

Predictive algorithms for early detection of retinopathy of prematurity

Stefano Piermarocchi,¹ Silvia Bini,¹ Ferdinando Martini,¹ Marianna Berton,¹ Anna Lavini,² Elena Gusson,³ Giorgio Marchini,³ Ezio Maria Padovani,⁴ Sara Macor,⁵ Silvia Pignatto,⁵ Paolo Lanzetta,⁵ Luigi Cattarossi,⁶ Eugenio Baraldi⁷ and Paola Lago⁷

¹Department of Ophthalmology, University of Padova, Padova, Italy

²University of Verona, Verona, Italy

³Ophthalmology Unit, Department of Neurological, Neuropsychological, Morphological and Movement Sciences, University of Verona, Verona, Italy

⁴Department of Pediatrics, Section of Newborn Intensive Care Unit, University Hospital of Verona, Verona, Italy

⁵Department of Medical and Biological Sciences- Ophthalmology, University of Udine, Udine, Italy

⁶Department of Pediatrics- Neonatology, University of Udine, Udine, Italy

⁷Department of Woman's and Child's Health, University of Padova, Padova, Italy

ABSTRACT.

Purpose: To evaluate sensitivity, specificity and the safest cut-offs of three predictive algorithms (WINROP, ROPScore and CHOP ROP) for retinopathy of prematurity (ROP).

Methods: A retrospective study was conducted in three centres from 2012 to 2014; 445 preterms with gestational age (GA) \leq 30 weeks and/or birthweight (BW) \leq 1500 g, and additional unstable cases, were included. No-ROP, mild and type 1 ROP were categorized. The algorithms were analysed for infants with all parameters (GA, BW, weight gain, oxygen therapy, blood transfusion) needed for calculation (399 babies).

Results: Retinopathy of prematurity (ROP) was identified in both eyes in 116 patients (26.1%), and 44 (9.9%) had type 1 ROP. Gestational age and BW were significantly lower in ROP group compared with no-ROP subjects (GA: 26.7 ± 2.2 and 30.2 ± 1.9 , respectively, $p < 0.0001$; BW: 839.8 ± 287.0 and 1288.1 ± 321.5 g, respectively, $p = 0.0016$). Customized alarms of ROPScore and CHOP ROP correctly identified all infants having any ROP or type 1 ROP. WINROP missed 19 cases of ROP, including three type 1 ROP. ROPScore and CHOP ROP provided the best performances with an area under the receiver operating characteristic curve for the detection of severe ROP of 0.93 (95% CI, 0.90–0.96, and 95% CI, 0.89–0.96, respectively), and WINROP obtained 0.83 (95% CI, 0.77–0.87). Median time from alarm to treatment was 11.1, 5.1 and 9.1 weeks, for WINROP, ROPScore and CHOP ROP, respectively.

Conclusion: ROPScore and CHOP ROP showed 100% sensitivity to identify sight-threatening ROP. Predictive algorithms are a reliable tool for early identification of infants requiring referral to an ophthalmologist, for reorganizing resources and reducing stressful procedures to preterm babies.

Key words: algorithms – CHOP ROP – retinopathy of prematurity – ROPScore – WINROP

Acta Ophthalmol. 2017; 95: 158–164

© 2016 Acta Ophthalmologica Scandinavica Foundation. Published by John Wiley & Sons Ltd

doi: 10.1111/aos.13117

Introduction

Retinopathy of prematurity (ROP) continues to be one of the primary causes of treatable childhood blindness, especially in the developing countries (Gilbert et al. 2005). During the last decade, the improvement of neonatal care has increased the survival rate of extremely immature babies. Evidence-based screening guidelines have been published through the years and may vary with the characteristics of the population of premature newborns. These guidelines are continuously revised to improve the identification of infants at risk of developing a sight-threatening form of ROP. In many Western countries, all infants with a gestational age (GA) \leq 30 weeks and birthweight (BW) \leq 1500 g are systematically included in screening programs with a fundus examination (Fierston et al. 2013). In middle income countries, type 1 ROP also affects more mature infants, who are thus included in the guidelines. However, the number of countries with no specific guidelines is still significant (Hellström et al. 2013).

Although the incidence of all forms of ROP (type 1 and mild ROP) among very preterm babies (GA \leq 28 weeks) has been reported up to 70%, the proportion of newborns who will

develop a severe type 1 ROP and require treatment is a much smaller percentage of infants. Type 1 ROP incidence may be variable, ranging from 5% to 35% (Early treatment for Retinopathy of Prematurity Cooperative Group 2003; Allegaert et al. 2004; Darlow et al. 2005; Markestad et al. 2005; Weber et al. 2005; Tommiska et al. 2007; Austeng et al. 2009; Cerman et al. 2014; Painter et al. 2015). Thus, only a limited percentage of the screened infants will eventually need treatment for a severe ROP. A delay in the identification of a type 1 ROP may rapidly lead to extensive tractional retinal detachment, dramatically jeopardizing all strategies to preserve vision. Prompt diagnosis and timely treatment are of paramount importance for good clinical practice and successful outcomes. Current guidelines for ROP screening are still mostly based on GA and BW (Fierson et al. 2013), with a large number of infants undergoing fundus examination. Fortunately, most mild ROP cases spontaneously resolve, as physiological retinal vascularization is completed.

