

Picterus® Jaundice Pro as a screening tool for neonatal jaundice



There are several publications (Vartdal et al 2014, Jimenez 2019, Aune et al 2020, Tusoy et al 2022, Aune et al 2023) on the development and validation of the Picterus® Jaundice Pro (JP) device; since publication of these studies, however, the product has been significantly developed and improved. In this paper we present the latest findings from studies conducted across various countries evaluating the performance of Picterus JP.

Picterus AS, September 2026.

Disclaimer: Some of the data presented herein are unpublished and are not yet subjected to formal peer review or public distribution. These findings are preliminary and may be subject to further analysis or revisions. For the unpublished studies, the collaborating authors have agreed for data to be included in this report.

ABSTRACT



Abstract

OVERVIEW

 Introduction and Background **Study Design** **Study population: Characteristics****TECHNICAL DEVICE PERFORMANCE**

 Technical Performance: Picterus vs TSB**SCREENING PERFORMANCE**

 Screening Performance Based on Individual Treatment Thresholds **Screening Performance: AAP Guidelines****CONCLUSIONS**

 Conclusions**REFERENCES**

 Bibliography**APPENDIX**

 Justification of methods: Technical Device Performance

Abstract

- Picterus® Jaundice Pro (Picterus JP) is a medical device designed to screen for neonatal jaundice using a smartphone app and a Picterus® Calibration Card. The device can be used by both parents and healthcare professionals, allowing for early detection and monitoring of jaundice even when the newborn is at home.
- This document presents the results from 8 different studies assessing Picterus JP performance in 790 newborns with various skin tones.
- Correlation results between Picterus and total serum bilirubin show that Picterus JP has a strong agreement with the reference gold standard (Pearson's $r = 0.78$, Spearman correlation = 0.80). Bland Altman analysis showed a small positive bias (9.4 $\mu\text{mol/L}$, (0.5 mg/dL)). The standard deviation of the bias was 53.5 $\mu\text{mol/L}$ (2.8 mg/dL), which is in line with existing transcutaneous devices.
- Picterus JP demonstrated high sensitivity (81%) and specificity (85%) for a threshold of 250 $\mu\text{mol/L}$

(approximately 15 mg/dL) in identifying newborns at risk of developing severe neonatal jaundice, ensuring timely intervention and reducing the risk of severe jaundice.

- Picterus JP showed a good screening potential for identifying newborns requiring further follow-up for diagnosis of neonatal jaundice with a blood test.
- These benefits highlight the potential of Picterus JP as a safe, effective, and accessible tool for screening, ultimately improving newborn health outcomes.

Introduction and Background

Background

Neonatal jaundice is a condition where newborns have yellowish skin due to high levels of bilirubin in their blood. It affects 60-80% of newborns within the first 48 to 72 hours after birth (Olusanya et al. 2018). While it is usually harmless, about 1 million newborns develop severe jaundice each year, which can cause serious problems if undetected or untreated. When bilirubin levels get too high, it can affect the newborns brain and cause permanent damage. Therefore, early detection, monitoring and treatment is important (Olusanya et al. 2014, Bhutani et al. 2015, Bhutani et al. 2011, Bhutani et al. 2010).

The standard test for neonatal jaundice diagnosis is the total serum bilirubin (TSB) test, which requires a blood sample and lab facilities. Non-invasive devices like transcutaneous bilirubinometers (TcB) can detect high bilirubin levels without needing a blood sample (Okwundu et al. 2023). However, these devices are expensive and not always available in health facilities or for home visits (Cohen et al. 2010, Engle et al. 2014, Van den Esker-Jonker et al. 2016, Dräger. Products and Solutions 2024, UNICEF. Supply Catalogue 2026). Many newborns are already at home when signs of jaundice start, typically around 48

hours after birth. As a result, visual examination is often used to check for neonatal jaundice both in the hospital and at home, but this method is not fully reliable and requires experience (Olusanya et al. 2018, Greco et al. 2016, Riskin et al. 2008).

Picterus® Jaundice Pro (Picterus JP) is a medical device that uses a smartphone to screen for neonatal jaundice. It is approved for use by both parents and healthcare professionals. The device uses a smartphone app and a Picterus® Calibration Card to estimate bilirubin levels from images of the skin of a newborn's chest (Figure 1). The Picterus® Calibration Card adjusts for different light conditions and camera sensors.

For more information, please refer to our [website](#).



Figure 1. The Picturus Calibration card is placed on the skin of the newborn's chest. The app automatically takes pictures when the card is detected. An image quality check and bilirubin calculation is performed and the result is displayed within seconds.

Study Design

Study Sites

We included eight cross-sectional studies conducted at different hospital sites (Table 1).

Table 1. Overview of clinical study sites, data collection period and included newborns.

Study site	Data collection period	Newborns (n)
Intermountain Medical Center, McKay-Dee Hospital, and Utah Valley Hospital, USA	February - July, 2024	199
University of Chicago, USA	March - October, 2024	72
Hospital General Zona 1 Oaxaca, Mexico	February 2023-June 2024	144

Candida Vargas Institute, Brazil	January - March, 2025	155
Gulu University & Teaching Hospitals, Uganda	February - April, 2025	102
University Children's Hospital in Skopje, North Macedonia	April - October, 2025	27
Multicentre, India	July - October, 2025	60
Imperial College London, UK	November - December, 2025	31
Total		790 newborns

Study participants

Newborns included in the analysis met the following criteria: age between 1 to 14 days, gestational age of at least 35 weeks, birth weight of 2000 g or more and requiring a blood sample either for clinically suspected jaundice or as part of clinical routine practice. Each newborn contributed one paired Picturus and TSB result.

Newborns were excluded if they had a condition or treatment that affected skin colour or appearance, other than jaundice. Infants

receiving or previously receiving phototherapy were also excluded.

Ethical approvals were obtained from the study sites' Ethics Committees prior to the study start. Written informed consent was collected for all newborns.

Data collection

Data were collected under the ambient conditions found in each of the settings. The following data on each newborn was collected: day and time of birth, birth weight, and gestational age.

All procedures were performed within a span of 60 minutes.

Total Serum Bilirubin (TSB)

TSB levels were measured with a blood sample obtained by either heel prick or venous puncture. Samples were immediately processed at the hospitals' laboratory by trained personnel. Different TSB analytical methods were used across study sites, depending on local availability and routine laboratory practice.

Picterus JP

To perform a Picterus JP scan, the Picterus Calibration Card was placed on a flat part of the newborn's chest and an image set, consisting of 3 images with and 3 without flash, was automatically captured on a smartphone using the Picterus JP app.

Study population: Characteristics

Clinical characteristics of the newborns included in the study are shown in Table 2.

Table 2. Clinical characteristics of newborns (n = 790).

