Retinopathy of prematurity screening criteria and workload implications at Tygerberg Children’s Hospital, South Africa: A cross-sectional study

E Visser Kift, MB ChB, BSc (Med Sci) Hons; N Freeman, MB ChB, FCOphth, MMed; C Cook, MB ChB, DO, MPH, FCOphth, FRCOphth; L Myer, MB ChB, PhD

1 Division of Clinical Pharmacology, Department of Medicine, Tygerberg Hospital and Faculty of Medicine and Health Sciences, Stellenbosch University, Tygerberg, Cape Town; Division of Epidemiology and Biostatistics, School of Public Health and Family Medicine, Faculty of Health Sciences, University of Cape Town; Previously affiliated to the Division of Ophthalmology, Department of Surgery, Tygerberg Hospital and Stellenbosch University, Tygerberg, Cape Town, South Africa
2 Division of Ophthalmology, Department of Surgery, Tygerberg Hospital and Faculty of Medicine and Health Sciences, Stellenbosch University, Tygerberg, Cape Town, South Africa
3 Division of Ophthalmology, Department of Surgery, Groote Schuur Hospital and Faculty of Health Sciences, University of Cape Town, South Africa
4 Division of Epidemiology and Biostatistics, School of Public Health and Family Medicine, Faculty of Health Sciences, University of Cape Town, South Africa

Corresponding author: E Visser Kift (eviskift@sun.ac.za)

Background. Screening guidelines for retinopathy of prematurity (ROP) used in high-income countries are not appropriate for middle-income countries, and screening requirements may vary even between units within one city.

Objective. To determine optimal ROP screening criteria, and its workload implications, for Tygerberg Children’s Hospital (TCH), Cape Town, South Africa.

Methods. This cross-sectional study included premature infants screened for ROP at TCH from 1 January 2009 to 31 December 2014. Logistic regression analysis for prediction and classification was performed. Predictors were birth weight (BW) and gestational age (GA). Endpoints were clinically significant ROP (CSROP) and type 1 ROP (T1ROP).

Results. Of 1 104 eligible infants, 33.4% had ROP (CSROP 9.1%, T1ROP 2.5%). All T1ROP infants received laser therapy. The number of screening examinations was inversely correlated with GA and BW. The number needed to screen to identify one infant requiring treatment was 41 (entailing 83 examinations, 4 screening hours, one technician and three doctors). Screening infants with a GA of ≤28 weeks or a BW of <1 000 g would have detected all infants with T1ROP but missed two outliers with CSROP. These outliers would only have been detected with a GA of ≥32 weeks or a BW <1 500 g.

Conclusions. Detection of infants with T1ROP is resource intensive. Larger infants require screening to include a few outliers, but they require fewer examinations than smaller infants. Making local screening criteria narrower on the basis of a limited evidence base may be dangerous. Risk factors for CSROP in larger infants need to be researched.


Retinopathy of prematurity (ROP) has an estimated global incidence of 20 000 infants per year. South Africa (SA) has joined the ‘third epidemic’ of ROP, which is attributed to both increasing survival of premature infants and poor oxygen management in neonatal intensive care units (NICUs). Oxygen management inadequacies in middle-income country NICUs lead to larger infants being at risk than in high-income countries.

In SA, an estimated 16 000 infants are at risk of ROP and require screening each year. In 1995, ROP accounted for 10.6% of cases of blindness in schools for the blind. The prevalence of ROP among infants with birth weights (BWs) <1 500 g in tertiary hospitals in SA remains low (any ROP 16.3 - 24.5%, clinically significant ROP (CSROP) 1.56 - 4.4% and sight-threatening ROP 0.6 - 2.9%).

The first SA ROP guideline was published in 2013. It recommends screening of infants with a gestational age (GA) of <32 weeks or a BW of <1 500 g, as well as larger infants (GA 32 - 35 weeks, BW 1 500 - 2 000 g) with an increased risk of ROP.

