tvst

Pediatric Ophthalmology & Strabismus

Multivariate Models to Diagnose Early Referral-Warranted Retinopathy of Prematurity With Handheld Optical Coherence Tomography

Alex T. Legocki[1], Aaron Y. Lee[1,2], Leona Ding[1], Yasman Moshiri[3], Emily M. Zepeda[4], Thomas B. Gillette[5], Laura E. Grant[6], Ayesha Shariff[7], Phanith Touch[1], Cecilia S. Lee[1,2], Kristina Tarczy-Hornoch[1,8], and Michelle T. Cabrera[1,8]

[1] Department of Ophthalmology, University of Washington, Seattle, WA, USA
[2] The Roger and Angie Karalis Johnson Retina Center, Seattle, WA, USA
[3] Department of Ophthalmology, Bascom Palmer Eye Institute, University of Miami, Miami, FL, USA
[4] Department of Ophthalmology, Dean McGee Eye Institute, University of Oklahoma, Oklahoma City, OK, USA
[5] Department of Ophthalmology, University of South Florida Eye Institute, Tampa, FL, USA
[6] Department of Ophthalmology, Millman-Derr Center for Eye Care, Rochester Hills, MI, USA
[7] Department of Ophthalmology, New Mexico Veterans Affairs Medical Center, University of New Mexico, Albuquerque, NM, USA
[8] Division of Ophthalmology, Seattle Children's Hospital, Seattle, WA, USA

Correspondence: Michelle T. Cabrera, OA.9.220, 4800 Sand Point Way NE, Seattle, WA 98105, USA.
e-mail: cabreram@uw.edu

Received: December 29, 2022
Accepted: April 23, 2023
Published: May 24, 2023

Keywords: retinopathy of prematurity (ROP); referral-warranted retinopathy of prematurity; retina; optical coherence tomography (OCT); pediatric ophthalmology

Citation: Legocki AT, Lee AY, Ding L, Moshiri Y, Zepeda EM, Gillette TB, Grant LE, Shariff A, Touch P, Lee CS, Tarczy-Hornoch K, Cabrera MT. Multivariate models to diagnose early referral-warranted retinopathy of prematurity with handheld optical coherence tomography. Transl Vis Sci Technol. 2023;12(5):26, https://doi.org/10.1167/tvst.12.5.26

Purpose: The purpose of this study was to create multivariate models predicting early referral-warranted retinopathy of prematurity (ROP) using non-contact handheld spectral-domain optical coherence tomography (OCT) and demographic data.

Methods: Between July 2015 and February 2018, infants $\leq$1500 grams birth weight or $\leq$30 weeks gestational age from 2 academic neonatal intensive care units were eligible for this study. Infants were excluded if they were too unstable to participate in ophthalmologic examination (2), had inadequate image quality (20), or received prior ROP treatment (2). Multivariate models were created using demographic variables and imaging findings to identify early referral-warranted ROP (referral-warranted ROP and/or pre-plus disease) by routine indirect ophthalmoscopy.

Results: A total of 167 imaging sessions of 71 infants (45% male infants, gestational age 28.2+/−2.8 weeks, and birth weight 995.6+/−292.0 grams) were included. Twelve of 71 infants (17%) developed early referral-warranted ROP. The area under the receiver operating characteristic curve (AUC) was 0.94 for the generalized linear mixed model (sensitivity = 95.5% and specificity = 80.7%) and 0.83 for the machine learning model (sensitivity = 91.7% and specificity = 77.8%). The strongest variables in both models were birth weight, image-based Vitreous Opacity Ratio (an estimate of opacity density), vessel elevation, and hyporeflective vessels. A model using only birth weight and gestational age yielded an AUC of 0.68 (sensitivity = 77.3% and specificity = 63.4%), and a model using only imaging biomarkers yielded 0.88 (sensitivity = 81.8% and specificity = 84.8%).

Conclusions: A generalized linear mixed model containing handheld OCT biomarkers can identify early referral-warranted ROP. Machine learning produced a less optimal model.

Translational Relevance: With further validation, this work may lead to a better-tolerated ROP screening tool.

Introduction

Retinopathy of prematurity (ROP) is a major cause of childhood blindness worldwide, characterized by abnormal vasoproliferation in the developing retina.[1] The diagnosis of ROP is made by clinical examination in premature infants, traditionally by binocular indirect ophthalmoscopy (BIO) with use of an eyelid speculum and scleral depressor. This examination is known to cause significant agitation and discomfort for the infants, demonstrated by deviations of cardiorespiratory metrics and pain scores, and this stress is particularly associated with the speculum and scleral depressor use.[2–8]

Wide-field digital retinal imaging (e.g. Retcam, Natus Medical Incorporated, Middleton, WI, USA) is a well-established alternative to BIO, but also requires contact with the globe. Unfortunately, studies have shown that Retcam photography is also associated with significant alterations in pain scores and vital signs, in some cases greater than BIO.[2–4,9]

Handheld optical coherence tomography (OCT) is an imaging tool used to capture high-resolution cross-sectional retinal and optic nerve images in the supine position. One direction of current research is imaging of the vascular-avascular junction in the retinal periphery, which once again requires a combination of eyelid speculum placement and often scleral depression use or direct contact between the eye and OCT probe.[10–12] However, Maldonado et al. introduced a technique in 2010 to acquire handheld OCT images in awake infants without the use of an eyelid speculum, allowing the infant to remain comfortable during image

