The joint use of human and electronic eye: visual assessment of jaundice and transcutaneous bilirubinometry

Daniele De Luca, Enrico Zecca, Antonio Alberto Zuppa, Costantino Romagnoli
Division of Neonatology, Department of Pediatrics, Catholic University of the Sacred Heart, Rome, Italy

SUMMARY: De Luca D, Zecca E, Zuppa AA, Romagnoli C. The joint use of human and electronic eye: visual assessment of jaundice and transcutaneous bilirubinometry. Turk J Pediatr 2008; 50: 456-461.

Our aim was to study the usefulness of jaundice visual assessment combined with skin bilirubin determination in 517 healthy newborns. Yellowness assessment was made and babies were included in three different bilirubin classes. Skin bilirubin and total serum bilirubin were determined within 10 minutes from the visual assessment. This latter led to underestimation of serum bilirubin in 16.7-40.4% and overestimation in 4.9-35.7% of newborns. Skin bilirubin measurement after the visual assessment decreased the risk of underestimation to 0-9.2% and the risk of overestimation to 2.1-11.1%. The majority of visual assessment errors were performed in the more lighted hours of the morning (75%), while the smallest number (39%) occurred during the afternoon. Skin bilirubin measurement significantly corrected these diagnostic errors ($p<0.001$, $p<0.02$) without differences during the day. Clinical estimate is unreliable for evaluating the need for serum bilirubin assay. Using the addition of skin bilirubin determination is a more advisable approach.

Key words: jaundice, yellowness, neonate, transcutaneous bilirubinometry.

Nearly 60% of term newborns become visibly jaundiced in the first week of life and total serum bilirubin (TSB) assay is actually one of the most frequent laboratory tests performed in healthy neonates[1]. In many centers, the diagnostic process is initially based on jaundice visual assessment (VA) by nurses and/or pediatricians. Guidelines by the American Academy of Pediatrics (AAP) advise clinicians to measure TSB when the jaundice "appears excessive for the infant's age"[1]. Unfortunately, no additional definition of clinically significant jaundice is given, and the AAP also states "visual estimation of bilirubin levels from the degree of jaundice can lead to errors"[1]. This happens because measurement of the skin yellowness, as a method to predict TSB levels, is unreliable. The complexity of bilirubin transfer from vascular space to the skin seems to be the main cause of this unreliability[2-4]. Moreover, concentration of bilirubin in the skin capillary beds and tissue is only one of the factors affecting the skin color: it also depends on the presence of other skin pigments, like melanin and hemoglobin. Because of the poor correlation between visually estimated jaundice and TSB, Moyer et al.[5] suggested that TSB testing should be based on risk factors for severe hyperbilirubinemia rather than on the clinical observation. Nevertheless, in a recent study, the reliability of mothers in examining their infants for jaundice was surprisingly high, whether determining the yellowness extent or using an Ingram icterometer[6].

Various transcutaneous devices have been proposed to evaluate the skin yellowness. Accuracy of first-generation bilirubinometers, based on a two-wavelength technique, was affected by skin pigmentation and other factors[7-12]. Many authors recently studied the accuracy of BiliCheck™ (BC; Respironics, Marietta, GA), a new transcutaneous bilirubinometer based on a multiwavelength spectral reflectance technology[7-10]. BC may provide an accurate measurement of skin bilirubin (SB), independent of the presence of other skin pigments, providing an automatic subtraction of the light reflected by skin chromogens among the whole spectrum of visible light[7-10]. Compared with the Ingram's icterometer, the BC provides a "real" SB measurement and not a generic estimate of yellowness[13].

The aim of this study was to verify if the combined use of yellowness VA and SB determination can be a reliable indicator for TSB measurement and a useful tool in the management of neonatal jaundice.

Material and Methods

This observational study was conducted in a maternity unit, randomly enrolling healthy term white newborns that had a TSB assayed for any reason. The Ethics Board of our University approved the study and informed consent was obtained from all parents. TSB was routinely measured if a baby appeared icteric; otherwise, TSB was determined in all babies just prior to hospital discharge. All babies had a postnatal age <72 hours and were full- or near-term (gestational age ≥35 weeks), appropriate for gestational age infants. Study neonates were delivered after an uneventful pregnancy, without asphyxia (Apgar score >7 at 1 minute and 5 minutes) and with no Rh or major ABO isoimmunization. Infants with previous TSB determinations or receiving phototherapy were excluded from the study to prevent observer bias.

