Received: 3 December 2019 | Revised: 25 March 2020 | Accepted: 27 March 2020

DOI: 10.1111/apa.15287

REGULAR ARTICLE

ACTA PÆDIATRICA NURTURING THE CHILD WILEY

Bilirubin estimates from smartphone images of newborn infants' skin correlated highly to serum bilirubin levels

Anders Aune[1] | Gunnar Vartdal[2] | Håkon Bergseng[3,4] | Lise Lyngsnes Randeberg[5] | Elisabeth Darj[1,6,7]

[1]Department of Public Health and Nursing, Norwegian University of Science and Technology, Trondheim, Norway

[2]Picterus AS, Trondheim, Norway

[3]Department of Pediatrics, St. Olav University Hospital, Trondheim, Norway

[4]Department of Laboratory Medicine, Children's and Women's Health, Norwegian University of Science and Technology, Trondheim, Norway

[5]Department of Electronic Systems, Norwegian University of Science and Technology, Trondheim, Norway

[6]Department of Obstetrics and Gynecology, St. Olav University Hospital, Trondheim, Norway

[7]Department of Women' and Children's Health, Uppsala University, Uppsala, Sweden

Correspondence
Anders Aune, NTNU, Faculty for Medicine and Health Sciences, Department of Public Health and Nursing, Post box 8905, N-7491 Trondheim, Norway.
Email: a.aune@ntnu.no

Funding information
This study was funded as a part of a PhD grant from the Norwegian University of Science and Technology to the first author.

Abstract

Aim: Neonatal jaundice is an important cause of morbidity and mortality, and identifying the condition remains a challenge. This study evaluated a novel method of estimating bilirubin levels from colour-calibrated smartphone images.

Methods: A cross-sectional prospective study was undertaken at two hospitals in Norway from February 2017 to March 2019, with standardised illumination at one hospital and non-standardised illumination at the other hospital. Healthy term-born infants with a normal birthweight were recruited up to 15 days of age. The main outcome measures were bilirubin estimates from digital images, plus total bilirubin in serum (TSB) and transcutaneous bilirubin (TcB).

Results: Bilirubin estimates were performed for 302 newborn infants, and 76 had severe jaundice. The correlation between the smartphone estimates and TSB was measured by Pearson's *r* and was .84 for the whole sample. The correlation between the image estimates and TcB was 0.81. There were no significant differences between the hospitals. Sensitivity was 100%, and specificity was 69% for identifying severe jaundice of more than 250 μmol/L.

Conclusion: A smartphone-based tool that estimated bilirubin levels from digital images identified severe jaundice with high sensitivity and could provide a screening tool for neonatal jaundice.

KEYWORDS

digital images, hyperbilirubinaemia, mobile health, neonatal jaundice, smartphone

1 | BACKGROUND

Neonatal jaundice is a common, mainly harmless and self-limiting condition that affects 60%-80% of newborn infants.[1] However, the condition is potentially dangerous as bilirubin can accumulate in the basal ganglia of the brain and cause permanent brain damage.[2] Such brain damage, known as kernicterus, can manifest as cerebral palsy, deafness, language difficulties or can be fatal in the worst cases. It is estimated that more than 100 000 infants worldwide die of jaundice every year and a large number survive with severe disabilities.[3] Three-quarters of the deaths are estimated to occur in the poorest regions of the world: sub-Saharan Africa and south Asia.[3] A major

Abbreviations: mHealth, mobile health; ROC, receiver operating characteristic; TcB, transcutaneous bilirubin; TSB, total serum bilirubin.

challenge in reducing the burden of neonatal jaundice is to identify children at risk at an early stage, so that effective treatment can be given.[4,5]

The gold standard for diagnosing neonatal jaundice is measuring total serum bilirubin (TSB), and an alternative, and well-established method, is transcutaneous bilirubinometry (TcB).[6] TcB has also been proven to work well in infants with dark skin.[7] However, transcutaneous devices and laboratory equipment are both expensive and blood analyses require both skilled personnel and puncturing the infant's skin.[5] Visual detection of jaundice is not recommended, as it has been associated with a high risk of missing severe cases.[8,9]

There has been an increasing focus on researching how mobile phone technology can be used in health care. Mobile phone health (mHealth) is of interest, because even areas with scarce economic resources can use it.[10] Studies published in the last few years have described novel ways of assessing neonatal jaundice by using digital images or smartphones, by adopting slightly different approaches.[11-16]

Based on previous research on the bio-optical properties of newborn skin, we have developed a physics-based system that estimates bilirubin from digital images.[17] The purpose of this study was to evaluate the accuracy of this new method.