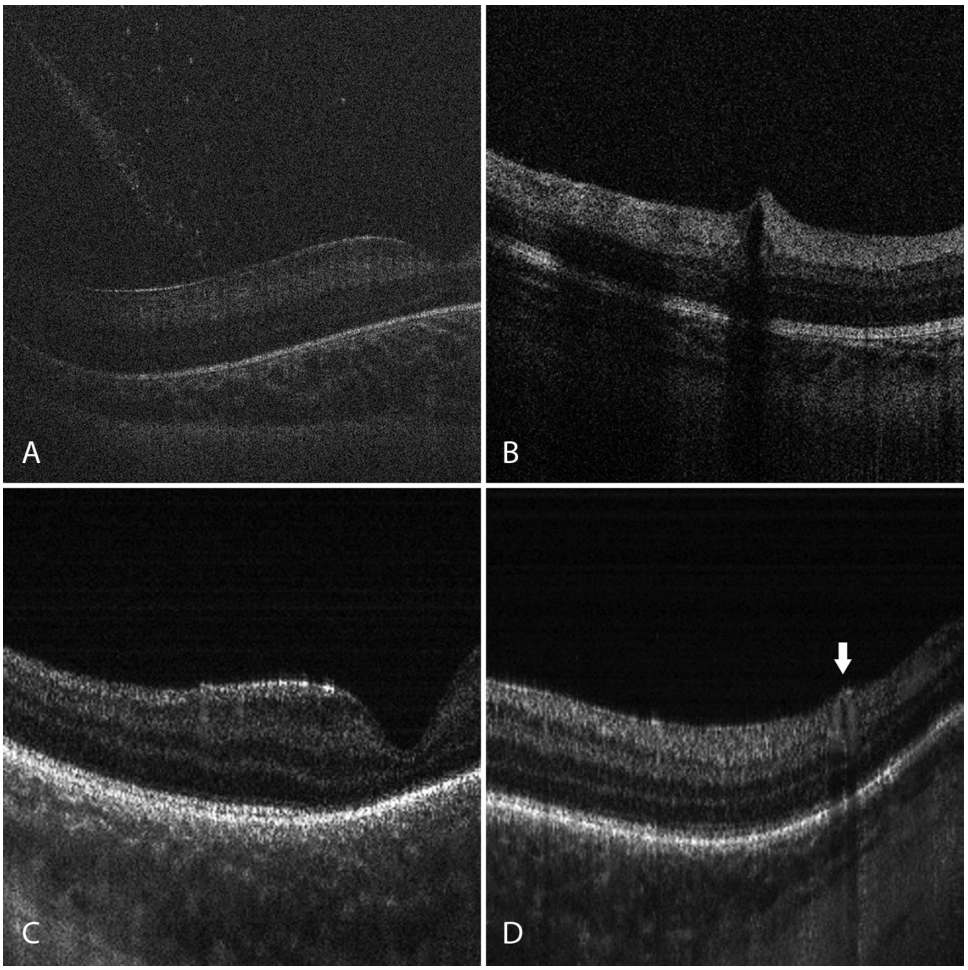

Figure 1. Optical coherence tomography volumes were reviewed for vitreoretinal findings associated with retinopathy of prematurity. These findings were combined into models to predict early referral-warranted retinopathy of prematurity. (**A**) Tractional bands were identified as straight bands at a steep angle to the retina, and punctate hyper-reflective vitreous opacities were used to calculate the vitreous opacity ratio in the five frames centered at the fovea.[16] (**B**) Vessel elevation described spiked vessels emerging from the inner retinal surface, and hyporeflective vessels were vessels with lumen hyporeflectivity. (**C**) Scalloped retinal layers of the inner and/or outer plexiform layers were identified, and (**D**) retinal spaces described hyporeflective spaces located adjacent to vessels (*white arrow*).[15]

acquisition.[13] Preterm infants have been shown to have diminished behavioral and cardiovascular manifestations of stress during such handheld OCT imaging compared to standard binocular indirect ophthalmoscopy.[8] In these cases, imaging does not capture the peripheral vascular-avascular junction and is instead limited to the posterior pole. In spite of limited field of view, recent work has associated awake handheld OCT vitreoretinal findings in these posterior areas with clinical ROP severity. These potential markers of ROP severity include elevated retinal vessels, scalloped retinal layers, hyporeflective vessel lumina, retinal spaces, punctate hyper-reflective vitreous opacities, and vitreous bands (Fig. 1).[14–16] Nonetheless, handheld OCT has never been considered as a potential ROP screening tool because none of these individual findings alone can predict an infant's risk of blindness.

The purpose of this study is to use statistical and machine learning tools to identify early referral-warranted ROP (eRW-ROP) from demographic and handheld spectral domain OCT (SD-OCT) imaging findings of the posterior retina.

Methods

This prospective, observational study enrolled neonates undergoing routine ROP examinations at the University of Washington and Seattle Children's Hospital (Seattle, WA, USA) neonatal intensive care units (NICUs). The study was approved by the University of Washington and Seattle Children's Hospital institutional review boards. All study procedures adhered to regulations of the Health Insurance Portability and Accountability Act and the tenets of the Declaration of Helsinki. The infants' parents or legal guardians gave informed consent prior to the initial handheld OCT imaging session.

Study Participants

Infants were eligible for participation if their gestational age was ≤30 weeks or if birth weight was ≤1500 grams. Pupils were dilated with 1% phenylephrine hydrochloride and 0.2% cyclopentolate ophthalmic solution (Cyclomydril; Alcon Laboratories, Geneva, Switzerland) prior to the imaging and clinical examinations. Each infant's medical record was reviewed for race/ethnicity, sex, gestational age, birth weight, clinical examination results (stage, zone, and presence of plus disease), and history of ophthalmic injections or laser treatment. Infants were excluded if they were deemed too unstable by the NICU medical team for safe BIO examination and/or handheld OCT imaging, if they had inadequate macular SD-OCT image quality, as determined by the image graders, or if they received ROP treatment prior to enrollment.

SD-OCT Imaging

Imaging was performed on the same day as each routine ROP examination using an Envisu C2300 handheld SD-OCT system (Leica Microsystems, Wetzlar, Germany) as previously described.[13] Infants were positioned supine without sedation, and handheld SD-OCT images were acquired by one of three trained imagers (authors A.S., L.E.G., and T.B.G.). Sucrose and/or a pacifier were given to infants as needed to aid compliance with the examination. Fingers were used to retract the eyelids; eyelid specula were not used during examinations, and no contact was made with the ocular surface. Attempts were made to acquire high-quality images of the fovea and optic nerve of both eyes for no longer than 15 minutes per imaging session per infant.

Clinical Examination

Routine BIO examinations were performed by one of three trained pediatric ophthalmologists (including authors M.T.C. and K.T.-H.) on the same day as the handheld SD-OCT image acquisition. BIO examinations, performed with an eyelid speculum and scleral depressor, were used to make all clinical decisions. Clinical examination findings were determined by the ophthalmologist examining the patient on that date. If clinically warranted according to routine care, subsequent examinations for an infant were performed 1 to 3 weeks apart. The pediatric ophthalmologists were masked to handheld SD-OCT imaging results.

