Retrospective Validation of the Postnatal Growth and Retinopathy of Prematurity (G-ROP) Criteria in a Japanese Cohort

AKIHIKO SHIRAKI, YOKO FUKUSHIMA, RYO KAWASAKI, HIROKAZU SAKAGUCHI, MIWA MITSUHASHI, HIROMI INEYAMA, YOSHIKAZU HATSUKAWA, AND KOHJI NISHIDA

• PURPOSE: We aimed to externally validate the performance of new screening criteria for retinopathy of prematurity (ROP) developed in the Postnatal Growth and Retinopathy of Prematurity (G-ROP) study among a Japanese cohort.

• DESIGN: Validation of screening criteria.

• METHODS: We reviewed premature infants screened for ROP between September 2009 and May 2017 at a single institution. The G-ROP criteria, except hydrocephalus, were applied as a prediction model for infants with both a known outcome of ROP and serial measurements of weight gain. We assessed sensitivity and specificity for treatment-requiring ROP, and reduction in the number of infants who receive ROP screening and in the number of retinal examinations.

• RESULTS: Of 692 premature infants screened for ROP, 537 had information of ROP outcome and weight gain. In this cohort, 81 infants required treatment for ROP; in 218 infants, ROP regressed spontaneously; and 238 infants did not develop any ROP. The G-ROP model reached a sensitivity of 100% (95% confidence interval [CI], 95.4%-100%) and specificity of 28.9% (95% CI, 24.9%-33.2%). No infants required any treatment for ROP before the date of risk determination. The number of infants requiring screening and the number of examinations would have been reduced by 24.5% and 12.9%, respectively.

• CONCLUSIONS: This is the first validation study of the G-ROP criteria in a developed country other than North America. The criteria demonstrated high sensitivity in this Japanese cohort, even though the criterion of hydrocephalus was excluded. (Am J Ophthalmol 2019;205:50–53. © 2019 Elsevier Inc. All rights reserved.)

Accepted for publication Mar 27, 2019.

From the Department of Ophthalmology, Osaka University Graduate School of Medicine, Osaka, Japan (A.S., Y.F., R.K., H.S., M.M., K.N.); and the Department of Ophthalmology, Osaka Women's and Children's Hospital, Osaka, Japan (H.I., Y.H.).

Inquiries to Yoko Fukushima, Department of Ophthalmology, E7, Osaka University Graduate School of Medicine, 2-2 Yamadaoka, Suita, Osaka 565-0871, Japan; e-mail: youko.fukushima@ophthal.med.osaka-u.ac.jp

THE INCIDENCE OF RETINOPATHY OF PREMATURITY (ROP) varies worldwide and depends on multiple factors, such as socioeconomic circumstances and race/ethnicity.[1,2] However, low gestational age (GA) and low birthweight (BW) are common risk factors for severe ROP globally.[3] Because severe ROP can result in childhood blindness, screening criteria are set with a safety margin for cutoff values of GA and BW, to avoid missing infants who will develop ROP.[3–6] Therefore, in most screened infants, repeated ophthalmic examinations would have been unnecessary. Recently, new screening criteria for ROP were proposed in the Postnatal Growth and Retinopathy of Prematurity (G-ROP) study.[7] Although the criteria are a promising model to improve screening efficiency while maintaining high sensitivity in North America, no validation studies have been reported. We applied the new criteria for Japanese premature infants retrospectively, to evaluate whether the model can be generalized across countries and racial/ethnic groups.

METHODS

• STUDY DESIGN: This was a retrospective cohort study of preterm infants who underwent eye examinations and had a known ROP outcome in the neonatal intensive care unit of a single institution in Japan from September 2009 to May 2017. The study was approved by the Ethical Review Board of Osaka Women's and Children's Hospital and was carried out in compliance with the principles of the Declaration of Helsinki (1964). Eligible participants were premature infants with GA <33 weeks, with BW ≤1800 g, or with being at high risk of ROP determined by a neonatologist. We excluded infants without a known ROP outcome, those without sequential measurements of daily weight, and those with any ocular diseases including retinovascular conditions besides ROP, such as familial exudative vitreoretinopathy or incontinentia pigment (Bloch-Sulzberger). For ROP screening, an initial fundus examination was performed at 29 or 30 weeks postmenstrual age or at 3 weeks chronological age, whichever was later. The diagnosis of ROP and indication of treatment for ROP followed the International Classification of ROP Revisited[8] and the Early Treatment for ROP Study,[9] respectively. The term "type 1

