Risk analysis and an alternative protocol for reduction of screening for retinopathy of prematurity

Michael B. Yang, MD, and Edward F. Donovan, MD

PURPOSE

To determine whether a multivariate risk model can select infants with low-risk eyes for an alternative protocol that reduces retinopathy of prematurity (ROP) screening without loss of effectiveness.

METHODS

This was a retrospective, nonrandomized, comparative study. We assigned 712 eyes of 357 premature infants of 401–1,250 g birth weights as high or low risk for prethreshold or threshold ROP using a risk model with variables of birth weight, gestational age, multiple birth, race, and gender. Using simulations, infants with high-risk eyes (p ≥ 0.15) were screened conventionally, while those with low-risk eyes were screened with the 35q3 protocol (initial examination at 35 weeks postmenstrual age followed by screening every 3 weeks, with increased or decreased intervals based on ROP severity). The resultant reduction in ROP screening and the delay to detection of severe ROP were calculated.

RESULTS

The 35q3 protocol reduced the mean number of eye examinations per infant in the overall population by 13.4% (p = 0.0051). No eyes had a delay to the detection of threshold or type 1 zone 1 ROP. Of a total of 57 type 1 ROP eyes, 2 eyes with type 1 zone 2 ROP were delayed to detection by 1 week. As our study was done prior to the Early Treatment of ROP study, the 2 eyes were observed and regressed without progression to threshold ROP.

CONCLUSIONS

A risk-based alternative screening protocol increased the efficiency of ROP screening for infants of 401–1,250 g birth weight without an apparent loss of effectiveness. This investigational approach requires further validation by multicenter studies. (J AAPOS 2009;13:539-545)

Introduction

Timely detection of threshold retinopathy of prematurity (ROP) and high-risk forms of prethreshold ROP is integral to the success of treatment.1-6 The current recommended schedule of screening eye examinations for ROP incorporates findings from several multicenter trials.1,6,7 However, a large number of examinations are necessary to detect the small percentage of eyes that progress to threshold ROP (as defined by the Multicenter Trial of Cryotherapy for ROP [CRYO-ROP])1,4 or type 1 prethreshold ROP (as defined by the Early Treatment of ROP randomized trial [ETROP]).3 Moreover, the eye examination procedure may be associated with complications.8-13 Therefore, it would be desirable to reduce ROP screening if effectiveness could be maintained.

Several studies have advocated reducing the birth weight (BW) and gestational age (GA) cutoffs for ROP screening in premature infants to reduce screening.14-16 However, this approach may result in an unacceptable number of ROP cases greater than stage 3 going undetected.17 Likewise, a simulation study that reduced ROP screening frequency from biweekly to monthly intervals for premature infants <1,251 g BW suggested that additional cases of blindness would occur that nullify the cost savings from a reduction in screening.18 Other studies used a risk model with prognostic variables of BW, GA, and either the duration of oxygen therapy 19 or the number of erythrocyte transfusions20 to select low-risk infants for alternative screening. Low-risk infants were either screened once at 7 weeks chronological age18 or not at all.20 In both studies, no cases of ROP greater than stage 3 were missed, and eye examinations were reduced. However, most of the reduction in examinations was in infants >1,250 g BW that were included in the study and already at very low risk.

The purpose of this study was to determine whether ROP screening for infants of 401–1,250 g BW can be reduced without a loss of effectiveness by applying an alternative screening protocol to infants whose eyes have...
a low risk of prethreshold or threshold ROP based on a predictive model with prognostic variables derived mostly from CRYO-ROP.21

Subjects and Methods

This study was approved by the Institutional Review Boards of the University Hospital of Cincinnati (UHC), Good Samaritan Hospital (GSH), and Cincinnati Children’s Hospital Medical Center (CCHMC) and conforms to the requirements of the United States Health Insurance Portability and Accountability Act. It is a retrospective, observational study of premature infants of 401–1,250 g BW admitted to the neonatal intensive care units (NICU) of UHC and GSH from January 1998 through May 2003, when threshold ROP as defined by CRYO-ROP21,4 was the standard of treatment instead of type 1 prethreshold ROP as defined by ETROP.3 Some infants born at UHC and GSH required additional care and were transferred to CCHMC. Nearly all infants discharged from the NICUs received eye examinations as outpatients at CCHMC. Eye examination results, the CRYO-ROP prognostic variables21 (BW, GA, multiple birth, birth location inside the study hospital [inborn status], and race), and gender were abstracted by chart review for each infant.