Key notes

- We studied 302 full-term, normal weight infants at up to 15 days of age to see if smartphone images calibrated with a colour card could detect neonatal jaundice as well as laboratory tests.
- This showed that bilirubin estimates from smartphone images were highly correlated to bilirubin levels in serum and standard transcutaneous measurements.
- The method showed high sensitivity (100%) and specificity (69%) and could provide a screening tool for severe jaundice.

2 | PATIENTS AND METHODS

2.1 | Technology

Researchers from Norwegian University of Technology and Science and Picterus AS, a spin-off company from the university, have developed a novel way of estimating bilirubin levels in newborn infants using colour analysis of digital images combined with physics-based modelling of light transport in skin. Based on previous research on the bio-optics of newborn skin,[18] we used a mathematical model that used diffusion theory to create a library of simulated reflectance spectra of newborn skin. In addition to bilirubin, multiple factors influence skin reflectance, such as skin thickness, haemoglobin and melanin levels.[19] These factors were built into the model, and the model was adapted so that it could be used with a smartphone camera, which only uses three colours: red, green and blue.

Changes in illumination influence how colours appear in an image, and a way to correct for these variances is to include an object with known properties in the image. For this purpose, we created a colour calibration card that was included in each photograph. The colours on this card were chosen to reflect light in the same way as the skin of newborn infants, and the card was printed using a method known as spectral printing.[20] This technique requires a printer that can print with seven inks but can be printed on standard paper with standard inks. We printed the cards using a HP DesignJet Z7300 printer on HP Premium Matte Photo paper (HP Inc). After printing, each card was measured with a colour spectrometer and the readings were stored and used for later calibration of the images. Figure 1 shows a suggested final overview of the system: Digital images are taken with a smartphone, colour calibrated using the colours on the calibration card, and then compared with a large database of colour and bilirubin pairs. Cloud storage could be used to store this database, or the database could be downloaded to the phone and then run offline.

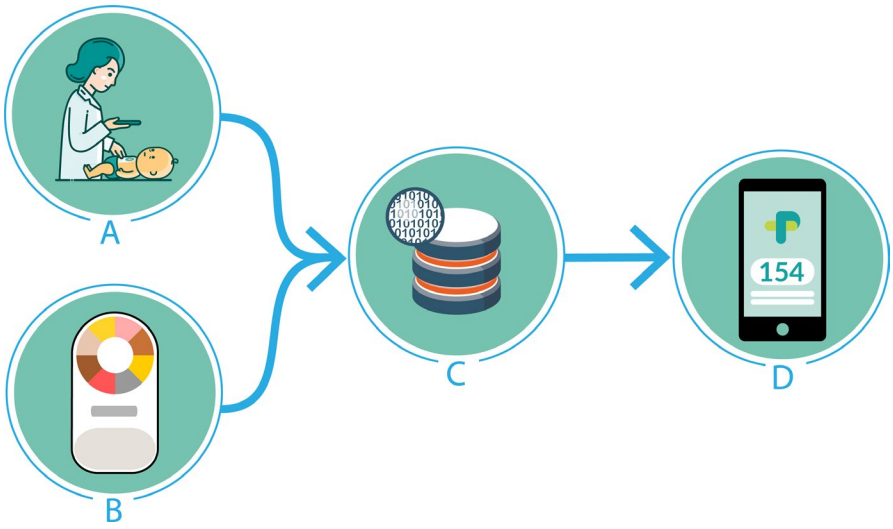

FIGURE 1 Suggested workflow of final application of the physics-based system to estimate bilirubin from digital images: (A) a photograph is taken of the newborn, (B) colours from the calibration card are used to calibrate the image, (C) calibrated colours are compared with a large database of colour and bilirubin pairs, (D) the matching bilirubin estimate is presented to the relevant user

2.2 | Cross-sectional study

We conducted a cross-sectional prospective study at two Norwegian hospitals: St Olav Hospital in Trondheim and Akershus University Hospital in Lørenskog. Both hospitals are university teaching hospitals, with 4000-5000 deliveries a year and neonatal intensive care units. The study took place at St Olav from February to August 2017, and at Akershus from February 2018 to March 2019.

