The Use of Postnatal Weight Gain Algorithms to Predict Severe or Type 1 Retinopathy of Prematurity
A Systematic Review and Meta-analysis

Sam Athikarisamy, MD; Saumil Desai, MD, DM; Sanjay Patole, DrPH; Shripada Rao, DM; Karen Simmer, PhD; Geoffrey C. Lam, MBBS

Abstract

IMPORTANCE The currently recommended method for screening for retinopathy of prematurity (ROP) is binocular indirect ophthalmoscopy, which requires frequent eye examinations entailing a heavy clinical workload. Weight gain–based algorithms have the potential to minimize the need for binocular indirect ophthalmoscopy and have been evaluated in different setups with variable results to predict type 1 or severe ROP.

OBJECTIVE To synthesize evidence regarding the ability of postnatal weight gain–based algorithms to predict type 1 or severe ROP.

DATA SOURCES PubMed, MEDLINE, Embase, and the Cochrane Library databases were searched to identify studies published between January 2000 and August 2021.

STUDY SELECTION Prospective and retrospective studies evaluating the ability of these algorithms to predict type 1 or severe ROP were included.

DATA EXTRACTION AND SYNTHESIS Two reviewers independently extracted data. This meta-analysis was performed according to the Cochrane guidelines and reported according to the Preferred Reporting Items for Systematic Reviews and Meta-analysis of Diagnostic Test Accuracy Studies (PRISMA-DTA) guidelines.

MAIN OUTCOMES AND MEASURES Ability of algorithms to predict type 1 or severe ROP was measured using statistical indices (pooled sensitivity, specificity, and summary area under the receiver operating characteristic curves, as well as pooled negative likelihood ratios and positive likelihood ratios and diagnostic odds ratios).

RESULTS A total of 61 studies (>37 000 infants) were included in the meta-analysis. The pooled estimates for sensitivity and specificity, respectively, were 0.89 (95% CI, 0.85-0.92) and 0.57 (95% CI, 0.51-0.63) for WINROP (Weight, IGF-1 [insulinlike growth factor 1], Neonatal, ROP), 1.00 (95% CI, 0.88-1.00) and 0.60 (95% CI, 0.51-0.69) for CHOPROP (Children's Hospital of Philadelphia ROP), 0.99 (95% CI, 0.97-1.00) and 0.49 (95% CI, 0.36-0.68) for G-ROP (Postnatal Growth and ROP), 0.60 (95% CI, 0.15-0.93) for ROPScore, 0.98 (95% CI, 0.94-0.99) and 0.35 (95% CI, 0.22-0.51) for CO-ROP (Colorado ROP). The original PINT (Premature Infants in Need of Transfusion) ROP study reported a sensitivity of 0.98 (95% CI, 0.91-0.99) and a specificity of 0.36 (95% CI, 0.30-0.42). The pooled negative likelihood ratios were 0.19 (95% CI, 0.13-0.27) for WINROP, 0.0 (95% CI, 0.00-0.01) for G-ROP, 0.10 (95% CI, 0.02-0.53) for CHOPROP, 0.03 (95% CI, 0.00-0.77) for ROPScore, and 0.07 (95% CI, 0.03-0.16) for CO-ROP. The pooled positive likelihood ratios were adequate sensitivity and negative likelihood ratios (<0.2), and pooled likelihood ratios were inadequate.

RESULTS The pooled estimates for sensitivity and specificity, respectively, were 0.89 (95% CI, 0.85-0.92) and 0.57 (95% CI, 0.51-0.63) for WINROP (Weight, IGF-1 [insulinlike growth factor 1], Neonatal, ROP), 1.00 (95% CI, 0.88-1.00) and 0.60 (95% CI, 0.51-0.69) for CHOPROP (Children's Hospital of Philadelphia ROP), 0.99 (95% CI, 0.97-1.00) and 0.49 (95% CI, 0.36-0.68) for G-ROP (Postnatal Growth and ROP), 0.60 (95% CI, 0.15-0.93) for ROPScore, 0.98 (95% CI, 0.94-0.99) and 0.35 (95% CI, 0.22-0.51) for CO-ROP (Colorado ROP). The original PINT (Premature Infants in Need of Transfusion) ROP study reported a sensitivity of 0.98 (95% CI, 0.91-0.99) and a specificity of 0.36 (95% CI, 0.30-0.42). The pooled negative likelihood ratios were 0.19 (95% CI, 0.13-0.27) for WINROP, 0.0 (95% CI, 0.00-0.01) for G-ROP, 0.10 (95% CI, 0.02-0.53) for CHOPROP, 0.03 (95% CI, 0.00-0.77) for ROPScore, and 0.07 (95% CI, 0.03-0.16) for CO-ROP. The pooled positive likelihood ratios were adequate sensitivity and negative likelihood ratios (<0.2), and pooled likelihood ratios were inadequate.

Key Points

Question Do postnatal weight gain–based algorithms have the potential to identify infants with type 1 retinopathy of prematurity (ROP) or severe ROP?

Findings This systematic review and meta-analysis that included 61 studies (>37 000 infants) found that weight gain–based algorithms have adequate sensitivity, ranging from 0.89 to 1.00, and negative likelihood ratios (<0.2). However, specificity and positive likelihood ratios were inadequate.

Meaning This study suggests that weight gain–based algorithms have adequate sensitivity and negative likelihood ratios and provide reasonable certainty that type 1 ROP or severe ROP is unlikely to develop (ie, the algorithm is useful for ruling out the disease).

Supplemental content

Author affiliations and article information are listed at the end of this article.
CONCLUSIONS AND RELEVANCE This study suggests that weight gain–based algorithms have adequate sensitivity and negative likelihood ratios to provide reasonable certainty in ruling out type 1 ROP or severe ROP. Given the implications of missing even a single case of severe ROP, algorithms with very high sensitivity (close to 100%) and low negative likelihood ratios (close to zero) need to be chosen to safely reduce the number of unnecessary examinations in infants at lower risk of severe ROP.

Introduction

Retinopathy of prematurity (ROP) is a disease of pathologic neovascularization affecting preterm infants. In early postnatal life, hyperoxia leads to suppression of vascular growth factors (phase 1). Subsequently, as retinal hypoxia sets in, there is an upsurge of vascular growth factors leading to unregulated vasoproliferation (phase 2). ROP either regresses spontaneously or continues to advance and can progress to cause retinal detachment and blindness if not detected and treated early. Infants with lower gestation and lower birth weight have a higher risk of developing ROP. Currently, at-risk infants are screened using repeated eye examinations (binocular indirect ophthalmoscopy [BIO]) starting at approximately 30 to 32 weeks' postmenstrual age and continuing until the retinal vasculature is fully mature (approximately 40 weeks' postmenstrual age). However, only fewer than 10% of screened infants need treatment for ROP.

