Validation of DIGIROP- Birth and DIGIROP- Screen for the discovery of retinopathy of prematurity requiring treatment in preterm births in Saudi Arabia

Huda K. Ahmedhussain, MD, Lina H. Raffa, MD, PhD, Amal M. Alosaimi, MD, Sarah K. Alessa, MD, Suzan Y. Alharbi, MD, Hashem Almarzouki, MD, Mansour A. AlQurashi, MD.

ABSTRACT

الأهداف: تقييم نموذجين DIGIROP لتنبؤ النوع الأول من اعتلال شبكية الخدج ومقارنتهما بخوارزميات أخرى معتمدة على الوزن في مجموعة من الأطفال الخدج في المملكة العربية السعودية.

المنهجية: تم تضمين 363 من الأطفال الخدج من وحدة العناية بحديثي الولادة و المولودين في الأسابيع 24-30 أسبوع من الحمل أو من كان وزنه عند الولادة 1500 غرام أو أقل في مستشفيين في جدة بين يناير 2015 وسبتمبر 2021. استخدم نموذج DIGIROP_Birth عمر الحمل عند الولادة، والجنس، ووزن الولادة، وعمر بداية حصول تغيرات اعتلال شبكية الخدج كعوامل تنبؤ، وتم تقدير المساحة تحت منحنى خاصية التشغيل المستقبلي (AUC) بفاصل ثقة 95%، والحساسية، والنوعية. تم حساب تقديرات مخاطر الإصابة بالاعتلال عبر نموذج DIGIROP_Screen في عمر ما بعد الولادة 6-14 أسبوعًا.

النتائج: كان متوسط عمر الحمل في العينة 27.94±1.6 أسبوعًا، ومتوسط وزن الولادة 1068.2±269.2 غرام. كانت حساسية نموذج DIGIROP_Birth 93.8% و النوعية 48.9%، والمساحة تحت المنحنى 0.70 (AUC)، والدقة 52.9%. بالنسبة لنموذج DIGIROP_Screen، تراوحت قيمةAUC للنماذج بين عمر ما بعد الولادة 6-14 أسبوعًا من 0.68 إلى 0.83 وتراوحت الحساسية من 73.3% إلى 96.8%. وأظهرت نماذجDIGIROP_Birth و DIGIROP_Screen أعلى دقة وقيمة AUC مقارنة بالخوارزميات الأخرى.

الخلاصة: أظهرت نماذج DIGIROP_Birth وDIGIROP_Screen قدرة تنبؤية عالية لخطر الإصابة بالنوع الأول من اعتلال شبكية الخدج في هذه المجموعة. يجب التحقق من قدرة هذه الأدوات على تحديد الأطفال ذوي الخطورة العالية وتجنب الفحص الروتيني في الأطفال ذوي الخطورة المنخفضة من خلال دراسات واسعة النطاق.

Objectives: To validate 2 DIGIROP prediction models for retinopathy of prematurity (ROP) type 1 and compare them to other weight-based algorithms in a premature Saudi Arabian infant cohort.

Methods: Preterm infants of 24-30 weeks' gestational age (GA) or body weight (BW) of ≤1500g who were admitted to the neonatal units of 2 Jeddah tertiary centers between January 2015 and September 2021 were included (N=363). The DIGIROP-Birth employed the birth GA, gender, birth weight, and age at ROP onset as predictors. The area under the receiver operating characteristic curve (AUC) with 95% confidence interval, specificity, and sensitivity were projected. The DIGIROP-Screen risk of risk were identified at 6-14 weeks postnatal age (PNA).

Results: The mean GA was 27.94±1.6 weeks and the mean BW was 1068.2±269.2 g. The DIGIROP-Birth had a sensitivity of 93.8%; specificity of 48.9%; AUC of 0.70; and accuracy of 52.9%. For DIGIROP-Screen, the AUC for models spanning PNA 6-14 weeks varied from 0.68-0.83, and sensitivity varied from 73.3-96.8%. The DIGIROP-Birth and DIGIROP-Screen showed the highest accuracy and AUC value in comparison to other ROP prediction models.

Conclusion: The 2 models demonstrated high predictive capacity for type 1 ROP risk assessment in this cohort. The potential of these tools for identifying high-risk infants and avoiding standard ROP screening in low-risk infants needs to be verified through large-scale studies.

Keywords: algorithms, CO-ROP, G-ROP, ROPscore, WINROP

Saudi Med J 2025; Vol. 46 (4): 345-351
doi: 10.15537/smj.2025.46.4.20240773

From the Department of Ophthalmology (Ahmedhussain, Raffa, Alessa), King Abdulaziz University Hospital, from the Faculty of Medicine (Ahmedhussain, Raffa, Alessa), King Abdulaziz University, from the Department of Obstetrics and Gynecology (Alosaimi); from the Department of Ophthalmology (Almarzouki); from the Department of Pediatrics (AlQurashi), Neonatology Division, King Abdulaziz Medical City, from the Department of Ophthalmology (Alharbi), Jeddah Eye Hospital, from the College of Medicine (Almarzouki, AlQurashi), King Saud bin Abdulaziz University for Health Sciences, and from King Abdullah International Medical Research Center (Almarzouki, AlQurashi), Ministry of National Guard Health Affairs, Jeddah, Kingdom of Saudi Arabia.

