COMMENTS AND RESPONSES

Evaluation of a System for Automatic Detection of Diabetic Retinopathy From Color Fundus Photographs in a Large Population of Patients With Diabetes

Response to Abràmoff et al.

We read with interest the recent article by Abràmoff et al.(1) but were disappointed in their conclusion that automated grading software could not be recommended for clinical practice.

Our group's published work (2) shows that automated grading of diabetic retinopathy, based on image-quality assessment and microaneurysm detection, can safely reduce the burden of grading in diabetic retinopathy screening programs. Comparing manual and automated grading against a reference standard grading of 14,406 images (from 6,722 patients), we found that our automated system attained a higher sensitivity for detection of patients requiring "full disease" grading than the manual graders. The automated system detected 97.9% of patients having referable diabetic retinopathy. Although the specificity of the automated system was lower than for manual graders, the grading workload was reduced and offered useful financial savings (3).

Screening is a means for reducing the risk of disease in the screened population, and, in practice, large-scale implementation means that there is a compromise between sensitivity and specificity. Hence a recommendation against using automated grading is only valid if it is shown that there is a higher performing and readily available alternate methodology. More specifically, it is important that an automated grading system is compared with what can be achieved by human experts who are routinely employed within a screening program. In the real world, such manual grading is imperfect. For example, we found that the full disease graders, whose job is to be highly specific, missed 18 of 330 cases of referable diabetic retinopathy (2).

Hence, our main criticism of the study by Abràmoff et al. is that the lack of a common reference standard resulted in insufficient evidence to draw their main conclusion, namely, that the automated grading software could not be recommended for clinical practice.

We also note two other factors that may have influenced the results and made them difficult to generalize. First, selection bias may have been a factor. The data were selected on the basis that patients previously shown to have diabetic retinopathy are not rescreened. While this may be the policy of the EyeCheck program, the data may not be regarded as "unselected" outside the context of this particular program. Second, the authors note that there seemed to be a slight effect associated with increasing camera resolution. However, the results show a variation in specificity from 22 to 83% depending on camera resolution. This suggests that performance may be greatly improved by using the higher resolution images.

We congratulate Abràmoff et al. on this study. However, we believe that the conclusions are not universally applicable. Our work shows that the automated analysis of retinal images does have an important role to play in diabetic retinal screening programs.

JOHN A. OLSON, MD[1]
PETER F. SHARP, PHD[2]
ALAN FLEMING, PHD[3]
SAM PHILIP, MD[1]

From the [1]Diabetes Retinal Screening Service, National Health Service Grampian, Aberdeen, Scotland; [2]Biomedical Physics and Bioengineering, University of Aberdeen, Aberdeen, Scotland; and the [3]Department of Radiology, University of Aberdeen, Aberdeen, Scotland.
Corresponding author: John Olson, john.olson@nhs.net.

● ●

References

1. Abràmoff MD, Niemeijer M, Suttorp-Schulten MSA, Viergever MA, Russell SR, van Ginneken B: Evaluation of a system for automatic detection of diabetic retinopathy from color fundus photographs in a large population of patients with diabetes. *Diabetes Care* 31:193-198, 2008
2. Philip S, Fleming AD, Goatman KA, Fonseca S, Mcnamee P, Scotland GS, Prescott GJ, Sharp PF, Olson JA: The efficacy of automated "disease/no disease" grading for diabetic retinopathy in a systematic screening programme. *Br J Ophthalmol* 91:1512-1517, 2007
3. Scotland GS, McNamee P, Philip S, Fleming AD, Goatman KA, Prescott GJ, Fonseca S, Sharp PF, Olson JA: Cost-effectiveness of implementing automated grading within the national screening programme for diabetic retinopathy in Scotland. *Br J Ophthalmol* 91:1518-1523, 2007