Published in final edited form as:

Ophthalmol Retina. 2017 ; 1(4): 347–354. doi:10.1016/j.oret.2016.12.005.

Intereye Agreement of Retinopathy of Prematurity from Image Evaluation in the Telemedicine Approaches to Evaluating of Acute-Phase ROP (e-ROP) Study

Gui-shuang Ying¹, Wei Pan¹, Graham E Quinn², Ebenezer Daniel¹, Michael X Repka³, and Agnieszka Baumritter² for the e-ROP Cooperative Group^{*}

¹Department of Ophthalmology, Scheie Eye Institute, University of Pennsylvania, Perelman School of Medicine, Philadelphia, Pennsylvania ²Division of Pediatric Ophthalmology, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania ³Wilmer Eye institute, Johns Hopkins University School of Medicine, Baltimore, Maryland

Abstract

Purpose—To determine the symmetry on retinal image grading of fellow eyes for retinopathy of prematurity (ROP) features (stage, zone and plus disease) and severity, to provide the basis for the within subject comparison for ROP trials.

Design—Secondary analyses of data from the Telemedicine Approaches to Evaluating of Acute-Phase ROP (e-ROP) Study.

Subjects—Infants with birth weight less than 1251g.

Methods—Infants underwent serial retinal imaging sessions in both eyes by certified imagers starting at 32-weeks postmenstrual age (PMA). Two trained non-physician readers graded each eye independently for ROP features in a 5 retinal-image set from each session. Discrepancies were adjudicated by a reading center supervisor. Readers were masked to all eye examination results, previous gradings of both eyes, current grading of the fellow eye, and demographic data.

Main Outcome Measures—The inter-eye agreement assessed using percent exact agreement and weighted kappa (K_w) for stage, zone, plus, referral-warranted ROP (RW-ROP, defined as presence of stage 3 or above, plus disease, or zone I ROP), and severity of ROP.

Corresponding Author: Gui-shuang Ying, PhD, 215-615-1514 (Phone), 215-615-1531 (Fax) gsyng@mail.med.upenn.edu 3535 Market Street, Suite 700, Philadelphia PA 19104.

^{*}A listing of the members of the e-ROP Cooperative Group is in the Appendix.

Reprints requests to: Gui-shuang Ying, PhD, University of Pennsylvania, 3535 Market Street, Suite 700, Philadelphia, PA 19104-3309

Part of this material was presented at the ARVO 2016 Annual Meeting.

This manuscript contains online-only material. The following should appear online-only: Tables 2, 4, 6, 8, 10 and the list for members of e-ROP Cooperative Group

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.

Results—Among 3918 image sessions in 1235 infants, the percent agreement (K_w) between paired eyes was 75.3% (0.65) for stage of ROP, 82.3% (0.68) for zone of ROP, 78.7% (0.51) for plus disease, 84.7% (0.56) for RW-ROP, and 72.7% (0.63) for severity of ROP. Similar inter-eye agreements were found when considering ROP features at the first image session, at the last image session, at any image session. Based on image evaluations from all sessions, 412 (33.4%) infants had ROP stage 3 or above, 148 (12.0%) had zone I ROP, 70 (5.7%) had plus disease, and 419 (33.9%) had RW-ROP in one or both eyes; symmetrical findings were present in 71.4% for ROP stage 3 or above, 56.8% for zone I ROP, 50.0% for plus disease, and 73.7% for RW-ROP.

Conclusions—Masked image evaluations by trained readers showed good inter-eye agreement in ROP characteristics, consistent with the high inter-eye agreement in ROP from clinical examinations by ophthalmologists in other studies. These data confirm that acute ROP is typically symmetrical and supports within subject comparison in ROP trials.

INTRODUCTION

Retinopathy of prematurity (ROP) is the leading cause of preventable childhood blindness. Though timely detection and treatment of potentially blinding ROP can be difficult in some areas of the US and Canada, ROP blindness is being found at epidemic rates in countries with rapidly developing neonatal care systems.^{1–3} The standard for detecting, staging and monitoring ROP is a dilated examination of the retina performed by an ophthalmologist skilled in binocular indirect ophthalmoscopy and recognition of ROP.⁴ But because fewer than 10% of infants examined require treatment, this is an inefficient use of skilled and costly personnel.^{5,6} Furthermore, there is some variation among clinicians in the evaluation of ROP using indirect ophthalmoscopy.⁷ Use of digital retinal images and telemedicine may offer an alternative solution to the identification of the small percentage of infants with referral-warranted ROP (RW-ROP, defined as presence of stage 3 or above, or plus disease, or zone I ROP) who need ophthalmology examinations to consider possible treatment.⁸ Telemedicine approaches to ROP have been evaluated in United States and other countries, and showed promise for use in the delivery of health care for premature infants.^{9–14}