Programs aimed at identifying those infants who are likely to develop type 1 ROP have led to the formulation of algorithms for ROP prediction. While the inclusion criteria of the current screening guidelines put all preterm babies at a common risk level, the algorithms take into account other postnatal factors, such as postnatal weight gain, thus better tailoring risk calculation. These algorithms may assist both ophthalmologists and neonatologists in planning an efficient policy for screening, thereby improving the management of the complex issues related to the prevention of the dire consequences of untreated ROP.

In this study, we hypothesized that the algorithms for ROP risk prediction are more effective at promptly and correctly identifying the infants who are at high risk of developing sight-threatening ROP, compared to current protocols. In view of this hypothesis, we attempted to validate three published algorithms: WINROP (Löfqvist et al. 2006, 2009; Hellström et al. 2009), ROPScore (Eckert et al. 2012) and CHOP ROP (Binenbaum et al. 2011, 2012) in a population of preterm newborns enrolled at three different centres.

Raw sensitivity (SE) and specificity (SP), with standard cut-off values, of the three algorithms were analysed. To optimize the clinical performances of the algorithms, newly adjusted cut-off values were also calculated.

Patient population

Three neonatal intensive care units (NICUs) of three university hospitals (Padova, Verona and Udine) participated in the study providing data on infants screened from 2012 through 2014. All babies with GA \leq 30 weeks and/or BW \leq 1500 g were studied. Newborns with GA $>$ 30 weeks or BW between 1500 and 2000 g, when clinically unstable, were included in the study as well. Clinical instability has been judged by the neonatologist of each NICU and was generally due to sepsis or respiratory problems. This resulted in the inclusion of 12 more children out of range. The mean number per year of admissions with GA \leq 32 weeks from 2012 to 2014 is 120 in Padova (mean GA = 29 weeks), 90 in Verona (mean GA = 30.2) and 115 in Udine (mean GA = 29 weeks). The design and execution of the study was conducted in accordance with the tenets of the Declaration of Helsinki, and ethics committee approval was obtained.

Patients and Methods

All infants underwent serial fundoscopic examinations, with an indirect ophthalmoscope after pupil dilation, performed by experienced ophthalmologists, dedicated to ROP screening visits. These specialists from the three centres were also independently tested and agreed on the ophthalmoscopic classification of ROP. An interobserver agreement was also undertaken with *k* statistic ($k > 0.9$). The timing of visits was conducted according to the current screening guidelines, as approved by American Academy of Pediatrics and American Academy of Ophthalmology (Fierson et al. 2013). All included infants were submitted to the first ophthalmoscopic examination from 4 to 6 weeks after birth, and not before the 31st week of postmenstrual age (PMA) with the exception of extremely preterm infants (<25 weeks), who began screening earlier. Retinopathy of prematurity was graded for each infant and classified as '1' corre-

sponding to type 1 ROP, and '2' corresponding to type 2 ROP (as defined by the ETROP study) (Fierson et al. 2013) and all other mild forms of ROP. To restrict a stressful procedure to the most critical infants, fundus photography was obtained by means of a customized fundus camera (RetCam Clarity Medical Systems, Inc., Pleasanton, CA, USA) only when at least stage 2 was identified.

Clinical data collected by all centres included GA, BW, ethnicity, weekly weights after birth, weight at the sixth week of life, blood transfusions, use of oxygen in mechanical ventilation (nasal cannula, continuous positive airway pressure, high-flow nasal cannula or nasal intermittent positive pressure ventilation). The time from first ROP diagnosis to laser treatment was also recorded. Three algorithms, WINROP, ROPScore and CHOP ROP, were calculated for each infant in a masked fashion by another ophthalmologist (MB) who was blinded for the correspondent ROP clinical findings. Sensitivity, specificity and the most efficient cut-off values of the algorithms were then calculated.

The WINROP is an online surveillance system, originally based on weekly postnatal weights and insulin growth factor (IGF)-1 serum levels (Löfqvist et al. 2006, 2009). The algorithm is run by the weekly upload of the weights until an alarm is called (Hellström et al. 2009). The algorithm classifies the risk into two categories: absence or presence of the risk of developing any ROP. In our study, the risk was stratified into four levels divided as follows: (A) no risk and no alarm; (B) low risk (GA $>$ 29 weeks and/or BW $>$ 850 g) and alarm at PMA \geq 32 weeks; (C) low risk and alarm at PMA $<$ 32 weeks; and (D) high risk (GA $<$ 29 weeks and/or BW $<$ 850 g) (Hellström 2009). The algorithm allows risk calculation only for infants with GA \leq 32 weeks. ROPScore is an easily accessible algorithm requiring BW, GA, weight at 6th week of life, presence or absence (up to the 6th week of life) of blood transfusion and oxygen in mechanical ventilation (Eckert et al. 2012). The formula is in public domain and reported on a Microsoft Excel spreadsheet. The score is calculated only once per infant and not before the 6th week of life. The values obtained create a continuous

numeric scale (range between 9.1 and 21.6; Eckert et al. 2012). The published alarm cut-off points were 11 for any stage of ROP and 14.5 for type 1 ROP. The CHOP (Children's Hospital Of Philadelphia) ROP model was developed from the PINT (Premature Infants in Need of Transfusion) ROP (Binenbaum et al. 2011, 2012). The formula is accessible in public domain and works with postnatal weight gain. The score is calculated weekly, starting from the second week of life, to avoid the physiological loss of weight that usually happens during the first days of life. The score is expressed as a decimal number, being 0.010 the standard alarm cut-off value, over which the infant is considered at risk. The algorithm aims to identify all infants at risk of developing type 1 ROP.

