

We have developed and implemented a deep-learning model using bidirectional long short-term memory (Bi-LSTM) neural network to enable automated triage of unstructured free-text paediatric magnetic resonance imaging (MRI) brain orders in conformance to the American College of Radiology (ACR) criteria² for appropriate utilisation of MRI. These ACR guidelines assist clinicians in the appropriate triaging of brain MRI orders for routine imaging, versus ultrafast MRI screening protocols for less appropriate orders.

After approval of waiver of consent from the Institution Review Board (CIRB reference number 2017/2078), data comprising 5,181 retrospective paediatric MRI brain orders (online Supplementary Table S1) extracted from 2006 to 2017 (excluding those with additional scans of other body parts and follow-up scans) were manually labelled for conformance to the ACR guidelines² under supervision of a senior paediatric radiologist. These were used as ground truth to develop a Bi-LSTM and other machine learning models to classify these free-text orders based on adherence to the ACR guidelines. Initially 2,470 orders from 2006 to 2013 were used for model training (80–20 training and validation split), and 2,711 orders from 2014 to 2017 for model testing, using receiver operating characteristics to measure model performance (online Supplementary Table S2). Another 50 orders from a 2020 audit were used for simulated implementation of the best performing model predicting MRI orders conforming to ACR guidelines,² comparing its performance against radiology staff with variable experience (including the aforesaid senior paediatric radiologist as gold standard), using Cohen's kappa statistics (online Supplementary Table S3). The model graphic user interface (Fig. 1) and details of its creation and testing are attached in the online Supplementary Materials.

The highest accuracy and area under the curve (AUC) were seen with the Bi-LSTM model (Supplementary Table S2). This model, utilised by a non-medical staff

such as a research assistant, has a kappa of 0.67, which shows evidence of a significant improvement compared to the kappa of 0.42 achieved by junior residents ($P=0.01$). It is comparable to the kappa of 0.68 seen in residents with several years of neuroradiology experience although it remains less than the kappa of 0.72 attained by a junior paediatric radiologist and experienced MRI radiographers (Supplementary Table S3) ($P<0.01$). An advantage of the Bi-LSTM model is its ability to map similar medical terms together while factoring sentence structure and context through a word vector matrix. In contrast, the bag-of-words traditional machine learning model requires more pre-processing steps and training time, results in a sparse dataset where each unique word represents a feature, and lacks the consideration of the contextual information derived from the sequence of words.

Model performance is likely dependent on both the model structure, and dataset size and complexity. Larger datasets with wide variability of data reflective of the real-life environment in which the model would be deployed are best for model creation. Machine learning is an evolving field and the numbers that constitute adequate sample sizes for developing prediction models are unclear, as evident from publications with sample sizes ranging from hundreds to millions.³⁻⁷

Incorporating a local interpretable model to explain each prediction outcome via a graphic user interface, builds confidence among non-medical or junior medical staff when protocoling MRI brain requests. In turn this will reduce the burden of MRI protocoling, increase productivity, and allow senior staff to focus on more pressing clinico-radiological issues. In addition, more objective criteria for MRI protocoling will reduce miscommunication and enhance radiology workflow efficiency.

Should busy physicians provide scanty information in their MRI orders, this may result in the model generating a low score for guideline adherence and the patient is triaged for an inappropriate ultrafast MRI brain protocol. Nevertheless, the radiology workflow provides an inherent additional safety net for patients with significant abnormalities, such as a mass. Radiographers, upon finding a mass on the initial MRI sequence, would

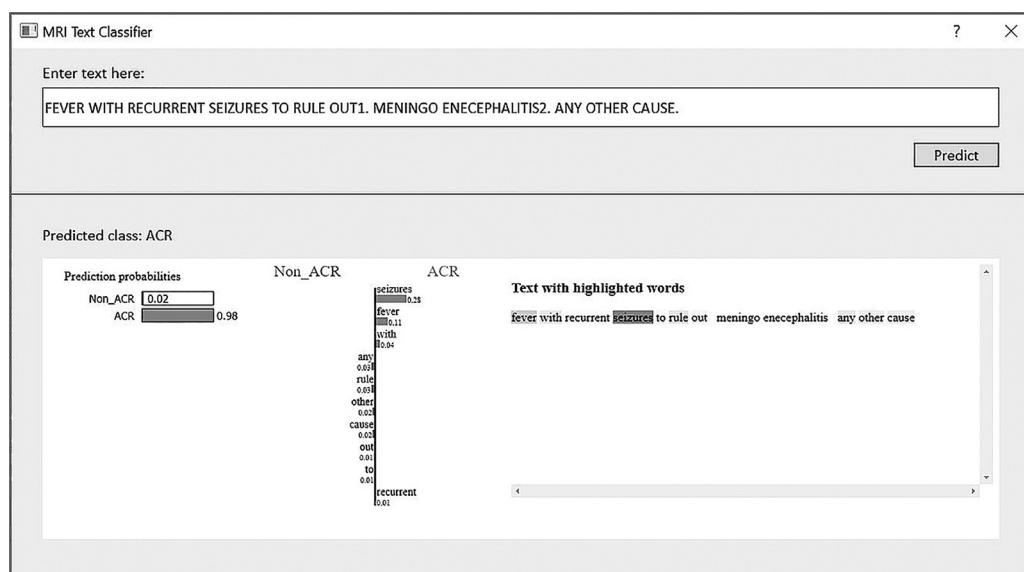


Fig. 1. User-friendly graphic user interface with case illustration of model explainer. Magnetic resonance imaging (MRI) indication on the electronic order form is copied and pasted into the textbox. Click of a button generates the percentage adherence to guidelines, with clear display of words that support the adherence to guidelines. An order with a generated percentage of less than 50% is deemed suitable for a screening ultrafast MRI brain protocol.

obtain radiologist clearance to convert the ultrafast protocol into a comprehensive MRI brain scan, including the addition of contrast-enhanced sequences.

