The Specificity of the WINROP Algorithm Can Be Significantly Increased by Reassessment of the WINROP Alarm

Pia Lundgren, Elisabeth Stoltz Sjöström, Magnus Domellöf, Lois Smith, Carolyn Wu, Deborah VanderVeen, Ann Hellström, Chatarina Löfqvist

Institute of Neuroscience and Physiology, Sahlgrenska Academy, University of Gothenburg, Gothenburg, and Department of Food and Nutrition, and Unit for Paediatrics, Department of Clinical Sciences, Umeå University, Umeå, Sweden; Department of Ophthalmology, Boston Children’s Hospital, Harvard Medical School, Boston, Mass., USA

Background: Retinopathy of prematurity (ROP) is a sight-threatening disease affecting extremely preterm infants. The introduction of new ROP screening surveillance systems, with higher sensitivity and specificity than established ROP screening guidelines, has the potential to reduce the number of stressful eye examinations in these infants.

Objectives: To improve the specificity of the WINROP (Weight, Insulin-like growth factor-I, Neonatal, ROP) surveillance system, identifying extremely preterm infants requiring treatment for ROP.

Methods: Two cohorts that had previously been subjected to WINROP analyses were included and reevaluated in this study. The weight at WINROP alarm for extremely preterm infants, born at gestational age <27 weeks, was reevaluated and by establishing ‘safe’ WINROP alarm weight limits, an intersample reassessment of WINROP alarm was performed. The two cohorts were as follows: (1) the Extremely Preterm Infants in Sweden Study (EXPRESS) cohort, infants born in Sweden during 2004–2007 (n = 407), and (2) extremely preterm infants in a North American cohort, born during 2006–2009 (n = 566).

Results: In the EXPRESS cohort, 12.5% (40/319) of the infants who previously received a WINROP alarm were now reassessed as having no alarm; the specificity of WINROP in EXPRESS increased from 23.9% (86/360) to 35.0% (126/360). In the North American cohort, 15.4% (81/526) were reassessed as having no alarm; the specificity increased from 8.5% (38/447) to 26.6% (119/447). The sensitivity persisted as 97.5% in EXPRESS (45/47) and 98.3% (117/119) in the North American cohort.

Conclusions: The specificity of the WINROP surveillance system for extremely preterm infants can be significantly improved by reassessment using the weight at WINROP alarm.

Key Words
Retinopathy of prematurity · Extremely preterm · WINROP system

Abstract
Background: Retinopathy of prematurity (ROP) is a sight-threatening disease affecting extremely preterm infants. The introduction of new ROP screening surveillance systems, with higher sensitivity and specificity than established ROP screening guidelines, has the potential to reduce the number of stressful eye examinations in these infants. Objectives: To improve the specificity of the WINROP (Weight, Insulin-like growth factor-I, Neonatal, ROP) surveillance system, identifying extremely preterm infants requiring treatment for ROP. Methods: Two cohorts that had previously been subjected to WINROP analyses were included and reevaluated in this study. The weight at WINROP alarm for extremely preterm infants, born at gestational age <27 weeks, was reevaluated and by establishing ‘safe’ WINROP alarm weight limits, an intersample reassessment of WINROP alarm was performed. The two cohorts were as follows: (1) the Extremely Preterm Infants in Sweden Study (EXPRESS) cohort, infants born in Sweden during 2004–2007 (n = 407), and (2) extremely preterm infants in a North American cohort, born during 2006–2009 (n = 566). Results: In the EXPRESS cohort, 12.5% (40/319) of the infants who previously received a WINROP alarm were now reassessed as having no alarm; the specificity of WINROP in EXPRESS increased from 23.9% (86/360) to 35.0% (126/360). In the North American cohort, 15.4% (81/526) were reassessed as having no alarm; the specificity increased from 8.5% (38/447) to 26.6% (119/447). The sensitivity persisted as 97.5% in EXPRESS (45/47) and 98.3% (117/119) in the North American cohort. Conclusions: The specificity of the WINROP surveillance system for extremely preterm infants can be significantly improved by reassessment using the weight at WINROP alarm.

Introduction
Retinopathy of prematurity (ROP) is a sight-threatening disease that primarily affects infants born very or extremely preterm [1]. To prevent blindness from retinal detachment, timely identification and treatment of those infants who are at risk of developing severe ROP are crucial. The infants’ eyes are screened with repeated eye examinations using indirect ophthalmoscopy and/or RetCam digital imaging. Tropicamide and phenylephrine
eye drops are instilled topically to dilate the pupil. A lid speculum and/or a depressor are often used to better identify peripheral retinal alterations. Eye examinations must be repeated until the retina is fully vascularized. Most ROP screening guidelines are based on low gestational age (GA) and low birth weight (BW) – well-known risk factors for potentially sight-threatening ROP that requires treatment [2, 3].

