ORIGINAL ARTICLE

Defining the limitations of transcutaneous bilirubin measurement in late preterm newborns

KL Fine[1], WA Carey[1], JAW Schuster[1], SC Bryant[2], WJ Cook[1] and BS Karon[3]

OBJECTIVE: The objective of this study is to determine the impact of postnatal age on the bias between transcutaneous (TcB) and total serum bilirubin (TSB), and evaluate a TcB screening protocol.

STUDY DESIGN: Preterm and term infants had paired TcB and TSB performed on days 1 to 3 of life; a subset of preterm infants had measurements on days 4 to 7. Sensitivity and specificity of TcB (plotted on an age-specific TSB nomogram) for prediction of high-intermediate (HIR) or high-risk TSB were calculated.

RESULTS: Median TcB bias was 2.6 and 2.5 mg dl^{-1} for term and preterm infants in the first 3 days of life, respectively. However, median bias was 2.2 mg dl^{-1} for preterm infants at 4 to 7 days of life. TcB in preterm infants predicted HIR or high-risk TSB with 94% sensitivity and 56% specificity.

CONCLUSION: TcB screening protocols developed for term infants can be used for late preterm infants in the first 3 days of life.

Journal of Perinatology (2017) **37,** 658–661; doi:10.1038/jp.2017.8; published online 16 February 2017

INTRODUCTION

Jaundice is a common and usually benign finding in newborns. However, severe consequences of hyperbilirubinemia, such as acute or chronic bilirubin encephalopathy (kernicterus), can result if left untreated. Neurologic consequences of hyperbilirubinemia in newborns can almost always be prevented by early detection and treatment. As a result, the American Academy of Pediatrics Subcommittee on Hyperbilirubinemia recommends that all newborns be assessed for hyperbilirubinemia before discharge from the birth hospital with either a total serum (TSB) or transcutaneous (TcB) bilirubin measurement.[1]

TcB screening is appealing given that it provides rapid results and does not require a blood draw. Research suggests that TcB is a useful tool in healthy term newborns, even in 'real world' settings removed from ideal study conditions, during the first several days of life.[2] Evidence also points to a different correlation between paired measurements of TSB and TcB in healthy term newborns, depending on when the paired measurements are performed (that is, during birth hospitalization versus at 4 to 7 days of life in the outpatient setting).[3] However, the effect of postnatal age on the TSB–TcB correlation has not been studied in late preterm infants, a population that often remains in the hospital setting through the first week of life.

The primary objective was to examine the relationship between TcB and TSB in late preterm infants compared with a healthy term infant population. We hypothesized that a TcB screening algorithm developed based upon the observed relationship between TcB and TSB in healthy term infants could safely and effectively be applied to late preterm infants in a level II nursery. The secondary objective was to characterize how the TcB–TSB relationship changes as a function of postnatal age in late preterm infants. We hypothesized that there would be an inverse relationship between TcB bias (defined as TcB–TSB difference) and postnatal age among late preterm infants.

METHODS

Subject selection

This was a prospective study of infants ⩾ 34 weeks gestational age (GA) in a level II nursery. Patients were eligible for the study if a TSB level was drawn in the first 3 days of life. In these infants, a TcB measurement was obtained within 1 h of the TSB. If enrolled infants required a second TSB measurement between 4 and 7 days of life, a second paired (within 1 h) set of TcB and TSB levels was obtained, in order to determine the effect of postnatal age on TcB bias. TcB values obtained on late preterm infants were not used to guide treatment decisions; all management was based on TSB measurements.

A retrospective study of healthy term infants in a level I nursery was performed to provide a comparison population. Infants were included if they were undergoing confirmatory TSB testing after an elevated TcB. Universal TcB screening of healthy newborn infants in the level I nursery began at Mayo Clinic in Rochester, MN, USA, in February 2009. The current protocol specifies that infants with adjusted TcB levels (obtained by subtracting 1 mg dl^{-1} from the TcB) in the high-intermediate (HIR; 75th to 95th percentile) or high-risk (HR; ⩾ 95th percentile) zones for age on an hour-specific nomogram undergo confirmatory TSB testing.[4] Infants were excluded if > 60 min elapsed between TcB and TSB measurements or if they had previously received phototherapy. Based upon results of our previous study,[3] we aimed for a minimum of 85 patients in each group to be able to determine whether median TcB bias differed between level I term and level II late preterm infants.