	Median [IQR]	Mean (SD)	Min – Max
Birthweight (grams)	3160 [2770 – 3500]	3156 (533)	2000 – 5700
Gestational age (weeks)	39.0 [37.3 – 40.0]	38.5 (1.5)	35 – 43
Age at inclusion (hours)	45 [29 – 84]	65 (51)	24 – 319
TSB (μmol/L)	145 [103 – 210]	164 (83)	8 – 571
(mg/dL)	8.5 [6.0 – 12.3]	9.6 (4.9)	0.5 – 33.4

Picterus JP ($\mu\text{mol/L}$)	149 [112 – 242]	173 (78)	39 – 362
(mg/dL)	8.7 [6.5 – 14.1]	10.1 (4.6)	2.3 – 21.2

IQR: Interquartile Range, SD: standard deviation, Min: minimum, Max: maximum, TSB: Total Serum Bilirubin, Picterus JP: Picterus® Jaundice Pro, $\mu\text{mol/L}$: micromoles per litre, mg/dL: milligram per decilitre.

The inclusion of newborns from sites spanning multiple continents with diverse populations ensured representation of all skin tones within the study cohort, which is important, since transcutaneous bilirubin assessments are affected by the presence of melanin (Engle et al. 2014, Maisels et al. 2004). For each newborn, a red-channel (R-channel) value was determined from Picterus images. The R-channel value is affected by the melanin content of a newborn's skin. The R-channel values in this data set range from 0.14 to 0.56; lower values indicate darker skin. Figure 2 shows a histogram of the R-channel values from the combined clinical studies.

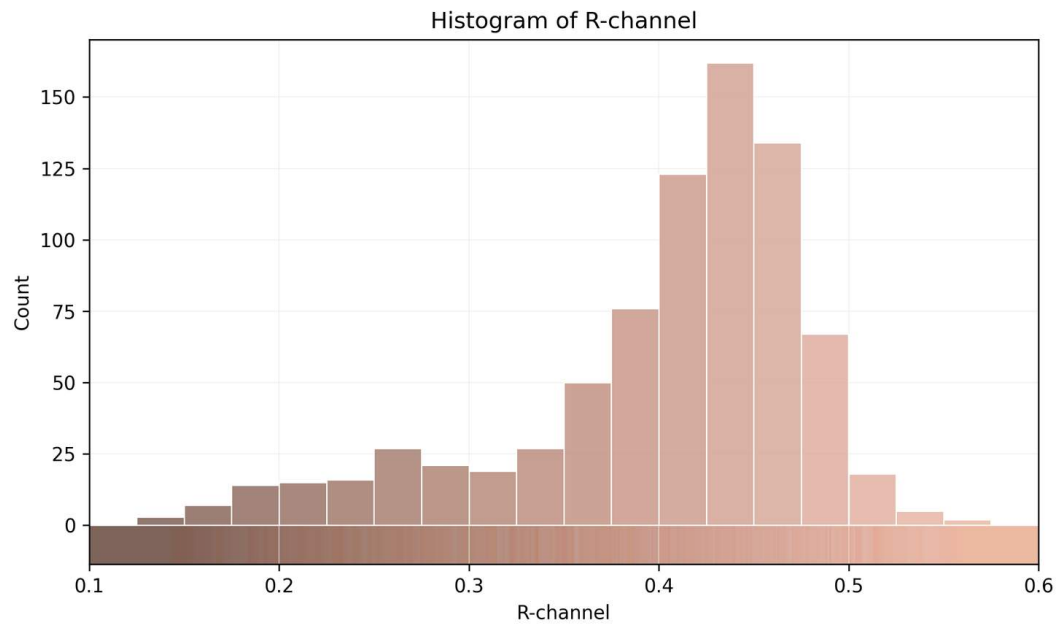


Figure 2. The R-channel values (x-axis) from Picterus® Jaundice Pro skin images were used as a proxy for newborn skin tone. Each histogram bar represents the number of samples within an R-channel bin. The bar colours and the gradient bar indicate the mean skin colour of the samples in the corresponding R-channel range.

Technical Performance: Picterus vs TSB

Method

Technical performance was assessed by comparing paired Picterus JP bilirubin results with total serum bilirubin (TSB). Agreement and correlation were evaluated using Pearson's r , Spearman (ρ) and Bland-Altman analysis. Diagnostic accuracy was assessed using receiver operating characteristic (ROC) analysis, including area under the curve (AUC) and Youden Index. Sensitivity and specificity were calculated at a threshold of $250 \mu\text{mol/L}$ ($\sim 15 \text{ mg/dL}$) by comparing Picterus JP classifications ("above"/"below" threshold) with corresponding TSB classifications. Sensitivity was defined as the proportion of newborns with $\text{TSB} \geq 250 \mu\text{mol/L}$ ($\sim 15 \text{ mg/dL}$) correctly identified, and specificity as the proportion with $\text{TSB} < 250 \mu\text{mol/L}$ ($\sim 15 \text{ mg/dL}$) correctly classified. Further methodological justification is provided in the Appendix ("Justification of methods: Technical Device Performance").

Correlation and Agreement

Figures 3 and 4 show the correlation and Bland-Altman plots for Picterus JP and TSB results of the combined clinical trials ($n = 790$). There was a strong statistically significant correlation between

Picterus JP results and TSB (Pearson's $r = 0.78$, $P < 0.0001$, Spearman $\rho = 0.80$) (Figure 3).

The Bland-Altman plot shows that Picterus JP has a mean difference of $9.4 \mu\text{mol/L}$ (0.5 mg/dL), this means that Picterus JP tends to slightly overestimate bilirubin values compared with TSB. The standard deviation of the mean difference was $53.5 \mu\text{mol/L}$ (2.8 mg/dL) and the 95% limits of agreement was $-95.4 \mu\text{mol/L}$ to $114.2 \mu\text{mol/L}$ (-5.6 mg/dL to 8.3 mg/dL).

Despite the Bland-Altman also shows that for values above $300 \mu\text{mol/L}$ there is a tendency for Picterus to underestimate the TSB result. This observation alone is not clinically meaningful, as clinical relevance is better reflected in the screening performance at decision thresholds (see Section "Screening Performance Based on Individual Treatment Thresholds"). In addition, a limitation is also the few data points collected at high bilirubin values, so more high values are needed to draw solid conclusions.