For each neonate, we performed jaundice VA, SB determination and TSB measurement within 10 minutes. Blood samples (150 µl) were drawn by heel puncture into two heparinized capillary tubes for micro-assay and were protected from light exposure to avoid bilirubin photo conversion. After centrifugation at 3000 rpm for 5 minutes, samples were analyzed with the direct spectrophotometer (Microbilimeter Twin Beam Plus mod.11144A73, Ginevri, Rome, Italy), routinely used for TSB assay in our Division. Direct spectrophotometry is a simple and accurate method for TSB assaying, requiring a small blood volume[14], which analyzes serum absorbance at two wavelengths (455 and 575 nm), making automatic subtraction of absorbance due to hemoglobin. Spectrophotometer was calibrated before each determination according to the manufacturer's recommendations. The same laboratory technician, who was blinded to both yellowness VA and SB values, did all TSB determinations. Skilled pediatric nurses performed blood samplings and VA of jaundice, whereas a fellow-neonatologist measured SB. According to the AAP guidelines[1], yellowness was evaluated in a well-lighted room, blanching the skin of a naked newborn to reveal the underlying color.

To evaluate jaundice, nurses recorded their VA, subdividing yellowness in three classes. Class A: estimated TSB value 6-8 mg/dl; class B: estimated TSB value 8.1-12 mg/dl; class C: estimated TSB value 12.1-15 mg/dl. These estimated TSB levels were chosen for their usefulness in the management of physiological jaundice in term newborns. All SB measurements were made according to the manufacturer's recommendations[15]: BC was calibrated before each measurement and all SB determinations were performed with five readings in different points of the neonatal forehead while the infant was in a quiet state. Locations of the readings were distanced from the hairline and free of any bruising, nevus, hemangioma or other skin anomalies. For each yellowness class, we calculated the number of VA underestimates and overestimates when a nurse ascribed a baby into an incorrect yellowness class, respectively below or above the real TSB value. Thereafter, we computed the number of errors corrigible by using SB measurement; we considered the SB measurement as true and we corrected the VA diagnostic error when SB was not in accord with VA but in the same class of TSB.

Sensitivity and specificity were calculated in order to evaluate the accuracy of both techniques for predicting different TSB ranges. Proportions were compared by χ^2 or McNemar's test, when appropriate. Continuous variables were contrasted by Student's t test. Data were analyzed with the statistical software SPSS for Windows, rel. 11.01 (SPSS Inc., Chicago, IL, USA) and p values <0.05 were considered to be statistically significant.

Results

Baseline characteristics of the study population are shown in Table I. A total of 517 neonates underwent TSB measurement, jaundice VA and SB determination during a period of six months ending in July 2005. A slight predominance of males was observed. Subdividing the study population by VA yellowness classes, general characteristics were similar apart from bilirubin levels.

Table II shows the diagnostic usefulness of jaundice evaluation: this was initially performed by means of VA alone and then by adding the SB measurement. Assessing jaundice only visually led to underestimation of TSB in 16.7-

Table I. Details of Study Population [Data shown as mean ± SD or as number (%)]

		Yellowness class		
	All newborns	Class A (6-8 mg/dl)	Class B (8.1-12 mg/dl)	Class C (12.1-15 mg/dl)
No. of newborns	517	358	141	18
Gestational age (weeks)	39.1 ± 1.5	39.1 ± 1.5	39 ± 1.4	39.3 ± 1.2
Birth weight (g)	3295 ± 426	3287 ± 424	3317.1 ± 440	3265 ± 167
Males	299 (57.8)	201 (56.1)	84 (59.6)	14 (77.8)
Skin bilirubin	8.3 ± 3.3	6.85 ± 2.7	12.07 ± 2.1	15.5 ± 1.3
Total serum bilirubin	8.4 ± 3.6	6.9 ± 2.5	11.6 ± 2.2	14.1 ± 2.1

