The Colorado–retinopathy of prematurity model (CO-ROP): postnatal weight gain screening algorithm

Jennifer H. Cao, MD,a Brandie D. Wagner, PhD,b Emily A. McCourt, MD,a Ashlee Cerda, MPH,a Stefan Sillau, PhD,b Alan Palestine, MD,a Robert W. Enzenauer, MD, MBA, MPH,a Rebecca B. Mets-Halgrimson, MD, MPH,a Miguel Paciuc-Beja, MD,a Jane Gralla, PhD,c Rebecca S. Braverman, MD,a and Anne Lynch, MD, MSPHa

PURPOSE To describe a novel retinopathy of prematurity (ROP) screening model incorporating birth weight, gestational age, and postnatal weight gain that maintains sensitivity but improves specificity in detecting all grades of ROP compared to current 2013 screening guidelines.

METHODS The medical records of 499 neonates from a single tertiary referral center who met the 2013 screening guidelines for ROP were retrospectively reviewed. Weekly weights were analyzed using standard logistic regression to determine the age at which the weekly net weight gain best predicted the development of ROP, which was designated as the postnatal weight gain criterion. The 2013 birth weight and gestational age criteria were included in an “and” fashion to form the CO-ROP model. Sensitivities and specificities in detecting high grade (type 1 and 2) and all grades of ROP were calculated.

RESULTS The CO-ROP model screens infants with a gestational age at birth of \( \geq 30 \) weeks and birth weight of \( \geq 1500 \) g and net weight gain of \( \geq 650 \) g between birth and 1 month of age. In our cohort, CO-ROP had a sensitivity of 100% (95% CI, 92.1%-100.0%) for high-grade (type 1 and 2) ROP and 96.4% (95% CI, 92.3%-98.7%) for all grades of ROP. It would reduce the number of infants screened by 23.7% compared to 2013 guidelines. Calibrating the model to detect only high-grade ROP would result in a 45.9% reduction in the total number of infants screened.

CONCLUSIONS CO-ROP is a simple model that maintains a statistically similar sensitivity in detecting all grades of ROP while significantly reducing the total number of required ROP screenings compared to 2013 guidelines. The study had a small sample size but shows promise for future research and clinical efforts. (J AAPOS 2016;20:19-24)

Retinopathy of prematurity (ROP) is a disease of abnormal proliferative vascularization of the immature retina and is a leading cause of vision impairment and blindness in children. The Multicenter Trial of Cryotherapy for Retinopathy of Prematurity (CRYO-ROP) study has shown that treatment of infants with severe ROP can decrease the incidence of unfavorable retinal outcomes by 50%. The Early Treatment for Retinopathy of Prematurity (ETROP) study found that treatment of select cases of prethreshold ROP eyes, specifically those meeting type 1 ROP criteria, could further reduce the rate of unfavorable outcomes compared to treatment at threshold severity. The most recent (January 2013) United States screening guidelines recommend screening “infants with a birth weight of \( \geq 1500 \) g, or gestational age of 30 weeks or less (as defined by the attending neonatologist), or selected infants with a birth weight between 1500 and 2000 g or gestational age of >30 weeks with an unstable clinical course, including those requiring cardiorespiratory support and who are believed by their attending pediatrician or neonatologist to be at high risk for ROP.” While these screening criteria are very sensitive, fewer than 10% of the total number of infants identified for screening are eventually treated for ROP. Identification of infants at risk of developing ROP involves serial dilated examinations, a laborious process that requires multiservice coordination and is distressing to the fragile, premature infant. The ROP screening burden (for infants, families, pediatricians, and ophthalmologists) continues to grow as medical advancements in neonatal care have enabled increasing...
numbers of the smallest premature infants to survive.\textsuperscript{9,10} Thus, there is interest in safely improving the efficiency of ROP screening.\textsuperscript{6,7,11-14}

Poor postnatal weight gain has come to the forefront as an additional predictor for the development of ROP.\textsuperscript{6,8,13,15-19} IGF-1 promotes fetal growth and is required for angiogenesis, VEGF signaling, and importantly, growth of the retinal vasculature.\textsuperscript{11,15-17} Poor postnatal weight gain serves as a surrogate measure for a deficiency in the somatic growth factor Insulin Growth Factor-1 (IGF-1) and is thought to be a proxy for overall health and perinatal clinical course.