2.3 | Study participants

Infants were eligible for inclusion if they were from 1 to 15 days of age, born at term at 37 weeks of gestation or more and with a normal birth weight of at least 2500 g. Newborn infants with signs of diseases other than jaundice, or who received advanced medical treatment, were not included. Infants receiving phototherapy were also excluded, as the treatment might have affected their TcB.[21]

We recruited newborn infants from the maternity wards or from breastfeeding outpatient clinics. In order to recruit infants with both high and low levels of bilirubin, we included both infants with clinically suspected jaundice, as well as a group of healthy infants that had their bilirubin level measured when newborn screening was performed. At Akershus, we also included infants that attended the outpatient clinic for suspected jaundice, but whose TcB levels did not indicate to obtain TSB levels.

2.4 | Data collection

We collected background data on each newborn infant, including their birth weight, gestational age and age at inclusion. We also classified the ethnicity of the parents, based on their own responses into: Caucasian, Middle Eastern, Asian, African, and other or unknown. When the infants were lying on an examination table, we placed the colour calibration card over the sternum and took photographs using a smartphone application designed for the purpose. All images were captured with Samsung Galaxy S7 smartphones (Samsung Electronics Co. Ltd). Images were uploaded to an external server for later analysis.

At St Olav, the images were obtained under standardised light conditions with illumination from two lamps with 46-watt halogen light bulbs, placed 60 cm on each side of the infant. Four images were obtained, three with a flash from three predefined distances—20, 30 and 40 cm—and one without a flash from 40 cm. At Akershus, the light conditions were not standardised, and images were obtained under ambient light conditions. We took six images, three with a flash and three without a flash, all from a distance of 30 cm.

The application used for data collection included software for recognising the calibration card. This was to ensure that images were taken at the same distance and angle on each infant. Based on initial experience from St Olav, this software was improved before data collection at Akershus.

The time from drawing blood sample to capturing images was set to maximum of 60 minutes. We obtained 600 µL of blood by heel prick or venipuncture, and this was then stored in light protected containers. The blood samples were analysed at St Olav by vanadate oxidation method using Siemens Advia Chemistry XPT (Siemens Healthcare GmbH). At Akershus they were analysed with a slide-adapted colorimetric method using Vitros 5.1 FS (Ortho Clinical Diagnostics Inc). At the same time as images were captured, we measured transcutaneous bilirubin with a standard device, either a Dräger JM-103 or Dräger JM-105 (Dräger Medical GmbH). The measurements were performed by trained personnel following hospital routines and in accordance with the manufacturers' instructions.

A bilirubin estimate was performed for each image, as described. The final estimate used in analysis was the average of the bilirubin estimates from the images taken with a flash and the images taken without a flash.

2.5 | Statistical analyses

The correlations between the bilirubin estimates from the images and the TSB and TcB results were evaluated using Pearson's *r*. Bland-Altman plots were created to evaluate potential bias in the difference between estimates and serum levels, as well as to evaluate confidence intervals for the difference. We drew receiver operating characteristic (ROC) curves to evaluate the sensitivity and specificity for the bilirubin estimates at different cut-off levels. A positive outcome for the ROC analysis was set at a TSB level of 250 µmol/L or more. This cut-off was chosen as several guidelines for neonatal jaundice have recommended obtaining TSB if TcB levels were ≥250 µmol/L.[1,22] Analyses were performed using IBM SPSS Statistics Version 25 (IBM Corp), and MedCalc Version 19.0.3 (MedCalc Software Ltd).

3 | RESULTS

The study initially comprised 342 newborn infants: 181 from St Olav and 161 from Akershus. In the initial sample, 76 infants had severe jaundice, defined as TSB level above 250 µmol/L, where 52 were recruited from St Olav and 24 from Akershus. TSB was missing in 86 participants, 10 participants due to technical problems when obtaining the sample, and for 76 participants there was no clinical indication to obtain TSB. A flow chart of participant recruitment is shown in Figure 2.