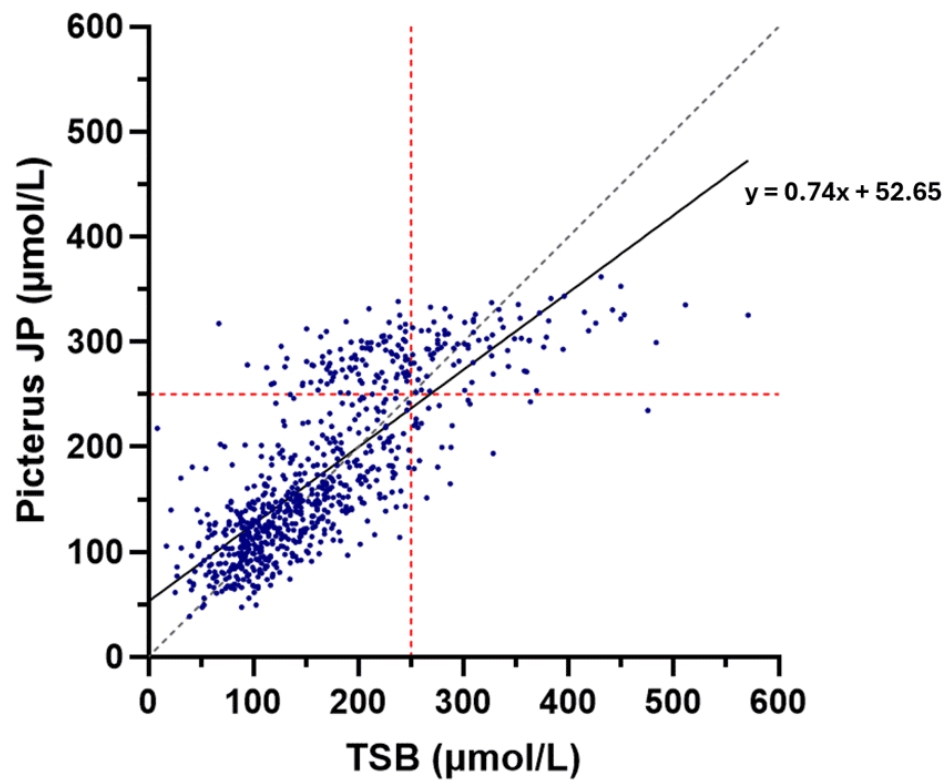


Figure 3. Correlation plot between Picterus JP (y-axis) and TSB (x-axis) in $\mu\text{mol/L}$ ($n=790$). Pearson's $r = 0.78$. The line of best fit is the black line, with equation indicated, the grey dashed line is the line of identity, and the red dotted lines represent the bilirubin threshold at 250 $\mu\text{mol/L}$. Picterus JP: Picterus® Jaundice Pro; TSB: total serum bilirubin; $\mu\text{mol/L}$: micromoles per litre

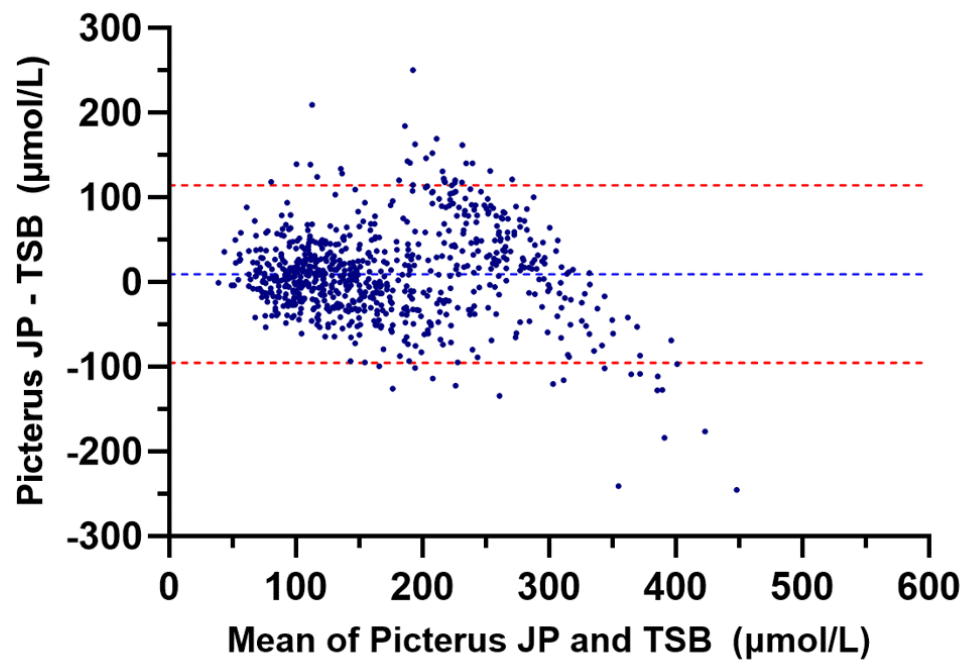


Figure 4. Bland-Altman plot of Picterus JP and TSB. X-axis shows the mean of Picterus JP and TSB; y-axis shows the differences between Picterus JP-TSB; The blue dotted line represents the mean difference between Picterus JP-TSB (9.4 $\mu\text{mol/L}$), the red-dotted lines represent the 95% limits of agreement (-95.4 $\mu\text{mol/L}$ to 114.2 $\mu\text{mol/L}$). Picterus JP: Picterus Jaundice Pro; TSB: total serum bilirubin; $\mu\text{mol/L}$: micromoles per litre

Sensitivity and Specificity at the 250 $\mu\text{mol/L}$ (~15 mg/dL) Threshold

At the 250 $\mu\text{mol/L}$ threshold, Picterus JP demonstrated a sensitivity of 81% and a specificity of 85%, indicating good overall classification performance. A sensitivity of 81% reflects the device's ability to correctly identify the majority of newborns with $\text{TSB} \geq 250 \mu\text{mol/L}$, while a specificity of 85% shows that most newborns with TSB below this threshold were correctly classified as not exceeding it.

It should be noted that this analysis at a fixed threshold does not reflect the full screening potential of Picterus JP, as individual newborns have different treatment thresholds depending on factors such as gestational age and the presence of neurotoxicity risk factors. This variability in clinical decision thresholds is addressed in the following sections where treatment thresholds according to AAP or NICE guidelines are applied individually for each newborn.

Receiver operating characteristic (ROC)

The ROC curve (Figure 5) illustrates the agreement between Picterus and total serum bilirubin (TSB) measurements across a range of decision cut-offs. The area under the curve (AUC) of 0.93 indicates excellent overall discrimination, demonstrating that Picterus effectively distinguishes between higher and lower bilirubin concentrations. The figure also shows the operating point corresponding to the predefined 250 $\mu\text{mol/L}$ threshold (sensitivity of 81%, specificity 85%), which reflects the selected threshold for the device performance analysis, even though it is not the statistically optimal ROC cut-off.

The Youden index ($J = 0.72$) identified 150 $\mu\text{mol/L}$ as the statistically optimal cut-off, corresponding to a sensitivity of 92% and a specificity of 80%. This threshold maximised the combined sensitivity and specificity and therefore represented the best overall discrimination in the study population. However, the cut-off is relatively low in the context of clinical decision-making, where treatment thresholds are

typically higher and individualised according to gestational age and other clinical factors. Consequently, the Youden index threshold is best interpreted as a statistical reference point rather than a clinically actionable threshold.

A cut-off of 105 $\mu\text{mol/L}$ was associated with 100% sensitivity, ensuring that all infants with TSB concentrations above the reference threshold were correctly identified in this dataset. This level of sensitivity was achieved at the expense of specificity, which was 58% (false positive rate 42%), resulting in a higher proportion of false positive classifications. In a screening context, however, maximising sensitivity is often prioritised to minimise the risk of missed cases. Consequently, the 105 $\mu\text{mol/L}$ threshold may be useful as a conservative screening threshold, although its lower specificity would be expected to increase the number of infants requiring confirmatory testing.