Animal experiments have demonstrated the importance of nutrition and insulin-like growth factor 1 (IGF-1) in the retinal vascular development. Insufficient activation of endothelial growth factor by IGF-1 can alter the development of the retinal vasculature. In the clinical setting, low postnatal weight gain is considered as a surrogate marker for slower-than-expected increases in serum IGF-1 levels. Based on this hypothesis, risk prediction models such as WINROP (Weight, IGF-1, Neonatal Retinopathy of Prematurity), G-ROP (Postnatal Growth and Retinopathy of Prematurity), PINT (Premature Infants in Need of Transfusion) ROP, CHOP (Children's Hospital of Philadelphia) ROP, ROPScore, and CO-ROP (Colorado Retinopathy of Prematurity) have been evaluated to see whether they can predict the development of significant ROP. These models have the potential advantage of reducing the number of BIO examinations. However, to our knowledge, these models have not been implemented in clinical practice because of their limited generalizability.

The rationale for this systematic review was to assess whether postnatal weight gain–based algorithms have the potential to predict the development of type 1 or severe ROP. This review gains further importance given the current COVID-19 pandemic, which has led to the curtailment of health services, including ROP screening programs, due to the limited availability of ophthalmologists and mobile screening teams. The aim of this systematic review was to synthesize evidence by pooling the diagnostic accuracy indices for postnatal weight gain–based algorithms for predicting type 1 or severe ROP in preterm infants.

Methods

Design and Registration

This meta-analysis was performed according to the Cochrane guidelines and reported according to the Preferred Reporting Items for Systematic Reviews and Meta-analysis of Diagnostic Test Accuracy.
Studies (PRISMA-DTA) guidelines. The study protocol was registered in PROSPERO (CRD42020172874).

Search Strategy
A systematic search of the PubMed, MEDLINE, Embase, and Cochrane Library databases was performed to identify studies published between January 2000 and August 2021. PubMed was searched using the standard terminology (eMethods in the Supplement). Similar terminology was used while searching other databases. We also searched the Cochrane Library, ClinicalTrials.gov, grey literature (on OpenGrey, Google Scholar, and MedNar). No language restrictions were applied. The reference lists of all publications were searched manually for additional studies.

Data Extraction
Two reviewers (S.A. and S.D.) independently collected data from the included studies. Study data were further verified by one of the reviewers (S.R.) who had not collected the study data earlier. A total of 61 studies were included for the final meta-analysis. The following data were collected from each study: year(s) the study was conducted, country, gestational age at birth, birth weight, weekly weight gain, total sample size, follow-up rates, tool used to diagnose ROP (indirect ophthalmoscopy or wide-field digital retinal imaging), prospective or retrospective design, high-income or low- to middle-income country (World Bank list of economies), and diagnostic indices (true positives, false positives, true negatives, and false negatives). The process of study selection is shown in the study flow diagram (eFigure 1 in the Supplement).

Criteria for Considering Studies for This Review
The outcomes examined were the sensitivity and specificity of postnatal weight gain–based algorithms to predict type 1 or severe ROP. Studies that used type 1 ROP or severe ROP as the target disease were included.

Both prospective and retrospective studies that met the following criteria were included: (1) retinopathy screening for preterm infants and (2) the ability of weight gain–based screening algorithms to predict type 1 or severe ROP. The following standard definitions were accepted for type 1 ROP: zone I (any stage ROP with plus disease), zone I (stage 3 with or without plus disease), or zone II (stage 2 or 3 ROP with plus disease).

Severe ROP was defined as any ROP in zone I, stage 2 ROP in zone II with plus disease, or any stage 3 ROP. The studies included described the diagnostic ability of weight gain–based algorithms to predict type 1 or severe ROP in comparison with the findings by ophthalmologists either by BIO or by wide-field digital retinal imaging, which were considered the reference standards. Two reviewers (S.A. and S.D.) independently decided on the eligibility of studies for inclusion in the systematic review. Differences in opinion were resolved by discussion among all reviewers.

Quality Assessment
The quality assessment of included studies consisted of the following 4 domains according to the revised version of the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2): patient selection, index test, reference standard, and flow and timing. In this review, index test refers to the “alarm” in any of the postnatal weight gain–based algorithms for type 1 or severe ROP. The included studies were assessed for risk of bias in each domain and for applicability concerns in the first 3 domains.

Statistical Analysis
Meta-analysis was conducted using STATA, version 16 (StataCorp) by one of us (S.R.) in the presence of 2 other authors (S.A. and S.D.). We used the bivariate mixed-effects regression model to derive the pooled sensitivity, specificity, positive likelihood ratio (PLR), and negative likelihood ratio (NLR) and the diagnostic odds ratio (DOR) and respective 95% CIs. A summary receiver operating
characteristic (ROC) curve was generated to display the results of individual studies. The following cutoffs were used for interpretation of the area under the ROC values: low (0.5-0.7), moderate (0.71-0.9), or high (>0.9) accuracy. The DOR is the ratio of the odds of positivity in the disease relative to the odds of positivity in those without the disease. The value of the DOR ranges from 0 to infinity, with higher values indicating better discriminatory test performance. A likelihood ratio of approximately 1 means the test result neither rules in nor rules out the condition. A PLR above 1 indicates increased evidence of disease; the higher from 1, the more chance of disease. An NLR below 0.1 is very strong evidence to rule out a disease.

Additional Analysis

There is uneven practice with regard to oxygen saturation targeting, blood transfusion thresholds, and higher incidence of fetal growth restriction in low- and middle-income countries; these factors could have an impact on the performance of these algorithms. Hence, sensitivity analyses were conducted separately for high-income and low- and middle-income countries when possible.

Results

Selected Studies

The electronic database search yielded 779 titles and abstracts and 3 additional studies by exploring additional sources. After removal of duplicates, 160 articles were screened based on the title and abstract, and 85 articles were selected. The full text was read for eligibility, of which 61 were included in the final analysis. The 61 studies (>37,000 infants) included WINROP (n = 36), G-ROP (n = 9), PINTROP (n = 1), CHOPROP (n = 6), ROPscore (n = 5), and CO-ROP (n = 4). Studies that used a different cutoff score for ROPscore were not included in the analysis. The general characteristics of the studies included in the systematic review are reported in eTables 1 to 6 in the Supplement. The methodological quality and applicability of the included studies were assessed according to the QUADAS-2 guidelines. All 61 studies evaluated one of the weight gain–based algorithms, and there was minimal risk of bias. However, the reference standards used in the studies were either BIO or wide-field digital retinal imaging.

Meta-analysis

WINROP

The pooled estimates from 36 studies (n = 11,500) for the sensitivity of the WINROP algorithm to predict type 1 or severe ROP was 0.89 (95% CI, 0.85-0.92), and the pooled estimate of the specificity was 0.57 (95% CI, 0.51-0.63) (Figure 1, Table). The summary area under the ROC curve was 0.82 (95% CI, 0.78-0.85) (Figure 2). The pooled PLR was 2.1 (95% CI, 1.8-2.4), and the pooled NLR was 0.19 (95% CI, 0.13-0.27) (Table 2 in the Supplement). The summary PLR and NLR for index test fell in the right lower quadrant of the likelihood matrix (Figure 3A). The pooled DOR of the WINROP algorithm was 11 (95% CI, 7-17).