The development of screening guidelines that are appropriate for all institutions in a middle-income country is challenging. ROP screening requirements are dependent on NICU quality of care, and this varies widely between units. Screening infants of higher weight may not be feasible in units where resources are limited. Most secondary-level hospital NICUs in SA do not even have access to screening facilities. In such NICUs, larger infants are likely to be at risk of ROP as a result of the level of care. In contrast, screening larger infants in SA tertiary institutions with high-quality neonatal care may represent inappropriate use of scarce resources. Ideally, ROP screening criteria should be based on local evidence.

Tygerberg Children’s Hospital (TCH), Western Cape Province, SA, reported a ROP prevalence of 31.1% (CSROP 7.1%) among preterm infants ventilated for respiratory distress (1986 - 1987). Preterm infants treated exclusively with non-invasive ventilation in the first week of life (2009 - 2010) had a ROP prevalence of 21.8% (CSROP 4.4%). Only 1.5% of the latter cohort required treatment, and no infant with a BW >1 250 g had CSROP. The results from these subgroup analyses are instructive, but insufficient to inform local screening policy and resource allocation.
Objective
A focused analysis of the existing TCH ROP screening database was undertaken to determine optimal screening criteria for ROP at TCH and to explore the workload implications of using different BW and GA cut-off values.

Methods
This was a cross-sectional study, based on a database analysis, of infants born in the drainage area of TCH. Attending TCH neonatologists identified premature infants requiring ROP screening, and registered them in the ROP database. ROP screening activities at TCH commenced before the release of national screening criteria and were governed by local data and resources at the time. Infants were examined and staged according to the International Classification of ROP (2005 revision)\(^1\)\(^5\) by a trained ophthalmologist using a 28-dioptre condensing lens and an indirect ophthalmoscope. The screening team consisted of one technician and three doctors (the supervising paediatric ophthalmologist and two registrars in training). Retcam III imaging was performed in selected cases. Examination findings were recorded in hospital records and the database. Infants were examined from 31 weeks corrected GA or 4 weeks after birth, whichever was later. Follow-up screens were performed 1 - 3-weekly until vascularisation of zone III was completed or the corrected GA of 41 weeks was reached. Infants with treatable stages of ROP (T1ROP) received laser therapy within 72 hours of diagnosis.

The sample size was fixed to the number of infants with information recorded in the database. Screening examinations with missing outcome data were excluded from the study. The recorded findings (zone, stage and presence of plus disease for each eye) were converted to binary outcome variables: any ROP, CSROP and T1ROP. CSROP was defined as ROP involving zone I, any stage 3 ROP or plus disease associated with any stage ROP in any zone. T1ROP was defined as any stage zone I ROP with plus disease, zone I stage 3 with or without plus disease, or zone II stage 2 or 3 with plus disease.

---

**Table 1. Numbers and proportions of infants developing ROP according to GA and BW group (N=621)**

<table>
<thead>
<tr>
<th>GA category (weeks)</th>
<th>Total, N</th>
<th>Any ROP, n (%)</th>
<th>CSROP, n (%)</th>
<th>T1ROP, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>24 - 25</td>
<td>15</td>
<td>8 (53.3)</td>
<td>4 (26.6)</td>
<td>2 (13.3)</td>
</tr>
<tr>
<td>26 - 27</td>
<td>221</td>
<td>77 (34.5)</td>
<td>28 (12.6)</td>
<td>7 (3.2)</td>
</tr>
<tr>
<td>28 - 29</td>
<td>292</td>
<td>88 (30.1)</td>
<td>20 (6.8)</td>
<td>5 (1.7)</td>
</tr>
<tr>
<td>30 - 32</td>
<td>79</td>
<td>14 (17.7)</td>
<td>3 (3.79)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>33 - 37</td>
<td>14</td>
<td>4 (28.6)</td>
<td>1 (7.1)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>Total</td>
<td>621</td>
<td>191 (30.8)</td>
<td>56 (9.0)</td>
<td>14 (2.3)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BW category (g)</th>
<th>Total, N</th>
<th>Any ROP, n (%)</th>
<th>CSROP, n (%)</th>
<th>T1ROP, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;600</td>
<td>10</td>
<td>3 (30.0)</td>
<td>2 (20.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>600 - 799</td>
<td>112</td>
<td>49 (43.8)</td>
<td>15 (13.4)</td>
<td>6 (5.4)</td>
</tr>
<tr>
<td>800 - 999</td>
<td>284</td>
<td>93 (32.7)</td>
<td>27 (9.5)</td>
<td>6 (2.1)</td>
</tr>
<tr>
<td>1 000 - 1 199</td>
<td>171</td>
<td>37 (21.6)</td>
<td>9 (5.3)</td>
<td>2 (1.2)</td>
</tr>
<tr>
<td>1 200 - 1 399</td>
<td>34</td>
<td>8 (23.5)</td>
<td>2 (5.9)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>≥1 400</td>
<td>10</td>
<td>1 (10.0)</td>
<td>1 (10.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>Total</td>
<td>621</td>
<td>191 (30.8)</td>
<td>56 (9.0)</td>
<td>14 (2.3)</td>
</tr>
</tbody>
</table>

