Accuracy of a Deep Learning System for Classification of Papilledema Severity on Ocular Fundus Photographs

Caroline Vasseneix, MD, Raymond P. Najjar, PhD, Xinxing Xu, PhD, et al., on behalf of the BONSAI Group

Cite as: Neurology® 2021;97:e369-e377. doi:10.1212/WNL.0000000000012226

Correspondence
Dr. Milea
dan.milea@snec.com.sg

Study Question
Can a deep learning system (DLS) accurately classify the severity of papilledema associated with increased intracranial pressure based on standard retinal fundus photographs?

What Is Known and What This Paper Adds
Papilledema severity at presentation is the most important predictor of subsequent visual outcomes, but traditional methods used to classify the severity of papilledema can be difficult to apply. This study shows that a DLS can accurately classify papilledema severities based on retinal fundus photographs.

Methods
For this prognostic model development study, the investigators used 2,103 mydriatic fundus photographs of papilledema from a multiethnic cohort of 965 patients with confirmed elevation of intracranial pressure to train a DLS to classify the severity of papilledema. These patients received care through 16 neuro-ophtalmology centers affiliated with the BONSAI consortium, which has a presence in 14 countries. A panel of experts classified 1,052 of the photographs as showing mild/moderate papilledema (MP) and the other 1,051 photographs as showing severe papilledema (SP). The investigators tested the performance of the DLS and 3 independent neuro-ophtalmologists with 214 photographs from 111 patients treated at 4 centers in the BONSAI consortium. The testing photographs included 92 that showed MP and 122 that showed SP. The primary outcome was the performance of the DLS at differentiating between MP and SP cases.

Results and Study Limitations
The DLS successfully differentiated between MP and SP cases (area under the receiver operating characteristic curve, 0.93; 95% confidence interval, 0.89–0.96), achieving levels of accuracy, sensitivity, and specificity that were comparable to the performance of the 3 neuro-ophtalmologist reviewers. This study provides Class II evidence that a deep learning system using mydriatic retinal fundus photographs accurately classified the severity of papilledema associated in patients with a diagnosis of increased intracranial pressure. In terms of study limitations, the present study is based on retrospective data, not taking into consideration visual function and optical coherence tomography data.

Study Funding and Competing Interests
This study was funded by the Singapore National Medical Research Council, the Duke–National University of Singapore Medical School, and the NIH. Some authors report consulting for healthcare companies. Go to Neurology.org/N for full disclosures.
Accuracy of a Deep Learning System for Classification of Papilledema Severity on Ocular Fundus Photographs

Caroline Vasseneix, Raymond P. Najjar, Xinxing Xu, et al.

Neurology 2021;97:e369-e377 Published Online before print May 19, 2021
DOI 10.1212/WNL.0000000000012226

This information is current as of May 19, 2021