Published in final edited form as:

J AAPOS. 2017 April ; 21(2): 141–145. doi:10.1016/j.jaapos.2017.01.001.

Comparison of strategies for grading retinal images of premature infants for referral warranted retinopathy of prematurity

David L. Rogers, MD^a, Don L. Bremer, MD^a, Rae R. Fellows, MEd^d, Agnieszka Baumritter, MS^c, Ebenezer Daniel, MPH, PhD^c, Max Pastilli, MEd, MS^c, Gui-shang Ying, PhD^c, and Graham E. Quinn, MD, MSCE^d on behalf of the e-ROP Cooperative Group*

^aDepartment of Ophthalmology, The Ohio State University College of Medicine, Columbus

^bDepartment of Ophthalmology, Nationwide Children's Hospital, Columbus, Ohio

^cDepartment of Ophthalmology, University of Pennsylvania Perelman School of Medicine, Philadelphia

^dOphthalmology, The Children's Hospital of Philadelphia, University of Pennsylvania Perelman School of Medicine, Philadelphia

Abstract

Purpose—To determine the accuracy of identifying referral-warranted retinopathy of prematurity (RW-ROP, defined as any zone I ROP, stage 3 or worse, or plus disease) from retinal image sets using three grading protocols: a single optic disk-centered image, a set of 3 horizontal images, and a 5-image set.

Methods—In this secondary analysis of images from the e-ROP study, a weighted sample of 250 image sets from 250 infants (125 with RW-ROP and 125 without RW-ROP) was randomly selected. The sensitivities and specificities for detecting RW-ROP and its components from a single disk center image, along with nasal and temporal retinal images, were calculated and compared with the e-ROP grading of RW-ROP of all 5 retinal images (disk center and nasal, temporal, superior, and inferior retinal images).

Results—RW-ROP was identified with a sensitivity of 11.2% (95% CI, 6.79%–17.9%) using a single disk center image, with a sensitivity of 70.4% (95% CI, 61.9%–77.9%) using 3 horizontal images, and a statistically higher sensitivity of 82.4% (95% CI, 75.0%–89.0%) using all 5 images ($P = 0.002$). The specificities were 100%, 86.4%, and 90.4%, respectively. For grading using 3 horizontal images, sensitivity was 14.3% for plus disease, 25% for zone I ROP, and 71.2% for stage 3 or worse compared to 40.8%, 50%, and 79.8% for grading using 5-image sets, respectively.

Correspondence: David L. Rogers, MD, Department of Ophthalmology, Nationwide Children's Hospital, 555 S. 18th St., Ste 4-C, Columbus, OH 43205, (david.rogers@nationwidechildrens.org).

*A complete listing of the e-ROP Cooperative Group is provided in eSupplement 1, available at jaapos.org.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conclusions—Both a single, disk-centered, posterior pole image and 3 horizontal images were less effective than a 5-image set in determining the presence of RW-ROP on qualitative grading by trained readers.

The use of telemedicine in retinopathy of prematurity (ROP) is increasing worldwide to identify at-risk premature infants.^{1–7} Early detection of serious disease with properly timed treatment can significantly reduce visual impairment.⁸ However, infants continue to go blind from ROP, often due to poor access to timely detection and treatment.

Implementation of an ROP telemedicine system has several challenges. Imaging the retina in premature infants requires special training and skill, as does the interpretation of these images. The recently published e-ROP observational cohort study¹ reported use of a protocol that required evaluation of a set of 5 standard retinal images and an external image to assess pupillary dilation of each eye.⁹ Retinal images included disk-center and peripheral images with the disk offset to allow views of the superior, inferior, nasal, and temporal retina. A 5-image set was used for grading by masked nonphysician readers in a central reading center and assessed for the presence of referral-warranted ROP (RW-ROP), defined as presence of plus disease, or stage 3 or greater ROP, or any ROP in zone I. The primary outcome was a comparison of the nonphysician readers' gradings to the results of the diagnostic examination by the e-ROP certified clinical center ophthalmologist. The results showed that an ROP telemedicine system was feasible and valid with high sensitivity and specificity for detection of RW-ROP.¹