---

Fig. 1. Overview of study population and outcomes.
Data were entered in a custom-designed Microsoft Excel spreadsheet, and analysed using Stata 13 (StataCorp, USA). Ethical approval was obtained from the ethics committees of the University of Cape Town (Ref. No. 924/2014) and Stellenbosch University and Tygerberg Hospital (both Ref. No. S14/10/218).

**Results**

Between 1 March 2009 and 28 February 2014, 1 104 infants were examined for ROP and the examination findings were recorded in the ROP database at TCH. Of the 2 643 ROP screening examinations (bookings) registered in the database, 406 examination findings (15.4%) were missing. Infants not arriving for scheduled examinations (7.1%) was the main reason for missing data (Fig. 1). Another explanation was inability to perform the examinations (2.8%) owing to poorly dilated pupils or clinical instability.

The median GA at birth was 28 weeks (interquartile range (IQR) 27 - 29, range 24 - 37); 95% of infants had a GA of <32 weeks. The median BW was 930 g (IQR 820 - 1 040, range 523 - 2 600). Median GAs and BWs were similar across the 5 years of screening. The median postnatal age (PNA) of the infants at the first screening examination was 5 weeks (IQR 4 - 7) and the postmenstrual age (PMA) was 33 weeks (IQR 32 - 35).

**Retinopathy of prematurity**

ROP (any ROP) was found in 33.4% of infants, CSROP in 9.1% and T1ROP in 2.5% (Fig. 1). The prevalence of ROP was highest in the subgroup of infants with a gestational age of <27 weeks (any ROP 44.9% (44/98) and CSROP 18.4% (18/98)). Overall, the median GA was 28 weeks in infants with any ROP and 27 weeks in those with CSROP. The median BW was 865 g in infants with any ROP and 826 g in those with CSROP. Rates of ROP according to GA and BW categories (among the subset of 621 infants with both GA and BW recorded in the database) are shown in Table 1. CSROP was detected in two infants with a GA >32 weeks or a BW >1 500 g. One of these infants had a BW of 866 g and a GA of 31 weeks. Pre-plus ROP was detected in zone II of the left eye at 34 weeks PMA (3 weeks PNA).

**Screening cut-points**

Among the 621 infants for whom both GA and BW were recorded, GA was significantly associated with CSROP (odds ratio (OR) 0.710, 95% confidence interval (CI) 0.579 - 0.872; \( p < 0.001 \)). The association between BW and CSROP was weaker (OR 0.998, 95% CI 0.996 - 0.9995; \( p = 0.013 \)). Both GA and BW were predictors of CSROP and T1ROP (Fig. 2). GA was not a significantly better predictor of CSROP than BW (\( p = 0.521 \)), and using both GA and BW in the prediction model was no better than using either GA (\( p = 0.412 \)) or BW (\( p = 0.181 \)). A GA cut-point of 30 weeks had a sensitivity of 97% and specificity of 9% in detecting CSROP. With this classification rule, two older infants (infant A: GA 35 weeks and BW 866 g; infant B: GA 31 weeks and BW 1 530 g) with CSROP would have been