Statistical analysis

The analysis was made by SAS® 9.3 software (SAS Institute, Cary, NC, USA). Tests were considered statistically significant with a p value of <0.05. A comparison of the infants' demographic characteristics in relation to the development of ROP was made by means of Student's t-test for independent samples. Sensitivity, specificity, positive predictive value (PPV) and negative predictive value of WINROP, ROPScore and CHOP ROP were calculated. The 95% confidence interval (CI) was calculated. The predictions of any ROP and type 1 ROP were calculated using the area of the receiver operating characteristic (ROC) curve and were compared by means of chi-square test. The best cut-offs of the three algorithms were also calculated. Whenever possible, a SE of 100% was chosen.

Results

Clinical observations

Overall, 445 preterm infants were screened in the three centres: 214 from Padova, 170 from Verona and 61 from Udine. Three hundred and ninety-nine infants had all data to obtain the algorithms and were thus eligible for the study. The main reason for exclusion was the transfer to other hospitals and the consequent loss of information, especially the weight gain. The demographic characteristics (particularly BW and GA) of the 46 excluded

infants were comparable to the subset used in the study. Their mean GA was 31.3 ± 2.3 weeks (range 26–37), and mean BW was 1318 ± 174.2 g (range 990–1745). Because of clinical instability, 12 children out of the range of screening guidelines were included (mean GA = 31.6 ± 0.6; mean BW = 1737.8 ± 194.7). As none of them developed any ROP, their influence for threshold calculation has been negligible. The most represented ethnicity was Caucasian (83.2%) followed by Black infants (13.0%) and Asiatics (3.8%). The infants from the NICU of Padova had the lowest mean GA and BW. Demographic characteristics are listed in Table 1. The mean PMA at first diagnosis of any ROP was 33.8 weeks (range 30.0–41.1). Retinopathy of prematurity was detected in both eyes in 116 of 399 patients (29.1%), and all infants had bilateral ROP. Type 1 ROP was present in 44 of 399 infants (11% of the entire population, 37.9% of those with any ROP). As expected, both GA and BW were significantly lower in ROP group compared with no-ROP babies (GA: 26.7 ± 2.2 and 30.2 ± 1.9 weeks, respectively ANOVA test, p < 0.0001; BW: 839.8 ± 287.0 and 1288.1 ± 321.5 g, respectively, ANOVA test adjusted for GA, p = 0.0016). The proportion of infants with ROP and type 1 ROP was inversely proportional

to GA. All babies born at the 23th week of GA developed ROP, and more than ¾ had type 1 ROP (Fig. 1). At the 28th week, less than half of the babies developed any ROP. A BW of <500 g was always associated with ROP and nine of 10 developed type 1 ROP (Fig. 2); 41.4% of newborns weighing at least 750 g developed any ROP.

When considering the mean weight gain at the 6th week, those without ROP had significantly higher values than those with ROP (percentage of weight gain compared to BW: 63.8 ± 19.2% versus 54.8 ± 21.7% respectively, ANOVA test adjusted for GA and BW, p = 0.0006). All type 1 ROP babies were treated with laser. Interestingly, almost all infants who developed type 1 ROP had received blood transfusion and oxygen via mechanical ventilation. No significant ethnic differences in rates of ROP were observed.

Algorithm results

Winrop

The algorithm calculation could be obtained for 377 infants (22 had a GA ≥ 32 weeks which prevented WINROP calculation). The algorithm correctly classified at risk (level B, C and D): 97 of 116 infants who developed any ROP and 41 of 44 who developed type 1 ROP. Nineteen

Table 1. Demographic characteristics of the three centres involved in the study.

	Centre			
	Padova	Verona	Udine	Total
Sample, n (%)	214 (48.1)	170 (38.2)	61 (13.7)	445 (100)
Gestational age, weeks				
Mean ± SD	28.7 ± 2.5*	30.0 ± 2.5	30.3 ± 2.3	29.4 ± 2.6
Minimum	23.0	23.4	24.0	23.0
Maximum	33.7	37.0	32.9	37.0
Birthweight, grams				
Mean ± SD	1094.7 ± 371.8*	1231.0 ± 329.4	1401.0 ± 382.6	1188.8 ± 372.2
Minimum	440	476	630	440
Maximum	2300	2000	2612	2612
Any ROP, n (%)				
No	131 (61.2)	146 (85.9)	52 (85.2)	
Yes	83 (38.8)	24 (14.1)	9 (14.8)	
Total	214 (100)	170 (100)	61 (100)	
Type 1 ROP, n (%)				
No	185 (86.4)	158 (95.1)	58 (95.1)	
Yes	29 (13.6)	12 (7.1)	3 (4.9)	
Total	214 (100)	170 (100)	61 (100)	

ANOVA = analysis of variance, ROP = retinopathy of prematurity.

* The comparison among mean values with ANOVA test (p < 0.0001), followed by Tukey's *post hoc* test for multiple comparison, showed mean gestational age and birth weight in Padova significantly inferior to Verona and Udine (p < 0.05), while no statistically significant difference between Verona and Udine was found.