The e-ROP study used the RetCam Shuttle (Clarity MSI, Pleasanton, CA) 130° ROP lens for imaging to obtain a wide-field retinal image that generally includes the anatomic area defined as zone I when centered on the disk. Given the definition of RW-ROP and the ability of the RetCam to provide images of the retina that include zone I and posterior zone II, it is reasonable to posit that a single posterior pole image taken with the wide-angle lens could provide the necessary information to determine whether RW-ROP is present.¹⁰ Using computer-assisted analysis to determine tortuosity and width, investigators have shown high sensitivity in predicting the presence of potentially serious ROP, generally in the presence of plus disease.^{11–19}

An image set required in the e-ROP study protocol included 5 retinal images. However, obtaining these 5 retinal images with good quality from premature babies can be challenging. Simplifying the imaging requirement by obtaining fewer images and decreasing the duration of imaging could make the ROP telemedicine system more tolerable for the infant and acceptable to the ROP community. Therefore, we hypothesized that obtaining selected retinal images in an ROP telemedicine system would be simpler than requiring a 5-image set yet would still provide similar sensitivity and specificity. The purpose of this study was to determine whether RW-ROP could be accurately identified by trained nonphysician readers using fewer images than the 5-image set used in the e-ROP study. To answer this question, we regraded a subset of image sets obtained in the e-ROP study using a single disk-centered wide-angle (130°) posterior pole image only and a set of 3 horizontal images and compared the sensitivity/specificity from these gradings to the gradings obtained from the 5-image set.

Methods

This is a secondary analysis of images from the e-ROP study to determine the sensitivity and specificity of RW-ROP using three grading protocols: a single optic disk-centered image, a set of 3 horizontal images, and a set of 5 images. The e-ROP study design, the main results and the image grading procedures have been previously published.^{9,20,21} For this analysis, we randomly selected image sets from 125 infants with RW-ROP and 125 infants without RW-ROP based on clinical examination in either eye. This sample size provides half width of 95% confidence intervals for sensitivity and specificity to be within 7% when sensitivity and specificity are 80% or higher. A 5-image set from one eye per infant per image session was selected, for a total of 250 image sets. For eyes with RW-ROP, the image set selected was from the first image session at which the clinical center ophthalmologist documented morphology consistent with RW-ROP and at which all 5 retinal images were obtained. For eyes without RW-ROP, an image session at which all 5 retinal images were obtained was randomly selected.

These selected image sets were independently graded by 2 of 3 trained nonphysician readers using the same standard grading protocol as the e-ROP grading of 5-image sets.²¹ Gratings from the two readers were compared and discrepancies were adjudicated by the reading center director. A total of 20 image sets were randomly selected for regrading by trained readers and went through the same adjudication process for discrepancies to assess grade-regrade and inter-grader agreement. The 5-image set was not regraded for this study; the initial gradings obtained of these images for the primary analysis in the e-ROP study were used for this study.

Statistical Analysis

The primary statistical analysis was the calculation of sensitivity and specificity for detecting RW-ROP from gradings of the posterior pole image, the 3-horizontal-image set, and the 5-image set, by comparing the image grading results to findings of clinical eye examinations. The 95% confidence intervals for the sensitivity and specificity was calculated using the Wilson method for small sample.²² We also calculated sensitivity and specificity for detecting each component of RW-ROP (plus disease, zone I ROP, ROP stage 3 or above).

The grade-regrade agreement and inter-grader agreement was assessed based on percent of agreement and weighted kappa.²³ All the statistical analyses were performed in SAS V9.4 (SAS Institute Inc, Cary, NC).

Results

Grading for RW-ROP

Image sets from one eye of 250 infants were included: 125 infants with RW-ROP and 125 infants without RW-ROP. In the group of 125 eyes with RW-ROP, the presence of RW-ROP was detected from grading of a single wide-field posterior pole image in 14 eyes with a sensitivity of 11.2% (95% CI, 6.79%–17.9%). When the 3-horizontal-image set was used, RW-ROP was identified in 88 eyes with a sensitivity of 70.4% (95% CI, 61.9%–77.9%). When the 5-image set was used, RW-ROP was identified in 103 eyes with a sensitivity of

82.4% (95% CI, 75.0%–89.0%), which was significantly higher than that from grading of 3-image set ($P = 0.002$, McNemar's test). The corresponding specificities were 100%, 86.4%, and 90.4%, respectively (Table 1).