Received 3rd September 2024. Accepted 20th March 2025.

Address correspondence and reprint request to: Dr. Lina H. Raffa, Department of Ophthalmology, King Abdulaziz University Hospital, Jeddah, Kingdom of Saudi Arabia. E-mail: lraffa@kau.med.sa
ORCID ID: https://orcid.org/0000-0001-8211-2353

In preterm infants, a disorder known as retinopathy of prematurity (ROP) affects the retinal blood vessels. If ROP is not diagnosed or treated on time, it can eventually lead to blindness. Therefore, its early prediction and treatment are crucial. According to the currently used ROP screening criteria, even infants at low risk are exposed to painful eye exams.[1] According to the standard criteria in Saudi Arabia, ROP screening is carried out in neonates with ≤1500 g birth weight (BW) or ≤32 weeks gestational age (GA).[2] However, according to a study carried out in 20 neonatal intensive care units (Level 3), only 7.7% of the 2188 Saudi Arabian newborns screened by the Saudi Arabian retinopathy of prematurity national telemedicine programme (SAROP) had ROP that required therapy.[3] From a healthcare and economic standpoint, individualized risk assessments would help optimize the time and frequency of screenings, particularly in rural regions, where ROP experts are scarce. In addition, more efficient timing of screening could minimize the number of low-risk infants who are tested and maximize the detection rate of high-risk individuals.

To reduce the number of infants undergoing unnecessary examinations and to focus on those at higher risk of sight-threatening retinopathy of prematurity (ROP), several prediction models and algorithms—such as the weight, insulin-like growth factor, neonatal ROP (WINROP) algorithm; the postnatal growth and ROP (G-ROP) criteria; the Colorado-ROP (CO-ROP) model; and the ROPScore—have been developed and validated in various populations.[4-8] Multiple algorithms have been evaluated in Saudi Arabia and have shown variable sensitivity and specificity. For example, Raffa et al[9] validated a Swedish tool called WINROP and showed that it had 100% sensitivity and 31.5% specificity in identifying type 1 ROP. However, this tool requires the weekly weight input into the website, and this might be time-consuming compared to other currently developed tools based on simple birth characteristics. Furthermore, Raffa et al[10] validated the G-ROP 2 algorithm with sensitivity value of 100% and a low specificity value of 16.7% and Alexandria ROP model (Alex-ROP) with a sensitivity value of 80% and a low specificity value of 41%, for the detection of ROP requiring treatment.[11]

Recently, Sahlgrenska Center for Pediatric Ophthalmology Research at the University of Gothenburg, Sweden, developed a new prediction model for infants birthed at 24-30 weeks' GA. They used a website that only includes birth characteristics (DIGIROP-Birth) and ROP screening data (DIGIROP-Screen).[12,13] The model calculates the percent risk (within the 95% confidence range) for sight-threatening ROP that requires treatment.[12,13] This tool presents an easier way to detect infants at risk compared to other published tools that require weight measurements at specific cut-off points.[4-8] These models have been validated in cohorts from Sweden, Germany, and the United States, with results published in the original articles. Further validations have been conducted in Portuguese and Chinese cohorts.[12-16]

In this study, we aim to further validate the DIGIROP prediction models by applying them in a cohort of premature infants from Saudi Arabia. We also compare their performance to other weight-based algorithms (WINROP, CO-ROP G-ROP and ROPScore) within the context of a developing country.

Methods. This study included all preterm infants with 24-30 weeks GA or ≤1500g BW were admitted to the neonatal units of 2 Jeddah tertiary care centers and completed ROP screening between January 2015 and September 2021. The study duration involved 473 children who underwent ROP screening, of which 110 were eliminated, leaving 363 infants in the final cohort. Infants with hydrocephalus (n=11), severe congenital abnormalities or hydrops (n=2), or those who had undergone intestinal surgery (n=11, such as for necrotizing enterocolitis) were excluded. Infants with GA of <24 weeks (n=4) or of >30 weeks (n=82) were not included.

This study adheres to the Declaration of Helsinki declarations and was authorized by King Abdullah International Medical Research Center, Jeddah, Saudi Arabia, institutional review board (approval no. JED-20-427780-86181) and King Abdulaziz University, Jeddah, Saudi Arabia (approval no. KAUH No 451-22). As this study was retrospective, informed consent was not required.