![Fig. 2. Receiver operating characteristic curves showing how prematurity predicts ROP.](image)

| Table 2. Workload implications and false negatives with application of different screening criteria (N=621) |
|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| screening criteria              | A: All infants                  | B: GA ≤32 weeks or BW <1 500 g | C: GA ≤30 weeks or BW <1 200 g |
| **Infants, N/n (%)**            | 621                             | 618 (99.5)                      | 610 (98.2)                      |
| **Examinations, N/n (%)**       | 1 385                           | 1 382 (99.8)                    | 1 367 (98.7)                    |
| **Examinations per infant, N/n (%)** | 2.23                 | 2.24 (100.5)                    | 2.24 (100.5)                    |
| **Infants with CSROP, N/n (%)** | 56                              | 56 (100.0)                      | 55 (98.2)                       |
| **Infants needed to screen to detect 1 infant with CSROP, N/n (%)** | 11                              | 11 (100.0)                      | 11 (100.0)                      |
| **Examinations required to detect 1 infant with CSROP, N/n (%)** | 25                              | 25 (100.0)                      | 25 (100.0)                      |
| **Infants with T1ROP**          | 14                              | 14 (100.0)                      | 14 (100.0)                      |
| **Infants needed to screen to detect one infant needing treatment, N/n (%)** | 44                              | 44 (100.0)                      | 44 (100.0)                      |
| **Examinations required to detect one infant needing treatment, N/n (%)** | 99                              | 99 (100.0)                      | 97 (98.0)                       |
| **CSROP missed (false negative), n** | -                               | 0                               | 1                               |
| **T1ROP missed (false negative), n** | -                               | 0                               | 0                               |
missed in the given dataset, but neither required treatment.

A less sensitive cut-point of ≥28 weeks would have missed nine infants, one of whom would have required treatment. At a cut-point of ≥1 200 g, the BW classification rule had a sensitivity of 96% and a specificity of 6%. With this classification rule, three infants with a BW ≥1 200 g and CSROP would have been missed, but none of them required treatment (infant B: GA 31 weeks and BW 1 530 g; infant C: GA 28 weeks and BW 1 200 g; infant D: GA 28 weeks and BW 1 220 g). A less sensitive cut-point of ≥1 000 g would have missed 12 infants (two with T1ROP). The last two rows of Table 2 summarise infants missed (false negatives) when BW and GA cut-offs are combined and applied. Screening criteria D (GA ≥28 weeks or BW <1 000 g) would have detected all infants with CSROP requiring treatment (T1ROP). However, the most mature infant in the dataset with CSROP that did not require intervention had a GA of 31 weeks and BW of 1 530 g and would have been detected with screening criteria B (GA ≤32 weeks or BW <1 500 g) only.

**Infants needing treatment**

Among the 1 104 infants with recorded screening examination outcomes, all with T1ROP received laser therapy. The 27 infants treated for ROP were ≥29 weeks GA or ≤1 060 g (median GA 27 weeks (IQR 26 - 28, range 24 - 29) and median BW 815 g (IQR 763 - 940, range 640 - 1 060)). The median PMA of the infants at the time of T1ROP detection was 35 weeks (IQR 33 - 38, range 31 - 43) and the PNA was 7 weeks (IQR 6 - 10, range 5 - 15). Fig. 3 reflects how infants needing treatment (T1ROP) changed over time, compared with those with CSROP and any ROP. Eleven of 27 infants (40.7%) were treated during 2012. Any ROP and CSROP gradually increased over the study period from 13.9% to 48.3% and from 4.2% to 13.2%, respectively. T1ROP had a baseline of 1.8% with a peak in 2012 (11/234 (4.7%) and returned to baseline in 2013.