WINROP (High-Income vs Low- to Middle-Income Countries)

The sensitivity analyses of the WINROP algorithm assessed its performance in high-income and low-to-middle-income countries. For high-income countries, the pooled estimate from 24 studies (n = 8543) for the sensitivity of the WINROP algorithm was 0.91 (95% CI, 0.85-0.95), and the pooled estimate for the specificity was 0.60 (95% CI, 0.53-0.66) (Table). The summary area under the ROC curve was 0.82 (95% CI, 0.78-0.85). The PLR was 2.3 (95% CI, 1.9-2.7), and the NLR was 0.15 (95% CI, 0.08-0.26).

For low- to middle-income countries, the pooled estimate from 12 studies (n = 2957) for the sensitivity of the WINROP algorithm was 0.85 (95% CI, 0.78-0.90), and the pooled estimate for the
specificity was 0.51 (95% CI, 0.39-0.64) (Table). The summary area under the ROC curve was 0.81 (95% CI, 0.78-0.84). The PLR was 1.7 (95% CI, 1.3-2.3), and the NLR was 0.29 (95% CI, 0.19-0.45).

G-ROP
The pooled estimates from 9 studies (n = 14120) reported a sensitivity of 1.00 (95% CI, 0.88-1.00) and a specificity (7 studies) of 0.60 (95% CI, 0.15-0.93) for G-ROP to predict type 1 or severe ROP (Figure 4; Table). The summary area under the ROC curve was 0.99 (95% CI, 0.98-1.00). The PLR was 2.5 (95% CI, 0.7-9.1), and the NLR was 0.00 (95% CI, 0.00-0.32) (eFigure 3 in the Supplement). The pooled DOR of the G-ROP algorithm to predict type 1 or severe ROP was 3523 (95% CI, 4-3155457). The summary PLR and NLR for index test fell in the left lower quadrant of the likelihood matrix (Figure 3B).

PINT ROP
The original PINT ROP study (n = 334) reported a sensitivity of 0.98 (95% CI, 0.91-0.99) and a specificity of 0.36 (95% CI, 0.30-0.42) for PINT ROP to predict type 1 or severe ROP (Table). The PLR...
### Table. Summary of ROP Prediction Algorithms Based on Postnatal Weight Gain

| Algorithm name and description | Components of algorithm | No. of studies (No. of infants) | Sensitivity | Specificity | Positive likelihood ratio | Negative likelihood ratio | Diagnostic odds ratio | AUROC | Strengths | Limitations |
|-------------------------------|-------------------------|---------------------------------|-------------|-------------|---------------------------|--------------------------|-----------------------|-------|-----------|------------|
| **WINROP**<sup>6</sup>        | Cumulative deviations statistical approach: deviations between expected and actual weight accumulated week to week; expected weight data derived from control infants with no or mild ROP; alarms when cumulative deviations exceed a threshold | Gestational age, birth weight, and weekly input of observed weight | 24 (8543) For high-income countries; 12 (2957) for low-to middle-income countries; 36 (11 500) for both high- and low-to middle-income countries | 0.91 (0.85-0.95) for high-income countries; 0.85 (0.78-0.90) for low-to middle-income countries; 0.89 (0.85-0.92) for both high- and low-to middle-income countries | 0.60 (0.53-0.66) for high-income countries; 0.51 (0.39-0.64) for low-to middle-income countries; 0.57 (0.51-0.63) for both high- and low-to middle-income countries | 7.3 (1.9-2.7) for high-income countries; 1.7 (1.3-2.3) for low-to middle-income countries; 2.1 (1.8-2.4) for both high- and low-to middle-income countries | 0.15 (0.08-0.26) for high-income countries; 0.29 (0.19-0.45) for low-to middle-income countries; 0.19 (0.13-0.27) for both high- and low-to middle-income countries | 0.60 (0.53-0.66) for high-income countries; 0.51 (0.39-0.64) for low-to middle-income countries; 0.57 (0.51-0.63) for both high- and low-to middle-income countries | 2.3 (1.9-2.7) for high-income countries; 1.7 (1.3-2.3) for low-to middle-income countries; 2.1 (1.8-2.4) for both high- and low-to middle-income countries | 15 (8-30) For high-income countries; 6 (3-11) for low-to middle-income countries; 11 (7-17) for both high- and low-to middle-income countries | Most widely studied in both high-income and low-to middle-income countries; provides risk assessment every week through a web-based application; algorithm will often identify infants as high- or low-risk well before eye examinations are started |
| **G-ROP**<sup>7</sup>          | Hybrid model; 6 components; infant qualifies for a retinal examination if any of the algorithm components is present | Gestational age <28 wk; birth weight <1051 g; weight gain over 3 growth periods (10-19 d: <120 g; 20-29 d: <180 g; 30-39 d: <170 g); hydrocephalus | 9 (14 120) | 1.00 (0.88-1.00) | 0.60 (0.15-0.93) | 2.5 (0.7-9.1) | 0.00 (0.00-0.32) | 3523 (4-3155 457) | 0.99 (0.98-1.00) | Validated in a multicenter cohort with a large sample size; high sensitivity and low negative likelihood ratio | Relatively new algorithm; more studies are needed to assess generalizability |
| **PINT ROP**<sup>8</sup>      | A simpler (logistic regression) predictive model; an alarm is triggered when the risk is >0.085 on the scale provided | Birth weight <1000 g; gestational age; daily weight gain rate (weight gain rate calculated from the current and prior week’s measurement) | 1 (334) | 0.98 (0.91-0.99) | 0.36 (0.30-0.42) | 1.55 (1.41-1.69) | 0.04 (0.01-0.29) | 38.75 (9.00-58.00) | NA | Simple paper-based nomogram; model evaluated risk on a weekly basis | Not widely validated |
| **CHOP ROP**<sup>9</sup>      | A simpler (logistic regression) predictive model; similar to PINT ROP but included infants with birth weight <1501 g; an alarm is triggered when the risk is >0.014 on the scale (nomogram) | Birth weight <1501 g; gestational age; daily weight gain rate (weight gain rate calculated from the current and prior week’s measurement) | 6 (2135) | 0.95 (0.71-0.99) | 0.52 (0.36-0.68) | 2.0 (1.5-2.6) | 0.10 (0.02-0.53) | 20 (4-99) | 0.75 (0.71-0.79) | Simple paper-based nomogram; model evaluated risk on a weekly basis | Poor generalizability; low sensitivity reported from low- to middle-income countries |

