AOS Thesis

Application of a Quantitative Vascular Severity Score in Retinopathy of Prematurity in the United States and India: New Insights Into Disease Epidemiology and Pathophysiology

J. PETER CAMPBELL, SAMEERA NAYAK, SUSAN OSTMO, AARON S. COYNER, PRAVEER SINGH, JAYASHREE KALPATHY-CRAMER, PARAG SHAH, RV PAUL CHAN, AND MICHAEL F. CHIANG

• PURPOSE: To compare the epidemiology and phenotypic characteristics of 2 populations of patients with retinopathy of prematurity (ROP) using a quantitative artificial intelligence (AI)–derived vascular severity score (VSS).

• METHODS: We retrospectively analyzed 2 data sets of ROP images: (1) The iROP data set including 1019 babies and 6552 eye-level examinations developed as part of a large US multicenter cohort study from 2012 to 2020. (2) 8281 consecutive ROP telemedicine examinations from 2363 babies in the Indian Aravind Eye Care System (AECS), from 2019 to 2020. We analyzed all eye examinations using the iROP DL system, which outputs a VSS from 1 to 9. The main outcome was the relationship between the VSS and the zone, stage, and extent of stage 3 in each data set, assessed using generalized estimating equations to account for the correlation from bilateral and repeated examinations.

• RESULTS: In both data sets, higher VSS was associated with higher stage ($P < .001$ for all stages in both data sets), higher extent of stage 3 (i-ROP, $P = .004$, and AECS, $P < .001$), and a more posterior zone for a given stage of ROP (i-ROP: stage 1, $P = .04$, stage 2, $P < .001$, and stage 3, $P < .001$; AECS: stage 1, $P < .001$, and stage 3, $P < .001$). For any given stage, the VSS was lower in the AECS data set than in the i-ROP data set (all $P < .001$).

• CONCLUSIONS: In both populations, the VSS was associated with more posterior zone and higher ROP stage/extent, though on average the VSS was lower in the AECS population. A correlation between measured VSS and intraocular vascular endothelial growth factor concentration might explain both findings. (Am J Ophthalmol 2026;290: 123–136.

INTRODUCTION

• A HISTORICAL LOOK AT RETINOPATHY OF PREMATURITY: Understanding the regional epidemiology of retinopathy of prematurity (ROP) has always depended on identifying a given region's evolution in the broader epidemiologic transition of the modern era, specifically as it reflects neonatal mortality and the survival of prematurely born infants.[1,2] In regions and times where premature infants do not survive, the incidence of ROP is low. As neonatal care, including resuscitation of earlier preterm infants, increases, the incidence of ROP rises significantly. Finally, as modern neonatal care practices mature, including careful monitoring and titration of oxygen, minimization of infection, and optimization of nutrition, the incidence of severe ROP levels out with the risk of ROP being primarily related to the degree of preterm birth.[1,2]

Looking back at the evolution of ROP in the 20th century, these transitions have been described in terms of 3 "epidemics." Although these are not true epidemics in the epidemiologic sense, the schema has become so ingrained

AJO.com NOTE: Publication of this article is sponsored by the American Ophthalmological Society.
Accepted for publication June 8, 2026.

From the Casey Eye Institute, Oregon Health & Science University, Portland, Oregon, USA (J.P.C., S.N., S.O., A.S.C.); Anant Bajaj Retina Institute, LV Prasad Eye Institute, India (S.N.); Department of Ophthalmology, University of Colorado School of Medicine, Aurora, Colorado, USA (P.S., J.K.-C.); Aravind Eye Care System, Department of Pediatric Retina, Coimbatore, India (P.S.); Illinois Eye and Ear Infirmary, University of Illinois at Chicago, Chicago, Illinois, USA (R.P.C.); National Library of Medicine and National Eye Institute, National Institutes of Health, Bethesda, Maryland, USA (M.F.C.).

Inquires to: J. Peter Campbell, Casey Eye Institute, Oregon Health & Science University, 545 SW Campus Drive, Portland, Oregon, 97239, USA; e-mail: campbelp@ohsu.edu

0002-9394/$36.00
https://doi.org/10.1016/j.ajo.2026.06.008

in our understanding of ROP that we will continue to use the term here. The first epidemic of what we now call ROP occurred with the advent of modern neonatal care in the mid-20th century in high-income regions and was characterized by babies developing "retrolental fibroplasia" (RLF) with total retinal detachments after surviving premature birth.[3,4] In the next few decades, as our understanding of the spectrum of disease and its risk factors, including exposure to oxygen, infection, etc., improved, the incidence of severe ROP decreased dramatically, before any proven treatment or formal disease classification. The major lesson from the first epidemic of ROP was the key role of primary prevention in reducing the number of children with retrolental fibroplasia.

The second epidemic of ROP evolved as modern neonatal care pushed the boundaries of viability to lower and lower gestational ages (GAs). In the first epidemic, babies born at 34 weeks might develop RLF, and babies born at less than 30 weeks were not surviving. Optimizing neonatal care and our understanding of ROP physiology changed both of these things; in the second epidemic, the 34-week baby was no longer developing severe ROP, but babies at younger and younger GAs were surviving. Conceptually, we remain in this phase of the disease and will remain so for the foreseeable future in higher-income countries, where babies as young as 22 weeks are now surviving.[2] The 2 largest risk factors in this era for severe ROP are the degree of prematurity and the exposure to oxygen necessary for survival.[5,6]

The so-called third epidemic of ROP reflects the broader epidemiologic transition that is occurring throughout low- and middle-income countries worldwide.[1,2,7] In these regions, not only is the rate of premature birth higher, but resource constraints limit the ability of emerging neonatal care units to provide optimal staffing and monitoring of infants, and the lack of screening means that many babies progress to retinal detachment and blindness, despite the advancements in ROP care in the last 50 years which have demonstrated that this outcome is nearly always avoidable. India is a country of >1 billion people and has one of the highest prematurity rates in the world. As health care conditions have improved in the last 3 decades, ophthalmologists in India have begun to see what might rightly be described as an emerging epidemic of retrolental fibroplasia.[8,9]

In many ways, the third epidemic and the first epidemic of ROP share a number of similarities; fewer premature babies were (are) developing severe ROP, in large part due to access to unregulated oxygen. As described previously, the solution to the third epidemic is multifaceted—first and foremost, improving primary prevention will reduce the population at risk and the incidence of severe ROP for any given GA. But secondarily, systems need to be developed to provide secondary prevention through ophthalmology-led ROP screening, either in person or via telemedicine, in each region where babies are surviving. There are a number of ways to accomplish this, but first it is important to take a step back and look at how we diagnose and classify ROP and how this has changed as the epidemiology of ROP has evolved.

• **DISEASE PHENOTYPES, CLASSIFICATION, AND TREATMENT:** The International Classification of ROP (ICROP) was first defined in the 1980s, in conjunction with the planning and execution of the first randomized clinical trial of cryotherapy (CRYO-ROP).[10,11] The ICROP defined the language that enabled clinicians to interpret and apply treatment consistently based on the ophthalmoscopic examination findings. ICROP defined the "zone," which is a method for describing the extent of retinal vascularization from the optic nerve to the ora serrata—a process that normally completes by term birth but is in various degrees of immaturity with premature birth. The "stage" of disease was defined as the degree of pathology at the vascular-avascular border. Plus disease was defined as marked dilation and tortuosity of the posterior retinal blood vessels, occasionally associated with neovascularization of the iris and/or an enlarged persistent tunica vasculosa lentis. Interestingly, this finding, now called plus disease, was one of the earliest recognized risk factors for subsequent retrolental fibroplasia in the earliest descriptions of the disease. The CRYO-ROP study used the definitions from ICROP to evaluate whether treatment of "threshold ROP" would lead to improved anatomical outcomes (fewer retinal detachments). Threshold ROP was defined as 5 contiguous clock hours of stage 3 or 8 noncontiguous clock hours of stage 3, in zone I or II, with plus disease. Subsequently, cryotherapy became the standard of care after the publication of this study in 1988.

The treatment "threshold" was re-evaluated in the Early Treatment for ROP (ETROP) study, published in 2003.[12] ETROP evaluated whether treatment of "pre-threshold" ROP might improve outcomes and found that treatment of "type 1" prethreshold ROP was advantageous over waiting for threshold disease to develop. This changed the treatment criteria to zone I or II ROP, any stage, with plus disease, or zone I, stage 3 (with or without plus disease).[12] Notably, this shift removed the only "measurable" or quantifiable biomarker for disease severity—the number of clock hours of stage 3 that defined threshold ROP—and replaced it with the simpler but more subjective heuristic that (apart from zone I, stage 3), treatment depended only on whether plus disease was present or absent.

