BMJ Paediatrics Open

Smartphone-based screening of neonatal jaundice in three populations in low and middle-income countries: a cross-sectional study

Anders Aune ,[1,2] Gabriela Jímenez-Díaz ,[1,3] Lobke Marijn Gierman,[3] Gunnar Vartdal,[3] Mónica Reyes Berlanga,[4,5] Jorge Tusoy,[1] Håkon Bergseng,[2,6] Sunila Shakya,[7] Elisabeth Darj[1,8]

To cite: Aune A, Jímenez-Díaz G, Gierman LM, *et al*. Smartphone-based screening of neonatal jaundice in three populations in low and middle-income countries: a cross-sectional study. *BMJ Paediatrics Open* 2025;**9**:e002242. doi:10.1136/bmjpo-2023-002242

Received 17 December 2024
Accepted 21 April 2025

For numbered affiliations see end of article.

Correspondence to
Dr Anders Aune; a.aune@ntnu.no

ABSTRACT

Background Neonatal hyperbilirubinemia (NHB) is a significant cause of morbidity and mortality, particularly in low and middle-income countries (LMICs). Transcutaneous bilirubinometers offer a non-invasive method for assessing NHB but have limited availability due to cost and maintenance requirements. Visual assessment of jaundice is shown to be inaccurate. Smartphone-based technologies have the potential to provide innovative and accessible healthcare solutions. This study aimed to evaluate the Picterus system, a smartphone-based tool for screening of NHB, in three non-Caucasian populations in LMICs.

Methods Between 2018 and 2022, cross-sectional studies were conducted in three countries: Mexico, Nepal and the Philippines. Newborns meeting the inclusion criteria were recruited, and data on demographic characteristics, skin type and visual assessment of jaundice were collected. Bilirubin levels were measured using both the Picterus system and total serum bilirubin (TSB) analysis. Correlation analyses, Bland-Altman plots and receiver operating characteristic (ROC) curves were used to evaluate the Picterus system.

Results A total of 416 infants were included in the analysis. The Picterus smartphone system demonstrated a significant positive correlation with TSB levels across all sites (r=0.76). The correlation coefficient was significantly higher in Mexico compared with Nepal and the Philippines. Bland-Altman plots showed limits of agreement ±89.2 µmol/L. Picterus values were underestimated in Mexico, whereas they were overestimated in Nepal and the Philippines. ROC analysis for detection of infants with TSB >225 µmol/L indicated that the Picterus system had higher sensitivity and specificity compared with visual assessment using the Kramer scale.

Discussion This study shows that the Picterus system can potentially be used in screening for neonatal jaundice in populations with moderate dark skin types. Further studies are needed before the system can be used in clinical practice.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Neonatal hyperbilirubinemia remains an important cause of morbidity and mortality globally.
⇒ Current available methods for identification are either too expensive or related to large uncertainties.
⇒ Previous clinical testing in newborns with fair skin types has shown that bilirubin values from a smartphone-based system are highly correlated to bilirubin measured in blood.

WHAT THIS STUDY ADDS

⇒ Clinical testing of a smartphone-based system to screen for neonatal hyperbilirubinemia in three different populations in Mexico, Nepal and the Philippines showed a high correlation between bilirubin values from the smartphone system to bilirubin measurements in blood.
⇒ The smartphone-based system was better than visual assessment to detect newborns with hyperbilirubinaemia.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ There is a potential to improve identification of jaundice by using novel smartphone-based tools.
⇒ Further studies are needed before the system can be used in clinical practice.

BACKGROUND

Jaundice is a common condition among newborns and is caused by elevated levels of bilirubin, hyperbilirubinemia.[1 2] Despite that most cases are harmless and self-limiting, newborn hyperbilirubinemia (NHB) remains an important cause of morbidity and mortality globally. When bilirubin accumulates in the brain, it can result in permanent damage as cerebral palsy and deafness or can even be fatal.[3]

In high-income settings, these events are rare,[4] but in low and middle-income countries (LMICs), NHB remains an important cause of newborn morbidity and mortality.[5 6] The challenges related to NHB in LMIC are multifactorial, and among these are correctly identifying newborns at risk.[7]

Assessing jaundice visually is challenging, with a large number of infants being misclassified.[8–10] Non-invasive assessment of jaundice using transcutaneous bilirubinometers (TcB) has previously been proven to be highly correlated to total serum bilirubin (TSB) levels in multiple regions, including Latin-America, South-East and South-East Asia.[11–13] TcBs have been available for decades, but price and need for maintenance and calibration are limiting factors for widespread distribution of these devices in LMIC.[14] Smartphone-based technologies have opened up for new and innovative ways of delivering healthcare, which could be especially relevant in low-income settings.[15]

Such a smartphone-based system, Picterus, was initiated and developed at the Norwegian University of Science and Technology and later patented.[16] Previously, we have reported results of clinical testing of the Picterus system.[17 18] The newborns in these studies have mainly been of Caucasian descent with parents with Fitzpatrick skin types I and II. Studies of the use of TcB on newborns with different ethnicities have shown that skin type could influence the results.[19 20] In this paper, we present the results of an evaluation of the Picterus system in three populations with a non-Caucasian ethnicity and different skin types in LMICs.

