Reliability of Transcutaneous Bilirubin Devices in Preterm Infants: A Systematic Review

AUTHORS: Gaurav Nagar, MD,[a] Ben Vandermeer, MSc,[b] Sandra Campbell, MLS,[c] and Manoj Kumar, MD, MSc[a]

[a]Division of Neonatology, Department of Pediatrics, [b]Alberta Research Center for Health Evidence, and [c]John W. Scott Health Sciences Library, University of Alberta, Edmonton, Alberta, Canada

Address correspondence to Manoj Kumar, MD, MSc (Clinical Epidemiology), Department of Pediatrics, Edmonton Clinical Health Academy, Room 3-528, 11405 87 Ave NW, Edmonton, AB, Canada T6G 1C9. E-mail: manojk@ualberta.ca

abstract

BACKGROUND AND OBJECTIVE: Transcutaneous bilirubin (TcB) devices are widely used for the estimation of serum bilirubin levels in term and near-term infants. Our objective was to review the diagnostic accuracy of TcB devices in preterm infants.

METHODS: Medline, Embase, Cochrane library, Cumulative Index to Nursing and Allied Health Literature, and Scopus were searched (from database inception date until December 2012). Additional citations were identified by using the bibliographies of selected articles and from conference proceedings. The studies were included if they compared TcB with total serum bilirubin in preterm infants before phototherapy and presented data as correlation coefficients or as Bland-Altman difference plots. Data were extracted by 1 reviewer and checked for accuracy by the second reviewer. An assessment tool (quality assessment of diagnostic accuracy studies) was used for risk of bias assessments.

RESULTS: Twenty-two studies met the inclusion criteria; 21 studies reported results as correlation coefficients, with pooled estimates of $r = 0.83$ for each site of measurement. Pooled estimates in infants <32 weeks' gestation were similar to the overall preterm population ($r = 0.89$ [95% confidence interval: 0.82–0.93]). For the 2 commonly used TcB devices (ie, JM103 and BiliCheck), the results were comparable at the forehead site, although the JM103 device exhibited better correlation at the sternum. Analysis of the Bland-Altman plots (13 studies) revealed negligible bias in measurement at the forehead or sternum site by using either the JM-103 or BiliCheck device; however, the JM-103 device exhibited better precision than the BiliCheck (SD for TcB − total serum bilirubin differences: 24.3 and 31.98 μmol/L, respectively).

CONCLUSIONS: The TcB devices reliably estimated bilirubin levels in preterm infants and could be used in clinical practice to reduce blood sampling. *Pediatrics* 2013;132:871–881

Jaundice is common in the newborn period, with a majority of neonates developing visible jaundice within the first few days of birth.[1,2] In a significant proportion, this development does not possess any risks to the newborn; however, up to 10% of term and 25% of near-term neonates develop significant serum bilirubin levels requiring treatment.[3] Optimal diagnosis and management of neonatal hyperbilirubinemia require clinical examination and estimation of serum bilirubin. On clinical examination, the newborn jaundice is noted to have cephalocaudal progression.[4] However, visual assessment of serum bilirubin has been shown to correlate poorly with measured bilirubin levels in recent studies.[2,5,6] Blood sampling for estimation of serum bilirubin is one of the most common tests ordered in the neonatal units. The blood sampling is often done by heel prick and is painful, with potential long-term consequences.[7]

Transcutaneous bilirubin (TcB) devices estimate serum bilirubin noninvasively. These devices work by directing light into the skin of the neonate and measuring the intensity of specific wavelengths returned.[8] The number of wavelengths varies depending on the TcB device. These devices have been shown to correlate well with serum bilirubin levels in term and near-term infants.[9–13] The American Academy of Pediatrics recommends the use of TcB devices for the evaluation of jaundice in infants >35 weeks' gestation.[14]

Hyperbilirubinemia in preterm infants is more prevalent, more severe, and its course more protracted than in term neonates,[15] likely from slower postnatal maturation of hepatic bilirubin uptake and conjugation mechanisms.[16] In addition, delay in initiation of enteral feedings may limit intestinal motility and bacterial colonization, resulting in enhancement of bilirubin enterohepatic circulation. Although the existing guidelines allow use of TcB devices for the evaluation of jaundice in term and near-term neonates,[14] the accuracy of TcB devices for estimation of serum bilirubin in preterm infants remains unclear.

The objective of the current systematic review was to assess the diagnostic accuracy of TcB devices compared with the total serum bilirubin (TSB) measurement in preterm infants during the neonatal period.

METHODS

Search and Selection

We executed a sensitive search strategy of the following databases: Medline, Embase, Cochrane library, Cumulative Index to Nursing and Allied Health Literature, and Scopus (from the date of inception of the database to December 2012) by using both Medical Subject Headings and key words such as: ((exp Infant, Newborn/ or exp Intensive Care Units, Neonatal/ or exp Neonatal Nursing/ or nicu.mp. or exp Infant, Newborn, Diseases/ or exp Infant, Premature/ or neonat*.mp. or exp Neonatal Screening/ or exp Premature Birth/ or pre-term. mp. or preterm*.mp. or post-term*. mp.) and (bilitest* or bilimed* or biliblitz* or bilicheck* or bilichek* or tcbr or icterometer* or bilirubinometer*).mp. and ((exp Hyperbilirubinemia, Neonatal/ or (bilirubin* or hyperbilirubin*). ti,ab. or exp Hyperbilirubinemia, Hereditary/ or exp Jaundice, Neonatal/) and transcutaneous* or non-invasive* or noninvasiv* or minolta or skin* or tissue).mp.)) and ((blood or capillar* or plasma* or prick* or "heel poke*" or serum or tsbr or tsb).mp. or exp Blood/or exp Capillaries/ or exp Serum/ or exp Plasma/)). Additional terminology and predefined database limits were added to restrict the references to those related to infants <1 month of age. English language restriction was applied. Conference proceedings and bibliographies of included studies were searched for additional studies.

Studies were included in the review if they enrolled preterm (<37 weeks' gestation) infants and compared TcB results with TSB estimation during the neonatal period. We excluded pilot studies (defined a priori as those enrolling ≤20 subjects), studies enrolling preterm infants along with term infants if they did not provide separate preterm data, and studies evaluating TcB devices in subjects receiving phototherapy (Supplemental Information).
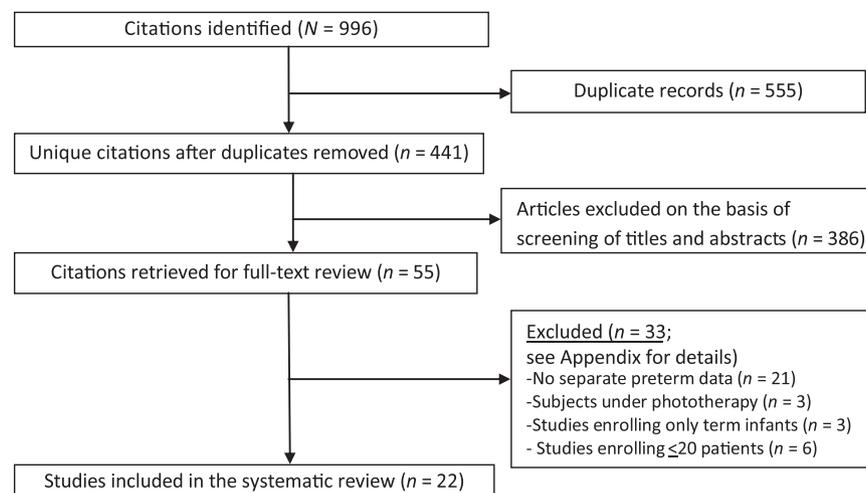

FIGURE 1
Flow of studies through the selection process.