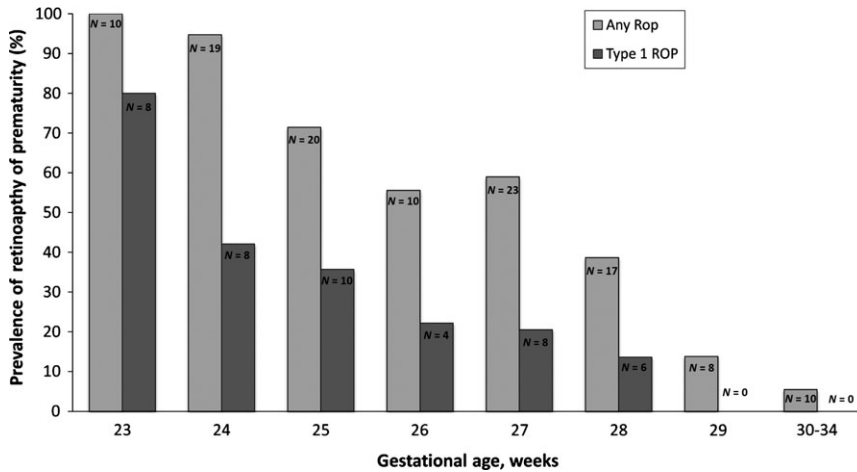


Fig. 1. Distribution of premature newborns with retinopathy of prematurity (ROP) according to gestational age (GA) at birth. Prevalence of any type of ROP and of type 1 ROP shows high correlation with GA.

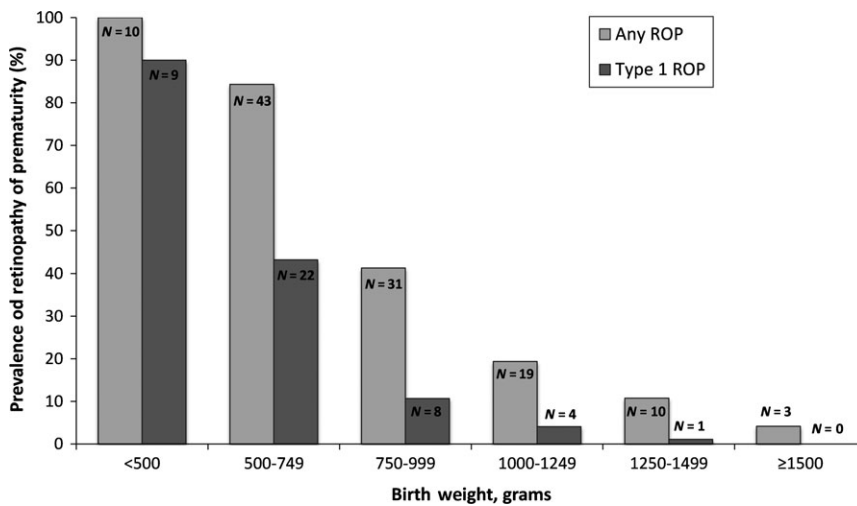


Fig. 2. Distribution of premature newborns with retinopathy of prematurity (ROP) according to weight at birth. Prevalence of any type of ROP and of type 1 ROP shows high correlation with weight at birth.

infants who were not considered at risk (no alarm) by WINROP developed ROP lesions, and three of them required laser treatment for type 1 ROP (Table 2). The median time from alarm to treatment was 11.1 weeks (range 6.9–19.9).

ROPscore

In our population, the algorithm values ranged from 6.6 to 20.5 (reference values between 9.1 and 21.6; Eckert et al. 2012). Setting the algorithm with the standard cut-offs (11 for any ROP and 14.5 for type 1 ROP), all babies with ROP and with type 1 ROP were identified. To maintain the highest SE (zero false negatives), but also to maximize SP, we adjusted the cut-off alarms and confirmed the value of 11, as ‘any ROP’

cut-off value, and introduced a customized value of 15.8, for type 1 ROP. This allowed false positives for type 1 ROP drop from 173 to 110. ROPscore correctly identified all 116 infants who developed any ROP and the 44 infants who developed type 1 ROP. Of the 202 infants of 399 (50.6%) who scored between 11 and 15.8, 16 developed ROP but none was type 1. All type 1 ROP infants scored over 15.8 (Table 2). The median time from alarm to treatment for type 1 ROP cases was 5.1 weeks (range 0.9–13.9).

Chop rop

This algorithm was originally designed to identify children with sight-threatening type 1 ROP (Binenbaum et al. 2012). In our sample, the values of the

algorithm ranged from 0 to 0.735. Although the algorithm was updated weekly, the first value for each infant was used for the statistical analysis. With a standard cut-off alarm value of 0.010, all the 44 babies who developed type 1 ROP could be identified. To increase SP, the alarm threshold was adjusted to 0.016 and all type 1 ROP were still identified, while false positives dropped from 141 to 115 (Table 2). The median time from alarm to treatment was 9.1 weeks (range 4.9–17.9).

Sensitivity and specificity

For WINROP, the category no alarm/no risk (A), compared with the presence of any risk (B, C and D), showed the best SE at identifying both infants with any ROP and infants with type 1 ROP; therefore, the stratification into four risk levels did not increase the SE of the algorithm in our population. At identifying any ROP, WINROP never reached the target SE of 100%, while ROPscore could identify all affected infants over the threshold 11. ROPscore provided the best performances for type 1 ROP for all parameters.

Receiver operating characteristic curve analysis for the detection of any ROP gave an area under the curve (AUC) of 0.89 for ROPscore, 0.88 for CHOP ROP that were significantly higher than WINROP (WINROP AUC = 0.81, chi-square test $p < 0.0001$) (Fig. 3A). For the detection of type 1 ROP, both ROPscore and CHOP ROP provided high AUC values (AUC = 0.93 for both algorithms), significantly higher than WINROP (AUC = 0.83, chi-square test $p < 0.0001$) (Fig. 3B).