(continued)
| Algorithm name and description | Components of algorithm | No. of studies (No. of infants) | Diagnostic indices, pooled estimates (95% CI) | Strengths | Limitations |
|-------------------------------|-------------------------|-------------------------------|------------------------------------------|-----------|-------------|
| ROPScore 10                   | Birth weight; gestational age; weight gain at a single time point (from birth to 6 wk) as a proportion of birth weight; oxygen use on ventilator; blood transfusion | 5 (1625)                      | Sensitivity: 0.99 (0.73-1.00) Specificity: 0.49 (0.03-0.74) Positive likelihood ratio: 1.9 (1.1-3.3) Negative likelihood ratio: 0.03 (0.00-0.77) Diagnostic odds ratio: 69 (2-2228) AUROC: 0.88 (0.84-0.90) | Takes into account other risk factors (blood transfusion, oxygen use on mechanical ventilation) | Once per child risk calculation at 6 wk of age; hence, could potentially miss infants developing aggressive posterior ROP; validated in low- to middle-income countries and not tested widely in high-income countries |
| CO-ROP 11                     | Gestational age <30 wk; birth weight <1500 g; weight gain at a single time point (from birth to 4 wk) <650 g | 4 (8082)                      | Sensitivity: 0.98 (0.94-0.99) Specificity: 0.35 (0.22-0.51) Positive likelihood ratio: 1.5 (1.2-1.9) Negative likelihood ratio: 0.07 (0.03-0.16) Diagnostic odds ratio: 22 (9-58) AUROC: 0.95 (0.93-0.97) | High sensitivity in the original cohort (100%) | Poor generalizability; low sensitivity reported in later cohorts from US and Canada |

Abbreviations: AUROC, area under the receiver operating characteristic curve; CHOPROP, Children’s Hospital of Philadelphia Retinopathy of Prematurity; CO-ROP, Colorado Retinopathy of Prematurity; G-ROP, Postnatal Growth and Retinopathy of Prematurity; NA, not applicable; PINTROP, Premature Infants in Need of Transfusion Retinopathy of Prematurity; ROP, retinopathy of prematurity; WINROP, Weight, Insulinlike Growth Factor 1, Neonatal Retinopathy of Prematurity.
was 1.55 (95% CI, 1.41-1.69), and the NLR was 0.04 (95% CI, 0.01-0.29). The DOR was 38.75 (95% CI, 9.00-58.00).

CHOP ROP
The pooled estimates from 6 studies (n = 2135) showed a sensitivity of 0.95 (95% CI, 0.71-0.99) and a specificity of 0.52 (95% CI, 0.36-0.68) for the CHOP ROP algorithm to predict type 1 or severe ROP (Table). The summary area under the ROC curve was 0.75 (95% CI 0.71-0.79). The PLR was 2.0 (95% CI, 1.5-2.6), and the NLR was 0.10 (95% CI, 0.02-0.53). The pooled DOR of the CHOP ROP algorithm was 20 (95% CI, 4-99).

ROPScore
The pooled estimates from 5 studies (n = 1625) showed a sensitivity of 0.99 (95% CI, 0.73-1.00) and a specificity of 0.49 (95% CI, 0.03-0.74) for ROPScore to predict type 1 or severe ROP (Table). The summary area under the ROC curve was 0.88 (95% CI, 0.84-0.90). The PLR was 1.9 (95% CI, 1.1-3.3), and the NLR was 0.03 (95% CI, 0.00-0.77). The pooled DOR of the ROPScore algorithm was 69 (95% CI, 2-2228).

CO-ROP
The pooled estimates from 4 studies (n = 8082) showed a sensitivity of 0.98 (95% CI, 0.94-0.99) and a specificity of 0.35 (95% CI, 0.22-0.51) for CO-ROP to predict type 1 or severe ROP (Table). The summary area under the ROC curve was 0.95 (95% CI, 0.93-0.97). The PLR was 1.5 (95% CI, 1.2-1.9), and the NLR was 0.07 (95% CI, 0.03-0.16). The pooled DOR of the CO-ROP algorithm was 22 (95% CI, 9-58).

Discussion
The incidence of blindness after ROP is increasing globally. Early identification and treatment of ROP are of paramount importance to prevent irreversible blindness. Current screening protocols that require frequent retinal examinations place an enormous workload on the health care system. The
current COVID-19 pandemic has further strained the health care system owing to staff shortages, bed shortages, and limited availability of transport. Also, there is a risk of COVID-19 transmission to the preterm infants undergoing screening for ROP due to close contact with screening staff. The Royal

Figure 3. Likelihood Matrix

A, Summary positive likelihood ratio (PLR) and negative likelihood ratio (NLR) for Weight, Insulinlike Growth Factor 1, Neonatal Retinopathy of Prematurity (WINROP) are shown in the right lower quadrant (RLQ). B, Summary PLR and NLR for Postnatal Growth and Retinopathy of Prematurity (G-ROP) are shown in the RLQ. LLQ indicates left lower quadrant; LRN, likelihood ratio negative; LRP, likelihood ratio positive; LUQ, left upper quadrant; and RUQ, right upper quadrant.

Figure 4. Pooled Estimate for Sensitivity and Specificity of Postnatal Growth and Retinopathy of Prematurity (G-ROP) Model

| Source | Sensitivity (95% CI) | Specificity (95% CI) | Less sensitivity | More sensitivity | Less specificity | More specificity |
|--------|----------------------|----------------------|------------------|-----------------|-----------------|-----------------|
| Ahmed et al,62, 2021 (UK cohort) | 1.00 (0.69-1.00) | 0.60 (0.15-0.93) | 0.17 (0.13-0.22) | 0.34 (0.28-0.41) | 0.42 (0.38-0.47) | 0.60 (0.15-0.93) |
| Ahmed et al,62, 2021 (Egypt cohort) | 1.00 (0.93-1.00) | 0.21 (0.13-0.31) | 0.15 (0.12-0.19) | 0.29 (0.25-0.33) | 0.42 (0.38-0.47) | 0.60 (0.15-0.93) |
| Almeida et al,34, 2021 | 0.91 (0.71-0.99) | 0.91 (0.91-1.00) | 1.00 (0.99-1.00) | Not available | Not available | Not available |
| Binenbaum et al,7, 2018 | 1.00 (0.99-1.00) | 0.91 (0.80-1.00) | 0.91 (0.80-1.00) | Not available | Not available | Not available |
| Carughi et al,61, 2021 | 1.00 (0.91-1.00) | 0.17 (0.13-0.22) | 0.15 (0.12-0.19) | 0.29 (0.25-0.33) | 0.42 (0.38-0.47) | 0.60 (0.15-0.93) |
| Yabas Kiziloglu et al,38, 2020 | 0.91 (0.76-0.98) | 0.91 (0.76-0.98) | 0.34 (0.28-0.41) | 0.42 (0.38-0.47) | 0.60 (0.15-0.93) | 0.60 (0.15-0.93) |
| Shiraki et al,56, 2019 | 1.00 (0.96-1.00) | 1.00 (0.96-1.00) | 0.17 (0.13-0.22) | 0.34 (0.28-0.41) | 0.42 (0.38-0.47) | 0.60 (0.15-0.93) |
| Wadley et al,59, 2020 | 1.00 (0.91-1.00) | 0.91 (0.91-1.00) | 0.91 (0.91-1.00) | 1.00 (0.99-1.00) | 0.91 (0.91-1.00) | 1.00 (0.99-1.00) |

Combined | 1.00 (0.68-1.00) | 0.60 (0.15-0.93) | 0.17 (0.13-0.22) | 0.34 (0.28-0.41) | 0.42 (0.38-0.47) | 0.60 (0.15-0.93) |

\( Q = 157.62; df = 8.00; P < .001; I^2 = 94.92 \) (95% CI, 92.79-97.06)

\( Q = 1906.94; df = 8.00; P < .001; I^2 = 99.50-99.66 \)

JAMA Network Open. 2021;4(11):e2135879. doi:10.1001/jamanetworkopen.2021.35879 (Reprinted) November 23, 2021 9/16

Downloaded From: https://jamanetwork.com/ by a Non-Human Traffic (NHT) User on 01/15/2022
Ophthalmologist College has issued a statement in this regard to overcome the challenges by rationalizing the ROP screening criteria by using evidence-based weight gain–based algorithms.12

Weight gain–based screening models, such as WINROP, G-ROP, PINTROP, CHOPROP, ROPS Score, and CO-ROP,6–11 have been evaluated in various studies to identify infants who should be referred for retinal examination. However, the 2018 policy statement from the American Academy of Pediatrics stated that the uses of such weight gain–based algorithms alone are not justified based on the current literature.3 Our systematic review increases the evidence base in this area by including more than 37,000 preterm infants from 61 studies in diverse settings across the world.