TABLE 1 Characteristics of the Included Studies

Author (Reference), Year	Population Characteristics; Ethnicity	M/N	TcB Site	TcB: Device	Comparison Method	TSB: Method	Maximum Interval Between Tests, min	Comments
Badiee et al[21] (2012)	Iran; Persian GA (range): 25–33 wk	63/63	F	BiliCheck	r, BA	DS	10	Separate data provided for ≤30 wk (18 infants) and >30 wk (45 infants) GA
Ahmed et al[20] (2010)	United Kingdom; mixed ethnicity GA (range): 26–34 wk	183/57	F	BiliCheck	r, BA	Diazo	15–30	Results for 1 assessment per subject also provided
Siu et al[34] (2010)	Hong Kong; Chinese GA (range): 27–34 wk	110/30	F, S	JM-103	r, BA	DS	10	110 measurements at each site
Stillova et al[36] (2009)	Slovakia; white GA (range): 25–31 wk	32/32	F, S, A	JM-103	r, BA	DS	10	32 measurements at each site
Schmidt et al[33] (2009)	United States; mixed ethnicity GA (3 subgroups): 24–28 wk / 29–31 wk / 32–34 wk	131/90 / 39/30 / 43/29 / 49/31	S	JM-103	r, BA	Diazo	45	Results for 1 assessment per subject also provided
Karen et al[26] (2009)	Switzerland; mixed GA (2 subgroups): 28–33 wk / 34–36 wk	68/51 / 21/13 / 47/38	S	Bilimed	r, BA	Diazo	15	Study also enrolled term infants; shown here are data for preterm infants
Namba and Kitajima[29] (2007)	Japan; Japanese GA: <34 wk	351/50	F	JM-103	r	DS	60	All subjects had birth weight <1500 g
Stillova et al[35] (2007)	Slovakia; white GA (range): 32–34 wk	44/44	F, S, A	JM-103	r, BA	DS	10	44 measurements at each site
De Luca et al[23] (2007)	Italy; ethnicity not mentioned GA (range): 30–36 wk	340/340	F	BiliCheck	r, BA	DS	10	
Sanpavat and Nuchprayoon[32] (2007)	Thailand; ethnicity not mentioned GA (range): 30–35 wk	249/196	F	JM-103	r, BA	DS	60	
Jangaard et al[25] (2006)	Canada; majority white GA (range): not stated Mean ± SD: 30.8 ± 2.5 wk	65/33	F	BiliCheck	BA	DS	TcB measured before or soon after TSB	Study had multiple groups; data shown here are for preterm infants not receiving phototherapy
Nanjundaswamy et al[31] (2005)	United States; ethnicity not mentioned GA (range): 25–35 wk	70/70	F	BiliCheck	r	DS	30	Shown here are prephototherapy data
Willems et al[39] (2004)	Holland; majority white GA (range): 26–29 wk	93/24	F	BiliCheck	r, BA	Diazo	30	Agreement statistic provided based on 1 data set per subject
Szabo et al[37] (2004)	Switzerland; majority white GA (range): 34–36 wk	107/69	F, S / S	BiliCheck JM-102	r, BA	DS	TSB measured soon after TcB	107 measurements at each site with Bilicheck device. JM-102 only tested at sternum (107 measurements)
Karolyi et al[27] (2004)	Germany; ethnicity not mentioned GA (range): 23–33 wk	212/124	S	JM-102	r	Diazo	10	All subjects had birth weight <1500 g (VLBW)
Yasuda et al[40] (2003)	Japan; ethnicity not mentioned GA (range): 27–36 wk	75/24	F	JM-103 JM-102	r, BA	DS	30	Study also enrolled term infants; data shown here are for preterm infants
Knüpfer et al[28] (2001)	Germany; majority white GA (range): 23–36 wk	245/135	F	BiliCheck	r	Diazo	60	Data shown here are for preterm infants not receiving phototherapy
Bhutani et al[9] (2000)	United States; mixed ethnicity GA (range): 35–36 wk	163/45	F	BiliCheck	r	HPLC	30	Study also enrolled term infants; data shown here are for preterm infants

TABLE 1 Continued

Author (Reference), Year	Population Characteristics; Ethnicity	M/N	TcB Site	TcB: Device	Comparison Method	TSB: Method	Maximum Interval Between Tests, min	Comments
Donzelli and Pratesi[24] (2000)	Italy; Caucasian GA (range): 24–36 wk	82/51	S	JM-102	r	DS	10	Shown here are prephototherapy data; study also provides data for one assessment per subject
Bhardwaj et al[22] (1989)	India; ethnicity not mentioned GA: <37 wk (range not stated): 3 subgroups with mean GA of 29.2, 33.2, and 35.1 wk	90/30	F	Minolta Air-Shields bilirubin meter	r	DS	TCB and TSB measured at the same time	The study had 3 equal groups of 10 subjects each with 3 reading per subject
Tan and Mylvaganam[38] (1988)	Singapore; Chinese GA: Mean ± SD: 29.9 ± 2.9 wk	614/40	F, S	Minolta Air-Shields bilirubin meter	r	DS	TSB measured soon after the TcB	614 measurements at each site
Palmer et al[30] (1982)	Australia; ethnicity not mentioned (fair-skinned infants) GA: <37 wk	30/30	S, A	Minolta bilirubin meter	r	Diazo	Not mentioned	Study also enrolled term infants; data shown here are for preterm infants not receiving phototherapy. Thirty measurements at each site

A, abdomen; BA, Bland-Altman difference plots; DS, direct spectrophotometry; F, forehead; GA, gestational age; HPLC, high-performance liquid chromatography; M, total number of paired measurements; N, number of subjects; S, sternum; VLBW, very low birth weight.

The outcome of interest was agreement statistic between TcB and TSB measurements, provided either as the correlation coefficient or as the mean and SDs of absolute difference plots (Bland-Altman difference plots).[17]

Data Extraction and Assessment of Risk of Bias

Titles, abstracts, and citations were independently assessed by 2 reviewers for inclusion based on predefined selection criteria. Data from included studies were extracted on a specifically designed data extraction form by 1 reviewer and checked for accuracy by a second reviewer. Risk of bias assessments were conducted according to the QUADAS-2 tool.[18] This tool consists of 4 key domains: patient selection, index test, reference standard, and flow and timing. Each study is assessed for risk of bias in each of the domains and for concerns regarding applicability in the first 3 domains. Disagreements were resolved by consensus among the members of the review team.