Discussion

In our study, two algorithms (ROPscore and CHOP ROP) correctly identified all newborns that developed any type of ROP and/or more aggressive type 1 ROP long before the appearance of retinal lesions. All algorithms adopted GA, BW and weight gain to quantify risk, with ROPscore also considering oxygen administration and blood transfusions. From this study, we may conclude that ROPscore and CHOP ROP, which allow cut-off alarm customization, seem to be the most accurate for early prediction

Table 2. Prediction of any ROP with WINROP and ROPScore (A) and prediction of type 1 ROP (TIROP) with WINROP, ROPScore and CHOP ROP (B).

(A)					
Algorithm	WINROP		ROPScore		
CUT-OFF VALUE	ALARM		11		
Correctly predicted ROP/True ROP	97/116		116/116		
SENSITIVITY % (95% CI)	83.6 (75.8–89.7)		100 (96.8–100)		
Correctly predicted No-ROP/True No-ROP	144/261		43/283		
SPECIFICITY % (95% CI)	55.2 (49.1–61.1)		15.2 (11.5–19.8)		
Correctly predicted ROP/All predicted ROP	97/214		116/356		
POSITIVE PREDICTIVE VALUE % (95% CI)	45.3 (38.8–52.0)		32.6 (27.9–37.6)		
Correctly predicted No-ROP/All predicted No-ROP	144/163		43/43		
NEGATIVE PREDICTIVE VALUE % (95% CI)	88.3 (82.5–92.4)		100 (91.8–100)		

(B)					
Algorithm	WINROP	ROPScore		CHOP ROP	
CUT-OFF VALUE	ALARM	14.5*	15.8 [†]	0.010*	0.016 [†]
Correctly predicted TIROP/True TIROP	41/44	44/44	44/44	44/44	44/44
SENSITIVITY % (95% CI)	93.2 (81.8–97.7)	100 (92.0–100)	100 (92.0–100)	100 (92.0–100)	100 (92.0–100)
Correctly predicted No-TIROP/True No-TIROP	160/333	182/355	245/355	214/355	240/355
SPECIFICITY % (95% CI)	48.1 (42.7–53.4)	51.3 (46.1–56.4)	69.0 (64.0–73.6)	60.3 (55.1–65.2)	67.6 (62.6–72.3)
Correctly predicted TIROP/All predicted TIROP	41/214	44/217	44/154	44/185	44/159
PPV % (95% CI)	19.6 (14.5–2.05)	20.3 (15.5–26.1)	28.5 (22.0–36.2)	23.8 (18.2–30.4)	27.7 (21.3–35.1)
Correctly predicted No-TIROP/All predicted No-TIROP	160/163	182/182	245/245	214/214	240/240
NPV % (95% CI)	98.2 (94.7–99.4)	100 (97.9–100)	100 (94.5–100)	100 (98.2–100)	100 (98.4–100)

PPV = positive predictive value, NPV = negative predictive value, ROP = retinopathy of prematurity.

The performances (sensitivity, specificity, PPV and NPV) are shown for each algorithm.

* Cut-off value from literature data.

[†] Customized cut-off values, based on our population.

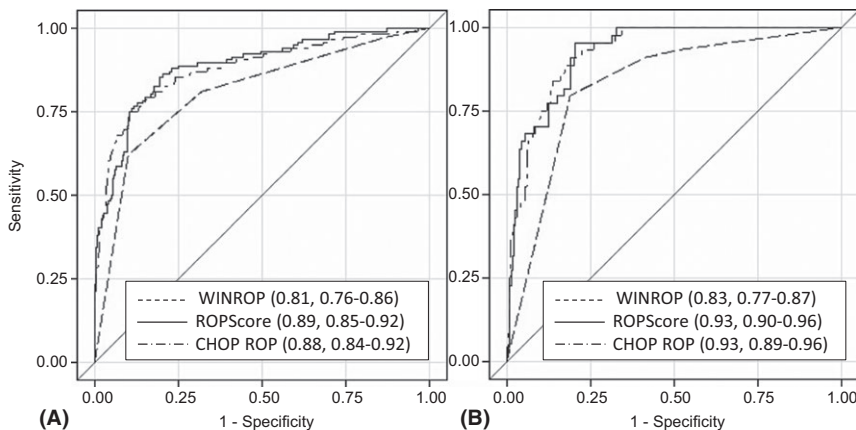


Fig. 3. Receiver operating characteristic (ROC) curves for detection of any type of retinopathy of prematurity (ROP) (A) and of type 1 ROP (B) according to the algorithms WINROP, ROPScore and CHOP ROP. Receiver operating characteristic curve analysis shows high AUC values both for detection of any type of ROP (A) and of type 1 ROP (B). For each algorithm, the value of AUC and the 95% CI are indicated in brackets. AUC, area under the curve; CI, confidence interval.

of a sight-threatening ROP. All tested algorithms led to the identification of an at-risk infant, with a level of accuracy from moderate to high, far in advance compared with the clinical detection of retinal lesions, with a

median time from alarm to diagnosis and treatment of severe ROP of 9.1 and 11.1 weeks, for CHOP ROP and WINROP, respectively. The need to collect data at the 6th week of life justifies a shorter time from alarm to

treatment for ROPScore (5.1 weeks) and may create a limitation in its use, for example, when an early and rapid form or ROP, that is aggressive posterior ROP, is present. Adjusted ROPScore showed the highest potential capability of safely reducing the need of a referral for an ophthalmological evaluation. For any type of ROP detection, 163 babies of 377 (43%) with WINROP and 43 of 399 (11%) for ROPScore would have skipped referral to the ophthalmologist. However, WINROP missed 19 cases (with three cases of type 1 ROP). Both ROPScore and CHOP ROP had zero false negative for type 1 ROP detection, thus hypothetically reducing ocular examination of 245 (61%) and 240 (60%) infants, respectively.