The ICROP was revisited shortly after the ETROP was published and introduced several new concepts, including increased recognition of the importance of zone I ROP, the introduction of aggressive posterior ROP (APROP), and the concept of "pre-plus" disease, all of which are important phenotypic characteristics that have certain epidemiologic and prognostic implications.[13] Zone I ROP was recognized to be at higher risk for progression and was related to the main demographic risk factor—GA. The lower the

GA, the more posterior the zone of ROP is, in general, *and* the more oxygen is required for resuscitation, on average. APROP was most common in zone I eyes and was defined both by its phenotypic appearance (flat neovascularization rather than the more traditional "staged" appearance and plus disease "out of proportion" to the degree of apparent stage) and functionally by its faster progression to retinal detachment. Finally, the introduction of pre-plus reflected the earliest acknowledgment that vascular changes associated with plus disease were not a binary disease characteristic.[14,15]

Therapies have evolved over time, both with and without clinical trials. Between CRYO-ROP and ETROP, indirect laser photocoagulation replaced cryotherapy as the standard of care, without a pivotal study comparing the 2 treatment modalities. Subsequent to ETROP, there have been several pivotal clinical trials for anti-vascular endothelial growth factor (anti-VEGF) treatments.[16-18] Although these have not changed our indications for treatment, anti-VEGF has become increasingly the first line for the highest-risk eyes (e.g., zone I and APROP eyes), with multiple studies demonstrating improved anatomic outcomes with anti-VEGF compared with laser in this population.[19-21]

The most recent iteration of ICROP was published in 2021 with several key goals: to be more broadly representative of the spectrum of pathology found both in high- and low-income regions (second and third epidemics), to introduce new concepts unique to the anti-VEGF era, such as reactivation, and to further elaborate on the spectrum of plus disease. APROP was renamed aggressive ROP (AROP) because, in the third epidemic regions, AROP can occur in zone II eyes with high oxygen exposure. This often leads to phenotypic appearances that differ from traditional zone I "APROP."[20,22] Whereas the 2005 ICROP introduced the concept of pre-plus, the 2021 edition[23] formally acknowledged that dilation and tortuosity reflect a spectrum of disease severity, a finding that was largely the result of a decade of literature identifying discrepancies in the diagnosis of plus disease, in part due to systematic bias and the challenges of consistently applying a reference standard image to clinical care.

• **THE IMPACT OF DIGITAL IMAGING:** The incorporation of digital imaging has provided a number of opportunities not only to improve ROP care but also to advance ROP research.[9,24-28] Clinically, it has enabled the ability to photodocument disease for the medical record and to compare over time, which has proven to be an important capability.[29,30] It has also enabled remote "tele-ROP" monitoring systems, such as SUNDROP[31] in the United States, KIDROP,[9] and the Aravind Eye Care System (AECS) ROP telemedicine programs in India, which have been force multipliers in enabling ROP screening especially in rural and low-income regions.[28,32,33]

Imaging has also led to a better understanding of the patterns of ROP diagnosis (and misdiagnosis). Although a full review of this area is beyond the scope of this paper, the early work exploring interobserver variability in plus disease is relevant. Because ETROP changed practice and elevated the importance of plus disease, it became more important to understand how this diagnosis was made in practice. Despite a photographic standard for plus disease, analysis of expert diagnostic patterns revealed a high degree of disagreement.[24,34] Subsequent qualitative analysis yielded insights that not only did clinicians disagree on the outcome, they often disagreed on the process of how to make the diagnosis.[35]

Over time, larger imaging data sets enabled larger studies of the pattern of expert behavior, and it became clear that a key underlying factor was that plus disease was a spectrum, with some experts over- and underdiagnosing compared with their colleagues, without any objective standard to apply.[36,37] It further became apparent that this bias was leading to real-world treatment differences,[27,38] complicated efforts to train ophthalmologists,[39] and may be subject to temporal drift—that is, the level of disease at which the average clinician diagnoses "plus" might be changing over time.[25] The presence of large data sets of ROP images also emerged as interest in machine learning and artificial intelligence (AI) in medicine was rapidly evolving, which led to attempts to classify and quantify plus disease.

• **APPLICATIONS OF AN ROP VASCULAR SEVERITY SCORE:** There have been multiple groups and methodological approaches focused on using AI for the diagnosis of plus disease.[40-43] The ROPTool was one of the earliest approaches that used a semi-autonomous method integrating user input and output quantitative assessments of tortuosity and dilation, which could then be compared with clinical diagnoses.[40,42] Subsequent work using feature extraction methods led to a classifier that could diagnose plus disease at least and clinical experts, but it required manual segmentation of the retinal vessels as input.[43] In 2018, using the Imaging and Informatics in ROP (i-ROP) data set, the i-ROP deep learning (i-ROP-DL) algorithm used a convolutional neural network (CNN) to classify 3-level plus disease, including as international experts, without requiring manual segmentation by training a separate network to extract the vessel segmentation.[44]

The i-ROP-DL algorithm was developed based on reference standard input images for no plus, pre-plus, and plus and outputted not only a diagnosis but also a 3-level probabilistic output that summed to 1. Using this output, the team converted the 3 probabilities (P) into a weighted "vascular severity score" (VSS) from 1 to 9, based on the rough equation 1 x P ("no plus") + 5 x P ("pre-plus") + 9 x P ("plus"). This resulted in a 1 to 9 scale where 1 translated to thin, narrow, and straight vessels and 9 reflected severe plus disease, typically AROP.[45,46] Validation studies demonstrated not only that it corresponded to the spectrum

of plus disease among experts,[47] but also that it increased with higher stage, higher extent of stage 3, and more posterior zone.[45] As a result, a single photo could yield a VSS that corresponded to the overall ICROP classification of an eye.[46,48,49] Subsequently, the VSS concept has been used for a number of different applications, including assistive diagnosis,[50] autonomous screening,[51] longitudinal disease monitoring,[46,48] and integrated risk prediction.[52,53]

For the purposes of this study, 2 specific applications of the VSS will be analyzed. First, although it is not surprising that the VSS correlates with plus disease diagnostic labels—indeed, it was trained on them—it was a novel finding to recognize that it correlated with more posterior zone and higher disease stage, all biomarkers of more severe ROP. However, the original description of this finding was published based on the same data set from which the algorithm itself was developed and reflected only a North American population. In this study, we chose to analyze a second data set of images from the AECS, Coimbatore ROP telemedicine program that reflects some of the phenotypic variability of both second and third epidemic ROP.

Second, because the VSS reflects a surrogate of the level of ROP in an infant, it also works for population-level epidemiologic assessment by averaging individual baby-level data into neonatal care unit (NCU) level or regional data.[54-56] Previous work demonstrated the value of an "institutional" VSS that could be used to compare the severity of ROP geographically and over time.[54-56] In the United States, these differences were associated with underlying demographic risk (mean GA and birthweight of various NCUs),[56] whereas in an older AECS population (2015-2017), we found stronger associations with local NCU resources, such as the presence and use of oxygen blenders and pulse oximeters.[55] We further found that the population-level VSS correlated with changing ROP epidemiology over time in South India between 2015 to 2017 and 2019 to 2020, strongly suggesting improvements in neonatal care in that region over a relatively short period of time.[54]

In this study, we use an eye-level VSS to compare differences in the quantitative relationship between VSS and ICROP classification across the 2 populations: the i-ROP cohort and the more recent AECS cohort. The central hypothesis is that the underlying relationship between vascular severity and higher disease stage and more posterior disease zone remains intact, despite differences in underlying ROP epidemiology and phenotypic characteristics.

METHODS

• **ETHICAL APPROVALS:** This study followed the tenets of the Declaration of Helsinki and was approved by the institutional review boards (IRB) at Oregon Health & Science University and participating i-ROP sites, as well as by the Aravind Eye Care System (AECS) IRB before the study began for prospective data collection. This specific study is a retrospective investigation of the existing data set and was covered by the original IRB at all sites. Analyses for this study were performed on deidentified clinical and imaging-derived variables. Informed consent was obtained from the parents of all babies in both countries, and no financial incentives were provided.

• **DATA SETS:** We analyzed 2 independent data sets: the i-ROP data set and the AECS data set.

i-ROP

From January 2012 to July 2020, the i-ROP Consortium assembled a longitudinal data set of serial retinal fundus images from infants undergoing routine ROP screening (GA < 31 weeks, birth weight [BW] <1501 g). Images were acquired using a RetCam3 (Natus) with 5 standard fields of view captured per session (posterior, nasal, temporal, inferior, and superior). Bedside ophthalmoscopic examinations were performed in parallel with standardized image review by expert readers, and a reference-standard diagnosis was derived by integrating these assessments, as previously described.[57] Images judged not acceptable for diagnosis by reader consensus, examinations diagnosed with stage 4 or 5 ROP, and post-treatment examinations were excluded.

AECS

The AECS data set was collected through a telemedicine screening program in India. Between August 2015 and October 2017, tele-medical examinations were performed using a platform designed for diabetic retinopathy telescreening. In 2018, iTelegen, a telemedicine and data management system specifically designed for ROP, was developed and implemented, with a full transition occurring in 2019. Therefore, the AECS data set was split into 2 time periods based on which system was used: August 2015 to October 2017 and February 2019 to December 2020, the latter of which was used for the primary analysis and labeled "AECS," unless otherwise specified. The 2015-2017 and 2019-2020 data sets screened infants across 18 and 48 neonatal care units (NCUs), respectively. Trained technicians acquired images using either a RetCam Shuttle (Natus) or a Forus 3nethra (Forus) from infants meeting the Indian screening guideline criteria (GA < 34 weeks, BW < 2000 g), capturing multiple retinal fields of view. For this data set, the telemedical diagnoses were used as the ground truth. The exclusion criteria were analogous to those of i-ROP.

• **QUANTITATIVE VSS ANALYSIS:** Vascular severity was quantified using an AI-derived VSS from the i-ROP-DL algorithm, using methods previously published and illustrated in Figure 1.[45] Specifically, after automated image quality assessment, we averaged the image-level VSS output from all available images for each eye examination, yielding a single eye-level VSS per examination.