Grading for Plus Disease

In the 250 eyes graded, a total of 49 eyes were found to have plus disease based on the examination performed by the clinical center ophthalmologist. Plus disease was identified in 7 eyes, with a sensitivity of 14.3%, when either the single posterior pole or 3-image set were used and in 20 eyes, with a sensitivity of 40.8%, when the 5-image set was used. The specificities for identifying plus disease were high (>97%) using each of these grading methods (Table 2).

Grading for Zone of ROP

A total of 20 eyes had zone I ROP based on the eye examination performed by the clinical ophthalmologist. Image grading identified zone I ROP correctly in only 1 (5%) of the single posterior pole image sets, in 5 (25%) of the 3-image sets, and in 10 (50%) using the 5-image sets, with corresponding specificities of 97%, 90%, and 86%, respectively (Table 2).

Grading for Stage 3 ROP or Worse

A total of 104 eyes were found to have stage 3 or worse ROP from the clinical examination. Image grading identified stage 3 or worse ROP in 7 eyes (6.7%) using the single posterior pole image, in 74 eyes (71.2%) using the 3-image set, and in 83 eyes (79.8%) using the 5-image set, with corresponding specificities of 100%, 81%, and 85% respectively (Table 2).

Grade-regrade and Inter-grader Agreement

Among 20 image sets selected for evaluating the grade-regrade agreement and inter-grader agreement using 3 horizontal images, the grade-regrade agreement based on the consensus grading of two readers is good for RW-ROP (80% agreement; weighted kappa = 0.71), and moderate to good for its components (>80% agreement; weighted kappa = 0.50–0.71).

The inter-grader agreement is lower for RW-ROP (50% agreement; weighted kappa = 0.22), 100% for plus disease (weighted kappa = 1.0), 45% for stage 3 or higher, and 70% for zone I ROP.

Discussion

We found that RW-ROP could not be identified accurately using a single wide-field posterior pole image or using the 3-image set that included only additional images that provided views of the temporal and nasal retina. When this set of 3 images was used, the sensitivity and specificity of identifying RW-ROP increased, but it was still below that found when the set of 5 retinal images was used. Our study showed that identifying RW-ROP using the e-ROP telemedicine system was best accomplished using a set of 5 images of the retina.

Plus disease was not identified consistently by the readers in the single posterior pole image or in the set of 3 images. Our results comparing these 2 image sets for the detection of plus

disease was identical for sensitivity and specificity (Table 2). Initially, it may seem difficult to explain why accurate grading of plus disease could not be made from a single image centered on the optic nerve and why, on final analysis, plus disease was best identified using the set of all 5 images. One characteristic of the study that might help us understand this issue is that we used a qualitative assessment of plus disease; therefore, multiple images may have been necessary to determine whether or not plus disease was present. Several studies document the utility of a quantitative assessment of the posterior pole.^{24–29} The role of a quantitative analysis of the posterior pole vessels in a telemedicine screening program will require further study. Another possible reason plus disease was not accurately identified is the temporary morphological changes in plus disease created by the contact digital imaging using the RetCam.³⁰ Zepeda-Romero and colleagues³⁰ described how pressure on the cornea from the RetCam lens could secondarily result in compression, a decrease in intra-arterial diameter, and even collapse of retinal vessels. This transient change in caliber of the retinal vessels captured in the retinal image could result in the image reader occasionally missing plus disease. One factor that argues against this hypothesis for the e-ROP study results is that imagers in this study were specifically trained to recognize that compression of the central retinal artery might change the appearance of vessel caliber. Finally, we considered the role of image or shadow artifact created by a poorly dilated pupil obscuring the central part of the photo, preventing good visualization of the posterior pole vessels. In these cases, the additional images in the 3- and 5-image sets may have provided improved visualization of the optic nerve outside of the shadow artifact and allow for better assessment of the posterior pole vasculature. However, this would be unlikely, because images with significant artifact would not have been acceptable.