Image Analysis

Only right eye findings prior to treatment for type 1 ROP were included in the analysis. The OCT volume scans of each right eye were reviewed by two trained graders (authors A.S. and E.M.Z.) using the system's software (Envivovue; Leica Microsystems, Wetzlar, Germany) for vitreoretinal findings (see Fig. 1). Vitreous bands were defined as coalescing linear opacities above the retina that were visible in at least three consecutive frames with intensity above the background, and were classified as tractional (straight bands at a steep angle that made a connection to the

retinal surface) and non-tractional (bands that did not appear to have a connection to the retina captured by the scan, and hovered parallel to the retina).[14] Two additional trained graders (authors A.T.L. and Y.M.) evaluated all OCT volume scans of each right eye for findings previously defined by Maldonado et al. as predictive of clinical plus disease: vessel elevation (spiked vessels emerging from the inner retinal surface, graded as mild or severe), scalloped retinal layers (mild = scalloping of the inner plexiform layer and severe = scalloping of the outer plexiform layer), hyporeflective vessels, and retinal spaces (hyporeflective spaces located between retinal vessels or adjacent to vessels).[15] As described by Maldonado et al.,[15] all regions of all available OCT volumes were included for assessment of the above findings, without specifying number of involved quadrants. Disagreement among the graders for any OCT finding was mediated by a third trained grader (author M.T.C.), and all graders were masked to the patient's clinical examination results.

A custom, open-sourced web browser-based application developed by one of the authors (A.Y.L.) was used to quantify the density of punctate hyperreflective vitreous opacities, a calculation known to be associated with ROP stage.[16] As previously described,[16] five adjacent B-scan frames centered on the fovea (approximate width = 256 microns) were selected in the highest quality macular scan for a given imaging session based on subjective assessment. These foveal and parafoveal frames were chosen due to their easily identifiable anatomy, which would allow for analysis of comparable vitreous volume and location among infants. Although opacities may have been seen outside these foveal frames, these were not included in this quantitative analysis in order to maintain uniformity in regions and volumes sampled. Two independent, masked, trained graders (authors T.B.G. and M.T.C.) manually segmented the vitreous opacities and vitreoretinal border in these frames, excluding vitreous above the optic nerve where hyaloid remnants are frequently seen. Punctate opacities were defined as having markedly higher signal than background with clear separation from the retina.[14] All completed segmentations were reviewed by an independent, trained grader (author P.T.), and inaccurate segmentations were sent back for repeat segmentation. To estimate vitreous opacity density (vitreous opacity ratio, shown to be associated with ROP stage[16]) in the setting of variable captured vitreous area, the number of vitreous opacities per unit vitreous area in pixels, as determined by the software, was averaged between the 2 graders and multiplied by 100,000,000.

Developing a Predictive Model for eRW-ROP: Generalized Linear Mixed Model

A multivariate model was developed to predict eRW-ROP, which was defined as same-day BIO-determined plus disease, zone I ROP, or ROP stage 3 or greater,[17] with the addition of any pre-plus disease to better capture at-risk infants. As multiple examinations per infant were included, a generalized linear mixed model approach was used to adjust for this. The model included variable OCT findings and postmenstrual age at imaging that were time-specific – that is, on the same date as the clinical examination – however, the birth weight and gestational age were infant-specific regardless of the date of examination. Variable selection was performed by backward elimination using the Akaike information criterion for model evaluation to predict eRW-ROP using candidate demographic variables (gender, birth weight in grams, gestational age in weeks, race/ethnicity, and postmenstrual age at imaging in weeks) and OCT variables (presence of tractional and non-tractional vitreous bands, vitreous opacity ratio, vessel elevation [none, mild, and severe], scalloped retinal layers [none, mild, and severe], presence of hyporeflective vessels, and presence of retinal spaces). Recognizing the known correlation between gestational age and birth weight, each of these two variables were individually excluded to assess impact on the model, ultimately creating a final model excluding one of these variables.

To understand whether OCT alone, without demographic variables, could feasibly diagnose referral-warranted ROP, a modified mixed model was created that incorporates only the OCT variables above (before backward elimination). Conversely, to assess the predictive power of infant demographics that are currently used for ROP screening criteria, a modified mixed model was created including only birth weight and gestational age.

Developing a Predictive Model for eRW-ROP: Machine Learning

Extreme gradient boosting, a machine learning tool (Python Software Foundation, version 3.8.1, http://www.python.org; XGBoost, https://github.com/dmlc/xgboost; scikit-learn),[18] was used to develop a regression model to similarly predict eRW-ROP based on the same candidate explanatory variables that were used for the generalized linear mixed model (gender, birth weight, gestational age, race/ethnicity, postmenstrual age at imaging, presence of tractional and non-tractional vitreous bands, vitreous opacity ratio, vessel

elevation, scalloped retinal layers, presence of hyporeflective vessels, and presence of retinal spaces). Shapley additive explanations (SHAP) values were generated to provide an estimate of the impact of each observation on the model's output.

Additional Statistical Analysis

A receiver operating characteristic (ROC) curve was created for each model, using clinical eRW-ROP as the reference standard. Areas under the curve (AUC) were calculated for each model. Sensitivity and specificity values were reported based on the maximum Youden index (J) for each model. The unweighted Cohen's kappa (k) statistic was used to evaluate interobserver agreement for the presence or absence of all OCT vitreoretinal findings (tractional vitreous bands, non-tractional vitreous bands, vessel elevation, scalloped retinal layers, hyporeflective vessels, and retinal spaces). For the vitreous opacity ratio, interobserver agreements for the manual counting of punctate hyper-reflective vitreous opacities were assessed by F1 score and agreements for the vitreoretinal border segmentation were assessed using a Dice coefficient (Python Software Foundation, version 3.7.4, http://www.python.org; scikit-learn; NumPy; OpenCV).[18–20] Other than vitreous opacity ratio calculations, all statistical analyses were performed using SAS version 9.4 (SAS Institute, Inc., Cary, NC, USA).