The screening procedures of the American Academy of Pediatrics (2013 and 2018) served as a basis for the ROP screening timing and criteria. These guidelines advise checking all newborns with <1500g BW or born before 30 weeks, as well as infants with 1500-2000g BW or born at or after 30 weeks GA who have an unstable clinical course or who have been identified as having a high risk of ROP by the attending pediatrician.[1] According to the ROP screening schedule, a retinal specialist or a trained pediatric ophthalmologist carried out the examination at 31 weeks GA or 4 weeks after

Disclosure. Authors have no conflict of interests, and the work was not supported or funded by any drug company.

birth, whichever came later. The ROP diagnosis and treatment suggestion guidelines are provided by the international classification of ROP revisited and the early treatment for ROP study.[17,18] According to previously proposed treatment guidelines (in the "efficacy of intravitreal bevacizumab for stage 3+ retinopathy of prematurity" [BEATROP] and "ranibizumab versus laser therapy for the treatment of very low BW infants with ROP" [RAINBOW] studies), ROP was addressed with laser therapy or anti-vascular endothelial growth (VEGF) factors.[18,19]

In case of ROP diagnosis, the initial eye exam date, ROP detection date, worst stage, worst zone, presence of plus disease, ROP prognosis, and therapy were all reported. The infants were followed up until ROP eventually resolved, with or without treatment, or, in the absence of ROP, until the retina had fully vascularized. The DIGIROP-Birth estimates were calculated online using the following variables: BW, GA (weeks and days), and gender for predicting early risk of ROP. The DIGIROP-Birth estimations of risk, ROP status, and age at ROP onset were included as input variables for DIGIROP-Screen.[12,13] The risk probability was calculated by the application to determine whether the infant needed to be released or screened for ROP. At postnatal age (PNA) of 6-14 weeks, DIGIROP-Screen estimates of risk for ROP treatment were computed.

For ROPscore algorithm calculation, the following variable data were collected retrospectively from the infants' medical records: GA, BW, proportionate weight gain, oxygen usage in nasal continuous positive airway pressure or mechanical ventilation, and blood transfusions up to PNA 6 weeks.[6] The score is calculated once for each infant and the specified cut-off point for ROP of any stage is 11 and high-risk ROP cut-off point is 14.5. We incorporated postnatal weight gain at one month set at ≤650 g to validate CO-ROP and at ≤400 g to validate high grade CO-ROP (HgCO-ROP).[5] The G-ROP model was also validated and consisted of the following ROP screening criteria: BW of <1051g; GA of <28 weeks; weight gain of <120 g over the second 10 days post-birth, <180g over the third 10 days post-birth, or <170g over the fourth 10 days post-birth; or hydrocephalus diagnosed through brain imaging scan.[4] Using the online WINROP software, weekly weight measurements were retroactively obtained and plotted along with the final ROP results to validate its diagnostic accuracy to detect type 1 ROP. Infants with missing weight inputs were excluded from that respective model comparison.

Statistical analysis. The data were statistically analyzed with The Statistical Package for the Social Sciences, version 23.0 (IBM Corp., Armonk, NY, USA). For nominal and categorical variables, percentages and counts were determined to define the properties of the study variables. Continuous data were represented means and standard deviations (SDs). The link between categorical variables was determined through the Chi-square test and Fischer's exact test. The post-hoc test of one-way ANOVA with the least significant difference (LSD) was utilized to compare more than 2 groups, for analyses of data that were considered to be normally distributed. For data that were not normally distributed, the Games-Howell test was used instead of the LSD test to compare several groups. Specificity, sensitivity, negative and positive predictive values, prevalence of disease, and accuracy were calculated in percentages for all the tools that were evaluated. The "exact" Clopper-Pearson 95% confidence intervals (CIs) were used for accuracy, specificity, and sensitivity. The null hypothesis was rejected if the conventional *p*-value was <0.05.

Results. Among 363 infants who were included in this study from 2 neonatal care centers, 48.2% were boys. The mean GA was 27.9±1.6 weeks. The mean BW was 1068.2±269.2g. Altogether, 125 (34.4%) of 363 infants developed ROP in this cohort. The model detected any stage ROP in 25.6% of the infants, and it spontaneously regressed in all cases. Type 1 ROP was present in 32 (8.8%) of the participants (mean GA: 26.2±1.7 weeks; mean BW: 786.6±167.3 g). The modality of treatment was laser therapy in 68.8%, intravitreal anti-VEGF with either bevacizumab or ranibizumab in 15.6%, combined therapy (laser and intravitreal anti-VEGF injection) in 12.5%, and surgical intervention in one eye in the case of one (3.1%) infant. The mean PNA at first ROP treatment was 10.91±3.8 weeks. The birth characteristics of the cohort of preterm infants with and without ROP are detailed in Table 1.