It was not possible to perform a bilirubin estimate for 36 infants as the image sets were incomplete or had technical errors. The technical errors included parts of the infant's clothes covering one or more of the calibration colours, the operator's fingers covering the lens, images being completely out of focus and problems relating to the card-recognition software that was used at St Olav. Furthermore, four infants were excluded because they were too old or had been miscoded. A separate analysis revealed that the infants who had

FIGURE 2 Recruitment of participants

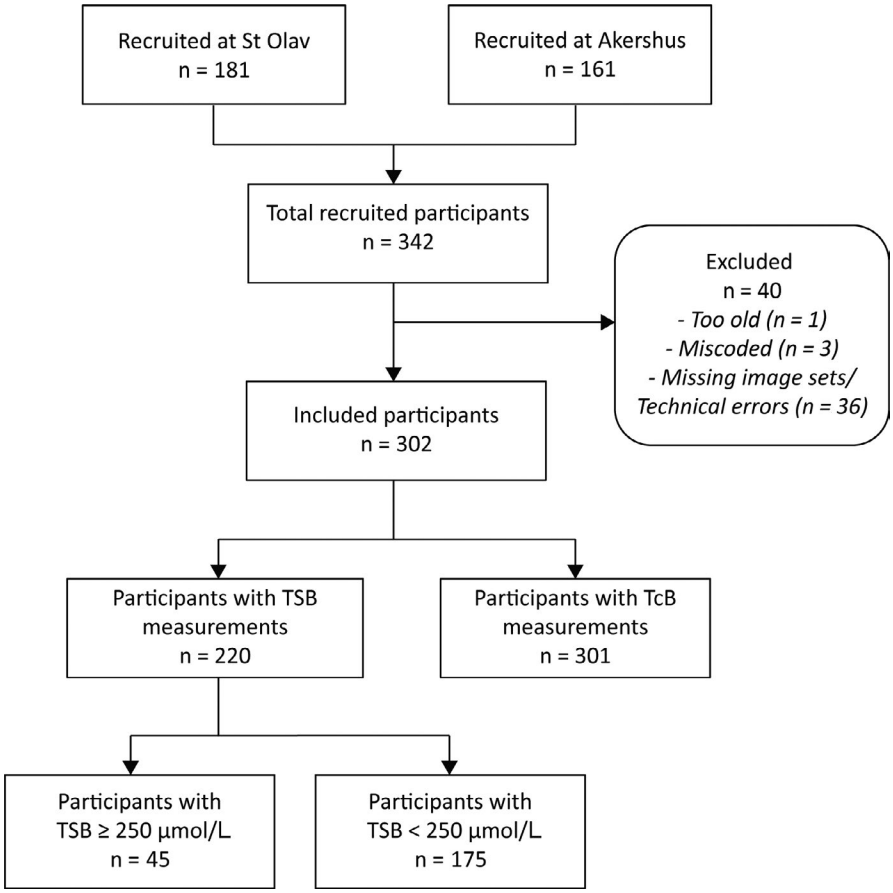

TABLE 1 Descriptive statistics of participants

	Total (n = 302)	St Olav (n = 144)	Akershus (n = 158)	*P*-value[a]
Birth weight (grams)	3530 ± 485	3552 ± 466	3511 ± 502	.5
Gestational age (weeks)	39.4 ± 1.3	39.3 ± 1.2	39.5 ± 1.5	.3
Age at inclusion (hours)	76 ± 43	66 ± 34	85 ± 48	<.001
Non-Caucasian	59 (20%)	13 (9%)	46 (29%)	<.001
TSB (µmol/L)	201 ± 75 (n = 220)	198 ± 70 (n = 136)	207 ± 83 (n = 84)	.4
TcB (µmol/L)	185 ± 70 (n = 301)	180 ± 62 (n = 144)	192 ± 76 (n = 157)	.06
Image estimates (µmol/L)	203 ± 73 (n = 302)	186 ± 65 (n = 144)	219 ± 76 (n = 158)	<.001
Severe jaundice (TSB > 250 µmol/L)	76	52	24	.8

[a]Chi-square for categorical variables and Student's *t* test for continuous variables.

to be excluded were not significantly different from the rest of the study sample. This left us with 302 newborn infants with complete image sets that could be used to perform bilirubin. Of these infants, 220 had a blood sample taken and 301 had their TcB tested. As the bilirubin estimates from the digital images are chosen from the database of skin simulation, we used a sample of 34 image sets from St Olav to optimise how the best matching estimate was chosen. The 34 infants to whom the images related were not included in the later analysis.

Descriptive statistics of the participants are presented in Table 1.