Only about 20% of extremely preterm infants, born before gestational week 27, screened for ROP require treatment [4]. The extremely preterm infants may be subject to large numbers of eye examinations, as recently reported in an extremely preterm cohort in Sweden; a median of 10 examinations were performed per infant (range 1–30) [5]. The ROP eye examination is a painful procedure [6]. Thus, to spare these fragile extremely preterm infants unnecessary ROP screening examinations it is crucial to find additional surveillance systems that identify infants at high risk of developing severe ROP requiring treatment.

Poor postnatal weight gain and low levels of insulin-like growth factor-I (IGF-I) have recently been established as strong predictors of severe ROP [7, 8]. WINROP (Weight, IGF-I, Neonatal, ROP) is an online surveillance system that identifies infants at risk of developing ROP requiring treatment through the use of weekly weight measurements. WINROP was developed in Gothenburg, Sweden. The first study enrolled Swedish preterm infants and presented that low postnatal serum levels of IGF-I and poor postnatal weight gain could predict whether an infant is at high risk of developing severe ROP [9]. After further development of the WINROP algorithm, weekly postnatal weight measurements alone without IGF-I measurement were found to be equally reliable in predicting this risk [10, 11], with the benefit of no blood draw.

WINROP has been available online since 2009 and use of the system is free of charge. The WINROP algorithm uses a weight reference model based on BW and weekly weight measurements from infants developing no ROP or mild ROP [9]. The WINROP algorithm estimates the differences between expected safe weekly weight gain and the infant’s observed actual weight gain. The values are calculated and accumulated. When the accumulated sum exceeds a limit, an alarm is signaled to indicate that the infant is at risk for severe ROP. The WINROP outcome is either no alarm or alarm. The alarm limits were constructed with the criteria of very high sensitivity in order to not miss any of the infants developing ROP requiring treatment. WINROP analyses cannot be used in infants with excessive edema or hydrocephalus because their weight gain does not reflect physiological weight gain.

Thus, by entering the infant’s BW and GA along with weekly weight measurements into WINROP, the infant’s risk of developing treatment-requiring ROP can be calculated. When using WINROP, weekly weight measurements must be performed until the infant is at postmenstrual age 35–36 weeks, even if WINROP does not signal with an alarm. Prior to that time, WINROP cannot evaluate whether the infant is completely free of risk.

WINROP has been evaluated in different preterm populations, both in Sweden and worldwide [12–17]. Adjustments to the WINROP algorithm have been suggested for specific populations in order to improve its sensitivity and specificity [18]. In one Swedish WINROP study including extremely preterm infants, the specificity of WINROP was found to be lower than in previous studies [17]. Many of the extremely preterm infants had poor postnatal weight development and thus many received a WINROP alarm. With the intention of improving the specificity of WINROP for extremely preterm infants, we retrospectively reviewed infants’ weight values at the time of the WINROP alarm.

**Methods**

In this study, the following two cohorts previously assessed by WINROP were utilized: (1) the population-based Extremely Preterm Infants in Sweden Study (EXPRESS) cohort comprising 407 infants born at GA <27 weeks in Sweden during 2004–2007 [17] and (2) a North American cohort comprising 1,706 infants born at...
In these previous studies the infants were conventionally screened for ROP, according to present screening guidelines. Retrospectively, WINROP analyses were performed by a person unaware of the ROP outcome. In this study only infants born from GA 23 weeks and 0 days to 26 weeks and 6 days in the North American cohort (n = 566) were enrolled to age-match the EXPRESS cohort.

Using the infants’ weight at the time of the WINROP alarm, we were able to establish a ‘safe’ alarm weight. The safe alarm weight was set above the maximum weight for the infants who developed ROP that would require treatment. Infants born at GA 26 weeks had a lower safe weight at the time of the alarm compared to infants born at GA 23–25 weeks, resulting in two subgroups of safe alarm weight limits.

Using the new constructed alarm weight limits, WINROP analyses were performed again, i.e. an intersample reassessment of the WINROP alarm in the same cohorts (the EXPRESS and North American cohorts) was performed. Infants with alarm weights below the safe weight limit remained at risk and infants with alarm weights above the limit were considered to have no alarm (fig. 1).

### Results

In the EXPRESS cohort, the median GA was 25 weeks and 4 days (23 weeks and 0 days to 26 weeks and 6 days) and the median BW was 784 g (348–1,315). In EXPRESS, 78.4% (319/407) of the infants received a WINROP alarm; of these, 12.5% (40/319) were now reassessed as having no alarm. This caused the specificity of WINROP in EXPRESS to increase from 23.9% (86/360) to 35.0% (126/360). In the North American cohort, the median GA

---

**Table 1.** WNROP sensitivity, specificity and positive and negative predictive values in identifying type 1 ROP in EXPRESS and North American cohorts before and after reassessment of WINROP alarm