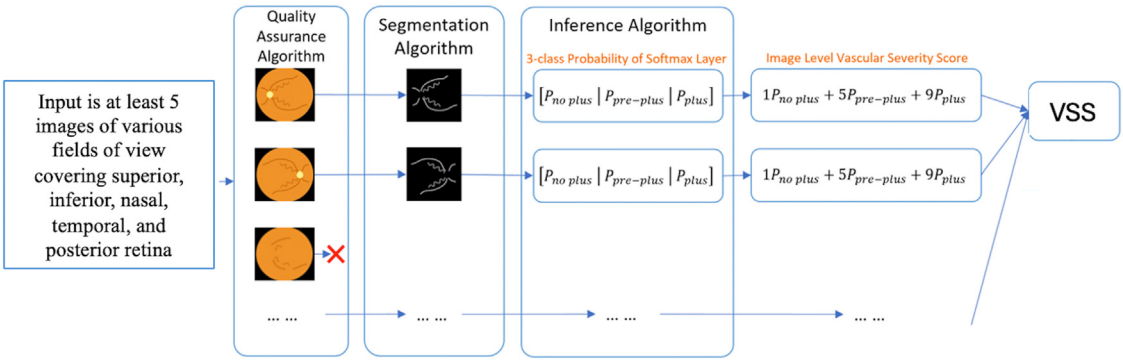

FIGURE 1. Process for creation of an eye-level VSS. All images from an eye examination are input into the algorithm. A quality step is performed as a preprocessing step that identifies the presence of an optic nerve in the image. This ensures the media are clear enough to visualize the disc, the image reflects the back of the eye, and the field of view is relatively posterior. A segmentation step is applied to images that pass the quality check. The algorithm then assigns probabilities to no plus, pre-plus, and plus, which are converted to a 1 to 9 scale for each image. All images are then averaged for an eye-level VSS. VSS = vascular severity score.

• **CROSS-SECTIONAL ANALYSIS:** Within each data set, we tested cross-sectional associations between examination-level VSS and ICROP definitions for zone, stage, plus, and the extent of stage 3. We evaluated associations of VSS with plus disease (normal, pre-plus, plus) and ROP category (no ROP, mild ROP, type 2 ROP, or any ROP with pre-plus, TR-ROP), and stage (0-3) after stratifying by zone (I-II). We also evaluated associations of VSS with clock-hour extent of stage 3 ROP, categorized as <3, 3 to 6, and >6 hours. Extent was measured during the ophthalmoscopic examination in the i-ROP data set and by manual image review by a physician with expertise in diagnosing ROP in India via telemedicine (SN).

• **LONGITUDINAL ANALYSIS:** Within each data set, we evaluated longitudinal patterns in examination-level VSS across PMA. PMA was binned into 2-week intervals from 30 to 40 weeks. Longitudinal models included PMA bin, treatment status (treatment-requiring [TR] vs non-TR), and their interaction to assess whether VSS trajectories differed over time between treated and untreated eyes.

• **STATISTICAL ANALYSIS:** All analyses were performed using R (version 4.5.1). The unit of analysis was either a patient or an eye exam. Baseline demographic characteristics were compared between cohorts at the infant level using Wilcoxon rank sum tests and 2-sample proportion tests, where appropriate. All inferential analyses used generalized estimating equations (GEE) to account for correlation from repeated exams within infants (and both eyes when available). Gaussian family models with an identity link and robust (sandwich) standard errors were fit, clustering on infant. Omnibus Wald tests and prespecified pairwise contrasts estimated from model-based marginal means with Holm adjustment for multiple comparisons were used. Statistical significance was defined as $P < .05$.

RESULTS

• **PATIENT-LEVEL DEMOGRAPHICS AND EXAMINATION-LEVEL FINDINGS:** Tables 1 and 2 highlight the patient-level and examination-level demographics in the 2 data sets for both TR and non-TR infants. The i-ROP data set included images from 6552 examinations from 1019 babies, and the AECS data set included images from 8281 exams from 2363 babies. Table 1 demonstrates substantial between-cohort differences in examination-level disease severity and topography. Overall, eyes in the AECS cohort were more likely to be in zone III on examination ($P < .001$), whereas i-ROP had a higher proportion of zone I eyes ($P < .001$). The proportion of eyes with TR-ROP was higher in i-ROP (201/6552 [3.2%]) compared with AECS (123/8281 [1.5%], $P < .001$). Table 2 demonstrates the demographic variables for TR and non-TR babies in both data sets. There are several notable differences. The mean GA and BW were higher in the AECS data set in both TR and non-TR groups and were significantly lower in TR babies in both data sets.

• **CROSS-SECTIONAL ANALYSIS:** Figure 2 demonstrates the relationship between the eye-level VSS and the ICROP classification of zone, stage, extent of stage 3 disease, plus disease, and overall category. There was a direct relationship between the VSS and the clinical diagnosis of stage in both data sets (Figure 2, A, all $P < .001$). In i-ROP, we found that for stages 1 ($P = .044$), 2 ($P < .001$), and 3 ($P < .001$), VSS was higher in more posterior zones (Figure 2, B). This finding held for stages 1 ($P < .001$) and 3 ($P < .001$) in the AECS data set. Traditionally, ROP classification has recorded the "extent" (number of clock hours) of each stage. In these data sets, we only had labels for the extent of stage 3 but found that the VSS increased with a higher extent of stage 3 in both i-ROP ($P = .004$) and

TABLE 1. Examination-Level Distribution of Zone, Stage, Plus, and Stage 3 Extent in i-ROP and AECS

Diagnosis	i-ROP, *n (%)*	AECS, *n (%)*
Category		
None	2683 (42.3)	6843 (83.0)
Mild	2328 (36.7)	981 (11.9)
Moderate	1138 (17.9)	296 (3.6)
TR	201 (3.2)	123 (1.5)
AROP[a]	46 (0.7)	84 (1.0)
Plus		
Normal	5260 (80.3)	7931 (95.8)
Pre-Plus	1086 (16.6)	306 (3.7)
Plus	206 (3.1)	44 (0.5)
Stage by zone		
Zone I	**335 (5.1)**	**88 (1.1)**
Stage 0	49 (0.7)	42 (0.5)
Stage 1	79 (1.2)	5 (0.1)
Stage 2	104 (1.6)	13 (0.2)
Stage 3	103 (1.6)	28 (0.3)
Zone II	**6106 (93.2)**	**3398 (41.0)**
Stage 0	2624 (40.0)	2066 (24.9)
Stage 1	1545 (23.6)	541 (6.5)
Stage 2	1571 (24)	682 (8.2)
Stage 3	366 (5.6)	109 (1.3)
Zone III	**111 (1.7)**	**4795 (57.9)**
Stage 0	90 (1.4)	4696 (56.7)
Stage 1	19 (0.3)	87 (1.1)
Stage 2	2 (0.1)	12 (0.1)
Stage 3	0 (0.0)	0 (0.0)
Extent of stage 3		
<3 clock hours	253 (53.9)	75 (54.7)
3–6 clock hours	191 (40.7)	58 (42.3)
>6 clock hours	25 (5.3)	4 (2.9)

[a]TR and AROP categories may overlap.
AECS = Aravind Eye Care System.

AECS ($P < .001$), as found in Figure 2, C. Unsurprisingly, there was a relationship between the VSS and the clinical diagnosis of plus disease (Figure 2, D, indeed, the VSS algorithm was developed from the plus labels of the i-ROP data set), but confirmation in a new data set is novel. Finally, when we evaluated VSS across the overall category of disease (none, mild, type 2/preplus, or type 1/treatment requiring), VSS increased with worsening ROP in both data sets (Figure 2, E, both $P < .001$). Interestingly, at all levels of comparison, the VSS was lower in the AECS data set compared with i-ROP, including stage (all $P < .001$, Figure 2, A), plus disease (all $P < .001$, Figure 2, D), and category ("None" $P = .002$, all other $P < .001$, Figure 2, E).

• **LONGITUDINAL ANALYSIS:** Longitudinal analysis demonstrated that eyes that were eventually treated had higher average VSS at all time points in both data sets (Figure 3). In i-ROP, eyes that eventually required treatment had higher VSS than eyes that did not in every PMA bin (all $P < .001$). A similar trend was found in AECS (PMA [38, 40), $P = .046$; all other $P < .001$). Figure 4 demonstrates the relationship between GA and PMA in eyes that developed TR-ROP in both data sets. There was a significant difference in underlying GA for babies who developed TR-ROP across all time points in this comparative analysis (all $P < .001$). The mean ± standard deviation age at treatment was higher in the i-ROP data set (11.8 ± 3.8 weeks of life) compared with AECS (6.3 ± 3.4 weeks of life, $P < .001$).

DISCUSSION

In this paper, we evaluate the relationship between an AI-derived VSS and the full ICROP classification (zone, stage, plus) in 2 distinct populations. There were several key findings. First, the epidemiology of ROP, including the relative risk of treatment for a given GA at birth, differs between the populations. This was assessed by reviewing demographics and treatment outcomes, as well as by evaluating population-level VSS characteristics. Second, in both populations, a higher VSS was associated with a higher disease stage, a more posterior zone, and a greater extent of stage 3, despite differences in disease epidemiology and phenotypes. Third, we identified population-level differences between the VSS and the overall stage/category of ROP in the 2 populations that merit further investigation—specifically, the VSS was lower in the Indian population for a given zone, stage, and extent of disease.