Plus disease or zone I ROP rarely occur in isolation, they often coexist with stage 3 ROP or above. Infants who are not identified as having plus disease can still be detected with RW-ROP if they were detected with one of the other RW-ROP components. Thus, our telemedicine imaging approach using a set of 5 images can achieve high sensitivity for detecting RW-ROP, although the sensitivity for plus disease alone is low. Our findings on the identification of plus disease are in contrast with those of a smaller study conducted by Lajoie and colleagues,¹⁰ who found that the detection of plus disease was comparable when single images or multiple images were used. It should be noted, that in their study, only temporal and nasal images were required, effectively comparing a single image with a set of 3 images. In addition, they used trained retina specialists as readers of the photographs.

Correctly identifying the zone of ROP was most accurately accomplished using the set of 5 images. This makes sense as a single posterior pole image that is centered on the optic disk will only show zone I and the most posterior portion of zone II. The additional images taken in the periphery allow for a broader view of zone II but are unable to capture the entire area of zone II without the addition of scleral depression, which e-ROP imagers did not do. One must also consider the application of the definition of zone I ROP. Although the zones of ROP are precisely defined in the international classification of ROP policy statement,^{8,31} the use of these definitions is determined at the bedside by the individual examiner.

It is also fair to say that ROP examinations can be stressful and uncomfortable for the infant. Temporary levels of distress occur during both the clinical examination and during retinal

imaging.^{32,33} It is possible that examiners over-called zone I in cases where they made this determination while the infant or the infant's eye was moving during the examination. On the other hand, the reader of the digital retinal image was able to apply a standard definition of zone I on the motionless retinal image with precision.

Overall image quality can also affect image grading. In the e-ROP study, retinal images were obtained by nonphysicians. These imagers were highly trained, underwent an extensive certification process, and worked in tertiary care centers where the volume of infants requiring ROP examinations was high. For these reasons, it is a fair assumption that these imagers had maximum proficiency and skill. In fact, the e-ROP study included 5,520 images sets, resulting in a total of 27,600 individual retinal images. When these images were graded for quality, 3% were missing, 91% were of adequate quality, and only 6% were poor quality.⁹ Therefore, it is unlikely that poor image quality played a role on image grading in our study.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Supported by a cooperative agreement grant (U10 EY017014) from the National Eye Institute of the National Institutes of Health, Department of Health and Human Services.

References

1. Lorenz B, Bock M, Müller HM, Massie NA. Telemedicine based screening of infants at risk for retinopathy of prematurity. *Stud Health Technol Inform.* 1999; 64:155–63. [PubMed: 10747534]
2. Lorenz B, Spasovska K, Elflein H, Schneider N. Wide-field digital imaging based telemedicine for screening for acute retinopathy of prematurity (ROP). Six-year results of a multicentre field study Graefes. *Arch Clin Exp Ophthalmol.* 2009; 247:1251–62.
3. Vinekar A, Gilbert C, Dogra M, et al. The KIDROP model of combining strategies for providing retinopathy of prematurity screening in underserved areas in India using wide-field imaging, telemedicine, non-physician graders and smart phone reporting. *Indian J Ophthalmol.* 2014; 62:41–9. [PubMed: 24492500]
4. Breidenstein BG, Dai S. Wide-field digital retinal imaging for retinopathy of prematurity screening in Australasia: a survey of practicing ophthalmologists. *J Pediatr Ophthalmol Strabismus.* 2014; 51:375–8. [PubMed: 25427308]
5. Fijalkowski N, Zheng LL, Henderson MT, et al. Stanford University Network for Diagnosis of Retinopathy of Prematurity (SUNDROP): five years of screening with telemedicine. *Ophthalmic Surg Lasers Imaging Retina.* 2014; 45:106–13. [PubMed: 24444469]
6. Weaver DT, Murdock TJ. Telemedicine detection of type 1 ROP in a distant neonatal intensive care unit. *J AAPOS.* 2012; 16:229–33. [PubMed: 22681938]
7. Skalet AH, Quinn GE, Ying GS, et al. Telemedicine screening for retinopathy of prematurity in developing countries using digital retinal images: a feasibility project. *J AAPOS.* 2008; 12:252–8. [PubMed: 18289897]
8. Fierson WM. American Academy of Pediatrics Section on Ophthalmology; American Academy of Ophthalmology; American Association for Pediatric Ophthalmology and Strabismus; American Association of Certified Orthoptists. Screening examination of premature infants for retinopathy of prematurity. *Pediatrics.* 2013; 131:189–95. [PubMed: 23277315]