METHODS

Clinical study design

The aim of this study was to compare bilirubin estimates from the Picterus system to TSB, as TSB is the gold standard in jaundice assessment of newborns. Cross-sectional clinical studies were performed at three different sites. (1) In Mexico, at Hospital Materno Infantil de Irapuato, Guanajuato, from July to August 2018. (2) In Nepal at Dhulikhel Hospital, Kathmandu University, from June to July 2019 and (3) in Philippines at Governor Celestino Gallares Memorial Hospital, Tagbilaran City, Bohol, in February 2022.

Recruitment

Participants were recruited from maternity wards in all three hospitals. For the data collection in Mexico, participants were in addition recruited at an outpatient clinic where newborns were brought for check-ups regarding breastfeeding, weight gain, clinically suspected jaundice or other health concerns.

Newborns were included when they were less than 14 days of age and had a blood sample drawn, either for clinically suspected jaundice or for other purposes. As recruitment was dependent on a clinical need for blood sampling, newborns were included at different days of age, up to 14 days, and not at a fixed time point.

In the Philippines, newborns were selected when they were born at term with normal birth weight (BW), defined as between 2500 and 4500 g and with an age >24 hours, whereas in Mexico and Nepal, newborns with gestational age (GA) >35 weeks and BW >1750 g were included.

Newborns showing signs of other diseases than jaundice or newborns receiving advanced medical treatment were excluded for participation. Newborns that had received phototherapy were also excluded, as phototherapy can influence skin colour.[21]

Data collection

Background data on all participants were collected, including date and time of birth, age (in hours) at inclusion, GA in weeks at birth and BW. Skin type of mother, and father if present, was assessed with the Fitzpatrick skin scale.[22] Newborns were undressed and placed on an examination table for further assessment. For the study in the Philippines, the skin type of the infants was in addition assessed with the Neomar scale.[23] The visual degree of jaundice was assessed by the Kramer scale.[24] Assessments were performed by trained, designated research assistants with experience of paediatric nursing on each site (n=1 in Mexico, n=2 in Nepal and n=1 in Philippines). All research assistants received specific training on how to use the smartphone system as well as on how to register and code the data. Manuals for data collection were developed and provided to all study sites. If there were uncertainties or disagreements on classification, attending paediatricians or other medical doctors were consulted for final classification.

A colour calibration card was placed on the chest of the newborn, and an image set with a total of six digital images was automatically obtained, three with flash, and three without flash using a version of the Picterus Jaundice Pro app for clinical research (Mexico and Nepal beta version, Philippines V.2.2.0, Picterus AS), installed on Samsung Galaxy S7 smartphones.[18] In the app, the image sets were obtained, were given a unique ID and uploaded to servers. Images were stored for later analysis, and no bilirubin value was communicated back to the research assistants. This setup was implemented, so potential results of the digital image analysis would not influence further follow-up and treatment of the infants.

Laboratory analysis

A blood sample was obtained with venipuncture within 60 min before or after obtaining the digital images. Samples were protected from sunlight before analysis. Analysis for TSB was performed on hospital laboratories on each site. In Mexico, samples were processed on Abbot ARCHITECT ci 4100 (Abbott, Abbott Park, IL, USA), and in Philippines on Selectra ProXL (ELITechGroup, Puteaux, France), and TSB was analysed utilising the diazo method. In Nepal, TSB was analysed by a dry chemistry method using a VITROS 350 System (Ortho Clinical Diagnostics, New Jersey).

Data analysis

All image sets were analysed using the same version of the automated image analysing software installed on the servers, where image quality was checked, and colours

Table 1 Descriptive statistics of participants

Site	Number	Birth weight* (grams)	Gestational age (weeks)	Age at inclusion† (hours)	TSB (µmol/L)	Fitzpatrick	
						Mother‡	Father§
All	416	3030±440	38.6±1.4	59.2±62.5	148.6±66.1	3.7±0.5	3.7±0.5
Mexico	165	3088±470	38.4±1.4	74.1±89.4	140.0±85.6	3.7±0.5	3.7±0.5
Nepal	182	2946±446	38.7±1.5	48.8±20.9	157.1±49.6	3.3±0.5	3.6±0.5
Philippines	69	3105±302	38.6±1.0	51.0±50.0	146.7±45.6	3.2±0.6	4.2±0.6

*ANOVA F-ratio 5.7, p=0.004, Tukey post-hoc tests p<0.05 when Nepal compared with Mexico and Philippines.
†ANOVA F-ratio 8.1, p<0.001, Tukey post-hoc tests p<0.05 when Mexico compared with Nepal and Philippines.
‡ANOVA F-ratio 41.0, p<0.001, Tukey post-hoc tests p<0.05 when Mexico compared with Nepal and Philippines.
§ANOVA F-ratio 12.7, p<0.001, Tukey post-hoc tests p<0.05 when Philippines compared with Mexico and Nepal.
ANOVA, Analysis of variance; TSB, total serum bilirubin.

are calibrated, and then converted to bilirubin values for each image set.