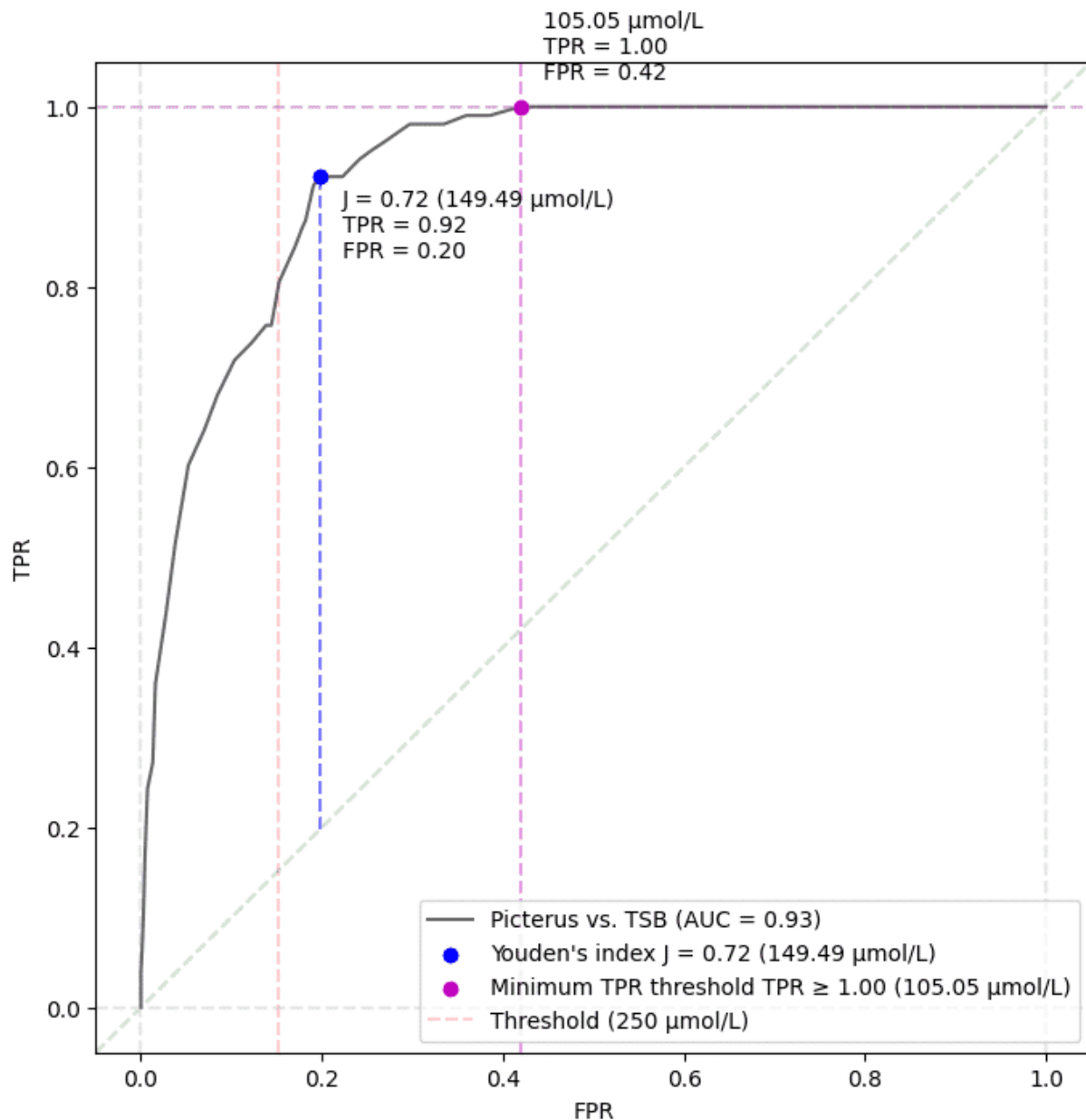


Figure 5. Receiver operating characteristic (ROC) curve illustrating the diagnostic performance of Picterus relative to total serum bilirubin (TSB) measurements across a range of cut-offs. The area under the curve (AUC) was 0.93. The Youden index ($J = 0.72$) identified an optimal cut-off of 150 µmol/L, corresponding to the threshold that maximised the sum of sensitivity and specificity. The threshold corresponding to 100% sensitivity was 105 µmol/L. The predefined 250 µmol/L threshold is shown for reference. TPR = true positive rate; FPR = false positive rate.

Affect of Skin Tone on Picterus Results

Since transcutaneous bilirubin results are known to be influenced by skin tone, we investigated the relationship between the R-channel value and the difference between paired Picterus and TSB results. Figure 6 shows the correlation between error (Picterus-TSB) and R-channel values. A weak but statistically significant negative correlation was observed (Pearson's $r = -0.13$, $P = 0.0003$; Spearman's $\rho = -0.18$). Linear regression analysis gave an r-squared value of 0.017, so only 1.7% of the variability in the precision of Picterus JP could be attributed to variability in the R-channel value. The equation for the regression line is: Difference (Picterus JP-TSB) = $45 - 88.5 \times (r \text{ channel})$. R-squared = 0.017. These results indicate that higher R-channel values (corresponding to lighter skin tones) are associated with a slight reduction in the error.

Given the small effect size, these findings suggest that skin tone does not impact follow-up decisions for Picterus screening results.

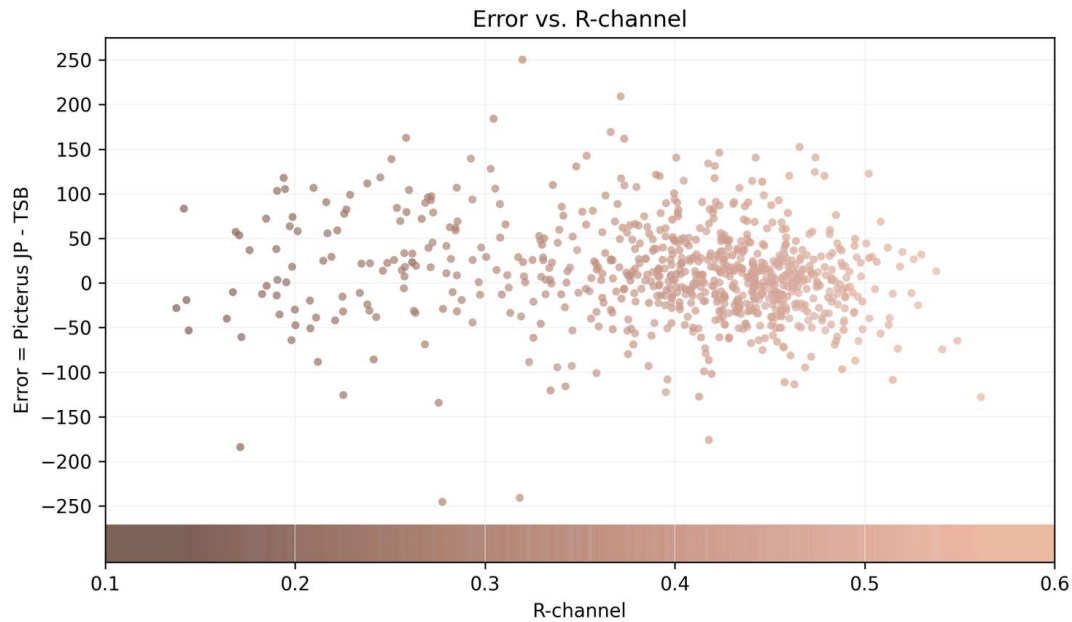


Figure 6. The error between paired Picterus and TSB results plotted against R-channel. The colour of each data point and the gradient bar represents the corresponding skin colour across the R-channel range. The equation of the best fit line (not shown): $y = -88.5x + 45.1$

Conclusions – Technical Performance

This technical performance review shows that Picterus Jaundice Pro performs reliably for screening neonatal hyperbilirubinemia. Its agreement pattern, including a conservative bias, supports its use as a screening tool.