Finally, the data from level II infants was compared with data previously obtained on healthy term infants with paired TcB and TSB measurements at outpatient follow-up appointments, to assess the impact of postnatal age on TcB bias.[3]

TcB and TSB measurements

TcB measurements were obtained on the infants' foreheads using a BiliChek device (Respironics, Marietta, GA, USA) according to the manufacturer's instructions. Blood samples were obtained by either capillary or venous blood draw. TSB was measured with a modified Diazo method using the Roche Total Bilirubin reagent (Roche Diagnostics, Indianapolis, IN, USA) run on a Roche Cobas c501 or c701

[1]Department of Pediatric and Adolescent Medicine, Mayo Clinic, Rochester, MN, USA; [2]Division of Biomedical Statistics and Informatics, Department of Health Science Research, Mayo Clinic, Rochester, MN, USA and [3]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA. Correspondence: Dr WJ Cook, Department of Pediatric and Adolescent Medicine, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA.
E-mail: Cook.walter@mayo.edu

instrument. Lithium heparin plasma (rather than serum) is the usual sample choice for laboratory bilirubin, though serum samples may also be drawn.

Outcomes and analysis

The primary study outcome was TcB bias between paired values from late preterm infants hospitalized in a level II nursery and a comparison cohort of healthy infants in a level I nursery. Median (interquartile range) TcB bias was calculated for each group (level I infants, level II infants with paired measurements in the first 1 to 3 days of life, level II infants with paired measurements at 4 to 7 days of life). Bland–Altman plots of the data were constructed to determine measurement bias. The Wilcoxon rank sum test was used to assess for differences in median TcB bias between groups. $P < 0.0167$ was considered statistically significant for adjusted multiple comparisons.

In addition, the effect of postnatal age on TcB bias was assessed by plotting the relationship between TcB bias and postnatal age and calculating best fit slopes and intercepts by general estimating equation. This was done using all paired TcB–TSB levels obtained from level II infants and compared to a dataset previously obtained on 88 healthy term infants with paired TcB-TSB measurements at outpatient follow-up appointments. General estimating equations are similar to linear regression, but adjust for the standard errors of the slope and intercept estimates when patients contribute more than one measurement to the analysis. For this analysis, a significant difference in slope and/or intercept between groups (level II versus outpatient) was defined as P-value < 0.05.

The clinical significance of differences between TcB and TSB was defined by risk level determination for development of hyperbilirubinemia according to the hour-specific serum nomogram produced by Bhutani et al.,[4] which plots bilirubin level as a function of postnatal age in hours. An internet-based electronic version (bilitool.org) of the nomogram was used to determine risk categories based on TcB and TSB measurements for each infant. Bilirubin levels exceeding the 75th percentile for age in hours are considered HIR, whereas bilirubin levels exceeding the 95th percentile for age are considered HR for development of hyperbilirubinemia. Sensitivity and specificity of a HIR or HR TcB value for predicting a HIR or HR TSB were then calculated. Score confidence intervals were used to estimate the 95% confidence interval of sensitivity and specificity based on the methods of Newcombe.[5]

The study was approved by the institutional review board at Mayo Clinic. Prospective study data were collected on infants born between August 2012 and December 2013, and hospitalized in the level II nursery; parents or guardians signed informed consent before participation in the prospective study. Retrospective comparison study data (also using BiliChek TcB and Roche Diazo TSB) were collected on infants born between January 2013 and June 2013, and hospitalized in the level I nursery.

RESULTS

Demographics

A total of 90 late preterm infants in a level II nursery provided 119 paired measurements (89 measurements from the first 3 days of life and 30 from days 4 to 7). Median (range) GA was 36^1 (34^0 to 37^5) weeks. Median (range) postnatal age at first paired TcB and TSB measurement was 40.6 (11.2 to 88.5) h, whereas median (range) age at second paired measurement was 93.7

(73.4 to 133.7) h. Ethnicity information obtained from the medical record indicated that 59 of infants were Caucasian, 5 were Asian, three were African American and 23 were 'other' or 'unknown'. For the retrospective comparison data in level I nursery patients, 170 paired TcB and TSB values were obtained from 170 infants ⩾37 weeks GA within the first 3 days of life. For these level I infants, median (range) postnatal age at paired TcB and TSB measurement was 31.3 (14.7 to 55.4) h. Ethnicity information was not available for the level I infants.

Relationship between TcB and TSB

TcB systematically overestimated TSB in both term infants (median TcB bias 2.6 mg dl^{-1}) and late preterm infants (median TcB bias 2.5 mg dl^{-1}) in the first 3 days of life. Median TcB bias did not differ between term infants and late preterm infants in the first 3 days of life ($P = 0.39$, Table 1). In contrast, median TcB bias in preterm infants at 4 to 7 days of life was significantly lower (2.2 mg dl^{-1}, $P = 0.002$, Table 1).