The general interpretation of the results of diagnostic accuracy studies are as follows: PLRs of more than 10 or NLRs less than 0.1 generate large and often conclusive changes in the posttest probability, PLRs from more than 5 to 10 or NLRs from 0.1 to 0.2 generate moderate shifts in posttest probability, PLRs from more than 2 to 5 or NLRs from more than 0.2 to 0.5 generate small changes in posttest probability, and PLRs from more than 1 to 2 or NLRs from more than 0.5 to 1 alter posttest probability to a very small degree.71,72 High sensitivity corresponds to a high negative predictive value and a low NLR and is the ideal property of a "rule-out" test.73,74 Because all of the algorithms that we evaluated had an NLR of less than 0.2 and a high sensitivity of 0.89 to 1.00, they are useful to rule out type 1 ROP if the algorithm had "no alarm."

Given that the sensitivity of G-ROP were better than WINROP and the sample size was larger, G-ROP may be more suitable as a rule-out test compared with WINROP. However, G-ROP is not widely evaluated outside high-income countries. High specificity corresponds to a high positive predictive value and a high PLR and is the ideal property of a "rule-in" test.73,74 Because all of the algorithms had a very low specificity and a very low PLR of 1.5 to 2.5, they may not be suitable for ruling in the disease.

Although the overall sensitivity (0.89) and NLR (0.19) for WINROP were reasonably adequate, they were significantly lower than the validation studies that were performed in the early 2010. The sensitivity was 1.00 in a Swedish cohort6 of 353 infants and a Boston cohort22 of 318 infants. A multicenter study in the US and Canada showed that the sensitivity of the WINROP algorithm was 0.986.24 The sensitivity values were lower in studies conducted in low- to middle-income countries. The publication by Lundgren et al35 in 2018 showed an association between the decrease in sensitivity and the change in oxygen target ranges that occurred in 2010. This could have been the explanation for the low sensitivity in low- to middle-income countries, where there is uneven practice with regard to oxygen targeting in preterm infants.44-55

The G-ROP model was originally developed from a retrospective cohort of 7483 infants from 29 North American centers that showed a sensitivity of 1.00 to identify type 1 ROP.7 This finding was further validated in a prospective cohort of 3981 infants in North America with a similar sensitivity of 1.00.57 When validated outside the North American cohort, the sensitivity was 1.00 in a small cohort of infants from Japan,56 the UK, and Egypt50; however, it decreased to approximately 0.91 in a Turkish and Portuguese cohort.42,58

Summary PLRs and NLRs for index tests fell in the right lower quadrant of the likelihood matrix for WINROP and the left lower quadrant for G-ROP. This finding suggests that diagnostic indices are better for G-ROP compared with WINROP. However, the strength of WINROP is that it has been tested extensively in nearly 36 studies across the world.

There are a significant number of infants in low- and middle-income countries who develop type 1 ROP and who are more mature (>28 weeks) and heavier (>1050 g).75 Hence, validating the G-ROP algorithm in these groups of infants will give additional information about the performance of the weight gain component of the G-ROP algorithm (weight gain over 3 growth periods: 10-19 days, 20-29 days, and 30-39 days).

The CHOPROP model was an updated version of the PINTROP model and had a sensitivity of 1.00 to identify type 1 ROP in the original cohort and a sensitivity of 0.985 to identify type 1 ROP in a multicenter cohort.9,62 Sensitivities of 0.667 and 0.549 in the CHOPROP model were reported from 2 different Indian cohorts.55,63 The ROPS score had additional risk factors, such as oxygen use in
mechanical ventilation and blood transfusion. In the original cohort, it had a sensitivity of 0.9583; however, when it was tested using different cutoff scores, the sensitivity was 0.50 to 1.00 in different cohorts. Proportional weight gain calculated at 6 weeks of postnatal life could potentially miss some of the infants who would develop aggressive posterior ROP within this time period.

This review summarizes the evidence for weight gain–based algorithms and their performance in predicting type 1 or severe ROP. This review has highlighted how the performance of these weight gain–based models has varied in their predictive values depending on the situation in which it was tested. Hesitancy in implementing these algorithms into clinical practice is justifiable because the focus has been on not missing a single child who may need treatment.

Postnatal weight gain–based models are not ideal for infants who show a pattern of nonphysiological weight gain (eg, hydrocephalus, sepsis, and patent ductus arteriosus). Future studies should endeavor to incorporate other risk factors, such as oxygen use and sepsis, into the algorithm to improve their diagnostic performance. Another area of research would be to validate these algorithms against a combined reference standard rather than BIO or wide-field digital retinal imaging individually.

Strengths and Limitations
This study has some strengths. To our knowledge, this is the first systematic review including a meta-analysis to evaluate the diagnostic accuracy of postnatal weight gain–based algorithms with a sample size of more than 37,000 infants. The study sample included infants from both high-income and low-to-middle-income countries, and additional analyses were performed in these subgroups. Furthermore, a standardized tool (ie, QUADAS-2) was used for the quality assessment. In addition, this systematic review and meta-analysis followed the recent PRISMA-DTA guidelines for transparent reporting.

This study also has some limitations. The weight gain–based algorithms rely strongly on accurate measurement of weight, which is very challenging in a neonatal intensive care setting and could potentially affect the performance of the algorithms. Apart from the index test, there were multiple heterogeneities present across the studies. These heterogeneities were partly explained by factors such as geographical background, workflow, and the choice of a reference standard, which in itself is known to have interobserver variability.

Conclusions
This systematic review and meta-analysis suggests that weight gain–based algorithms have adequate sensitivity and negative likelihood ratios to provide reasonable certainty in ruling out type 1 or severe ROP. Given the implications of missing even a single case of severe ROP, algorithms with very high sensitivity (close to 100%) and low negative likelihood ratio (close to zero) need to be chosen. These algorithms have the potential to reduce the number of unnecessary examinations for infants at lower risk of severe ROP. Future studies should endeavor to incorporate additional clinical parameters (eg, oxygen use and sepsis), which could potentially improve the diagnostic indices of these algorithms.

