Published in final edited form as:

J Pediatr Ophthalmol Strabismus. 2015 March 1; 52(2): 85–92. doi:10.3928/01913913-20150114-02.

Evaluation of the Accuracy of Grading Indirect Ophthalmoscopy Video Images for Retinopathy of Prematurity Screening

Sasapin G. Prakalapakorn, MD, MPH¹, David K. Wallace, MD, MPH¹, Riana S. Dolland, FCOphth, MMed², and Sharon F. Freedman, MD¹

¹Duke University Eye Center, Durham, North Carolina, USA

²University of the Witwatersrand, Johannesburg, South Africa

Abstract

Purpose—To determine whether digital retinal images obtained using a video indirect ophthalmoscopy system (Keeler) can be accurately graded for zone, stage, and plus or pre-plus disease and used to screen for type 1 retinopathy of prematurity (ROP).

Methods—We retrospectively reviewed the charts of 114 infants who had retinal video images acquired using the Keeler system during routine ROP examinations. Two masked ophthalmologists (1 expert and 1 non-expert in ROP screening) graded these videos for image quality, zone, stage, and pre-plus or plus disease. We compared the ophthalmologists' grades of the videos against the clinical examination results, which served as the reference standard. We then determined the sensitivity/specificity of 2 predefined criteria for referral in detecting disease requiring treatment (i.e. type 1 ROP).

Results—Of images the expert considered fair or good quality (n=68), the expert and non-expert correctly identified zone (75% vs. 74%, respectively), stage (75% vs. 40%, respectively), and the presence of pre-plus or plus disease in 79% of images. Expert and non-expert judgment of prethreshold disease, pre-plus or plus disease had 100% sensitivity and 75% vs. 79% specificity, respectively, for detecting type 1 ROP. Expert and non-expert judgment of pre-plus or plus disease had 92% vs. 100% sensitivity and 77% vs. 82% specificity, respectively, for detecting type 1 ROP.

Conclusions—High-quality retinal video images can be read with high sensitivity and acceptable specificity to screen for type 1 ROP. Grading for pre-plus or plus disease alone may be sufficient for the purpose of ROP screening.

Corresponding author and address for reprint requests: S. Grace Prakalapakorn MD, MPH, Assistant Professor of Ophthalmology, Duke Eye Center, 2351 Erwin Rd, DUMC 3802, Durham, NC 27710, phone 919-684-0584, fax 919-684-6096, grace.prakalapakorn@duke.edu.

Meeting Presentations: Presented in part as an oral presentation at the 2014 American Association for Pediatric Ophthalmology and Strabismus Annual Meeting, Rancho Mirage, CA, April 2014.

Financial Conflict of Interest: No authors have any financial, commercial, or proprietary interest in the materials presented herein.

Introduction

Retinopathy of prematurity (ROP) is a significant cause of childhood blindness in many middle income countries,¹ and is the leading preventable and treatable cause of childhood blindness in the United States (US).² While the burden of childhood blindness from ROP could be reduced by appropriate screening and treatment, there are many barriers to effective ROP screening including both a shortage of ophthalmologists skilled at ROP screening³ and a lack of access to these ophthalmologists. While current guidelines in the US state that ROP screening examinations should be performed by an ophthalmologist trained in ROP screening using binocular indirect ophthalmoscopy,³ alternative screening approaches are needed since the supply of ophthalmologist ROP screeners is unlikely to meet the need in many parts of the world.¹

A previous study found that the Vantage Plus LED Digital Binocular Indirect Ophthalmoscope system (Keeler Instruments Inc, Broomall, PA), a binocular indirect ophthalmoscope with an integrated camera that can capture and store still and/or dynamic digital images during the examination, could capture still images of sufficient quality to demonstrate the presence of posterior pole disease (pre-plus or plus disease).⁴ Because the field of view obtained by the Keeler system is similar to that seen during the examination with binocular indirect ophthalmoscopy, one could infer that the Keeler system would be able not only to capture valuable information about the posterior pole, but also the retinal periphery. If the Keeler system could capture images of both the posterior pole, and the retinal periphery that were of sufficient quality to demonstrate the zone and stage of ROP, this system would have potential as a tool for both ROP screening and teaching.