TABLE 2. Demographics of TR ROP in Both Populations

	Not Treated			Treated		
	i-ROP	AECS	*p* value	i-ROP	AECS	*p* value
Patients, *n (%)*	865 (84.9)	2289 (96.9)	<0.001	154 (15.1)	74 (3.1)	<0.001
Gestational age, *mean weeks ± SD*	27.7 ± 2.3	33.5 ± 2.8	<0.001	24.9 ± 1.3	30.4 ± 2.5	<0.001
Birthweight, *mean weeks ± SD*	1017 ± 320	1732 ± 438	<0.001	661 ± 173	1291 ± 291	<0.001

AECS = Aravind Eye Care System.

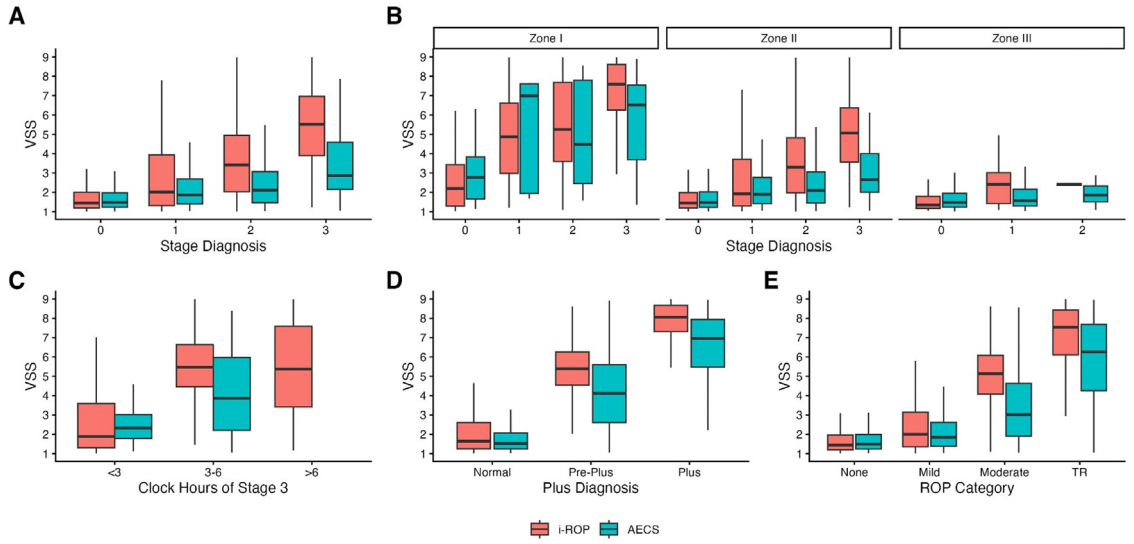

FIGURE 2. Cross-sectional comparisons of i-ROP and AECS data sets across ROP risk factors. (A) VSS increased with worsening ROP stage in both data sets (all $P < .001$), with consistently lower VSS in AECS than i-ROP at a given stage (all $P < .001$). (B) For a given stage, VSS tended to be lower in more posterior zones. In i-ROP, this was significant for stage 1 ($P = .044$), 2 ($P < .001$), and 3 ($P < .001$), and in AECS for stages 1 ($P < .001$) and 3 ($P < .001$). (C) Among stage 3 examinations, VSS increased with greater clock-hour extent of disease (i-ROP $P = .004$, AECS $P < .001$). (D) VSS increased with plus disease severity (all $P < 0.001$) and (E) overall ROP severity, with i-ROP demonstrating higher VSS than AECS within severity strata (all $P < .001$). AECS = Aravind Eye Care System; VSS = vascular severity score.

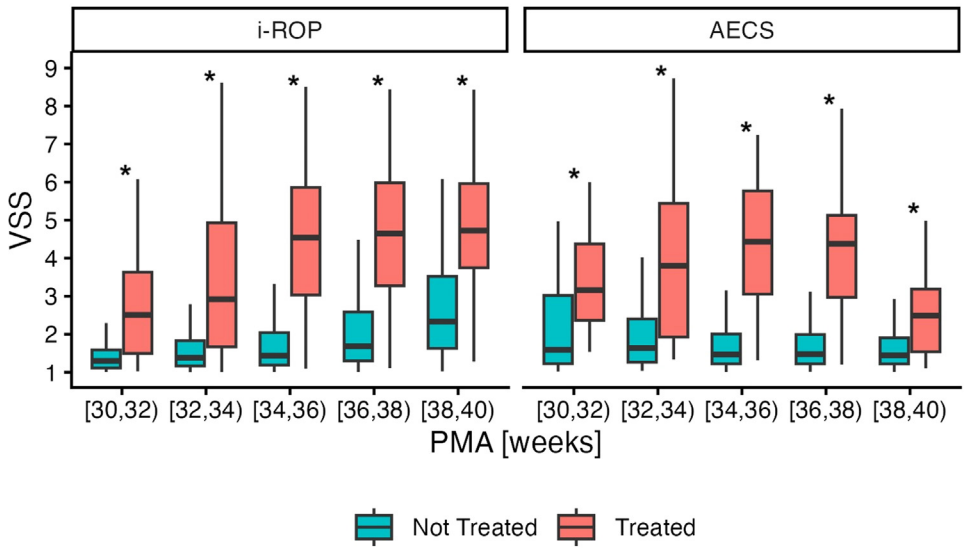

FIGURE 3. Longitudinal analysis of vascular severity scores in treated vs non-treated eyes in both data sets. In both data sets, eyes that eventually required treatment had higher VSS at all time points (all $P < .001$, except AECS PMA [38,40] $P = .046$).

• **GEOGRAPHIC AND TEMPORAL TRENDS IN ROP EPIDEMIOLOGY:** These data highlight some of the key differences in disease epidemiology between the United States and India, which are illustrative of high- and low-middle-income countries in general. The population being screened in India was older and heavier at birth, which should imply lower demographic risk, and treated at a younger postnatal age. There is a circularity here as the screening criteria are different (<34 weeks and 2 Kg in India vs <30 weeks and 1.5 Kg in the United States).[58] This explains both the higher BW/GA in the AECS population, as well as the higher proportion of examinations in zone III with stage 0 (as the retina is more vascularized at the time of birth in an older population). But the screening criteria are different because the risk of severe ROP is higher in India for a given BW/GA. This can be found in Table 2 which demonstrates that the mean GA at birth of eyes that develop TR-ROP in AECS remains >30 weeks, which is barely within US

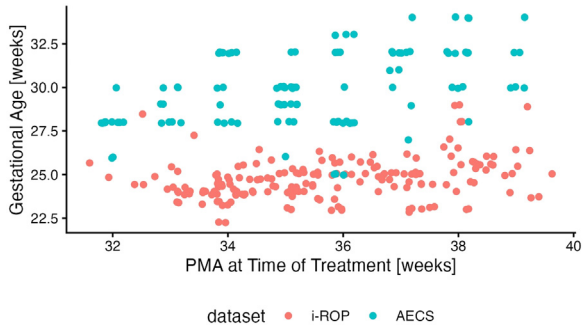

FIGURE 4. Scatterplot of gestational age at birth (GA) vs PMA at time of treatment. There was a significant difference in underlying GA for babies who developed treatment-requiring ROP ($P < .001$). This translated to a mean ± standard deviation (SD) post-natal age at time of treatment in i-ROP of 11.8 ± 3.8 weeks compared with the 6.3 ± 3.4 weeks in AECS ($P < .001$). AECS = Aravind Eye Care System; GA = gestational age; PMA = postmenstrual age.

screening criteria, and, as Figure 3 reveals, treatment occurs at a much earlier time point relative to birth.

It is well established that these observed epidemiologic differences are due in large part to resource constraints that make optimization of oxygen exposure challenging.[1,7] Higher exposure to oxygen than is necessary for resuscitation and life increases the population at risk by increasing the relative risk of ROP for a given BW/GA. This means that regions such as India, which already have a higher proportion of premature births, also have a higher risk of ROP for each preterm baby, and fewer ROP clinicians. AECS, along with the KIDROP team, have made remarkable progress utilizing ROP telemedicine to improve the effective screening rate across India in the last 10 to 20 years.[8,33,54,59] Perhaps most significantly, there is evidence that these programs are having a measurable impact on reducing the incidence of severe ROP.[54] Specifically, in the AECS ROP telemedicine program, in an approximately 5-year period from 2015 to 2020, the relative risk of developing moderate to severe ROP dropped by more than 50% for a baby of any given BW/GA, due presumably in large part to education, feedback, and optimization of early oxygen exposure.[54] Moreover, this was measurable not only by manually reviewing the chart for treatment records but also by using quantitative AI-based approaches to measure individual, neonatal-care unit level, and regional levels of ROP. Nonetheless, the persistent differences in ROP risk between the incidence of severe ROP for a given BW/GA and the timing of the onset of severe disease relative to birth suggest that there are ongoing opportunities to reduce the geographic health disparities in ROP risk.