ARTICLE INFORMATION
Accepted for Publication: September 29, 2021.
Published: November 23, 2021. doi:10.1001/jamanetworkopen.2021.35879
Open Access: This is an open access article distributed under the terms of the CC-BY License. © 2021 Athikarisamy S et al. JAMA Network Open.
Corresponding Author: Sam Athikarisamy, MD, Neonatal Directorate, Perth Children's Hospital, 15 Hospital Ave, Nedlands, Western Australia (sam.athikarisamy@health.wa.gov.au).
Author Affiliations: Neonatal Directorate, Perth Children's Hospital and King Edward Memorial Hospital for
Women, Perth, Australia (Athikarisamy, Desai, Patole, Rao); School of Medicine, University of Western Australia, Crawley, Australia (Athikarisamy, Patole, Rao, Simmer); Department of Ophthalmology, Perth Children's Hospital, Perth, Australia (Lam); Centre for Ophthalmology and Visual Science, University of Western Australia, Crawley, Australia (Lam).

Author Contributions: Drs Athikarisamy and Desai had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Athikarisamy, Desai, Rao, Simmer, Lam.

Acquisition, analysis, or interpretation of data: Athikarisamy, Desai, Patole, Rao.

Drafting of the manuscript: Athikarisamy, Desai, Rao.

Critical revision of the manuscript for important intellectual content: All authors.

Statistical analysis: Athikarisamy, Desai, Rao.

Supervision: Patole, Rao, Simmer.

Conflict of Interest Disclosures: None reported.

REFERENCES

1. HellströmA, Smith LE, Dammann O. Retinopathy of prematurity. Lancet. 2013;382(9902):1445-1457. doi:10.1016/S0140-6736(13)60178-6

2. Early Treatment For Retinopathy Of Prematurity Cooperative Group. Revised 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

3. 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. doi:10.1542/peds.2018-3061

4. Hellström A, Engström E, Härd AL, et al. Postnatal serum insulin-like growth factor I deficiency is associated with retinopathy of prematurity and other complications of premature birth. Pediatrics. 2003;112(5):1016-1020. doi:10.1542/peds.112.5.1016

5. Löfqvist C, Andersson E, Sigurdsson J, et al. Longitudinal postnatal weight and insulin-like growth factor I measurements in the prediction of retinopathy of prematurity. Arch Ophthalmol. 2006;124(12):1711-1718. doi:10.1001/archopht.124.12.1711

6. Hellström A, Härd 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. doi:10.1542/peds.2008-2697

7. Binenbaum G, Bell EF, Donohue P, et al; G-ROP Study Group. 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. doi:10.1001/jamaophthalmol.2018.2753

8. Binenbaum G, Ying GS, Quinn GE, et al; Premature Infants in Need of Transfusion Study Group. A clinical prediction model to stratify retinopathy of prematurity risk using postnatal weight gain. Pediatrics. 2011;127(3):e607-e614. doi:10.1542/peds.2010-2240

9. 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. doi:10.1001/archophthalmol.2012.2524

10. Eckert GU, Fortes Filho JB, Maia M, Prociannon RS. A predictive score for retinopathy of prematurity in very low birth weight preterm infants. Eye (Lond). 2012;26(3):400-406. doi:10.1038/eye.2011.334

11. 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. doi:10.1016/j.jaapos.2015.10.017

12. The Royal College of Ophthalmologists. RCOphth: Management of ophthalmology services during the Covid pandemic. Accessed June 30, 2021. https://www.rcophth.ac.uk/wp-content/uploads/2020/03/RCOpth-Management-of-Ophthalmology-Services-during-the-Covid-pandemic-FINAL.pdf

13. Deeks JJ, Wisniewski S, Davenport C. Chapter 4: Guide to the contents of a Cochrane Diagnostic Test Accuracy Protocol. In: Deeks JJ, Bossuyt PM, Gatsonis C, eds. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy: Version 1.0.0. The Cochrane Collaboration; 2013.

14. PRISMA: Transparent Reporting of Systematic Reviews and Meta-Analyses. PRISMA-DTA checklist. Accessed June 30, 2021. http://www.prisma-statement.org/documents/PRISMA%20DTA%20Checklist.pdf
15. Athikarisamy SE, Desai S, Rao S, Patole S. National Institute for Health Research: PROSPERO: International prospective register of systematic reviews. Weight gain based algorithm WINROP to predict type 1 or severe retinopathy of prematurity (ROP) in preterm infants: a systematic review and meta-analysis. Accessed June 30, 2021. https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=172874

16. The World Bank. The world by income and region. Accessed June 30, 2021. https://datatopics.worldbank.org/world-development-indicators/the-world-by-income-and-region.html

17. University of Bristol. QUADAS-2: background document. Accessed June 30, 2021. https://www.bristol.ac.uk/media-library/sites/quadas/migrated/documents/background-doc.pdf

18. STATA. New in STATA 16. Accessed June 30, 2021. https://www.stata.com/stata16/

19. Fischer JE, Bachmann LM, Jaeschke R. A reader’s guide to the interpretation of diagnostic test properties: clinical example of sepsis. *Intensive Care Med*. 2003;29(7):1043-1051. doi: 10.1007/s00134-003-1761-8

20. Šimundić AM. Measures of diagnostic accuracy: basic definitions. *EJIFCC*. 2009;19(4):203-211.

21. 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

22. Wu C, Vanderveen DK, Hellström A, Löfqvist C, Smith LE. Longitudinal postnatal weight measurements for the prediction of retinopathy of prematurity. *Arch Ophthalmol*. 2010;128(4):443-447. doi:10.1001/archophthalmol.2010.31

23. Flückiger S, Bucher HU, Hellström A, Lövquist C, Sturm V, Arri SJ. The early postnatal weight gain as a predictor of retinopathy of prematurity. Article in German. *Klin Monbl Augenheilk*. 2011;228(4):306-310. doi:10.1055/s-0031-1273217

24. Wu C, Löfqvist C, Smith LE, VanderVeen DK, Hellström A; WINROP Consortium. Importance of early postnatal weight gain for normal retinal angiogenesis in very preterm infants: a multicenter study analyzing weight velocity deviations for the prediction of retinopathy of prematurity. *Arch Ophthalmol*. 2012;130(8):992-999. doi:10.1001/archophthalmol.2012.243

25. Choi JH, Löfqvist C, Hellström A, Heo H. Efficacy of the screening algorithm WINROP in a Korean population of preterm infants. *JAMA Ophthalmol*. 2013;131(1):62-66. doi:10.1001/jamaophthalmol.2013.566

26. Lundgren P, Stolz SJöström E, Domnellöf M, et al. WINROP identifies severe retinopathy of prematurity at an early stage in a nation-based cohort of extremely preterm infants. *PLoS One*. 2013;8(9):e73256. doi:10.1371/journal.pone.0073256

27. Eriksson L, Lidén U, Löfqvist C, Hellström A. WINROP can modify ROP screening praxis: a validation of WINROP in populations in Sörmland and Västmanland. *Br J Ophthalmol*. 2014;98(7):964-966. doi:10.1136/bjophthalmol-2013-304617