A Bland–Altman plot shows that for the level II infants, the distribution of biases differed between the first (1 to 3 days of life) and second (4 to 7 days of life) paired measurements (Figure 1). There was more variability in TcB bias with the second measurements when compared to the first measurements. Only one of 89 TcB measurements obtained in the first three days of life underestimated TSB, whereas 5 of 30 paired measurements (4 to 7 days of life) underestimated TSB (Figure 1).
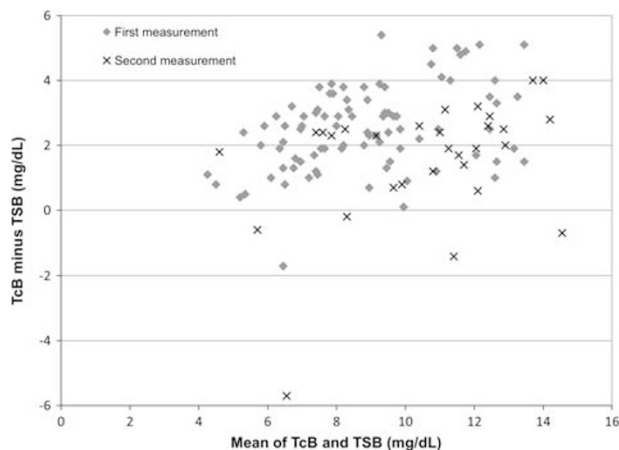

Figure 1. Bland–Altman plot of transcutaneous (TcB) and total serum (TSB) bilirubin for paired measurements obtained in the first 3 days of life (first measurement) and paired measurements obtained at 4–7 days of life (second measurement). The difference between TcB and TSB (bias) is shown on the y axis as a function of the mean of TcB and TSB values (x axis).

Table 1. Median TcB and TSB values				
	Median TcB (mg dl^{-1})	Median TSB (mg dl^{-1})	Median (IQR) TcB bias (mg dl^{-1})	P-value
Level I infants ($n = 170$)	10.2	7.5	2.6 (2.0–3.3)	Reference
Level II infants 1–3 days life ($n = 89$)	10.0	7.2	2.5 (1.6–3.4)	0.39
Level II infants 4–7 days life ($n = 30$)	12.2	10.3	2.2 (0.8–2.6)	0.002

Abbreviations: TcB, transcutaneous bilirubin, TSB = total serum bilirubin, IQR = interquartile range. P values are for significance of difference in median bias, compared to median bias for level I infants. Median TcB and TSB values among infants in the level I nursery, infants in the level II nursery with paired TcB and TSB measurements performed at 1–3 days of life and a subset of the same level II infants with paired TcB and TSB measurements performed at 4–7 days of life.

Table 2. Concordance between transcutaneous and serum bilirubin

Serum bilirubin	Transcutaneous bilirubin		
	Low or low-intermediate risk	High-intermediate or high risk	Total
Low or low-intermediate risk	40	31	71
High-intermediate or high risk	1	15	16

All results classified as either low, low-intermediate, high-intermediate or high risk based on infant's age of life in hours and bilirubin level using the Bhutani nomogram.
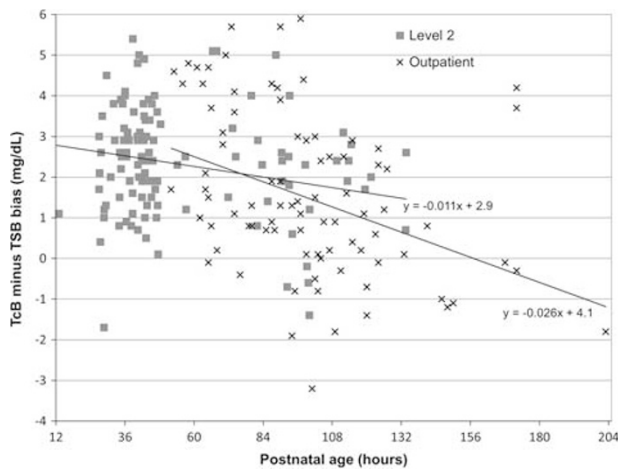

Figure 2. Plot of TcB bias (*y* axis) as a function of postnatal age (*x* axis) for level II infants with paired TcB and TSB measurements performed in the first week of life and for a population of term infants with paired measurements performed at outpatient follow-up visits (data obtained from previous publication[3]). The best fit slopes and intercepts obtained by general estimating equations are shown for both groups.

