Improving parkinsonism diagnosis with machine learning

Multiple systems atrophy and progressive supranuclear palsy are neurodegenerative disorders that often mimic Parkinson’s disease, which can lead to misdiagnosis.¹ Accurate diagnosis is crucial to patient care and categorisation of patients for clinical trials.² These issues highlight the need for a disease-specific biomarker to improve diagnostic accuracy.³ In The Lancet Digital Health, Archer and colleagues⁴ address this deficiency using non-invasive imaging and machine learning to distinguish Parkinson’s disease from atypical parkinsonism and multiple systems atrophy from progressive supranuclear palsy.

This study used free water and free-water-corrected fractional anisotropy (FA⁵) measures from diffusion-weighted MRI to develop disease-specific machine learning comparisons to differentiate forms of parkinsonism. The authors titled this automated pipeline the Automated Imaging Differentiation in Parkinsonism (AID-P) and tested models including diffusion-weighted MRI measurements from several brain regions, Movement Disorders Society Unified Parkinson’s Disease Rating Scale part III (MDS-UPDRS III), or both diffusion-weighted MRI and MDS-UPDRS III.

The diagnostic performance of a test can be assessed using the area under the curve (AUC), for which a high number indicates high sensitivity and specificity with the maximum value being 1.0. In this work, the AUC for AID-P models including diffusion-weighted MRI outperformed the MDS-UPDRS III only models in all comparisons. For differentiation of Parkinson’s disease from atypical parkinsonism, the diffusion-weighted MRI plus MDS-UPDRS III model showed a significantly higher (p=0.0001) AUC (0.962) relative to the MDS-UPDRS III only model (0.775), as did the diffusion-weighted MRI only model (0.955). Similarly, the diffusion-weighted MRI plus MDS-UPDRS III (AUC 0.897) and the diffusion-weighted MRI only models (AUC 0.926) significantly outperformed the MDS-UPDRS III only model (AUC 0.582; p=0.0001) in distinguishing multiple systems atrophy from progressive supranuclear palsy.

This work extends the findings from previous machine learning neuroimaging studies⁶ with the inclusion of the largest cohort of patients to date (n=1002) collected across 17 international sites using a variety of MRI scanner vendors. Although scan parameters were roughly matched across sites, they were not identical. Site was not included as a covariate in the models, and supplemental analysis showed that harmonisation of the diffusion-weighted MRI data did not have a substantial effect on the performance of the machine learning models. Thus, this fully automated procedure promises to be highly generalisable.

In terms of the imaging techniques, the free water and FA, analysis has been shown to be powerful for use as a biomarker in Parkinson’s disease,⁷ with changes in the posterior substantia nigra differentiating people with early stage Parkinson’s disease from healthy controls and longitudinal changes reflecting disease progression. The scan time of less than 12 min, although rather lengthy for clinical application, is reasonable compared with PET, with greater accessibility and no radioactive tracers. In the future, exploration of multi-shell diffusion-weighted MRI techniques⁸ might provide greater insights and biological interpretability.

One caveat for these findings is that the Parkinson’s disease cohort had fairly advanced disease severity. Because misdiagnosis is most likely to occur early in the disease,⁴ whether the AID-P would be applicable to the patients for which it is most relevant is unknown. The diffusion-weighted MRI only model did not effectively distinguish healthy controls from Parkinson’s disease (AUC 0.350). Although the MDS-UPDRS III alone is highly effective for detecting parkinsonism (AUC 0.983 for the MDS-UPDRS III only model for Parkinson’s disease compared with healthy controls), it is not yet clear whether the AID-P will be effective for patients with Parkinson’s disease with lower disease severity. Nonetheless, the mean disease duration for patients with Parkinson’s disease at most sites was less than 5 years, and the variables that contributed most to the models were consistent with established literature on neurodegeneration in early Parkinson’s disease, multiple systems atrophy and progressive supranuclear palsy,⁹,¹⁰ providing some confidence that the findings might be generalisable to individuals earlier on in their disease course.

An additional concern is that classification was based on clinical diagnosis, when post-mortem neuropathological confirmation remains the gold standard.² However, post-mortem classification data...
was available for five participants, for which the AID-P agreed with 100% accuracy. Ideally, future work would assess the accuracy of the AID-P in early stage patients with post-mortem confirmation and longitudinal data. Ultimately, availability of the AID-P for global use, possibly using cloud computing as suggested by the authors, could provide an essential tool to clinicians to aid with decisions regarding clinical care, to researchers to better understand fundamental differences between these related syndromes, and to clinical trials to select appropriate populations.

Shawna Abel, *Shannon Kolind
Department of Medicine (Neurology), University of British Columbia, Vancouver, Canada (SA); and Department of Medicine (Neurology), Department of Physics & Astronomy, and Department of Radiology, University of British Columbia, Vancouver, BC, Canada (SK)

SK is supported by a Scholar Award from the Michael Smith Foundation for Health Research and reports research grants from Roche, Genzyme, the Multiple Sclerosis Society of Canada, the Natural Sciences and Engineering Research Council of Canada, Vancouver Coastal Health Research Institute, and the Milan & Maureen Ilich Foundation and personal fees from Genzyme, outside the submitted work. SA declares no competing interests.

Copyright © 2019 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

1 Hughes AJ, Daniel SE, Ben-Shlomo Y, Lees AJ. The accuracy of diagnosis of parkinsonian syndromes in a specialist movement disorder service. Brain 2002; 125: 861–70.
2 Adler CH, Beach TG, Hentz JG, et al. Low clinical diagnostic accuracy of early vs advanced Parkinson disease: clinicopathologic study. Neurology 2014; 83: 406–12.
3 Archer DB, Bricker JT, Chu WT, et al. Development and validation of the automated imaging differentiation in parkinsonism (AID-P): a multicentre machine learning study. Lancet Digital Health 2019; published online Aug 27. http://dx.doi.org/10.1016/S2589-7500(19)30105-0.
4 Tripathi M, Tang CC, Feigin A, et al. Automated differential diagnosis of early parkinsonism using metabolic brain networks: a validation study. J Nucl Med 2016; 57: 60–6.
5 Scherfler C, Göbel G, Müller C, et al. Diagnostic potential of automated subcortical volume segmentation in atypical parkinsonism. Neurology 2016; 86: 1242–49.
6 Yang J, Borcu RG, Vaillancourt DE. Longitudinal progression markers of Parkinson’s disease: current view on structural imaging. 2018; Curr Neurol Neurosci Rep 18: 83.
7 Pasternak O, Shenton ME, Westin C-F. Estimation of extracellular volume from regularized multi-shell diffusion MRI. Med Image Comput Comput Assist Interv 2012; 15: 305–12.
8 Andica C, Kamagata K, Hatano T, et al. Neurite orientation dispersion and density imaging of the nigrostriatal pathway in Parkinson’s disease: retrograde degeneration observed by tract-profile analysis. Parkinsonism Relat Disord 2018; 51: 55–60.
9 Dickson DW, Rademakers R, Hutton ML. Progressive supranuclear palsy: pathology and genetics. Brain Pathol 2007; 17: 74–82.
10 Koga S, Aoki N, Uitti RJ, et al. When DLB, PD, and PSP masquerade as MSA: an autopsy study of 134 patients. Neurology 2015; 85: 404–12.