Results

Study Participants

A total of 111 premature infants were recruited for the study, of which 40 were excluded. Twenty were excluded due to poor image quality (attributable to difficult alignment of the probe with the infant's direction of gaze). Poor image quality was defined as lack of visualization of all retinal laminations and/or inadequate signal strength to identify all OCT findings of interest, or absence of five high-quality B-scan frames centered at the fovea for calculation of the vitreous opacity ratio. Eight infants were withdrawn by their guardians due to concern for infant stress, two were determined to be too hemodynamically unstable by the NICU medical team, two were withdrawn by the parents who did not want to pursue follow-up, one was withdrawn after death due to unrelated respiratory failure, five were excluded due to lost imaging data, and two were excluded due to prior treatment for type 1 ROP. The remaining 71 infants underwent 176 imaging sessions. Infants were generally visibly comfortable for all handheld OCT imaging sessions, however, these data were not recorded as part of this study. After excluding an additional 9 right eye imaging sessions among 4 infants that occurred following treatment for type 1 ROP, the remaining 167 right eye imaging sessions from 71 infants were included in the study. Of these infants, 32 (45%) were boys. Mean birth

Table. Summary of Demographic and Clinical Data

Characteristic	*N* (%) Unless Otherwise Specified
All subjects	71 (100)
Postmenstrual age at time of imaging, mean (SD), wk[a]	38.4 (4.8)
Demographics	
Birth weight, mean (SD), grams	995.6 (292.0)
Gestational age, mean (SD), wk	28.2 (2.8)
Gender, *n* male	32 (45)
Race/ethnicity	
White	39 (55)
Hispanic	12 (17)
Black	5 (7)
Asian	4 (6)
Native American	2 (3)
Pacific Islander	1 (1)
Other[b]	8 (11)
Indirect ophthalmoscopic examination findings[c]	
Stage of ROP	
0	40 (56)
1	12 (17)
2	8 (11)
3	11 (16)
Plus disease status	
No plus disease	63 (89)
Pre-plus disease	4 (6)
Plus disease	4 (6)
eRW-ROP[d]	12 (17)
Type 1 ROP	4 (6)

SD, standard deviation; ROP, retinopathy of prematurity; eRW-ROP, early referral-warranted retinopathy of prematurity.

[a]Among 167 imaging sessions.

[b]Patients who did not have a race/ethnicity listed in their medical records.

[c]All findings are for right eyes only. Worst disease among all imaging sessions is shown.

[d]The eRW-ROP was defined as the presence of plus disease, zone I ROP, or ROP stage 3 or greater,[17] with the addition of pre-plus disease to better capture high-risk infants. This includes patients counted elsewhere in the table.

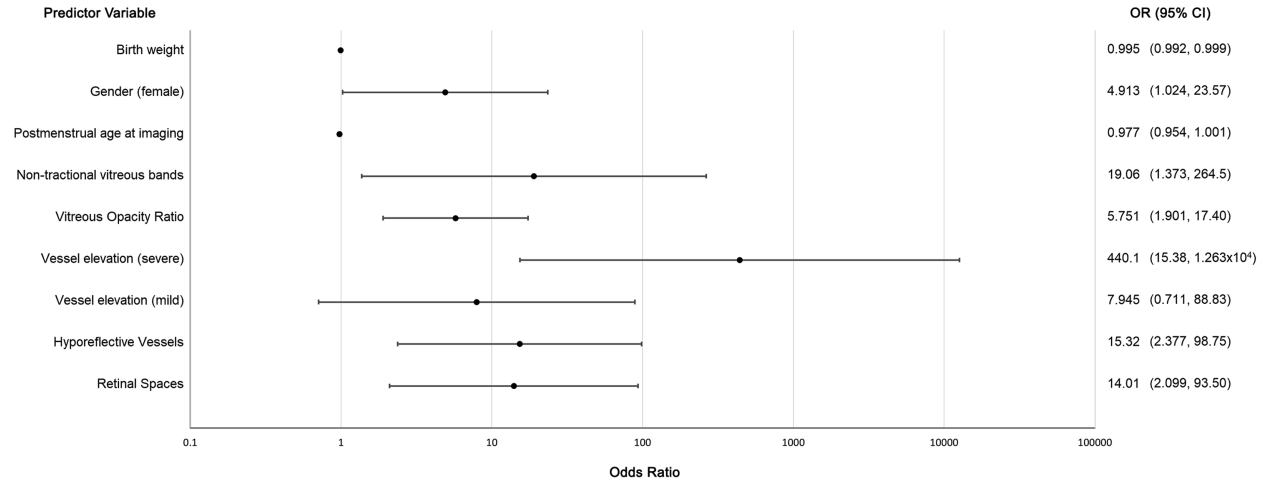

Figure 2. Forest plot of estimated odds ratios for predictors of clinical early referral-warranted retinopathy of prematurity. A generalized linear mixed model approach was used with backward stepwise elimination. Gestational age was not included to avoid the known correlation with birth weight. The final mixed model included handheld spectral domain optical coherence tomography findings and demographic variables listed. Odds ratios are displayed on a logarithmic x-axis scale, with 95% confidence interval error bars. OR, odds ratio; CI, confidence interval.

weight was 995.6 (±292.0) grams, mean gestational age was 28.2 (±2.8) weeks, and mean postmenstrual age at imaging was 36.4 (±3.5 weeks). Thirty-nine of the 71 infants (55%) were White, 12 (17%) were Hispanic, 2 (3%) were Native American, 5 (7%) were Black, 1 (1%) was Pacific Islander, 4 (6%) were Asian, and 8 (11%) were other. The mean number of imaging sessions per infant was 2.4 (±1.6, range = 1-8) over a mean of 4.6 (±4.2) weeks. Twelve of the 71 infants (17%) developed eRW-ROP in the right eyes at mean 1.7 (±4.2) weeks after their first examination, and 4 (5.6%) developed type 1 ROP in the right eyes which required treatment. All infants that required treatment received laser photocoagulation. Demographic and clinical data are summarized in the Table.

