Applicability of the ROPScore as a Predictive Algorithm for Early Detection of Retinopathy of Prematurity

Ricardo Figueiredo1; Tiago Moreira Sarmento2; João Garrido3; M. Inês Nunes Marques4; Teresa Almeida5; Sara Carrasquinho6
1Department of Ophthalmology, Hospital de Espírito Santo de Évora, Évora, Portugal
2Department of Pediatrics, Hospital de Espírito Santo de Évora, Évora, Portugal

ABSTRACT

Purpose: To assess the accuracy of the ROPScore algorithm as a predictor of retinopathy of prematurity (ROP) by the second week of life.

Materials and Methods: Retrospective cohort study of 239 preterm infants with a gestational age (GA) ≤32 weeks and/or birthweight (BW) ≤1500 g. No ROP, any stage of ROP and severe ROP requiring treatment were categorized. ROPScore was calculated in the second week of life using the following parameters: GA, BW, weight by the second week of life, use of oxygen in mechanical ventilation and use of blood transfusions. Sensitivity, specificity and positive (PPV) and negative (NPV) predictive values were calculated. The best cut-offs of the algorithm were calculated and, whenever possible, a sensitivity of 100% was chosen.

Results: Mean BW was 1241.6±310.0 g and mean GA was 29.8±3.4 weeks. Of the 239 infants, ROP was identified in 101 (42.3%) and 12 (11.9%) had severe ROP requiring treatment. Mean ROPScore was 15.39±1.94 in the any stage of ROP cohort and 17.52±1.80 in the severe ROP cohort. The sensitivity of the algorithm was 96.0% and 100% for any stage of ROP and severe ROP, respectively. NPV was 92.3% and 100% for any stage of ROP and severe ROP, respectively.

Conclusion: ROPScore is a simple public domain algorithm that can be applied as early as by the second week of life. It may be used as a reliable screening tool for ROP and optimize examination timings, but further validation is required.

Keywords: retinopathy of prematurity; ROP; ROPScore; algorithm

RESUMO

Objetivo: Avaliar a precisão do algoritmo ROPScore como preditor de retinopatia da prematuridade (ROP) à segunda semana de vida.

Material e Métodos: Estudo de coorte retrospectivo de 239 prematuros com idade gestacional (IG) ≤32 semanas e/ou peso de nascimento (PN) ≤1500g. Os doentes foram categorizados como
sem ROP, ROP de qualquer estádio e ROP grave com necessidade de tratamento. Calculou-se o ROPSscore à segunda semana de vida utilizando os parâmetros: IG, PN, peso à segunda semana de vida, uso de oxigênio em ventilação mecânica e transfusões sanguíneas. Foram calculados a sensibilidade, especificidade e valores preditivos positivo (VPP) e negativo (VPN). Calcularam-se os melhores pontos de corte do algoritmo e escolhida uma sensibilidade de 100% sempre que possível.

Resultados: O PN médio foi de 1241.6±310.0g e a IG média foi de 29.8±3.4 semanas. Dos 239 prematuros, identificou-se ROP em 101 (42.3%) e 12 (11.9%) apresentavam ROP grave com necessidade de tratamento. O ROPSscore médio foi de 15.39±1.94 na coorte de ROP de qualquer estádio e de 17.52±1.80 na coorte de ROP grave. A sensibilidade do algoritmo foi de 96.0% e de 100% para ROP de qualquer estádio e ROP grave, respectivamente. O VPN foi de 92.3% e 100% para ROP de qualquer estádio e ROP grave, respectivamente.

Conclusão: O ROPSscore é um algoritmo simples, de domínio público, que pode ser aplicado a partir da segunda semana de vida. Pode ser utilizado como método confiável de triagem para o desenvolvimento de ROP e optimizar os tempos de observação, apesar de ser necessária validação adicional.

Palavras-chave: retinopatia da prematuridade; ROP; ROPSscore; algoritmo

INTRODUCTION

Retinopathy of prematurity (ROP) is a vasoproliferative retinal disorder that affects premature infants, and is the leading cause of preventable childhood blindness in high- and middle-income countries, although it may be underdiagnosed in developing countries. A number of risk factors for ROP have been described, the most consistent being low birth weight (BW) and gestational age (GA); others include postnatal weight gain, prolonged oxygen therapy, pulmonary diseases, blood transfusions, sepsis and hemodynamic changes. Oxygen-dependent growth factors, such as vascular endothelial growth factor (VEGF), and deficiency of non-oxygen dependent growth factors, such as insulin-like growth factor-1 (IGF-1) also play a primary role in the pathophysiology of ROP.