Adherence to guidelines decreased across the years, with 76% adherence on the initial training dataset (2006–2013) dropping to 61% on the test dataset (2014–2017) and 40% adherence on the 2020 dataset. Paediatric MRI brain orders may deviate from evidence-based guidelines for a variety of reasons, such as uncertainties in the clinical assessment of an uncooperative child, pressure from anxious caregivers, and the potential for malpractice suits due to missed pathology. In practice, even though an MRI order may not comply with the radiological guidelines, a negative MRI scan does allay anxiety and promptly narrows management options for the patient.

Ultrafast or abbreviated protocols are increasingly proposed for suitable patients as an alternative to the full MRI protocol due to shorter acquisition time while maintaining diagnostic accuracy,⁸ ensuring effective utilisation of expensive MRI equipment. Screening protocols can be both clinically efficient and cost-effective when used for MRI brain requests with a low probability of having brain abnormalities in both children⁹ and adults.¹⁰

The primary limitation of our study is that the models were trained on labelled paediatric datasets from a single hospital and implemented within the same hospital.

External validity would be required to deploy the model beyond the confines of our hospital. Our model could also be further trained on an adult population, incorporating radiological terms from a different case mix, which would increase model robustness and generalisability. Future work could incorporate cross-institutional data and the cost-effectiveness analysis to better evaluate the efficacy of the model in improving radiology practices, and the correlation of outcomes of automated triaging of free-text MRI orders with actionable abnormality in the final MRI report.

In conclusion, radiology departments can consider leveraging on artificial intelligence to create user-friendly and explainable triage models. These can help the non-medical or junior staff determine if MRI orders are appropriate and facilitate workflow to help cope with an overwhelming demand for this expensive and limited imaging resource.

Funding

This study was funded by Health Services Research and Analytics Technologies for SingHealth grant (HEARTS 2017/024) and KKH NMRC Centre Grant Programme (NMRC/CG/M002/2017_KKH).

Acknowledgements

The authors would like to acknowledge the contributions and support of the initial members of the study team, especially Khine Nwe Win, Kathy Low, Edna Aw, Emily Chai, Tsang Wing Sze, Audrey Ling, Conceicao Edwin Philip and Amanda Choo for their help towards the curation of the datasets and the initial preliminary analysis of the study.

REFERENCES

1. Liew CJ, Krishnaswamy P, Cheng LT, et al. Artificial Intelligence and Radiology in Singapore: Championing a New Age of Augmented Imaging for Unsurpassed Patient Care. *Ann Acad Med Singap* 2019;48:16-24.
2. American College of Radiology. ACR–ASNR–SPR practice parameter for the performance and interpretation of magnetic resonance imaging (MRI) of the brain, 2013. <https://www.asnr.org/wp-content/uploads/2019/06/MR-Brain.pdf>. Accessed 14 April 2017.
3. Chen MC, Ball RL, Yang L, et al. Deep learning to classify radiology free-text reports. *Radiology* 2018;286:845-52.
4. Hassanpour S, Langlotz CP, Amrhein TJ, et al. Performance of a Machine Learning Classifier of Knee MRI Reports in Two Large Academic Radiology Practices: A Tool to Estimate Diagnostic Yield. *AJR Am J Roentgenol* 2017;208:750-3.
5. Huhdanpaa HT, Tan WK, Rundell SD, et al. Using Natural Language Processing of Free-Text Radiology Reports to Identify Type 1 Modic Endplate Changes. *J Digit Imaging* 2018;31:84-90.
6. Cheng LT, Zheng J, Savova GK, et al. Discerning tumor status from unstructured MRI reports--completeness of information in existing reports and utility of automated natural language processing. *J Digit Imaging* 2010;23:119-32.
7. Lakhani P, Kim W, Langlotz CP. Automated detection of critical results in radiology reports. *J Digit Imaging* 2012;25:30-6.
8. Ahamed SH, Lee KJ, Tang PH. Role of a modified ultrafast MRI brain protocol in clinical paediatric neuroimaging. *Clin Radiol* 2020;75:914-20.
9. Expert Panel on Pediatric Imaging: Hayes LL, Palasis S, et al. ACR Appropriateness Criteria® Headache-Child. *J Am Coll Radiol* 2018;15:S78-90.
10. Lim JLL, McAdory LE, Tang PH, et al. Appropriateness of MRI brain orders: Application of American and British guidelines. *J Neurol Sci* 2020;414:116874.

Phua Hwee Tang ^{1FRCR}, Alwin Yaoxian Zhang ^{2MSc},
 Sean Shao Wei Lam ^{2PhD}, Marcus Eng Hock Ong ^{3FRCS},
 Ling Ling Chan ^{4FRCS}

¹ Department of Diagnostic and Interventional Imaging, KK Women's and Children's Hospital, Singapore

² Health Service Research Centre, SingHealth, Singapore

³ Emergency Medicine, Singapore General Hospital, Singapore

⁴ Department of Diagnostic Imaging, Singapore General Hospital, Singapore

Correspondence: Dr Phua Hwee Tang, Department of Diagnostic and Interventional Imaging, KK Women's and Children's Hospital, 100 Bukit Timah Road, Singapore 229899.
 Email: tang.phua.hwee@singhealth.com.sg