Data Analysis

A meta-analysis was performed on the available data from both the correlation coefficients between measurements of TcB and TSB and the Bland-Altman difference plots. All correlations were first converted to Fisher z scores before being pooled. The resulting pooled Fisher z scores were then transformed back into standard correlation coefficients for ease of interpretation. For Bland-Altman difference plots, we pooled the mean TcB − TSB differences and variance across eligible studies for estimation of bias and SDs, respectively, by using methods as described by Peyton and Chong.[19]

A priori subgroup analyses were planned to explore the influence of the site of TcB measurement and the type of TcB device used. Data pertaining to all TcB

TABLE 2 Risk of Bias Assessments of the Included Studies

Author (Year)	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Badiee et al[21] (2012)	✓	✓	✓	✓	✓	✓	✓
Ahmed et al[20] (2010)	✓	?	✓	✓	✓	✓	✓
Siu et al[34] (2010)	?	?	✓	?	✓	✓	✓
Stillova et al[36] (2009)	✓	?	?	✓	✓	✓	✓
Schmidt et al[33] (2009)	✓	?	?	✓	✓	✓	✓
Karen et al[26] (2009)	✓	?	✓	✓	✓	✓	✓
Namba and Kitajima[29] (2007)	✓	✓	✓	✓	✓	✓	✓
Stillova et al[35] (2007)	✓	?	✓	✓	✓	✓	✓
De Luca et al[23] (2007)	✓	✓	✓	✓	✓	✓	✓
Sanpavat and Nuchprayoon[32] (2007)	✓	?	✓	✓	✓	✓	✓
Jangaard et al[25] (2006)	✓	✓	✓	✓	✓	✓	✓
Nanjundaswamy et al[31] (2005)	✓	✓	✓	✓	✓	✓	✓
Willems et al[39] (2004)	✓	?	✓	✓	✓	✓	✓
Szabo et al[37] (2004)	✓	?	✓	✓	✓	?	✓
Karolyi et al[27] (2004)	✓	?	✓	✓	✓	?	✓
Yasuda et al[40] (2003)	✓	✓	✓	✓	✓	?	✓
Knüpfer et al[28] (2001)	✓	?	✓	✓	✓	✓	✓
Bhutani et al[9] (2000)	✓	?	✓	✓	✓	✓	✓
Donzelli and Pratesi[24] (2000)	?	✓	✓	✓	✓	?	✓
Bhardwaj et al[22] (1989)	?	✓	✓	✓	✓	X	✓
Tan and Mylvaganam[38] (1988)	✓	?	✓	✓	✓	X	✓
Palmer et al[30] (1982)	✓	✓	✓	✓	✓	X	✓

✓, low risk; X, high risk; ?, unclear risk.

devices were included for most of the analyses; however, for comparison of individual TcB devices, we restricted our analyses to the 2 commonly used devices in current clinical practice (ie, JM-103 and BiliCheck). Additional sensitivity analyses were planned to assess the accuracy of TcB devices in infants <32 weeks' gestational age and for single measurements per enrolled subject data.

Meta-analyses were performed by using Review Manager Version 5.2 software (The Nordic Cochrane Centre, The Cochrane Collaboration, 2011; Copenhagen, Denmark). I^2 statistic was calculated for each analysis to quantify heterogeneity across studies. Forest plots were created by using SPlus Software version 3.4 (TIBCO Software Inc, Palo Alto, CA).

RESULTS

We identified 22 studies[9,20–40] providing 3527 paired measurements of TcB and TSB in 1628 patients who fulfilled the inclusion criteria (Fig 1). The baseline characteristics of the included studies are presented in Table 1. Studies varied in terms of the gestational age of the participants (23–36 weeks); site of TcB measurement (forehead: 16 studies; sternum: 10 studies; abdomen: 3 studies); TcB device used (BiliCheck: 9 studies; JM-103: 7 studies; JM-102: 4 studies; other devices: 4 studies); method used for serum bilirubin measurement (direct spectrophotometry: 14 studies; Diazo method: 7 studies; high-performance liquid chromatography: 1 study); and agreement statistic used for comparison (correlation coefficient: 21 studies; Bland-Altman difference plots: 13 studies [with 12 studies reporting results by both methods]).

Results for risk of bias assessments on the included studies are provided in Table 2. Using the QUADAS-2 tool, the majority of the included studies were assessed as low risk for bias with respect to patient selection, index test, reference standard, and flow and timing. Most studies conducted the TcB and TSB estimations within a short interval of time (within ≤30 minutes: 17 studies [77%]; within 1 hour: 21 studies [95%]). For applicability, 3 studies[22,30,38] were assessed as high risk for applicability concerns as an index test because the TcB devices used in those studies are no longer used in clinical practice.

Meta-analysis of Correlation Coefficients

Twenty-one studies[9,20–24,26–40] provided results for correlation coefficients, and the pooled estimates according to the site of measurement were as follows: forehead[9,20–23,28,29,31,32,34–40] (16 studies): 0.83 (95% confidence interval [CI]: 0.80–0.86); sternum[24,26,27,30,33–38] (10 studies): 0.83 (95% CI: 0.76–0.87); and abdomen[30,35,36] (3 studies): 0.83 (95% CI: 0.72–0.90) (Fig 2). There were no subgroup differences with respect to