• **VSS REFLECTS DISEASE SEVERITY IN BOTH POPULATIONS:** Despite these differences, the fundamental relationship between vascular severity and the zone, stage, and extent of stage 3 remained consistent in both data sets. Specifically, the higher the stage, the higher the vascular severity (Figure 2). This relationship was first described a few years ago[45,47] and has several potential explanations, including increased shunting with a higher stage leading to higher flow, as well as the direct effects of VEGF. The more interesting relationship is the association with a more posterior zone and higher VSS. This also raises the question of a hemodynamic explanation vs a VEGF-driven explanation, because a more posterior zone implies more avascular retina secreting VEGF. Without the ability to directly measure VEGF in vivo, differentiating these explanations has been challenging, though it is interesting to hypothesize how and why a more posterior zone might lead to higher retinal blood flow and VSS for a given degree of ROP stage. Future work involving hemodynamic modeling may be helpful to better understand this relationship.[60]

• **DIFFERENCES IN VASCULAR SEVERITY BETWEEN POPULATIONS:** The last key finding was that, on average, the measured vascular severity, for any degree of ROP, was lower in India compared with the United States. There are several potential explanations for this, which are challenging to tease apart. First, with the higher GA at baseline, the mean "zone II" might be more anterior in the India data set than in the US data set, which is associated with lower VSS, as described previously. In the future, automated methods of measuring the extent of retinal vascularization or the area of vascularized retina may be helpful.[61] A second potential association, related to the underlying differences in neonatal care, is that higher oxygen exposure, on average, might lead to phenotypic differences characterized by lower VSS on average. Indeed, it is well established that AROP is due, in part, to higher oxygen exposure. Given the remaining differences in epidemiology and demographic risk, it may be that, for all zone/stage/plus combinations, the vessels are slightly thinner and less tortuous compared with disease in the United States. Figure 5 highlights 2 examples of AROP that illustrate some of the similarities and differences between the historical zone I APROP and the more heterogeneous AROP phenotypes.

There are several sources of bias that could also produce these results. For example, approximately half of the AECS data set used a different camera system (Forus 3Nethra), which in theory could introduce bias in the VSS calculation, although prior work evaluating autonomous ROP screening by the i-ROP-DL algorithm on the 3Nethra did not yield false negatives (which might occur with a camera-specific systematic bias), suggesting that this is not likely the main explanation.[51,62] Furthermore, it is well established that there can be both individual and group-related differences in disease classification that affect research comparisons.[34,63,64] This has been well described in prior literature, both for plus disease and for stage, even among the ICROP committee members,[47] in a way that has affected prior ROP multinational clinical trials.[27] If, for example, on average

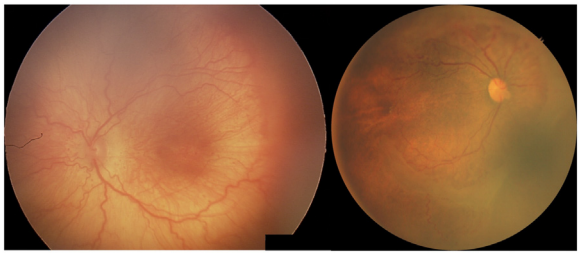

FIGURE 5. Examples of AROP in the i-ROP (left) and AECS (right) data sets. There are several notable differences between the epidemiology of these two babies and the disease phenotypes. The baby on the left was born at 24 weeks gestational age, whereas the baby on the right was born at 35 weeks gestational age. At that age, AROP only occurs in the setting of unregulated oxygen. The left image reflects the traditional zone I appearance of "APROP" with plus disease out of proportion to stage appearance, strongly suggesting the presence of reticulated networks of flat neovascularization. The right image on the other hand, while it appears to be relatively posterior, has a vascular loop inferiorly that suggests the retina was originally vascularized more anteriorly, and then experienced significant vaso-occlusion of the retina leading to these large peripheral shunt vessels and a more posterior appearance. Also note that qualitative assessment of the vascular severity would rank the left image higher than the right, suggesting that quantitative approaches might have the same bias resulting in lower measured vascular severity in regions where oxygen exposure is higher. The measured vascular severity score of the left image was 8.6, compared with 6.9 on the right. AROP = aggressive retinopathy of prematurity.

AECS clinicians diagnose disease as stage 2 that might have been diagnosed as stage 1 in the i-ROP data set, this would result in a lower VSS for the same actual disease level. This could be assessed through a centralized reference standard diagnosis process in future work. It may also be possible to apply quantitative methods to the image-based diagnosis of stage and directly compare clinical labels to quantitative assessments of stage in the future.[65–67]

• **A UNIFYING HYPOTHESIS:** There is a plausible unifying hypothesis that explains not only the underlying relationship between VSS and zone and stage but also the relative differences between the 2 populations. If the changes associated with the spectrum of plus disease, and therefore the VSS, correlate with the intraocular VEGF concentration, then both key findings are explainable and suggest that the VSS might be a surrogate biomarker for intraocular VEGF concentration. There are several lines of evidence to suggest that this hypothesis is plausible. First, although the ability to assay for VEGF in vivo in babies is limited, there are data to suggest that systemic and aqueous VEGF are higher in babies with more severe ROP, including AROP.[68-72] Second, the clinical response to anti-VEGF agents strongly suggests that a VEGF-mediated mechanism at least partially explains plus disease, especially the improved outcomes in eyes with AROP. We have previously revealed that the VSS decreases rapidly within days post-treatment,[73] and faster with anti-VEGF than with laser,[49] and rises again with reactivation requiring retreatment,[73] all of which make sense if there is a high correlation between VSS and VEGF concentration. Thus, the relationship between VSS and more posterior ROP, higher disease stage, and higher extent of stage 3, and the highest in APROP/AROP might make sense if VEGF and VSS are highly correlated.[45-47]

The second part of the hypothesis relates to the underlying 2-phase model for ROP and the relative implications of oxygen and VEGF levels in each phase. Specifically, although early oxygen exposure (in phase I) leads to more vasoconstriction and capillary injury, the relative oxygen supply and demand reverse in phase II, where ischemic avascular retina leads to higher VEGF release. Increasing oxygen exposure in phase II, in theory, may be therapeutically beneficial by reducing VEGF secretion and causing neovascularization to recede. There are preclinical data from the oxygen-induced retinopathy model suggesting this is true.[74-76] Furthermore, this was the hypothesis for the STOP-ROP trial, which randomized eyes with threshold ROP to higher or lower oxygen concentrations. Although the study failed to meet its primary end point, it did suggest an ocular benefit in the subgroup of pre-threshold eyes (more similar to what would become known as type 1 after ETROP), which has led to ongoing speculation and individualized treatment regimens at some institutions.[77-79]

The heterogeneity in oxygen exposure in the AECS population thus creates a natural experiment where, on average, the AECS population has higher oxygen exposure in both phase I and phase II ROP. Higher exposure in phase I leads to a higher population at risk, higher incidence and severity of ROP for a given GA, and the overall differences in epidemiology described previously (and illustrated in Table 2 and Figures 4 and 5). Higher exposure in phase II, however, might lead to lower VEGF, on average, in staged ROP and lower VSS for a given zone/stage combination. Testing this hypothesis directly will be challenging; however, there are several testable predictions that would support it. (1) Analysis of populations demographically similar to the United States will have VSS distributions similar to the i-ROP cohort, and vice versa. (2) Image-based analysis of a modern "STOP-ROP like" intervention of selected intentional hyperoxia in phase II would reveal reduced VSS in type 2 ROP in the intervention group compared with the control. (3) Over time, these differences will diminish as neonatal practice differences between high and low-middle-income countries become more similar.

We went back to older AECS data to test this third hypothesis. Previous work had demonstrated that the proportion of eyes in the region around AECS, Coimbatore, developing severe ROP had decreased in the time frame between 2 cohorts (2015-2017 and 2020-2021), strongly suggesting improvements in neonatal care in that short time

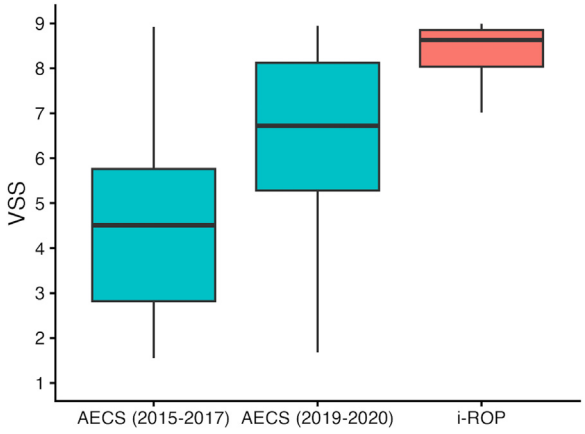

FIGURE 6. VSS for 3 populations of eyes with AROP. In this post hoc analysis, we compared the VSS for all eyes diagnosed with "APROP" or "AROP" adding 1 additional population, and older data set of images from the AECS. Left: AECS 2015-2017. Middle: AECS 2019-2020. Right: i-ROP 2012-2020. This was done to test the hypothesis that the differences between populations might decrease over time with improvements in neonatal care, such as what has been occurring in the AECS population in the last decade. All pairwise comparisons were significantly different, $P < .001$, confirming this hypothesis, and supporting the hypothesis that the VSS likely correlates with VEGF activity in the eye, which is lower in the setting of higher oxygen exposure in phase II. AECS = Aravind Eye Care System; AROP = aggressive retinopathy of prematurity; VSS = vascular severity score.

period. This provided a natural opportunity to test the hypothesis that these differences in VSS would diminish over time. Figure 6 demonstrates the distribution of VSS for eyes diagnosed with APROP/AROP in the original "old" AECS cohort vs the more recent AECS cohort vs the i-ROP cohort. As predicted, the VSS distribution for eyes with APROP/AROP was significantly lower in the earlier time frame, which had less well-regulated oxygen exposure, and thus the more recent AECS cohort is more similar to the i-ROP cohort than it had been 5 years earlier ($P < .001$ for all comparisons). These data are supportive of the hypothesis that the lower VSS is due at least in part to higher oxygen exposure in phase II ROP, presumably mediated by oxygen-induced suppression of intraocular VEGF levels.