The primary purpose of this study was to determine whether digital video images of the retina obtained using an indirect ophthalmoscopy imaging system could be accurately graded for zone and stage of ROP and the presence of plus or pre-plus disease. The secondary aim was to determine whether these digital retinal video images could be accurately graded to detect the presence of disease requiring treatment (i.e. type 1 ROP) by comparing two predefined criteria for referral.

Patients and Methods

This study was approved by the Duke Health System Institutional Review Board and conformed to the requirements of the United States Health Insurance Portability and Privacy Act. A retrospective chart review was performed of infants screened for ROP in the Duke University Neonatal Intensive Care Unit (NICU). At our institution, infants are screened for ROP per recommended guidelines at the time of screening.⁵ All examinations were performed by one of two pediatric ophthalmologists (SFF or DKW), both experienced ROP examiners. As part of routine ROP screening, we digitally recorded every examination using the Vantage Plus LED Digital Binocular Ophthalmoscope System and a 28 D condensing lens. Follow-up examinations occurred according to current published guidelines at the time of the examination.^{5,6} The presence or absence of ROP and the zone, stage, and presence or absence of plus or pre-plus disease were documented for each eye according to current international classification guidelines.^{7,8} We extracted demographic data including birth

date, gestational age (GA), birth weight (BW), and date of ROP examinations. Post-menstrual age (PMA) was calculated based on examination date and birth date. Clinical examination findings were recorded as the lowest zone and highest stage ROP documented for each eye during a given examination session.

To be eligible for inclusion in this study, infants had to have been hospitalized in the Duke University NICU, screened for ROP from November 1, 2009 to November 16, 2011, have digital video retinal images acquired using the Keeler system during their routine ROP examination(s), and have BW <1500g or GA $\geq$ 30 weeks. We excluded Infants if they had received laser or anti-vascular endothelial growth factor (VEGF) treatment prior to having an examination recorded by the Keeler system during the study period.

We chose video images from one examination date for each infant. We enhanced the sample with images of stage 3 ROP by preferentially choosing examination dates where infants were diagnosed as having stage 3 ROP. If an infant had stage 3 ROP on more than one examination date, then of those examination dates, one date was randomly chosen.

Otherwise, for each infant, an examination date was randomly selected from all examination dates in which there was a video recording for that infant. Any examination performed after laser or anti-vascular endothelial growth factor treatment was excluded from our study.

After the examination date was chosen, one of the authors (SGP) reviewed and edited all videos recorded by the Keeler system for the infant on that date using a video editing program (Windows Movie Maker 2.6, Microsoft, Redmond, WA). All videos were edited in order to remove extraneous non-retinal images (e.g. images of placing or removing the lid speculum, etc.) or images in which the examiner was pointing at ocular pathology. Only images from the right eye were included for each infant.

After the videos were edited, these images were randomly numbered and presented to the graders without any accompanying demographic or clinical information. Masked to demographic information and clinical examination findings, two ophthalmologists, one expert (SFF) and one non-expert in ROP screening (RSD), independently reviewed the videos and evaluated them for (1) image quality, (2) zone, (3) stage, and (4) the presence of pre-plus or plus disease. The non-expert in ROP screening was a general ophthalmologist who received about 6 months of training in ROP screening during 4 years of ophthalmology residency, had no additional fellowship training, and had worked for about 7 months after residency prior to participating in this study. The non-expert did not have any specific training to participate in this study. In this study, “pre-plus disease” was defined according to the international classification of ROP (ICROP)-revisited guidelines definition of “vascular abnormalities of the posterior pole that are insufficient for the diagnosis of plus disease but that demonstrate more arterial tortuosity and more venous dilatation than normal”⁷ and “plus disease” was defined as the presence in $\geq$ 2 quadrants of the eye of sufficient vascular dilation and tortuosity as compared to a standard photograph.^{7,8} Based on the ability of the grader to determine the stage of ROP in the video selected for each infant, image quality was graded as follows: “good,” a video in which the grader could easily discern the stage of ROP; “fair,” a video in which it was difficult to clearly discern the stage of ROP; and “poor,” a video in which the grader was unable to discern the stage of ROP.

The zone (I, II, III) and stage (none, 1, 2, or 3) of ROP were defined according to ICROP-revisited guidelines.⁷ Because the expert performed ROP screening during 50% of the time period of image acquisition, she pledged to recuse herself from grading any videos that she recognized.