Potential differences between the participants from each study site were evaluated by analysis of variance (ANOVA), and a level of 0.05 was considered significant. Correlations between image bilirubin values and TSB were evaluated using Pearson's r, both for all sites, as well as on each specific site. Bland-Altman plots were used to evaluate potential bias in the difference between image bilirubin values and TSB. We developed receiver operating characteristic (ROC) curves for identification of NHB defined as TSB >225 µmol/L. Youden index was determined for each site. Statistical analyses were performed using MedCalc Software V.20.115.

Patient and public involvement

Patients and/or the public were not involved in the design, or conduct, or reporting or dissemination plans of this research.

Ethics

Participation in this study was based on written informed assent from one of the parents. Assent forms were translated to relevant languages for each study site. The studies were all approved by the relevant ethical committees.

RESULTS

A total of 436 infants were recruited. Twenty infants were excluded from data analysis as they either did not fulfil the inclusion criteria (GA or BW), had no TSB result recorded or were missing the image set, leaving a total of 416 participants. The flow of recruitment is shown in figure 1. The lower number of included participants in the Philippines was due to strengthened measures related to the COVID pandemic, which limited possibilities for recruitment.

Descriptive statistics of the participants are shown in table 1.

The participants recruited in Nepal had a significantly lower BW when compared with those recruited in Mexico and in the Philippines, and the participants in Mexico had a higher age at inclusion, when compared with the two other sites.

Assessment of Fitzpatrick scale showed that the skin type of the mothers was significantly different between the sites, with a larger proportion of the mothers in Mexico with a higher Fitzpatrick score. Of all mothers, 94% had Fitzpatrick skin type score 3 or 4. For the skin type assessment of the fathers, the fathers in Philippines had a significantly higher Fitzpatrick skin type score, compared with Mexico and Nepal. The assessment of the participants in Philippines with the Neomar scale showed that of the 68 newborns, 78% were assessed to grade 2 and 19% to grade 3. One infant was graded as 1 and one as grade 4.

Table 2 shows the visual assessment of jaundice using Kramer scale. There was a significant difference between the groups, with a larger proportion of infants in Nepal with a higher Kramer scale as compared with the other sites.

Scatter plots of TSB versus Picterus bilirubin values for all the sites combined, as well as for each site individually, are presented in figure 2 with the relevant correlations presented by Pearson's r.

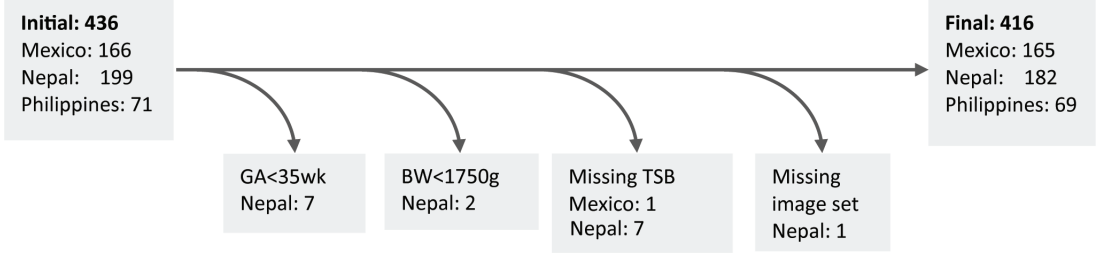

Figure 1 Study flow diagram. BW, birth weight; GA, gestational age; TSB, total serum bilirubin.

Table 2 Visual assessment of jaundice by Kramer scale

	Kramer scale					
Site	0	1	2	3	4	5
All	148 (36%)	71 (17%)	84 (20%)	52 (13%)	46 (11%)	15 (4%)
Mexico	78 (47%)	34 (21%)	34 (21%)	16 (10%)	3 (2%)	0 (0%)
Nepal*	49 (27%)	9 (5%)	33 (18%)	35 (19%)	41 (23%)	15 (8%)
Philippines	21 (30%)	28 (41%)	17 (25%)	1 (1%)	2 (3%)	0 (0%)

*ANOVA F-ratio 45.4, p<0.001, Tukey post-hoc tests p<0.05 compared with Mexico and Philippines.
ANOVA, Analysis of variance.

A positive significant correlation was found in all the analysis, with an overall correlation coefficient of 0.76. The subanalyses on each site demonstrated a significantly stronger correlation in Mexico compared with Nepal and the Philippines (p<0.001 and p=0.002). No difference was found between Nepal and the Philippines.

Bland-Altman plots for all sites combined and for each site are presented in figure 3. Overall, the Picterus values overestimated TSB levels with a mean difference of 9.5 µmol/L between Picterus and TSB. There was a significant difference between the sites, as Picterus values in Mexico underestimated TSB levels with a mean difference of 6.2 µmol/L, whereas in Nepal and Philippines, they were overestimated, with 21.5 µmol/L and 15.5 µmol/L, respectively.

