



Rapid grading of fundus photos for diabetic retinopathy using crowdsourcing (#5356)

Christopher J. Brady MD,¹ Andrea C. Villanti MPH PhD,² Jennifer Pearson MPH PhD,²

Thomas Kirchner PhD,² Omesh P. Gupta MD MBA,¹ Chirag P. Shah MD MPH

1. Wills Eye Hospital, MidAtlantic Retina, Philadelphia, PA 2. Schroeder Institute for Tobacco Research and Policy Studies, Legacy, Washington, DC 3. Ophthalmic Consultants of Boston, Boston, Massachusetts

Building a Corps of Citizen Scientists: *Turker Comments*

"This HIT was very good and a nice break from all of the bubbling surveys. Thank you!"

"I have learn about diabetes little bit"

"I really liked seeing the pics of the eye, very interesting"

Future Directions:

- Further refine AMT methodology and interface.
- Expansion of AMT model to large public image datasets.
 - Messidor (n=1200)
- Seek mechanisms to facilitate use of this tool in prospective screening programs.

Research supported by the
J. Arch McNamara Research Fund
of Wills Eye Hospital

Contact:

Christopher J Brady, MD brady@jhmi.edu

Purpose

To develop and validate a novel method for fundus photo grading.

Methods

An interface for fundus photo classification was developed for Amazon.com's crowd-sourcing interface (AMT), including 6 annotated training images.

Phase I:

- 19 expert-graded images were posted to AMT for grading by anonymous workers (AWs), with 10 repetitions per photo for an initial proof-of-concept
- AW's were asked to sequentially grade images as:
 - Normal/Abnormal
 - Normal/Mild to moderate/Severe
 - Normal/Mild/Moderate/Severe

Phase II

- 1 image from each of 4 grading categories posted for 500 unique AW interpretations to determine the ideal number of graders.

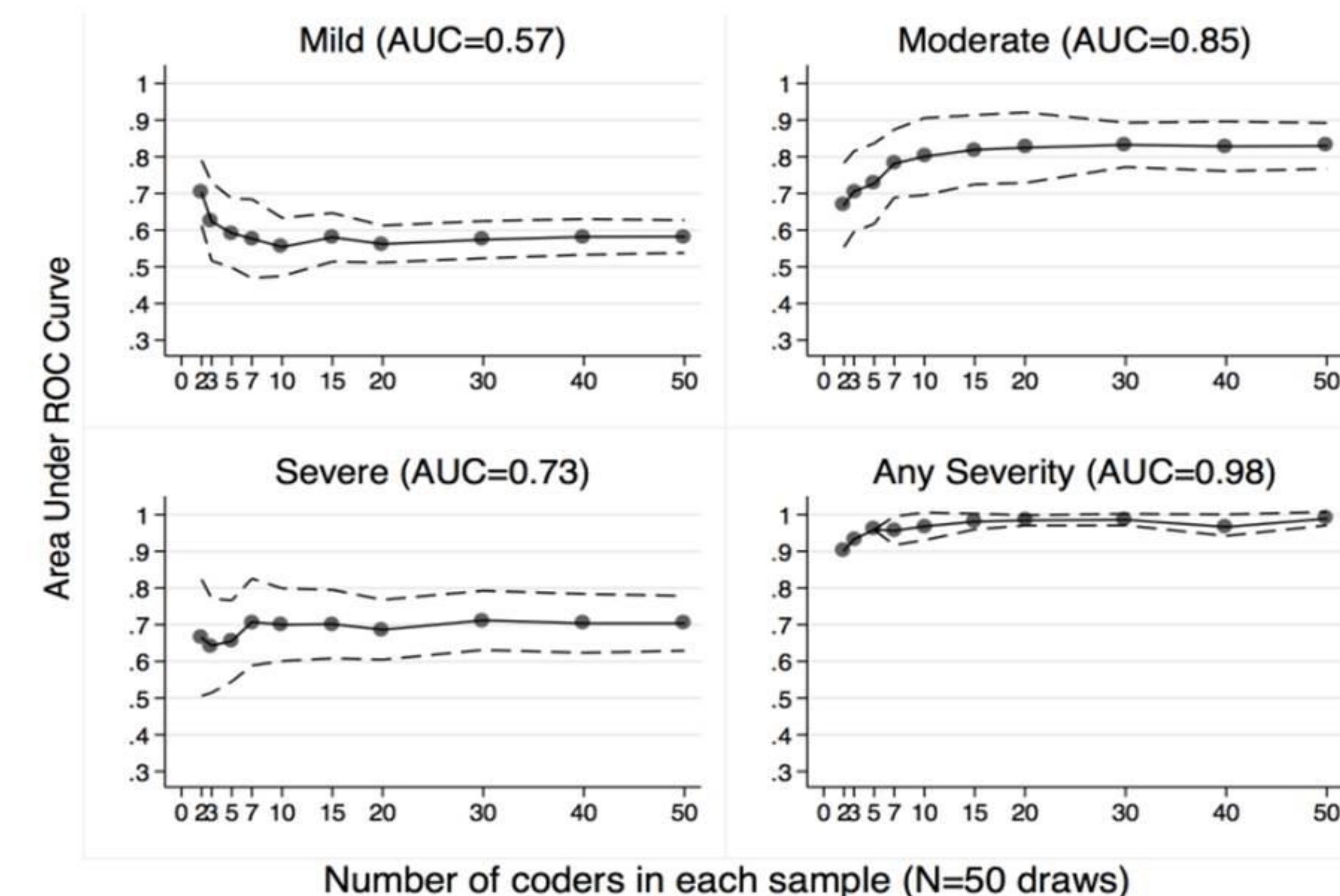
This is a photo of the inside of the eye. We are looking to label eyes as healthy or unhealthy with respect to diabetes. This task is difficult, but all good faith efforts will be accepted. Please answer both questions even if you are not sure. To close pop-out, move cursor left, or [tap here](#) on iPad.

Hover for examples:
[Normal](#)
[Abnormal - Mild](#)
[Abnormal - Moderate](#)
[Abnormal - Severe Example 1](#)
[Abnormal - Severe Example 2](#)
[Abnormal - Severe Example 3](#)

1) Rate this eye:
NOTE: The color of the photo is not a marker of disease.
 Normal (Healthy)
 Abnormal (mild, moderate, severe disease)

2) Is this photo good enough quality to grade?
 Yes
 No

Please leave any comments or feedback you have about this HIT or project:



Mean Time to complete ratings (in seconds)			
	Normal/Abnormal (Two-category rating)	Normal/Mild to moderate/ Severe (Three-category rating)	Normal/Mild/Moderate/ Severe (Four-category rating)
Mean time to complete HITs (95% CI)	25.16 (21.93 - 28.38)	50.87 (43.18 - 58.55)	54.52 (46.15 - 62.88)
Effective hourly wage	\$14.31	\$7.08	\$6.60

Results

Phase I:

- Sensitivity and specificity for any abnormality vs. normal were consistent across the three batches at 93.6-96.4% and 66.7-68.9%
- Average time to grade each image was 25 seconds, including time to review training images.

Phase II:

- Maximum AUC was reached after 7 graders for normal versus abnormal (AUC=0.98).
- AUC was asymptotic after 10 graders for mild, moderate, or severe vs. normal.

Conclusions

- With minimal training, the AMT workforce can rapidly and correctly categorize fundus photos of diabetic patients as normal or abnormal.
- This represents a novel and inexpensive (\$1/image) means to screen for diabetic retinopathy.
- Possibilities for microenterprise in emerging markets.