The high sensitivity (90%) and specificity (87%) of trained readers' image evaluation of RW-ROP in the large multi-center Telemedicine Approaches to Evaluating of Acute-Phase ROP (e-ROP) Study¹⁵ suggests utility of remote evaluation of retinal image for clinical management of ROP, for studying the ROP natural history or for conducting ROP treatment trials. However, further understanding of image evaluation findings is needed, including the symmetry of ROP features (stage, zone, and plus disease) in eyes of premature infants. High symmetry of ROP in retinal images will provide strong support for the within subject comparison in ROP trials. The symmetry of ROP in image evaluations also could help improve the image evaluation protocol, especially on deciding whether grading of images from both eyes instead of just one eye per infant is necessary. In the Cryotherapy for Retinopathy of Prematurity (CRYO-ROP) Study, ROP features evaluated by ophthalmologists using indirect ophthalmoscopy were highly correlated. For example, when ROP status in fellow eyes was categorized as no ROP, less than prethreshold ROP, prethreshold ROP, threshold ROP, or unable to determine, the ROP agreement between fellow eyes ranged from 82% to 89% for exams at 32 to 42 weeks' postmenstrual age.¹⁶

However, this high inter-eye symmetry may be related to observer bias because the same ophthalmologist examined both eyes at the same visit. In the e-ROP study, trained readers graded ROP in images from both eyes independently, masked to all respective eye examination results, previous grading of both eyes, the current grading of the fellow eye, and the demographic data. These masked image grading data provide a unique opportunity to evaluate the symmetry of ROP features in retinal images in an unbiased manner.

To provide support for the within subject comparison in ROP clinical trials and the optimal implementation of telemedicine in ROP, this study is to determine the symmetry of ROP features (stage, zone and plus disease) and ROP severity between fellow eyes in the evaluation of retinal images from at-risk infants.

METHODS

This is a secondary analysis of data from the e-ROP Study, a multi-center observational cohort study to evaluate the validity, reliability, feasibility, safety and relative cost-effectiveness of a digital imaging system to identify infants with RW-ROP. The details of the design, methods, and primary results for the e-ROP Study have been published.^{15;17;18} Major features of the study and image evaluation procedure are described here.

Study Subjects

The study enrolled 1284 infants with birth weight (BW) less than 1251g from 13 North American centers from May 25, 2011, through October 31, 2013. Informed consent was obtained from the parent/guardian of eligible infants. The study protocol and informed consent process were approved by institutional review boards of participating centers.

Study Procedure

Infants underwent serial ROP imaging in both eyes using the RetCam Shuttle® (Clarity Medical Systems, Pleasanton, CA).¹⁹ From each eye, a trained imager obtained an image set that included an external image to assess pupillary dilation, along with five retinal views: disc center and four disc off-center views showing the inferior, superior, temporal and nasal retina. Imaging sessions began at 32 weeks postmenstrual age (PMA), and follow-up imaging was conducted at least every other week unless medically contraindicated until the ophthalmologist noted: mature retinal vessels, immature zone III on two occasions at least 7 days apart, ROP regressed or regressing on two visits at least 7 days apart, treatment for severe ROP, or the infant reached 40 weeks PMA with no ROP or only stage 1 or 2 ROP.^{19;20}

Image Evaluation

Image sets taken from the each clinical center were uploaded by a secure web-based portal for remote evaluation by three non-physician trained readers in the image reading center at the University of Pennsylvania. Each image set was independently graded by two trained readers following a standard grading protocol, and discrepancies were adjudicated by the Reading Center Supervisor.¹⁷ Readers and the adjudicator were masked to results of all

diagnostic examinations, previous gradings of both eyes, current grading of the fellow eye, and demographic data.

All readers determined for each quadrant, whether the posterior pole vessels were normal or sufficiently abnormal to be designated as pre-plus or plus disease. In addition, the reader determined zone of vascularization or the zone in which morphologic features consistent with ROP were present. ROP was determined by the presence of a demarcation line (stage 1), a ridge (stage 2), extraretinal neovascularization (stage 3), or retinal detachment (stage 4 or 5). When the image was not of sufficient quality to grade specific ROP features, it was recorded as cannot grade (CG).

