Appendix A: Details of the studies included in this exceptional surveillance review of jaundice CG98

Studies on risk factors for Hyperbilirubinaemia

Study Details	Population	Cases	Controls	Reference Standard	Outcomes
Hamadneh et al. 2016 Retrospective study Jan 2013 – Nov 2015	886 women with at least two previous cesareans who delivered by cesarean at 37 weeks of pregnancy or later	Group 1 505 (57.0%) delivered at 37 weeks	Group 2 381 (43.0%) delivered at 38 weeks or later		In a multivariate analysis, neonatal jaundice was more common in group 2 (adjusted odds ratio 2.1, 95% confidence interval 1.7–2.7; P=0.035).

Studies on identifying hyperbilirubinaemia in babies with darker skin tones

Study Details	Population	Index test(s)	Reference Standard	Outcomes
Studies reporting visual assessment for jaundice				
Dionis et al. 2021 Cross-sectional study	Neonates of black descent	Kramer's method	Total serum bilirubin (TSB)	Prevalence of neonatal jaundice: 49.8% by Kramer's method

Study Details	Population	Index test(s)	Reference Standard	Outcomes
June 2020 – July 2020 Tanzania			Cut-offs: unclear Presence of jaundice	63.5% by TSB sensitivity 70.5, specificity 86.1, PPV 89.8, and NPV 62.6%, +LR 5.07 and –LR 0.34. Diagnostic accuracy of Kramer's method 76.1% and a moderate agreement with TSB (cohen kappa κ = 0.524, P<0.001)
Bhutani 2019 Prospective cohort study. Bangladesh	Neonates >2000g.	visual icterometer ("Bili-ruler")	total serum bilirubin (TSB) Cut-offs: TSB ≥11 mg/dL, TSB >17 mg/dl	The visual Bili-ruler performed well compared with TSB ≥11 mg/dL, 84.5% (95% CI, 79.1%–90.3) and 83.2% (95% CI, 76.1%–90.3%), for sensitivity and specificity, respectively, and 5.04 (3.29–7.71) and 0.184 (0.126–0.268) for positive and negative likelihood ratios, respectively. For TSB >17 mg/dl, Bili-ruler performed moderately well, 87.8 (95% CI, 80.9–95) and 66.5 (95% CI, 59.6–73.3), for sensitivity and specificity, respectively, and 4.91 (3.53–6.83) and 0.224 (0.131–0.382) for positive and negative likelihood ratios, respectively.
Olusanya et al. 2017 Nigeria	2492 mother-infant pairs	Two-color icterometer **(Bilistrip™)**	Total bilirubin in serum (TSB) and transcutaneous bilirubin (TcB) Cut-offs: TcB thresholds ≥10mg/dL, ≥12mg/dL, ≥15mg/dL, ≥17mg/dL	347 (13.9%) were Dark Yellow For TcB thresholds (≥10mg/dL, ≥12mg/dL, ≥15mg/dL, and ≥17mg/dL). Bilistrip™ showed increasing sensitivity (47.0% - 92.6%) and negative predictive value (NPV) (91.4% - 99.9%). Among neonates with TSB measurements (n = 124), Bilistrip™ was associated also with increasing sensitivity (86.8% - 100%) and NPV (62.5% - 100%).

Study Details	Population	Index test(s)	Reference Standard	Outcomes
Kittiarpornpon et al. 2020 Prospective study Bangkok	180 mothers	Maternal visual assessment using infants' palm skin colour (dermal icterus zones)	Total bilirubin in serum (TSB) and transcutaneous bilirubin (TcB) Cut-offs: hyperbilirubinaemia ≥239.4 µmol/L (14 mg/dL) or requiring phototherapy	Detecting hyperbilirubinaemia requiring phototherapy: Sensitivity (95% CI): 91.7% (73.0–99.0) NPV (95% CI): 96.6% (87.9–99. Identifying hyperbilirubinaemia: Sensitivity: 92.9% (76.5–99.1) NPV: 96.6% (87.9–99.1) The accuracy of maternal report of dermal zones for serum bilirubin levels was only 44.5%.
Singh et al. 2022 Prospective comparative diagnostic study [India]	188 samples from 134 unique patients	"Color Card" initially by yellow color shades that fall into 4 bilirubin categories, i.e. TSB up to 7 mg/dl, 7.1 to 12 mg/dl, 12.1 to 18 mg/dl and >18 mg/dl	total serum bilirubin (TSB) by diazo method. Cut-offs: <7 mg/dl and >18 mg/dl.	The specificity, negative predictive value and accuracy of the color card for the observations made by observer 1 comparing with lab TSB were >95% for clinically important categories of <7 mg/dl and >18 mg/dl. The overall accuracy of color card in measuring various TSB ranges varied from 75% to 96.8%. The agreement between two observers was 85.6% (Cohen's kappa co-efficient: 0.61, *p*-value: .0001) overall and was 92.3%, 86%, 84%, 81.2% for each of the four bilirubin categories in ascending order.
Studies reporting diagnostic accuracy for Kejian 8000 (KJ-8000) in darker skin tones				
Afjeh et al. 2015 Prospective cross-sectional study Tehran	613 neonates weighing ≥ 1,800 g with gestational age of ≥ 35 weeks	Transcutaneous bilirubin test (TcB) **Kejian 8000 (KJ-8000)**	Total serum bilirubin (TSB) Cut-offs:	491 (80%) revealed high TcB 398/491 neonates revealed high total serum bilirubin (TsB)