Handheld SD-OCT Findings

Among 71 infants included in the study, right eye handheld SD-OCT imaging revealed that 6 infants (8%) had evidence of tractional vitreous bands, 20 (28%) had non-tractional vitreous bands, and 17 (24%) had macular punctate hyper-reflective vitreous opacities (within the five B-scans centered at the fovea). Among the 17 infants with macular vitreous opacities, the maximum vitreous opacity ratio reached by each infant was mean 1.12 (±1.13). Agreement between 2 graders in counting opacities revealed an F1 score of 0.82 (±0.36) and a Dice coefficient for vitreous segmentation of 0.97 (±0.04). With regard to the VASO criteria outlined by Maldonado et al.,[15] 53 (75%) demonstrated vessel elevation (43/53 [81%] mild and 10/53 [19%] severe), 38 (54%) had scalloped retinal layers (31/38 [82%] involving inner plexiform layer, and 7/38 [18%] involving outer plexiform layer), 11 (15%) had hyporeflective vessels, and 12 (17%) had retinal spaces. Cohen's kappa coefficient for inter-observer agreement was 0.61, rated as "substantial" agreement.[21] Agreement among all OCT findings was 79.8%.

Generalized Linear Mixed Model

The logistic mixed model consisted of the following nine predictors from demographic and handheld OCT imaging data: birth weight, gestational age, postmenstrual age at imaging, non-tractional vitreous bands, vitreous opacity ratio, vessel elevation, scalloped retinal layers, hyporeflective vessels, and retinal spaces. The model eliminated gender, race/ethnicity, and tractional vitreous bands as part of the backward elimination process. The AUC was 0.96, with a sensitivity of 90.9% and a specificity of 96.6% (J = 0.88). Due to the known correlation between birth weight and gestational age, each of these two variables were manually excluded from the model in turn, which resulted in minimal change to the predictive power (AUC = 0.91 after birth weight removed, and 0.94 after gestational age removed). Therefore, the model without gestational age was selected as the "best" model to avoid this correlation (Fig. 2). This final model eliminated race/ethnicity and tractional vitreous bands as part of the backward elimination process. The strongest predictors of eRW-ROP in that model

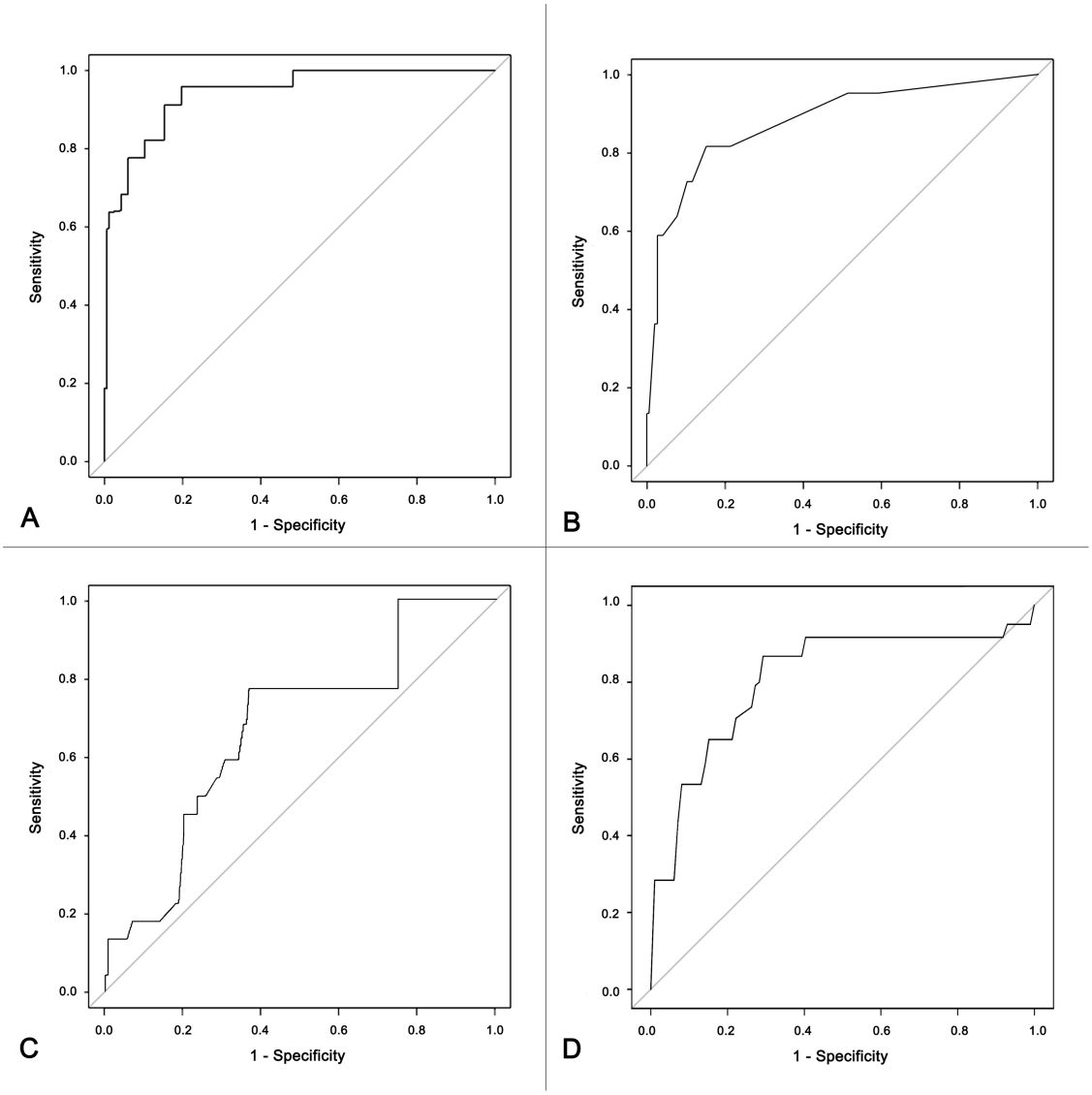

Figure 3. Receiver operating characteristic curves of multivariate models to predict clinical early referral-warranted retinopathy of prematurity. (**A**) A generalized linear mixed model was developed, with candidate explanatory variables that included demographic data and handheld spectral domain optical coherence tomography findings. Gestational age was not included to avoid the known correlation with birth weight. The final multivariate model included eight predictors: birth weight, gender, postmenstrual age at imaging, non-tractional vitreous bands, vitreous opacity ratio, vessel elevation, hyporeflective vessels, and retinal spaces. Area under the curve was 0.94, sensitivity 95.5%, and specificity 80.7%. (**B**) A mixed model consisting of only OCT imaging variables had an area under the curve of 0.88, sensitivity 81.8%, and specificity 84.8%. (**C**) A mixed model that included only birth weight and gestational age as explanatory variables had an area under the curve of 0.68, sensitivity 77.3%, and specificity 63.4%. (**D**) Regression was also performed using machine learning tools (extreme gradient boosting), resulting in an area under the curve of 0.83, sensitivity 91.7%, and specificity 77.8%.