Table II. Diagnostic Usefulness of VA Alone and Corrected by SB Measurement [Data shown as number (%)]

	Yellowness class		
	Class A (n: 358)	Class B (n: 141)	Class C (n: 18)
Underestimate by VA	125 (34.9)°	57 (40.4)#	3 (16.7)
Underestimate by VA + SB	20 (5.6)°	13 (9.2)#	0
p	°<0.001	#<0.001	ns
Overestimate by VA	128 (35.7)*	7 (4.9)	3 (16.7)
Overestimate by VA + SB	26 (7.2)*	3 (2.1)	2 (11.1)
p	*<0.001	NS	NS

VA: Visual assessment. SB: Skin bilirubin. NS: Not significant.

40.4% and to overestimation in 4.9-35.7% of newborns. SB measurement after yellowness VA decreased the risk of underestimation to 0-9.2% and that of overestimation to 2.1-11.1%. SB determination significantly reduced both diagnostic errors when the estimated bilirubin level was ≤8 mg/dl ($p<0.001$) and significantly decreased the risk of underestimation when the estimated bilirubin level was 8.1-12 mg/dl ($p<0.001$). SB determination zeroed the underestimation and decreased the risk for overestimation in the highest yellowness range, although statistical significance was not reached.

Table III reports the number of diagnostic errors of jaundice VA alone and jointly with SB, considering three different daily periods of observation. The vast majority of measurements were made in the morning (from 7:00 to 13:00) when rate of incorrect VAs was 61-75%, while less than 40% of errors occurred in the afternoon. The number of errors was significantly different between the three periods. SB measurement significantly reduced the number of VA diagnostic errors, independently from the daily hours of observation.

Table IV reports the diagnostic accuracy of either jaundice VA alone and VA corrected by SB. Sensitivity was high for both methods when TSB ranged from 6 to 8 mg/dl, while the specificity increased from 35.5% to 75.3% using SB determination after the jaundice VA. For TSB levels higher than 8 mg/dl, sensitivity decreased and specificity increased

Table III. Diagnostic Errors of VA Alone and with SB Measurement by Hours of Determination [Data shown as number (%)]

Hours	Number of newborns	VA errors	VA + SB errors	p
7:00-9:00	396	242 (61.1)*°	51 (12.9)	<0.001
9:01-13:00	93	70 (75.2)#°	10 (10.7)	<0.001
15:00-18:00	28	11 (39.3)#*	3 (10.7)	<0.02

°$p<0.02$; *$p<0.05$; #$p<0.001$.
VA: Visual assessment. SB: Skin bilirubin.

Table IV. Diagnostic Accuracy of Jaundice VA Alone and Corrected by SB Measurement, for Predicting TSB Ranges [Data shown as percent (%)]

		TSB Ranges to be predicted (mg/dl)		
		6-8	8.1-12	12.1-15
VA	Sensitivity	93.2	37.6	5.7
	Specificity	35.5	81.8	99.1
VA + SB	Sensitivity	98.5	87.8	30.8
	Specificity	75.3	95.7	100

VA: Visual assessment. SB: Skin bilirubin.

with increasing TSB level for both methods of evaluation. Correction made by the SB measurement gave very good sensitivity and specificity at each TSB range.

Discussion

In the last 10 years, four studies have examined the accuracy of clinical estimate and its usefulness in the diagnosis of neonatal jaundice[5,16-18]. Although clinicians usually assume that jaundice is a reliable clinical finding, existing data do not support this hypothesis. Madlon-Kay[16] found that the correlation between jaundice VA and TSB is less than optimal if estimates were performed by physicians or nurses (r=0.55 and r=0.52, respectively). Level of correlation seems to increase when parents perform VA but it remains too low (r maximum=0.7) to be clinically useful. Riskin et al.[17] stated that VA performed by an experienced clinician is a reliable method to estimate the jaundice. However, Pearson's coefficient between VA and TSB was 0.68 with high inter- and intra-observer variations. This paper has also been criticized because VA accuracy could be significantly affected by the observer's experience and this training should be strictly defined[19]. Moreover, because of the well-known variation of color perception among individuals[20], it seems unreasonable to expect two observers to agree. The low inter-observer agreement has been confirmed by a recent comparative work by Szabo et al.[18]. Clinical assessment was easily affected by skin color and ambient lighting and these authors described the jaundice VA as significantly worse for predicting TSB compared to transcutaneous bilirubinometry. Indeed, Moyer et al.[5] definitively demonstrated that agreement between observers is not much better than would be predicted by chance. Since no observer performed better than others for predicting TSB, this study confirmed that the variability between yellowness and TSB level is peculiar to each infant and mainly depends on bilirubin transfer to the skin and on skin structure itself.