28. Piyasena C, Dhalwal C, Russell H, et al. Prediction of severe retinopathy of prematurity using the WINROP algorithm in a birth cohort in South East Scotland. *Arch Dis Child Fetal Neonatal Ed*. 2014;99(1):F29-F33. doi:10.1136/adc.2013.304101

29. Ko CH, Kuo HK, Chen CC, et al. Using WINROP as an adjuvant screening tool for retinopathy of prematurity in southern Taiwan. *Am J Perinatol*. 2015;30(2):149-154. doi:10.1055/s-0034-1376389

30. Pierrarcocchi S, Bini S, Martini F, et al. Predictive algorithms for early detection of retinopathy of prematurity. *Acta Ophthalmol*. 2017;95(2):158-164. doi:10.1111/aos.13117

31. Jung JI, Wagner BD, McCourt EA, et al. Validation of WINROP for detecting retinopathy of prematurity in a North American cohort of preterm infants. *JAAPOS*. 2017;21(3):229-233. doi:10.1016/j.jaapos.2017.05.004

32. Jagla M, Peterko A, Oleśirska K, Szymońska I, Kwinta P. Prediction of severe retinopathy of prematurity using the WINROP algorithm in a cohort from Malopolska: a retrospective, single-center study. *Dev Period Med*. 2017;21(4):336-343.

33. Timkovic J, Pokryvkova M, Janurova K, Barinova D, Polackova R, Masek P. Evaluation of the WinROP system for identifying retinopathy of prematurity in Czech preterm infants. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub*. 2017;161(1):111-116. doi:10.5507/bp.2016.061

34. Ali E, Al-Shafouri N, Hussain A, Baier RJ. Assessment of WINROP algorithm as screening tool for preterm infants in Manitoba to detect retinopathy of prematurity. *Paediatr Child Health*. 2017;22(4):203-206. doi:10.1093/pch/pox053

35. Lundgren P, Härd AL, Wilde Å, Löfqvist C, Smith LEH, Hellström A. Implementing higher oxygen saturation targets reduced the impact of poor weight gain as a predictor for retinopathy of prematurity. *Acta Paediatr*. 2018;107(5):767-773. doi:10.1111/apa.14049
36. Chaves-Samaniego MJ, Gómez Cabrera C, Chaves-Samaniego MC, et al. Multicenter validation study of the WINROP algorithm as a method for detecting retinopathy of prematurity. J Matern Fetal Neonatal Med. 2020;33(8):1302-1306. doi:10.1080/14767058.2018.1517325

37. Ueda K, Miki A, Nakai S, Yanagisawa S, Nomura K, Nakamura M. Prediction of severe retinopathy of prematurity using the weight gain, insulin-like growth factor 1, and neonatal retinopathy of prematurity algorithm in a Japanese population of preterm infants. Jpn J Ophthalmol. 2020;64(2):223-227. doi:10.1007/s10384-019-00709-z

38. Binwale M, Weiner A, Sardesai S, Cayabyab R, Barton L, Ramanathan R. Early postnatal weight gain as a predictor for the development of retinopathy of prematurity. J Matern Fetal Neonatal Med. 2019;32(3):429-433. doi:10.1080/14767058.2017.1381902

39. Wirth M, Naud A, Caputo G, Hascoët JM. Algorithm for predicting threshold retinopathy of prematurity is insufficient and fundus examinations are still needed before 31 weeks. Acta Paediatr. 2019;108(6):1049-1054. doi:10.1111/apa.14654

40. Desai S, Athikarisamy SE, Lundgren P, Simmer K, Lam GC. Validation of WINROP (online prediction model) to identify severe retinopathy of prematurity (ROP) in an Australian preterm population: a retrospective study. Eye (Lond). 2021;35(5):1334-1339. doi:10.1038/s41433-020-1094-7

41. Raffa LH, Alessa SK, Alamri AS, Malaikah RH. Prediction of retinopathy of prematurity using the screening algorithm WINROP in a Saudi cohort of preterm infants. Saudi Med J. 2020;41(6):622-627. doi:10.15537/smj.2020.6.25127

42. Almeida AC, Sandina T, Azevedo R, et al. Retrospective comparison between growth and retinopathy of prematurity model versus WINROP model. Can J Ophthalmol. 2021;56(4):418-424. doi:10.1016/j.jcjo.2021.02.030

43. Fernández-Ramón R, Follana-Neira I, Ruiz-Sancho MD. Validation of WINROP algorithm as a screening tool for retinopathy of prematurity in a northern Spanish cohort. Int J Retina. 2021;4(1). doi:10.35479/ijretina.2021.vol004.iss001.137

44. Hård AL, Löfqvist C, Fortes Filho JB, Procianoy RS, Smith L, Hellström A. Predicting proliferative retinopathy in a Brazilian population of preterm infants with the screening algorithm WINROP. Arch Ophthalmol. 2010;128(11):1432-1436. doi:10.1001/archophthalmol.2010.255

45. Zepeda-Romero LC, Hård AL, Gomez-Ruiz LM, et al. Prediction of retinopathy of prematurity using the screening algorithm WINROP in a Mexican population of preterm infants. Arch Ophthalmol. 2012;130(6):720-723. doi:10.1001/archophthalmol.2012.215

46. Sun H, Kang W, Cheng X, et al. The use of the WINROP screening algorithm for the prediction of retinopathy of prematurity in a Chinese population. Neonatology. 2013;104(2):127-132. doi:10.1159/000351297

47. Koçak N, Niyaş L, Artirk N. Prediction of severe retinopathy of prematurity using the screening algorithm WINROP in preterm infants. JAAPOS. 2016;20(6):486-489. doi:10.1016/j.jaapos.2016.08.008

48. Bautista GC, Ventura RH. Local validation of WINROP, an online screening tool for retinopathy of prematurity. Philipp J Ophthalmal. 2018;42(1):15-18.

49. Kesting SJ. Prediction of retinopathy of prematurity using the WINROP screening algorithm in a South African population (doctoral dissertation). May 2018. Accessed June 30, 2021. https://wiredspace.wits.ac.za/bitstream/handle/10539/25296/Kesting%201527528.pdf?sequence=1&isAllowed=y

50. Sanghi G, Narang A, Narula S, Dogra MR. WINROP algorithm for prediction of sight-threatening retinopathy of prematurity: initial experience in Indian preterm infants. Indian J Ophthalmol. 2018;66(1):110-113. doi:10.4103/ijo.IJO_486_17

51. Kadir NA, Ahmad SS, Ghani SA, Bastion ML. Validation of the WINROP screening algorithm among preterm infants in East Malaysia. Asian J Ophthalmal. 2019;16(4):250-259. doi:10.3519/asjop.v16i4.402

52. Lim ZD, Oo KT, Tai ELM, Shatriiah I. Efficacy of WINROP as a Screening Tool for Retinopathy of Prematurity in the East Coast of Malaysia. Clin Ophthalmal. 2020;14:1101-1106. doi:10.2147/OPHT.2147820