were severe vessel elevation ($P < 0.001$), vitreous opacity ratio ($P = 0.002$), hyporeflective vessels ($P = 0.004$), retinal spaces ($P = 0.006$), and birth weight ($P = 0.006$). The AUC was 0.94 (Fig. 3A), with a sensitivity of 95.5% and a specificity of 80.7% (J = 0.76).

The mixed model consisting of only OCT variables included vitreous opacity ratio, vessel elevation, hyporeflective vessels, and retinal spaces. Tractional vitreous bands, non-tractional vitreous bands, and scalloped retinal layers were removed by backward elimination. This model had an AUC of 0.88 (Fig. 3B), with a sensitivity of 81.8% and a specificity of 84.8% (J = 0.67).

The mixed model that included only the explanatory variables birth weight and gestational age had a reduced AUC of 0.68 (Fig. 3C), with a sensitivity of 77.3% and a specificity of 63.4% (J = 0.41).

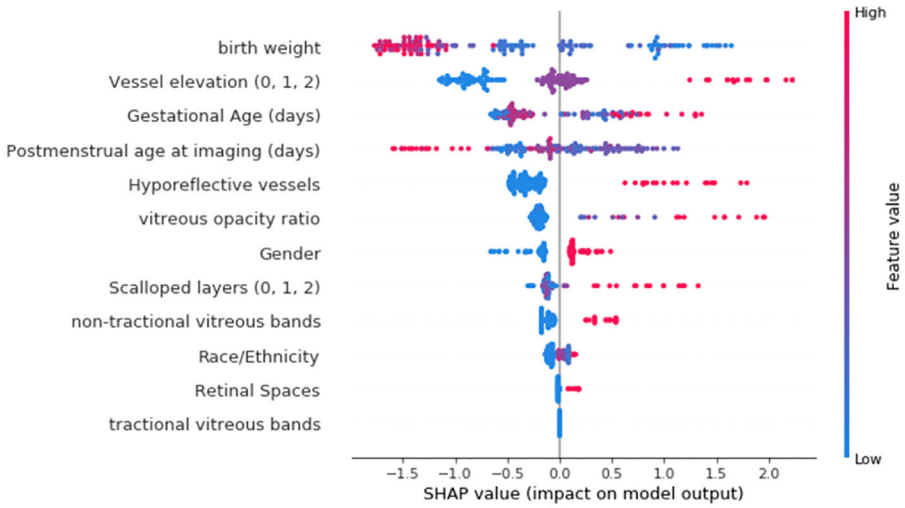

Figure 4. Plot of each imaging session's SHAP value for all variables in the extreme gradient boosting model. Extreme gradient boosting by machine learning resulted in a regression model to predict early referral-warranted retinopathy of prematurity. Predictor variables are organized along the y-axis by overall significance to the final model. X-axis indicates positive or negative impact and scale. Colors represent the relative value of the data points (from high to low). SHAP, Shapley additive explanations.

Machine Learning

The extreme gradient boosting approach yielded a multivariate model with AUC 0.83 (Fig. 3D), sensitivity 91.7%, and specificity 77.8% (J = 0.69) for predicting eRW-ROP. The most impactful variables based on SHAP values were birth weight, vessel elevation, gestational age, postmenstrual age at the time of imaging, hyporeflective vessels, and vitreous opacity ratio (Fig. 4).

Discussion

Utilizing 167 imaging sessions from 71 premature infants at a single site, we developed models to identify eRW-ROP using only posterior retinal handheld SD-OCT findings and demographic data, with the best being the generalized linear mixed model (AUC = 0.94, sensitivity = 95.5%, and specificity = 80.7%). To our knowledge, this study is the first to comprehensively diagnose early high-risk ROP through handheld OCT (PubMed search terms: optical coherence tomography AND retinopathy of prematurity), although further validation is needed.

The gold standard ROP screening examination includes BIO[22] and, more recently, wide-field retinal photography.[23] These approaches involve interpretation of both central and peripheral retinal pathology using techniques that are distressing to the infants. This stress has been associated with the use of an eyelid speculum or scleral depressor.[2–9] Investigations into the utility of OCT in diagnosing ROP have focused either on the posterior vitreoretinal interface or the peripheral vascular-avascular junction. In the latter, techniques are limited by use of eyelid speculum, scleral depression, and/or direct contact between the imaging probe and the eye, leading to infant distress.[10–12] In contrast, handheld OCT of the posterior retina requires neither ocular contact nor an eyelid speculum. In this study and others,[8] infants generally remained comfortable throughout OCT image acquisition. Ultimately, OCT imaging of the posterior vitreoretinal interface may reduce the number of distressing, traditional examinations currently necessary to screen preterm infants for ROP.

Other studies have also sought to reduce the number of infants receiving such disruptive examinations by improving on existing screening criteria for ROP. Recently, the Postnatal Growth and Retinopathy of Prematurity study (G-ROP) demonstrated that consideration of postnatal growth may reduce the number of infants that require ophthalmoscopy by as much as 15% to 30%.[24–26] In contrast, instead of creating screening criteria, as in G-ROP, we propose that OCT imaging could be analogous to digital wide-field retinal imaging, similarly identifying a subset of infants who would require ophthalmoscopy, but with the advantage of significantly reduced cardiorespiratory disturbance to the infants and potentially greater quantitative data. New screening criteria, such as G-ROP, have the potential to work synergistically with OCT, by creating a reduced number of infants requiring OCT screening. OCT then provides quantitative evaluation of the

posterior segment in these infants, allowing for more specific information on ROP risk to reduce the number of ophthalmoscopy examinations even further. Such an approach requires further validation, however.