Recent epidemiologic studies demonstrate that ROP still represents an important cause of potentially treatable and possibly preventable childhood blindness worldwide (Hameed et al. 2004; Good & Gendron 2005; Lad et al. 2008, 2009; Gunn

et al. 2012; Tan et al. 2012; Hellström et al. 2013). Improvements of intensive care procedures in premature infants have led to an increase in survival of extremely preterm and low-weight newborns, which is still associated with a high incidence of this devastating disease, along with the appearance of new more aggressive forms of ROP (Miller et al. 2014).

Most scientific communities strongly recommend following strict screening guidelines to allow early diagnosis and prompt treatment, before the appearance of sight-threatening complications. According to standard screening criteria of GA and BW, a large number of preterm infants are screened when only a small number actually develop type 1 ROP requiring treatment. Other factors, such as weight gain after birth or systemic complications that have been recognized as potential triggers for progression to more aggressive forms, are normally not taken sufficiently into account. The physical stress of an invasive fundus examination to these fragile infants, the high cost of customized fundus cameras and the relative scarcity of experienced ophthalmologists capable of diagnosing and treating the disease are still today the major drawbacks of the current policies for GA- and BW-based ROP detection models.

New strategies in most countries are being adopted to offer more suitable solutions for a prompt diagnosis and a timely treatment. One is telemedicine, where centralized reading centres are created and collect retinal images from neonatal units, where a local efficient network of experienced eye consultants and costly equipment cannot be afforded (Fierson et al. 2015; Ying et al. 2015). These centres provide a prompt at-distance expert diagnosis and may also suggest surgical options in many clinical situations.

Predictive algorithms may help neonatologists to better target fundus examination at higher risk babies. In our series, all algorithms showed good performances at recognizing infants who developed type 1 ROP, with ROPScore and CHOP ROP showing 100% SE.

ROP pathogenesis is a multifactorial process: a first hypoxic preclinical phase is followed by a vasoproliferative phase. Insulin growth factor (IGF)-1, whose levels in preterm infants are

lower, plays a permissive role for vascular endothelial growth factor (VEGF) to act for normal retinal vascularization. As IGF-1 levels begin to rise, the abnormally accumulated VEGF elicits a massive vasoproliferation within the hypoxic retina. Low postnatal IGF-1 levels are well correlated with low weight gain during the early weeks after birth. Thus, simple measurements of weight gain reflect the serum IGF-1 values of each infant (Hellström et al. 2009). Although GA and BW represent the most important factors driving the need for referral to expert ophthalmologists, all algorithms also consider weight gain.

Other conditions might influence the course of ROP. Among these, supplemental oxygen and blood transfusions are both included in the ROPScore. Anaemia, sepsis, apnoea, necrotizing enterocolitis and intraventricular haemorrhage have also been associated with severe ROP, but they do not appear in any algorithms. It is currently held that these factors act by lowering the IGF-1 and are implicitly measured with weight gain (Hellström et al. 2009). Some factors related to the variable racial and GA distribution of infants among centres as well as differences in healthcare systems may partially account for the need of customizing cut-off alarms. Although this disease still represents an important cause of childhood blindness in Western countries, the small number of ROP expert ophthalmologists is a growing challenge around the world. For those centres that have to manage a high number of preterm babies and cannot rely on a constant ophthalmological support, combined use of predictive algorithms and low-cost high-quality fundus cameras, might represent a suitable solution for successful ROP screening and proper management. The ideal scenario would be that each neonatal unit includes a retinal camera as part of its essential equipment (Gilbert et al. 2015); after identification through algorithms of at-risk babies, dedicated personnel at the neonatal unit (i.e. neonatologist, nurse or trained technician) collects the images and makes the decision on the correct management. This would mean that, in those units without constant ophthalmological support, the ophthalmologist would intervene if an infant develops what can be

defined as a 'referral warranted ROP'. This technology has the potential to greatly increase the coverage of programs and to share with the neonatal team the responsibility for screening (Vasalaki et al. 2015).

In clinical practice, the risk stratification operated by the algorithms would greatly reduce the number of infants undergoing screening visits and the workload for ophthalmologists. In this hypothesis, in our population, the reduction would have meant 61% fewer infants undergoing fundus examination for ROPScore, 60% for CHOP ROP and 43% for WINROP, although WINROP incorrectly classified not at risk three cases who developed type 1 ROP.

Some limits to the use of the algorithms might be evident in low-income countries. In these areas, the performances of the algorithms are usually worse than in high-income countries (Hård et al. 2010; Zepeda-Romero et al. 2012; Choi et al. 2013; Sun et al. 2013; Ko et al. 2015). Moreover, as WINROP does not calculate the risk for infants with GA > 32 weeks, its use is limited, especially in developing nations, where older preterm infants are at risk of developing even sight-threatening ROP (Hård et al. 2010; Zepeda-Romero et al. 2012). Validation studies of WINROP are still ongoing. For example, a recent study has shown that a reassessment of WINROP can significantly improve the specificity of the algorithm (Lundgren et al. 2015).