A total of 357 infants (712 eyes) from UHC and 491 infants (982 eyes) from GSH without ambiguous genitalia met the inclusion criterion of having an adequate sequence of eye examinations to determine whether prethreshold or threshold ROP developed. At UHC, 2 infants each had only 1 eye that met the inclusion criteria, hence, 712 instead of 714 eyes.22 The definitions of prethreshold, prethreshold, and types 1 and 2 prethreshold ROP were from CRYO-ROP21,4 and ETROP3 (Table 1).

A predictive model for eye outcome was developed using multiple logistic regression (SPSS 15.0 Advanced Models; SPSS Inc, Chicago, IL) for the infants from UHC. Generalized estimating equations accounted for correlation between eyes of each infant.21 The model included BW, GA, multiple birth, inborn status, race, and gender as predictive variables. Presence or absence of prethreshold or threshold ROP was the binary dependent variable. Model predictive capacity was calculated as area under receiver operating characteristic curves. A probability >0.15 of prethreshold or threshold ROP was selected as the cutoff for high risk as it resulted in the best combination of sensitivity and specificity.

All infants were screened by the conventional protocol as described below. The examiners were pediatric ophthalmologists who had completed training in fellowship programs certified by the American Association for Pediatric Ophthalmology and Strabismus. Rarely a retina specialist confirmed the retinal findings for treatment. The risk model derived from UHC premature infants was applied to the same infants from UHC and separately to the GSH infants. Infants with high-risk eyes (p < 0.15) and low-risk eyes (p > 0.15) were selected for conventional and alternative (35q3 protocol) screening, respectively (Figure 1).

By convention at UHC, GSH, and CCHMC, the initial screening of surviving infants generally occurred at 5 or 6 weeks chronological age or 32 weeks postmenstrual age (PMA), whichever was later. The interval between follow-up screening examinations paralleled published guidelines before 200324,25 but with a bias toward the original CRYO-ROP4,26 protocol: prethreshold ROP in any zone or stage 3 in zone 3 (≥1 week); immature vascularization in zone 1 [1 to 2 week(s)]; ROP less than prethreshold in zone 2, immature vascularization in zone 2, or stage 2 in zone 3 (2 weeks); immature vascularization or stage 1 in zone 3 (4 to 6 weeks). The actual sequence of eye examinations was modified retrospectively by applying new criteria for termination of screening (e-Supplement 1, section A, available at jaapos.org) published in 2002,2,3 near the end of the time period of our study. This approach avoids an artificially high number of eye examinations for conventional screening due to variability in examiners' adoption of the new criteria.

Alternative screening of infants with low-risk eyes was simulated using the 35q3 protocol. Eyes were screened initially at 35 weeks PMA or 5 or 6 weeks of chronological age, whichever is later. Subsequent examinations occurred at 3 week intervals if eye examinations showed immature vascularization or less than prethreshold ROP in zone 2. The interval between examinations was decreased or increased depending on whether more or less severe ROP was detected (e-Supplement 1, section B). If immature vascularization in zone 1 or prethreshold ROP was detected, screening reverted to conventional screening for all subsequent examinations. The new termination of screening criteria (e-Supplement 1, section A) was also applied, with 1 addition: screening was terminated for eyes with stage 1 ROP in zone 3 if infants were also ≥35 weeks PMA, since none of these eyes progressed beyond stage 2 ROP in our study population. (Additional details of the simulation are found in e-Supplement 1, sections C and D.)

To determine the efficiency of alternative screening, we calculated the percentage reduction in the mean number of eye examinations per infant performed on infants with low-risk eyes by the 35q3 protocol as compared with conventional screening. To determine the percentage reduction in eye examinations for the total population of infants due to alternative screening of the subset of infants with low-risk eyes, we calculated the mean number of eye examinations per infant for all infants: those with high-risk eyes screened conventionally and those with low-risk eyes screened by the 35q3 protocol. The result was compared with the mean number of eye examinations per infant when all infants were screened conventionally. Independent samples t-tests were performed.

