

Peter Szabo · Martin Wolf · Hans Ulrich Bucher
Jean-Claude Fauchère · Daniel Haensse
Romaine Arlettaz

Detection of hyperbilirubinaemia in jaundiced full-term neonates by eye or by bilirubinometer?

Received: 15 March 2004 / Revised: 21 July 2004 / Accepted: 28 July 2004 / Published online: 8 September 2004
© Springer-Verlag 2004

Abstract The aim of this study was to compare predictions of hyperbilirubinaemia by eye, performed by trained physicians and nurses, with predictions obtained using two commercial bilirubinometers. Jaundice was assessed in 92 white and 48 non-white healthy full-term neonates using three non-invasive methods and by total serum bilirubin as the reference method. Clinical assessment of cephalocaudal progression of jaundice was carried out independently by a physician and by nurses. Simultaneously, the Minolta Airshields JM-102 was applied on the sternum, the BiliCheck on both the forehead and the sternum, and finally, serum bilirubin concentrations were determined. The Minolta JM-102 showed the best performance with $r^2=0.90$, an intraclass correlation coefficient (ICC) of 0.93, and a 95% confidence interval (CI) of ± 4 units (approx. $56 \mu\text{mol/l}$). The BiliCheck performed slightly better on the forehead than over the sternum with $r^2=0.90$, an ICC of 0.88, and a CI of $\pm 62 \mu\text{mol/l}$. Assessment of jaundice by eye was least accurate with $r^2=0.74$, an ICC of 0.67, and a CI of ± 1.5 zones (corresponding to $\pm 75 \mu\text{mol/l}$). Skin pigmentation and ambient light both adversely affected non-invasive bilirubin estimation. **Conclusion:** All three non-invasive methods are well suited for estimation of serum bilirubin but show large confidence intervals. In healthy term newborns, hyperbilirubinaemia ($> 250 \mu\text{mol/l}$) can be safely ruled out by eye if jaundice does not reach the abdomen or the extremities (Kramer zones 1 and 2), with < 22 units ($< 230 \mu\text{mol/l}$) for the Minolta JM-102, or with a cut-off of $190 \mu\text{mol/l}$ for the BiliCheck. If these respective thresholds are exceeded, serum bilirubin concentrations should be measured.

Keywords Clinical assessment · Hyperbilirubinaemia · Jaundice · Transcutaneous bilirubinometry

Abbreviations CI: 95% confidence interval · ICC: intraclass correlation coefficient · ROC: receiver operating characteristic

Introduction

Hyperbilirubinaemia, defined as a serum bilirubin concentration $> 250 \mu\text{mol/l}$, is a frequent problem in term neonates. As a consequence, serum bilirubin measurement is the standard method to assess jaundiced neonates. It is, unfortunately, invasive, painful, and costly. To overcome these drawbacks, non-invasive methods of bilirubin measurement have been proposed. Since Yamanouchi described a transcutaneous bilirubinometer in 1980 [22], several devices have been developed to non-invasively assess serum bilirubin concentration [16, 17, 21, 22]. These instruments have been shown to give results that correlate well with serum bilirubin values, but exhibit large scattering [1]. Consequently, with the application of non-invasive bilirubinometers, the number of serum measurements could be reduced and painful heel pricks avoided but not completely abolished [5, 15]. Warnings have been published for the use of bilirubinometers in newborns with dark and yellow skin and in bright sunlight [23].

The easiest available and cheapest non-invasive method for assessment of jaundice is the human eye. As early as 1969, Kramer [13] had published a study which correlated the clinically observed cephalocaudal advancement of jaundice with the values of unconjugated serum bilirubin. Although in clinical practice the human eye is used to screen babies for jaundice, this method has, to our knowledge, not yet been systematically compared with non-invasive bilirubinometers.

P. Szabo · M. Wolf · H. U. Bucher (✉) · J.-C. Fauchère
D. Haensse · R. Arlettaz
Neonatology Clinic, University Hospital,
Frauenklinikstrasse 10, 8091 Zürich,
Switzerland
E-mail: buh@usz.ch
Tel.: +41-1-2555340
Fax: +41-1-2554244

The aim of our study was therefore to compare the clinical assessment of jaundice using the method of Kramer and two methods of transcutaneous bilirubinometry (Minolta JM-102 and BiliCheck) with serum bilirubin values. The second aim was to investigate the effect of skin colour and ambient light on the three non-invasive methods.