The magnitude and variability of measurement differences are consistent with those reported for established transcutaneous bilirubin (TcB) devices (Engle et al. 2014), indicating that Picterus JP performs within the range of accepted non-invasive bilirubin screening technologies. These findings support the technical validity of the

camera-based approach as an alternative means of bilirubin assessment.

As with all TcB-based screening methods, the clinical utility of Picterus JP is best assessed within individualized, guideline-based treatment frameworks that account for gestational age and additional risk factors, as outlined in the following section "Screening Performance Based on Individual Treatment Thresholds".

Screening Performance Based on Individual Treatment Thresholds

Method

Neonatal hyperbilirubinaemia is managed using individualised treatment thresholds, which vary according to gestational age, postnatal age, and the presence of neurotoxicity risk factors. As a result, there is no single bilirubin concentration that universally defines clinically significant jaundice. Instead, clinical decision-making is based on whether an infant's bilirubin level exceeds their specific treatment threshold as defined by established guidelines.

For this reason, evaluation of screening performance must be aligned with the clinical decision context. Use of a fixed bilirubin cut-off (e.g., 250 $\mu\text{mol/L}$) does not reflect how jaundice is managed in practice and may lead to biased estimates of device performance. Specifically, a fixed threshold would not account for infants with lower treatment thresholds (e.g., preterm or high-risk infants) and may overestimate performance in infants with higher thresholds. Therefore, screening accuracy is most appropriately assessed based on the device's ability to correctly classify infants relative to their individualised follow-up and treatment thresholds. In this study, this approach was applied

using AAP (Kemper et al. 2022) guideline thresholds for newborns with neurotoxicity risk factors.

Picterus screening performance according to the NICE guidelines has been independently evaluated, and the white paper can be downloaded from the [Clinical Evidence](#) section of the Picterus website.

Use of TSB as Reference Standard

Total serum bilirubin (TSB) is the accepted gold standard measurement for assessing bilirubin concentration and forms the basis of follow-up and treatment decisions in clinical guidelines. In this analysis, TSB is used to define the reference outcome for screening performance.

For each paired Picterus Scan and TSB measurement:

- The infant's individual follow-up threshold is determined according to the relevant clinical guideline.
- The measured TSB value is compared to this threshold to establish the reference classification:
 - Above follow-up threshold → *follow-up indicated* ("Follow-up = yes")
 - Below follow-up threshold → *no follow-up indicated* ("Follow-up = no")

Picterus results are then evaluated against this reference classification. Each result is categorised as a true positive, false positive, true negative, or false negative based on agreement with the TSB-derived outcome (Table 3).

- **True Positive (TP):** Picterus Scan result above follow-up threshold, and TSB result above follow-up threshold.
- **False Positive (FP):** Picterus Scan result above follow-up threshold, and TSB result below follow-up threshold.
- **True Negative (TN):** Picterus Scan result below follow-up threshold, and TSB result below follow-up threshold.
- **False Negative (FN):** Picterus Scan result below follow-up threshold, and TSB result above follow-up threshold.

Table 3. Categorization of results based on agreement with gold standard. TSB = Total serum bilirubin.

	TSB above threshold	TSB below threshold
Picterus result above threshold	True positive (TP)	False positive (FP)
Picterus result below threshold	False negative (FN)	True negative (TN)

Screening Performance: AAP Guidelines

The Picterus screening performance was calculated using the same approach as for standard TcB devices according to individual thresholds for newborns with one or more hyperbilirubinemia neurotoxicity risk factors, as outlined in the AAP 2022 guidelines for managing hyperbilirubinemia in newborns with a gestational age of 35 weeks or more (Kemper et al. 2022). The presence of risk factors was applied to all newborns to ensure a more conservative approach because we lacked specific information on individual risk factors for some of the newborns.

According to the guidelines, a total of 74 newborns had TSB results that exceeded their phototherapy threshold.

Paired Picterus and TSB results were collected for all newborns. Picterus screening results and their corresponding TSB results were independently classified as either requiring follow-up or not requiring follow-up according to the criteria below. The follow-up action indicated by Picterus was then compared with that indicated by the paired TSB result to assess agreement. For the purposes of this analysis, the TSB result was considered the reference standard for determining the appropriate follow-up action.

Each of the screening results were defined as "follow-up = yes", or "follow-up = no" according to the following criteria:

1. Follow-up indicated (Follow-up = yes). Picterus/TSB Result:

- a. Greater than or equal to 15mg/dL, or
- b. Greater than or equal to the phototherapy threshold, or
- c. within 0-3 mg/dL below phototherapy threshold.

2. No follow-up indicated (Follow-up = no). Picterus/TSB Result:

- a. below 15 mg/dL, and
- b. more than 3 mg/dL below phototherapy threshold.

Finally, any false negative results were further analysed to see whether newborns had TSB values exceeding their individual phototherapy thresholds:

- **TSB: Phototherapy = yes:** the newborn's TSB result was above their phototherapy threshold
- **TSB: Phototherapy = no:** the newborn's TSB result was below their phototherapy threshold

These results are visualised on Figure 7.