The remote evaluation of image by trained readers showed good reproducibility. The percent agreement was above 90% for grading RW-ROP and its components (i.e., stage 3 or above, plus disease, zone I ROP). The weighted kappa for RW-ROP was 0.72 (95% CI: 0.52 – 0.93) for inter-reader agreement and 0.77 (95% CI: 0.57 – 0.97) for grade re-grade agreement.¹⁷

Statistical Analysis

Using data from all image gradings, we evaluated the inter-eye agreement for ROP features (ROP stage, zone, plus disease,), RW-ROP and severity of ROP (type 1, type 2, mild ROP, no ROP) using agreement statistics including percent of exact agreement and weighted Kappa (K_w). We consider inter-eye agreement slight for $K_w < 0.2$, fair for $K_w 0.21 - 0.40$, moderate for $K_w 0.41 - 0.60$, substantial for $K_w 0.61 - 0.80$, and almost perfect for $K_w > 0.80$.²¹ For assessing inter-eye agreement, we categorized ROP stage into 4 levels (No ROP, stage 1 or 2, stage 3 or above, cannot grade), ROP zone into 4 levels (No ROP, zone II ROP, zone I ROP, cannot grade), and posterior pole status into 4 levels (normal, preplus, plus, cannot grade). We also evaluated the inter-eye agreement for RW-ROP, and severity of ROP. The RW-ROP was defined as the presence of plus disease, or zone I ROP, or ROP stage 3 or greater.⁸ Severity of ROP was classified as type 1 ROP or type 2 ROP (as defined by ETROP²²), ROP not type 1 or 2 (i.e., mild ROP), no ROP. For these inter-eye agreement analyses, we considered ROP features at all image sessions combined, at the first image session, at the last image session, most severe ROP features across all image sessions and by PMA weeks. In addition, among infants with specific ROP features (ROP stage 3 or above, zone I ROP, plus disease) detected at any image session in one or both eyes, we assessed their symmetry at image session of their first detection and at all image sessions. All statistical analyses were performed in SAS (version 9.4, SAS Institute Inc., Cary, NC).

RESULTS

The e-ROP study enrolled 1284 infants, 49 (3.8%) infants were discharged prior to the first image session. Therefore, this analysis included 1235 premature infants (mean BW of 865g, mean gestation age of 27 weeks) who underwent a total of 3918 image sessions. The mean (SD) number of image sessions per infant was 3.2 (2.0). The mean PMA (SD) was 34 (1.9) weeks at the first image session, and 37 (2.6) weeks at the last image session.

Inter-eye agreement for ROP stage

Among 3918 image sessions, both eyes were judged in 2947 sessions to be symmetrical: no ROP in 1914 (49%) sessions, stage 1 or 2 ROP in 511 (13%) sessions, stage 3 or above in 510 (13%) sessions, and not gradable in 12 (0.3%) sessions. In 109 (2.8%) of image sessions, one eye had ROP stage 3 or above, while the fellow eye did not have any ROP. The percent of exact agreement was 75.3% with a K_w of 0.65 (95% CI: 0.63 – 0.67) (Table 1). Similar inter-eye agreement was found at the first image session (percent agreement = 79.4%, K_w = 0.65 (0.61 – 0.69)), and slightly higher inter-eye agreement at the last image session (percent agreement = 82.5%, K_w = 0.72 (0.69 – 0.76)). When the highest ROP stage was considered across all image sessions, 584 (47.3%) infants had no ROP in both eyes, 110 (8.9%) infants had ROP stage 1 or 2, and 294 (23.8%) infants had ROP stage 3 or above in both eyes, while image evaluations of all sessions found 15 (1.2%) infants had stage 3 or above in one eye at at least one session, but never had any ROP detected in the fellow eye. Their inter-eye agreement was 83.0% with K_w of 0.77 (0.75 – 0.80, Table 1). The inter-eye agreement of ROP stage did not differ substantially across PMA weeks (range = 32 to 42 weeks), with agreement ranging from 67% to 78%, and K_w ranging from 0.54 to 0.70 (Table 2, online).

Inter-eye agreement for ROP Zone

When the ROP location was graded as zone I ROP, zone II ROP, no ROP, and cannot grade, the percent of agreement was 82.3%, with K_w 0.68 (95% CI: 0.66 – 0.70) in all sessions analysis, and only in 11 (0.3%) image sessions, one eye had zone I ROP while their the fellow eye did have any ROP (Table 3). Similar results were found at the first image session (percent agreement 82.1% and K_w 0.64 (0.60–0.69)), and slightly higher at the last image session (86.6% and K_w 0.73 (0.69 – 0.77)). When the lowest ROP zone was considered across all image sessions, 84 (6.8%) infants had zone I ROP in both eyes and 346 (28.0%) infants had zone II ROP in both eyes, with percent agreement of 85.9% and K_w 0.75 (0.72 – 0.78) (Table 5). Only 3 (0.2%) infants had zone 1 ROP in one eye while did not have any ROP in their fellow eye. The inter-eye agreement was higher with older PMA, with percent of agreement ranges from 75.4% (K_w = 0.54) for 32 PMA weeks to 86.3% (K_w = 0.74) for 41–42 PMA weeks (Table 4, online).

