NICE National Institute for Health and Care Excellence

2023 exceptional surveillance of jaundice in newborns under 28 days (NICE guideline CG98)

Surveillance report

Published: 31 October 2023

www.nice.org.uk

Contents

Surveillance decision

We will change:

- Recommendation 1.2.5 to include text to acknowledge that jaundice may be harder to detect visually in darker skin tones.
- Recommendation 1.7.1 to be more flexible around urine culture.

No further areas of the guideline will be updated at this time.

This surveillance proposal does not impact on the previous decision to update the NICE guideline on total serum bilirubin (TSB) thresholds for starting phototherapy or exchange transfusion in all term babies with neonatal hyperbilirubinaemia due to variability between assays from different manufacturers, please see the exceptional surveillance report from April 2023.

Reason for the surveillance review

This exceptional surveillance review was triggered by the Health Services Safety Investigations Board (HSIB) report on the detection of jaundice in newborn babies following an incident of delayed diagnosis due to visual inspection not detecting clinical jaundice in a premature newborn of black African ethnicity.

While considering the impact of the HSIB report, 2 further triggers were identified: NHS Race and Health Observatory review of neonatal assessment and concerns from the Royal College of Paediatrics and Child Health (RCPCH) as to the extent to which the NICE guideline covers conjugated Jaundice.

HSIB report

The HSIB report details a baby who was initially assessed as a clinically stable premature baby. Subsequently, a routine blood sample taken 2 hours after birth indicated that bilirubin was present and the level was recorded. The clinical report was seen by 1 of the staff members, who observed the high level of bilirubin and possible need for treatment. This member of staff was then called to attend an emergency and no action was taken.

The baby was subsequently examined by the clinical team and was not noted to be jaundiced. Routine observations such as temperature, heart rate and skin colour were documented at approximately 2-hourly intervals. No visual signs of jaundice were documented.

Another blood sample was taken when the baby was 2 days old, again showing a high level of bilirubin, but the result was not acted upon. In the next 2 days, no visual signs of jaundice were documented.

At 5 days of age, a change in skin colour and visual signs of jaundice were documented. Bilirubin levels were checked and found to be high. Phototherapy was commenced. Over the next 3 days, bilirubin levels returned to acceptable levels and the baby was discharged home.

The HSIB report highlighted the issue of delay in the diagnosis of jaundice in newborns with black or brown skin-colour, in whom visual signs were not noted. This was attributed to the darker skin tone of the infant. The HSIB report also highlighted there is a need to explore other risk factors of jaundice including prematurity, to enable early identification.

As babies with darker skin tones were a significant subgroup in the 2016 update of the NICE guideline, this exceptional surveillance review also investigates whether there is new evidence on the accuracy of various tests (clinical history and examination, urine/stool examination, icterometer and transcutaneous bilirubin [TcB] levels) in recognising neonatal jaundice or hyperbilirubinaemia in the subgroup of babies with darker skin tones (review question 3), in whom initial visual inspection could be difficult or unreliable.

NHS Race Health Observatory report

In July 2023, the NHS Race Health Observatory published a report into findings from a formal review of neonatal assessments and the potential for disadvantaging darker skin tones due to practice based on White European babies.

The report recommends that guidelines are reviewed with respect to neonatal assessment by skin colour specifically for jaundice to draw attention to differences in assessment techniques for Black, Asian and ethnic minority neonates.

RCPCH concerns

The RCPCH contacted NICE about challenges in the screening and assessment of conjugated jaundice and the impact of the NICE guideline recommendations on practice. The British Society of Paediatric Gastroenterology, Hepatology and Nutrition (BSPGHAN) carried out a clinical audit involving 14 NHS trusts exploring the screening pathway for prolonged jaundice and concluded that the NICE guidelines do not sufficiently distinguish between early and prolonged jaundice which may inadvertently be leading to variable practices in the diagnostic work up of babies over 14 days of age with jaundice symptoms.

Exceptional surveillance review summary

Methods

The exceptional surveillance process consisted of:

- Considering the HSIB report on detection of jaundice in newborns (January 2023) and the NHS Race Health Observatory report.
- Literature searches on risk factors of hyperbilirubinaemia, accuracy of visual assessment and various tests (clinical history and examination, urine/stool examination, icterometer and TcB levels) in identifying neonatal jaundice in babies with darker skin tones.
- Assessing the new evidence and information against current recommendations to determine whether or not to update sections of the guideline, or the whole guideline.
- Reviewing the BSPGHAN audit document highlighted by RCPCH.