Substantial improvements in neonatal care have been made in the recent past and, while improving survival rates in preterm infants with very low weight, they have also increased the incidence of ROP. To prevent blindness caused by ROP, appropriate screening of infants at risk, prompt diagnosis and timely treatment is critical for successful outcomes. Current screening guidelines are based primarily on BW and GA, and recommend screening all infants with BW of 1500 g or less and/or GA of 32 weeks or less. These criteria, while having a high sensitivity to detect patients needing treatment, result in an excessive number of examinations, which may be costly and can lead to instability in patients with other comorbidities.

The increasing number of infants requiring screening for ROP, added to the shortfall availability of physicians, have led to the development of predictive models and algorithms to identify and screen the infants who are at the highest risk for development of severe ROP requiring treatment. These algorithms take in account risk factors other than GA and BW to allow a better tailoring of risk stratification, and may be used to increase efficiency in the screening of preterm infants, having shown improved specificity to detect high-risk patients.

ROPSscore is a simple scoring algorithm proposed by Eckert et al to predict severe ROP. It only needs to be calculated once per infant at the 2nd or 6th weeks of life, using a public domain Microsoft Excel spreadsheet with easily obtainable parameters: GA, BW, weight by the second week of life, use of oxygen in mechanical ventilation and use of blood transfusions. The purpose of
this study is to assess the accuracy of the ROPScore algorithm as a predictor of ROP by the second week of life in a secondary center in Portugal.

MATERIALS AND METHODS

Study design

We performed a retrospective cohort analysis of all very low birth weight (VLBW) infants admitted and screened for ROP between January 2008 and March 2019. The data used for this study was collected at the Department of Ophthalmology and at the Neonatal Unit of the Department of Pediatrics of the Hospital do Espírito Santo de Évora, Évora, Portugal.

The study adhered to the principles of the Declaration of Helsinki and board approval was obtained from the institutional ethics committee.

Study subjects

According to the Portuguese nationwide screening program for ROP, preterm infants with a GA ≤ 32 weeks and/or BW ≤ 1500 g, as well as premature infants with BW < 2000 g with prolonged need for supplementary oxygen therapy, critically ill newborns or newborns who were submitted for major surgery were screened. All infants were initially screened at 31 to 33 weeks of GA or 4 to 6 weeks of chronological age, whichever was later, and then following accordingly to exam findings. Fundus examinations were performed and recorded with RetCam® (Natus Medical Incorporated, USA), after pupillary dilation with tropicamide 0.5% and phenylephrine 2.5%.

The only exclusion criterion used was infant death before completing 6 weeks of life or before reaching 45 weeks of corrected GA.

Study procedures

Data regarding gender, GA, BW, weight by the second week of life, use of oxygen in mechanical ventilation, use of blood transfusions were retrospectively collected through analysis of medical records by the second week of life of the subjects. Clinical outcomes were defined and categorized as no ROP, any stage of ROP and severe ROP requiring treatment, always corresponding to the highest stage of ROP observed during patient follow-up. Staging was recorded according to the International Classification of Retinopathy of Prematurity Revised23 and severe ROP requiring treatment was defined according to the Early Treatment for Retinopathy of Prematurity trial15 as type 1 ROP (stage 2 or 3 in zone II with plus disease, stage 3 in zone I with or without plus disease, or stage 1 or 2 disease in zone I with plus disease), or worse ROP.

ROPScore was calculated using GA, BW, weight by the second week of life, use of oxygen in mechanical ventilation, and use of blood transfusions in the public domain Microsoft® Excel spreadsheet (Figure 1).

Statistical analysis

Statistical analysis was performed using the Statistical Package for the Social Sciences (SPSS), version 21.0 for Windows (SPSS Inc., Chicago, IL, USA).