In an effort to improve the efficiency of ROP examinations investigators who developed the Weight, IGF-1, Neonatal ROP (WINROP)\textsuperscript{13} algorithm and the Children’s Hospital of Philadelphia postnatal weight gain, birth weight, and gestational age ROP risk (CHOP-ROP) model\textsuperscript{6} incorporated weekly postnatal weight gain, birthweight, and gestational age into their prediction models. These models reduced the total number of infants needing to be screened by 77% and 49%, respectively.\textsuperscript{6,8} Further validation studies have shown a 20%-74.5% reduction in total number of infants screened.\textsuperscript{9,13,21-23}

To use either model, weight gain must be calculated and the prediction model applied weekly for each infant until the need for an ROP examination is indicated. In either model, infants need to be followed serially, sometimes up to 10 weeks or more, before an alarm is triggered. These weekly calculations can become time consuming in the setting of busy, high-volume neonatal intensive care units. Furthermore, both WINROP and CHOP-ROP are designed to identify only high-grade ROP with limited analysis on sensitivity of low-grade ROP. Infants with lower-grade ROP are at increased risk for short-term\textsuperscript{24,25} and long-term\textsuperscript{10,24,26,27} visual sequelae associated with prematurity compared to their premature cohorts without ROP. Reduced overall sensitivity for any degree of ROP may limit a models’ use in predicting future visual disability.

The purpose of our study was to propose a new screening model (CO-ROP) incorporating birthweight, gestational age, and postnatal weight gain for ROP. Our goal was to create a set of simple criteria that can accurately predict the risk of developing \textit{any} grade of ROP while significantly reducing the total number of infants screened for ROP. We focused on identifying a postnatal weight, gained within a specific point in time that could be used as additional screening criterion in a manner similar to the national 2013 guidelines.

**Subjects and Methods**

This study was approved by the Colorado Institutional Review Board and complied with the US Health Insurance Portability and Accountability Act of 1996. The medical records of neonates admitted to the neonatal intensive care units (NICU) at the University of Colorado Hospital and Children’s Hospital of Colorado who underwent ROP screening examinations between June 2008 and December 2011 and had a known ROP outcome and available weight data were reviewed retrospectively. Only infants who met 2013 screening criteria were included in the data set.

Patients with type 1 ROP (stage 1 or 2 ROP in zone I with plus disease, stage 3 ROP zone I with or without plus disease, or stage 2 or 3 ROP in zone II with plus disease) were treated with retinal laser photocoagulation in accordance with recommendations from ETROP.\textsuperscript{1} Type 2 ROP, for which observation was indicated, was also defined according to ETROP criteria: stage 1 or 2 ROP in zone I without plus, or stage 3 ROP in zone II.\textsuperscript{4} The initial review was conducted by one ophthalmology resident (JHC). Following this review all cases of type 1 and type 2 ROP and a random selection of records with no ROP were reviewed by a fellowship-trained pediatric ophthalmologist (EAM) to ensure that there was no misclassification of the outcome and that outcome was characterized accurately. Two discrepancies were adjudicated by a second fellowship-trained pediatric ophthalmologist (RSB). Of note, 4 infants developed stage 3, zone III ROP with dilatation and tortuosity of the posterior pole vessels characteristic of plus disease. Based on clinical judgment of the treating ophthalmologist, these neonates underwent laser photocoagulation. However, because they did not meet the official definition of type 1 ROP, they were classified as type 2 ROP.

**Data Collection**

The following information was collected from the medical record: sex, gestational age, stage, zone, and presence of plus disease at each examination, and any treatment required. Weekly weight data (4 weeks = chronological day of life 28 = 1 month, etc) was collected from birth to time of first ROP examination (1 month of age or 30 weeks corrected gestational age, whichever was later).