The TSB values ranged from 15 to 460 µmol/L, and the bilirubin estimates from the smartphone application ranged from 8 to 374 µmol/L. The newborn infants at Akershus were significantly older at inclusion, and the proportion of non-Caucasian participants

was larger compared to those at St Olav. Bilirubin estimates from the images taken at Akershus were also significantly higher than those taken at St Olav. Correlations between the bilirubin estimates from the images and the TSB and TcB levels are presented in Table 2.

The correlation between image estimates and TSB was significantly lower than the correlation between TcB and TSB. There was no significant difference in correlation between the two hospitals, and the correlations did not change when we adjusted for the study site in a partial correlation analysis. The correlation in the Caucasian subgroup was higher than the non-Caucasian group. We created ROC curves to evaluate the sensitivity and specificity of the bilirubin estimates. A positive outcome in the ROC analysis was defined as a TSB ≥ 250 μmol/L, and 45 infants had a TSB above this cut-off. A scatter plot, Bland-Altman plot and ROC curve are presented in Figure 3.

The mean difference between the image estimates and the serum levels was −0.2 μmol/L, indicating no systematic overestimation or underestimation of bilirubin levels. The standard error of the estimate was 41 μmol/L. We did not find any significant correlation between the standard error of the estimate and TSB levels. The Bland-Altman plots showed no sign of systematic overestimation or underestimation, or any trend in error related to TSB levels. The ROC curve showed an area under the curve of 0.925. Youden's index was 208 μmol/L and gave a 100% sensitivity and a 69% specificity. Setting a cut-off at 225 μmol/L decreased the sensitivity to 87% and increased the specificity to 77%.

4 | DISCUSSION

This study evaluated a new method of estimating bilirubin levels in newborn infants based on colour analysis of smartphone images of newborn skin. We recruited newborn infants with a wide range of bilirubin levels and found that bilirubin estimates from the smartphone application were highly correlated to bilirubin serum levels with Pearson's *r* of .84. Correlation was significantly higher in the Caucasian subgroup. However, this interpretation should be treated with caution as the number of non-Caucasian infants was only 20%. The ROC analysis showed that the screening tool detected severe jaundice with high sensitivity, while maintaining high specificity.

A method for bilirubin screening in newborn infants needs to be robust to changes in ambient light. Therefore, we tested the application in both standardised illumination conditions as well as in a real-life hospital setting with random illumination. We found similar correlations, which shows that the calibration technique corrected for changes in ambient light. A potential arena for the use of a screening tool for neonatal jaundice is in remote and out-of-hospital settings. Illumination in hospitals is often in the form of fluorescent light or LED lights, whereas in a home environment the type of illumination typically varies. Hence, it is important that the tool gives reliable estimates under varying illumination.

An advantage of the method described in this paper is that it uses standard features that are incorporated into even low-end smartphones. The calibration cards are printed using a method that involves special measurements and calibration of the printer, but they are printed on standard paper and with standard ink. The cost for producing one single-use card was estimated to be 10 American cents, and the cost of the working hours related to production and measurement was estimated to be 70 cents, giving a total estimated cost of 80 cents per card. Cost for one laminated card that can be sanitised and used multiple times was estimated at 1 USD. Even if distribution costs are added, the system can be implemented at an affordable cost, also in low-income settings.

Other studies have used attachments to mobile phones. Munkholm et al[13] attached a dermatoscope, and Sufian et al[23] used custom-developed plastic housing with a built-in light. A group of researchers in the UK described a procedure without attachments but used the mobile phone screen to control illumination and the front camera to take images of the sclera.[11] This approach could lower the barrier for implementation but is dependent upon the smartphone having the required features.

In our study, we used a physics-based simulation model to estimate bilirubin levels from smartphone images. Previous studies have based bilirubin estimates on machine-learning techniques and found high correlations to TSB.[15] In machine learning, the computer is trained to recognise predefined classes of data according to given

TABLE 2 Correlations between image estimates, TcB and TSB measurements

	Image estimates vs TSB		Image estimates vs TcB		TcB vs TSB	
Group	Pearson's *r*	95% CI	Pearson's *r*	95% CI	Pearson's *r*	95% CI
Overall	.84 (n = 185)[a]	0.79-0.88	.81 (n = 266)	0.76-0.85	.91 (n = 201)	0.88-0.93
St Olav	.83 (n = 101)[b]	0.76-0.88	.85 (n = 109)[c]	0.79-0.90	.91 (n = 117)	0.87-0.94
Akershus	.85 (n = 84)	0.78-0.90	.79 (n = 157)	0.73-0.84	.92 (n = 84)	0.87-0.94
Caucasian	.87 (n = 142)[d]	0.82-0.90	.83 (n = 207)	0.78-0.86	.93 (n = 156)	0.91-0.95
Non-Caucasian	.75 (n = 43)	0.58-0.86	.73 (n = 59)	0.58-0.83	.84 (n = 45)	0.72-0.91

[a]*P* = 0.003 compared with TcB vs TSB.