For further details about the process and the possible update decisions that are available, see ensuring that published guidelines are current and accurate in developing NICE guidelines: the manual.

Search and selection strategy

The searches retrieved 1,330 studies for risk factors of hyperbilirubinaemia and 1,730 studies for accuracy of visual assessment and various tests in identifying neonatal jaundice in babies with darker skin tones, published between February 2015 and February 2023. Searches for accuracy of visual assessment and various tests were based on the

search strategies previously used in 2015 and can be found in appendix D in the original evidence review as review question 3.

No review protocol was available for the risk factors question in the original 2010 guideline. Hence, the following criteria were used to select studies on risk factors:

- Prospective or retrospective cohort studies (with a control group); case-control studies.
- Any variables that defined by the studies as risk factor.
- Hyperbilirubinaemia as the outcome of interest (defined by TSB or needing phototherapy).
- Multivariate analysis was conducted in the studies with confounders adjusted.

The review protocol of review question 3 in the 2016 update was used to select studies on accuracy of visual assessment and various tests/examinations in identifying hyperbilirubinaemia in babies with darker skin tones.

See appendix A for details of all evidence considered, and references.

Information considered in this exceptional surveillance review

Based on the selection criteria, 1 study is included for risk factors of hyperbilirubinaemia (recommendation 1.2.1) and 16 studies are included on accuracy of visual assessment and various tests/examinations in identifying hyperbilirubinaemia in babies with darker skin tones (5 studies on visual assessment [recommendation 1.2.5] and 11 studies on TcB using various devices [consensus-based threshold table]).

Risk factors for all babies

Studies on risk factors for hyperbilirubinaemia

One study was included looking at risk factors for hyperbilirubinaemia. Hamadneh et al. (2016) compared 2 groups of women who were delivered by caesarean section and who have had at least 2 previous caesarean sections. Group 1 (n=505) was women who

underwent caesarean section at 37 weeks and group 2 (n=381) was women who underwent caesarean section at 38 weeks or later. Results from multivariate analysis showed that neonatal jaundice was more prevalent in group 2 in which babies were delivered at 38 weeks or later (adjusted odds ratio 2.1, 95% confidence interval [CI] 1.7 to 2.7; p=0.035).

Limitations of the identified evidence

This study focused on babies who were delivered by caesarean section and whose mothers have had at least 2 caesarean sections previously. The study compared delivery on week 37 with weeks 38 or later. The sample size was moderate in both groups. However, as the study population was specific to women who underwent caesarean it is not likely to be generalisable to a broader population with heterogenous demographics beyond caesarean delivery. The study included retrospective data collection with multivariate analysis which may be subject to selection bias.

Impact statement

Recommendation 1.2.1 in the NICE guideline states that, gestational age under 38 weeks, a previous sibling with neonatal jaundice, mother's intention to breastfeed exclusively and visible jaundice in first 24 hours of life as risk factors for hyperbilirubinaemia. The guideline recommends gestational age of under 38 weeks as a risk factor whereas evidence published by Hamadneh et al. (2016) reported gestational age of 38 weeks and above delivered by caesarean in women with 2 previous caesarean deliveries. This is inconsistent with the existing recommendation. However, it is worth noting that the evidence originates from 1 study only with limitations and hence it is not sufficient to trigger an update for the recommendation.

Identifying hyperbilirubinaemia in babies with darker skin tones

In total, 16 studies were identified exploring identification of hyperbilirubinaemia in babies. 5 studies investigated accuracy of visual assessment of jaundice in babies with 'assumed' darker skin tones compared to TSB (Singh et al. 2022; Olusanya et al. 2017; Kittiarpornpon et al. 2020; Dionis et al. 2021; Bhutani 2019). 10 studies investigated accuracy of TcB using various devices compared with TSB or serum bilirubin (Starwociz et al. 2019; Maya-Enero et al. 2021; Afjeh et al. 2015; Alsaedi 2016; Gunaseelan et al. 2017; Sharma et al. 2022, Mohamed et al. 2022; Shihadeh et al. 2016; Vilanueva-Uy et al. 2022; Chimhini et al. 2018). One study explored the Bilistick system to estimate TSB (Greco et al. 2018).