For the purposes of this study, an infant born at 30 weeks and 6 days was counted as 30 weeks’ gestational age. An “unstable clinical course” was defined as the presence of at least one of the following comorbidities of prematurity: necrotizing enterocolitis, positive pressure ventilation >14 days, sepsis (blood or CSF culture positive sepsis or serious culture-proven infection, excluding urinary tract infection and tracheitis), patent ductus arteriosus requiring surgical ligation, intraventricular hemorrhage grade 3 or higher, or a genetic abnormality (confirmed genetic disorder, or clinical suspicion warranting a chromosomal or microarray analysis).

**Statistical Analysis**

Infants were subdivided into 3 categories based on highest grade of ROP detected in any eye at any examination: high grade (type 1 or 2 ROP), low grade (ROP not meeting type 1 or 2 criteria), or no ROP. Demographic information was compared using \( \chi^2 \) tests for categorical variables and analysis of variance for continuous variables. To identify the weekly net postnatal weight gain that provided the best prediction for the development of any type ROP, the postnatal week (both unadjusted and adjusted for birthweight and gestational age cutoffs) with the highest c-index (area under
A total of 499 neonates met inclusion criteria. Thirty infants (6%) developed type 1 ROP, 15 (3.0%) developed type 2 ROP, 122 (24.4%) developed low-grade ROP, and 332 (66.5%) did not develop any ROP. The baseline demographics for our cohort appear in Table 1.

Model performance was assessed by calculating sensitivities and specificities for detection of high-grade (type 1 and 2), low-grade (ROP not meeting type 1 or type 2 criteria), and overall ROP. Corresponding 95% confidence intervals were calculated using exact Clopper-Pearson type test. Statistical analysis was performed using SAS version 9.4 software (SAS Institute Inc, Cary, NC).

Results

A total of 499 neonates met inclusion criteria. Thirty infants (6%) developed type 1 ROP, 15 (3.0%) developed type 2 ROP, 122 (24.4%) developed low-grade ROP, and 332 (66.5%) did not develop any ROP. The baseline demographics for our cohort appear in Table 1.

Due to transfer of out-born infants into our tertiary referral system, there were 42/1996 (2%) missing weekly weights; none of the infants were missing weights at 1 month of age. For this analysis, all available weights were included. The mean net weight gain at 1 month of age was 225.3 g (95% CI, 205.2–245.3 g) for high-grade (type 1 or 2) ROP and 293.1 g (95% CI 272.1–314.1 g) for any ROP, both of which were significantly lower than the mean 1 month net weight gain for no ROP (mean, 449.4 g; 95% CI, 429.9–468.9 g; \( P < 0.001 \)).

The weight gain at 1 month (28 days) of age was most predictive of the receiver operator characteristic curve) from a logistic regression was chosen as the cutoff time point. A cutoff value for postnatal weight gain was determined by identifying the value at which 100% sensitivity was preserved for high grade ROP with appropriately high sensitivity for low grade ROP. In line with previous models designed to detect high grade ROP only, a subanalysis was performed on the infants who developed high grade ROP (type 1 or type 2).

Model performance was assessed by calculating sensitivities and specificities for detection of high-grade (type 1 and 2), low-grade (ROP not meeting type 1 or type 2 criteria), and overall ROP. Corresponding 95% confidence intervals were calculated using exact Clopper-Pearson type test. Statistical analysis was performed using SAS version 9.4 software (SAS Institute Inc, Cary, NC).

The area inside the box represents the interquartile range (25%-75%). The median and mean are denoted by a line and a circle, respectively. The whiskers extend 1.5 of the interquartile range from the box; observations outside of this range are displayed as points.