BiliCheck™ (BC) method has been described as unaffected by both inter/intra-observer variability[10] and different skin structure, and our data show that its use coupled with jaundice VA can improve the accuracy of clinical estimate. We only enrolled healthy term white infants in our study, with SB measurements made solely on the forehead, and our considerations may only concern this population. One limitation of our study could be that TSB measurements were performed by direct spectrophotometry and not by high performance liquid chromatography (HPLC), which is considered the gold standard for TSB assay[8]. However, some authors recently reported that direct spectrophotometry provides a better agreement with HPLC than other techniques[14]. Moreover, Bhutani et al.[7] demonstrated that the agreement between SB and TSB determined by a variety of laboratory methods is similar to the agreement between SB and HPLC.

In our study, the observers were pediatric nurses with several years of full-time experience in the well-baby nursery, and we asked them to categorize yellowness within TSB ranges that are useful in our routine management of neonatal jaundice. Nevertheless, we found a significant number of diagnostic errors, confirming that clinical estimate is an inaccurate method to evaluate neonatal jaundice. The main concern can be raised about the underestimation rate. In fact, to ascribe a neonate into the yellowness class B leads to an underestimate risk of more than 40%, meaning that 40% of babies with TSB >12 mg/dl would be unrecognized,

underestimating jaundice that could be no more physiological. Surprisingly, a lot of VA errors were found in the more lighted hours of the morning (about 75% of errors from 9:00 to 13:00), while the lowest number of errors (39%) occurred from 15:00 to 18:00. Even though the number of neonates sampled in the afternoon is small, a better room illumination does not seem to improve the jaundice VA. Diagnostic errors are fewer in the highest yellowness class (TSB >12 mg/dl) in which the vast majority of yellow babies are correctly identified as severely jaundiced. This is due to the fact that the gap between TSB levels and yellowness decreases with the increase of large amounts of bilirubin in the skin[5].

The use of BC after the jaundice VA can correct the diagnostic error, significantly reducing the underestimation rate and decreasing the number of overestimates in the lowest yellowness class. An improvement can also be observed at higher bilirubin levels but it does not reach statistical significance, probably due to the small number of cases.

As we chose yellowness classes in function of TSB levels useful for the management of physiological jaundice, it has to be emphasized that the major improvement in the jaundice estimate was obtained in classes A and B (from 6 to 12 mg/dl). In these classes, SB determination after the nurses' VA reduced the underestimate rate of 77-84% and the overestimate rate of 57-80%. Avoiding the bilirubin underestimation within these levels could easily help to identify physiological jaundice and those babies who need a TSB determination. The improvement of jaundice estimate due to SB determination is unrelated to the hour of observation and it demonstrates that BC is not affected by different room illumination, and sets BC apart from the first-generation bilirubinometers[21].

Improving the diagnosis of neonatal jaundice is actually an urgent need[22,23]. It demands a reduction both in the management costs and in the kernicterus incidence that is now increasing.

BiliCheck is useful to assess the need for TSB assay[1,24], whereas clinical estimate cannot be recommended as the only method to evaluate this need. Our study is the first to demonstrate the usefulness of the joint use of jaundice VA and BC.

Acknowledgements

We are grateful to all nurses of our maternity unit and to the laboratory technician, for performing blood samplings and analysis.