• **LIMITATIONS:** There are several limitations to this analysis. First, the VSS algorithm was developed on the i-ROP data set, based on the plus disease labels, so the association between VSS and plus disease is definitional, rather than a new association. However, the associations with zone, stage, extent, and all the findings in the AECS are novel. Second, as described, the difference in camera type in the India data set may affect the results in ways that are difficult to determine, as we did not have accurate labels in the data set for the camera used for each examination. Third, the lack of a reference standard diagnosis in the AECS data set (which has single clinician labels only) might obfuscate true diagnostic similarities or differences, given the wide interobserver variability described previously. Fourth, although in this paper we generalize to differences between a US population and an Indian population, it is not accurate to generalize to the entire country, where the epidemiology and demographic risk remain highly heterogeneous. If anything, the AECS population might be more similar to the US population than other regions of India, given the documented recent improvements in ROP incidence in that region, specifically.[54] Finally, whereas the AECS data set was a true population-level data set of all babies within those given neonatal centers during that period of time, the i-ROP data set was part of a larger research study involving genetics, imaging, and informatics analysis and thus did not include all babies within the population.

CONCLUSION

In this paper, we have reviewed the epidemiology of ROP geographically and over time and the impact that digital imaging has on our understanding of ROP diagnosis, including its influence on the evolution of ICROP over time. It also highlights several ways in which ROP epidemiology, classification, imaging, and AI might interact in the future.

ROP epidemiology. As neonatal care continues to improve worldwide, our shared goal is to reduce geographic disparities in ROP risk and blindness through improved primary prevention, consistent ROP screening, and timely treatment. This paper adds to the evidence base that not only are we seeing this improvement occur in the last decade, but we can also measure and quantify the differences using technologies like the AI-derived VSS used here. But this is only half of the story—it is very likely that the implementation of the AECS ROP telemedicine program itself, which enabled a 30,000-foot view of the heterogeneity of outcomes between NCUs in the same region,[54,55] and 1:1 communication and education as a result, directly influenced the improvements in neonatal care in that region.[54,55] Moreover, from a global ophthalmology perspective, given the significantly larger population at risk and fewer ophthalmologists, imaging-enabled ROP telemedicine will be key to reducing the number of babies who become blind from ROP, as there are simply not enough ophthalmologists for in-person ROP screening to work effectively.[80,81]

ROP pathophysiology. Advances in ophthalmic imaging have yielded pathophysiologic insights into nearly every (adult) retinal disease; ROP will be no exception. As future technologies, such as AI, ultra-widefield cameras, and OCT, become more widely used in research and clinical settings, we will learn more about

ROP physiology. For example, this paper argues that higher exposure to oxygen in phase II ROP leads to lower vascular severity (less plus disease), suggestive of a lower VEGF concentration. This finding does not necessarily mean that, as STOP-ROP hypothesized, ROP structural and functional outcomes will be better—an issue that should continue to be investigated through controlled clinical trials. It does suggest, however, that we ought to incorporate imaging and quantitative image analysis in such studies and continues to raise important questions about the fundamental balance of oxygen, ROP, and mortality in prematurity.

ROP classification. The members of the 2021 ICROP 3 committee were committed to formalizing a classification system that could be applied anywhere in the world with an ophthalmologist and an indirect ophthalmoscope. This goal remains admirable, especially given how prevalent ROP is and will continue to be worldwide. However, there are very few diseases in ophthalmology, or medicine in general, where our classification schema continues to depend on physical examination alone. Generally, disease diagnosis and classification are slowly moving from subjective and qualitative to objective and quantitative, and ROP will likely follow the same course, with advances in AI, OCT, and multimodal integrated informatics risk modeling identifying unique "phenotypes" of disease with specific prognostic implications. Future ICROP committees will need to evaluate the evidence for the added value of imaging and other multimodal data with respect to disease classification; however, the increased use of ophthalmic imaging for ROP screening, diagnosis, and research will certainly continue to lead to an evolution in ICROP, as it has already.

In summary, we specifically compared the relationship between the VSS and the overall severity of ROP in an eye within and between populations. We found that a higher VSS was associated with more posterior ROP, a higher stage of ROP, and a higher extent of stage 3, and that the VSS was lower in the AECS population compared with iROP. Integrating all the findings together, we presented a unifying hypothesis, based on several levels of evidence, that the VSS may correlate with intraocular VEGF concentration and thus represents an imaging biomarker of the severity and risk of ROP progression. Finally, we presented several testable hypotheses for future evaluation of this thesis and were able to confirm one of them by looking back at prior AECS data, which demonstrated that as neonatal care improves, the populations are becoming more similar over time. It is an important reminder that although we all endeavor to optimize the ophthalmological care of patients at risk for ROP-related blindness, as we learned from the first epidemic of ROP, the most impactful thing we can do is work with our neonatology colleagues to reduce geographic disparities in the incidence of severe ROP in all preterm babies.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

J. PETER CAMPBELL: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing. **SAMEERA NAYAK:** Data curation, Formal analysis, Writing – review & editing. **SUSAN OSTMO:** Data curation, Methodology, Project administration, Writing – review & editing. **AARON S. COYNER:** Data curation, Formal analysis, Investigation, Methodology, Validation, Writing – review & editing. **PRAVEER SINGH:** Data curation, Formal analysis, Writing – review & editing. **JAYASHREE KALPATHY-CRAMER:** Conceptualization, Formal analysis, Methodology, Writing – review & editing. **PARAG SHAH:** Conceptualization, Data curation, Funding acquisition, Project administration, Writing – review & editing. **RV PAUL CHAN:** Conceptualization, Formal analysis, Resources, Writing – review & editing. **MICHAEL F. CHIANG:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing – review & editing.

Funding/Support: This work was supported by grants R01 EY019474, R01 EY031331, and P30 EY010572 from the National Institutes of Health (Bethesda, MD), by unrestricted departmental funding and a Career Development Award (JPC) from Research to Prevent Blindness (New York, NY), and the Malcolm Marquis Innovation Fund. These funding organizations had no role in the design or conduct of this research.

Financial Disclosures: The i-ROP DL system has been licensed to Siloam Vision by Oregon Health & Science University and Massachusetts General Hospital, which may result in royalties for Drs. Chan, Campbell, Coyner, and Kalpathy-Cramer in the future. Drs. Campbell, Chan, and Kalpathy-Cramer are equity owners of Siloam Vision. Dr. Coyner is a consultant for Siloam Vision. Dr. Chan is a consultant for Alcon (Fort Worth, TX), AbbVie (Chicago, IL), and EyePoint Pharmaceuticals (Watertown, MA). Dr. Campbell was previously a consultant to Ikerian AG (Bern, Switzerland). The author attests meeting the current ICMJE criteria for authorship.

Acknowledgments: Members of the i-ROP Research Consortium were essential to the development of the i-ROP database used in this study. The AECS data set would not have been possible without the dedication and work of the entire AECS ROP telemedicine team, including photographers, coordinators, and readers. The authors also acknowledge Will Goldring and Karyn Jonas for their contributions to developing the iTelegen system.

Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper, other than what is disclosed above.