Table 3 presents the mean differences, SD of the difference and 95% limits of agreements (LoA) between the two methods. The overall SD was 45.5 µmol/L, resulting in a 95% LoA of ±89.2 µmol/L. There was a tendency towards a lower SD for the difference for the Philippines, but this was not significant.

ROC curves for identification of newborns with TSB levels above 225 µmol/L are presented in figure 4. This level was chosen to get a relevant number of participants with a positive outcome, making the ROC analyses more robust. A total of 52 cases (13%) of the participants had levels higher than 225 µmol/L.

In table 4, the statistics on the ROC analyses are presented. Both overall and in each individual site, the area under the curve (AUC) was high. The cut-off levels according to the Youden index are different between the sites.

The AUC of the ROC curve for visual identification of hyperbilirubinemia defined as TSB >225 µmol/L by Kramer scale was 0.816 (95% CI 0.775 to 0.852) and was significantly lower than the AUC of Picterus values (p<0.0001). The cut-off level according to the Youden index for Kramer scale was 1, resulting in a sensitivity of 96.2% and specificity of 59.6%.

DISCUSSION

In this study, we show that bilirubin values determined with the Picterus system in three different populations with skin types of mainly Fitzpatrick III and IV are correlated to TSB levels. The results support that this smartphone technology shows potential to serve as a screening method for neonatal jaundice not only in newborns of

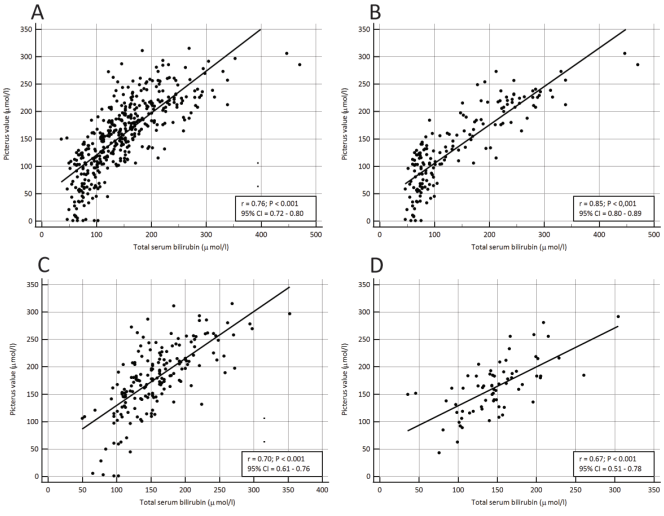

Figure 2 Scatter plots of total serum bilirubin (TSB) (µmol/L) vs Picterus values (µmol/L) with correlations. (A) shows all sites combined, (B) shows Mexico, (C), Nepal and (D) the Philippines.

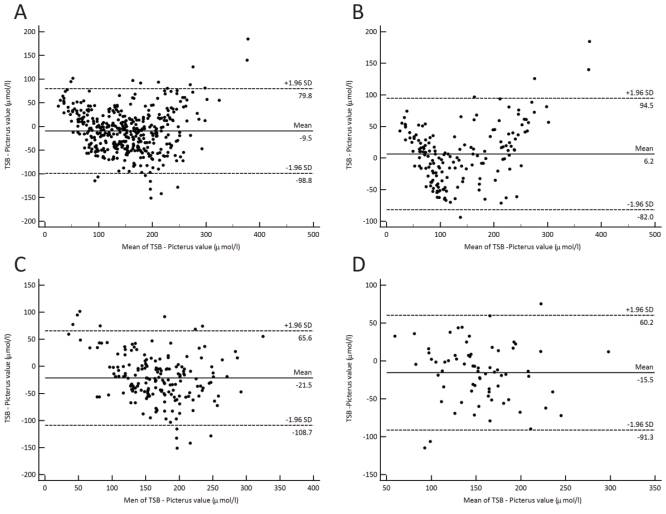

Figure 3 Bland-Altman plots of the difference between Picterus and TSB values. (A) shows all sites combined, (B) shows Mexico, (C) Nepal and (D) the Philippines. TSB, total serum bilirubin.

Table 3 Mean and 95% limits of agreement between TSB and Picterus values (µmol/L)

Site	Mean	SD	95% limits of agreement	
			Lower	Upper
All	−9.5	45.5	−98.8	79.8
Mexico*	6.2	45.0	−82.0	94.5
Nepal	−21.5	44.5	−108.7	65.6
Philippines	−15.5	38.6	−91.3	60.2

*ANOVA F-ratio 18.2, p<0.001, Tukey post-hoc tests p<0.05 when Mexico compared with Nepal and Philippines.
TSB, total serum bilirubin.