The DIGIROP-Birth showed 48.9% specificity, 93.8% sensitivity, area under the curve (AUC) of 0.70, and 52.9% accuracy. For DIGIROP-Screen, the AUCs across multiple models spanning PNA 6-14 weeks varied from 0.68-0.83, and the sensitivity varied from 73.3-96.8%. At PNA 6 weeks, the accuracy was 49.3%, and at PNA 14 weeks, it was 83.2% (Table 2). Figure 1 displays the receiver operating characteristic (ROC) curves for DIGIROP-Birth and DIGIROP-Screen at PNA 6-14 weeks. Out of the 110 infants (43% of the total screened) who were considered to have low treatable ROP risk according to DIGIROP-Birth, only one infant developed type 1 ROP. Moreover, out of 103 (41%) who were considered to have low treatable ROP

Table 1 - Baseline characteristics of the study cohort by ROP status (N=363).

Variables	Types of ROP			*P*-values
	No ROP	Treatable ROP	Any stage ROP	
Total	238 (65.6)	32 (8.8)	93 (25.6)	-
Age (years), mean±SD	4.18±1.6*	3.56±1.7*	4.77±1.8†	<0.001*,‡
GA (weeks), mean±SD	28.37±1.4*	26.16±1.7†	27.47±1.6‡	<0.001*,§
BW (grams), mean±SD	1145.12±248.6*	786.56±167.3†	968.23±252.1‡	<0.001*,§
Probability (n=368), mean±SD	1.63±3.7*	11.79±11.1†	4.92±7.9‡	<0.001*,§
Gender				
Male (n=175)	116 (66.3)	14 (8.0)	45 (25.7)	0.868
Female (n=188)	122 (64.9)	18 (9.6)	48 (25.5)	
GA				
24 (n=14)	2 (14.3)	6 (42.9)	6 (42.9)	
25 (n=20)	9 (45.0)	8 (40.0)	3 (15.0)	
26 (n=38)	13 (34.2)	6 (15.8)	19 (50.0)	
27 (n=52)	31 (59.6)	4 (7.7)	17 (32.7)	<0.001†
28 (n=78)	56 (71.8)	3 (3.8)	19 (24.4)	
29 (n=99)	75 (75.8)	5 (5.1)	19 (19.2)	
30 (n=62)	52 (83.9)	0 (0.0)	10 (16.1)	
Worst ROP stage either eye				
1 (n=73)	0 (0.0)	0 (0.0)	73 (100)	
2 (n=36)	0 (0.0)	17 (47.2)	19 (52.8)	
3 (n=15)	0 (0.0)	14 (93.3)	1 (6.7)	<0.001†
4 (n=1)	0 (0.0)	1 (100)	0 (0.0)	
N/A (n=238)	238 (100)	0 (0.0)	0 (0.0)	

Values are presented as numbers and percentages (%). *Significant using One-Way ANOVA test at <0.05 level. †Significant using Fisher's Exact test at <0.05 level. ‡Post-Hoc test = LSD. §Post-Hoc test = Games-Howell. ROP: retinopathy of prematurity, GA: gestational age, BW: body weight, N/A: not available, SD: standard deviation

Table 2 - Sensitivity, specificity, positive predictive value, negative predictive value, model accuracy, and area under the receiver operating characteristic curve with 95% confidence interval for DIGIROP-Birth and DIGIROP-Screen for type 1 retinopathy of prematurity risk prediction (N=363).