Although the introduction of the algorithms is still in a preliminary phase and it cannot yet substitute current screening guidelines, it may significantly help to reduce the number of missed or late diagnoses. One limitation of our study data is the sample size. Most authors recommend larger sample size studies to build more precise algorithms and narrower 95% CIs. As ROP is a potentially blinding disease, we cannot miss even one case of type 1 ROP (Binenbaum 2013). Another point is the limited ethnic variability of our population. Some authors recently demonstrated that there are ethnic differences in IGF-1 serum levels; specifically, infants born from black mother have significantly lower levels of IGF-1 at 32 and 33 weeks than other infants (Reddy et al. 2015). These data reinforce the

idea that the prediction of ROP risk is a race-specific phenomenon. In this hypothesis, we recommend that any ROP-team who starts using ROPScore or CHOP ROP would first validate those algorithms in its own population, to achieve the best and safest performances.

In conclusion, the predictive algorithms represent a promising and suitable tool to recognize preterm babies requiring referral to specialized ophthalmologist for prompt treatment. One major point is that infants who will develop a sight-threatening ROP can be identified through algorithms long before the appearance of retinal lesions. Even though the ultimate safety and the benefits of this strategy must still be thoroughly validated, it may offer new opportunities for therapies aiming at preventing this devastating disease.

References

Allegaert K, de Coen K, Devlieger H & EpiBel Study Group (2004): Threshold retinopathy at threshold of viability: the EpiBel Study. *Br J Ophthalmol* **88**: 239–242.

Austeng D, Källen KB, Ewald UW, Jacobsson PG & Hölmstrom GE (2009): Incidence of retinopathy of prematurity in infants born before 27 weeks gestation in Sweden. *Arch Ophthalmol* **127**: 1315–1319.

Binenbaum G (2013): Algorithms for the prediction of retinopathy of prematurity based on postnatal weight gain. *Clin Perinatol* **40**: 261–270.

Binenbaum G, Ying GS, Quinn GE, Dreiseitl S, Karp K, Roberts RS, Kirpalani H & Premature Infants in Need of Transfusion Study Group (2011): A clinical prediction model to stratify retinopathy of prematurity risk using postnatal weight gain. *Pediatrics* **127**: e607–e614.

Binenbaum G, Ying GS, Quinn GE, Huang J, Dreiseitl S, Antigua J, Foroughi N & Abbasi S (2012): The CHOP postnatal weight gain, birth weight, and gestational age retinopathy of prematurity risk model. *Arch Ophthalmol* **130**: 1560–1565.

Cerman E, Balci SY, Yenice OS, Kazokoglu H, Celiker H & Eraslan M (2014): Screening for retinopathy of prematurity in a tertiary ophthalmology department in Turkey: incidence, outcomes, and risk factors. *Ophthalmic Surg Lasers Imaging Retina* **45**: 550–555.

Choi JH, Löfqvist C, Hellström A & Heo H (2013): Efficacy of the screening algorithm WINROP in a Korean population of preterm infants. *JAMA Ophthalmol* **131**: 62–66.

Darlow BA, Hutchinson JL, Henderson-Smart DJ, Donoghue DA, Simpson JM, Evans NJ & Australian and New Zealand Neonatal Network (2005): Prenatal risk factors for severe retinopathy of prematurity among very preterm infants of the Australian and New Zealand Neonatal Network. *Pediatrics* **115**: 990–996.

Early treatment for Retinopathy of Prematurity Cooperative Group (2003): Revised indications for the treatment of retinopathy of prematurity: results of the Early Treatment for Retinopathy of Prematurity randomized trial. *Arch Ophthalmol* **121**: 1684–1694.

Eckert GU, Fortes Filho JB, Maia M & Procianny RS (2012): A predictive score for retinopathy of

prematurity in very low birth weight preterm infants. *Eye* **26**: 400–406.

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 (2013): Screening examination of premature infants for retinopathy of prematurity. *Pediatrics* **131**: 189–195.

Fierson WM, Capone A Jr & American Academy of Pediatrics Section on Ophthalmology, American Academy of Ophthalmology, American Association of Certified Orthoptists (2015): Telemedicine for evaluation of retinopathy of prematurity. *Pediatrics* **135**: e238–e254.

Gilbert C, Fielder A, Gordillo L, Quinn G, Semiglia R, Visintin P & Zin A; International NO-ROP Group (2005): Characteristics of infants with severe retinopathy of prematurity in countries with low, moderate, and high levels of development: implications for screening programs. *Pediatrics* **115**: e518–e525.

Gilbert C, Wormald R, Fielder A et al. (2015): Potential for a paradigm change in the detection of retinopathy of prematurity requiring treatment. *Arch Dis Child Fetal Neonatal Ed* **101**: F6–9.

Good WV & Gendron RL (2005): Retinopathy of prematurity: gone today, here tomorrow? *Clin Experiment Ophthalmol* **33**: 339–340.

Gunn DJ, Cartwright DW & Gole GA (2012): Incidence of retinopathy of prematurity in extremely premature infants over an 18-year period. *Clin Experiment Ophthalmol* **40**: 93–99.

Hameed B, Shyamanur K, Kotecha S, Manktelow BN, Woodruff G, Draper ES & Field D (2004): Trends in the incidence of severe retinopathy of prematurity in a geographically defined population over a 10-year period. *Pediatrics* **113**: 1653–1657.

Hård AL, Löfqvist C, Fortes Filho JB, Procianny RS, Smith L & Hellström A (2010): Predicting proliferative retinopathy in a Brazilian population of preterm infants with the screening algorithm WINROP. *Arch Ophthalmol* **128**: 1432–1436.