Visual assessment of jaundice

Five studies used visual assessment for determining hyperbilirubinaemia, these included the use of visual icterometer (Bili-ruler, Bhutani 2019), Kramer's method (Dionis et al. 2021;), maternal visual assessment (Kittiarpornpon et al. 2020), colour card (Singh et al. 2022) or two colour icterometer (Bilistrip, Olusanya et al. 2017). These studies were carried out in Bangladesh, Tanzania, Nigeria, Thailand and India. We have assumed the majority of the babies in these studies have darker skin tones.

Bhutani (2019) reported the use of visual icterometer (Bili-ruler) at TSB ≥11 mg/dL, with 84.5% (95% CI, 79.1% to 90.3%) and 83.2% (95% CI, 76.1% to 90.3%), sensitivity and specificity, respectively. For TSB >17 mg/dl, Bili-ruler performed moderately well, 87.8 (95% CI, 80.9 to 95) and 66.5 (95% CI, 59.6 to 73.3), for sensitivity and specificity, respectively.

Olusanya et al. (2017) utilised two colour icterometer; dark yellow and light yellow, (Bilistrip). The mean TcB in infants with dark yellow was significantly higher than those with light yellow. Bilistrip™ was associated with increasing sensitivity (47.0% to 92.6%) and NPV (91.4% to 99.9%) against increasing TcB thresholds (≥10mg/dL, ≥12mg/dL, ≥15mg/dL, and ≥17mg/dL; n=124).

Kittiarpornpon et al. (2020) compared maternal visual assessment with TSB for detecting hyperbilirubinaemia requiring phototherapy (sensitivity, 91.7%, 95% CI: 73.0 to 99.0) and for identifying hyperbilirubinaemia (sensitivity, 92.9% 95% CI: 76.5 to 99.1).

Singh et al. (2022) used 'Colour Card' initially by yellow colour shades that fall into 4 bilirubin categories, for example, TSB up to 7 mg/dl, 7.1 to 12 mg/dl, 12.1 to 18 mg/dl and >18 mg/dl. The overall accuracy of colour card in measuring various TSB ranges varied from 75% to 96.8%. The agreement between 2 observers was 85.6% (Cohen's kappa co-efficient: 0.61, p=0001) overall and was 92.3%, 86%, 84%, 81.2% for each of the 4 bilirubin categories in ascending order.

Kramer's method was used by Dionis et al. (2021) in babies of black descent and showed sensitivity of 70.5% and specificity of 86.1%.

Limitations of the studies on visual assessment

One of the limitations of this evidence was that it is 'assumed' the study populations have darker skin tones as the studies were conducted in countries with majority of population

as non-Caucasians. This assumption may not be true. Another limitation was that different devices or methods and different thresholds for TSB were used across the 5 studies to conduct visual assessment and therefore no direct conclusions can be drawn based on these different methods and different TSB thresholds. Also, a key limitation of these studies was that there was no comparator group of babies with lighter skin tones, hence it is impossible to determine the utility of visual assessment to identify jaundice in babies in darker skin tones and how it differs to babies with lighter skin tones. All the included study populations had high TSB, clinical opinion would indicate that jaundice would be visible at 5mg/dL and as such limit the external validity of findings from these studies to the UK context.

Impact statement for visual assessment studies

Due to variations in the studies, it is unclear whether the new evidence challenges the existing recommendations. All the studies were carried out in babies assumed to have darker skin tones; however, absence of comparator group with babies who have lighter skin tones makes it impossible to conclude that the accuracy of visual assessment will vary depending on skin tone based on this evidence. Therefore, the evidence does not support an update on existing recommendations.

Transcutaneous bilirubinometers to assess jaundice

11 studies investigated the accuracy of TcB in comparison to TSB using various devices (KJ-8000, Drager JM103, Drager JM 105, BiliChek, Bilistick).

Nine of these studies were conducted on babies with 'assumed' darker skin tones (conducted in countries with majority of population as non-Caucasians), with no comparator group of babies with lighter skin tones. Two studies compared babies with darker skin tones versus lighter skin tones.

Kejian 8000 device

One study used Kejian 8000 to determine TcB in Tehran (Afjeh et al. 2015).