We used SAS 9.3 (SAS Institute Inc, Cary, NC) for all statistical analysis. Prior to the commencement of this study, a sample size calculation indicated that in order to appropriately power our study to detect a sensitivity of 0.80 (95% CI, 0.725-0.875), a sample size of 114 was required. Before analyzing the data, we defined the “reference standard” as the diagnosis of ROP by indirect ophthalmoscopy during the clinical examination. For the primary analysis of accuracy, we evaluated the ability of each grader to accurately identify the zone and stage of ROP and the presence or absence of pre-plus or plus disease on reviewing the Keeler videos compared to the “reference standard” (i.e. clinical examination diagnosis). In our secondary analysis, we determined the accuracy (i.e. sensitivity and specificity) of 2 predefined criteria for referral in detecting disease requiring treatment (i.e. type 1 ROP). Because we wanted to evaluate if grading the zone and stage of ROP on reviewing the videos would increase the sensitivity and specificity of screening for type 1 ROP compared to the grading for the presence of pre-plus or plus disease alone, our first criteria for referral was defined as the presence of prethreshold disease, pre-plus, or plus disease and our second criteria for referral was defined as the presence of pre-plus or plus disease. Type 1 ROP is the presence of stage 3 in zone I, any stage ROP with plus disease in zone I, or stage 2 or 3 with plus disease in zone II.⁶ Prethreshold disease is any ROP in zone I, stage 2 in zone II with plus disease, or stage 3 in zone II.⁹

Results

A total of 114 infants were included (median GA, 26 weeks [range, 23-33], median BW, 840 grams [range, 450-2300], median PMA at examination, 35 weeks [range, 29-46]). As diagnosed by indirect ophthalmoscopy, our enhanced sample of images was comprised of 14% with retinal vascularization that ended in zone I, 74% in zone II, and 12% in zone III; 15% with stage 1, 20% stage 2, and 26% stage 3 ROP; and 15% with pre-plus and 9% with plus disease (Table 1).

Based on the ability of each grader to determine the stage of ROP in the video, the expert judged 60% (n=68) and the non-expert 55% (n=63) of the videos to have fair or good image quality (Table 1). Of the videos the expert felt had fair or good image quality, this set of videos contained a higher percentage of those from the entire set of videos with zone II > I > III, stage 3 > 2 > 1 > immature vasculature, and plus disease > pre-plus disease > a normal posterior pole (Table 1). The expert did not recognize any of the videos used in this study and thus did not recuse herself from grading any videos. The videos were graded at least 1.5 years after the expert acquired the images. Of the videos the non-expert felt had fair or good image quality, this set of videos contained a higher percentage of those from the entire set of videos with zone I > III > II, stage 3 > 1 > 2 > immature vasculature, and plus disease > pre-plus disease > a normal posterior pole (Table 1).

Of those videos that the expert felt were of fair or good quality, the expert and non-expert correctly identified zone (75% vs. 74%, respectively), stage (75% vs. 40%, respectively; Figure 1), and the presence of pre-plus or plus disease in 79% of the videos (Table 2).

Using the reference standard of indirect ophthalmoscopy-reported type 1 ROP, the sensitivity of grading Keeler videos for the presence of prethreshold disease, pre-plus or plus disease for both the expert and non-expert was 100% and the specificity was 75% and 79%, respectively (Table 3).

Using the reference standard of indirect ophthalmoscopy-reported type 1 ROP, the sensitivity of grading Keeler videos for pre-plus or plus disease was 92% for the expert and 100% for the non-expert and the specificity was 77% for the expert and 82% for the non-expert (Table 4). Of the cases of clinically-diagnosed type 1 ROP by indirect ophthalmoscopy, the non-expert graded all cases as either pre-plus or plus, while the expert judged one case as having a normal posterior pole. On further examination of this case, we had the expert evaluate the video of the other eye from the same examination session and the expert graded the video as having good quality and identified the eye has having stage 3 ROP in zone I with pre-plus disease.

Looking at only the videos judged by the expert to be of good or fair quality, the inter-grader reliability was 68% ($\kappa=0.1$) for zone, 56% ($\kappa=0.4$) for stage of ROP, 79% ($\kappa=0.6$) for the presence or absence of pre-plus or plus disease, 91% ($\kappa=0.8$) for the presence of prethreshold disease, pre-plus or plus disease, and 88% ($\kappa=0.7$) for the presence of pre-plus or worse disease on Keeler video images.