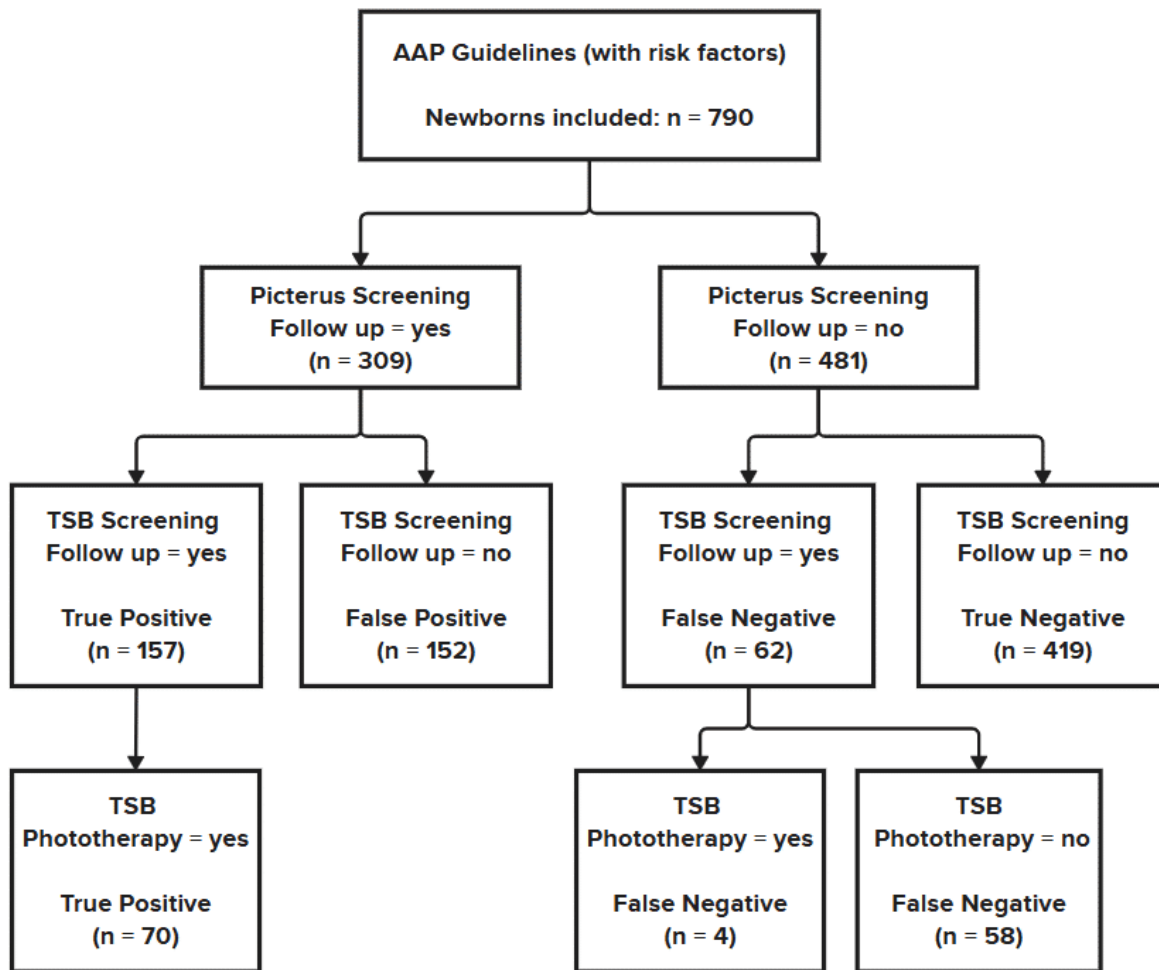


Figure 7. Flow diagram of result classifications for newborns according to AAP guidelines with risk factors.

Summary AAP Screening Performance

- According to the guidelines, a total of 74 newborns had TSB results that exceeded their phototherapy treatment threshold. Of these 74 newborns:
 - Picterus correctly identified 70/74 (95%) as needing follow-up.

- For 4/74 (5%) of newborns with TSB exceeding their AAP phototherapy thresholds, Picterus failed to identify the need for follow-up. The difference between the Picterus result and follow-up thresholds in these newborns were 0.5 mg/dL, 0.9 mg/dL, 1.6 mg/dL, 1.9 mg/dL and might be explained by the fact that Picterus JP measures bilirubin transcutaneously, whereas treatment decisions are based on serum bilirubin levels. Additional variability may arise from for example blood sample handling.
- Picterus correctly ruled out further follow-up for 419 newborns.
- Picterus correctly recommended follow-up in 157 newborns, of whom 70 were over the phototherapy threshold.

Using the general AAP follow-up criteria, Picterus achieved a sensitivity of 71.7%, specificity of 73.4%, positive predictive value of 50.8%, and negative predictive value of 87.1%. When the analysis was restricted to newborns with TSB concentrations above the phototherapy threshold, sensitivity increased to 94.6% and the negative predictive value to 99.2%.

Applying a Picterus "Safety Margin" for 100% sensitivity

Safety margins can be applied to Picterus bilirubin estimates to increase sensitivity. In the analysis presented below, a 3 mg/dL safety margin was added to all Picterus values before comparison with the AAP phototherapy thresholds. A margin of 3 mg/dL was selected

because it is consistent with the safety margin commonly recommended when interpreting transcutaneous bilirubin (TcB) measurements near treatment thresholds. This approach resulted in 100% sensitivity for identifying newborns with TSB concentrations above the phototherapy threshold, although the increased sensitivity was accompanied by a reduction in specificity and a higher number of follow-up recommendations.

The screening results for AAP guidelines with a 3 mg/dL safety margin added to Picterus values are shown on Figure 8.

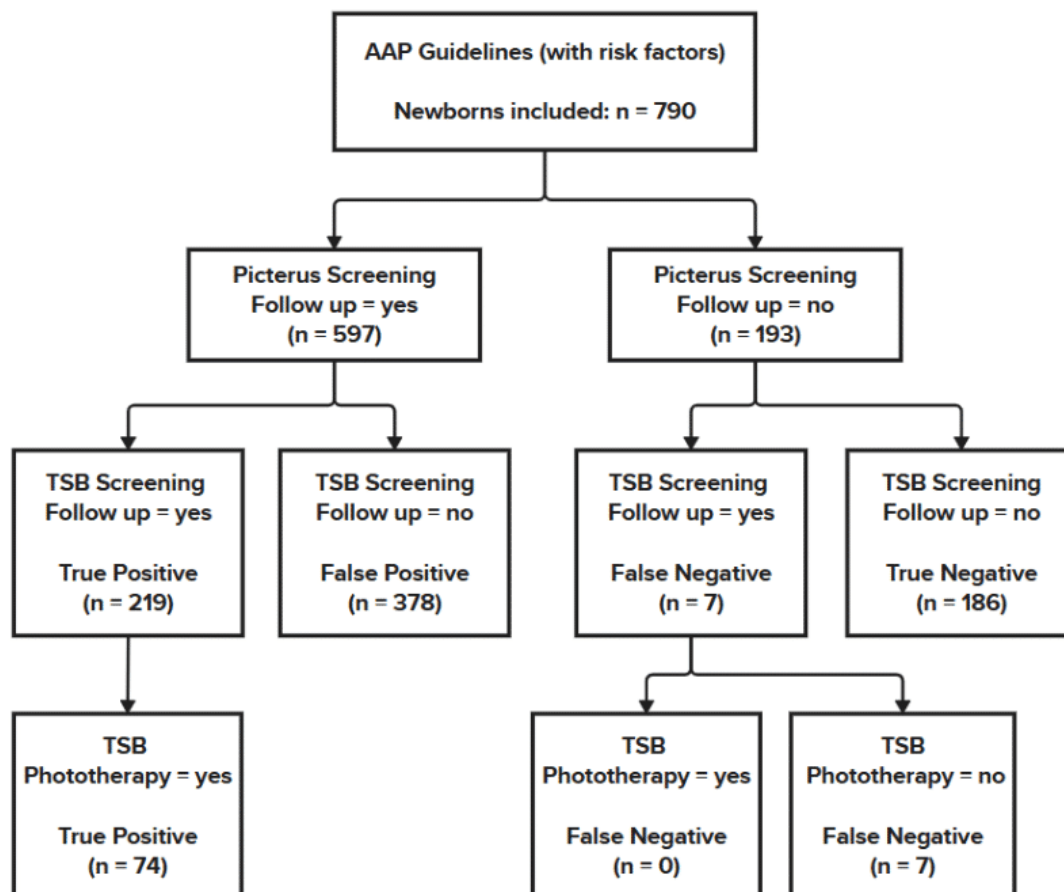


Figure 8. Flow diagram of result classifications for newborns according to AAP guidelines with risk factors. A safety margin of 3 mg/dL was applied to all Picterus values prior to analysis.