**Workload implications**

The 1 104 infants required 2 237 screening examinations (Fig. 1). The median number of examinations per infant was two (IQR 1 - 3, range 1 - 11), with 97.3% requiring fewer than six examinations. The median number of examinations for infants not needing treatment was two (IQR 1 - 3, range 1 - 10), as opposed to six (IQR 4 - 9, range 1 - 11) for treated infants. Thirteen infants (13/27, 48.1%) had reached the criterion for treatment at the first examination. However, the complete management of those treated involved four or more examinations in the majority of cases (20/27, 74.1%) (Fig. 1). The number needed to screen to identify one infant needing treatment was 4.1. Screening 41 infants entailed 83 examinations (4 screening hours) by the screening team of one technician and three doctors (a paediatric ophthalmologist and two registrars in training).

The number of examinations required was inversely correlated with GA and BW (Spearman correlation coefficient −0.20 and −0.27; p<0.001). Table 2 shows the workload implications of applying different screening criteria (options B – D) to the subset of infants with both GA and BW recorded (option A). The narrower the screening criteria, the smaller the number of infants to be screened but also the higher the number of screening examinations per infant. The workload, expressed as the total number of examinations, therefore does not decrease in proportion to the number of infants screened. For example, the screening criteria in Table 2, option D, are GA ≤28 weeks or BW <1 000 g. Narrowing the screening criteria to option D would lead to a 14% reduction (6 fewer infants) in the number to be screened to detect one infant needing treatment, but only an 11% (not 14%) reduction (11 fewer examinations) in the number of examinations required to detect one needing treatment. Such a reduction in screening criteria would have decreased the workload by 91 fewer infants to be screened (149 fewer examinations) at the expense of missing 2 infants with CSROP.

**Discussion**

The prevalence of CSROP among infants with GAs <27 weeks at birth was 18.4% (18/98). Comparison with the ROP burden elsewhere is difficult because of the substantial variability in study designs, gestational ages of infants, survival rates and treatments used. Estimates of the prevalence of severe ROP from population-based studies range from 5% to 35% in high-income countries with equivalent NICUs.[18] The terms low-, middle- and high-income country do not always reflect on the quality of care in a particular country.[19] In countries such as SA, India and Brazil, the standard of care may vary widely between individual NICUs. Zin et al.[20] showed that in Rio de Janeiro the better NICUs were able to use narrower guidelines than other units in the same city.[21] The prevalence of T1ROP in this middle-income city ranged from 2.1% to 7.8% among NICUs.[22] The TCH T1ROP prevalence of 2.5% compares with the estimate at the lower end of the range, which was measured at the Rio de Janeiro federal government research institute.

As expected, the smaller and younger the infants, the higher the rates of ROP (Table 1). A significant number of infants >32 weeks and ≥1 200 g were examined. This group included only infants at high risk for ROP as identified and referred for screening by neonatologists. This explains the higher than expected yield of ROP in this category. However, few of them had CSROP and none required treatment. One infant with CSROP had a BW of 1 530 g and another a GA of 35 weeks. However, when BW and GA are taken into account simultaneously, as in the SA national ROP screening guideline (GA ≤32 weeks or BW <1 500 g),[23] no infants with CSROP requiring follow-up fell outside the screening criteria. Neither of the two more mature infants mentioned reached the criteria for laser treatment. This indicates that neonatal care at TCH is well established and that narrower guidelines may be applicable for this specific unit.
When planning ROP screening services, it must be kept in mind that it is the immature and sick infants that require the most work.[16] In this study the majority of more mature infants needed only one or two examinations. Including more mature infants in the screening programme does not increase the workload to the same extent as screening very immature infants.

Screening criteria D (GA ≤28 weeks or BW <1 000 g) would have detected all infants needing treatment. Screening criteria D plus referral of all infants at increased risk of ROP may have included the two more mature infants with outlying measurements. However, reducing GA and BW for inclusion in screening should be undertaken with caution owing to a relatively small evidence base.[16] Further investigation towards specific risk factors, or reasons for referral, is required in this cohort of infants. In addition, research is required to determine whether using this lower GA and BW plus a ‘sickness’ criterion (risk factors) would be an effective and efficient method of including the outliers. For example, prediction methods based on GA, BW and postnatal weight gain (WINROP, ROPScore and CHOP ROP) are employed to reduce ROP screening workload in high-income settings. Poor generalisability limits their use.[16]

Another tool has been developed by Binenbaum,[17] which can also be used in more mature infants at risk for ROP in countries with developing neonatal care systems (personal communication, Prof. C Gilbert, 10 February 2016).