parents with fair skin but possibly also on newborns of parents with moderate dark skin types. The system showed a high sensitivity and specificity to detect infants with TSB levels exceeding 225 µmol/L, and the system was better than visual assessment performed by Kramer scale. The findings in this study are in line with previous testing of this system in a population of newborns with fair skin types[17] as well as other studies on smartphone-based assessment of jaundice.[25–27] Strengths of this study include that data were collected in hospital environments in three different countries, all categorised as LMICs. A total of 436 participants were recruited and only one infant had a missing image set, suggesting that the method of obtaining images is robust and can be feasible in a clinical setting. Further strengths are that results from the smartphone system were compared both to TSB as well as visual assessment of jaundice, and methods were compared with both correlation analyses, Bland Altman plot and ROC curves for identification of severe jaundice. There was a significant positive correlation between the Picterus bilirubin values and TSB on all study sites. A comparison between the three sites showed inconsistent findings, as Pearson correlation was significantly higher in Mexico compared with the two other sites, but we found no significant differences in the 95% LoA of the difference between TSB and Picterus values between the sites. However, in contrast to the correlation analysis, we found a trend towards narrower LoA in the Philippines. Correlation factors are sensitive to the range of the values as well as outliers in the data set.[28] The data collection in Mexico had a larger variation in TSB values as compared with the two other study sites and this could potentially explain this contradictory finding. The larger variation in TSB values could be due to differences in recruitment, as we in Mexico were able to recruit participants from the outpatient clinic where older newborns and newborns with suspected jaundice attended.

As correlation coefficients give no information on the actual difference between two methods, it is argued that Bland-Altman plots can provide more relevant insight for clinical decision-making.[29] The LoA of ±89.2 µmol/L found in this study are wider than reported in previously testing of the Picterus system as well as by other means of transcutaneous bilirubin estimations and are wider than recommended safety margins for the use of TcB in neonatal jaundice.[1 17 30] This could have clinical implications, with unwanted effects on clinical decision-making. A limitation of our study is that we did not define acceptable LoA for clinical use *a priori* to performing the study as suggested by Bland and Altman and should be conducted in future studies.

We performed ROC analyses of the system and found a sensitivity of 86.5% and a specificity of 83.5%. A limitation of this study is that due to the low number of infants with high bilirubin levels, we had to set the definition of positive case for the ROC analyses to TSB >225 µmol/L. This level is lower than we have in our previous studies as well as in other studies of detection of neonatal jaundice, which would limit direct comparison of our present results. The cut-off levels related to the Youden index varied between the sites, but these comparisons should be taken with precaution as the numbers of positive cases were few.

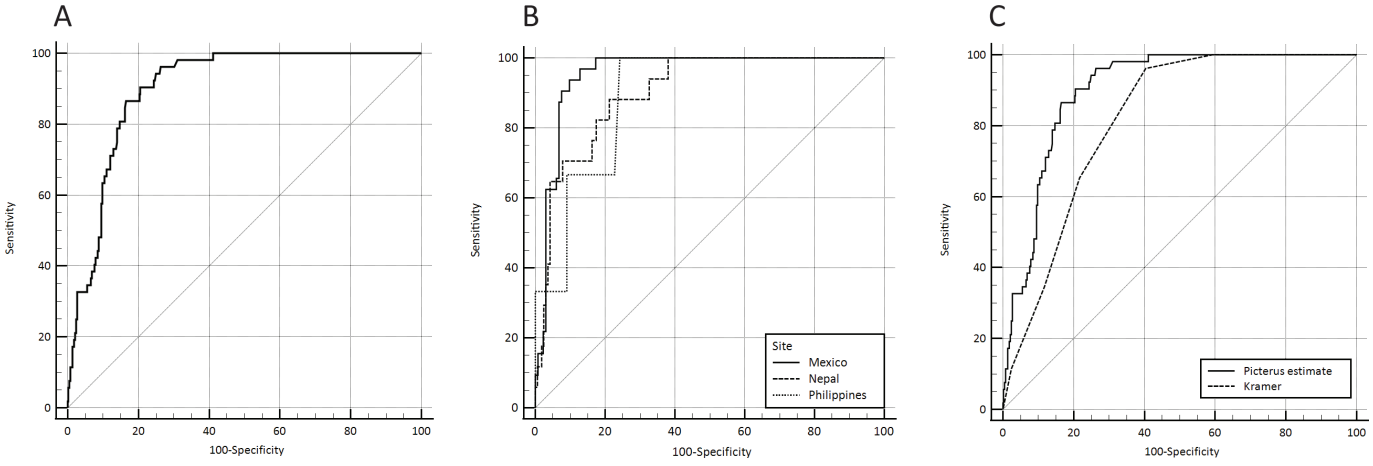

Figure 4 ROC curves for the performance of the Picterus system to identify infants with TSB>225 µmol/L. (A) shows all sites combined. (B) shows the different sites and (C) a comparison of Picterus values vs Kramer scale for all sites. ROC, receiver operating characteristic; TSB, total serum bilirubin.

Table 4 ROC curve by study site

Site	Positive cases	Negative cases	AUC	95% CI	Youden index cut-off	Sensitivity	Specificity
All	52 (13%)	364	0.900	0.868 to 0.927	>207.7	86.5	83.5
Mexico	32 (19%)	133	0.954	0.910 to 0.981	>178.8	96.9	87.2
Nepal	17 (9%)	165	0.903	0.850 to 0.942	>211.2	88.2	78.8
Philippines	3 (4%)	66	0.891	0.793 to 0.954	>184	100	75.8

ROC, receiver operating characteristic.