Discussion

This study found that digital video images of the retina of fair or good quality obtained using the Keeler system could be graded for ROP. Compared with the entire set of video images, those videos the expert and non-expert considered to have fair or good image quality were more likely to include eyes with stage 3 or plus disease (Table 1). It is unclear whether the clinical examiner dwelled more on viewing severe rather than mild disease pathology and/or that more severe ROP was more distinct and thus, easier to capture on imaging.

While the ROP expert was able to accurately grade zone, stage, and the presence of pre-plus or plus disease when grading the video images, the non-expert had a harder time identifying the stage of ROP (Table 2). Looking at zone, neither grader misclassified zone by more than one zone (i.e. zone I disease may have been graded as zone I or II, but never zone III disease and zone III disease may have been graded as zone II or III, but never zone I disease). For stage, the expert never misclassified stage by more than 1 stage, but the non-expert had a harder time identifying stage 3 ROP. For the presence of pre-plus or plus disease, neither grader misclassified disease by more than one category (i.e. “no posterior pole disease” may have been graded as no disease or pre-plus disease, but never plus disease; conversely, plus disease may have been graded as pre-plus or plus disease, but never as no posterior pole disease). More importantly, plus disease was always identified as pre-plus or worse disease by both the expert and the non-expert.

Looking at the Keeler system as a true “screening tool” for ROP, we wished to evaluate whether, in addition to grading for the presence of pre-plus or plus disease, the grading of zone and stage of ROP would further enhance the identification (or decrease the chance of false negative classification) of infants with type 1 ROP. In a previous study, we found that ROP experts could grade still images of the posterior pole acquired by the Keeler system with high accuracy for pre-plus or plus disease compared to the results of the clinical examination and that the grading of pre-plus or plus disease was highly sensitive and specific for the presence of plus disease.⁴ We wanted to evaluate how inclusive we needed to be in our screening criteria in order to make the grading of videos captured by the Keeler system an acceptable “screening test” for ROP (i.e. whether the grading of zone and stage of ROP, in addition to the grading of the presence of pre-plus or plus disease, would further increase the sensitivity and specificity of screening for type 1 ROP). Thus in our secondary analysis, we compared the sensitivities and specificities of two scenarios. In the first scenario (which we will refer to as “posterior pole only”), we found that if we had graders evaluate video images in only one eye for pre-plus or plus disease, the expert missed one case of type 1 ROP (sensitivity = 92%), while the non-expert was able to correctly identify all cases of type 1 ROP. Both graders had high specificity (77% for the expert and 82% for the non-expert) for identifying type 1 ROP. In the second scenario (which we will refer to as “posterior pole and periphery”), we found that if we had the graders evaluate the videos in one eye for zone, stage, and the presence of pre-plus or plus disease to diagnose the presence of prethreshold disease, pre-plus or plus disease, both the expert and non-expert had 100% sensitivity, while their specificity was 75% (expert) and 79% (non-expert), for identifying type 1 ROP.

Looking at the two scenarios above, both showed a high sensitivity and specificity for identifying infants with treatment-requiring (type 1) ROP compared to the clinical examination, suggesting that the Keeler system shows promise as an ROP screening tool. A good “screening test” must have a high sensitivity so that those with treatment-requiring disease are not missed. In the posterior pole only scenario, of the cases of clinically-diagnosed treatment-requiring (type 1) ROP by indirect ophthalmoscopy, one case (1/12, 8%, an eye with stage 3 in zone I with pre-plus disease) was judged to have a normal posterior pole by review of the video by the expert (Table 4); the video images were judged by that grader to be of fair quality. So, we also had the expert evaluate the video captured by the Keeler system of the other eye on the same infant from the same examination session and she judged the video to be of good quality and to have stage 3 ROP in zone I with pre-plus disease. In this example, the infant would have failed screening criteria if both eyes were evaluated because of the presence of pre-plus disease noted in the left eye on review of the infant’s video images. Because ROP can present asymmetrically in the same individual, it is important that each infant have both eyes examined in a true “screening” scenario, so that the more severely affected eye is not missed on screening. Thus, if screening criteria using the Keeler system required the evaluation of video images from both eyes and image quality to be fair or good, the presence of pre-plus or plus disease in either eye would trigger a standard diagnostic examination by an ophthalmologist trained in ROP screening using indirect ophthalmoscopy to not only re-evaluate the infant for the presence of pre-plus or plus disease, but also for the zone and stage of ROP in the retinal periphery. Further research

is needed to evaluate the true sensitivity and specificity of screening for type 1 ROP if both eyes are evaluated simultaneously and prospectively.