The predictions of any stage of ROP and severe ROP were calculated using the area of the receiver operation characteristic (ROC) curve. Sensitivity, specificity and positive (PPV) and negative (NPV) predictive values of the ROPScore were calculated. The best cut-offs of the algorithm were calculated and, whenever possible, a sensitivity of 100% was chosen. Student t-test and the chi-squared test were used to compare the groups. Tests were considered statistically significant with a p value of <0.05.
RESULTS

A total of 239 infants (121 male) met the criteria for inclusion in the study. Mean BW was 1241.6 ± 310.0 g and mean GA was 29.8 ± 3.4 weeks in the total cohort. The prevalence of any stage of ROP was 101/239 (42.3%) and the incidence of severe ROP requiring treatment was 12/239 (11.9%). Mean ROPscore was 15.39 ± 1.94 in the any stage of ROP group and 17.52 ± 1.80 in the severe ROP group, the higher the score, the higher the risk for developing ROP. The baseline demographics and clinical characteristics of the cohort are described in Table 1.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Total Cohort</th>
<th>Any Stage of ROP</th>
<th>Severe ROP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients, n (% of total)</td>
<td>239 (100)</td>
<td>101 (42.3)</td>
<td>12 (11.9)</td>
</tr>
<tr>
<td>Male gender, n (%)*</td>
<td>232 (50.6)</td>
<td>49 (48.5)</td>
<td>7 (58.3)</td>
</tr>
<tr>
<td>BW, g*</td>
<td>1241.6 ± 310.0</td>
<td>1042.4 ± 248.9</td>
<td>958.5 ± 321.2</td>
</tr>
<tr>
<td>GA, weeks*</td>
<td>29.8 ± 3.4</td>
<td>28.4 ± 2.1</td>
<td>26.3 ± 1.5</td>
</tr>
<tr>
<td>O₂ in mechanical ventilation, n (%)</td>
<td>174 (72.8)</td>
<td>96 (95.0)</td>
<td>12 (100)</td>
</tr>
<tr>
<td>Blood transfusions, n (%)</td>
<td>27 (11.3)</td>
<td>20 (19.8)</td>
<td>9 (75)</td>
</tr>
<tr>
<td>ROPscore (range)*</td>
<td>13.60 ± 2.46</td>
<td>15.39 ± 1.94</td>
<td>17.52 ± 1.80</td>
</tr>
</tbody>
</table>

Table 1 - Demographic and clinical characteristics of the patients by group.
*Data are expressed as mean ± standard deviation. BW: birth weight; GA: gestational age; ROP: retinopathy of prematurity.

The accuracy of the ROPscore as a predictor of the onset of any stage of ROP and severe ROP was determined by ROC curves (Figure 2), and cut-off points were obtained for sensitivity and specificity. The area under the curve (AUC) was 0.878 and 0.819 for any stage of ROP and severe ROP, respectively. The best cut-off point established for any stage of ROP was 13.7 (86.1% sensitivity and 78.8% sensitivity), while that for severe ROP was 16.3 (75.0% sensitivity and 79.8% specificity). As a sensitivity of 100% was prioritized, when possible, to detect true positives, a cut-off of 11.4 was chosen for any stage of ROP (96.0% sensitivity and 35.0% specificity) and of 14.7 for severe ROP (100% sensitivity and 39.3% specificity). Sensitivity, specificity, PPV and NPV for any stage of ROP and severe ROP are described in Table 2.

![ROC Curve - Any stage of ROP](image1)

Figure 2 - Receiver operating characteristic (ROC) curves for detection of any type of retinopathy of prematurity (ROP) (left) and severe ROP (right), according to ROPscore algorithm. For each curve, the value of area under the curve (AUC) and the 95% confidence interval are indicated, the latter in brackets.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Any Stage of ROP</th>
<th>Severe ROP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity, % (95% CI)</td>
<td>96.0 (89.8-98.7)</td>
<td>100.0 (71.3-100.0)</td>
</tr>
<tr>
<td>Specificity, % (95% CI)</td>
<td>35.0 (27.6-43.4)</td>
<td>39.3 (29.8-49.7)</td>
</tr>
<tr>
<td>Positive predictive value, % (95% CI)</td>
<td>52.2 (44.7-59.5)</td>
<td>18.2 (10.1-30.0)</td>
</tr>
<tr>
<td>Negative predictive value, % (95% CI)</td>
<td>92.3 (80.6-97.5)</td>
<td>100.0 (87.7-100.0)</td>
</tr>
</tbody>
</table>