9. Quinn GE, Ying GS, Daniel E, et al. e-ROP Cooperative Group. Validity of a telemedicine system for the evaluation of acute-phase retinopathy of prematurity. *JAMA Ophthalmol.* 2014; 132:1178–84. [PubMed: 24970095]
10. Lajoie A, Koreen S, Wang L, et al. Retinopathy of prematurity management using single-image vs multiple-image telemedicine examinations. *Am J Ophthalmol.* 2008; 146:298–309. [PubMed: 18547536]
11. Rabinowitz MP, Grunwald JE, Karp KA, Quinn GE, Ying GS, Mills MD. Progression to severe retinopathy predicted by retinal vessel diameter between 31 and 34 weeks of postconception age. *Arch Ophthalmol.* 2007 Nov.125:1495–500. [PubMed: 17998510]
12. Wilson CM, Cocker KD, Moseley MJ, et al. Computerized analysis of retinal vessel width and tortuosity in premature infants. *Invest Ophthalmol Vis Sci.* 2008; 49:3577–85. [PubMed: 18408177]
13. Grunwald L, Mills MD, Johnson KS, et al. The rate of retinal vessel dilation in severe retinopathy of prematurity requiring treatment. *Am J Ophthalmol.* 2009; 147:1086–91. [PubMed: 19286152]
14. Shah DN, Karp KA, Ying GS, Mills MD, Quinn GE. Image analysis of posterior pole vessels identifies type 1 retinopathy of prematurity. *J AAPOS.* 2009; 13:507–8. [PubMed: 19840733]
15. Shah DN, Wilson CM, Ying GS, et al. Semiautomated digital image analysis of posterior pole vessels in retinopathy of prematurity. *J AAPOS.* 2009; 13:504–6. [PubMed: 19840732]
16. Ghodasra DH, Karp KA, Ying GS, et al. Risk stratification of preplus retinopathy of prematurity by semiautomated analysis of digital images. *Arch Ophthalmol.* 2010; 128:719–23. [PubMed: 20547949]
17. Wallace DK, Freedman SF, Hartnett ME, Quinn GE. Predictive value of pre-plus disease in retinopathy of prematurity. *Arch Ophthalmol.* 2011; 129:591–6. [PubMed: 21555612]
18. Shah DN, Wilson CM, Ying GS, et al. Comparison of expert graders to computer-assisted image analysis of the retina in retinopathy of prematurity. *Br J Ophthalmol.* 2011; 95:1442–5. [PubMed: 21270432]
19. Ghodasra DH, Thuangtong A, Karp KA, et al. The rate of change in retinal vessel width and tortuosity in eyes at risk for retinopathy of prematurity. *J AAPOS.* 2012; 16:431–6. [PubMed: 23084379]
20. Quinn GE. e-ROP Cooperative Group. Telemedicine approaches to evaluating acute-phase retinopathy of prematurity: study design. *Ophthalmic Epidemiol.* 2014; 21:256–67. [PubMed: 24955738]
21. Daniel E, Quinn GE, Hildebrand PL, et al. e-ROP Cooperative Group. Validated System for Centralized Grading of Retinopathy of Prematurity: Telemedicine Approaches to Evaluating Acute-Phase Retinopathy of Prematurity (e-ROP) Study. *JAMA Ophthalmol.* 2015; 133:675–82. [PubMed: 25811772]
22. Wilson EB. Probable inference, the law of succession, and statistical inference. *J Am Stat Assoc.* 1927; 22:209–12.
23. Cohen J. Weighed kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin.* 1968; 70:213–20. [PubMed: 19673146]
24. Wallace DK, Freedman SF, Zhao Z, Jung SH. Accuracy of ROPTool vs individual examiners in assessing retinal vascular tortuosity. *Arch Ophthalmol.* 2007; 125:1523–30. [PubMed: 17998514]
25. Wallace DK, Quinn GE, Freedman SF, Chiang MF. Agreement among pediatric ophthalmologists in diagnosing plus and pre-plus disease in retinopathy of prematurity. *J AAPOS.* 2008; 12:352–6. [PubMed: 18329925]
26. Wallace DK. Computer-assisted quantification of vascular tortuosity in retinopathy of prematurity. *Trans Am Ophthalmol Soc.* 2007; 105:594–615. [PubMed: 18427631]
27. Shah DN, Karp KA, Ying GS, Mills MD, Quinn GE. Image analysis of posterior pole vessels identifies type 1 retinopathy of prematurity. *J AAPOS.* 2009; 13:507–8. [PubMed: 19840733]
28. Shah DN, Wilson CM, Ying GS, et al. Semiautomated digital image analysis of posterior pole vessels in retinopathy of prematurity. *J AAPOS.* 2009; 13:504–6. [PubMed: 19840732]
29. Grunwald L, Mills MD, Johnson KS, et al. The rate of retinal vessel dilation in severe retinopathy of prematurity requiring treatment. *Am J Ophthalmol.* 2009; 147:1086–91. 1091.e1–2. [PubMed: 19286152]