Subjects and methods

Patient selection

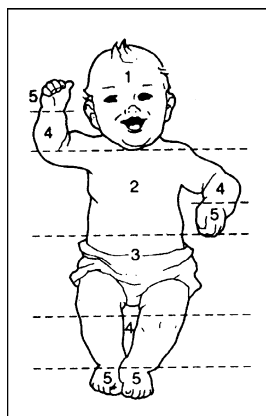
Full-term healthy newborns, with a birth weight of at least 2000 g and not older than 6 days, were recruited from July 1st 2002 until June 30th 2003 in the maternity ward. Exclusion criteria were: haemolysis, jaundice within the first 36 h, and phototherapy.

Clinical assessment of jaundice

The five dermal zones of the cephalocaudal progression of icterus, described by Kramer were used [13]. These dermal zones and the corresponding ranges of serum bilirubin values are given in Fig. 1. Jaundice progressing to zone 5 corresponds to a serum bilirubin of $>250 \mu\text{mol/l}$. All infants were assessed independently by the primary investigator and by the nurse in charge.

The Minolta JM-102 bilirubinometer

The Minolta Airshields JM-102 bilirubinometer is based on a two-wavelength analysis and gives a reflectance unit (index) [11, 12]. The manufacturer recommends that each institution establish a conversion table adapted to its specific patient mix and reference method. The corresponding values obtained in a pilot study are shown in Table 1. The equation for converting the Minolta JM-102 index values to units of $\mu\text{mol/l}$ is: $y = x \times 16 - 124$.



Dermal Zone	Mean \pm SD $\mu\text{mol/l}$
1	101 \pm 5
2	152 \pm 29
3	201 \pm 31
4	257 \pm 29
5	> 257

Fig. 1 Correlation between icteric dermal zones (Kramer) and serum bilirubin values

Table 1 Guide values for the conversion of Minolta JM-102 index numbers to serum bilirubin equivalents

Index	Mean serum bilirubin equivalent ($\mu\text{mol/l}$)	Serum bilirubin equivalent ($\mu\text{mol/l}$, mean + 2SD)
11	52	106
12	68	122
13	84	138
14	100	154
15	116	170
16	132	186
17	148	202
18	164	218
19	180	234
20	196	250
21	213	267
22	229	283
23	245	299
24	261	315
25	277	331
26	293	347
27	309	363
28	325	379

Following the manufacturers instructions, two readings were taken from the sternum and only the higher one included [7, 12]. Care was taken to avoid skin areas that were bruised, excessively hairy, or hypermelanotic.

The BiliCheck bilirubinometer

The BiliCheck system averages the spectra of five replicate measurements at one site to give a bilirubin estimate in $\mu\text{mol/l}$ (or mg/dl) that is based on reflectance data analysed at multiple wavelengths [2]. Measurements were performed on the forehead and on the sternum.

Serum bilirubin assay

For each comparison, two samples of 40 μl blood, collected in capillaries from a heel prick, were centrifuged for 10 min and the serum bilirubin measured with a Bilimeter II (Pfaff, Neuburg, Germany), a standard photometer using two wavelengths (455 nm and 575 nm). This instrument was calibrated with test solutions and has a technical error of $\pm 1.3\%$ for values up to 350 $\mu\text{mol/l}$. In a previous study, its performance was compared to HPLC and was shown to correlate well with HPLC values [17]. The mean value of the two capillaries was used for comparison.

Study procedure

Before starting the investigation, nurses on the maternity ward were instructed in the clinical assessment of the progression of jaundice using the five dermal zones described by Kramer [13]. A nurse routinely checked each infant for jaundice at intervals no greater than 12 h.

If jaundice reached the thighs (zone 3), the primary investigator undertook the same visual assessment of icterus without knowing the nurse's result and performed measurements with the two transcutaneous devices. The ambient light, whether daylight or fluorescent light (Osram Lumilux warm white L36 W31) was recorded. Infants were also included in the study if jaundice did not reach zone 3 by 72 h; in these infants a blood sample was taken for bilirubin measurement together with the routine metabolic screening. Thus, no additional heel prick was required. For all blood samplings, the nurse administered an oral sweetener for pain relief [6] before pricking the pre-warmed heel and taking the two capillary blood samples. Blood sampling occurred within 10 min after the application of the three non-invasive methods. The sampled blood was protected from light and analysed within 30 min.