Study Details	Population	Index test(s)	Reference Standard	Outcomes
			Only in those with higher TcB, TsB ≥ 5 mg/dL (not clear)	TcB has 81% positive predictive value (PPV) in diagnosis of hyperbilirubinemia. Correlation of TcB and TsB in highest rate is equal to 72% (P value < 0.001)
Studies reporting diagnostic accuracy for Draeger JM 103 in darker skin tones				
Gunaseelan et al. 2017 India	Neonates of gestational age more than 35 weeks and weighing more than 2 kg (icteric and late preterm babies) – 400 paired measurements	Transcutaneous bilirubin (TcB) **Draeger JM 103**	Total serum bilirubin (TSB) was measured if the initial TcB level was higher than the 50th centile in Bhutani's nomogram Cut-offs: Low-risk, medium-risk and high-risk for phototherapy	TcB was significantly correlating with TSB (P < 0.001) in both low-risk and medium-risk thresholds for phototherapy. TcB had a sensitivity and negative predictive value of 100% each, a specificity of 56%, and a positive predictive value of 23%. For high-risk cases, using the 75th centile as cutoff, the sensitivity and negative predictive value were reduced to 88% and 97.0%, respectively.
Villanueva-Uy et al. 2022 Philippines	1,412 stable, full-term infants (≥37 weeks age of gestation)	transcutaneous bilirubin (TcB) levels were determined at the 3rd, 6th, 12th, 24th, 36th, 48th, 72nd, 96th, and 120th hour of life (HOL) **Dräger-Minolta JM-103**	Simultaneous measurement of TcB and total serum bilirubin (TsB) on a subset of 106 infants Cut-offs: unclear	Correlation coefficients were high between TsB and forehead TcB (r^2 = 0.88), and between TsB and sternal TcB (r^2 = 0.91).
Chimhini et al. 2018 01 August and 30 November 2015	283 infants	Paired transcutaneous (forehead and sternum)	Paired serum bilirubin measurements	Correlation between serum and transcutaneous bilirubin (sternum): 0.77

Study Details	Population	Index test(s)	Reference Standard	Outcomes
Zimbabwe	Median gestational age was 38 weeks (range 28–42) Median postnatal age was 3 days (range 0–10). 115 preterm	**Draeger JM 103**	Cut-offs: unclear	Correlation between serum and transcutaneous (forehead):0.72. Preterm babies correlation for sternum: 0.77 forehead: 0.75. Term babies correlation for sternum: 0.76 Forehead: 0.70 Bland-Altman plot of serum versus transcutaneous measurements showed agreement between the tests. The ROC curves showed that the accuracy of the two diagnostic tests were good with no significant difference between the two, p = 0.2954. The sensitivity for the sternum site was 76%, specificity 90%, (PPV: 70 and NPV: 92) Sensitivity for forehead site was 62%, specificity 95% (PPV 80 and NPV 90)
Shihadeh et al. 2016 Prospective study Bahrain	88 newborns 128 paired measurements	Transcutaneous bilirubin (TcB) **Dräger JM 103™ device**	Simultaneous total serum bilirubin (TSB) Cut-offs: not given	The correlation between paired measurements were 0.75 (p<0.0005). The mean difference was 1.09 SD 2.16mg/dL (ranging from 6.18 to 7.00)
Studies reporting diagnostic accuracy for Drager JM 105 in darker skin tones				
Sharma et al. 2022 Govt. RDBP Jaipuria Hospital	120 Newborn babies up to the 10th postnatal day of life with visually found jaundice	TcB was measured over mid-sternum **Dräger JM 105™ device**	Simultaneous total serum bilirubin (TSB) measurements	Pearson's correlation coefficient was 0.892 (p<0.001). The average error in evaluating hyperbilirubinemia with TcB compared to TSB was 0.101, with limits of agreement between −3.73 and +3.55v(Bland-Altman analysis). The AUOC at three TSB levels (>10 mg/dl, >12 mg/dl, and >15 mg/dl) was 0.860, 0.892, and 0.849.