0002-9394/$36.00
https://doi.org/10.1016/j.ajo.2019.03.027

TABLE 1. Characteristics of the 537 Premature Infants in the Study

	TR-ROP	Non TR-ROP	
		Regressed ROP	No ROP
N	81	218	238
Gestational age, wk			
Mean (SD)	25.2 (1.9)	27.5 (2.6)	32.2 (3.1)
Median (range)	24.8 (22.2-34.1)	27.4 (22.2-35.8)	31.5 (25.0-41.5)
Birthweight, g			
Mean (SD)	691 (250)	873 (283)	1507 (590)
Median (range)	626 (394-2002)	847 (432-2070)	1404 (422-4062)
Male, n (%)	32 (39.5)	117 (53.6)	117 (49.1)

ROP = retinopathy of prematurity; TR-ROP = treatment-requiring retinopathy of prematurity.

TABLE 2. Prediction of Retinopathy of Prematurity Using 5 Criteria Developed in the G–ROP Study

	TR-ROP (N = 81)	Non TR-ROP (N = 456)	Any ROP (N = 299)	No ROP (N = 238)
Alarm-positive	81	324	275	130
Sensitivity, % (95% CI)	100 (95.4-100)	NA	91.9 (88.3-94.5)	NA
Alarm-negative	0	132	24	108
Specificity, % (95% CI)	NA	28.9 (24.9-33.2)	NA	45.3 (39.1-51.7)

NA = not applicable; ROP = retinopathy of prematurity; TR-ROP = treatment-requiring retinopathy of prematurity.

ROP" was defined by the criteria after the Early Treatment for ROP study. Demographic and ophthalmologic data, including sex, BW, GA, all available weight measurements, dates of retinal examinations, ROP stage and zone at every examination, and any treatments performed for ROP, were collected from the medical records.

• **OUTCOMES AND STATISTICAL ANALYSIS:** The original G-ROP criteria consist of 6 alarm levels, including 5 quantitative thresholds and 1 qualitative description. To increase ease of use for ophthalmologists, in this study, we first used the following 5 criteria of quantitative thresholds, except hydrocephalus, as a prediction model: GA <28 weeks; BW <1051 g; weight gain (WG) between postnatal day 10 and 19 <120 g; WG between postnatal day 20 and 29 <180 g; or WG between postnatal day 30 and 39 <170 g. The WGs were calculated by using the absolute values of weight at the indicated postnatal days. These criteria were applied for eligible infants to determine whether each infant had a greater risk than the 5 alarm levels. The performance of the criteria was tested by calculating sensitivity and specificity for treatment-requiring ROP (TR-ROP) and any ROP. The 95% confidence intervals (CIs) for measures of sensitivity and specificity were calculated using the Wilson score method. To evaluate the effects of ROP management, we calculated the reduction under the current screening setting in the number of infants requiring screening for ROP and in the number of retinal examinations. For infants who did not meet any of the criteria, all of those infants and all of their examinations were considered toward the reduction in examinations. For infants who met any 1 or more of the criteria, any examination performed prior to the date that they met 1 of the criteria were counted toward the reduction in examinations, but all further examinations were not counted toward the reduction in examinations. Those infants were not considered toward the reduction in number of infants requiring any examinations. JMP for Mac software, version 13.0 (SAS Institute Inc, Cary, North Carolina, USA) was used to perform all statistical analyses.

RESULTS

OF 692 INFANTS WHO RECEIVED RETINAL EXAMINATIONS, 537 were eligible in this study. Table 1 summarizes the characteristics of all infants. Any ROP developed in 299 infants (55.6%), of whom 81 (15%) required treatment, and 218 (40.5%) regressed spontaneously. The median GA was 29.1 weeks (range, 22.2-41.5 weeks) and the median BW was 986 g (range, 394-4062 g) in this cohort.