Harmful levels of bilirubin are further higher than the level set for our ROC analyses, and it can hence be argued that an assessment of the ability of the system to detect newborns needing treatment is more relevant for clinical practice. A major limitation of this study is that there were few infants needing phototherapy. Further studies with adequate power to evaluate this performance need to be conducted before the system is taken into use in patients.

Another limitation of our study is that we were only able to include participants who had a blood sample performed for a clinical reason and could not include otherwise healthy newborns. Sensitivity and specificity analyses are highly dependent on prevalence of the disease in the targeted population, and testing of a screening device should be carried out in a sample representative of the whole population. In our case, this would imply testing otherwise healthy newborns without signs of jaundice. Drawing blood samples without a clinical indication would not be approved from an ethical point of view, and it was not possible to execute. We aimed to limit this weakness by including newborns that had blood samples obtained for other reasons than suspected jaundice, but this could still influence the assessment of false positives.

A similar limitation regarding the participants in our study is that we did not recruit newborns born before 35 weeks of gestation, and only 18 newborns were born before gestational week 37. It is known that preterm born infants are at a higher risk of developing jaundice, and the system should be tested in this population.

Furthermore, follow-up of newborn infants after receiving phototherapy is an important part of jaundice care. As phototherapy affects TcB readings, a limitation of our study is that we did not include patients that had received phototherapy.[31] This should be conducted in future studies.

Bland-Altman plots showed that the values from the Picterus system underestimated TSB levels in the Mexican newborns while they overestimated TSB levels in both Nepal and the Philippines. The systematic shift of the Picterus bilirubin values between the study sites could origin from multiple factors. Even if the vast majority of the participants had parents with comparable Fitzpatrick skin types of III and IV, it is known that skin composition and melanin distribution vary between different ethnicities and could influence the analysis.[32] It is further known that newborns of Asian ethnicity have naturally higher bilirubin levels compared with other ethnicities.[33–35] In our study, the average TSB levels in Nepal and the Philippines were higher, but this does not explain the overestimation and underestimation of TSB levels, as this shift was seen both in high and low TSB values.

Furthermore, large discrepancies of TSB values have been observed between laboratories.[36 37] The magnitude of these reported between-laboratory differences is larger than the differences observed in our study. Thus, between-laboratory differences of TSB values may potentially explain our findings.[38] It is a limitation of the study that standardised samples to evaluate this possibility were not included in the study design.

This study was performed to illustrate testing of a smartphone app on newborns in other countries, than where it was developed and initially tested.[17] The aim was to evaluate if it was possible to use the system in various settings, including in LMIC, and on newborns with various skin types. To involve three countries, using different languages, diverse hospitals, to receive ethical clearance, train health providers, took unfortunately longer time than we hoped.

Other means of TcB were not available at the hospitals where the studies were performed and limits our possibilities to perform comparisons between the Picterus system and other TcBs. Visual assessment of neonatal jaundice is known to be challenging, with varying sensitivity and specificity, as well as large interobserver difference.[8–10] Since Picterus is an objective measurement, it is reason to believe that the result will be less dependent of the training and experience of the observer. However, we did not assess this. Further studies on usability, repeatability and reproducibility should be conducted to evaluate potential influence on different circumstances and different users on the performance.

One of the aims behind the development of this system was to develop a system that was substantially less expensive than currently available systems for transcutaneous bilirubinometry. All smartphone-based technologies in healthcare share an advantage in running on already existing platforms and devices. This could enable a rapid scale-up and low maintenance costs, as the phones can be easily replaced if damaged. The method tested in this study requires a physical colour calibration card, which sets some limitations on distribution, but these can be manufactured at a substantially lower cost compared with

TcB. The Picterus app is currently not available in either of the countries where the studies in this paper took place. In other markets, the system can be purchased at €8 for a calibration card, limited to a single newborn, and a cost per scan of €8.[39] How these costs would compare to other TcB devices was not part of this study.

We only performed direct assessment of the skin type of the newborns with the Neomar scale in the Philippines.[23] For the studies in Mexico and Nepal, the skin type assessments were performed on the parents of the infants, as the Fitzpatrick skin scale is intended for use on adults, and not valid for infants. In future studies of the Picterus system, the Neomar scale should be used.

CONCLUSION

We conclude that smartphone technologies show potential to support in jaundice assessment in populations with moderate dark skin types in LMICs. Further studies are needed before the system can be taken into use in clinical practice, including studies on newborns with high bilirubin levels, studies on reproducibility and reliability as well as the potential clinical implications of a system with the limits of agreement found in this study. Finally, the system should be tested on all newborns, including those with the highest melanin levels.