Inter-eye agreement for preplus/plus disease

The posterior pole vessels were graded as normal, preplus, plus or cannot grade. Among all 3918 image sessions, both eyes had normal posterior pole vessels in 2396 (61%) image sessions, preplus in 639 (16%), and plus in 40 (1%) image sessions. In only one image session, one eye had plus disease while the fellow eye had normal posterior pole vessels. The percent of agreement was 78.7% with a K_w of 0.51 (0.48 – 0.54) (Table 5). Similar inter-eye agreement was found at the first image session (percent agreement = 83.8%, K_w = 0.54 (0.48 – 0.60)) and at the last image session (percent agreement = 78.7%, K_w = 0.57 (0.52 – 0.61)). When the worst posterior pole disease was considered across all image sessions, both eyes had normal posterior vessels in 605 (49%) infants, preplus in 282 (22.8%) infants, and plus in 35 (2.8%) infants, while none of infants had plus disease in one eye with normal posterior pole vessels in the fellow eye, with percent agreement of 77.5%

and K_w of 0.59 (0.54 – 0.63) (Table 5). The inter-eye agreement was similar across all PMA weeks (Table 6, online).

Inter-eye agreement for RW-ROP

Among all image sessions, both eyes had RW-ROP in 590 (15.1%) sessions, and only one eye had RW-ROP in 508 (13.0%) image sessions, with percent agreement of 84.7% and K_w of 0.56 (0.53 – 0.59) (Table 7). The percent agreement was 88.4% with K_w 0.60 (0.54–0.66) at the first image session, and was higher at the last image session (percent agreement 91.0% and K_w 0.70 (0.65 – 0.75)). When presence of RW-ROP in any image session was considered across all image sessions, RW-ROP was present in both eyes of 309 (25%) infants and only in one eye of 99 (8.0%) infants, with percent agreement 91.6% and K_w 0.77 (0.73 – 0.80) (Table 7). The inter-eye agreement seems consistent across PMA weeks with percent agreement approximately 82–87% for most PMA weeks and K_w 0.51 to 0.63 except lower agreement in PMA 41 to 42 weeks (Table 8, online).

Inter-eye agreement for severity of ROP

When ROP severity was classified as type 1 ROP, type 2 ROP, mild ROP, no ROP and cannot grade, the percent agreement between two eyes was 72.7% with K_w 0.63 (95% CI: 0.61–0.66). Only in 2 (0.05%) image sessions, one eye had type 1 ROP while the fellow eye did not have any ROP when all image sessions were considered (Table 9). The percent agreement was 77% with K_w 0.64 (0.60 – 0.68) at the first image session, and 81% with K_w 0.72 (0.69 – 0.76) at the last session. When the worst ROP was considered across all image sessions, 80 (6.5%) infants had type 1 ROP in both eyes, 62 (5.0%) infants had type 1 ROP only in one eye (their fellow eye status is type 2 ROP in 55 eyes, mild ROP in 6 eyes and no ROP in one eye), and 153 (12.4%) had type 2 ROP in both eyes, with the percent agreement of 79% and K_w 0.75 (0.72 – 0.78) (Table 9). The inter-eye agreement seems consistent across PMA weeks for both percent agreement (64% to 75%) and K_w (0.54 to 0.69) (Table 10, online).

Fellow eye status when RW-ROP or its components were first detected in an image session

The symmetry of RW-ROP and its components at the image session of its first detection and across all image sessions is shown in Table 11. Among 419 infants detected with RW-ROP in one or both eyes from image evaluation at mean PMA of 35 weeks, 197 (47.0%) were bilateral at its first image detection, 112 (26.7%) were unilateral at its first detection, but became bilateral at a later image session with mean interval of 15 days. For each RW-ROP components at its first detection in images, the percent of bilateral was 44% for stage 3 or above ROP, 41% for Zone 1 ROP and 40% for plus disease.

DISCUSSION

We evaluated the inter-eye agreement of ROP classification features and severity of ROP from masked non-physician trained reader's gradings of retinal images taken during the acute phase of ROP. We found substantial inter-eye agreement (weighted kappa >0.60). For the key determination of the presence of RW-ROP, the agreement was 85%. An earlier study