Figure 1. Net weight gain of ≤650 g between birth and 1 month of age. A net weight gain of ≤650 g between birth and 1 month of age was chosen as the screening weight cutoff. This cutoff captured all but 2 far outlier infants who developed any grade of ROP. The 400 g cutoff captured all infants who developed high-grade ROP (type 1 or type 2). The area inside the box represents the interquartile range (25%-75%). The median and mean are denoted by a line and a circle, respectively. The whiskers extend 1.5 of the interquartile range from the box; observations outside of this range are displayed as points.
the development of ROP for both the unadjusted and after adjusting for birthweight and gestational age. Both 28 day c-indices are significantly greater than the c-index obtained with birthweight and gestational age alone (Table 2). At 1 month of age, a weight gain cutoff of 650 g was determined after inspection of the distributions across the ROP groups (Figure 1).

The net weight gain cutoff at 4 weeks of age was added to the standard 2013 birthweight (≤1500 g) and gestational age (≤30 weeks) screening model as an “and” criteria. We refer to this revised screening criteria as the CO-ROP screening model (CO-ROP), which suggests ROP screening for infants who are born at ≤30 weeks gestational age and have a birthweight ≤1500 g and achieve net weight gain ≤650 g between birth and 1 month of age.

A secondary analysis of net weight gain at 1 month of age for infants who developed high-grade ROP suggests that a net weight gain cutoff of ≤400 g may serve as a “high risk” alarm for developing high-grade ROP (hgCO-ROP). Table 3 shows that the CO-ROP model had a sensitivity of 100% (95% CI, 92.1%-100%) for high-grade ROP and 96.4% (95% CI, 92.3%-98.7%) for any ROP. The CO-ROP model had a specificity of 33.7% (95% CI, 28.7%-39.1%) for infants who did not develop any ROP (Figure 2). The hgCO-ROP model had a sensitivity of 100% (95% CI, 92.1%-100%) for high-grade ROP, 81.4% (95% CI, 74.7%-87.0%) for any ROP. Importantly, Table 4 shows that the CO-ROP model would have reduced the total number of infants screened by 23.7% compared to 2013 screening guidelines. The infants that developed ROP who were missed by the CO-ROP model all developed low-grade ROP that did not require treatment.

The hgCO-ROP model had a sensitivity of 100% (95% CI, 92.1%-100.0%) for high-grade ROP, 81.4% (95% CI 74.7-82.0%) for any ROP and would have reduced screening to 229/499 (45.9%) infants compared to 2013 screening guidelines (Table 5).

**Discussion**

The CO-ROP model suggests ROP screening for infants who are born at a gestational age ≤30 weeks and have a birthweight ≤1500 g and achieve a net postnatal weight gain ≤650 g between birth and 1 month of age. This model identified 100% of infants who developed high-grade ROP and 96.5% of infants who developed any degree of ROP. Implementing the CO-ROP model would have reduced the total number of infants requiring ROP screening by 23.7% compared to 2013 screening guidelines.

To identify the postnatal weight gain screening criterion, we chose to analyze weekly weights up to 1 month of age, because we were interested in developing a practical model that would identify at-risk infants in a timely manner (1 month) without delaying ROP screening any later than current screening protocol (1 month chronological age or 30 weeks’ gestational age, whichever is later). Individual weekly weight gain measurements for all infants who met 2013 screening criteria were compared using the c-index (area under the receiver operator characteristic curve).
The weight that was most predictive of the development of ROP was at 1 month of age (28 days). The c-index for net weight gain at 28 days, both independently, and after adjusting for birth weight and gestational age, was significantly greater than the c-index obtained with birth weight and gestational age alone (Table 2). All infants who developed type 1 or type 2 in our study ROP (net weight gain range 85–355 g) fell well within our chosen screening criteria of ≤650 g between birth and 1 month of age. A secondary analysis of net weight gain for infants who developed high grade ROP (type 1 or 2 ROP) suggests that a net weight gain ≤400 g cutoff may serve as a “high risk” alarm for developing high grade ROP. Doing so would have decreased the total number of infants screened by 46%.