REFERENCES

1. Gilbert C, Malik ANJ, Nahar N, et al. Epidemiology of ROP update—Africa is the new frontier. *Semin Perinatol*. 2019;43(6):317–322. doi:10.1053/j.semperi.2019.05.002.
2. Wong ES, Choy RW, Zhang Y, et al. Global and regional trends in retinopathy of prematurity. *JAMA Ophthalmol*. 2026;144(2):139–147. doi:10.1001/jamaophthalmol.2025.5103.
3. Owens WC, fibroplasia Owens EURetrolental. *Am J Public Health Nations Health*. 1950;40(4):405–408. doi:10.2105/AJPH.40.4.405.
4. Hartnett ME. Advances in understanding and management of retinopathy of prematurity. *Surv Ophthalmol*. 2017;62(3):257–276. doi:10.1016/j.survophthal.2016.12.004.
5. Kim SJ, Port AD, Swan R, Campbell JP, Chan RVP, Chiang MF. Retinopathy of prematurity: a review of risk factors and their clinical significance. *Surv Ophthalmol*. 2018;63(5):618–637. doi:10.1016/j.survophthal.2018.04.002.
6. Chen JS, Anderson JE, Coyner AS, et al. Quantification of early neonatal oxygen exposure as a risk factor for retinopathy of prematurity requiring treatment. *Ophthalmol Sci*. 2021;1(4):100070. doi:10.1016/j.xops.2021.100070.
7. 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(2):77–82. doi:10.1016/j.earlhumdev.2007.11.009.
8. Shah PK, Narendran V, Saravanan VR, et al. Fulminate retinopathy of prematurity—clinical characteristics and laser outcome. *Indian J Ophthalmol*. 2005;53(4):261–265. doi:10.4103/0301-4738.18908.
9. 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, tele-medicine, non-physician graders and smart phone reporting. *Indian J Ophthalmol*. 2014;62(1):41–49. doi:10.4103/0301-4738.126178.
10. The Committee for the Classification of Retinopathy of PrematurityAn international classification of retinopathy of prematurity. *Arch Ophthal*. 1984;102(8):1130–1134. doi:10.1001/archopht.1984.01040030908011.
11. Multicenter trial of cryotherapy for retinopathy of prematurity: preliminary results. *Arch Ophthal*. 1988;106(4):471–479. doi:10.1001/archopht.1988.01060130517027.
12. Early Treatment For Retinopathy Of Prematurity Cooperative GroupRevised indications for the treatment of retinopathy of prematurity: results of the early treatment for retinopathy of prematurity randomized trial. *Arch Ophthalmol*. 2003;121(12):1684–1694. doi:10.1001/archopht.121.12.1684.
13. International Committee for the Classification of Retinopathy of PrematurityThe International Classification of Retinopathy of Prematurity revisited. *American Medical Association*. 2005;123(7):991–999. doi:10.1001/archopht.123.7.991.
14. Wallace DK, Freedman SF, Hartnett ME, Quinn GE. Predictive value of pre-plus disease in retinopathy of prematurity. *Arch Ophthal*. 2011;129(5):591–596. doi:10.1001/archophthalmol.2011.63.
15. Capone A, Ells AL, Fielder AR, et al. Standard image of plus disease in retinopathy of prematurity. *Arch Ophthal*. 2006;124(11):1669–1670. doi:10.1001/archopht.124.11.1669-c.
16. Mintz-Hittner HA, Kennedy KA, Chuang AZBEAT-ROP Cooperative Group. Efficacy of intravitreal bevacizumab for stage 3+ retinopathy of prematurity. *N Engl J Med*. 2011;364(7):603–615. doi:10.1056/NEJMoa1007374.
17. Stahl A, Lepore D, Fielder A, et al. Ranibizumab versus laser therapy for the treatment of very low birthweight infants with retinopathy of prematurity (RAINBOW): an open-label randomised controlled trial. *Lancet*. 2019;394(10208):1551–1559. doi:10.1016/S0140-6736(19)31344-3.
18. Stahl A, Sukgen EA, Wu WC, et al. Effect of intravitreal aflibercept vs laser photocoagulation on treatment success of retinopathy of prematurity: the FIREFLEYE randomized clinical trial. *JAMA*. 2022;328(4):348–359. doi:10.1001/jama.2022.10564.
19. Patel NA, Acaba-Berrocal LA, Hoyek S, et al. Comparison in retreatments between bevacizumab and ranibizumab intravitreal injections for retinopathy of prematurity: A multicenter study. *Ophthalmology*. 2023;130(4):373–378. doi:10.1016/j.ophtha.2022.11.012.
20. Shah PK, Subramanian P, Venkatapathy N, Chan RVP, Chiang MF, Campbell JP. Aggressive posterior retinopathy of prematurity in two cohorts of patients in South India: implications for primary, secondary, and tertiary prevention. *J AAPOS*. 2019;23(5):264.e1–264.e4. doi:10.1016/j.jaapos.2019.05.014.
21. Sankar MJ, Sankar J, Chandra P. Anti-vascular endothelial growth factor (VEGF) drugs for treatment of retinopathy of prematurity. Cochrane Neonatal Group, ed. *Cochrane Database Syst Rev*. 2018;1(5):CD009734. doi:10.1002/14651858.CD009734.pub3.
22. Sanghi G, Dogra MR, Katoch D, Gupta A. Aggressive posterior retinopathy of prematurity in infants ≥ 1500 g birth weight. *Indian J Ophthalmol*. 2014;62(2):254–257. doi:10.4103/0301-4738.128639.
23. Chiang MF, Quinn GE, Fielder AR, et al. International classification of retinopathy of prematurity. *Ophthalmology*. 2021;128(10):e51–e68. doi:10.1016/j.ophtha.2021.05.031.
24. Chiang MF, Jiang L, Gelman R, Du YE, Flynn JT. Interexpert agreement of plus disease diagnosis in retinopathy of prematurity. *Arch Ophthalmol*. 2007;125(7):875–880. doi:10.1001/archopht.125.7.875.
25. Moleta C, Campbell JP, Kalpathy-Cramer J, et al. Plus Disease in Retinopathy of Prematurity: diagnostic Trends in 2016 vs. 2007. *Am J Ophthalmol*. 2017;176:70–76. doi:10.1016/j.ajo.2016.12.025.
26. Binenbaum G, Stahl A, Coyner AS, et al. P score: A reference image-based clinical grading scale for vascular change in retinopathy of prematurity. *Ophthalmology*. 2024;131(11):1297–1303. doi:10.1016/j.ophtha.2024.05.019.
27. Fleck BW, Williams C, Juszczak E, et al. An international comparison of retinopathy of prematurity grading performance within the Benefits of Oxygen Saturation Targeting II trials. *Eye*. 2017;123(1):1–7. doi:10.1038/eye.2017.150.
28. Kemper AR, Wallace DK, Quinn GE. Systematic review of digital imaging screening strategies for retinopathy of pre-

maturity. *Pediatrics*. 2008;122(4):825–830. doi:10.1542/peds.2007-3667.
29. Rosenblatt TR, Ji MH, Vail D, et al. Key factors in a rigorous longitudinal image-based assessment of retinopathy of prematurity. *Sci Rep*. 2021;11(1):5369. doi:10.1038/s41598-021-84723-7.
30. Cole ED, Park SH, Kim SJ, et al. Variability in plus disease diagnosis using single and serial images. *Ophthalmol Retina*. 2022;6(12):1122–1129. doi:10.1016/j.oret.2022.05.024.
31. Fijalkowski N, Zheng LL, Henderson MT, et al. (SUNDROP): five years of screening with telemedicine. *Ophthal Surg Lasers Imaging Retina*. 2014;45(2):106–113. doi:10.3928/23258160-20140122-01.
32. Sabri K, Ells AL, Lee EY, Dutta S, Vinekar A. Retinopathy of prematurity: A global perspective and recent developments. *Pediatrics*. 2022;150(3). doi:10.1542/peds.2021-053924.
33. Vinekar A, Jayadev C, Bauer N. Need for telemedicine in retinopathy of prematurity in middle-income countries. *JAMA Ophthalmol*. 2015;133(3):360–361. doi:10.1001/jamaophthalmol.2014.4913.
34. Chiang MF, Gelman R, Jiang L, Martinez-Perez ME, Du YE, Flynn JT. Plus disease in retinopathy of prematurity: an analysis of diagnostic performance. *Trans Am Ophthalmol Soc*. 2007;105:73–84.
35. Hewing NJ, Kaufman DR, Chan RVP, Chiang MF. Plus disease in retinopathy of prematurity. *JAMA Ophthalmol*. 2013;131(8):1026–1027. doi:10.1001/jamaophthalmol.2013.135.
36. Campbell JP, Kalpathy-Cramer J, Erdogmus D, et al. Plus disease in retinopathy of prematurity: a continuous spectrum of vascular abnormality as basis of diagnostic variability. *Ophthalmology*. 2016;123(11):2338–2344. doi:10.1016/j.ophtha.2016.07.026.
37. Kalpathy-Cramer J, Campbell JP, Erdogmus D, et al. Plus disease in retinopathy of prematurity: improving diagnostic variability by ranking disease severity and using quantitative image analysis. *Ophthalmology*. 2016;123(11):2345–2351. doi:10.1016/j.ophtha.2016.07.020.
38. Choi RY, Brown JM, Kalpathy-Cramer J, et al. Variability in plus disease identified using a deep learning-based retinopathy of prematurity severity scale. *Ophthalmol Retina*. 2020;4(10):1016–1021. doi:10.1016/j.oret.2020.04.022.
39. Chan RV, Patel SN, Ryan MC, et al. The global education network for retinopathy of prematurity (gen-rop): development, implementation, and evaluation of a novel tele-education system. *(An American Ophthalmological Society Thesis)*. *Trans Am Ophthalmol Soc*.. 2015;113:T2.
40. Wallace DK. Computer-assisted quantification of vascular tortuosity in retinopathy of prematurity [an American Ophthalmological Society thesis]. *Trans Am Ophthalmol Soc*. 2007;105:594–615.
41. Scruggs BA, Chan RVP, Kalpathy-Cramer J, Chiang MF, Campbell JP. Artificial intelligence in retinopathy of prematurity diagnosis. *Transl Vis Sci Technol*. 2020;9(2):5. doi:10.1167/tvst.9.2.5.
42. Abbey AM, Besirli CG, Musch DC, et al. Evaluation of screening for retinopathy of prematurity by ROPtool or a lay reader. *Ophthalmology*. 2016;123(2):385–390. doi:10.1016/j.ophtha.2015.09.048.
43. Ataer-Cansizoglu E, Bolon-Canedo V, Campbell JP, et al. Computer-based image analysis for plus disease diagnosis in retinopathy of prematurity: image analysis features associated with expert diagnosis. *Transl Vis Sci Technol*. 2015;4(6):5. doi:10.1167/tvst.4.6.5.
44. Brown JM, Campbell JP, Beers A, et al. Automated diagnosis of plus disease in retinopathy of prematurity using deep convolutional neural networks. *JAMA Ophthalmol*. 2018;136(7):803–810. doi:10.1001/jamaophthalmol.2018.1934.
45. Campbell JP, Kim SJ, Brown JM, et al. Evaluation of a deep learning-derived quantitative retinopathy of prematurity severity scale. *Ophthalmology*. 2021;128(7):1070–1076. doi:10.1016/j.ophtha.2020.10.025.
46. Bellsmith KN, Brown J, Kim SJ, et al. Aggressive posterior retinopathy of prematurity: clinical and quantitative imaging features in a large North American cohort. *Ophthalmology*. 2020;127(8):1105–1112. doi:10.1016/j.ophtha.2020.01.052.
47. Campbell JP, Chiang MF, Chen JS, et al. Artificial intelligence for retinopathy of prematurity: validation of a vascular severity scale against international expert diagnosis. *Ophthalmology*. 2022;129(7):e69–e76. doi:10.1016/j.ophtha.2022.02.008.
48. Taylor S, Brown JM, Gupta K, et al. Monitoring disease progression with a quantitative severity scale for retinopathy of prematurity using deep learning. *JAMA Ophthalmol*. 2019;137(9):1022–1028. doi:10.1001/jamaophthalmol.2019.2433.
49. Gupta K, Campbell JP, Taylor S, et al. A quantitative severity scale for retinopathy of prematurity using deep learning to monitor disease regression after treatment. *JAMA Ophthalmol*. 2019;137(9):1029–1036. doi:10.1001/jamaophthalmol.2019.2442.
50. Coyner AS, Young BK, Ostmo SR, et al. Use of an artificial intelligence-generated vascular severity score improved plus disease diagnosis in retinopathy of prematurity. *Ophthalmology*. 2024;131(11):1290–1296. doi:10.1016/j.ophtha.2024.06.006.
51. Coyner AS, Murickan T, Oh MA, et al. Multinational external validation of autonomous retinopathy of prematurity screening. *JAMA Ophthalmol*. 2024;142(4):327–335. doi:10.1001/jamaophthalmol.2024.0045.
52. Coyner AS, Oh MA, Shah PK, et al. External validation of a retinopathy of prematurity screening model using artificial intelligence in 3 low- and middle-income populations. *JAMA Ophthalmol*. 2022;140(8):791–798. doi:10.1001/jamaophthalmol.2022.2135.
53. Coyner AS, Chen JS, Singh P, et al. Single-examination risk prediction of severe retinopathy of prematurity. *Pediatrics*. 2021;148(6):e2021051772. doi:10.1542/peds.2021-051772.
54. deCampos-Stairiker MA, Coyner AS, Gupta A, et al. Epidemiologic evaluation of retinopathy of prematurity severity in a large telemedicine program in India using artificial intelligence. *Ophthalmology*. 2023;130(8):837–843. doi:10.1016/j.ophtha.2023.03.026.
55. Campbell JP, Singh P, Redd TK, et al. Applications of artificial intelligence for retinopathy of prematurity screening. *Pediatrics*. 2021;147(3):e2020016618. doi:10.1542/peds.2020-016618.
56. Hanif A, Lu C, Chang K, et al. Federated learning for multicenter collaboration in ophthalmology: implications for clinical diagnosis and disease epidemiology. *Ophthalmol Retina*. 2022;6(8):650–656. doi:10.1016/j.oret.2022.03.005.