[b]Not significant compared with Akershus (*P* = .7).

[c]Not significant compared with Akershus (*P* = .14)

[d]*P* = 0.045 compared with non-Caucasians.

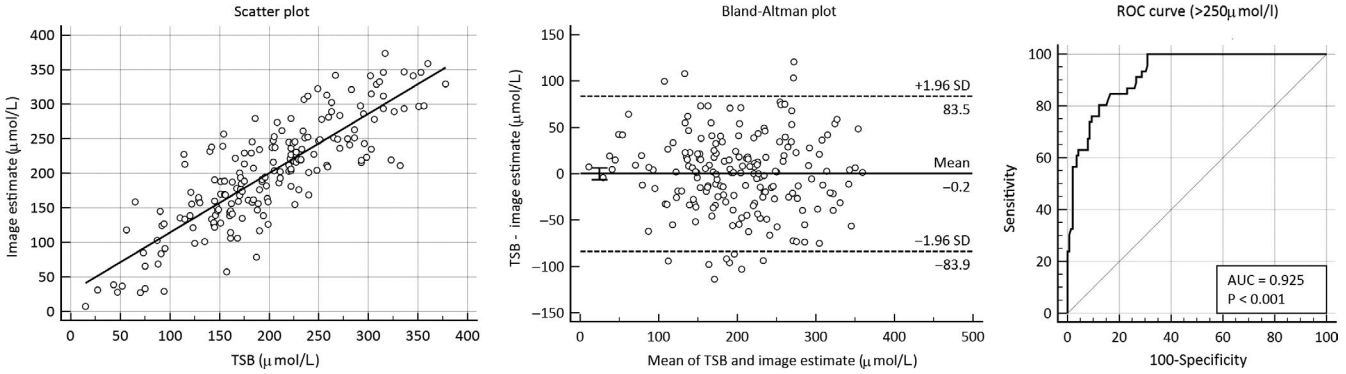

FIGURE 3 Scatter plot, Bland-Altman plot and ROC curve

rules. The model is trained by giving the computer a vast amount of labelled data. In this case, it might be given bilirubin values and images of skin. When the model is trained, it can be used to recognise cases like those it was trained on. With an unlimited data set which covers all possible cases, this method can be highly accurate. However, such data sets are difficult to collect in medicine, and even more so in neonatology.[24] If the final machine-learning algorithm were presented for an image outside the training set, it might either not be able to classify it at all, or it might assign it to the wrong group. This might result in difficulties estimating the bilirubin level, which would result in a high risk of it giving false predictions.

Rong et al conducted a Chinese study on BiliScan machine-learning software (Shenzen Beishen Healthcare Technology Co), which provided bilirubin estimates from smartphone images, with a correlation of $r = .79$.[14] Swarna et al tested the BiliScan application on 35 Indian infants and found that the correlation was substantially lower at $r = .6$.[16] The difference between these two studies might have been because the software had not been trained on Indian infants. However, the study from India was based on a very small sample and therefore the above-mentioned interpretations should be treated with caution. A physics-based method reduces the risk of false predictions in such situations. This is because variations such as skin type could be built into the model. However, if the parameters in the model are wrong, this could result in systematic errors in the bilirubin predictions.

The results of our study were limited by the low number of non-Caucasian participants. As one of the main properties of melanin is to protect the skin from ultraviolet and short wavelength visible light, a higher level of melanin could affect a bilirubin estimate based on digital images. Ethnicity has been shown to be a factor in skin bilirubin kinetics.[25] In our study, the correlation was lower in non-Caucasian infants compared with infants in other ethnic groups. We did not have a sufficient sample size for further analysis between ethnic groups in the non-Caucasian group. As the vast majority of neonatal mortality caused by hyperbilirubinaemia occurs in sub-Saharan Africa and South Asia,[3,25,26] studies in populations with higher melanin contents are needed.