The study was a prospective cross-sectional study that included 613 babies with gestational age of 35 weeks or above. It reported high TcB in 491 babies (those with TcB ≥5 mg/dL in first 24 hours and >8 mg/dL in second 24 hours). Serum sample was taken from those 491 babies with high TcB, and 398 out of 491 babies also had high TSB, with a

correlation of 72% (p<0.001).

The study selectively checked TSB in babies with high TcB in the first and second 24 hours and reported positive predictive values of 81% in diagnosis of hyperbilirubinaemia. The study does not suggest what TcB threshold is most accurate against TSB for babies with 'assumed' darker skin tones, or whether TcB differs compared to babies with lighter skin tones. These findings are not conclusive nor sufficient to determine the utility of KJ-8000 on babies with darker skin tones.

Drager JM 103 device

Four studies compared TcB measured by Drager JM 103 with TSB in babies with 'assumed' darker skin tones (Gunaseelan et al. 2018; Villanueva-Uy et al. 2022; Chimhini et al. 2018; Shihadeh et al. 2016). Two studies included babies of gestational age between 35 weeks and 37 weeks (Gunaseelan et al. 2018; Villanueva-Uy et al. 2022). One study included 283 babies with median gestational age of 38 weeks (25 to 42 weeks), of which 115/283 (41%) were preterm babies (Chimhini et al. 2018). The fourth study did not report gestational age (Shihadeh et al. 2016). These studies were carried out in India, Philippines, Bahrain and Zimbabwe, where the majority of populations are 'assumed' to have darker skin tones.

Three studies compared paired/simultaneous measurements between TcB and TSB (Gunaseelan et al. 2018; Villanueva-Uy et al. 2022; Shihadeh et al. 2016). The correlation measures between TcB (sternum) and TSB was 0.91, and between TcB (forehead) and TSB 0.88 (Villanueva-Uy et al. 2022). A study by Gunaseelan et al. (2018) did not report the correlation value for paired measurements of TcB and TSB but a statistical significance level of p<0.001 was reported for TcB at low-risk and medium-risk thresholds of phototherapy. Shihadeh et al. (2016), reported correlation between paired measurements of TcB and TSB as 0.75 (p<0.0005) with a mean difference of 1.09±2.16 mg/dL (range of differences: 6.18 mg/dL to 7.00 mg/dL; Shihadeh et al. 2016).

Chimhini et al. (2018) also carried out subgroup analysis of correlation for those 115 preterm babies, with the correlation of 0.77 for sternum TcB and 0.70 for forehead TcB with TSB. For the other 168 term babies, correlation was 0.76 for sternum and 0.70 for forehead. Bland-Altman plot showed agreement between TcB and TSB in this study (no further details from abstract). The other two studies did not report Bland-Altman results.

Overall, the evidence identified using the Drager JM 103 device is unclear, inconsistent and does not suggest a clear threshold of TcB for babies with 'assumed' darker skin tones

in comparison to babies with lighter skin tones. The populations are of heterogenous gestational age and paired correlations were sometimes only available for subsets of babies.

Drager JM 105 device

Two studies used Drager JM-105 to determine the accuracy of TcB against TSB in India and Malaysia (Sharma et al. 2022; Mohamed et al. 2022). The correlation between TcB and TSB was 0.892 ($p<0.001$; Mohamed et al. 2022). Sharma et al. (2022) reported correlation of TSB separately for forehead TcB and sternum as 0.82 and 0.80 respectively.

The average error in evaluating hyperbilirubinemia with TcB compared to TSB was 0.101, with limits of agreement between minus 3.73 and plus 3.55 (Sharma et al. 2022). Mohamed (2022) also reported that TcB underestimates TSB with a mean difference of 10.10 µmol/L at the forehead and 9.27 µmol/L at the sternum.

The area under the curve at 3 TSB levels (>10 mg/dl, >12 mg/dl, and >15 mg/dl) was 0.860, 0.892, and 0.849 (Sharma et al. 2022). A good discriminations ability was observed for both the TcB forehead (ROC curve = 89.8%) and sternum (ROC curve = 89.7%) at a TSB level of 205 µmol/L (Mohamed et al. 2022).