Clinical concordance between TcB and TSB measurements in the level II nursery

The clinical concordance between TcB and TSB for the level II late preterm infants was determined by assessing risk category associated with both TcB and TSB levels using the Bhutani nomogram and *bilitool.org* online assessment tool. In level II infants who had paired TcB and TSB measurements performed in the first 1 to 3 days of life, an HIR or HR TcB value predicted an HIR or HR TSB value with a sensitivity of 94% (95% confidence interval 72–99%) and specificity of 56% (95% confidence interval 45–67%) (Table 2). There were not a sufficient number of paired measurements at 4 to 7 days of life to allow accurate estimation of clinical concordance of second TcB measurements.

Relationship between postnatal age and TcB bias

The relationship between TcB bias and postnatal age was plotted to visually compare effects of postnatal age on TcB bias. Data from the current study (paired TcB and TSB measurements in level II infants) were compared to data from a previous publication[3] (paired TcB and TSB measurements obtained at an outpatient follow up visit from a population of term infants). In both studies, TcB bias decreased by 0.01–0.03 mg dl^{-1} h^{-1} over the first 7 days of life. Using general estimating equations to obtain best fit slope

and intercepts, the slopes of this effect (bias versus postnatal age) did not differ significantly between infant populations ($P = 0.10$) nor did the *y* intercepts ($P = 0.15$) (Figure 2).

DISCUSSION

Universal bilirubin screening prior to discharge from the birth hospitalization has been recommended by the American Academy of Pediatrics.[1] There is evidence to suggest that TcB measurements can provide reasonable estimates of TSB in healthy newborns.[2] The primary aim of our study was to determine whether the same universal TcB screening protocol used in a level I nursery, developed based on the observed relationship of TcB and TSB measurements for healthy term infants in our institution, could be safely applied to late preterm (≥34 weeks GA) infants in a level II nursery.

The sensitivity and specificity of HR or HIR TcB for predicting HR or HIR TSB were 94% and 56%, respectively, which is very similar to the sensitivity and specificity observed in a previous study of level I term infants.[6] In our previous study of term infants, it was necessary to subtract 1 mg dl^{-1} from TcB measurements to achieve desired performance criteria of ≥95% sensitivity and ≥50% specificity for prediction of HIR or HR TSB. If the TcB values were used without adjustment, specificity was reduced to the point that few blood draws were avoided (due to TcB falling into the HIR or HR zones when the paired TSB was low or low-intermediate risk).[6] Interestingly, although the median bias for level II infants was only slightly (and not statistically significantly) smaller than that observed for level I infants, subtraction of 1 mg dl^{-1} was not necessary to achieve similar sensitivity and specificity. Importantly, no preterm infants with high risk TSB results were misclassified as low intermediate or low risk by TcB screening.

Overall, the results demonstrate that a universal TcB screening protocol can be adapted for infants in a level II nursery (late preterm newborns) in the first 3 days of life. Although we did not have enough paired values to calculate sensitivity and specificity of TcB risk prediction for values obtained at 4 to 7 days of life, decreasing TcB bias (resulting in more frequent underestimation of TSB) and increasing variability between TcB and TSB may make this screening less reliable.

A number of previous studies have examined the utility of TcB in predicting TSB values in preterm infants. These studies have reached differing conclusions, with some finding significant and potentially dangerous underestimation of TSB in preterm infants,[7–9] whereas others have concluded that TcB accurately predicts TSB.[10–12] Consistent with our findings, one recent meta-analysis concluded that TcB was reliable in preterm infants.[13] Differences in range of GA and ethnicity of infants, statistical methods used to compare TcB and TSB, laboratory bilirubin assays used as reference methods and other factors may underlie discrepancies in previous studies. One important variable that may contribute to varying conclusions might be the postnatal age at which TcB measurements were obtained, which has been addressed in only a few studies.

One limitation to our study is that we examined only forehead TcB measurements in order to directly compare TcB results in level II late preterm infants to those obtained on level I term infants. A previous study found that as postnatal age increased, the accuracy of forehead TcB measurements decreased. These authors found that the correlation between sternum TcB and TSB was better than forehead measurements for infants older than four days of life, although a different TcB device was used.[14] Further studies are necessary to determine whether different sampling sites can mitigate the impact of postnatal age on TcB bias.

Another study found that forehead TcB measurements performed near windows (exposure to daylight) were significantly lower than those performed away from windows, and proposed that environmental factors (amount of natural light exposure) during TcB measurement may impact the relationship between

TcB and TSB.[15] These two studies used different TcB devices, so it is not clear whether these findings are directly applicable to the BiliChek device. A large multi-site study using different TcB and TSB methods also found an independent relationship between increasing postnatal age and TcB bias.[2] Although the impact of postnatal age on the accuracy of TcB measurements remains incompletely understood, there is growing evidence to suggest that TcB bias decreases as a function of increasing postnatal age.