According to 5 criteria from the G-ROP, we considered that an alarm was raised when an infant met at least 1 criterion (Table 2). The model predicted 81 of 81 infants with TR-ROP (sensitivity, 100%; 95% CI, 95.4%-100%) and 275 of 299 infants with any ROP (sensitivity, 91.9%; 95% CI, 88.3%-94.5%). Specificity for prediction of infants with non TR-ROP and infants with no ROP was 28.9% (95% CI, 24.9%-33.2%) and 45.3% (95% CI, 39.1%-51.7%), respectively. The number of infants requiring screening for ROP would have decreased from 537 to 405 infants (24.5%). The number of examinations among all 537 infants would have decreased from 4089 to 3561, a reduction of 528 examinations (12.9%). The model correctly identified in advance all infants requiring any treatment for ROP; however, in 131 of 132 infants who did not meet these criteria, the model had not yet determined their risk status when the initial retinal examination was scheduled, under the current screening schedules.

The components of the criteria were evaluated according to the distribution of BW and GA in infants with TR-ROP (Figure). In this cohort, 78 of 81 infants with TR-ROP already met more than 1 criterion at birth and the remaining 3 infants were accurately categorized as high risk according to WG. Of 299 infants with any ROP, 257 infants were categorized as high risk according to GA and BW. Although 3 criteria of WG triggered an alarm for 18 infants with any ROP, these WG criteria missed 24 infants (8.0%) with any ROP. All 24 infants had low-grade ROP of less than type 2 ROP, of whom only 2 infants were diagnosed with hydrocephalus and 15 infants (62%) were male.

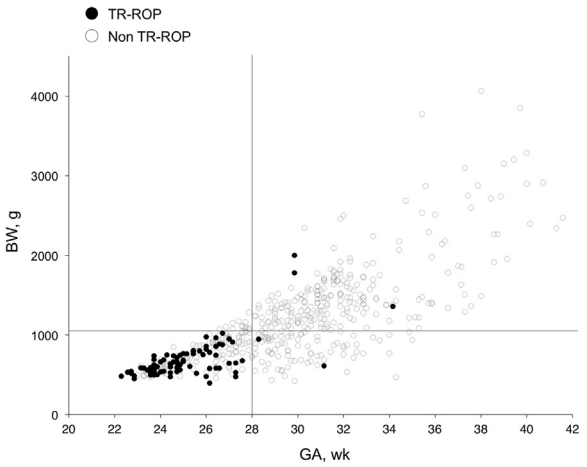

FIGURE. Distribution of infant subsets based on gestational age (GA) and birthweight (BW) thresholds of the G-ROP study criteria. The lines define the thresholds for BW and GA. TR-ROP = treatment-requiring retinopathy of prematurity.

DISCUSSION

WE VALIDATED THE G-ROP CRITERIA IN A JAPANESE cohort. Whereas there were only 233 (3.1%) Asian infants among the 7483 infants in the G-ROP study,[7] the present cohort of 537 Japanese/East Asian infants showed reproducible high performance of the criteria. The model estimated a reduction in the number of infants requiring screening of about 25%. However, infants had already been examined when their risk status was eventually determined at postnatal day 39 according to the G-ROP criteria. Therefore, we should consider the timing of the initial examination in clinical application.

There were more male than female infants with low-grade ROP who were missed in the model. Because growth in weight is different between male and female preterm infants,[10] WG specific for sex could be better for identifying at-risk infants with any ROP. Among infants with TR-ROP, the criteria of WG successfully identified 3 infants as being at risk despite having not met the criteria of GA and BW, which might be sufficiently high performance for practical use. The criteria for detecting slow WG are indeed only intended to capture the very few higher BW and older GA infants who develop TR-ROP. The vast majority of infants developing such severe disease have low BW and GA. Therefore, capturing these 3 specific infants is exactly what the 3 extra WG criteria are meant to do.

The screening criteria in this study were a bit higher than those in North America with regard to the GA and BW cut-off levels of 33 weeks and 1800 g in this study, vs 30 weeks and 1500 g in the United States and Canada. This strengthens the importance of performing such a validation, as there could be differences in the characteristics of infants being examined in Japan and North America, as a result. The fact that the criteria still had high sensitivity suggests that these criteria are a promising model.

There are several limitations to consider in this study. Although these data were collected retrospectively, this is not a limitation in the study because the data could generally be collected reliably even if retrospectively. BW, GA, and weight measurements are routinely collected in the neonatal intensive care unit, as are ROP examination results. In this study, we did not use hydrocephalus as a criterion. One reason that the remaining 5 criteria had 100% sensitivity is that the median BW was smaller among infants in this study than those in the G-ROP study (986 g in this study vs 1070 g in the G-ROP study). The BW and 3 WG thresholds of the G-ROP criteria should have sufficient safety margins for premature Japanese infants. If we use stricter criteria, it may be necessary to include the source of unhealthy nonphysiologic WG, such as hydrocephalus and late-onset circulatory collapse, as a criterion. The small sample size and inclusion of cases from a single institution are also limitations. Larger cohorts are required to verify reproducibility and reliability.