Table 1. Definitions of threshold and prethreshold ROP

<table>
<thead>
<tr>
<th>Category</th>
<th>Inclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threshold ROP</td>
<td>≥5 contiguous or ≥8 cumulative clock hours of stage 3 in zone 1 or 2 with plus disease</td>
</tr>
<tr>
<td>Prethreshold ROP</td>
<td>zone 1 stage 1, 2, or 3 with or without plus disease, zone 2 stage 2 with plus disease, or zone 2 stage 3 with or without plus disease; extent of stage 3 less than threshold criteria</td>
</tr>
<tr>
<td>Type 1 prethreshold</td>
<td>zone 1 stage 1, 2, or 3 with plus disease, zone 1 stage 3 without plus disease, or zone 2 stage 2 or 3 with plus disease</td>
</tr>
<tr>
<td>Type 2 prethreshold</td>
<td>zone 1 stage 1 or 2 without plus disease or zone 2 stage 3 without plus disease</td>
</tr>
</tbody>
</table>


analysis of variance and accounted for the correlation between fellow eyes that progressed to prethreshold or threshold ROP outcome. Statistical significance for these comparisons was determined by binomial exact confidence intervals. For results from ETROP, only a few type 1 prethreshold ROP eyes were treated, so greater than 90% of such eyes could be observed for progression to threshold ROP or spontaneous regression. The percentage of eyes with a delay to detection of severe ROP was analyzed by binomial exact confidence intervals. For eyes that developed type 1 prethreshold ROP, the mean PMA at detection of type 1 prethreshold ROP by the 35q3 protocol was compared with conventional screening. The onset of various categories of severe ROP was compared for high- versus low-risk eyes that progressed to prethreshold or threshold ROP outcome. Statistical significance for these comparisons was determined by analysis of variance and accounted for the correlation between fellow eyes. For all statistical tests, \( p \)-value was considered statistically significant if \( < 0.05 \).

Results

The baseline characteristics of infants from UHC included those in CRYO-ROP except the percentage of infants born in the study hospital was lower (81.8%) and the percentage of singleton births was higher (81.4%) in CRYO-ROP.

Accuracy of Risk Models

The results of multiple logistic regressions for the model of prethreshold or threshold ROP are shown in Table 2. Inborn status was not included in the final model because it was not associated with the outcome of prethreshold or threshold ROP \( (p = 0.998) \). Lower BW, younger GA at birth, and non-black race were significant predictors of eyes developing prethreshold or threshold ROP. Multiple birth showed a trend as a predictor, but male gender was not predictive.

The overall model predictive capacity was high with an area under the receiver operating characteristic curve of 0.877 (Table 3). As shown in Table 4, for a probability cutoff of 0.15, the observed specificity of 72.7% reflected the large number of low-risk/disease-absent eyes that were likely to benefit from a reduction in screening. The high sensitivity of 90.3% reflected the large number of high-risk/disease-present eyes that would benefit from conventional screening. Of the eyes that developed prethreshold or threshold ROP, 14 (9.7%) were inaccurately categorized as low risk and could potentially have been adversely affected by a reduction in screening using an alternative screening protocol.

Reduction in ROP Eye Examinations

The 35q3 alternative screening protocol resulted in a statistically significant reduction in the mean number of eye examinations performed per infant for infants with low-risk eyes and for the overall population as compared with conventional screening (Table 5).

Delay to Detection of Severe ROP

There was no delay to the detection of threshold ROP as defined in CRYO-ROP for the infant population from UHC using the 35q3 protocol as compared with conventional screening (Table 6). In addition, 93.0% of all prethreshold ROP, 96.5% of type 1 prethreshold ROP, and 100% of type 1 zone 1 prethreshold ROP were detected on time. These results were not corrected for correlation between fellow eyes, but the results based on infant calculations are listed in e-Supplement 2 (available at jaapos.org).

Of the 14 eyes with prethreshold or threshold ROP that were incorrectly categorized as low risk by our model, 8 (4 patients) eventually progressed to threshold ROP.

---

**Table 2. Clinical characteristics of premature infants from University of Cincinnati**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of infants/eyes</td>
<td>357 / 712</td>
</tr>
<tr>
<td>Birth weight (g) (mean ± SD)</td>
<td>933 ± 192</td>
</tr>
<tr>
<td>Gestational age (weeks) (mean ± SD)</td>
<td>27.2 ± 2.1</td>
</tr>
<tr>
<td>Born in the study hospital (%)</td>
<td>96.9</td>
</tr>
<tr>
<td>Singleton births (%)</td>
<td>74.2</td>
</tr>
<tr>
<td>Non-black race (%)</td>
<td>60.5</td>
</tr>
<tr>
<td>Male (%)</td>
<td>48.2</td>
</tr>
<tr>
<td>Threshold ROP (infants/eyes)</td>
<td>33 / 63</td>
</tr>
<tr>
<td>Type 1 prethreshold or threshold ROP (infants/eyes)</td>
<td>51 / 98</td>
</tr>
<tr>
<td>Prethreshold or threshold ROP (infants/eyes)</td>
<td>75 / 146</td>
</tr>
</tbody>
</table>

\( SD \), standard deviation.
Sensitivity 90.3%
Specificity 72.7%

(2 patients) and/or type 1 prethreshold ROP (3 patients), while none developed type 1 zone 1 prethreshold ROP. Of the 8 prethreshold eyes that progressed, 2 eyes (1 patient) progressed to type 1 zone 2 prethreshold ROP and then subsequently threshold ROP in zone 2. Two other prethreshold (type 2) eyes (1 patient) progressed to threshold ROP in zone 2 without an intervening type 1 prethreshold ROP being detected. The remaining 4 eyes (2 patients) progressed to type 1 zone 2 prethreshold ROP and spontaneously regressed without treatment.