While a good “screening test” must have a high sensitivity so that those with treatment-requiring disease are not missed, high specificity is also desirable, so that those without disease are not subjected to unnecessary examinations. Both scenarios described above had high specificities for ruling out type 1 ROP when it was not present according to the clinical examination (Tables 3 and 4).

While the graders showed poor inter-grader reliability for grading zone and stage of ROP, they both showed high inter-grader reliability for grading for both the presence of prethreshold disease, pre-plus or plus disease; and the presence of pre-plus or plus disease (both as a dichotomous and trichotomous variable). The inter-grader reliability for grading for the presence of pre-plus or plus disease was similar to that reported in a previous study evaluating the ability to grade for the presence of pre-plus or plus disease in still images of the posterior pole acquired by the Keeler system.⁴

If it served adequately as a pure “screening test”, examination of only the posterior pole for the presence of pre-plus or plus disease, and not also of the zone and stage of ROP, would provide certain advantages. Namely, “posterior pole only” screening (i.e. evaluating only for the presence or absence of pre-plus or plus disease in the posterior pole) would save time and perhaps be less stressful on the infants requiring screening. In addition, the posterior pole is easier to image than the retinal periphery, can be captured in a still (which has a smaller file size) vs. a dynamic video image, and lends itself to a wider variety of imaging modalities (including other narrow-field imaging devices such as the NIDEK NM200-D camera¹⁰ (Nidek Inc., Gamagori, Japan) and Pictor camera¹¹ (Volk Optical Inc., Mentor, OH)), as well as to a wider expertise-level amongst potential imagers and graders. As shown by our study, it may be more difficult for a non-expert, compared to an ROP expert, to accurately identify stage of ROP from review of video images. Therefore, by omitting assessment of stage of ROP from the image grading process, one could theoretically improve reliability of graders identifying infants needing a diagnostic examination by an ophthalmologist with binocular indirect ophthalmoscopy. This “posterior pole disease only” grading strategy might, in turn, allow acceptable image grading (for treatment-requiring ROP) by those with a wider range of training/expertise.

This study’s findings must be considered in light of a number of limitations. With respect to graders’ ability to identify the stage of ROP, many videos in this study were judged to have poor image quality. As this was a retrospective study, the videos evaluated were not captured for the purposes of this study. Thus, the quality of videos included in this study likely underestimates the quality obtainable if the acquisition of videos was purely for ROP “screening” (rather than the clinical teaching) purposes. Also, all videos were obtained by two pediatric ophthalmologists who are experienced ROP examiners. A prospective study evaluating the ability of ophthalmologists with varying degrees of expertise to acquire videos of adequate quality for ROP screening to include all screened infants, with a protocol in place for reimaging those with poor quality video images within a reasonable time frame is needed. Also, our study only evaluated images acquired from one eye of an infant on one

examination date. As ROP can have an asymmetric presentation in the same individual, we feel it is important that each infant have both eyes examined in a true “screening” scenario.

The limited number of ophthalmologists trained and willing to screen for ROP underscores the need for a true ROP “screening test” to help decrease the burden of screening on these experts, and hopefully thereby also to increase access to screening for infants at risk of ROP. Because high-quality video images of the retina captured by the Keeler system can be graded for the presence of type 1 ROP with high accuracy, less experienced ophthalmologists able to use the indirect ophthalmoscope to capture high-quality video images, but less confident in their diagnosis of ROP, could help screen for type 1 ROP with the guidance of ROP experts from a distance. The ROP experts could evaluate videos captured by non-experts to rule out an urgent need for a bedside examination by the expert without traveling to the bedside, which could decrease the amount of time an expert ROP examiner spends performing ROP examinations (both at the bedside and from a distance). The results of this and a previous study⁴ suggest that the Keeler or comparable system may be suitable not only for remote ROP screening, but also to help educate ophthalmologists in the nuances of ROP evaluation, especially in the identification of stage of ROP and the presence of pre-plus or plus disease.