- According to the guidelines, a total of 74 newborns had TSB results that exceeded their phototherapy treatment threshold. Of these 74 newborns:
 - Picterus correctly identified 74/74 (100%) as needing follow-up.
- Picterus correctly ruled out further follow-up for 186 newborns.
- Picterus correctly recommended follow-up in 219 newborns, of whom 74 were over the phototherapy threshold.

Using the general AAP follow-up criteria, Picterus achieved a sensitivity of 96.9%, specificity of 33.0%, positive predictive value of 36.7%, and negative predictive value of 96.4%. When the analysis was restricted to newborns with TSB concentrations above the phototherapy threshold, applying a 3 mg/dL safety margin resulted in 100% sensitivity and 100% negative predictive value.

Conclusions

Overall, across the combined dataset (n = 790), Picterus JP demonstrated robust analytical and clinical performance. The device showed strong agreement with TSB ($r = 0.78$) and a small positive bias ($9.4 \mu\text{mol/L}$), consistent with a conservative, safety-oriented measurement profile that minimizes the risk of underestimation. Observed variability (SD $53.5 \mu\text{mol/L}$ [2.8 mg/dL]) was in line with established TcB devices, supporting the technical validity of the approach.

At a clinically relevant threshold, Picterus JP achieved good classification performance (sensitivity 81%, specificity 85%) and, importantly, high screening sensitivity when evaluated against individualized treatment thresholds. The device correctly identified 95% (AAP) of newborns requiring phototherapy, indicating reliable detection of clinically significant hyperbilirubinemia within guideline-based care pathways. Applying a safety margin of 3 mg/dL increased sensitivity for detecting newborns needing phototherapy to 100%.

This is further supported by excellent discrimination on ROC analysis (AUC = 0.93), confirming the strong inherent ability of the method to differentiate between clinically relevant bilirubin levels.

Taken together, these findings indicate that Picterus JP provides clinically meaningful and safety-oriented screening performance, comparable to existing TcB technologies, with reliable identification of newborns requiring further evaluation.

Importantly, Picterus JP is designed as a screening tool, and its clinical value should be interpreted in that context. Good screening performance demonstrates that small numerical differences between Picterus estimates and TSB measurements do not impact clinical decision-making, as only a small number of infants exceeding individual phototherapy thresholds were missed. A potential limitation is the risk of false positive results, where overestimation by Picterus may lead to additional confirmatory testing. However, the non-invasive nature of Picterus enables frequent and repeat measurements, which supports closer monitoring without burden to the infant. In situations where guidelines recommend repeat screening, the ability to easily perform serial measurements with Picterus represents a practical and accessible option for follow-up. Although not formally evaluated in this study, repeated measurements may also facilitate assessment of bilirubin trends over time, which could further support clinical management.

In summary, the findings indicate that Picterus JP combines strong analytical performance with clinically applicable screening capability, while also enabling more flexible, patient-centred models of care, supporting its role as a screening tool for neonatal jaundice.

Bibliography

Bibliography

The following articles, publications, and other sources are referenced throughout the white paper. Here listed in alphabetical order

- Amos RC, Jacob H, Leith W. Jaundice in newborn babies under 28 days: NICE guideline 2016 (CG98). Archives of disease in childhood - Education & practice edition. 2017;102(4):207-9.
- Aune A, Vartdal G, Bergseng H, Randeberg LL, Darj E. Bilirubin estimates from smartphone images of newborn infants' skin correlated highly to serum bilirubin levels. Acta Paediatrica, International Journal of Paediatrics. 2020;109(12):2532-8.
- Aune A, Vartdal G, Jimenez Diaz G, Gierman LM, Bergseng H, Darj E. Iterative Development, Validation, and Certification of a Smartphone System to Assess Neonatal Jaundice: Development and Usability Study. JMIR Pediatrics and Parenting. 2023;6:e40463.
- Bhutani VK J-HL. The clinical syndrome of bilirubin-induced neurologic dysfunction. Semin Perinatol. 2011;2011(35):101-3.
- Bhutani VK, Wong RJ, Vreman HJ, Stevenson DK. Bilirubin production and hour-specific bilirubin levels. J Perinatol. 2015;35(9):735-8.

- Bhutani VK, Zipursky A, Blencowe H, Khanna R, Sgro M, Ebbesen F, 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.
- Cohen RS, Wong RJ, Stevenson DK. Understanding neonatal jaundice: a perspective on causation. *Pediatr Neonatol.* 2010;51(3):143-8.
- Diagnóstico y Tratamiento de la Ictericia Neonatal. Guía de Práctica Clínica: evidencias y recomendaciones. 2019 ed. Mexico: CENETEC; 2019.
- Dräger. Products and Solutions 2024 [cited 2024. Available from: https://www.draeger.com/en-us_us/SearchResults?s=Dr%C3%A4ger+Jaundice+Meter+JM-105.
- Engle WD, Jackson GL, Engle NG. Transcutaneous bilirubinometry. *Seminars in Perinatology.* 2014;38(7):438-51.
- Greco C, Arnolda G, Boo N-Y, Iskander IF, Okolo AA, Rohsiswatmo R, et al. Neonatal Jaundice in Low- and Middle-Income Countries: Lessons and Future Directions from the 2015 Don Ostrow Trieste Yellow Retreat. *Neonatology.* 2016;110(3):172-80.
- Jimenez Diaz G. Validation of a new smartphone app to assess neonatal jaundice in a Mexican population. [Master thesis]. Trondheim, Norway: Norwegian University of Science and Technology; 2019.
- Kemper AR, Newman TB, Slaughter JL, Maisels MJ, Watchko JF, Downs SM, et al. Clinical Practice Guideline Revision: Management of Hyperbilirubinemia in the Newborn Infant 35 or More Weeks of Gestation. *Pediatrics.* 2022;150(3).
- Maisels MJ, Ostrea EM, Jr, Touch S, Clune SE, Cepeda E, Kring E, et al. Evaluation of a new transcutaneous

bilirubinometer. *Pediatrics*. 2004;113(6):1628-35.