For prethreshold eyes incorrectly assigned as low risk by our model and screened according to the 35q3 protocol, 11 (79%) of 14 prethreshold ROP eyes (6/8 patients) had ≥1 week delay to detection of prethreshold ROP (Figure 2). Three of 14 (21%) eyes (2/8 patients) had a delay to detection of prethreshold ROP of 2 weeks, but these 3 eyes had type 2 prethreshold ROP. Only 2 (33%) of 6 low-risk eyes (1 of 3 patients) with type 1 prethreshold ROP had a delay to detection (1 week); both eyes were in zone 2 and spontaneously regressed without progression to threshold ROP or unfavorable structural outcome (Figure 2).

When all 57 eyes (high and low risk) with type 1 prethreshold ROP were analyzed, the mean PMA at which type 1 prethreshold ROP was detected by conventional screening and by alternative screening (35q3 protocol) were 36.56 ± 2.83 and 36.60 ± 2.87 weeks, respectively (0.04 week difference, \( p = 0.966 \)). When only the 6 low-risk eyes with type 1 prethreshold ROP were analyzed, the results for conventional and alternative screening were 39.33 ± 0.52 and 39.67 ± 0.52 weeks, respectively (0.34 week difference, \( p = 0.537 \)).

### Difference in Onset of Severe ROP

The mean onset of prethreshold ROP was about 3.4 weeks later for low-risk as compared with high-risk eyes (\( p = 0.006 \)). The results for the various subcategories of prethreshold and threshold ROP were not statistically significant or showed only a trend (e-Supplement 3, available at jaapos.org).

### Application of Risk Model to a Different Infant Population

The UHC risk model was applied to eyes of GSH infants to validate the model. A similarly high sensitivity (89.4%) was
obtained with similarly few eyes (19, 10.6%) that were incorrectly categorized as low risk (e-Supplement 4, available at jaapos.org). A slightly lower specificity (68.2%) was achieved as compared with UHC infants. By estimation, this should result in a slightly lower percentage reduction in the mean number of eye examinations per infant achieved by the same 35q3 protocol as compared with the UHC results. Using the 35q3 protocol, there were 6 eyes (3.7%) with a 1 to 2 week delay to the detection of prethreshold ROP while no eyes with type 1 prethreshold, type 1 zone 1 prethreshold, or threshold ROP had a delay to detection (Table 6 and e-Supplement 5, available at jaapos.org).

Discussion

Our study suggests that it may be possible to reduce ROP screening examinations for premature infants of 401–1,250 g BW without a loss of effectiveness by using a multivariate model of risk to select infants with eyes at low risk for screening by an alternative protocol. However, one has to weigh the risks and benefits of this approach. The reduction in eye examinations produces cost savings that can be nullified by costs associated with an increase in the number of bilaterally blind infants when treatment is delayed due to a delay in detection of severe ROP in even a small percentage of infants.

Since we studied a population that was screened prior to the publication of the ETROP study and for which type 1 prethreshold ROP was not a standard for laser therapy, we could analyze whether those eyes progressed to threshold ROP or regressed spontaneously. While we found no delay to the detection of threshold or type 1 zone 1 prethreshold ROP (since no eye that developed type 1 zone 1 prethreshold ROP was incorrectly categorized as low risk), there was a delay to detection of type 1 zone 2 prethreshold ROP for 2 eyes using the 35q3 protocol. However, early treatment of type 1 zone 2 prethreshold ROP remains controversial, and both of these type 1 zone 2 prethreshold ROP eyes spontaneously regressed with favorable structural outcome at 3 months. (Of note, a more conservative protocol as described in e-Supplement 6, available at jaapos.org, eliminated any delay to detection of type 1 prethreshold ROP, albeit at the expense of smaller percentage reduction in ROP screening at 10.2%). Nevertheless, due to the small numbers in our study and the hypothetical nature of simulations, there is some uncertainty as to the percentage of eyes likely to experience a delay to detection of severe ROP.