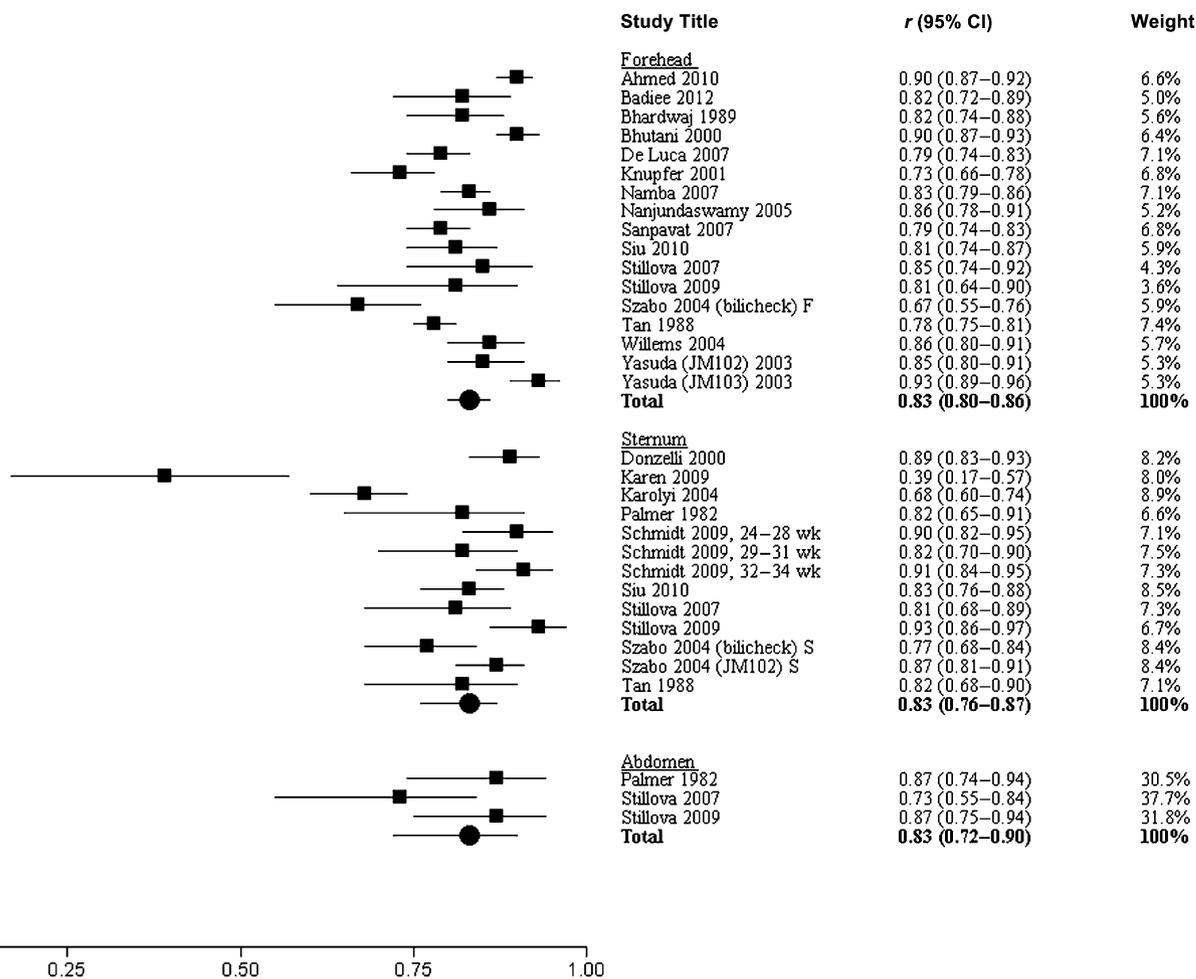| Study Title	r (95% CI)	Weight
Forehead		
Ahmed 2010	0.90 (0.87–0.92)	6.6%
Badiee 2012	0.82 (0.72–0.89)	5.0%
Bhardwaj 1989	0.82 (0.74–0.88)	5.6%
Bhutani 2000	0.90 (0.87–0.93)	6.4%
De Luca 2007	0.79 (0.74–0.83)	7.1%
Knupfer 2001	0.73 (0.66–0.78)	6.8%
Namba 2007	0.83 (0.79–0.86)	7.1%
Nanjundaswamy 2005	0.86 (0.78–0.91)	5.2%
Sanpavat 2007	0.79 (0.74–0.83)	6.8%
Siu 2010	0.81 (0.74–0.87)	5.9%
Stillova 2007	0.85 (0.74–0.92)	4.3%
Stillova 2009	0.81 (0.64–0.90)	3.6%
Szabo 2004 (bilicheck) F	0.67 (0.55–0.76)	5.9%
Tan 1988	0.78 (0.75–0.81)	7.4%
Willems 2004	0.86 (0.80–0.91)	5.7%
Yasuda (JM102) 2003	0.85 (0.80–0.91)	5.3%
Yasuda (JM103) 2003	0.93 (0.89–0.96)	5.3%
Total	**0.83 (0.80–0.86)**	**100%**
Sternum		
Donzelli 2000	0.89 (0.83–0.93)	8.2%
Karen 2009	0.39 (0.17–0.57)	8.0%
Karolyi 2004	0.68 (0.60–0.74)	8.9%
Palmer 1982	0.82 (0.65–0.91)	6.6%
Schmidt 2009, 24–28 wk	0.90 (0.82–0.95)	7.1%
Schmidt 2009, 29–31 wk	0.82 (0.70–0.90)	7.5%
Schmidt 2009, 32–34 wk	0.91 (0.84–0.95)	7.3%
Siu 2010	0.83 (0.76–0.88)	8.5%
Stillova 2007	0.81 (0.68–0.89)	7.3%
Stillova 2009	0.93 (0.86–0.97)	6.7%
Szabo 2004 (bilicheck) S	0.77 (0.68–0.84)	8.4%
Szabo 2004 (JM102) S	0.87 (0.81–0.91)	8.4%
Tan 1988	0.82 (0.68–0.90)	7.1%
Total	**0.83 (0.76–0.87)**	**100%**
Abdomen		
Palmer 1982	0.87 (0.74–0.94)	30.5%
Stillova 2007	0.73 (0.55–0.84)	37.7%
Stillova 2009	0.87 (0.75–0.94)	31.8%
Total	**0.83 (0.72–0.90)**	**100%**

FIGURE 2
Pooled estimates of correlation coefficients according to the site of TcB measurement.

site of TcB measurement. There was significant heterogeneity noted in the pooled estimates from the forehead and sternum sites. In predefined sensitivity analysis, when data for a single measurement per enrolled subject were used for analysis, the pooled estimates were no more heterogeneous (forehead site: 0.83 [95% CI: 0.79–0.85], I^2 = 16%; sternum site: 0.87 [95% CI: 0.81–0.90], I^2 = 42%) (Fig 3).

Figure 4 shows the meta-analysis results of correlation coefficients for the 2 most widely used devices in current clinical practice. For the BiliCheck device, 8 studies[9,20,21,23,28,31,37,39] provided data for correlation coefficients (all measured at the forehead site except for 1 study, which also provided

data for the sternum site),[37] and the pooled estimate was 0.83 (95% CI: 0.77–0.88) for measurement at the forehead compared with 0.77 (95% CI: 0.68–0.84) at the sternum. For the JM-103 device, 7 studies provided data for correlation coefficients, and the pooled estimate for measurement at the forehead site was 0.85 (6 studies[29,32,34–36,40] [95% CI: 0.80–0.89]) and for the sternum site, it was 0.87 (4 studies[33–36] [95% CI: 0.82–0.91]).

Figure 5 shows pooled estimates from 5 studies[20,21,33,36,39] that provided separate data for infants <32 weeks' gestation. The results show a trend toward a slightly better correlation coefficient (r = 0.89 [95% CI: 0.82–0.93]) than overall estimates in the preterm

population with no significant heterogeneity noted. Two of these studies provided separate data for subjects ≤28 weeks' gestation and reported correlation coefficients of 0.92 (Schmidt et al[33]) and 0.94 (Ahmed et al[20]).

Bland-Altman (TcB − TSB Differences) Plot Analysis

Thirteen studies[20,21,23,25,26,32–37,39,40] provided results as an analysis of the Bland-Altman difference plots. The studies were published between the years 2003 and 2012, and the majority used either JM-103 (6 studies) or Bili-Check (6 studies) devices except for the study by Karen et al,[26] which used Bilimed. In addition, 2 studies
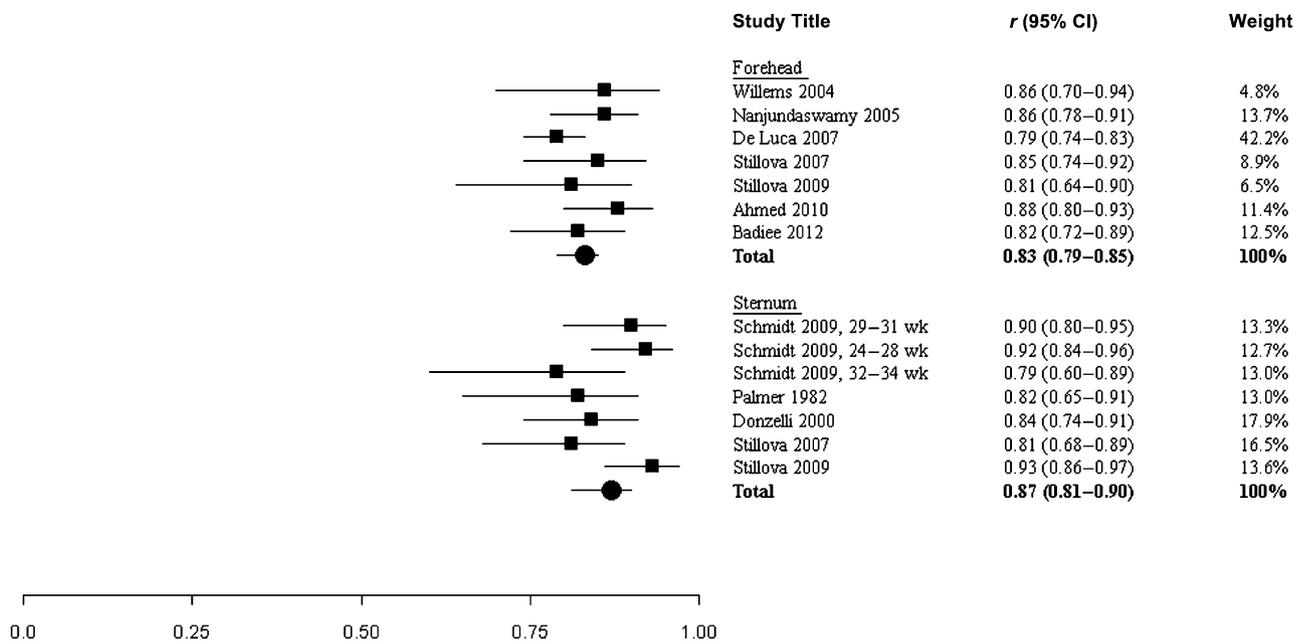
Study Title	r (95% CI)	Weight
Forehead		
Willems 2004	0.86 (0.70–0.94)	4.8%
Nanjundaswamy 2005	0.86 (0.78–0.91)	13.7%
De Luca 2007	0.79 (0.74–0.83)	42.2%
Stillova 2007	0.85 (0.74–0.92)	8.9%
Stillova 2009	0.81 (0.64–0.90)	6.5%
Ahmed 2010	0.88 (0.80–0.93)	11.4%
Badiee 2012	0.82 (0.72–0.89)	12.5%
Total	**0.83 (0.79–0.85)**	**100%**
Sternum		
Schmidt 2009, 29–31 wk	0.90 (0.80–0.95)	13.3%
Schmidt 2009, 24–28 wk	0.92 (0.84–0.96)	12.7%
Schmidt 2009, 32–34 wk	0.79 (0.60–0.89)	13.0%
Palmer 1982	0.82 (0.65–0.91)	13.0%
Donzelli 2000	0.84 (0.74–0.91)	17.9%
Stillova 2007	0.81 (0.68–0.89)	16.5%
Stillova 2009	0.93 (0.86–0.97)	13.6%
Total	**0.87 (0.81–0.90)**	**100%**