Both studies were carried out in babies with 'assumed' darker skin tones and neither included a comparator by skin tone. Overall, the sample size was relatively small in both studies (120 and 130 babies). Sharma (2022) included babies with visually identifiable jaundice and the other study had babies with jaundice that required TSB determination (Mohamed et al. 2022). Potential heterogeneity in outcomes arises from timing of assessment, gestational age and variation in TSB thresholds which are not aligned with UK clinical practice. Both studies reported an underestimation of 10µmol/L for accuracy of TcB against TSB.

BiliChek device

One study compared TcB measured by BiliChek with TSB in jaundiced term (37 to 42 weeks) babies (n=665) in Saudi Arabia (Alsaedi, 2016). Paired values of TcB and TSB indicated a correlation of r=0.84 (CI: 0.82 to 0.86; $p<0.001$). The results showed that TcB overestimated TSB with a mean difference of 17 µmol/L, 95% CI of 40 ± 77 µmol/L. The sensitivity was 83% and specificity 71%. Again the study did not report a specific cut-off of TSB but indicated that bilirubin level at low and above intermediate risk zone was

considered significant (definition of low and intermediate risk zone was not reported in abstract). The findings are only reported in (assumed) darker skin tones babies and not compared with lighter skin tone babies, therefore it is uncertain whether the overestimation of TcB will apply specifically to babies with darker skin tones.

Bilistick system

One study used the Bilistick system to measure TcB in 1,458 babies in 17 hospitals from Nigeria, Egypt, Indonesia, and Vietnam (Greco et al. 2018). TSB level measured by Bilistick system correlated well with the lab result in all four countries, with positive predictive value (PPV) of 92.5% and a negative predictive value (NPV) of 92.8%. As with previous studies, it was 'assumed' that babies were with darker skin tones based on the country where data was collected and assessed. Results were not stratified based on skin tones or ethnicity. No other findings were reported. Due to the limitations of the study, especially reporting bias, the evidence on the Bilistick system is not sufficient to conclude its accuracy.

Studies comparing babies with darker skin tone versus lighter skin tone

Two studies which compared TcB with TSB in different skin tone populations were identified, Maya-Enero et al. (2021) used the Drager JM105 device for TcB and Starwociz et al. (2019) used the KJ 8000 TcB device.

Maya-Enero et al (2021), included 1,359 babies who were assigned to different groups at 24 hours of life according to Neomar's skin colour scale (Maya-Enero et al. 2020) into 4 categories: light (colour 1) n=337, medium-clear (2) n=750, medium-dark (3) n=249, and dark (4) n=23. Correlation between TcB and TSB range was reported as r^2 = 0.908 to 0.956, with slight differences between darker and lighter skin tones. Correlation for colour 1 was 0.935 (95% CI 0.921; 0.947), for colour 2 was 0.924 (95% CI 0.913; 0.933), for colour 3 was 0.908 (95% CI 0.887; 0.926), and for colour 4 was 0.956 (95% CI 0.914; 0.978). Bland-Altman plots also showed differences in mean bilirubin bias depending on skin colour with the bias increasing from colour 1 to colour 4. The difference increased gradually from colours 1 to 4 and was statistically significantly different between colours 1 and 2, between colours 1 and 3, between colours 1 and 4, and between colours 2 and 3, and between colours 2 and 4 (p value <0.001), but not between colours 3 and 4.

In a multiple linear regression, serum bilirubin was associated with TcB (β=0.88, 95% CI 0.87; 0.90, p <0.001). Considering colour 1 as reference, β coefficients for colour 2 were β

= -0.22, (95% CI -0.40 to -0.04), p=0.019; for colour 3 β= -0.81, (95% CI -1.04 to -0.58), p <0.001; and for colour 4 β= -0.80, (95% CI -1.32 to -0.28), p=0.003. Bland Altman plot for all skin colours showed that generally TcB underestimated TSB for TcB >15 mg/dl (for TcB 15: LoA, 95% CI -1.73 [-6.02; 2.55]) but overestimated it for TcB ≤15 mg/dL (for TcB 15: LoA, 95% CI 1.06 [-3.87; 1.75]). The overestimation was greater in darker skin tones, with mean overestimation, 0.7 mg/dL for light; 1.08 mg/dL for medium light; 1.89 mg/dL for medium dark; and 1.86 mg/dL for dark; p<.001 for light versus medium dark or dark. This study concluded that TcB is more likely to overestimate TSB in darker skin toned babies compared with neonates with lighter skin tone.