Our study only included newborn infants born at term. TcB has also been correlated with serum levels in preterm born infants.[27,28] As premature infants are more likely to have severe jaundice, it is important to conduct studies in this population in the future.

The largest potential use of the screening tool for neonatal jaundice is in situations where only visual assessment of jaundice is currently available. An advantage of a smartphone-based system over visual assessment is that little or no training is needed to perform the estimates. A limitation of our study was that we did not collect data on visual assessment of jaundice and the comparative potential benefits of the new tool could not be evaluated.

Previous studies have shown that the accuracy of visual detection of jaundice can be influenced by illumination, which varies during the day.[29] Our study tested the system under both standardised and non-standardised light conditions and found similar correlations in the two settings. However, a limitation of the study is that we did not collect data on light intensity while obtaining the images.

Our data were collected from two hospitals that used different methods for measuring TSB. Studies have shown that there are large variations in bilirubin measurements between different laboratory systems and this could have influenced the results.[30]

We had to make changes in the software that we used to obtain the images after data collection at St Olav, and the refined software was used at Akershus. Ideally, we would use the same setup for both locations. As the changes were made to improve the usability and quality of the images, we assume that if the same software was used at both locations, we could potentially have improved the correlations. Finally, only one brand and model of smartphone was used in the study.

5 | CONCLUSION

Based on the findings from the study, a new smartphone-based tool that estimates bilirubin levels from digital images can be used to screen for neonatal jaundice in Caucasian newborn infants. The bilirubin estimates from the images were highly correlated to TSB levels. The tool showed a sensitivity of 100% to identify participants with severe jaundice, defined as TSB above 250 µmol/L, and a specificity of 69%.

ACKNOWLEDGEMENTS

We thank Dr Anne Lee Solevåg for help with arranging data collection at Akershus University Hospital, the participating nurses and midwives who assisted with the data collection and the parents who allowed their infants to take part.

CONFLICT OF INTEREST

Aune, Vartdal and Randeberg are co-founders of Picterus AS, the company with the rights for the further development of the technology described in this study for commercial use. The other authors have no potential conflicts of interest to disclose.

ETHICAL APPROVAL

The study was approved by the Regional Ethical Committee of South East Norway (reference number 2014/619) and by the data protection officers at both hospitals. Participation in the study was based on receiving written informed consent from the parents of the newborn infants.

ORCID

Anders Aune https://orcid.org/0000-0001-6757-7398
Lise Lyngsnes Randeberg https://orcid.org/0000-0003-2608-3759
Elisabeth Darj https://orcid.org/0000-0002-8311-4956