FIGURE 3
Pooled estimates of correlation coefficients from the studies reporting data for a single reading/subject.

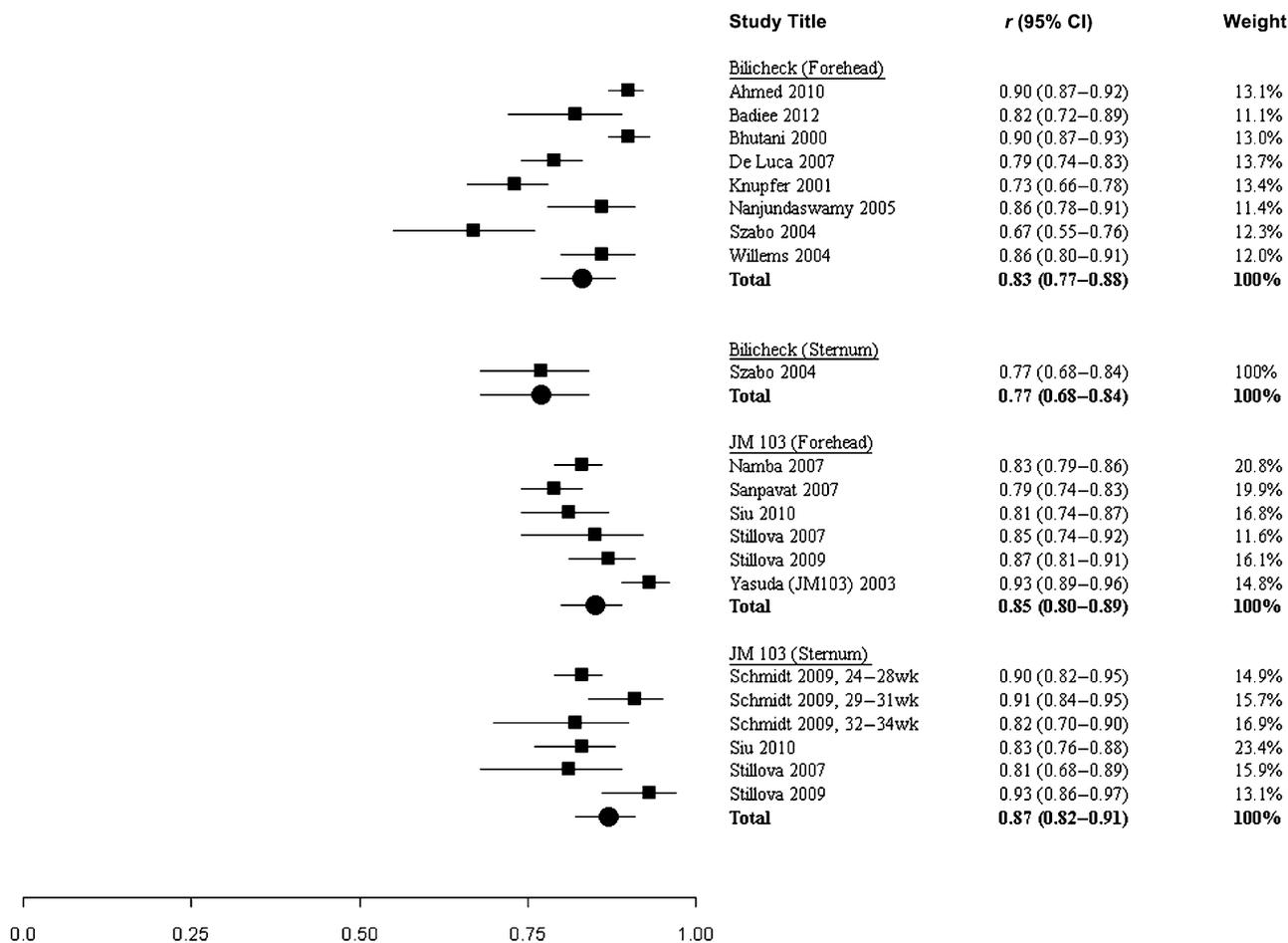| Study Title	r (95% CI)	Weight
Bilicheck (Forehead)		
Ahmed 2010	0.90 (0.87–0.92)	13.1%
Badiee 2012	0.82 (0.72–0.89)	11.1%
Bhutani 2000	0.90 (0.87–0.93)	13.0%
De Luca 2007	0.79 (0.74–0.83)	13.7%
Knupfer 2001	0.73 (0.66–0.78)	13.4%
Nanjundaswamy 2005	0.86 (0.78–0.91)	11.4%
Szabo 2004	0.67 (0.55–0.76)	12.3%
Willems 2004	0.86 (0.80–0.91)	12.0%
Total	**0.83 (0.77–0.88)**	**100%**
Bilicheck (Sternum)		
Szabo 2004	0.77 (0.68–0.84)	100%
Total	**0.77 (0.68–0.84)**	**100%**
JM 103 (Forehead)		
Namba 2007	0.83 (0.79–0.86)	20.8%
Sanpavat 2007	0.79 (0.74–0.83)	19.9%
Siu 2010	0.81 (0.74–0.87)	16.8%
Stillova 2007	0.85 (0.74–0.92)	11.6%
Stillova 2009	0.87 (0.81–0.91)	16.1%
Yasuda (JM103) 2003	0.93 (0.89–0.96)	14.8%
Total	**0.85 (0.80–0.89)**	**100%**
JM 103 (Sternum)		
Schmidt 2009, 24–28wk	0.90 (0.82–0.95)	14.9%
Schmidt 2009, 29–31wk	0.91 (0.84–0.95)	15.7%
Schmidt 2009, 32–34wk	0.82 (0.70–0.90)	16.9%
Siu 2010	0.83 (0.76–0.88)	23.4%
Stillova 2007	0.81 (0.68–0.89)	15.9%
Stillova 2009	0.93 (0.86–0.97)	13.1%
Total	**0.87 (0.82–0.91)**	**100%**

FIGURE 4
Pooled estimates of correlation coefficients for the BiliCheck and JM-103 devices.