In conclusion, high-quality video images of the retina obtained by the Keeler system can be read with high sensitivity and specificity to screen for type 1 ROP. The Keeler system holds promise as a tool for ROP screening and teaching.

Acknowledgments

Financial Support:

Dr. Prakalapakorn is supported by NIH K23EY024268 and an unrestricted grant from Research to Prevent Blindness to the Duke Eye Center. The funding organizations had no role in the design or conduct of this research or the decision to submit this report for publication.

References

1. Gilbert C. Retinopathy of prematurity: a global perspective of the epidemics, population of babies at risk and implications for control. *Early human development*. 2008; 84:77–82. [PubMed: 18234457]
2. Kong L, Fry M, Al-Samarraie M, Gilbert C, Steinkuller PG. An update on progress and the changing epidemiology of causes of childhood blindness worldwide. *J AAPOS*. 2012; 16:501–7. [PubMed: 23237744]
3. Kemper AR, Wallace DK. Neonatologists’ practices and experiences in arranging retinopathy of prematurity screening services. *Pediatrics*. 2007; 120:527–31. [PubMed: 17766525]
4. Prakalapakorn SG, Freedman SF, Wallace DK. Evaluation of an indirect ophthalmoscopy digital photographic system as a retinopathy of prematurity screening tool. *J AAPOS*. 2014; 18:36–41. [PubMed: 24568980]
5. Screening examination of premature infants for retinopathy of prematurity. *Pediatrics*. 2006; 117:572–6. [PubMed: 16452383]
6. Early Treatment For Retinopathy Of Prematurity Cooperative G. Revised indications for the treatment of retinopathy of prematurity: results of the early treatment for retinopathy of prematurity randomized trial. *Arch Ophthalmol*. 2003; 121:1684–94. [PubMed: 14662586]
7. The International Classification of Retinopathy of Prematurity revisited. *Arch Ophthalmol*. 2005; 123:991–9. [PubMed: 16009843]

8. Capone A Jr, Ells AL, Fielder AR, et al. Standard image of plus disease in retinopathy of prematurity. *Arch Ophthalmol*. 2006; 124:1669–70. [PubMed: 17102031]
9. Multicenter trial of cryotherapy for retinopathy of prematurity. Preliminary results. Cryotherapy for Retinopathy of Prematurity Cooperative Group. *Arch Ophthalmol*. 1988; 106:471–9. [PubMed: 2895630]
10. 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]
11. Prakalapakorn SG, Wallace DK, Freedman SF. Retinal imaging in premature infants using the Pictor noncontact digital camera. *J AAPOS*. 2014; 18:321–6. [PubMed: 25173892]

Figure 1.
Representative Keeler images of (a) stage 1, (b) stage 2, and (c) stage 3 retinopathy of prematurity.

Table 1

Retinopathy of Prematurity (ROP) Diagnosis by Indirect Ophthalmoscopy and Image Quality (IQ) of Keeler Videos

	Diagnosis by Indirect Ophthalmoscopy of all images (n=114)	Images judged by expert in ROP screening to have fair or good IQ* (n=68)	Images judged by a non-expert in ROP screening to have fair or good IQ* (n=63)
Zone	No. (% of all images)	No. (% of those with ROP in that Zone)	
I	16 (14)	7/16 (44)	11/16 (69)
II	84 (74)	56/84 (67)	44/84 (52)
III	14 (12)	5/14 (36)	8/14 (57)
Stage	No. (% of all images)	No. (% of those with ROP of that Stage)	
None	44 (39)	18/44 (41)	18/44 (41)
1	17 (15)	8/17 (47)	10/17 (59)
2	23 (20)	15/23 (65)	12/23 (52)
3	30 (26)	27/30 (90)	23/30 (77)
Posterior Pole disease	No. (% of all images)	No. (% of those with that Degree of posterior pole disease)	
None	87 (76)	44/87 (51)	42/87 (48)
Pre-plus	17 (15)	14/17 (82)	11/17 (65)
Plus	10 (9)	10/10 (100)	10/10 (100)
Total	114	68/114 (60)	63/114 (55)

* Image quality graded based on the ability of the grader to determine the stage of ROP in the video selected from one examination date for each infant. "Good" was defined as a video in which the grader could easily discern the stage of ROP; "fair" as a video in which it was difficult to clearly discern the stage of ROP; and "poor" as a video in which the grader was unable to discern the stage of ROP.