Maya-Enero et al (2021) included only 23 babies in the dark skin group (colour 4), and Bland-Altman results may not be comparable between colour 4 and colours 1, 2, and 3. The study included 1,549 paired SB/TcB measurements (379 for colour 1, 828 for colour 2, 308 for colour 3, and 34 for colour 4) and mentioned that some babies had more than 1 determination of SB/TcB but all paired measurements were included. This could result in duplication of data points and might affect estimates of outcomes. Only healthy babies were included in the study with mean gestational age of 39 weeks. This study is based in only 1 hospital and hence the results might only be relevant and representative of the population in the area.

Starwociz et al (2019) carried out a prospective study comparing TcB and serum bilirubin in Caucasian (n=76) and non-Caucasian (n=24) preterm (<32 weeks gestation) and term babies. Overall correlation was 0.8 (p<0.0001) between TcB and serum bilirubin levels. Correlation in term babies was r=0.82 and preterm babies r=0.49. The Caucasian group had stronger correlation r=0.84 compared to the non-Caucasian group r=0.71. Overall, the bias was -5.9 μmol/L (95% CI: -101, 89) which was not evenly spread, and TcB tending to overestimate at lower serum bilirubin levels and underestimate at higher levels of serum bilirubin. TcB was overestimated and less precise in non-Caucasian babies.

The study sample size was small and thus outcome measures may not be comparable. There was also underrepresentation of term babies compared to preterm babies and multiple readings were collected more from preterm babies. Due to such imbalances, the bias and precision might have been affected.

Overall impact statement (TcB devices)

NICE's guideline currently recommends visual inspection as the initial assessment approach for suspected jaundice, followed by dependent on risk measurement of serum

bilirubin or TcB. In light of the HSIB and NHS Race Observatory reports, this exceptional review explored whether TcB could play a role as initial assessment for babies with darker skin tones in whom visual inspection may be difficult or unreliable.

No conclusive evidence was available on accuracy of TcB using various devices in babies with darker skin tones. The evidence had a lack of comparator groups of babies with lighter skin tones. Additionally, different thresholds for TSB across studies and a number of studies did not report cut-offs used, limiting conclusions that can be drawn. It is worth noting that the study population of babies with darker skin tones was 'assumed' in these studies, given the countries where these studies were conducted.

Only 2 studies compared babies with darker versus lighter skin tones in the same setting. One used Drager JM 105 (Maya-Enero et al. 2021) and another used KJ 8000 (Starwociz et al. 2019) for TcB measurements. Maya-Enero et al. (2021) used Neomar's neonatal skin colour scale with 4 categories: light, medium-light, medium-dark and dark. These categories were defined solely based on skin colour regardless of ethnicity. Both studies reported a trend for overestimation of TSB in darker skin tones, however the evidence is of limited quality to allow definitive conclusions to be drawn on the role of TcB in identifying jaundice in babies with darker skin tones in whom visual inspection may be difficult or unreliable and thus at this point in time does not impact on the NICE guideline.

Prolonged Jaundice

The RCPCH highlighted that BSPGHAN carried out a clinical audit around diagnosis of consolidated jaundice. The audit involved 14 NHS trusts, and 746 babies. The average day of assessment was 21 days (11 to 56). There was heterogeneity in the investigations performed on the babies. Some trusts took a minimalist approach with a split bilirubin test. Some trusts included a urine sample and in many cases had sample contamination and so required repeat sampling. Some trusts were carrying out G6PD testing, AAT testing or Direct antiglobulin testing (coombs test). No babies were unwell babies or had serious complications. No trusts in the audit were performing clotting screening in their diagnostic work up. Hence the RCPCH are questioning if the pathway outlined in section 1.7 of the NICE guideline is overly complex in the management of prolonged jaundice. Concerns were particularly raised around the practicalities of obtaining urine samples from babies and the necessity in the clinical work up for prolonged jaundice.

Information considered in previous surveillance of

this guideline

In 2023, an exceptional surveillance review was conducted, triggered by 2 external communications relating to variability of bilirubin assays from different manufacturers. The communications suggested variations in TSH threshold values for newborns with hyperbilirubinaemia dependent on the assay used. The outcome of that surveillance review indicated that the NICE guideline would be updated. The update will focus on TSB thresholds for starting phototherapy or exchange transfusion in all term babies with neonatal hyperbilirubinaemia. This may involve updating the threshold table for the management of hyperbilirubinaemia.