Author affiliations
[1]Faculty of Medicine and Health Sciences, Department of Public Health and Nursing, Norwegian University of Science and Technology, Trondheim, Norway
[2]Department of Pediatrics, St Olav's Hospital HF, Trondheim, Norway
[3]Picterus AS, Trondheim, Norway
[4]Department of Pediatrics, Mexican Institute of Social Security, Mexico City, Mexico
[5]Maternal and Infant Hospital in Irapuato, Guanajuato, Mexico
[6]Department of Clinical and Molecular Medicine, Norwegian University of Science and Technology, Trondheim, Norway
[7]Department of Obstetrics and Gynecology, Kathmandu University School of Medical Sciences, Kathmandu, Nepal
[8]Department of Women's and Children's Health, Faculty of Medicine, Uppsala Universitet, Uppsala, Sweden

Acknowledgements We thank Marivic S. Deluna for help with arranging data collection in the Philippines, all nurses and midwives who assisted with the data collection in Nepal and Mexico, and all parents who allowed their infants to take part in these studies.

Contributors AA, GV, HB and ED. Data collection: AA, GJD, MRB, JT and SS. Data management and analysis: AA, GJ-D, LMG and GV. Visualisation: AA. All authors contributed to the interpretation of the results and reviewed both the original and revised manuscripts. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work. AA is the guarantor for this study.

Funding This study was funded from the Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, St. Olav's Hospital, Trondheim University Hospital, grant number 17/10533-131, and Picterus AS.

Competing interests AA and GV are founders and previous shareholders and employees at Picterus AS, the company currently commercialising the Picterus Jaundice app. GJD and LMG are both currently shareholders and employees at Picterus AS. The other authors have no conflicts of interest to disclose.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Consent obtained from parent(s)/guardian(s).

Ethics approval This study involves human participants and was approved by Mexico: Regional Ethics and Research Committee in Hospital Materno Infantil of León, Guanajuato. Number: 0294/2018. Nepal: Nepal Health Research Council (NHRC) and the Kathmandu University School of Medical Sciences—Institutional Review Committee. No: 571/2018. The Philippines: Research Ethics Committee of Governor Celestino Gallares Memorial Hospital. Number: 2021-20. Norway: Regional Committee for Medical and Health Research Ethics (REC) in Norway (Number: 2018/1001, 2018/1608 and 2021/322016). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; internally peer-reviewed.

Data availability statement Data are available upon reasonable request. Data are available upon reasonable request. Newborn images might contain identifiable information and are not publicly available.

ORCID iDs
Anders Aune http://orcid.org/0000-0001-6757-7398
Gabriela Jímenez-Díaz http://orcid.org/0000-0002-5944-7427

REFERENCES

1 Amos RC, Jacob H, Leith W. Jaundice in newborn babies under 28 days: NICE guideline 2016 (CG98). Archives of disease in childhood. *Education and Practice* 2017;102:207–9.
2 Hansen T. The epidemiology of neonatal jaundice. *Pediatric Medicine* 2021;5:18.
3 Maisels MJ. Managing the jaundiced newborn: a persistent challenge. *CMAJ* 2015;187:335–43.
4 Alkén J, Håkansson S, Ekéus C, *et al*. Rates of Extreme Neonatal Hyperbilirubinemia and Kernicterus in Children and Adherence to National Guidelines for Screening, Diagnosis, and Treatment in Sweden. *JAMA Netw Open* 2019;2:e190858.
5 Olusanya BO, Teeple S, Kassebaum NJ. The Contribution of Neonatal Jaundice to Global Child Mortality: Findings From the GBD 2016 Study. *Pediatrics* 2018;141:e20171471.
6 Bhutani VK, Zipursky A, Blencowe H, *et al*. Neonatal hyperbilirubinemia and Rhesus disease of the newborn: incidence and impairment estimates for 2010 at regional and global levels. *Pediatr Res* 2013;74 Suppl 1:86–100.
7 Olusanya BO, Ogunlesi TA, Slusher TM. Why is kernicterus still a major cause of death and disability in low-income and middle-income countries? *Arch Dis Child* 2014;99:1117–21.
8 Moyer VA, Ahn C, Sneed S. Accuracy of clinical judgment in neonatal jaundice. *Arch Pediatr Adolesc Med* 2000;154:391–4.
9 Darmstadt GL, Hamer DH, Carlin JB, *et al*. Validation of visual estimation of neonatal jaundice in low-income and middle-income countries: a multicentre observational cohort study. *BMJ Open* 2021;11:e048145.
10 Szabo P, Wolf M, Bucher HU, *et al*. Detection of hyperbilirubinaemia in jaundiced full-term neonates by eye or by bilirubinometer? *Eur J Pediatr* 2004;163:722–7.
11 Oldak D, García G, Gonzalez EE, *et al*. Reproducibility of BiliCare Transcutaneus Bilirrubin Meter in Mexican Newborns. *Int J Pediatr* 2019;2019:3812152.
12 Villanueva-Uy MET, G Uy H, Amarillo MLE. Applicability of the hour of life approach in hyperbilirubinemia among Filipino term infants. *Front Pediatr* 2022;10:990919.
13 Kumar V, Kahlon PS, Singh P, *et al*. Utility of Transcutaneous Bilirubinometer in Tertiary Care Hospital. *J Nepal Paedtr Soc* 2017;37:72–8.
14 Slusher TM, Zipursky A, Bhutani VK. A global need for affordable neonatal jaundice technologies. *Semin Perinatol* 2011;35:185–91.
15 Istepanian RSH, Laxminarayan S, Pattichis CS. M-Health: emerging mobile health systems. Boston, MA: Springer Science+Business Media, Inc, 2006.
16 Vartdal G, Randeberg L, Aune A, *et al*. Inventors, picterus as, assignee. Image based bilirubin determination, EP3393353B1. European Patent Office patent; 2016. Available: https://patents.google.com/patent/EP3393353B1/en:Method
17 Aune A, Vartdal G, Bergseng H, *et al*. Bilirubin estimates from smartphone images of newborn infants' skin correlated highly to serum bilirubin levels. *Acta Paediatr* 2020;109:2532–8.