of inter-eye correlation of ROP findings from study-certified ophthalmologist eye examinations in the CRYO-ROP study found a high degree of concordance in the severity and location of ROP, presence of plus disease and sectors of stage 3 ROP in the two eyes of infants.¹⁶ In particular, their overall percent of agreement for ROP severity (defined as no ROP, less than prethreshold, prethreshold, threshold, unable to determine) from PMA 32 to 42 weeks was 85%, which is higher than our agreement of 73% for ROP severity (defined as no ROP, mild ROP, type 2 ROP, type 1 ROP, can not grade). The higher correlation for the clinical examination could be due to the lack of masking of the examiner from the fellow eye findings from the current visit and exam findings from previous visits. The lower inter-eye agreement for retinal image gradings could be due to differences in the quality of the images (both focus and fields available for evaluation) from the two eyes, the mis-interpretation of images from one eye, inherent inter-reader and grade-regrade variations of grading images by trained readers (as images from two eyes of an image session were not necessarily assigned to the same readers). In spite of the limitations of each approach, the substantial inter-eye agreement in fellow eyes of an infant supports the hypothesis that development of acute phase of ROP is mainly determined by the systemic factors (such as prematurity, genetic factors) rather than ocular factors.²³

The high symmetry of ROP found in image evaluation from this study and in clinical eye exams from a previous study¹⁶ provides strong support for the within-subject design for ROP trials, where two eyes of an infant are randomized with one eye for treatment and fellow eye for control, and ROP outcomes (evaluated either by images or by exam) are compared between treated eye to control eye of the same infant. Such design is efficient in that it requires smaller sample size than the between-subject comparison trials, and also minimizes the confoundings from infant-level factors.

The burden for undertaking timely screening of at-risk infants for ROP is large, particularly in developing countries.^{13;24} The effectiveness of ROP screening for large number of infants might be relieved by adopting a telemedicine approach with images taken by trained local imagers with remote grading done by non-physician readers. Thus, it is important to investigate approaches for improving the efficiency of telemedicine of ROP.²⁵ Given the high concordance in ROP between two eyes as reported CRYO-ROP,¹⁶ a screener could improve the efficiency of image capture and grading by taking and grading images from one eye only. Our study found that the inter-eye agreements in ROP features and severity in images are substantial but not perfect, and the findings in this study do not support grading of only one eye for ROP screening. For example, the percent of inter-eye agreement for ever developing RW-ROP across all image sessions was 92%, with 4% of infants detected with RW-ROP in left eye only and another 4% infants detected with RW-ROP in right eye only. This would mean that a ROP screening system imaging a single eye would miss the detection of RW-ROP in at least 4% of infants. It also indicates that the imagers need to assure that quality images are obtained from both eyes as often as possible.

In recent years, the use of retinal images for ROP clinical care and research has become widespread.¹³ There are many advantages in using images, including the ability to transfer images remotely for review by qualified readers to allow greater precision in ROP determination, and provide a permanent record of images to allow physician the ability to

evaluate change in the patient. Use of retinal images will likely benefit the study of the natural history, the ROP status at treatment and outcome in clinical trials, and the description of ROP progression at various time points. Our findings of substantial inter-eye agreement need to be taken into consideration when designing observational studies and clinical trials that use ROP image evaluation as an outcome measure.

In summary, masked image evaluations by trained readers showed substantial inter-eye agreement in ROP features and severity, slightly lower than reported from unmasked diagnostic examinations by ophthalmologists. This finding supports the potential value of retinal images in the clinical management and research of ROP. The high inter-eye symmetry of ROP supports the within subject comparison in ROP clinical trials.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Supported by National Eye Institute of the National Institutes of Health, Department of Health and Human Services U10 EY017014 and R21EY025686.

Reference List

1. Blencowe H, Lawn JE, Vazquez T, et al. Preterm-associated visual impairment and estimates of retinopathy of prematurity at regional and global levels for 2010. *Pediatr Res*. 2013; 74(Suppl 1): 35–49. [PubMed: 24366462]
2. Gilbert C, Fielder A, Gordillo L, et al. Characteristics of infants with severe retinopathy of prematurity in countries with low, moderate, and high levels of development: implications for screening programs. *Pediatrics*. 2005; 115:e518–e525. [PubMed: 15805336]
3. Gilbert C. Retinopathy of prematurity: a global perspective of the epidemics, population of babies at risk and implications for control. *Early Hum Dev*. 2008; 84:77–82. [PubMed: 18234457]
4. Screening examination of premature infants for retinopathy of prematurity. *Pediatrics*. 2006; 117:572–6. [PubMed: 16452383]
5. Kemper AR, Wallace DK. Neonatologists' practices and experiences in arranging retinopathy of prematurity screening services. *Pediatrics*. 2007; 120:527–31. [PubMed: 17766525]
6. Kemper AR, Wallace DK, Quinn GE. Systematic review of digital imaging screening strategies for retinopathy of prematurity. *Pediatrics*. 2008; 122:825–30. [PubMed: 18829807]
7. Reynolds JD, Dobson V, Quinn GE, et al. Evidence-based screening criteria for retinopathy of prematurity: natural history data from the CRYO-ROP and LIGHT-ROP studies. *Arch Ophthalmol*. 2002; 120:1470–6. [PubMed: 12427059]
8. Ells AL, Holmes JM, Astle WF, et al. Telemedicine approach to screening for severe retinopathy of prematurity: a pilot study. *Ophthalmology*. 2003; 110:2113–7. [PubMed: 14597517]
9. Weaver DT. Telemedicine for retinopathy of prematurity. *Curr Opin Ophthalmol*. 2013; 24:425–31. [PubMed: 23872819]
10. 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]
11. Prakash 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]