Accurate GA estimates are not routinely available in SA. GA in our study population was based on the date of the last normal menstruation, early ultrasound or the new Ballard score calculation at birth. BW was as good as GA in predicting CSROP, and using GA in addition to BW was not better at predicting CSROP. However, as was illustrated above with infant B in the dataset, using GA when it is available acts as a safety net to include those with high BW for GA (outliers).

Zin et al.[13] showed that NICUs with a <80% survival rate of infants with a BW of <1 500 g should screen infants with a GA of ≤35 weeks or a BW of ≤500 g, and that NICUs with a survival rate of ≥80% could adopt narrower screening criteria of GA <32 weeks or BW ≤1 500 g.[13] The TCH survival rate of infants with a BW of <1 500 g gradually increased from 2009 to 2014 and was 89.7% on average during the study period (personal communication, Prof. J Smith, 29 July 2015). We also showed that screening infants with GA <32 weeks or BW ≤1 500 g, then that NICUs with a <1 500 g infant survival rate of >80%. In third-epidemic settings, the infant mortality rate is therefore a useful proxy to indicate which infants should be examined. These findings need to be validated.

The Sweden ROP register (SWEDROP) showed that the incidence of severe ROP did not change substantially over time.[19] Increased survival balanced against improved neonatal intensive care can perhaps explain this. In our study, the numbers of infants seen increased over time and the screening programme became more efficient in detecting any ROP and CSROP over time. The prevalence of T1ROP was 2% each year across the study period, with the exception of a peak of 5% in 2012. An increase of very immature infants at high risk for T1ROP plus the sudden increase in numbers of very low-birth-weight admissions (745 in 2012 and 801 in 2013, as opposed to 626 in 2011 and 632 in 2010) can explain the peak. An improvement in overall neonatal care may explain the return to the norm in 2013.

Study limitations and research opportunities

Missing data limit the data analysis and interpretation in this study. The exact ROP screening programme coverage is not known, but our impression is that it is insufficient. Not all infants weighing <1 500 g were screened. Infants weighing ≥2 00 g were screened only if they had additional risk factors for ROP. A significant proportion (15.2%, 405/2 670) of the examinations had missing outcomes and were excluded from analyses. Either GA or BW was not recorded for 43.8% (483/1 104) of infants, but thorough exploratory analyses revealed no clear pattern in missingness of data. Improved routine data collection is needed for audit purposes and to establish the ROP programme coverage.

Further studies are required to establish whether in resource-scarce environments more mature infants should be screened routinely, or whether screening more mature infants should rather be based on risk scores. Short- and long-term visual outcomes of infants are not recorded in the database. Study of these outcomes would guide scarce resource allocation. Digital imaging of the retina is employed in TCH ROP screening ward rounds, but its role in ROP screening workload reduction has yet to be established locally.

Conclusions

Unnecessary ROP screening examination should be avoided because it is resource intensive, requires specially trained ophthalmologists and is stressful for infants. Neonatal management is continually changing, so screening criteria should be revised accordingly. Local audit is an essential component that informs this revision. Local guidelines should ideally be based on prospective studies.

At TCH no infants needing treatment were larger than 1 060 g or older than 29 weeks, and no infant requiring follow-up for CSROP fell outside the current national screening criteria for GA (<32 weeks or BW <1 500 g).[12] However, making the local screening criteria narrower on the basis of a limited evidence base may be dangerous. More mature infants require fewer screening examinations than immature infants. A modest broadening of current local inclusion criteria and screening infants of ≥32 weeks at increased risk of ROP should not result in an unmanageable workload. More research is required to develop screening criteria for the identification of more mature infants at risk of ROP in resource-limited settings.

Consent.

All authors consent to publication.

Acknowledgements.

Prof. J Smith, Dr L Visser, Prof. C Gilbert.

References


Accepted 9 March 2016.