Consistent with our previous study of term infants, we found that postnatal age had a significant impact on the relationship between TcB and TSB. We previously speculated that either environmental exposure to light after hospital discharge, changes in skin thickness related to hydration and/or weight gain after hospital discharge or other factors could account for decreasing TcB bias observed as a function of postnatal age.[3] In the current study we assessed the relationship between TcB bias and postnatal age by repeating paired measurements on the same infants at 1 to 3 days and 4 to 7 days of life. The fact that infants in this study remained in a controlled environment (level II nursery) over the study period allows us to conclude that environmental variables alone (exposure to ambient light after hospital discharge) are not responsible for decreasing TcB bias observed with increasing postnatal age. Further studies are necessary to understand which physiological factors (skin hydration, skin thickness, weight gain and others) are responsible for decreasing TcB bias.

CONCLUSION

TcB bilirubin screening protocols developed for use in healthy term newborns in the first 3 days of life can be safely applied to a population of late preterm infants. However, the relationship between TcB and TSB changes as a function of postnatal age, which may limit the utility of screening these infants after three days of life.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1 Maisels MJ, Bhutani VK, Bogen D, Newman TB, Stark AR, Watchko JF. Hyperbilir-ubinemia in the newborn infant ⩾ 35 weeks' gestation: an update with clarifications. *Pediatrics* 2009; **124**: 1193–1198.

2 Taylor JA, Burgos AE, Flaherman V, Chung EK, Simpson EA, Goyal NK *et al.* Discrepancies between transcutaneous and serum bilirubin measurements. *Pediatrics* 2015; **135**: 224–231.

3 Wickremasinghe AC, Karon BS, Cook WJ. Accuracy of neonatal transcutaneous bilirubin measurement in the outpatient setting. *Clin Pediatr (Phila)* 2011; **50**: 1144–1149.

4 Bhutani VK, Johnson L, Sivieri EM. Predictive ability of a predischarge hour-specific serum bilirubin for subsequent significant hyperbilirubinemia in healthy term and near-term newborns. *Pediatrics* 1999; **103**: 6–14.

5 Newcombe RG. Improved confidence intervals for the difference between binomial proportions based on paired data. *Stat Med* 1998; **17**: 2635–2650.

6 Karon BS, Teske A, Santrach PJ, Cook WJ. Evaluation of the BiliChek noninvasive bilirubin analyzer for prediction of serum bilirubin and risk of hyperbilirubinemia. *Am J Clin Pathol* 2008; **130**: 976–982.

7 Ebbesen F, Rasmussen LM, Wimberley PD. A new transcutaneous bilirubinometer, BiliCheck, used in the neonatal intensive care unit and the maternity ward. *Acta Paediatr* 2002; **91**: 203–211.

8 Jangaard KA, Curtis H, Goldbloom RB. Estimation of bilirubin using BiliChek™, a transcutaneous bilirubin measurement device: effects of gestational age and use of phototherapy. *Paediatr Child Health* 2006; **11**: 79–83.

9 Schmidt ET, Wheeler CA, Jackson GL, Engle WD. Evaluation of transcutaneous bilirubinometry in preterm neonates. *J Perinatol* 2009; **29**: 564–569.

10 Fouzas S, Karatza AA, Skylogianni E, Mantagou L, Varvarigou A. Transcutaneous bilirubin levels in late preterm neonates. *J Pediatr* 2010; **157**: 762–766.

11 Qualter YM, Allen NM, Corcoran JD, O'Donovan DJ. Transcutaneous bilirubin—comparing the accuracy of BiliChek® and JM 103® in a regional postnatal unit. *J Matern Fetal Neonatal Med* 2011; **24**: 267–270.

12 Ahmed M, Mostafa S, Fisher G, Reynolds TM. Comparison between transcuta-neous bilirubinometry and total serum bilirubin measurements in preterm infants < 35 weeks gestation. *Ann Clin Biochem* 2010; **47**: 72–77.

13 Nagar G, Vandermeer B, Campbell S, Kumar M. Reliability of transcutaneous bilirubin devices in preterm infants: a systematic review. *Pediatrics* 2013; **132**: 871–881.

14 Yamauchi Y, Yamanouchi I. Transcutaneous bilirubinometry: effect of postnatal age. *Acta Paediatr Jpn* 1991; **33**: 663–667.

15 Yamauchi Y. Factors affecting transcutaneous bilirubin measurement: effect of daylight. *Acta Paediatr Jpn* 1991; **33**: 658–662.