Machine learning has served as an exciting new tool for data analysis in ophthalmology,[27] however, this study's extreme gradient boosting yielded a less optimal model (AUC = 0.83, sensitivity = 91.7%, and specificity = 77.8%) compared to the generalized linear mixed model. Limitations of machine learning include skewed data analysis when based on smaller, less representative datasets or inadequate representation from rare pathologies in the dataset.[27] Our results highlight the importance of model exploration using traditional statistical approaches in addition to machine learning, particularly when datasets are limited.

The OCT variables used for our models were previously identified potential predictors of ROP risk located in the posterior pole. Vessel elevation, scalloped retinal layers, hyporeflective vessels, and retinal spaces were described by Maldonado et al. to collectively diagnose plus disease by imaging the posterior retina.[15] We previously established the vitreous opacity ratio (an estimate of density of punctate hyperreflective vitreous opacities) and vitreous bands as OCT findings correlating with ROP stage and plus disease, respectively.[16] Future exploration into other OCT-based biomarkers of ROP using emerging technologies is warranted, such as handheld OCT angiography,[28] choroidal measurements,[29] and wide-field OCT imaging of the ROP ridge.[30,31] Further automation in image analysis would decrease subjectivity and improve efficiency for OCT analysis of all such biomarkers.

As in previous weight-based models of ROP risk,[25,32–34] demographic factors, such as birth weight or gestational age, were included to improve the performance of the models in this study. Nonetheless, demographics were relatively weak drivers of these models (see Figs. 2, 4). Excluding the demographics entirely with a model using only OCT variables resulted in a good although lower AUC of 0.88 (see Fig. 3B), with a sensitivity of 81.8% and a specificity of 84.8%. In contrast, an algorithm containing birth weight and gestational age alone resulted in a significant drop in predictive power (AUC = 0.68, with a sensitivity of 77.3% and a specificity of 63.4%; see Fig. 3C).

This study is limited by a single site population without a separate validation set for the models. A validation set could not be achieved with adequate sampling of severe ROP in the timeframe of the current study. The AUC is therefore not indicative of the models' performance in any dataset other than that used to derive the models. Nonetheless, the primary purpose of this work is to serve as a proof-of-concept for the utility of handheld OCT of the posterior retina to assess ROP risk in preterm infants, rather than to create a final model validated for clinical use. Once validated, future clinical use could involve reducing the number of patients that require distressing screening examinations. A cutoff for the predictive formula would be set at 100% sensitivity (53% specificity in the model presented here) to avoid missing any patients with vision-threatening disease. Limitations in handheld SD-OCT image quality in awake preterm infants should also be acknowledged. Twenty of 111 (18%) infants were excluded due to inadequate image quality (attributable to difficult alignment of the probe with the infant's direction of gaze), suggesting that this device would face challenges in real-life clinical application in awake infants. However, use of this device can still reduce how many infants receive ophthalmoscopic examinations, and recent advances in newer handheld OCT technology have resulted in consistently high image quality for awake infant vitreoretinal analysis.[28,35] Validation of these algorithms with such newer devices are needed. Finally, although this study is limited by lack of quantitative vitreous opacity analysis outside the macula, this previously validated approach allows for uniform anatomic sampling among infants.[16]

In summary, this study demonstrated that multivariate models using demographic and handheld OCT data can identify eRW-ROP in premature infants. The use of handheld OCT as a better-tolerated early ROP screening tool would reduce examination morbidity in this vulnerable population. Validation of future models in large multicenter studies is necessary before considering clinical use.

Acknowledgments

The authors thank Nicole Mattson, Luke Johnson, Calvin Lee, Jordan Sandhu (all University of Washington undergraduate student volunteers), and Hannah Walsh (Yale undergraduate student volunteer) for serving as research assistants for this study.

Supported by grants from The Knights Templar Eye Foundation (M.T.C.), the Latham Vision Research Innovation Award (M.T.C.), Violet Sees (M.T.C.), and the National Institutes of Health (grant numbers K23EY029246 [A.Y.L.], K23EY024921, and R01AG060942 [C.S.L.]). This work was also supported by an unrestricted grant and Career Development Award from Research to Prevent Blindness, and a National Institutes of Health CORE Grant

(grant number EY001730) to the University of Washington Department of Ophthalmology. The funding organizations had no role in the study design, data collection, analysis and interpretation of data, writing of the report, or the decision to submit the article for publication.

Disclosure: **A.T. Legocki**, None; **A.Y. Lee**, Santen (F), Carl Zeiss Meditec (F), Novartis (F), Genentech (C), Topcon (C), Verana Health (C), Johnson and Johnson (C), Gyroscope (C), Microsoft (F), Regeneron (F); **L. Ding**, None; **Y. Moshiri**, None; **E.M. Zepeda**, None; **T.B. Gillette**, None; **L.E. Grant**, None; **A. Shariff**, None; **P. Touch**, None; **C.S. Lee**, None; **K. Tarczy-Hornoch**, None; **M.T. Cabrera**, None