Type 1 ROP	Sensitivity	Specificity	Disease prevalence	Positive predictive value	Negative predictive value	Accuracy	Area under the curve
DIGIROP-Birth	93.8 (79.2-99.2)	48.9 (43.4-54.5)	8.8 (6.1-12.2)	15.1 (13.4-16.9)	98.8 (95.5-99.7)	52.9 (47.6-58.1)	0.70 (0.58-0.82)
DIGIROP-Screen PNA6w	93.8 (79.2-99.2)	45.0 (39.5-50.5)	8.9 (6.1-12.3)	14.2 (12.7-15.9)	98.7 (95.1-99.7)	49.3 (44.0-54.6)	0.68 (0.56-0.81)
DIGIROP-Screen PNA7w	96.8 (83.3-99.9)	45.0 (39.5-50.5)	8.6 (5.9-12.0)	14.2 (12.9-15.7)	99.3 (95.5-99.9)	49.4 (44.2-54.7)	0.68 (0.56-0.81)
DIGIROP-Screen PNA8w	89.7 (72.7-97.8)	57.5 (51.9-62.9)	8.1 (5.5-11.4)	15.7 (13.5-18.1)	98.4 (95.6-99.5)	60.1 (54.8-65.2)	0.75 (0.64-0.86)
DIGIROP-Screen PNA9w	80.8 (60.7-93.5)	66.0 (60.6-71.1)	7.3 (4.8-10.6)	15.8 (12.9-19.3)	97.8 (95.2-99.0)	67.0 (61.9-71.9)	0.75 (0.62-0.88)
DIGIROP-Screen PNA10w	80.0 (59.3-93.2)	72.6 (67.5-77.4)	7.1 (4.6-10.3)	18.2 (14.6-22.4)	98.0 (95.6-99.1)	73.2 (68.2-77.7)	0.82 (0.72-0.92)
DIGIROP-Screen PNA11w	85.7 (63.7-97.0)	74.5 (69.4-79.1)	6.0 (3.8-9.0)	17.7 (14.3-21.7)	98.8 (96.6-99.6)	75.1 (70.3-79.6)	0.83 (0.73-0.93)
DIGIROP-Screen PNA12w	94.1 (71.3-99.9)	75.1 (70.0-79.7)	4.9 (2.9-7.8)	16.3 (13.5-19.6)	99.6 (97.4-99.9)	76.0 (71.2-80.4)	0.83 (0.74-0.93)
DIGIROP-Screen PNA13w	73.3 (44.9-92.2)	83.6 (79.1-87.4)	4.4 (2.5-7.1)	16.9 (12.1-23.1)	98.6 (96.7-99.4)	83.1 (78.8-86.9)	0.75 (0.59-0.91)
DIGIROP-Screen PNA14w	75.0 (42.8-94.5)	83.5 (79.1-87.4)	3.5 (1.8-6.1)	14.3 (10.0-20.0)	98.9 (97.2-99.6)	83.2 (78.8-87.1)	0.79 (0.65-0.94)

Values are presented as numbers with 95% confidence intervals. ROP: retinopathy of prematurity, PNA: Postnatal age in weeks

risk according to the DIGIROP-Screen model at PNA 6 weeks, the same infant developed type 1 ROP. The infant was a female birthed at 29 weeks with 938 g BW and had intraventricular hemorrhage and severe anemia that necessitated multiple blood transfusions; she also required mechanical ventilation and had a prolonged

hospitalization period of 112 days and 9 days of parenteral nutrition.

Out of the 393 infants, 272 were included in the subset analysis to compare the DIGIROP with the other ROP prediction models focusing on those with complete data available for an accurate comparison (Table 3). However, G-ROP 1 and 2, included only 146 infants. Compared to other models for the prediction of type 1 ROP, DIGIROP-Birth and DIGIROP-Screen at PNA 6 weeks demonstrated the most accuracy (47.1% and 44.5%) and highest AUC (0.68 and 0.65). The ROC curves for the estimations from the 2 models are shown in Figure 2. The ROPscore, WINROP, G-ROP 1 and G-ROP 2 models had a sensitivity of 100%, but their specificity were 28.5% for ROPscore, 20.2% for WINROP, 15.2% for G-ROP 1, and 9.1% for G-ROP 2 and AUC were 0.63 for ROPscore, 0.61 for WINROP, 0.58 for G-ROP 1, and 0.55 for G-ROP 2 were lower to those of the DIGIROP-Screen model. In comparison to DIGIROP-Screen, CO-ROP (AUC=0.51) and HgCO-ROP (AUC=0.56) had considerably lower sensitivity and specificity (Table 3).

Discussion. The present study evaluates the DIGIROP-Birth and DIGIROP-Screen models for the identification of infants at risk for type 1 ROP. These tools have now been validated in a cohort from Saudi Arabia, providing an important starting point for their potential application in ROP detection in developing countries. In our cohort, both models perfomed well, with AUC values of 0.68-0.83 and accuracy ranging from 49.3-83.2% that improved with increase in PNA. In comparison, Chen et al[15] evaluated DIGIROP-Birth model in a Chinese population and reported a less satisfactory performance, with an AUC of 0.634. This difference could be explained by the disparities in BW and GA between the 2 cohorts: that is, our cohort included younger infants with lower BW (1068.2±269.2g) and gestational age (27.9±1.6 weeks), while the Chinese cohort had a mean BW of 1237.0±236.9g and GA of 28.8±1.3 weeks.[15] Furthermore, Ana et al[14] reported a moderate accuracy of 0.70 in their Portuguese population. This was lower than that reported in the original evaluation studies by Pivodic et al[13] in the DIGIROP-Birth development model, as they reported an AUC of 0.87 in the US cohort, 0.90 in the European group, and 0.93 in the Swedish cohort.[16] The differences in the performance of DIGIROP-Birth across studies could be attributable to differences in race and ethnicity between the studied populations.