REFERENCES

1. Amos RC, Jacob H, Leith W. Jaundice in newborn babies under 28 days: NICE guideline 2016 (CG98). *Arch Dis Child Educa Pract*. 2017;102:207-209.
2. Hansen TWR, Bratlid D. Bilirubin and brain toxicity. *Acta Paediatr*. 2008;75:513-522.
3. 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.
4. 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;99:1117-1121.
5. Slusher TM, Zipursky A, Bhutani VK. A global need for affordable neonatal jaundice technologies. *Semin Perinatol*. 2011;35:185-191.
6. Carceller-Blanchard A, Cousineau J, Delvin EE. Point of care testing: transcutaneous bilirubinometry in neonates. *Clin Biochem*. 2009;42:143-149.
7. Afanetti M, Eleni Dit Trolli S, Yousef N, Jrad I, Mokhtari M. Transcutaneous bilirubinometry is not influenced by term or skin color in neonates. *Early Human Dev*. 2014;90:417-420.
8. Keren R, Tremont K, Luan X, Cnaan A. Visual assessment of jaundice in term and late preterm infants. *Arch Dis Child Fetal Neonat Ed*. 2009;94:F317-F322.
9. Riskin A, Tamir A, Kugelman A, Hemo M, Bader D. Is visual assessment of jaundice reliable as a screening tool to detect significant neonatal hyperbilirubinemia? *J Pediatr*. 2008;152 782-787. e1-2.
10. Sondaal SF, Browne JL, Amoakoh-Coleman M, et al. Assessing the effect of mHealth interventions in improving maternal and neonatal care in low- and middle-income countries: a systematic review. *PLoS ONE*. 2016;11:e0154664.
11. Leung TS, Outlaw F, MacDonald LW, Meek J. Jaundice Eye Color Index (JECI): quantifying the yellowness of the sclera in jaundiced neonates with digital photography. *Biomed Opt Express*. 2019;10:1250-1256.
12. Aydin M, Hardalac F, Ural B, Karap S. Neonatal jaundice detection system. *J Med Syst*. 2016;40:166.
13. Munkholm SB, Krogholt T, Ebbesen F, Szecsi PB, Kristensen SR. The smartphone camera as a potential method for transcutaneous bilirubin measurement. *PLoS ONE*. 2018;13:e0197938.
14. 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.
15. Taylor JA, Stout JW, de Greef L, et al. Use of a smartphone app to assess neonatal jaundice. *Pediatrics*. 2017;140:e20170312
16. Swarna S, Pasupathy S, Chinnasami B, Manasa DN, Ramraj B. The smart phone study: assessing the reliability and accuracy of neonatal jaundice measurement using smart phone application. *Int J Contemp Pediatr*. 2017;5:285-289.
17. Vartdal G, Randeberg Lise L, Aune A, Kringstad A, Picterus ASImage Based Bilirubin Determination. EP patent EP 3393353 A4. 2019 2016/12/15.
18. Randeberg LL, Bruzell Roll E, Norvang Nilsen LT, Christensen T, Svaasand LO. In vivo spectroscopy of jaundiced newborn skin reveals more than a bilirubin index. *Acta Paediatr*. 2005;94:65-71.
19. De Luca D, Engle W, Jackson G. *Transcutaneous Bilirubinometry*. New York, NY: Nova Science Publishers Inc; 2013.
20. Le Moan S, Urban P. Spectral printing with a CMYKRGB printer: a closer look. *Proceedings of the Color and Imaging Conference*; 2014:131-135.
21. Nagar G, Vandermeer B, Campbell S, Kumar M. Effect of phototherapy on the reliability of transcutaneous bilirubin devices in term and near-term infants: a systematic review and meta-analysis. *Neonatology*. 2016;109:203-212.
22. Bratlid D, Nakstad B, Hansen TW. National guidelines for treatment of jaundice in the newborn. *Acta Paediatr*. 2011;100:499-505.
23. Sufian AT, Jones GR, Shabeer HM, Elzagzoug EY, Spencer JW. Chromatic techniques for in vivo monitoring jaundice in neonate tissues. *Physiol Measure*. 2018;39:095004.
24. Sendak M, Gao M, Nichols M, Lin A, Balu S. Machine learning in health care: a critical appraisal of challenges and opportunities. *EGEMS*. 2019;7:1-4.
25. De Luca D, Jackson GL, Tridente A, Carnielli VP, Engle WD. Transcutaneous bilirubin nomograms: a systematic review of population differences and analysis of bilirubin kinetics. *Arch Pediatr Adolesc Med*. 2009;163:1054-1059.
26. Slusher TM, Zamora TG, Appiah D, et al. Burden of severe neonatal jaundice: a systematic review and meta-analysis. *BMJ Paediatrics Open*. 2017;1:e000105.
27. Kitsommart R, Yangthara B, Wutthigate P, Paes B. Accuracy of transcutaneous bilirubin measured by the BiliCare device in late preterm and term neonates. *J Matern Fetal Neonatal Med*. 2016;29:3641-3645.
28. De Luca D, Zecca E, de Turris P, Barbato G, Marras M, Romagnoli C. Using BiliCheck for preterm neonates in a sub-intensive unit: diagnostic usefulness and suitability. *Early Human Dev*. 2007;83:313-317.
29. De Luca D, Zecca E, Zuppa AA, Romagnoli C. The joint use of human and electronic eye: visual assessment of jaundice and transcutaneous bilirubinometry. *Turk J Pediatr*. 2008;50:456-461.
30. van Imhoff DE, Dijk PH, Weykamp CW, Cobbaert CM, Hulzebos CV. Measurements of neonatal bilirubin and albumin concentrations: a need for improvement and quality control. *Eur J Pediatr*. 2011;170:977-982.

How to cite this article: Aune A, Vartdal G, Bergseng H, Lyngsnes Randeberg L, Darj E. Bilirubin estimates from smartphone images of newborn infants' skin correlated highly to serum bilirubin levels. *Acta Paediatr*. 2020;109:2532–2538. https://doi.org/10.1111/apa.15287