References

1. Blencowe H, Moxon S, Gilbert C. Update on blindness due to retinopathy of prematurity globally and in India. *Indian Pediatr*. 2016;53(Suppl 2):S89–S92.
2. Mehta M, Adams GG, Bunce C, Xing W, Hill M. Pilot study of the systemic effects of three different screening methods used for retinopathy of prematurity. *Early Hum Dev*. 2005;81(4):355–360.
3. Moral-Pumarega MT, Caserio-Carbonero S, De-La-Cruz-Bertolo J, Tejada-Palacios P, Lora-Pablos D, Pallas-Alonso CR. Pain and stress assessment after retinopathy of prematurity screening examination: Indirect ophthalmoscopy versus digital retinal imaging. *BMC Pediatr*. 2012;12:132.
4. Mukherjee AN, Watts P, Al-Madfai H, Manoj B, Roberts D. Impact of retinopathy of prematurity screening examination on cardiorespiratory indices: A comparison of indirect ophthalmoscopy and Retcam imaging. *Ophthalmology*. 2006;113(9):1547–1552.
5. Rush R, Rush S, Nicolau J, Chapman K, Naqvi M. Systemic manifestations in response to mydriasis and physical examination during screening for retinopathy of prematurity. *Retina*. 2004;24(2):242–245.
6. Laws DE, Morton C, Weindling M, Clark D. Systemic effects of screening for retinopathy of prematurity. *Br J Ophthalmol*. 1996;80(5):425–428.
7. Mitchell AJ, Green A, Jeffs DA, Roberson PK. Physiologic effects of retinopathy of prematurity screening examinations. *Adv Neonatal Care*. 2011;11(4):291–297.
8. Mangalesh S, Sarin N, McGeehan B, et al. Preterm infant stress during handheld optical coherence tomography vs binocular indirect ophthalmoscopy examination for retinopathy of prematurity. *JAMA Ophthalmol*. 2021;139(5):567–574.
9. Dhaliwal CA, Wright E, McIntosh N, Dhaliwal K, Fleck BW. Pain in neonates during screening for retinopathy of prematurity using binocular indirect ophthalmoscopy and wide-field digital retinal imaging: a randomised comparison. *Arch Dis Child Fetal Neonatal Ed*. 2010;95(2):F146–F148.
10. Nguyen TP, Ni S, Khan S, et al. Advantages of widefield optical coherence tomography in the diagnosis of retinopathy of prematurity. *Front Pediatr*. 2021;9:797684.
11. Nguyen TP, Ni S, Ostmo S, et al. Association of optical coherence tomography-measured fibrovascular ridge thickness and clinical disease stage in retinopathy of prematurity. *JAMA Ophthalmol*. 2022;140(11):1121–1127.
12. Scruggs BA, Ni S, Nguyen TP, et al. Peripheral OCT assisted by scleral depression in retinopathy of prematurity. *Ophthalmol Sci*. 2022;2(1):100094.
13. Maldonado RS, Izatt JA, Sarin N, et al. Optimizing hand-held spectral domain optical coherence tomography imaging for neonates, infants, and children. *Invest Ophthalmol Vis Sci*. 2010;51(5):2678–2685.
14. Zepeda EM, Shariff A, Gillette TB, et al. Vitreous bands identified by handheld spectral-domain optical coherence tomography among premature infants. *JAMA Ophthalmol*. 2018;136(7):753–758.
15. Maldonado RS, Yuan E, Tran-Viet D, et al. Three-dimensional assessment of vascular and perivascular characteristics in subjects with retinopathy of prematurity. *Ophthalmology*. 2014;121(6):1289–1296.
16. Legocki AT, Zepeda EM, Gillette TB, et al. Vitreous findings by handheld spectral-domain oct correlate with retinopathy of prematurity severity. *Ophthalmol Retina*. 2020;4(10):1008–1015.
17. Ells AL, Holmes JM, Astle WF, et al. Telemedicine approach to screening for severe retinopathy of prematurity: a pilot study. *Ophthalmology*. 2003;110(11):2113–2117.
18. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. *J Machine Learning Res*. 2011;12:2825–2830.
19. Oliphant TE. *A guide to NumPy*: 2nd Edition. In: Watson K, Ed. Austin TX: Continuum Press, a division of Continuum Analytics, Inc. 2006.
20. Bradski G. The OpenCV Library. *Dr Dobb's Journal of Software Tools*. 2000:122–125.

21. McHugh ML. Interrater reliability: The kappa statistic. *Biochem Med (Zagreb)*. 2012;22(3):276–282.
22. 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*. 2018;142(6):e20183061.
23. Chiang MF, Melia M, Buffenn AN, et al. Detection of clinically significant retinopathy of prematurity using wide-angle digital retinal photography: A report by the American Academy of Ophthalmology. Research Support, Non-U.S. Government. *Ophthalmology*. 2012;119(6):1272–1280.
24. Ahmed ISH, Aclimandos W, Azad N, et al. The postnatal growth and retinopathy of prematurity model: A multi-institutional validation study. *Ophthalmic Epidemiol*. 2021;29:296–301.
25. Binenbaum G, Bell EF, Donohue P, et al. Development of modified screening criteria for retinopathy of prematurity: primary results from the postnatal growth and retinopathy of prematurity study. *JAMA Ophthalmol*. 2018;136(9):1034–1040.
26. Binenbaum G, Tomlinson LA, de Alba Campomanes AG, et al. Validation of the postnatal growth and retinopathy of prematurity screening criteria. *JAMA Ophthalmol*. 2020;138(1):31–37.
27. Ting DSW, Pasquale LR, Peng L, et al. Artificial intelligence and deep learning in ophthalmology. *Br J Ophthalmol*. 2019;103(2):167–175.
28. Moshiri Y, Legocki AT, Zhou K, et al. Handheld swept-source optical coherence tomography with angiography in awake premature neonates. *Quant Imaging Med Surg*. 2019;9(9):1495–1502.
29. Moreno TA, O'Connell RV, Chiu SJ, et al. Choroid development and feasibility of choroidal imaging in the preterm and term infants utilizing SD-OCT. *Invest Ophthalmol Vis Sci*. 2013;54(6):4140–4147.
30. Chen X, Mangalesh S, Dandridge A, et al. Spectral-domain OCT findings of retinal vascular-avascular junction in infants with retinopathy of prematurity. *Ophthalmol Retina*. 2018;2(9):963–971.
31. Campbell JP, Nudleman E, Yang J, et al. Handheld optical coherence tomography angiography and ultra-wide-field optical coherence tomography in retinopathy of prematurity. *JAMA Ophthalmol*. 2017;135(9):977–981.
32. Binenbaum G, Ying GS, Quinn GE, et al. The CHOP postnatal weight gain, birth weight, and gestational age retinopathy of prematurity risk model. *Arch Ophthalmol*. 2012;130(12):1560–1565.
33. Hellström A, Hård AL, Engström E, et al. Early weight gain predicts retinopathy in preterm infants: new, simple, efficient approach to screening. *Pediatrics*. 2009;123(4):e638–e645.
34. Cao JH, Wagner BD, McCourt EA, et al. The Colorado-retinopathy of prematurity model (CO-ROP): postnatal weight gain screening algorithm. *J AAPOS*. 2016;20(1):19–24.
35. Scoville NM, Legocki AT, Touch P, et al. Vitreous opacities in infants born full-term and preterm by handheld swept-source optical coherence tomography. *J AAPOS*. 2022;26(1):20.e1–20.e7.