REFERENCES

1. American Academy of Pediatrics. Sub-Committee on Hyperbilirubinemia. Management of hyperbilirubinemia in the newborn infant 35 or more weeks of gestation. Pediatrics 2004; 114: 297-316. Available at: http://pediatrics.aappublications.org/cgi/content/full/114/1/297 (Accessed May 24, 2006).
2. Knudsen A, Brodersen R. Skin colour and bilirubin in neonates. Arch Dis Child 1989; 64: 605-609.
3. Ebbesen F. The relationship between the cephalocaudal progress of clinical icterus and the serum bilirubin concentration in newborn infants without blood type sensitisation. Acta Obstet Gynecol Scand 1975; 54: 329-332.
4. Knudsen A. The cephalocaudal progression of jaundice in newborns in relation to the transfer of bilirubin from plasma to skin. Early Hum Dev 1990; 22: 23-28.
5. Moyer VA, Ahn C, Sneed S. Accuracy of clinical judgement in neonatal jaundice. Arch Pediatr Adolesc Med 2000; 154: 391-394.
6. Madlon-Kay DJ. Maternal assessment of neonatal jaundice after hospital discharge. J Fam Pract 2002; 51: 445-448.
7. 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: http://www.pediatrics.org/cgi/content/full/106/2/e17. (Accessed May 24, 2006).
8. Rubaltelli FF, Gourley GR, Loskamp N, et al. Transcutaneous bilirubin measurement: a multicenter evaluation of a new device. Pediatrics 2001; 107: 1264-1271.
9. 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.
10. Bertini G, Rubaltelli FF. Non-invasive bilirubinometry in neonatal jaundice. Semin Neonatol 2002; 7: 129-133.
11. Schumacher RE. Noninvasive measurements of bilirubin in the newborn. Clin Perinatol 1990; 17: 417-435.
12. Dai J, Parry DM, Krahn J. Transcutaneous bilirubinometry: its role in the assessment of neonatal jaundice. Clin Biochem 1997; 30: 1-9.
13. Gosset IH. A perspex icterometer for neonates. Lancet 1960; 1: 87-88.
14. Kazmierczak SC, Robertson AF, Catrou PG, Briley KP, Kreamer BL, Gourley GR. Direct spectrophotometric method for measurement of bilirubin in newborns: comparison with HPLC and an automated diazo method. Clin Chem 2002; 48: 1096-1097.
15. Respironics, BiliCheck: Clinical indication for use. User instruction manual. Marietta (GA): Respironics; 1997.

16. Madlon-Kay DJ. Recognition of the presence and severity of newborn jaundice by parents, nurses, physicians and icterometer. Pediatrics 1997; 100: Available at: www.pediatrics.org/cgi/content/full/100/3/e3 (Accessed May 24, 2006).
17. Riskin A, Kuglman A, Abend-Winger M, Green M, Hemo M, Bader D. In the eye of beholder: how accurate is clinical estimation of jaundice in newborns? Acta Paediatr 2003; 92: 574-576.
18. Szabo P, Wolf M, Bucher HU, Haensse D, Fauchere JC, Arlettaz R. Assessment of jaundice in preterm neonates: comparison between clinical assessment, two transcutaneous bilirubinometers and serum bilirubin values. Acta Paediatr 2004; 93: 1491-1495.
19. Sarici SU, Candemir G, Alpay F. How reliable is clinical assessment in neonatal jaundice? Acta Paediatr 2004; 93: 857-858.
20. Hurlbert A. Visual perception: deciphering the colour code. Nature 1991; 349: 191-193.
21. Yamaouchi Y, Yamanouchi I. Factors affecting transcutaneous bilirubin measurement: effect of daylight. Acta Paediatr Jpn 1991; 33: 658-662.
22. Felc Z. Improvement of conventional transcutaneous bilirubinometry results in term newborn infants. Am J Perinatol 2005; 22:173-179.
23. Ross G. Hyperbilirubinemia in the 2000s: what should we do next? Am J Perinat 2003; 20: 415-424.
24. Briscoe L, Clark S, Yoxall CW. Can transcutaneous bilirubinometry reduce the need for blood tests in jaundiced full term babies? Arch Dis Child Fetal Neonatal Ed 2002; 86: F190-F192.