

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 Olson et al.

**W**e thank Olson et al. (1) for their close reading of our recent study (2), where we examined sensitivity and specificity of automated diabetic retinopathy detection and demonstrated an area under the receiver operating characteristic curve of 0.87. A limited, 500-patient sample of all 10,000 photographic exams was examined by multiple, masked experts. We felt uncomfortable recommending a system for clinical practice for which patient safety compared to an accepted (gold) standard could not be established, concluding that it should be tested against widely accepted clinical standards, if practical. We have recently presented studies of an improved algorithm on a new, larger dataset of 15,000 exams with an area under the curve of 0.90 (3). Most of these results support the work of Olson and colleagues (4), although the fact that their system only detects small hemorrhages and microaneurysms is a serious limitation in our view, and sensitivity and specificity based upon the photographic interpretation by a single reader is unlikely to become widely accepted.

Failure to detect large, rare, and/or advanced lesions deserves disproportionate attention. If a patient with isolated neovascularization of the disc, <1:5,000 in our series, were to be missed by a system but would not have been missed by a

person, that is a failure likely to lead to vision loss or blindness for that patient, potential litigation, and a backlash against implementation of automated detection.

Groups translating automated diabetic retinopathy detection into clinical practice operate in environments that differ on regulatory, legal, budgetary, and reimbursement aspects, but we disagree that “a recommendation against automated grading is only valid if it is shown that there is a higher performing and readily alternative methodology” (1). The currently established practice is human expert reading, and the burden of proof is therefore on the new system to be introduced, which is automated reading. For automated reading to gain widespread acceptance, no shortcuts regarding safety concerns will likely be permitted by regulatory agencies, payers, and patients.

One study’s entry criteria may be perceived as another’s selection bias. The target population of the EyeCheck project consists of patients who had not been previously identified to have diabetic retinopathy. In most settings, as patients are identified with diabetic retinopathy they are referred for evaluation or treatment, removing them from the screened population. To establish any other inclusion criteria would have constituted selection, affecting the potential application of this data to current clinical practice.

The potential positive effect of camera resolution on algorithmic performance is intriguing, although with less costly cameras presently offering at least 1,024 × 1,024 pixels, this debate may be self-limiting. We believe that comparison of algorithms to standardized datasets (<http://roc.healthcare.uiowa.edu>) as well as to the gold standard are required and should include: 1) demonstration in a prospective multicenter study of similar or better detection; on populations with defined race and ethnicity distributions, 2) acceptable comparison of detection to standard multifield stereo photographs read according to the Early Treatment of Diabetic Retinopathy Study standard; and 3) sensitivity/specificity analysis with standard and severity-weighted receiver operating characteristic curves.

In summary, we agree that automated detection of diabetic retinopathy can

make the prevention of blindness and vision loss objective, more accessible, and more cost-effective, provided safety issues are not overlooked.

MICHAEL D. ABRÀMOFF, MD, PHD<sup>1,2,3</sup>  
MEINDERT NIEMEIJER, PHD<sup>3</sup>  
MARIA S.A. SUTTORP-SCHULTEN, MD, PHD<sup>4</sup>  
MAX A. VIERGEVER, PHD<sup>5</sup>  
STEPHEN R. RUSSELL, MD<sup>1,2</sup>  
BRAM VAN GINNEKEN, PHD<sup>5</sup>

From the <sup>1</sup>Retina Service, Department of Ophthalmology and Visual Sciences, University of Iowa Hospitals and Clinics, Iowa City, Iowa; the <sup>2</sup>Department of Veterans Affairs, Iowa City VA Medical Center, Iowa City, Iowa; the <sup>3</sup>Department of Electrical and Computer Engineering, University of Iowa, Iowa City, Iowa; the <sup>4</sup>Ophthalmology Service, Onze Lieve Vrouwe Gasthuis (OLVG), Amsterdam, the Netherlands; and the <sup>5</sup>Image Sciences Institute, University Medical Center Utrecht, Utrecht, the Netherlands.

Corresponding author: Michael D. Abramoff, [michael-abramoff@uiowa.edu](mailto:michael-abramoff@uiowa.edu).

DOI: 10.2337/dc08-0952

© 2008 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. See <http://creativecommons.org/licenses/by-nc-nd/3.0/> for details.

•••••  
**References**

1. Olson JA, Sharp PF, Fleming A, Philip S: Evaluation of a system for automatic detection of diabetic retinopathy from color fundus photographs in a large population of patients with diabetes: response to Abramoff et al. (Letter). *Diabetes Care* 31: e63, 2008. DOI: 10.2337/dc08-0827
2. Abramoff MD, Niemeijer M, Suttorp-Schulten MS, 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
3. Abramoff MD, Van Ginneken B, Suttorp MSA, Russell SR, Niemeijer M: Improved computer aided detection of diabetic retinopathy evaluated on 10,000 screening exams (Abstract). *Invest Ophthalmol Vis Sci* 49:2735, 2008
4. 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