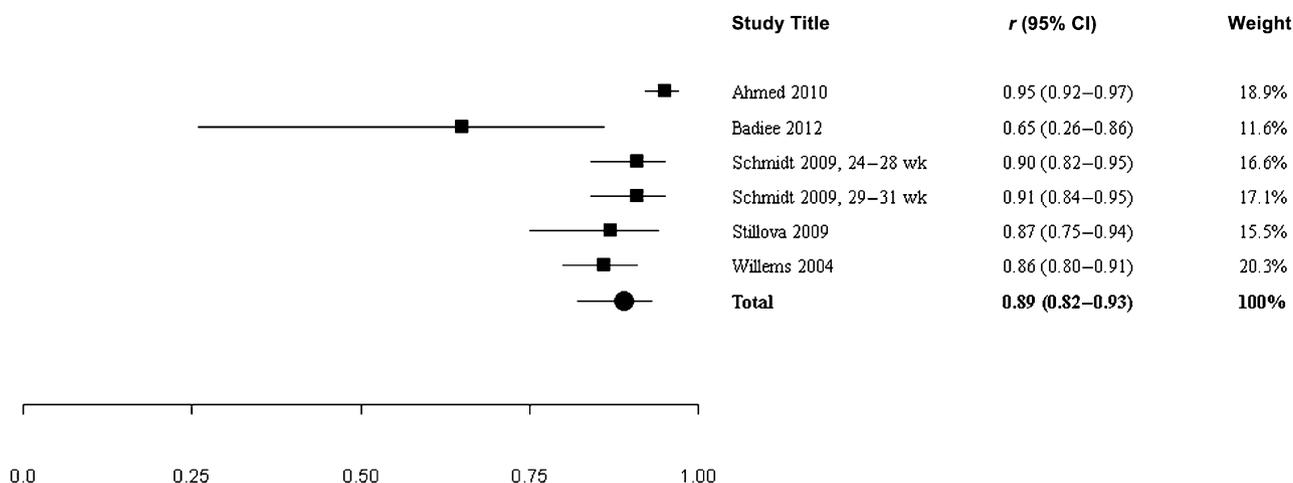**FIGURE 5**
Pooled estimates of correlation coefficients for studies reporting data for infants <32 weeks' gestation.

Study Title	r (95% CI)	Weight
Ahmed 2010	0.95 (0.92–0.97)	18.9%
Badiee 2012	0.65 (0.26–0.86)	11.6%
Schmidt 2009, 24–28 wk	0.90 (0.82–0.95)	16.6%
Schmidt 2009, 29–31 wk	0.91 (0.84–0.95)	17.1%
Stillova 2009	0.87 (0.75–0.94)	15.5%
Willems 2004	0.86 (0.80–0.91)	20.3%
Total	**0.89 (0.82–0.93)**	**100%**

evaluated the JM-102 device along with the BiliCheck and JM-103.[37,40] The results of TcB − TSB differences, along with their SDs for individual studies, are listed in Table 3.

Table 4 shows pooled data for the bias estimates and precision according to the study characteristics. The bias estimates were negligible at both the forehead and the sternum sites, with a comparable precision noted across both sites (pooled estimate of SD at forehead: 29.46 μmol/L; pooled estimate of SD at sternum: 26.06 μmol/L). In terms of devices used, the 2 commonly used TcB devices showed comparable bias; however, the JM-103 was noted to be more precise compared with the BiliCheck (pooled estimate of SD: 24.3 and 31.98 μmol/L, respectively).

DISCUSSION

We identified a large number of studies evaluating the diagnostic accuracy of various TcB devices in a preterm population. The results of this review found that TcB measurements correlate reasonably well with the serum bilirubin estimation in premature infants, particularly for the 2 widely used TcB devices in practice (ie, BiliCheck and JM-103). The accuracy of these 2 devices was similar for the measurement at the forehead site; however, JM103 exhibited better correlation with TSB for measurement at the sternum (P = .02). The analysis of absolute TcB − TSB difference plots revealed minimal bias in measurements, irrespective of the site and the device used, although the JM103 device showed a slightly better precision compared with the BiliCheck device. When the data from subjects born at ≤32 weeks' gestation were analyzed separately, the results were comparable to the diagnostic accuracy in the overall preterm population.

To the best of our knowledge, this is the first systematic review looking at the

TABLE 3 Results of Bland-Altman Difference Plots for the Studies Providing Those Data

Study	TcB Device	Measurements/ No. of Subjects	Bias[a] (μmol/L)	Precision (SD)
Site of measurement: forehead				
Badiee et al[21] 2012	BiliCheck	63/63	7.4	25.7
Ahmed et al[20] 2010	BiliCheck	183/57	28.08	31.26
Siu et al[34] 2010	JM-103	110/30	9.3	27.1
Stillova et al[36] 2009	JM-103	32/32	5.1	6.8
Stillova et al[35] 2007	JM-103	44/44	−22.4	4.2
De Luca et al[23] 2007	BiliCheck	340/340	−18.1	34.2
Sanpavat and Nuchprayoon[32] 2007	JM-103	249/196	5.1	25.7
Jangaard et al[25] 2006	BiliCheck	65/33	−3.8	33.57
Willems et al[39] 2004	BiliCheck	93/24	−4.92	27.15
Szabo et al[37] 2004	BiliCheck	107/69	−8	33
Yasuda et al[40] 2003	JM-103	75/24	7.48	27.2
Yasuda et al[40] 2003	JM-102	75/24	17.1	35.2
Site of measurement: sternum				
Siu et al[34] 2010	JM-103	110/30	20.1	27.6
Stillova et al[36] 2009	JM-103	32/32	12	6.8
Schmidt et al[33] (24–28 wk) 2009	JM-103	39/30	−18.7	32.3
Schmidt et al[33] (29–31 wk) 2009	JM-103	43/29	−13.6	22.1
Schmidt et al[33] (32–34 wk) 2009	JM-103	49/31	−17	27.2
Karen et al[26] (28–33 wk) 2009	Bilimed	21/13	−8	38
Karen et al[26] (34–36 wk) 2009	Bilimed	47/38	16	45.5
Stillova et al[35] 2007	JM-103	44/44	1.7	5.4
Szabo et al[37] 2004	BiliCheck	107/69	10	31
Szabo et al[37] 2004	JM-102	107/69	56	28

[a] TcB − TSB difference.