30. Zepeda-Romero LC, Martinez-Perez ME, Ruiz-Velasco S, Ramirez-Ortiz MA, Gutierrez-Padilla JA. Temporary morphological changes in plus disease induced during contact digital imaging. *Eye*. 2011; 25:1337–40. [PubMed: 21760627]
31. International Committee for the Classification of Retinopathy of Prematurity. The International Classification of Retinopathy of Prematurity revisited. *Arch Ophthalmol*. 2005; 123:991–9. [PubMed: 16009843]
32. Grabska J, Walden P, Lerer T, et al. Can oral sucrose reduce the pain and distress associated with screening for retinopathy of prematurity? *J Perinatol*. 2005; 25:33–5. [PubMed: 15343351]
33. Moral-Pumarega MT, Caserío-Carbonero S, De-La-Cruz-Bértolo J, Tejada-Palacios P, Lora-Pablos D, Pallás-Alonso CR. Pain and stress assessment after retinopathy of prematurity screening examination: indirect ophthalmoscopy versus digital retinal imaging. *BMC Pediatr*. 2012; 12:132. [PubMed: 22928523]

Table 1

Sensitivity and specificity of 1-image, 3-image, and 5-image grading

No. images	Eyes with RW-ROP from clinical exam (n = 125)		Eyes without RW-ROP from clinical exam (n = 125)	
	Image grading positive in RW-ROP	Sensitivity (95% CI)	Image grading negative in RW-ROP	Specificity (95% CI)
1	14	11.2% (6.79%–17.9%)	125	100% (97.0%–100%)
3	88	70.4% (61.9%–77.9%)	108	86.4% (79.3%–91.3%)
5	103	82.4% (75.0%–89.0%)	113	90.4% (84.0%–94.4%)

CI, confidence interval; *RW-ROP*, referral-warranted retinopathy of prematurity.

Table 2

Sensitivity and specificity of RW-ROP by Component from 1-image, 3-image, and 5-image grading sets

Components of RW-ROP	Eyes with specific RW-ROP components based on clinical exam		Eyes without specific RW-ROP components based on clinical exam	
	Image grading positive in RW-ROP component	Sensitivity (95% CI)	Image grading negative in RW-ROP component	Specificity (95% CI)
Plus disease	(N = 49)		(N = 201)	
1 image	7	14.3% (7.1%–26.7%)	200	99.5% (97.2%–99.9%)
3 images	7	14.3% (7.1%–26.7%)	200	99.5% (97.2%–99.9%)
5 images	20	40.8% (28.2%–54.8%)	195	97.0% (93.6%–98.6%)
Zone I ROP	(N = 20)		(N = 230)	
1 image	1	5.0% (8.9%–23.6%)	223	97.0% (93.9%–98.5%)
3 images	5	25.0% (11.2%–46.9%)	207	90.0% (85.4%–93.2%)
5 images	10	50.0% (29.9%–70.1%)	198	86.1% (81.0%–90.0%)
ROP stage 3 or worse	(N = 104)		(N = 146)	
1 image	7	6.7% (3.3%–13.3%)	146	100% (97.4%–100%)
3 images	74	71.2% (61.8%–79.0%)	118	80.8% (73.7%–86.4%)
5 images	83	79.8% (71.1%–86.4%)	124	84.9% (78.2%–89.8%)

CI, confidence interval; *RW-ROP*, referral warranted retinopathy of prematurity.