The 2019 Surveillance review of the guideline identified no evidence in relation to care of babies with prolonged jaundice. At that time some topic experts questioned the threshold table validity post 14 days. Additionally, a topic expert highlighted hypothyroidism as a cause of prolonged jaundice and suggested that testing should be added. No evidence on this issue was identified but recommendation 1.7.1 was editorially refreshed to make reference to congenital hypothyroidism screening.

Information considered when developing the guideline

Risk factors of risk of hyperbilirubinaemia

This section falls under the guideline review question 'What are the factors associated with an increased risk of hyperbilirubinaemia' and forms recommendation 1.2.1, which was developed in 2010. During the development of this recommendation, 10 studies were included. Evidence from good-quality studies showed that 4 factors are independently associated with an increased risk of hyperbilirubinaemia; gestational age <38 weeks, jaundice within 24 hours of birth, increase in severity of clinically apparent jaundice and intention to breastfeed exclusively. Other factors such as cephalohaematoma, vacuum delivery, male sex or race showed no statistically significant association with hyperbilirubinaemia. The committee noted that it is commonly believed that bruising, cephalohaematoma and vacuum delivery all contribute towards development of hyperbilirubinaemia, but the evidence was inconclusive to include those as risk factors.

This recommendation has not been updated since 2010.

Assessment of hyperbilirubinaemia

The guideline review question 'What is the accuracy of various tests (clinical history and examination, urine/stool examination, icterometer and TcB levels) in recognising neonatal jaundice or hyperbilirubinaemia?' was updated in 2016. The evidence considered at that time comprised of 25 studies (with another 7 studies from the 2010 original guideline, total 32 included studies), but only 2 studies specified babies of different skin tones (Wainer 2009; Karen 2009).

The committee noted that the evidence for the review question was limited: 1 study on clinical history and examination, no evidence on urine and stool examination: no evidence on icterometer and 31 studies on TcB. The committee noted that there were a number of limitations in the evidence base; 15/32 studies screened all newborns despite absence of clinical symptoms which resulted in lack of generalisability; inappropriate reference method; unclear information on whether phototherapy was given before the study; unclear reporting of infant's postnatal age; high uncertainty of the accuracy estimates due to wide confidence intervals. In addition to that, a range of diagnostic measures were reported using different devices measuring TcB. The diagnostic devices used in the studies varies and hence generates heterogenous readings of TcB with no obvious differences in accuracies when compared to TSB. Due to these reasons, the committee noted that the evidence was not sufficient to inform recommendations on diagnostic accuracies of tests to determine hyperbilirubinemia in newborns.

The committee also noted that the evidence on babies with different skin tones was still limited to inform recommendations based on skin tones.

Care of babies with prolonged jaundice

The guideline recommendations on care of babies with prolonged jaundice was largely based on clinical consensus at the time of development. Two case series (1 from the UK and 1 from Turkey) described some investigation and demographic correlations in prolonged jaundice cohorts. The Committee recognised the importance of clinical examination and key investigations including total and conjugated bilirubin, urine culture and where appropriate G6PD testing.

See ensuring that published guidelines are current and accurate in developing NICE guidelines: the manual for more details on our consultation processes.

Stakeholder consultation

Stakeholders were consulted and asked if they agreed with our proposals to change,

- Recommendation 1.2.5 to include text to acknowledge that jaundice may be harder to detect visually in darker skin tones.
- Recommendation 1.7.1 to be more flexible around urine culture.

Eleven stakeholders commented on the proposed changes: 3 professional societies, 2 hospital trusts, 1 royal college, 4 charities and 1 consultant neonatologist. For recommendation 1.2.5, 8 stakeholders agreed with the changes, 1 disagreed and 2 stated 'non-applicable'. The stakeholder that disagreed with the changes commented on practice issue of a case report in HSIB, which is outside the scope of the guideline. For recommendation 1.7.1, 7 stakeholders agreed with the changes, 3 stated 'non-applicable' and 1 did not comment.

No additional evidence was submitted by the stakeholders that is relevant to the following topic areas of interest:

- differential diagnostic accuracies of bilirubin measurement in neonates of varying skin tones,
- the diagnostic value of urine culture in the management of prolonged jaundice,
- risk factors for neonatal jaundice.