In conclusion, the criteria of the G-ROP study showed good accuracy as an exclusion tool in ROP screening among our Japanese cohort. The findings of this validation study suggest that the G-ROP criteria could be generalized to different racial/ethnic cohorts in high-income countries.

FUNDING/SUPPORT: THIS STUDY WAS SUPPORTED IN PART BY THE JAPAN SOCIETY FOR THE PROMOTION OF SCIENCE KAKENHI [Grant Number JP17K17859], with additional funding from the Takeda Science Foundation (Japan). Financial Disclosures: Yoko Fukushima: Grant (Otsuka). Ryo Kawasaki: Endowed chair (Topcon); Lecture fees (Alcon Pharma, Bayer, Novartis, Senju, Pfizer, Kowa, Takeda, Astellas, Santen, Nitto Medic, and Topcon); Consultant (Novartis, Novo Nordisk, Roche, Office Future, Predictive Analytic Group, and MICIN). Hirokazu Sakaguchi: Endowed chair (Menicon); Lecture fees (Alcon Pharma, Santen, and Alcon). Kohji Nishida: Grants (Alcon, AMO, Bayer, HOYA, Kowa, Menicon, MSD, Novartis, Otsuka, Pfizer, Rhoto, Santen, Senju, Topcon, and Wakamoto); Lecture fees (Alcon, Bayer, Boehringer Ingelheim, Chuo Sangio, HOYA, Johnson & Johnson, Kowa, Novartis, Otsuka, Pfizer, Santen, SEED, and Senju). The following authors have no financial disclosures: Akihiko Shiraki, Miwa Mitsuhashi, Hiromi Ineyama, and Yoshikazu Hatsukawa. All authors attest that they meet the current ICMJE criteria for authorship.

Other Acknowledgments: The authors thank Analisa Avila, ELS, of Edanz Group (www.edanzediting.com/ac) for editing a draft of this manuscript.

REFERENCES

1. Blencowe H, Lawn JE, Vazquez T, Fielder A, Gilbert C. Preterm-associated visual impairment and estimates of retinopathy of prematurity at regional and global levels for 2010. *Pediatr Res* 2013;74(Suppl 1):35–49.
2. Husain SM, Sinha AK, Bunce C, et al. Relationships between maternal ethnicity, gestational age, birth weight, weight gain, and severe retinopathy of prematurity. *J Pediatr* 2013;163(1):67–72.
3. Chan-Ling T, Gole GA, Quinn GE, Adamson SJ, Darlow BA. Pathophysiology, screening and treatment of ROP: a multidisciplinary perspective. *Prog Retin Eye Res* 2018;62:77–119.
4. Wilkinson AR, Haines L, Head K, Fielder AR. UK retinopathy of prematurity guideline. *Early Hum Dev* 2008;84(2):71–74.
5. 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. Screening examination of premature infants for retinopathy of prematurity. *Pediatrics* 2013;131(1):189–195.
6. Holmstrom GE, Hellstrom A, Jakobsson PG, Lundgren P, Tornqvist K, Wallin A. Swedish national register for retinopathy of prematurity (SWEDROP) and the evaluation of screening in Sweden. *Arch Ophthalmol* 2012;130(11):1418–1424.
7. Binenbaum G, Bell EF, Donohue P, et al. Development of modified screening criteria for retinopathy of prematurity: primary results from the Postnatal Growth and Retinopathy of Prematurity Study. *JAMA Ophthalmol* 2018;136(9):1034–1040.
8. International Committee for the Classification of Retinopathy of Prematurity. The International Classification of Retinopathy of Prematurity revisited. *Arch Ophthalmol* 2005;123(7):991–999.
9. Early Treatment for Retinopathy of Prematurity Cooperative Group, Good WV, Hardy RJ, et al. Final visual acuity results in the early treatment for retinopathy of prematurity study. *Arch Ophthalmol* 2010;128(6):663–671.
10. Fenton TR. A new growth chart for preterm babies: Babson and Benda's chart updated with recent data and a new format. *BMC Pediatr* 2003;3:13.