Table 2

Keeler Video Indirect Ophthalmoscopy Images considered to be of Fair or Good and image quality* by an Expert in Retinopathy of Prematurity (ROP) Screening (n=68)

Diagnosis by Indirect Ophthalmoscopy					
Grader	Zone [†]	I	II	III	
Expert in ROP Screening	I	3	1	0	
	II	4	47	4	
	III	0	8	1	
Non-expert in ROP Screening	I	5	9	0	
	II	2	45	5	
	III	0	2	0	
	Stage [‡]	None	1	2	3
Expert in ROP Screening	None	16	2	2	0
	1	2	6	2	0
	2	0	0	11	8
	3	0	0	0	18
Non-expert in ROP Screening	None	13	4	1	1
	1	4	4	6	7
	2	0	0	7	16
	3	0	0	0	3
	Posterior Pole disease [§]	None	Pre-plus	Plus	
Expert in ROP Screening	None	40	4	0	
	Pre-plus	4	7	3	
	Plus	0	3	7	
Non-expert in ROP Screening	None	41	5	0	
	Pre-plus	3	8	5	
	Plus	0	1	5	

* Image quality graded based on the ability of the grader to determine the stage of ROP in the video selected from one examination date for each infant. "Good" was defined as a video in which the grader could easily discern the stage of ROP; "fair" as a video in which it was difficult to clearly discern the stage of ROP; and "poor" as a video in which the grader was unable to discern the stage of ROP.

[†] The expert and non-expert graders misjudged zone I as zone II disease in 6% and 3% of all videos, respectively, zone II as zone I disease in 1% and 13% of all videos, respectively, zone II as zone III disease in 12% and 3% of all videos, respectively, and zone III as zone II disease in 6% and 7% of all videos, respectively.

[‡] The expert misidentified stage 3 as stage 2 ROP in 12% and the non-expert misidentified stage 3 disease as no ROP, stage 1 or stage 2 ROP in 1%, 10%, and 24% of all videos, respectively.

[§] The expert and non-expert graders misgraded plus as pre-plus disease in 4% and 7% of all videos, respectively, and neither grader misjudged the videos as having no posterior pole disease.

Table 3

Accuracy of identifying Type 1 Retinopathy of Prematurity (ROP) by grading Keeler video images of the retina for the presence of prethreshold disease,[†] pre-plus or plus disease compared to the reference standard[‡]

Grader	Presence of prethreshold disease, [†] pre-plus or plus disease	Reference Standard [‡]	
		Type 1 ROP	No Type 1 ROP
Expert in ROP screening	Present	12	14
	Absent	0	42
Non-expert in ROP screening	Present	12	12
	Absent	0	44
Total: No. (%)		12 (18)	56 (82)

[†]Prethreshold disease is the presence of any ROP in zone I, stage 2 in zone II with plus disease, stage 3 in zone II.¹⁰

[‡]Diagnoses of Type 1 ROP by indirect ophthalmoscopy on the clinical examination

Table 4

Accuracy of identifying Type 1 Retinopathy of Prematurity (ROP) by grading the presence of absence of posterior pole disease on Keeler video images of the retina compared to the reference standard*

Grader		Reference Standard*	
		Type 1 ROP	No Type 1 ROP
Expert in ROP screening	Plus [†]	8	2
	Pre-plus [‡]	3	11
	Normal	1	43
Non-expert in ROP Screening	Plus [†]	6	0
	Pre-plus [‡]	6	10
	Normal	0	46
Total: No. (%)		12 (18%)	56 (82%)

* Diagnoses of Type 1 ROP by indirect ophthalmoscopy on the clinical examination

[†] Image(s) in which there was at least 2 quadrants of the eye with sufficient vascular dilation and tortuosity present as compared to a standard photograph.^{7,8}

[‡] Image(s) in which there was “vascular abnormalities of the posterior pole that are insufficient for the diagnosis of plus disease but that demonstrate more arterial tortuosity and more venous dilatation than normal”.⁷