Table 2 - Prediction of any type of ROP and severe ROP with ROPscore.
*CI: confidence interval; ROP: retinopathy of prematurity.
DISCUSSION

In this study, ROPScore correctly identified all 101 infants who developed ROP and the 12 infants who developed severe ROP. Using the standard cut-off’s proposed by Eckert et al.\textsuperscript{21} (11 for any stage of ROP and 14.5 for severe ROP), all babies with ROP and severe ROP would also be identified in our study. However, adjusting the cut-off according to our results (11.4 for any stage of ROP and 14.7 for severe ROP) allowed false positives to drop from 96 to 89 with any stage of ROP and from 59 to 54 with severe ROP. The NPV calculated in this study indicated that the probability of an infant to not develop any stage of ROP or severe ROP if they scored ≤ 11.4 and ≤ 14.7, respectively, would be 92.3% and 100.0%, respectively. If the ROPScore was applied to this population, 50 of the preterm infants would not need fundus examination with the same frequency, which would theoretically result in a decrease of 20.9% of the total number of examinations to detect ROP.

ROPScore has been tested in a number of studies.\textsuperscript{11,20,21,24,25} In Brazil, where the ROPScore was developed, two prospective studies\textsuperscript{11,21} have demonstrated a sensitivity of 87.5-94.0% and 95.4-96.0%, while a retrospective study\textsuperscript{24} demonstrated a sensitivity of 98.6% and 100.0%, for any stage of ROP and severe ROP, respectively. In a 9-year retrospective analysis in Japan,\textsuperscript{25} ROPScore alone demonstrated an AUC of 0.935 (95% CI, 0.904-0.963), which was improved further with the addition of the predicting variables late-onset circulatory collapse and continuous positive airway pressure. In Italy, a multicenter study\textsuperscript{20} reported a sensitivity of 100% for predicting severe ROP, with an AUC of 0.93 (95% CI, 0.90-0.96). While some of ROPScore outcomes may be superior in these studies when compared to ours, population characteristics varied significantly, with most studies having lower incidences\textsuperscript{11,20,21} of any stage of ROP and severe ROP (17.6-26.1%, 5.0-12.1%, respectively), while some having higher incidences\textsuperscript{24,25} (46.7-68.3%, 17.0-18.2%, respectively), and no other study included infants with a BW > 1500 g. This adds to the importance of adjusting the cut-offs according to each population for improvement of outcomes.

Several other predictive models have been developed for prediction of clinically significant ROP,\textsuperscript{19} hereby presented in chronological order.

The Safety-Index Model, published in 1996,\textsuperscript{26} was based on BW, GA, and the number of days on oxygen. As the application of the model would have missed ROP in 23 infants, including 1 with stage 3 ROP, the authors then modified the model, which would not miss any case of ROP and would reduce the number of examinations by 10\.\textsuperscript{19}

The Termote Model, proposed in 2005,\textsuperscript{27} used BW, GA and number of erythrocyte transfusions as risk factors. Sensitivity and specificity to detect severe ROP were 100% and 24.5%, respectively, and 23% of the infants could have been excluded from screening.

The Weight, IGF-1, Neonatal, Retinopathy of Prematurity Model (WINROP) is an online surveillance system, originally published in 2006,\textsuperscript{28} based on weekly recordings of postnatal weights and IGF-1 serum levels until an alarm is called. This model was later modified\textsuperscript{7} (WINROP 2), eliminating serum IGF-1 levels from analysis, to reduce costs and stress on infants. While this model has been validated by several studies,\textsuperscript{19} some have shown that it does not perform well in low-income countries, where older preterm infants are at risk, as it does not calculate the risk for infants with a GA > 32 weeks.\textsuperscript{29-32}

The Yang Model, published in 2009,\textsuperscript{33} is a multivariate risk model that used BW, GA, weight gain, multiple birth, race and gender as variables, and presence or absence of prethreshold or threshold ROP as outcomes. Results demonstrated a reduction of 13% in the number of examinations, but it missed 9 cases of severe ROP.