12. Fijalkowski N, Zheng LL, Henderson MT, et al. Stanford University Network for Diagnosis of Retinopathy of Prematurity (SUNDROP): four-years of screening with telemedicine. *Curr Eye Res.* 2013; 38:283–91. [PubMed: 23330739]
13. Fierson WM, Capone A Jr. Telemedicine for evaluation of retinopathy of prematurity. *Pediatrics.* 2015; 135:e238–e254. [PubMed: 25548330]
14. Athikarisamy SE, Patole S, Lam GC, et al. Screening for retinopathy of prematurity (ROP) using wide-angle digital retinal photography by non-ophthalmologists: a systematic review. *Br J Ophthalmol.* 2015; 99:281–8. [PubMed: 25001320]
15. Quinn GE, Ying GS, Daniel E, et al. Validity of a telemedicine system for the evaluation of acute-phase retinopathy of prematurity. *JAMA Ophthalmol.* 2014; 132:1178–84. [PubMed: 24970095]
16. Quinn GE, Dobson V, Biglan A, et al. Correlation of retinopathy of prematurity in fellow eyes in the cryotherapy for retinopathy of prematurity study. The Cryotherapy for Retinopathy of Prematurity Cooperative Group. *Arch Ophthalmol.* 1995; 113:469–73. [PubMed: 7710397]
17. Daniel E, Quinn GE, Hildebrand PL, et al. 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]
18. Telemedicine approaches to evaluating acute-phase retinopathy of prematurity: study design. *Ophthalmic Epidemiol.* 2014; 21:256–67. [PubMed: 24955738]
19. Quinn GE, Ying GS, Daniel E, et al. Validity of a telemedicine system for the evaluation of acute-phase retinopathy of prematurity. *JAMA Ophthalmol.* 2014; 132:1178–84. [PubMed: 24970095]
20. Telemedicine approaches to evaluating acute-phase retinopathy of prematurity: study design. *Ophthalmic Epidemiol.* 2014; 21:256–67. [PubMed: 24955738]
21. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977; 33:159–74. [PubMed: 843571]
22. The International Classification of Retinopathy of Prematurity revisited. *Arch Ophthalmol.* 2005; 123:991–9. [PubMed: 16009843]
23. Hellstrom A, Smith LE, Dammann O. Retinopathy of prematurity. *Lancet.* 2013; 382:1445–57. [PubMed: 23782686]
24. Vinekar A, Jayadev C, Mangalesh S, et al. Role of tele-medicine in retinopathy of prematurity screening in rural outreach centers in India - a report of 20,214 imaging sessions in the KIDROP program. *Semin Fetal Neonatal Med.* 2015; 20:335–45. [PubMed: 26092301]
25. Fierson WM. Screening examination of premature infants for retinopathy of prematurity. *Pediatrics.* 2013; 131:189–95. [PubMed: 23277315]

Symmetry of ROP stage between 2 eyes at all image sessions (N=3918 image sessions) and symmetry of ROP stage across all image sessions (N=1235 infants)

Table 1

ROP stage of left eye at an image session	ROP stage of right eye at an image session			
	No ROP	Stage 1 or 2	Stage 3 or above	Total
No ROP	1914 (48.9%)	146 (3.7%)	46 (1.2%)	2130 (54.4%)
Stage 1 or 2	225 (5.7%)	515 (13.1%)	205 (5.2%)	953 (24.3%)
Stage 3 or above	63 (1.6%)	208 (5.3%)	510 (13.0%)	792 (20.2%)
Cannot grade	20 (0.5%)	6 (0.2%)	5 (0.1%)	43 (1.1%)
Total	2222 (56.7%)	875 (22.3%)	766 (19.6%)	3918
Percent agreement = 75.3% Weighted kappa (95% CI) =0.65 (0.63, 0.67)				
Highest ROP stage of left eye of across all image sessions	Highest ROP stage of right eye of across all image sessions			
	No ROP	Stage 1 or 2	Stage 3 or above	Total
No ROP	584 (47.3%)	30 (2.4%)	6 (0.5%)	626 (50.7%)
Stage 1 or 2	65 (5.3%)	110 (8.9%)	41 (3.3%)	221 (17.9%)
Stage 3 or above	9 (0.7%)	52 (4.2%)	294 (23.8%)	360 (29.1%)
Cannot grade	13 (1.1%)	3 (0.2%)	5 (0.4%)	28 (2.3%)
Total	671 (54.3%)	195 (15.8%)	346 (28.0%)	1235
Percent agreement = 83.0% Weighted kappa (95% CI) =0.77 (0.75, 0.80)				