Hellström A, Hård AL, Engström E, Niklasson A, Andersson E, Smith L & Löfqvist C (2009): Early weight gain predicts retinopathy in preterm infants: new, simple, efficient approach to screening. *Pediatrics* **123**: e638–e645.

Hellström A, Smith LE & Dammann O (2013): Retinopathy of prematurity. *Lancet* **382**: 1445–1457.

Ko CH, Kuo HK, Chen CC, Chen FS, Chen YH, Huang HC, Fang PC & Chung MY (2015): Using WINROP as an adjuvant screening tool for retinopathy of prematurity in Southern Taiwan. *Am J Perinatol* **30**: 149–154.

Lad EM, Nguyen TC, Morton JM & Moshfeghi DM (2008): Retinopathy of prematurity in the United States. *Br J Ophthalmol* **92**: 320–325.

Lad EM, Hernandez-Boussard T, Morton JM & Moshfeghi DM (2009): Incidence of retinopathy of prematurity in the United States: 1997 through 2005. *Am J Ophthalmol* **148**: 451–458.

Löfqvist C, Andersson E, Sigurdsson J, Engström E, Hård AL, Niklasson A, Smith LE & Hellström A (2006): Longitudinal postnatal weight and insulin-like growth factor I measurements in the prediction of retinopathy of prematurity. *Arch Ophthalmol* **124**: 1711–1718.

Löfqvist C, Hansen-Pupp I, Andersson E, Holm K, Smith LE, Ley D & Hellström A (2009): Validation of a new retinopathy of prematurity screening method monitoring longitudinal postnatal weight and insulin-like growth factor I. *Arch Ophthalmol* **127**: 622–627.

Lundgren P, Stoltz Sjöström E, Domellöf M, Smith L, Wu C, VanderVeen D, Hellström A & Löfqvist C (2015): The specificity of the WINROP algorithm can be significantly increased by reassessment of the WINROP alarm. *Neonatology* **108**: 152–156.

Markestad T, Kaarsen PI, Rønnestad A et al. & Norwegian Extreme Prematurity Study Group (2005): Early death, morbidity, and need of

treatment among extremely premature infants. *Pediatrics* **115**: 1289–1298.

Miller MM, Revenis ME, Lai MM et al. (2014): Risk and clinical course of retinopathy of prematurity in 78 infants of gestational age 22–25 weeks. *J AAPOS* **18**: 266–270.

Painter SL, Wilkinson AR, Desai P, Goldacre MJ & Patel CK (2015): Incidence and treatment of retinopathy of prematurity in England between 1990 and 2011: database study. *Br J Ophthalmol* **99**: 807–811.

Reddy MA, Patel HI, Karim SM, Lock H, Perry L, Bunce C, Kempley S & Sinha AK (2015): Reduced utility of serum IGF-1 levels in predicting retinopathy of prematurity reflects maternal ethnicity. *Br J Ophthalmol* **100**: 501–504.

Sun H, Kang W, Cheng X et al. (2013): The use of the WINROP screening algorithm for the prediction of retinopathy of prematurity in a Chinese population. *Neonatology* **104**: 127–132.

Tan SZ, Dhaliwal C, Becher JC & Fleck B (2012): Trends in the incidence of retinopathy of prematurity in Lothian, south-east Scotland, from 1990 to 2009. *Arch Dis Child Fetal Neonatal Ed* **97**: F310–F311.

Tommiska V, Heinonen K, Lehtonen L, Renlund M, Saarela T, Tammela O, Virtanen M & Fellman V (2007): No improvement in outcome of nationwide extremely low birth weight infant populations between 1996–1997 and 1999–2000. *Pediatrics* **119**: 29–36.

Vasalaki M, Patel H, Crame M, Solinap Y & Reddy MA (2015): Improvement in retinopathy of prematurity screening standards after the development of a multidisciplinary dedicated team. *Acta Ophthalmol* **93**: e88–e89.

Weber C, Weninger M, Klebermass K et al. (2005): Mortality and morbidity in extremely preterm infants (22 to 26 weeks of gestation): Austria 1999–2001. *Wien Klin Wochenschr* **117**: 740–746.

Ying GS, Quinn GE, Wade KC, Repka MX, Baumritter A, Daniel E & e-ROP Cooperative Group (2015): Predictors for the development of referral-warranted retinopathy of prematurity in the telemedicine approaches to evaluating acute-phase retinopathy of prematurity (e-ROP) study. *JAMA Ophthalmol* **133**: 304–311.

Zepeda-Romero LC, Hård AL, Gomez-Ruiz LM et al. (2012): Prediction of retinopathy of prematurity using the screening algorithm WINROP in a Mexican population of preterm infants. *Arch Ophthalmol* **130**: 720–723.

Received on October 20th, 2015.
Accepted on April 6th, 2016.

Correspondence:
Silvia Bini, MD
Via Giustiniani 2
35128 Padova
Italy
Tel: 0039 049 8212110
Fax: 0039 049 8755168
Email: silvia.bini@studenti.unipd.it

The authors wish to thank Dr. Fabiano Cavarzeran (Department of Neuroscience, University of Padova, Padova Italy) for his critical support to the statistical analysis of the results of the study. The authors also thank Dr. Gaia Martini and Dr. Caterina Mansoldo (Ophthalmology Unit, Department of Neurological, Neuropsychological, Morphological and Movement Sciences, University of Verona, Verona, Italy) and Dr. Paolo Biban (Department of Neonatal and Paediatric Intensive Care, University of Verona, Verona, Italy) for the contribution to the critical analysis of the data.