In the present cohort, DIGIROP-Screen showed increased accuracy as the PNA increased from 49.3% at PNA 6 weeks up to 83.2% at PNA 14 weeks (AUC=0.68-0.79). The best performance (AUC=0.83) was found at PNA 11 and 12 weeks (sensitivity =85.7% and 94.1%, specificity, 74.5% and 75.1%). Although the AUC is slightly lower than Pivodic et al's findings[13] in their development cohort (n=6991, 0.93) and validation cohort (N=1241, 0.92), our cohort

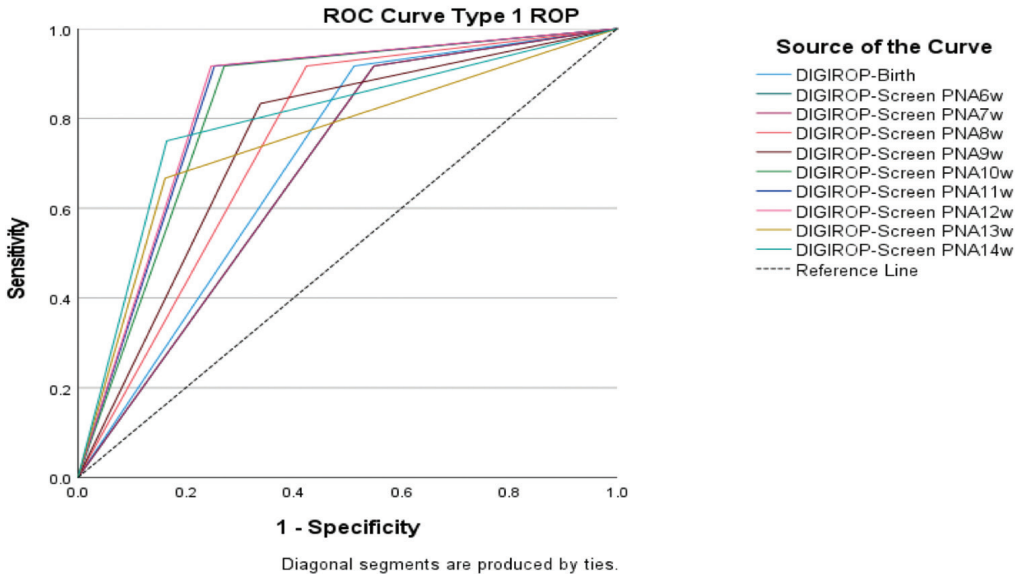

Figure 1 - Receiver operating characteristic (ROC) curves for type 1 retinopathy of prematurity (ROP) risk estimates obtained by DIGIROP-Birth and DIGIROP-Screen at postnatal age 6-14 weeks. PNA: postnatal age

Table 3 - Comparison of DIGIROP-Screen with other existing retinopathy of prematurity prediction models.

Type 1 ROP	Total	Sensitivity	Specificity	Disease prevalence	PPV	NPV	Accuracy	AUC
DIGIROP-Birth	272	94.7 (74.0-99.9)	43.5 (37.3-49.8)	7.0 (4.3-10.7)	11.2 (9.8-12.8)	99.1 (94.2-99.9)	47.1 (41.0-53.2)	0.68 (0.54-0.82)
DIGIROP-Screen PNA6w	272	94.7 (74.0-99.9)	40.7 (34.6-47.0)	7.0 (4.3-10.7)	10.7 (9.4-12.2)	99.0 (93.8-99.9)	44.5 (38.5-50.6)	0.65 (0.51-0.79)
ROPSCORE alarm	272	100 (82.4-100)	28.5 (23.0-34.5)	7.0 (4.3-10.7)	9.5 (8.9-10.2)	100 (95.0-100)	33.5 (27.9-39.4)	0.63 (0.49-0.77)
CO-ROP	272	94.7 (74.0-99.9)	5.1 (2.8-8.6)	7.0 (4.3-10.7)	7.0 (6.3-7.7)	92.9 (64.2-99.0)	11.4 (7.9-15.8)	0.51 (0.34-0.69)
HgCO-ROP	272	89.5 (66.9-98.7)	19.4 (14.7-24.8)	7.0 (4.3-10.7)	7.7 (6.6-9.0)	96.1 (86.6-98.9)	24.3 (19.3-29.8)	0.56 (0.39-0.72)
WINROP Alarm	272	100 (82.4-100)	20.2 (15.4-25.6)	7.0 (4.3-10.7)	8.6 (8.1-9.1)	100 (93.0-100)	25.7 (20.7-31.4)	0.61 (0.46-0.75)
G-ROP 1	146	100 (76.8-100)	15.2 (9.5-22.4)	9.6 (5.3-15.6)	11.1 (10.4-11.8)	100 (83.2-100)	23.3 (16.7-31.0)	0.58 (0.42-0.73)
G-ROP 2	146	100 (76.8-100)	9.1 (4.8-15.3)	9.6 (5.3-15.6)	10.5 (10.0-11.0)	100 (73.5-100)	17.8 (12.0-25.0)	0.55 (0.38-0.71)