Symmetry of ROP zone between 2 eyes at all image sessions (N=3918 image sessions) and lowest ROP zone across all image sessions (N=1235 infants)

Table 3

ROP zone of left eye at an image session	ROP zone of right eye at an image session			
	No ROP	Zone II ROP	Zone I ROP	Total
No ROP	1914 (48.9%)	186 (4.7%)	9 (0.2%)	2130 (54.4%)
Zone II ROP	288 (7.4%)	1170 (29.9%)	110 (2.8%)	1577 (40.3%)
Zone I ROP	2 (0.1%)	41 (1.0%)	128 (3.3%)	174 (4.4%)
Cannot grade	18 (0.5%)	8 (0.2%)	0 (0.0%)	37 (0.9%)
Total	2222 (56.7%)	1405 (35.9%)	247 (6.3%)	3918
Percent agreement = 82.3% Weighted kappa (95% CI) = 0.68 (0.66, 0.70)				
Lowest ROP zone of left eye across all image sessions	Lowest ROP zone of right eye across all image sessions			
	No ROP	Zone II ROP	Zone I ROP	Total
No ROP	584 (47.3%)	35 (2.8%)	1 (0.1%)	626 (50.7%)
Zone II ROP	73 (5.9%)	346 (28.0%)	37 (3.0%)	471 (38.1%)
Zone I ROP	2 (0.2%)	19 (1.5%)	84 (6.8%)	107 (8.7%)
Can not grade	12 (1.0%)	7 (0.6%)	3 (0.2%)	31 (2.5%)
Total	671 (54.3%)	407 (33.0%)	125 (10.1%)	1235
Percent agreement = 85.9% Weighted kappa (95% CI) = 0.75 (0.72, 0.78)				

Symmetry of preplus/plus between 2 eyes at all image sessions (N=3918 image sessions) and worst preplus/plus across all image sessions (N=1235 infants)

Table 5

		Preplus/plus of right Eye at an image session				
Preplus/plus of left eye at an image session		Normal	Preplus	Plus	Cannot grade	Total
Normal		2396 (61.2%)	369 (9.4%)	1 (0.0%)	15 (0.4%)	2781 (71.0%)
Preplus		362 (9.2%)	639 (16.3%)	22 (0.6%)	7 (0.2%)	1030 (26.3%)
Plus		0 (0.0%)	30 (0.8%)	40 (1.0%)	1 (0.0%)	71 (1.8%)
Cannot grade		16 (0.4%)	12 (0.3%)	0 (0.0%)	8 (0.2%)	36 (0.9%)
Total		2774 (70.8%)	1050 (26.8%)	63 (1.6%)	31 (0.8%)	3918
Percent agreement = 78.7% Weighted kappa (95% CI) =0.51 (0.48, 0.54)						
		Worst preplus/plus of right eye across all image sessions				
Worst preplus/plus of left eye across all image sessions		Normal	Preplus	Plus	Cannot grade	Total
Normal		605 (49.0%)	115 (9.3%)	0 (0.0%)	6 (0.5%)	726 (58.8%)
Preplus		121 (9.8%)	282 (22.8%)	12 (1.0%)	8 (0.6%)	423 (34.3%)
Plus		0 (0.0%)	20 (1.6%)	35 (2.8%)	1 (0.1%)	56 (4.5%)
Cannot grade		7 (0.6%)	12 (1.0%)	2 (0.2%)	9 (0.7%)	30 (2.4%)
Total		733 (59.4%)	429 (34.7%)	49 (4.0%)	24 (1.9%)	1235
Percent agreement = 77.5% Weighted kappa (95% CI) =0.59 (0.54, 0.63)						

Table 7

The symmetry of RW-ROP between 2 eyes at all image sessions (N=3918 image sessions) and across all image sessions (N=1235 infants)