- Okwundu CI, Olowoyeye A, Uthman OA, Smith J, Wiysonge CS, Bhutani VK, et al. Transcutaneous bilirubinometry versus total serum bilirubin measurement for newborns. *Cochrane Database of Systematic Reviews*. 2023(5).
- Olusanya BO, Kaplan M, Hansen TWR. Neonatal hyperbilirubinaemia: a global perspective. *Lancet Child Adolesc Health*. 2018;2(8):610-20.
- 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(12):1117-21.
- Riskin A TA, 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(6):782-7.
- Thomas, David H et al. "Total bilirubin assay differences may cause inconsistent treatment decisions in neonatal hyperbilirubinaemia." *Clinical chemistry and laboratory medicine* vol. 60,11 1736-1744. 29 Aug. 2022, doi:10.1515/cclm-2022-0749
- Tusoy J, Odland JØ, Aune A, Jimenez-Diaz G. A mHealth application as a screening tool for neonatal jaundice in Filipino neonates [Master thesis]: Norwegian University of Science and Technology; 2022.
- UNICEF. Supply Catalogue; 2026. Bilirubin meter product number: S0002644 Available from: <https://supply.unicef.org/>
- Van den Esker-Jonker B, den Boer L, Pepping RMC, Bekhof J. Transcutaneous Bilirubinometry in Jaundiced Neonates: A Randomized Controlled Trial. *Pediatrics*. 2016;138(6):e20162414.



Vartdal G. Development of a Smartphone-based diagnostic Tool for Jaundice. In: Stokke BT, Randeberg LL, Aune A, Ramstad T, editors.: NTNU; 2014.

Justification of methods: Technical Device Performance

Gold standard for hyperbilirubinemia diagnosis

When assessing the technical performance of a device such as Picterus, its results must be compared against a gold standard reference. The gold standard represents the most accurate and accepted method for determining the true bilirubin level, and serves as the benchmark for evaluating correctness.

Total serum bilirubin (TSB) is considered the gold standard for the assessment of neonatal hyperbilirubinemia. This designation reflects its long-standing clinical use, analytical accuracy, and its central role in evidence-based guidelines (e.g. AAP, NICE) that define thresholds for follow-up, phototherapy, and exchange transfusion.

TSB is measured from a blood sample using standardized laboratory methods (e.g. diazo-based assays or spectrophotometry). These methods directly quantify circulating bilirubin and are traceable to reference standards, allowing consistent interpretation across institutions and studies. Importantly, all major treatment nomograms

and decision thresholds are defined in terms of TSB, making it the clinical reference against which alternative measurement techniques are evaluated.

Using TSB to assess Picterus Technical Device Performance

When evaluating the technical device performance of Picterus, TSB serves as the comparator reference method. Both methods provide a result reflecting combined (total) bilirubin, without distinguishing between direct and indirect fractions. Picterus estimates bilirubin levels non-invasively using smartphone-based imaging and algorithms, whereas TSB represents the laboratory-confirmed serum concentration.

Technical Device Performance assessment typically involves:

- **Paired measurements**, where Picterus results are compared with TSB values obtained within a defined time window.
- **Statistical agreement analyses**, such as bias (mean difference), limits of agreement (Bland–Altman), correlation, and error metrics.

Limitations of Using TSB as the Reference Standard

While TSB is widely accepted as the reference standard, its limitations should be acknowledged when assessing the technical performance of Picterus. Picterus estimates bilirubin levels non-invasively via

transcutaneous imaging, whereas TSB represents a direct biochemical measurement; as such, some degree of variation between the two methods is expected. In addition, TSB measurements are subject to variability arising from differences in laboratory methods and calibration, as well as pre-analytical factors such as sample handling (e.g., haemolysis or light exposure) and inter-laboratory variation (Thomas et al., 2022). These factors introduce measurement uncertainty in the reference standard itself, thereby limiting the achievable agreement for any comparator device. Consequently, observed differences between Picterus and TSB reflect the combined uncertainty of both methods and should be considered when interpreting measures such as bias and limits of agreement.

Screening methods

Screening for neonatal jaundice is essential for the early identification of newborns at risk of elevated bilirubin levels. The aim of screening is to enable close monitoring and timely intervention when necessary, while also minimizing unnecessary blood draws.

Different screening methods may be used independently of each other, or in combination with each other. They each have their advantages and disadvantages.

Visual assessment remains widely used due to its simplicity and zero cost. However, it is subjective and highly dependent on clinician experience, ambient lighting, and skin tone. It has consistently demonstrated limited accuracy for detecting elevated bilirubin levels

and may fail to identify infants requiring follow-up, particularly in non-ideal conditions. As such, it is generally considered insufficient as a standalone screening method.

Transcutaneous bilirubin (TcB) provides a non-invasive, quantitative estimate of bilirubin and is well established in hospital-based screening. TcB improves detection compared to visual assessment and reduces the need for blood sampling. However, access to TcB devices may be limited outside hospital settings due to cost, availability, and maintenance requirements, which can restrict use in primary care, early discharge pathways, or home follow-up.

Picterus, a smartphone-based solution, is designed to provide an accessible and objective alternative for bilirubin assessment. By combining image-based analysis with standardized calibration, it enables quantitative screening using widely available devices.

Picterus demonstrates performance comparable to TcB across clinically relevant bilirubin ranges, while offering greater flexibility for use in community and home settings. This supports its potential to extend access to reliable jaundice screening beyond traditional clinical environments.

Cut-off values and thresholds for device performance

When evaluating a screening tool such as Picterus, it is essential to clearly define device performance and how it is measured. In this context, device performance refers to how well Picterus can correctly

classify infants according to whether their bilirubin level is above or below a clinically relevant threshold. This is typically expressed using metrics such as:

- **Sensitivity:** the proportion of true high bilirubin cases correctly identified
- **Specificity:** the proportion of low bilirubin cases correctly identified
- **Positive and negative predictive values**
- **Overall accuracy**

To calculate these metrics, results must be converted into binary outcomes (e.g., “above threshold” vs. “below threshold”). This is why a cutoff value is required: it defines what constitutes a clinically meaningful “positive” case. Without such a threshold, it is not possible to determine whether a measurement is correct or incorrect in a clinically interpretable way.

Why 250 $\mu\text{mol/L}$ (approximately 15 mg/dL) is a sensible threshold for evaluating technical device performance

A threshold of 250 $\mu\text{mol/L}$ (approximately 15 mg/dL) total serum bilirubin (TSB) represents a clinically meaningful and methodologically appropriate cutoff for evaluating technical Picterus performance.

First, this level aligns with relevant treatment and follow-up ranges in international guidelines (e.g., AAP, NICE). The 250 $\mu\text{mol/L}$ threshold targets the range where bilirubin transitions from physiological to potentially harmful levels in most newborns. This is reflected in the recommendations for follow-up and treatment, marking it as a general cut-off for all screening results. This makes it particularly relevant for screening, where the goal is to identify infants who require follow-up or intervention. Evaluating performance at this level therefore reflects real-world clinical decision-making.

Second, this cutoff is valuable from a methodological perspective. It supports comparison with existing studies, as many evaluations of TcB and other non-invasive tools use thresholds in a similar range. This allows technical Picterus performance to be benchmarked against established technologies.

In conclusion, demonstrating performance at 250 $\mu\text{mol/L}$ highlights Picterus' technical ability to function as an effective screening and triage tool.