57. Ryan MC, Ostmo S, Jonas K, et al. Development and evaluation of reference standards for image-based telemedicine diagnosis and clinical research studies in ophthalmology. *AMIA Annu Symp Proc*. 2014;2014:1902–1910.
58. 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.
59. Vinekar A, Mangalesh S, Jayadev C, Gilbert C, Dogra M, Shetty B. Impact of expansion of telemedicine screening for retinopathy of prematurity in India. *Indian J Ophthalmol*. 2017;65(5):390–395. doi:10.4103/ijo.IJO_211_17.
60. Koogler B, Young B, Peter Campbell J, Guidoboni G. Analysis of vessel tortuosity and its impact on hemodynamics in retinopathy of prematurity. *Invest Ophthalmol Vis Sci*. 2023;64(8) 1237-1237.
61. Sutter DA, Woodward MK, Jackson J, et al. Optical-coherence tomography derived quantitative measurement of extent of vascularization ("zone") in retinopathy of prematurity. *Ophthalmol Sci*. 2025(100912). doi:10.1016/j.xops.2025.100912.
62. Cole E, Valikodath NG, Al-Khaled T, et al. Evaluation of an artificial intelligence system for retinopathy of prematurity screening in Nepal and Mongolia. *Ophthalmol Sci*. 2022;2(4):100165. doi:10.1016/j.xops.2022.100165.
63. Quinn GE, Ells A, Capone A, et al. Analysis of discrepancy between diagnostic clinical examination findings and corresponding evaluation of digital images in the telemedicine approaches to evaluating acute-phase retinopathy of prematurity study. *JAMA Ophthalmol*. 2016;134(11):1263–1270. doi:10.1001/jamaophthalmol.2016.3502.
64. Campbell JP, Ryan MC, Lore E, et al. Diagnostic discrepancies in retinopathy of prematurity classification. *Ophthalmology*. 2016;123(8):1795–1801. doi:10.1016/j.ophtha.2016.04.035.
65. Sutter DA, Woodward MK, Jackson J, et al. Topographic features in retinopathy of prematurity with en face ultra-widefield optical coherence tomography. *Ophthalmol Retina*. 2026;10(1):88–94. doi:10.1016/j.oret.2025.07.003.
66. Burt SS, Coyner AS, Roti EV, et al. Automated quantification of retinopathy of prematurity stage via ultrawidefield OCT. *Ophthalmol Sci*. 2025;5(2):100663. doi:10.1016/j.xops.2024.100663.
67. Nguyen TTP, 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. doi:10.1001/jamaophthalmol.2022.4173.
68. Hellgren G, Löfqvist C, Hård AL, et al. Serum concentrations of vascular endothelial growth factor in relation to retinopathy of prematurity. *Pediatr Res*. 2016;79(1-1):70–75. doi:10.1038/pr.2015.181.
69. Liang T, Qian Z, Tao Y, et al. The relationship between the aqueous VEGF level and the severity of type 1 retinopathy of prematurity. *J Clin Med*. 2022;11(18):5361. doi:10.3390/jcm11185361.
70. Sjöbom U, Hellqvist T, Humayun J, et al. Circulating VEGF-A levels in relation to retinopathy of prematurity and treatment effects: a systematic review and meta-analysis. *Ophthalmol Sci*. 2024;4(6):100548. doi:10.1016/j.xops.2024.100548.
71. Sato T, Wada K, Arahori H, et al. Serum concentrations of bevacizumab (avastin) and vascular endothelial growth factor in infants with retinopathy of prematurity. *Am J Ophthalmol*. 2012;153(2) 327-333.e1. doi:10.1016/j.ajo.2011.07.005.
72. Guaiquil VH, Hewing NJ, Chiang MF, Rosenblatt MI, Chan RVP, Blobel CP. A murine model for retinopathy of prematurity identifies endothelial cell proliferation as a potential mechanism for plus disease. *Invest Ophthalmol Vis Sci*. 2013;54(8):5294–5302. doi:10.1167/iovs.12-11492.
73. Eilts SK, Pfeil JM, Poschkamp B, et al. Assessment of retinopathy of prematurity regression and reactivation using an artificial intelligence-based vascular severity score. *JAMA Netw Open*. 2023;6(1):e2251512. doi:10.1001/jamanetworkopen.2022.51512.
74. Kim C, D'Amore P, Connor K. Revisiting the mouse model of oxygen-induced retinopathy. *Eye Brain*. 2016;8:67–79. doi:10.2147/EB.S94447.
75. Penn JS, Thum LA. Oxygen-induced retinopathy in the rat. *Oxygen Radic Biol Med*. 1988:1025–1028. doi:10.1007/978-1-4684-5568-7_168.
76. Smith LE, Wesolowski E, McLellan A, et al. Oxygen-induced retinopathy in the mouse. *Invest Ophthalmol Vis Sci*. 1994;35(1):101–111.
77. The STOP-ROP Multicenter Study GroupSupplemental Therapeutic Oxygen for Prethreshold Retinopathy Of Prematurity (STOP-ROP), a randomized, controlled trial. I: Primary outcomes. *Pediatrics*. 2000;105(2):295–310. doi:10.1542/peds.105.2.295.
78. Gaynon MW. Rethinking STOP-ROP: is it worthwhile trying to modulate excessive VEGF levels in prethreshold ROP eyes by systemic intervention? A review of the role of oxygen, light adaptation state, and anemia in prethreshold ROP. *Retina*. 2006;26(7):S18–S23. doi:10.1097/01.iae.0000244292.86627.1e.
79. Mills MD. STOP-ROP results suggest selective use of supplemental oxygen for prethreshold ROP. *Arch Ophthalmol*. 2000;118(8):1121–1122. doi:10.1001/archopht.118.8.1121.
80. Resnikoff S, Felch W, Gauthier TM, Spivey B. The number of ophthalmologists in practice and training worldwide: a growing gap despite more than 200,000 practitioners. *Br J Ophthalmol*. 2012;96(6):783–787. doi:10.1136/bjophthalmol-2011-301378.
81. Young BK, Cole ED, Shah PK, et al. Efficacy of smartphone-based Telescreening for retinopathy of prematurity with and without artificial intelligence in India. *JAMA Ophthalmol*. 2023;141(6):582–588. doi:10.1001/jamaophthalmol.2023.1466.