	RW-ROP of right eye at an image session			
RW-ROP of left eye at an image session	No RW-ROP	RW-ROP	Cannot grade	Total
No RW-ROP	2713 (69.2%)	262 (6.7%)	37 (0.9%)	3012 (76.9%)
RW-ROP	246 (6.3%)	590 (15.1%)	12 (0.3%)	848 (21.6%)
Cannot grade	36 (0.9%)	7 (0.2%)	15 (0.4%)	58 (1.5%)
Total	2995 (76.4%)	859 (21.9%)	64 (1.6%)	3918
Percent agreement = 84.7% Weighted kappa (95% CI) =0.56 (0.53, 0.59)				
RW-ROP of left eye across all image session	RW-ROP of right eye across all image sessions			
No RW-ROP	772 (62.5%)	44 (3.6%)	14 (1.1%)	830 (67.2%)
RW-ROP	55 (4.5%)	309 (25.0%)	6 (0.5%)	370 (30.0%)
Cannot grade	21 (1.7%)	5 (0.4%)	9 (0.7%)	35 (2.8%)
Total	848 (68.7%)	358 (29.0%)	29 (2.3%)	1235
Percent agreement = 91.6% Weighted kappa (95% CI) =0.77 (0.73, 0.80)				

Table 9

The symmetry of severity of ROP between 2 eyes at all image sessions (N=3918 image sessions) and across all image sessions (N=1235 infants)

Severity of ROP at left eye at an image session	Severity of ROP of right eye at an image session					
	No ROP	Mild ROP	Type 2 ROP	Type 1 ROP	Cannot grade	Total
No ROP	1914 (48.9%)	137 (3.5%)	55 (1.4%)	0 (0.0%)	24 (0.6%)	2130 (54.4%)
Mild ROP	221 (5.6%)	456 (11.6%)	182 (4.6%)	26 (0.7%)	9 (0.2%)	894 (22.8%)
Type 2 ROP	62 (1.6%)	173 (4.4%)	368 (9.4%)	61 (1.6%)	10 (0.3%)	674 (17.2%)
Type 1 ROP	2 (0.1%)	10 (0.3%)	58 (1.5%)	98 (2.5%)	3 (0.1%)	171 (4.4%)
Cannot grade	23 (0.6%)	6 (0.2%)	6 (0.2%)	1 (0.0%)	13 (0.3%)	49 (1.3%)
Total	2222 (56.7%)	782 (20.0%)	669 (17.1%)	186 (4.7%)	59 (1.5%)	3918
Percent agreement = 72.7% Weighted kappa (95% CI) = 0.63 (0.61, 0.66)						
Most severe ROP of left eye across all image sessions	Most severe ROP of right eye across all image sessions					
	No ROP	Mild ROP	Type 2 ROP	Type 1 ROP	grade	Total
No ROP	584 (47.3%)	29 (2.3%)	7 (0.6%)	0 (0.0%)	6 (0.5%)	626 (50.7%)
Mild ROP	64 (5.2%)	104 (8.4%)	30 (2.4%)	5 (0.4%)	6 (0.5%)	209 (16.9%)
Type 2 ROP	7 (0.6%)	43 (3.5%)	153 (12.4%)	28 (2.3%)	17 (1.4%)	248 (20.1%)
Type 1 ROP	1 (0.1%)	1 (0.1%)	27 (2.2%)	80 (6.5%)	3 (0.2%)	112 (9.1%)
Can not grade	15 (1.2%)	5 (0.4%)	5 (0.4%)	2 (0.2%)	13 (1.1%)	40 (3.2%)
Total	671 (54.3%)	182 (14.7%)	222 (18.0%)	115 (9.3%)	45 (3.6%)	1235
Percent agreement = 79.2% Weighted kappa (95% CI) = 0.75 (0.72, 0.78)						

Table 11

The symmetry of RW-ROP and its components at the image session of its first detection and across all image sessions

	RW-ROP	Stage 3 or above ROP	Zone 1 ROP	Plus disease
Number of Infants with specific ROP in one or both eyes	419	412	148	70
Mean (SD) of PMA weeks at its first detection in images	35.0 (2.3)	35.2 (2.2)	33.9 (1.7)	36.7 (2.3)
Unilateral (%)	110 (26.3%)	118 (28.6%)	64 (43.2%)	35 (50.0%)
Bilateral (%)	309 (73.7%)	294 (71.4%)	84 (56.8%)	35 (50.0%)
At the same image session (%)	197 (47.0%)	179 (43.5%)	61 (41.2%)	28 (40.0%)
Not at the same image session	112 (26.7%)	115 (27.9%)	23 (15.5%)	7 (10.0%)
Interval among those bilateral not detected at the same image session (days) *				
Mean (SD)	15.0 (9.5)	14.3 (8.9)	10.1 (6.2)	7.1 (3.4)

* The interval was calculated days between date when RW-ROP or its components was first detected in one eye, and the date of subsequent first detection in their fellow eye.

Masked image evaluations by trained readers showed good inter-eye agreement in ROP stage, zone and plus disease, supporting the within-subject design for ROP clinical trials.