Our study is distinguished by the inclusion of infants with low-grade ROP. Previously proposed algorithms have been fitted to detect only high-grade ROP. Reduced sensitivity for low-grade ROP limits the use of these previous algorithms as an adjunct to current screening criteria. Although infants with low-grade ROP do not eventually require laser photocoagulation, infants who develop low-grade ROP are at risk of visual sequelae including but not limited to increased incidence of refractive error (particularly myopia), strabismus, and amblyopia compared to their low-birth-weight cohorts who do not develop any ROP.25,27 Furthermore, recognition of long-term vitreoretinal sequelae have begun to emerge as the earliest surviving preterm infants are now entering their fifth and sixth decades of life.10,26,27 It may be premature to implement ROP screening algorithms that eliminate identification of all ROP stages and zones (high and low grade) until further knowledge regarding long-term visual risk are better understood.

The CO-ROP model provides a distinct advantage in its ease of use. With dichotomous criteria for gestational age, birthweight, and postnatal weight gain, our formula eliminates the need for longitudinal weight calculations and complex statistical algorithms requiring a nomogram or calculation program for clinical use. Instead this is replaced with a one-time, one-step, simple calculation in which a ROP alarm is produced in a timely fashion (by 1 month of age) and can be practically implemented by neonatologists and ophthalmologists. Weight at 1 month of age is a single data point that is already collected and readily available. A single weight measurement also eliminates complications associated with gathering longitudinal weight data for infants, which can be a particular challenge for infants transferred from outside facilities. Furthermore, the CO-ROP model provides a one-step, two-tier alarm (ROP alarm at 650 g and a “high risk” ROP alarm at 400 g), further risk stratifying infants for the development of ROP.

The study has some limitations. Like previously proposed models, the CO-ROP model was developed from a single tertiary academic hospital, where infants have a higher ROP risk profile. This cohort does not represent the average demographic of the national at-risk neonatal population. As discussed by Binenbaum and colleagues,11 any proposed ROP screening model is limited by the small cohort of infants with severe ROP. Adoption of any revised ROP screening model will require a high degree of assurance that not a single infant with type 1 ROP will be missed. Therefore, a study including hundreds of infants with severe ROP will be needed to raise the lower boundary of the 95% confidence interval for high-grade ROP to a level closer to 99%-100%.5,11 Furthermore, our findings are not applicable to infants in developing nations where there exists differences in health care systems, patient demographics, and a cohort of older and larger infants who develop ROP and who may represent a different risk ROP profile.11,21 Another limitation of the CO-ROP model is that it screens based on lower-than-predicted weight gain. Thus, an outlier infant with higher-than-average weight gain due to nonphysiologic reasons (edema, sepsis, hydrocephalus) could theoretically be missed. In our study, no infant with severe ROP was missed, and post hoc analysis of the outlier infants with the highest weight gain did not yield any indication of nonphysiologic weight gain. Of the current proposed postnatal weight gain models, the WIN-ROP model is unique in its identification of such infants by automatically screening any infant who gains more than 400 g in a single week.

Our findings should be confirmed with further large-cohort, multi-institutional studies. In the interim we strongly emphasize that the CO-ROP model in its current formulation, as with any of the alternatively proposed novel ROP screening algorithms, can only serve as an adjunct to current ROP screening guidelines. Application of any novel ROP screening model prior to extensive study places one at risk of missing infants with ROP and thus at this stage, these models should be used only with intent to further stratify infants at risk of developing ROP.
The CO-ROP model suggests ROP screening for infants who are born at a gestational age of ≤30 weeks and have a birthweight ≤1500g and achieve a net postnatal weight gain ≥650 g between birth and 1 month of age. Using the CO-ROP model did not miss any infants with high-grade ROP and maintained acceptable sensitivity for infants with any degree of ROP. The CO-ROP model would have reduced the total ROP screening burden by 23.7%. Calibrating to detect for high-grade ROP only, the high-grade CO-ROP model would have reduced the total number of infants screened by 45.9%. The CO-ROP, is a simple, objective, user-friendly model incorporating postnatal weight gain as a one-step, 2-tiered risk stratification that eliminates the need for longitudinal weekly calculations or use of a nomogram or complex proprietary formula.

References