Five stakeholders commented on a health inequalities issue, highlighted the NHS Race Health Observatory report, where it identified that Black, Asian and minority ethnic parents can experience difficulties in accessing or receiving care for themselves or their babies. The report is summarised and highlighted in this surveillance review and the health equality impact assessment.

See appendix B for details of stakeholder comments and responses.

Equalities

An equalities and health inequalities assessment was completed during this surveillance review. See appendix C for details.

Overall decision

Risk factors and diagnostics for neonatal jaundice

The HSIB report highlights concerns about risk factors for jaundice. Only 1 poor quality study was identified concerning risk factors for neonatal hyperbilirubinemia. As such, there is insufficient new evidence to update the recommendations.

Both the HSIB and NHS Race Observatory reports raise questions about the diagnostic pathway for suspected jaundice. NICE's guideline takes a stepped approach to diagnostic assessment from visual inspection to TcB or TSB dependent on the age and gestational age of the babies. The evidence identified in this review was inconclusive about the impact of skin tone on diagnostic accuracy of various approaches to detect jaundice/ measure bilirubin, largely due to studies not having comparator groups. At present the recommendations in the guideline are uniform regardless of skin tone due to lack of evidence in 2016. Therefore, it remains uncertain whether the visual assessment pathway entry point results in different outcomes in babies with darker skin tones compared to those with lighter skin tones.

As babies with darker skin tones was an identified subgroup in the 2016 update, this exceptional surveillance review also explored whether there is new evidence to suggest the diagnostic accuracy of TcB may vary depending on different skin tones compared with standard reference to TSB. Based on the evidence identified and assessed in this exceptional surveillance review, there is very limited evidence (2 studies) to suggest that TcB overestimates the TSB in babies with darker skin tones compared to those with lighter skin tones, although a conclusion cannot be drawn as to exactly how large this overestimation is. In conclusion there is currently insufficient new evidence to update the consensus-based bilirubin threshold table for a specific subgroup of babies with darker skin tones.

The HSIB report raised 2 areas of concern for NICE to explore, the reliability of visual signs to detect jaundice in newborns, particularly in babies with darker skin tones and risk factors for jaundice. The evidence identified in this surveillance review does not support changing the recommendations at this time. Given the overlap in concerns about case recognition potentially being variable dependent on skin tone raised by both NHS Race Observatory and HSIB we will add a bullet to recommendation 1.2.5 to highlight that visual inspection may be harder in darker skin tones. The new recommendation will read:

When looking for jaundice (visual inspection):

- check the naked baby in bright and preferably natural light
- examine the sclerae and gums, and press lightly on the skin to check for signs of jaundice in 'blanched' skin
- be aware that hyperbilirubinemia may be harder to see visually in darker skin.

Prolonged jaundice

The BSPGHAN audit highlights that the guideline recommendations in section 1.7 are not being implemented uniformly in NHS trusts included in the audit for the management of prolonged jaundice. When the recommendations were developed it was an evidence sparse area and searches were not performed as part of this surveillance review. The audit also highlights no impact on outcomes of differential implementation of investigations. The implementation challenges for urine culture raised by BSPGHAN are acknowledged and given the lack of evidence when the recommendations were developed, we will change the wording of recommendation 1.7.1 to reflect the implementation challenges of carrying out urine culture, the lack of evidence and concerns of resource and impact on patient and carer experience. The revised recommendation will read:

In babies with a gestational age of 37 weeks or more with jaundice lasting more than 14 days, and in babies with a gestational age of less than 37 weeks and jaundice lasting more than 21 days:

- look for pale chalky stools and/or dark urine that stains the nappy
- measure the conjugated bilirubin
- carry out a full blood count
- carry out a blood group determination (mother and baby) and DAT (Coombs' test). Interpret the result taking account of the strength of reaction, and whether mother received prophylactic anti-D immunoglobulin during pregnancy
- consider a urine culture if there is clinical suspicion of a urinary tract infection
- ensure that routine metabolic screening (including screening for congenital hypothyroidism) has been performed.

Additionally, as part of their work, the BSPGHAN has developed a simplified pathway related to the audit which the RCPCH are considering endorsing into a clinical guideline. Given the specialist society interest in this area we will not update the recommendations on care of prolonged jaundice beyond the changes regarding urine culture.

For further details and a summary of all evidence identified in surveillance, see appendix A.

ISBN: 978-1-4731-5500-8