TABLE 4 Pooled Estimates of Bias and Precision for Studies Reporting Data as Bland-Altman Difference Plots

Variable	No. of Studies	No. of Measurements	No. of Subjects	Bias[a] (μmol/L)	Precision (SD)
According to site of measurement[b]					
Forehead	11	1361	912	−0.06	29.46
Sternum	5	424	265	3.8	26.06
According to device used					
JM-103	6	827	522	3.14	24.30
Bilicheck	6	958	655	1.10	31.98
<32-wk GA data	4	303	170	1.57	26.86

GA, gestational age.
[a] TcB − TSB difference.
[b] Analysis restricted to data from JM-103 and BiliCheck devices.

diagnostic accuracy of TcB devices in a preterm population. We present pooled data for bias and precision estimates along with the more commonly used measure of the correlation coefficient, as the latter typically describes the strength of a relation between 2 variables rather than agreement between them.[17] Thus, the clinical utility of correlation coefficient data is limited because they intuitively do not provide information regarding expected differences between the measurements conducted on a given patient by 2 separate tests.

The pooled estimates of bias noted here are comparable to the accuracy of these devices in a term population,[41–44] in whom the use of these devices has been shown to result in a marked decrease in blood sampling for assessment of neonatal jaundice.[45,46] However, there is a lower threshold for the initiation of phototherapy for preterm infants, with certain guidelines providing specific cutoffs for each gestational week according to the postnatal age.[47] Thus, the information from this systematic review should be incorporated in clinical practice, taking into consideration the thresholds for phototherapy in preterm infants. In our opinion, based on the data presented here, a TcB reading ≥50 μmol/L below the phototherapy threshold for an infant could be considered safe for not initiating phototherapy in an otherwise well preterm, without the need for TSB estimation from the laboratory. Similarly, a TcB reading above the phototherapy threshold may be sufficient grounds to initiate phototherapy without the invasive test in most situations. The latter recommendation is made despite knowing that some of these infants may be classified as below the phototherapy threshold based on TSB results because those infants are still likely to be reasonably close to the threshold.

Our review is not without limitations. First, we were unable to include a few studies that enrolled preterm infants along with the term infants. These studies did not provide comparison data for preterm populations separately in their publication and upon contacting the principal author of the study. The majority of these excluded studies enrolled near-term infants in whom the accuracy of TcB devices is not debated. Second, the estimates provided for the very preterm population are based on limited data. Although several studies enrolled subjects ≤32 weeks' and ≤28 weeks' gestation, only a few studies provided comparison data separately for these subpopulations. However, it is reassuring that the estimates for these subpopulations were comparable to the overall estimates in the preterm population. Third, several of the included studies also provided results of the relationship between TcB and TSB as a linear equation with slope and intercept. We did not pool results of these data because it would be difficult to interpret such information in clinical practice. Fourth, we did not include studies that provided data for infants under phototherapy or postphototherapy. Thus, the results of our review cannot be applied to those situations. Fifth, we did not apply a formal test to check for publication bias; however, the funnel plot of the included studies did not reveal any obvious asymmetry suggestive of missing studies with poor correlation coefficients. Lastly, several of the studies included in the meta-analysis provided multiple readings from each patient enrolled, leading to the statistical risks of dependency of data. We separately compiled studies that provided results for 1 data point for each subject for our sensitivity analysis (Fig 3). The results were similar to the overall results with no significant heterogeneity noted.

CONCLUSIONS

The results of this systematic review found that the currently available TcB devices, particularly JM-103 and BiliCheck, measure TSB values in preterm populations with reasonable accuracy, including in infants <32 weeks' gestational age. The performance of these devices in preterm populations is similar to those in term and near-term infants. Incorporating the use of TcB devices in clinical practice, as per our suggestions outlined here, could help reduce the need for blood sampling for the management of neonatal jaundice in preterm infants. The results of this review do not apply to preterm infants undergoing phototherapy or postphototherapy.

REFERENCES

1. Bhutani VK, Stark AR, Lazzeroni LC, et al; Initial Clinical Testing Evaluation and Risk Assessment for Universal Screening for Hyperbilirubinemia Study Group. Pre-discharge screening for severe neonatal hyperbilirubinemia identifies infants who need phototherapy. *J Pediatr*. 2013;162(3):477–482.e1

2. Keren R, Tremont K, Luan X, Cnaan A. Visual assessment of jaundice in term and late preterm infants. *Arch Dis Child Fetal Neonatal Ed*. 2009;94(5):F317–F322

3. Sarici SU, Serdar MA, Korkmaz A, et al. Incidence, course, and prediction of hyperbilirubinemia in near-term and term newborns. *Pediatrics*. 2004;113(4):775–780

4. Kramer LI. Advancement of dermal icterus in the jaundiced newborn. *Am J Dis Child*. 1969;118(3):454–458

5. Moyer VA, Ahn C, Sneed S. Accuracy of clinical judgment in neonatal jaundice. *Arch Pediatr Adolesc Med*. 2000;154(4):391–394

6. Szabo P, Wolf M, Bucher HU, Fauchère JC, Haensse D, Arlettaz R. Detection of hyperbilirubinaemia in jaundiced full-term neonates by eye or by bilirubinometer? *Eur J Pediatr*. 2004;163(12):722–727

7. Anand KJ. Pain, plasticity, and premature birth: a prescription for permanent suffering? *Nat Med*. 2000;6(9):971–973

8. el-Beshbishi SN, Shattuck KE, Mohammad AA, Petersen JR. Hyperbilirubinemia and transcutaneous bilirubinometry. *Clin Chem*. 2009;55(7):1280–1287

9. Bhutani VK, Gourley GR, Adler S, Kreamer B, Dalin C, Johnson LH. Noninvasive measurement of total serum bilirubin in a multiracial predischarge newborn population to assess the risk of severe hyperbilirubinemia. *Pediatrics*. 2000;106(2). Available at: www.pediatrics.org/cgi/content/full/106/2/e17

10. Engle WD, Jackson GL, Stehel EK, Sendelbach DM, Manning MD. Evaluation of a transcutaneous jaundice meter following hospital discharge in term and near-term neonates. *J Perinatol*. 2005;25(7):486–490

11. Kazmierczak SC, Robertson AF, Briley KP, Kreamer B, Gourley GR. Transcutaneous measurement of bilirubin in newborns: comparison with an automated Jendrassik-Grof procedure and HPLC. *Clin Chem*. 2004;50(2):433–435

12. Maisels MJ, Ostrea EM Jr, Touch S, et al. Evaluation of a new transcutaneous bilirubinometer. *Pediatrics*. 2004;113(6):1628–1635

13. Rubaltelli FF, Gourley GR, Loskamp N, et al. Transcutaneous bilirubin measurement: a multicenter evaluation of a new device. *Pediatrics*. 2001;107(6):1264–1271

14. American Academy of Pediatrics Subcommittee on Hyperbilirubinemia. Management of hyperbilirubinemia in the newborn infant 35 or more weeks of gestation. *Pediatrics*. 2004;114(1):297–316

15. Watchko JF, Maisels MJ. Jaundice in low birthweight infants: pathobiology and outcome. *Arch Dis Child Fetal Neonatal Ed*. 2003;88(6):F455–F458

16. Cashore WJ. Bilirubin and jaundice in the micropremie. *Clin Perinatol*. 2000;27(1):171–179, vii

17. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307–310

18. Whiting PF, Rutjes AW, Westwood ME, et al; QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529–536

19. Peyton PJ, Chong SW. Minimally invasive measurement of cardiac output during surgery and critical care: a meta-analysis of accuracy and precision. *Anesthesiology*. 2010;113(5):1220–1235

20. Ahmed M, Mostafa S, Fisher G, Reynolds TM. Comparison between transcutaneous bilirubinometry and total serum bilirubin measurements in preterm infants <35 weeks gestation. *Ann Clin Biochem*. 2010;47(pt 1):72–77