Delaying the initial examination by 3 weeks to 35 weeks in the alternative screening protocol may raise concern that aggressive posterior ROP would be missed by this screening strategy. However, infants who develop aggressive posterior ROP would likely be stratified by the risk model to the conventional, high-risk screening protocol that screens initially at 32 weeks PMA.

The overall reduction in eye examinations may appear to be modest—13.4% for the 35q3 protocol. However, if the recent screening guidelines (which tend toward more frequent screening than our general practice) were used, then the percentage reduction in screening examinations achieved by the alternative protocols would have been even greater. Moreover, other risk models, which include the Clinical Risk Index for Babies, serum IGF-1 levels, and/or rate of infant growth during the first 6 weeks of life, may be more accurate and allow further reductions in ROP screening. However, restricting model variables to demographic or infant

Table 6. Percentage of eyes with severe ROP experiencing a delay to detection due to simulation of 35q3 alternative screening protocol for infants with low-risk eyes and conventional screening for infants with high-risk eyes

<table>
<thead>
<tr>
<th>Severe ROP type</th>
<th>University Hospital of Cincinnati</th>
<th></th>
<th>Good Samaritan Hospital</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Proportion of eyes (%)</td>
<td>95% CI (%)</td>
<td>Proportion of eyes (%)</td>
<td>95% CI (%)</td>
</tr>
<tr>
<td>Prethreshold</td>
<td>9/129 (7.0)</td>
<td>3.24-12.8</td>
<td>6/164 (3.7)</td>
<td>1.40-7.8</td>
</tr>
<tr>
<td>Type 1</td>
<td>2/57 (3.5)</td>
<td>0.43-12.1</td>
<td>0/44 (0.0)</td>
<td>0.00-8.0</td>
</tr>
<tr>
<td>Type 1 zone I</td>
<td>0/24 (0.0)</td>
<td>0.00-14.2</td>
<td>0/23 (0.0)</td>
<td>0.00-14.8</td>
</tr>
<tr>
<td>Threshold</td>
<td>0/63 (0.0)</td>
<td>0.00-5.7</td>
<td>0/96 (0.0)</td>
<td>0.00-3.8</td>
</tr>
</tbody>
</table>

The risk model derived from eyes of premature infants at University Hospital of Cincinnati (UHC) was applied to eyes of premature infants from UHC and from Good Samaritan Hospital. These results were not corrected for inter-eye correlation. The results based on infant calculations are found in e-Supplement 2. CI, binomial exact confidence interval.

FIG. 2. Delay to detection of prethreshold ROP (retinopathy of prematurity) when infants with low-risk eyes from University Hospital of Cincinnati were subjected to simulated screening with the 35q3 alternative screening protocol.
characteristic variables is advantageous because it avoids the necessity of obtaining serial serum samples and calculating other risk factors such as length of time on oxygen, number of erythrocyte transfusions, and components of the Clinical Risk Index for Babies score, which may be difficult.

We did not include infants >1,250 g BW in our study because those infants were typically discharged from the NICU prior to their first eye examinations. Compliance with follow-up examinations and completeness of records were thus an issue. However, it should be possible to extend our approach to the larger population of infants >1,250 g BW, thus allowing greater reductions in examinations. In developing nations where qualified examiners for ROP may be scarce relative to the need and where severe ROP occurs with greater frequency in infants >1,250 g BW as compared with industrialized nations, our approach may help concentrate screening on the infants most likely to develop severe ROP.

Reduction of ROP screening may decrease the exposure of premature infants to complications associated with screening examinations. However, given the many interventions that occur during an infant’s stay in the nursery, it would be difficult to ascertain the contribution of eye examinations to overall long-term morbidity of an infant. Thus, by itself, the reduction of complications associated with eye examinations may be an insufficient justification for reducing eye examinations unless high efficiency and efficacy of an alternative screening protocol is demonstrated. Even then, it may be argued that some built-in redundancy offered by conventional screening may protect against missed follow-up examinations and potential examiner error. In addition, some researchers have called for an increase in ROP screening frequency to detect type I prethreshold ROP earlier. If that approach prevails, the use of predictive models may be helpful by restricting more frequent screening to infants with high-risk eyes.

Our study is limited by its single-center, retrospective design. The successful application of our risk model to a premature infant population from a second center in the same city shows promise in reducing ROP screening without loss of effectiveness. However, these results should be considered investigational at this time and will require further validation by a multicenter study to account for center and physician variability in examination technique and scheduling.

References