18 Aune A, Vartdal G, Jimenez Diaz G, *et al*. Iterative Development, Validation, and Certification of a Smartphone System to Assess Neonatal Jaundice: Development and Usability Study. *JMIR Pediatr Parent* 2023;6:e40463.
19 Varughese P, Krishnan L, Ravichandran K. Does color really matter? Reliability of transcutaneous bilirubinometry in different skin-colored babies. *Indian J Paediatr Dermatol* 2018;19:315.
20 Olusanya BO, Imosemi DO, Emokpae AA. Differences Between Transcutaneous and Serum Bilirubin Measurements in Black African Neonates. *Pediatrics* 2016;138:e20160907.
21 Nagar G, Vandermeer B, Campbell S, *et al*. Effect of Phototherapy on the Reliability of Transcutaneous Bilirubin Devices in Term and Near-Term Infants: A Systematic Review and Meta-Analysis. *Neonatology* 2016;109:203–12.
22 Fitzpatrick TB. The validity and practicality of sun-reactive skin types I through VI. *Arch Dermatol* 1988;124:869–71.
23 Maya-Enero S, Candel-Pau J, Garcia-Garcia J, *et al*. Validation of a neonatal skin color scale. *Eur J Pediatr* 2020;179:1403–11.
24 Kramer LI. Advancement of dermal icterus in the jaundiced newborn. *Am J Dis Child* 1969;118:454–8.
25 Taylor JA, Stout JW, de Greef L, *et al*. Use of a Smartphone App to Assess Neonatal Jaundice. *Pediatrics* 2017;140:e20170312.
26 Aydın M, Hardalaç F, Ural B, *et al*. Neonatal Jaundice Detection System. *J Med Syst* 2016;40:166.
27 Rong ZH, Luo F, Ma LY, *et al*. Evaluation of an automatic image-based screening technique for neonatal hyperbilirubinemia. *Zhonghua Er Ke Za Zhi* 2016;54:597–600.
28 Janse RJ, Hoekstra T, Jager KJ, *et al*. Conducting correlation analysis: important limitations and pitfalls. *Clin Kidney J* 2021;14:2332–7.
29 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–10.
30 Bratlid D, Nakstad B, Hansen TWR. National guidelines for treatment of jaundice in the newborn. *Acta Paediatr* 2011;100:499–505.
31 Joy C, Collyer TA, McMahon K. Accuracy of transcutaneous bilirubinometry in term infants after phototherapy: a prospective observational study. *J Matern Fetal Neonatal Med* 2024;37:2295808.
32 Rawlings AV. Ethnic skin types: are there differences in skin structure and function? *Int J Cosmet Sci* 2006;28:79–93.
33 Bentz MG, Carmona N, Bhagwat MM, *et al*. Beyond “Asian”: Specific East and Southeast Asian Races or Ethnicities Associated With Jaundice Readmission. *Hosp Pediatr* 2018;8:269–73.
34 Setia S, Villaveces A, Dhillon P, *et al*. Neonatal jaundice in Asian, white, and mixed-race infants. *Arch Pediatr Adolesc Med* 2002;156:276–9.
35 De Luca D, Jackson GL, Tridente A, *et al*. Transcutaneous bilirubin nomograms: a systematic review of population differences and analysis of bilirubin kinetics. *Arch Pediatr Adolesc Med* 2009;163:1054–9.
36 van Imhoff DE, Dijk PH, Weykamp CW, *et al*. Measurements of neonatal bilirubin and albumin concentrations: a need for improvement and quality control. *Eur J Pediatr* 2011;170:977–82.
37 Thomas DH, Warner JV, Jones GRD, *et al*. Total bilirubin assay differences may cause inconsistent treatment decisions in neonatal hyperbilirubinaemia. *Clin Chem Lab Med* 2022;60:1736–44.
38 Berska J, Bugajska J, Sztefko K. Newborns bilirubin concentration determined by different methods in relation to hematocrit and albumin level. *J Med Biochem* 2020;39:171–7.
39 Picterus AS. Order picterus Jaundice pro here. 2025. Available: https://picterus.com/hcp/order-here/