21. Badiee Z, Mohammadizadeh M, Shamee M. Diagnostic usefulness of transcutaneous bilirubinometry in very preterm newborns. *Int J Prev Med*. 2012;3(4):262–265

22. Bhardwaj HP, Narang A, Bhakoo ON. Evaluation of Minolta jaundicemeter and icterometer for assessment of neonatal jaundice. *Indian Pediatr*. 1989;26(2):161–165

23. De Luca D, Zecca E, de Turris P, Barbato G, Marras M, Romagnoli C. Using BiliCheck for preterm neonates in a sub-intensive unit: diagnostic usefulness and suitability. *Early Hum Dev*. 2007;83(5):313–317

24. Donzelli G, Pratesi S. Transcutaneous bilirubinometry in healthy preterm newborns. *Clin Biochem*. 2000;33(6):505–508

25. Jangaard K, Curtis H, Goldbloom R. Estimation of bilirubin using BiliChektrade mark, a transcutaneous bilirubin measurement device: effects of gestational age and use of phototherapy. *Paediatr Child Health (Oxford)*. 2006;11(2):79–83

26. Karen T, Bucher HU, Fauchère JC. Comparison of a new transcutaneous bilirubinometer (Bilimed) with serum bilirubin measurements in preterm and full-term infants. *BMC Pediatr*. 2009;9:70

27. Karolyi L, Pohlandt F, Muche R, Franz AR, Mihatsch WA. Transcutaneous bilirubinometry in very low birthweight infants. *Acta Paediatr*. 2004;93(7):941–944

28. Knüpfer M, Pulzer F, Braun L, Heilmann A, Robel-Tillig E, Vogtmann C. Transcutaneous bilirubinometry in preterm infants. *Acta Paediatr*. 2001;90(8):899–903

29. Namba F, Kitajima H. Utility of a new transcutaneous jaundice device with two optical paths in premature infants. *Pediatr Int*. 2007;49(4):497–501

30. Palmer DC, Zenner EM, Drew JH. Transcutaneous bilirubinometry: use in Australia. *Aust Paediatr J*. 1982;18(4):273–276

31. Nanjundaswamy S, Petrova A, Mehta R, Hegyi T. Transcutaneous bilirubinometry in preterm infants receiving phototherapy. *Am J Perinatol*. 2005;22(3):127–131

32. Sanpavat S, Nuchprayoon I. Transcutaneous bilirubin in the pre-term infants. *J Med Assoc Thai*. 2007;90(9):1803–1808

33. Schmidt ET, Wheeler CA, Jackson GL, Engle WD. Evaluation of transcutaneous bilirubinometry in preterm neonates. *J Perinatol*. 2009;29(8):564–569

34. Siu LY, Siu LW, Au SK, et al. Evaluation of a transcutaneous bilirubinometer with two optical paths in Chinese preterm infants. *HK J Paediatr*. 2010;2(15):132–140

35. Stillova L, Matasova K, Mikitova T, Stilla J, Kolarovszka H, Zibolen M. Evaluation of transcutaneous bilirubinometry in preterm infants of gestational age 32-34 weeks. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub*. 2007;151(2):267–271

36. Stillova L, Matasova K, Zibolen M, Stilla J, Kolarovszka H. Transcutaneous bilirubinometry in preterm neonates. *Indian Pediatr*. 2009;46(5):405–408

37. Szabo P, Wolf M, Bucher HU, Haensse D, Fauchère JC, Arlettaz R. Assessment of jaundice in preterm neonates: comparison between clinical assessment, two transcutaneous bilirubinometers and serum bilirubin values. *Acta Paediatr*. 2004;93(11):1491–1495

38. Tan KL, Mylvaganam A. Transcutaneous bilirubinometry in preterm very low birthweight infants. *Acta Paediatr Scand*. 1988;77(6):796–801

39. Willems WA, van den Berg LM, de Wit H, Molendijk A. Transcutaneous bilirubinometry with the Bilicheck in very premature newborns. *J Matern Fetal Neonatal Med*. 2004;16(4):209–214

40. Yasuda S, Itoh S, Isobe K, et al. New trans-cutaneous jaundice device with two optical paths. *J Perinat Med.* 2003;31(1):81–88

41. Lodha R, Deorari AK, Jatana V, Paul VK. Non-invasive estimation of total serum bilirubin by multi-wavelength spectral reflectance in neonates. *Indian Pediatr.* 2000;37(7):771–775

42. Sanpavat S, Nuchprayoon I. Noninvasive transcutaneous bilirubin as a screening test to identify the need for serum bilirubin assessment. *J Med Assoc Thai.* 2004;87(10): 1193–1198

43. Sanpavat S, Nuchprayoon I. Comparison of two transcutaneous bilirubinometers—Minolta AirShields Jaundice Meter JM103 and Spectrx Bilicheck—in Thai neonates. *Southeast Asian J Trop Med Public Health.* 2005;36 (6):1533–1537

44. Teran CG, Mohamed T, Casey J. Transcutaneous bilirubinometry: comparison of two multiwave-length devices in healthy term newborns. *Eur J Pediatr.* 2011;170(11):1485–, author reply 1487

45. Grohmann K, Roser M, Rolinski B, et al. Bilirubin measurement for neonates: comparison of 9 frequently used methods. *Pediatrics.* 2006;117(4):1174–1183

46. Maisels MJ, Kring E. Transcutaneous bilirubinometry decreases the need for serum bilirubin measurements and saves money. *Pediatrics.* 1997;99(4): 599–601

47. Recognition and treatment of neonatal jaundice. National Institute for Health and Care Excellence (NICE) 2010. Available at: www.nice.org.uk/CG98. Accessed May 29, 2013

GENES AND OBESITY: *A friend of mine has wrestled with being overweight for years. She is very careful about her diet, exercises maniacally, and limits her alcohol consumption. Despite these efforts, her BMI has always been > 30. Her weight does not seem to be related to lack of self-control and is not easily understood. My friend's weight problems may be due to her genes.*

As reported in The New York Times *(July 19, 2013), researchers have known for a long time that some aspects of weight gain are hereditary. Twins raised apart tend to have the same weight, while adopted children tend to have the body mass of the biologic parents, not their adopted parents. Evidence has accumulated over time to suggest that dozens of genes may be involved in increasing appetite, and new research suggests that at least one gene is associated not only with appetite but with a change in mammalian metabolic rate.*

The investigators developed knockout mice deficient in brain- and body-expressed "melanocortin receptor accessory protein 2" (MRAP2). When allowed to eat as much food as desired, the MRAP2 deficient mice were voracious and quickly became extremely obese. When MRAP2 deficient mice were fed the same number of calories as normal mice however, only the MRAP2 deficient mice became obese. The MRAP2 mice had to be fed 10% to 15% fewer calories in order to demonstrate the same weight gain as their normal siblings. How MRAP2 controls weight gain is not entirely understood, but researchers suspect that MRAP2 regulates "melanocortin 4 receptor" (Mc4r), a protein previously implicated in mammalian obesity. Without MRAP2 production, appropriate appetite and energy metabolism regulated by Mc4r is impaired. Interestingly, four children in a registry of 500 severely obese children were found to have alterations in the MRAP2 gene while none of the healthy controls in the same study did.

While still very preliminary, researchers are now looking for alterations in the MRAP2 gene that could lead to partial expression and hence, explain some of the variance in weight gain among people consuming the same number of calories. While it will not help my friend lose weight, I think she will appreciate learning that obesity is not always about loss of self-control.

Noted by WVR, MD