<table>
<thead>
<tr>
<th>Alarm status</th>
<th>Original EXPRESS WINROP study</th>
<th>EXPRESS WINROP after reassessment of alarm</th>
<th>Original North American WINROP study</th>
<th>North American WINROP after reassessment of alarm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ROP categories</td>
<td>ROP categories</td>
<td>ROP categories</td>
<td>ROP categories</td>
</tr>
<tr>
<td></td>
<td>Type 1 ROP, n</td>
<td>ROP type 1, n</td>
<td>ROP type 1, n</td>
<td>ROP type 1, n</td>
</tr>
<tr>
<td></td>
<td>Non-type 1 ROP, n</td>
<td>Non-type 1 ROP, n</td>
<td>Non-type 1 ROP, n</td>
<td>Non-type 1 ROP, n</td>
</tr>
<tr>
<td></td>
<td>Total, n</td>
<td>Total, n</td>
<td>Total, n</td>
<td>Total, n</td>
</tr>
<tr>
<td></td>
<td>Predictive value, %</td>
<td>Predictive value, %</td>
<td>Predictive value, %</td>
<td>Predictive value, %</td>
</tr>
<tr>
<td></td>
<td>PPV</td>
<td>PPV</td>
<td>PPV</td>
<td>PPV</td>
</tr>
<tr>
<td></td>
<td>NPV</td>
<td>NPV</td>
<td>NPV</td>
<td>NPV</td>
</tr>
<tr>
<td>Sensitivity, %</td>
<td>95.7 (84.3–99.2)</td>
<td>95.7 (84.3–99.2)</td>
<td>98.3 (93.4–99.7)</td>
<td>98.3 (93.4–99.7)</td>
</tr>
<tr>
<td>Specificity, %</td>
<td>23.9 (19.6–28.7)</td>
<td>35.0 (30.1–40.2)</td>
<td>8.5 (6.1–11.5)</td>
<td>26.6 (22.6–31.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PPV 14.1 (10.6–18.5)</td>
<td>PPV 16.1 (12.1–73.0)</td>
<td>PPV 22.2 (18.8–26.0)</td>
<td>PPV 26.3 (22.3–30.7)</td>
</tr>
<tr>
<td></td>
<td>NPV 97.7 (91.2–99.6)</td>
<td>NPV 98.4 (93.9–99.7)</td>
<td>NPV 95.0 (81.8–99.1)</td>
<td>NPV 98.3 (93.4–99.7)</td>
</tr>
</tbody>
</table>

Values in parentheses are 95% CI. NPV = Negative predictive value; PPV = positive predictive value.
was 25 weeks and 2 days (23 weeks and 26 weeks and 6 days) and the median BW was 723 g (380–1,210).

In the North American cohort, 92.9% (526/566) of the infants received a WINROP alarm; of these, 15.4% (81/526) were reassessed as having no alarm. The specificity increased from 8.5% (38/447) to 26.6% (119/447). The sensitivity was not affected in either cohort – 97.5% in EXPRESS (45/47) and 98.3% (117/119) in the North American cohort (table 1). The majority of the infants born extremely preterm who could be reassessed as having no WINROP alarm were born at GA 26 weeks and had an early alarm (fig. 1).

Discussion

By using the weight at WINROP alarm, the specificity of the WINROP algorithm, identifying infants at high risk of developing ROP requiring treatment, can be improved for infants born extremely preterm.

WINROP has a high sensitivity but relatively low specificity for extremely preterm infants (GA <27 weeks). Poor postnatal weight gain is a common problem for extremely preterm infants [17], consequently leading to many infants receiving a WINROP alarm. However, by analyzing the maximum weight at the time of the WINROP alarm for the group of infants requiring ROP treatment, we constructed safe alarm weight limits. The infants with a weight above the safe alarm weight limit could be safely reassessed as having no WINROP alarm. Thus, the specificity of the WINROP surveillance system could be significantly improved with persistent high sensitivity.

The majority of the infants who could be reassessed as having no WINROP alarm were born at GA 26 weeks and had an early alarm. These more mature infants appeared to have more variable postnatal weight gain and therefore their alterations in alarm weights were easier to detect. By analyzing the WINROP alarm weights of extremely preterm infants, we obtained an additional parameter for calculating the postnatal growth of the infants and predicting their risk of severe ROP that requires treatment.

Approximately 20% of extremely preterm infants screened for ROP require treatment. Since the extremely preterm infants are so fragile, it is especially important to reduce their number of stressful ROP eye examinations [6, 19].

Even though improvements in the specificity of ROP prediction are valuable, ways to prevent ROP are needed [20]. Early intervention studies using nutritional and hormonal approaches are ongoing [21].

We propose that other predictive models, using poor postnatal weight gain, could be improved by analyzing in particular the weight gain of extremely preterm infants. Consequently, we invite other groups to validate our findings. Before introducing the new WINROP alarm weight limits in the web-based WINROP surveillance system, further validation is warranted.

Acknowledgments

This study was supported by the following: the Swedish Medical Research Council (grants 2008-2842, 2011-2432), Government grants (ALFGB-137491), VINNOVA (grants 2009-01152, 2009-00221), the Anna-Lisa and Bror Björnsson Foundation, the Wilhelm and Martina Lundgren Science Foundation, the Herman Svensson Foundation, the Children’s Hospital Ophthalmology Foundation, the NIH/NEI (EY022275, EY017017, P01 HD18655), the Lowy Medical Research Institute, and European Commission FP7 Project 305485 (PREVENT-ROP).

References