Values are presented as numbers with 95% confidence intervals. PNA : postnatal age in weeks, ROP: retinopathy of prematurity, CO-ROP: Colorado retinopathy of prematurity, HgCO-ROP: high grade Colorado-retinopathy of prematurity model, WINROP: weight insulin-like growth factor 1 neonatal retinopathy of prematurity, G-ROP: postnatal growth and retinopathy of prematurity, PPV: positive predictive value, NPV: negative predictive value, AUC: area under the curve

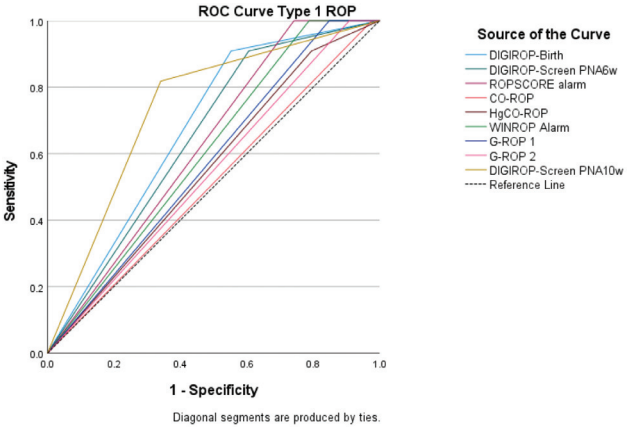

Figure 2 - Receiver operating characteristic (ROC) curves for type 1 retinopathy of prematurity (ROP) risk estimates obtained by DIGIROP-Birth and DIGIROP-Screen at PNA 6 weeks in comparison to other existing algorithms.

showed a higher specificity from PNA 6 weeks (45%) to PNA 14 weeks (83.5%) compared to the Dev and Val groups, for which the specificity was 48-75.3% and 46.3-72.1%.[13] In the latest Swedish cohort, Pivodic[16] reported DIGIROP-Screen AUC values between 0.93-0.97 for 6-14 weeks' PNA, with an accuracy of 51.4-76.5% and a specificity of 48.9-76.3%. Thus, the performance level of the tool was maintained in the subsequent Swedish cohort, too. Importantly, this tool appears to demonstrate good performance in our cohort from Saudi Arabia, indicating that it might have potential for application in this population.

Compared to other algorithms for identifying type 1 ROP that were examined in our cohort, that is, ROPScore, WINROP, CO-ROP, HgCO ROP and G-ROP 1 and 2, DIGIROP-Birth and DIGIROP-Screen appeared to have better accuracy and specificity. That is, they performed the best compared to all the other tools. In agreement with our findings, Raffa et al[9] also previously reported that the WINROP, G-ROP, and Alex-ROP tools had lower specificity than the DIGIROP tools in this cohort.[9-11,20] Further, they found that while WINROP and G-ROP have excellent sensitivity (100%) for identifying all type 1 ROP infants who require treatment, the specificity was much lower than that of the DIGIROP tools.[9,11]

Study strengths & limitations. The present findings demonstrate that DIGIROP-Birth and DIGIROP-Screen surpass previously validated algorithms for ROP prediction in Saudi Arabia, but they did not achieve 100% sensitivity, as they missed one infant with type 1 ROP. Advantages of DIGIROP includes its ability to reduce the need for frequent eye examinations by nearly a half allowing for better resource allocation by identifying high risk infants. Additionally, its user-friendly interface and integration of simple clinical parameters further enhance its utility making it a valuable tool for clinicians managing ROP in neonatal care setting. Another strength of the study is that it compared DIGIROP across several ROP prediction models in the same cohort providing a comprehensive evaluation of its performance. However, one limitation is that, specifically for the G-ROP algorithm, only a smaller sample of 146 infants was available for comparison with the other models, and conclusion must be drawn with caution. Another limitation of our study is the inclusion of different ethnicities of the involved population. Although the mixed ethnicity in the studied group is representative of the diversity in

Saudi Arabia, it is similar to the large cohort included in both the European and US cohorts on whom the validation models showed high predicative abilities, therefore these models might still be generalizable for different individuals of different ethnicities.[12] In the future, these tools need to be validated in larger populations to obtain more generalizable assessments.

In conclusion, we have studied and validated a new clinical decision support tool for ROP screening in the Saudi Arabian population, with the potential to reduce unnecessary examinations in infants. The accuracy showed promising release of nearly half of the screened infants with 93.8% sensitivity and 48.9% specificity. The findings are promising, but these algorithms need to be studied further in a larger population in order to confirm their applicability for the screening of high- and low-risk infants for type 1 ROP in Saudi Arabia.