The Denmark Retinopathy of Prematurity Model, published in 2011,\textsuperscript{34} used BW and GA as predicting variables and presence or absence of treatment-demanding ROP as outcomes, with positive results. However, treatment-demanding ROP was threshold ROP in several cases, while current treatment guidelines use type 1 ROP.\textsuperscript{19}

The Premature Infants in Need of Transfusion Retinopathy of Prematurity (PINT-ROP) model, published in 2011,\textsuperscript{35} used GA, BW and daily weigh gain rate calculated from the current and previous weeks’ weights. It correctly predicted all infants with severe ROP except one, who did not require treatment, and it would have resulted in 30% fewer infants examined.

The Children’s Hospital of Philadelphia Retinopathy of Prematurity (CHOP-ROP) model, published in 2012,\textsuperscript{36} is a simple clinical model that uses nomograms that can be accessed in public domain. It originated from the PINT-ROP model, and uses GA, BW and daily weight gain rate as variables, starting from the second week of life.
The Netherlands Retinopathy of Prematurity Study Model (NEDROP), described in 2013, included the predictive variables of infants with GA < 30 weeks, BW < 1250 g, or both; or GA 30–32 weeks, BW of 1250–1500 g, or both, with one or more of the following risk factors: artificial ventilation, necrotizing enterocolitis, sepsis, postnatal glucocorticoids, or cardiotonia. Sensitivity and specificity to detect severe ROP were 100% and 20%, respectively, and would have reduced the number of infants requiring examinations by 20%.

In 2015, Ying et al proposed the Telemedicine Approaches to Evaluating Acute-Phase Retinopathy of Prematurity (e-ROP) model. The predicting variables were gender, race or ethnicity, BW, GA, quadrants of preplus, stage of ROP, presence of retinal hemorrhage, degree of respiratory support, and weight gain. The model had a 96% sensitivity and 53% specificity for detecting referral-warranted ROP, but no internal or external validation has been performed.

More recently, in 2016, the Colorado-Retinopathy of Prematurity (CO-ROP) model was proposed, using GA ≤ 30 weeks, GW ≤ 1500 g and weight gain ≤ 650 g in 4 weeks, with good results, but further validation is required in other populations.

In a multicenter study in Italy that compared WINROP, CHOP-ROP and ROPScore, both ROPScore and CHOP-ROP showed 100% sensitivity to detect sight-threatening ROP, whereas WINROP missed 19 cases of ROP, including three cases of type 1 ROP. The fact that several of the described models require longitudinal diary or weekly inputs or a significant number of variables restricts its usefulness in daily clinical practice, when compared with ROPScore. While ROPScore used other parameters in addition to BW and GW, these are easily accessible – use of oxygen therapy with mechanical ventilation, the need for blood transfusions, and BW by the second or sixth postpartum week – and only need to be registered once in a transversal way. This allows to access the risk before the first ophthalmological examination, facilitating the follow-up of infants who are at a greater risk.

Our study has some limitations. First, the introduction of the algorithm is still in a preliminary phase and it is not meant to substitute current screening guidelines. As further validation studies are published, ROPScore may be helpful to reduce the number of late or missed diagnosis. Second, the retrospective nature limits the power of the study. However, ROPScore has already been evaluated in prospective cohorts with good results. Third, the sample size is relatively small, and a larger sample would help to adjust the algorithm to our population, as well as narrow the 95% CI. Fourth, current Portuguese screening guidelines also include premature infants with BW < 2000 g with prolonged need for supplementary oxygen therapy, critically ill newborns or newborns who were submitted for major surgery. To date, no other study has evaluated the utility of using ROPScore in premature infants with BW of 1500–2000 g, which may justify the relatively inferiority of the ROPScore in our study. Nonetheless, it represents the true screened population in our center.

In conclusion, ROPScore is a simple noninvasive screening tool for the prediction of ROP, that can be applied as early as by the second week of life and only needs to be calculated once. Although the application of prediction algorithms is still in its early stages, ROPScore appears to a reliable model for ROP screening and for optimization of examination timings. This study adds to the validation of this model using real-world data from a Portuguese population. Prospective studies are needed to further validate the algorithm using data from different settings.

REFERENCES


27. Termote JUM, Donders ART, Schalij-Delfos NE, et al. Can Screening for Retinopathy of Prematurity Be


**CONTACT**

Ricardo Figueiredo  
Department of Ophthalmology  
Hospital do Espírito Santo de Évora  
Largo Senhor da Pobreza  
7000-811 – Évora  
E-mail: ricardoamfigueiredo@gmail.com

None of the authors has any financial/conflicting interests to disclose.