Study Details	Population	Index test(s)	Reference Standard	Outcomes
Mohamed et al. 2022 cross-sectional study Malaysia	130 jaundiced neonates requiring serum bilirubin determination from day 2 to day 7 of life. (Malay neonates)	Transcutaneous bilirubin (TcB) **Dräger Jaundice Meter JM-105**	Total serum bilirubin (TSB) Cut-offs: 205 µmol/L	TcB underestimates TSB with a mean difference of 10.10 µmol/L at the forehead and 9.27 µmol/L at the sternum. A positive linear relationship was observed between TSB with TcB forehead (r = 0.82) and TcB sternum (r = 0.80). A good discriminations ability was observed for both the TcB forehead (receiver operating characteristics [ROC] curve = 89.8%) and sternum (ROC curve = 89.7%) at a TSB level of 205 µmol/L. The sensitivity ranges from 84.4% to 85.3%, while the specificity ranges from 77.4% to 76.4%.
Study reporting diagnostic accuracy for BiliChek in darker skin tones				
Alsaedi 2016 Prospective cohort study Jan 2009 – Dec 2012 Saudi Arabia	665 newborns Mean age 44.2 +/- 21 hour.	transcutaneous bilirubin test (TcB) **BiliChek®**	Total serum bilirubin (TSB) Paired values of TcB and TSB Cut-offs: Unclear (paired values of TcB and TSB)	Mean TSB:147 +/- 45 mumol/L Mean TcB: 156 +/- 50 mumol/L Correlation TcB and TSB (r: 0.84; 95% [CI] = 0.82-0.86; p<0.001) The TcB tends to overestimate TSB The TcB was sensitive, but less specific. The TcB sensitivity was 83% and specificity was 71% to predict TSB during the first 72 hours of life for the whole study group. PPV: 63%, NPV: 87%.
Studies reporting diagnostic accuracy for Bilistick system in darker skin tones				
Greco et al. 2018 April 2015- November 2016	1458 newborns	point-of care **Bilistick System (BS)**	Total serum bilirubin (TSB)	TSB level measured by BS agreed (p < .0001) with the lab result in all four countries.

Study Details	Population	Index test(s)	Reference Standard	Outcomes
17 hospitals from Nigeria, Egypt, Indonesia, and Vietnam			Cut-offs: Unclear	The diagnostic performance of BS showed a positive predictive value (PPV) of 92·5% and a negative predictive value (NPV) of 92·8%.
Studies comparing babies with darker skin tone versus lighter skin tone				
Maya-Enero et al. 2021 prospective, observational study Spain	1359 newborns were assigned to a color group at 24 h of life according to Neomar's skin color scale which has four categories: light (color 1)=337, medium-clear (2)= 750, medium-dark (3)=249, and dark (4)=23	Transcutaneous bilirubin (TcB) **Drager JM 105™**	Serum bilirubin (SB) by calorimetric method by diazotation Cut-offs: Unclear (paired TcB/SB measurements)	Correlation between TcB and serum bilirubin was very good (R2 = 0.908–0.956), globally and by color group, with slight differences between darker and lighter skin colors. Pearson correlation coefficient for color 1 was 0.935 (95% CI 0.921; 0.947), for color 2 0.924 (95% CI 0.913; 0.933), for color 3 0.908 (95% CI 0.887; 0.926), and for color 4 0.956 (95% CI 0.914; 0.978) Bland-Altman biases increased with the color scale, from − 0.70 (95% CI − 3.82;2.42) for color 1 to − 1.08 (95% CI − 3.98;1.82) for color 2, and until −1.89 (95% CI − 5.09;1.30) and − 1.86 (95% CI − 5.11;1.38) for colors 3–4, respectively. The study not only supports the reliability of TcB to assess SB regardless of skin color, but also supports the fact that TcB tends to overestimate SB in a higher degree in dark-skinned neonates.
Starowicz et al. 2019 Prospective study Nov 2015 – July 2017 Australia	201 infants (416 paired samples) with different ethnicity and gestational age	Transcutaneous bilirubin (TcB) **Kejian 8000 (KJ-8000)**	Serum bilirubin (SBR) Using ABL90 FLEX blood gas analyser or End-Point Diazo assay Spectrophotmetric Assay	There was a strong correlation between TcB and SBR with a Pearson correlation coefficient of 0.8 (<0.00001). Caucasian group: r=0.84 Non-Caucasian group r=0.71

Study Details	Population	Index test(s)	Reference Standard	Outcomes
	Caucasian origin =76 Non-Caucasian = 24		Cut-offs: Unclear (paired serum bilirubin and TcB)	The bias was −5.9 µmol/L (95% CI: −101, 89) (Bland Altman) The bias was not evenly spread, with TcB tending to overestimate at lower SBR levels and underestimate at higher SBR levels. Infants <32 weeks' gestation had a poor correlation of 0.48. Non-Caucasian infants were more likely to have TcB overestimation, and measurements were less precise. As a screening tool using local guidelines, the KJ-8000 had a sensitivity, specificity, positive predictive value and negative predictive value of 83, 53, 20 and 96%, respectively, and is predicted to avoid blood tests in 48% of infants screened.