Acknowledgment. *The authors gratefully acknowledge Oxford Science Editing Ltd. (www.oxfordscience.org) for their English language editing.*

References

1. Fierson WM. Screening examination of premature infants for retinopathy of prematurity. *Pediatrics* 2018; 142: e20183061.
2. Al Amro SA, Al Aql F, Al Hajar S, Al Dhibi H, Al Nemri A, Mousa A, et al. Practical guidelines for screening and treatment of retinopathy of prematurity in Saudi Arabia. *Saudi J Ophthalmol* 2018; 32: 222-226.
3. Al Amro SA, Al Ghamdi SM, Abouammoh MA, Al Aql F, Ahmad K, Alsulaiman SM. Saudi Arabian retinopathy of prematurity national telemedicine program: achievements and challenges. *Saudi J Ophthalmol* 2022; 36: 278-282.
4. Binenbaum G, Bell EF, Donohue P, Quinn G, Shaffer J, Tomlinson LA, 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: 1034-1040.
5. Cao JH, Wagner BD, McCourt EA, Cerda A, Sillau S, Palestine A, et al. The Colorado-retinopathy of prematurity model (CO-ROP): postnatal weight gain screening algorithm. *J AAPOS* 2016; 20: 19-24.
6. Eckert GU, Fortes Filho JB, Maia M, Procianoy RS. A predictive score for retinopathy of prematurity in very low birth weight preterm infants. *Eye (Lond)* 2012; 26: 400-406.
7. Hellström A, Hård AL, Engström E, Niklasson A, Andersson E, Smith L, et al. Early weight gain predicts retinopathy in preterm infants: new, simple, efficient approach to screening. *Pediatrics* 2009; 123: e638-e645.
8. Löfqvist C, Andersson E, Sigurdsson J, Engström E, Hård AL, Niklasson A, et al. Longitudinal postnatal weight and insulin-like growth factor I measurements in the prediction of retinopathy of prematurity. *Arch Ophthalmol* 2006; 124: 1711-1718.
9. Raffa LH, Alessa SK, Alamri AS, Malaikah RH. Prediction of retinopathy of prematurity using the screening algorithm WINROP in a Saudi cohort of preterm infants. *Saudi Med J* 2020; 41: 622-627.
10. Raffa LH, Akeely OM, Alariefy SA, Alharbi FA, Alkhateeb MA, Khan MA. Retrospective validation of the Alexandria retinopathy of prematurity model in preterm infants in Saudi Arabia. *J Clin Neonatol* 2023; 12: 7-12.
11. Raffa LH, Bugshan N, Aljuhani R, Alghoribi R, Alharbi A, Aljohani W. Retrospective validation of the postnatal growth and retinopathy of prematurity (G-ROP) criteria in a developing country. *Niger J Clin Pract* 2023; 26: 217-222.
12. Pivodic A, Hård AL, Löfqvist C, Smith LEH, Wu C, Bründer MC, et al. Individual risk prediction for sight-threatening retinopathy of prematurity using birth characteristics. *JAMA Ophthalmol* 2020; 138: 21-29.
13. Pivodic A, Johansson H, Smith LEH, Hård AL, Löfqvist C, Yoder BA, et al. Development and validation of a new clinical decision support tool to optimize screening for retinopathy of prematurity. *Br J Ophthalmol* 2022; 106: 1573-1580.
14. Almeida AC, Borrego LM, Brízido M, de Figueiredo MB, Teixeira F, Coelho C, et al. DIGIROP efficacy for detecting treatment-requiring retinopathy of prematurity in a Portuguese cohort. *Eye (Lond)* 2022; 36: 463-469.
15. Chen S, Wu R, Chen H, Ma W, Du S, Li C, et al. Validation of the DIGIROP-birth model in a Chinese cohort. *BMC Ophthalmol* 2021; 21: 236.
16. Pivodic A, E H Smith L, Hård AL, Löfqvist C, Almeida AC, Al-Hawasi A, et al. Validation of DIGIROP models and decision support tool for prediction of treatment for retinopathy of prematurity on a contemporary Swedish cohort. *Br J Ophthalmol* 2023; 107: 1132-1138.
17. International Committee for the Classification of Retinopathy of Prematurity. The international classification of retinopathy of prematurity revisited. *Arch Ophthalmol* 2005; 123: 991-999.
18. Mintz-Hittner HA, Kennedy KA, Chuang AZ. Efficacy of intravitreal bevacizumab for stage 3+ retinopathy of prematurity. *N Engl J Med* 2011; 364: 603-615.
19. Stahl A, Lepore D, Fielder A, Fleck B, Reynolds JD, Chiang MF, et al. Ranibizumab versus laser therapy for the treatment of very low birthweight infants with retinopathy of prematurity (RAINBOW): an open-label randomised controlled trial. *Lancet* 2019; 394: 1551-1559.
20. Raffa L, Alamri A, Alosaimi A, Alessa S, Alharbi S, Ahmedhussain H, et al. Validation of 3 weight gain-based algorithms as a screening tool to detect retinopathy of prematurity: a multicenter study. *Indian J Ophthalmol* 2023; 71: 2555-2560.