53. Bai YC, Wu R, Chen SZ, et al. Efficacy of the WINROP algorithm for retinopathy of prematurity screening in Southern China. Int J Ophthalmal. 2021;14(1):127-132. doi:10.18240/ijo.2021.01.18

54. Sute SS, Jain S, Chawla D, Narang S. Use of an online screening algorithm—Weight, Insulin-derived growth factor 1, Neonatal Retinopathy of Prematurity (WINROP) for predicting retinopathy of prematurity in Indian preterm babies. Indian J Ophthalmal. 2021;69(5):1214-1218. doi:10.4103/ijo.IJO_1521_20

55. Thomas D, Madathil S, Thukral A, et al. Diagnostic accuracy of WINROP, CHOP-ROP and ROPScore in detecting type 1 retinopathy of prematurity. Indian Pediatr. 2021;58(10):915-921.
56. Shiraki A, Fukushima Y, Kawasaki R, et al. Retrospective validation of the Postnatal Growth and Retinopathy of Prematurity (G-ROP) criteria in a Japanese cohort. Am J Ophthalmol. 2019;205:50-53. doi:10.1016/j.ajo.2019.03.027

57. Binenbaum G, Tomlinson LA, de Alba Camponomes AG, et al; Postnatal Growth and Retinopathy of Prematurity (G-ROP) Study Group. Validation of the Postnatal Growth and Retinopathy of Prematurity screening criteria. JAMA Ophthalmol. 2020;138(1):31-37. doi:10.1001/jamaophthalmol.2019.4517

58. Yabas Kiziloglu O, Coskun Y, Akman I. Assessment of the G-ROP study criteria for predicting retinopathy of prematurity: results from a tertiary centre in Turkey. Int Ophthalmol. 2020;40(7):1647-1652. doi:10.1007/s10792-020-01332-5

59. Wadley S, Recko M, Samson S. Validation of the postnatal growth and retinopathy of prematurity screening criteria. Proc (Bayl Univ Med Cent). 2020;33(4):546-549. doi: 10.1080/08998280.2020.1793709

60. Ahmed ISH, Aclimandos W, Azad N, et al. The Postnatal Growth and Retinopathy of Prematurity Model: a multi-institutional validation study. Ophthalmic Epidemiol. Published online June 17, 2021. doi:10.1080/09286586.2021.1939885

61. Carughi S, Scaramuzzi M, Cable MG, et al. Validation of the Postnatal Growth and Retinopathy of Prematurity screening criteria: a retrospective Italian analysis. Eur J Ophthalmol. Published online April 22, 2021. doi:10.1177/11206721211011362

62. Binenbaum G, Ying GS, Tomlinson LA; Postnatal Growth and Retinopathy of Prematurity (G-ROP) Study Group. Validation of the Children's Hospital of Philadelphia Retinopathy of Prematurity (CHOP ROP) model. JAMA Ophthalmol. 2017;135(8):871-877. doi:10.1001/jamaophthalmology.2017.2295

63. Doshi S, Desai S, Nanavati R, Kabra N, Martin S. Children's Hospital of Philadelphia Score to predict severe retinopathy in Indian preterm infants. Eye (Lond). 2019;33(9):1452-1458. doi:10.1038/s41433-019-0431-1

64. Sun H, Dong Y, Liu Y, et al. Using ROPScore and CHOP ROP for early prediction of retinopathy of prematurity in a Chinese population. Ital J Pediatr. 2021;47(1):39. doi: 10.1186/s13052-021-00991-z

65. Figueiredo R, Sarmento TM, Garrido J, Marques MI, Almeida T, Carrasquinho S. Applicability of the ROPScore as a predictive algorithm for early detection of retinopathy of prematurity. Ophthalmologio. 2020;44(2). doi:10.48560/rspo.18704

66. Gulkas S, Ozcan Y, Gumus H. Validity and reliability of ROPScore scoring method to predict the severity of retinopathy of prematurity in infant authors. Preprint posted online March 7, 2021. doi:10.22541/au.161515047.77230490/v1

67. Cao JH, Wagner BD, Cerda A, et al. Colorado Retinopathy of Prematurity Model: a multi-institutional validation study. J AAPOS. 2016;20(3):220-225. doi:10.1016/j.jaapos.2016.01.017

68. Huang JM, Lin X, He YG, Cao JH. Colorado Retinopathy of Prematurity Screening Algorithm (CO-ROP): a validation study at a tertiary care center. J AAPOS. 2017;21(2):152-155. doi:10.1016/j.jaapos.2017.02.009

69. McCourt EA, Ying GS, Lynch AM, et al; G-ROP Study Group. Validation of the Colorado Retinopathy of Prematurity Screening Model. JAMA Ophthalmol. 2018;136(4):409-416. doi:10.1001/jamaophthalmol.2018.0376

70. Lucio KCDV, Bentlin MR, Augusto ACL, et al. The ROPScore as a Screening Algorithm for Predicting Retinopathy of Prematurity in a Brazilian Population. Clinics (Sao Paulo). 2018;73:e377. doi:10.6061/clinics/2018/e377

71. Jaeschke R, Guyatt GH, Sackett DL; Evidence-Based Medicine Working Group. Users' guides to the medical literature, III: how to use an article about a diagnostic test: B, what are the results and will they help me in caring for my patients? JAMA. 1994;271(9):703-707. doi:10.1001/jama.1994.03510330081039

72. Bhandari M, Montori VM, Swiontkowski MF, Guyatt GH. Users’ guide to the surgical literature: how to use an article about a diagnostic test. J Bone Joint Surg Am. 2003;85(6):1133-1140. doi:10.2106/00044623-200306000-00027

73. Rubinstein ML, Kraft CS, Parrott JS. Determining qualitative effect size ratings using a likelihood ratio scatter matrix in diagnostic test accuracy systematic reviews. Diagnosis (Berl). 2018;5(4):205-214. doi:10.1515/dx-2018-0061

74. Florkowski CM. Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: communicating the performance of diagnostic tests. Clin Biochem Rev. 2008;29(suppl 1):S83-S87.

75. Gilbert C, Fielder A, Gordillo L, et al; International NO-ROP Group. Characteristics of infants with severe retinopathy of prematurity in countries with low, moderate, and high levels of development: implications for screening programs. Pediatrics. 2005;115(5):e518-e525. doi:10.1542/peds.2004-1180
SUPPLEMENT.
eFigure 1. Flow Diagram for Study Selection Process
eFigure 2. The Positive (PLR) and Negative Likelihood Ratio (NLR) for WINROP Algorithm
eFigure 3. The Positive (PLR) and Negative Likelihood Ratio (NLR) for G-ROP Model

eMethods.
eTable 1. Characteristics of Included Studies of WINROP Algorithm, From High Income Countries
eTable 2. Characteristics of Included Studies Evaluating WINROP Algorithm, From Low and Middle-Income Countries
eTable 3. Characteristics of Included Studies Evaluating G-ROP (Postnatal Growth and ROP)
eTable 4. Characteristics of Included Studies Evaluating CHOP ROP
eTable 5. Characteristics of Included Studies Evaluating ROP Score Based on Cumulative Scores
eTable 6. Characteristics of Included Studies Evaluating Colorado ROP