Data analysis

The statistical analysis was performed using SPSS (Version 11.5). We compared the different methods (non-invasive measurements versus serum bilirubin) using linear regression analysis, an intraclass correlation coefficient (ICC) [14, 18], a Bland-Altman analysis [4], and receiver operator characteristics (ROC) [10]. The ICC evaluates the level of agreement between different measurement methods. ICC = 1 indicates perfect agreement while ICC = 0 indicates no agreement at all. In contrast to the correlation coefficient (r) of the regression analysis, the ICC reflects systematic differences between methods [14]. For example, if one method always yields double the value of the other method, the conventional correlation coefficient shows a value of 1, which reflects the perfect linear relation. The ICC for such an example is, however, smaller than 1 and thus represents the mismatch. For the Bland-Altman analysis, Kramer zones (Fig. 1) and Minolta units (Table 1) were converted into $\mu\text{mol/l}$ and the difference non-invasive values - serum values" plotted against serum values. This analysis was used to calculate the 95% confidence intervals (CI) (± 2 standard deviations) and to investigate the influence of skin colour and ambient light. To compare the performance of the three non-invasive tests, sensitivity, specificity, and the area under the ROC curve [10] were calculated. The study protocol was approved by our local ethics committee and parental consent was obtained for each infant.

Results

Patients

A total of 140 infants born between 37 0/7 and 41 6/7 gestational weeks (median 39 weeks) were included in this study, of whom 92 were white (caucasian) and 48 were non-white (18 of Asian, 30 of Indian or African

origin). The median birth weight was 3320 g (range 2050–4400 g). Ten infants (7%) were growth retarded with a birth weight < 10th percentile.

Comparison between non-invasive methods and serum bilirubin

All three non-invasive methods correlated highly with serum bilirubin (Fig. 2) with the two bilirubinometers

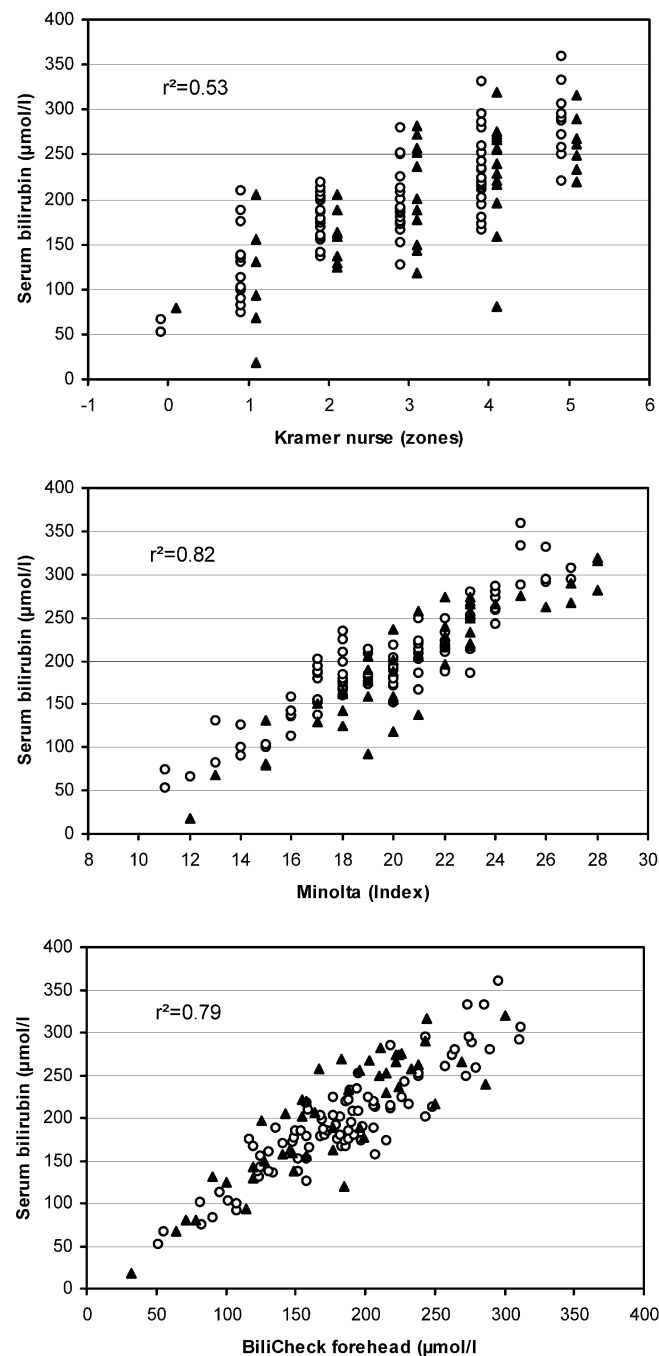


Fig. 2 Results of non-invasive methods plotted against serum bilirubin values. White infants (open circles), non-white infants (solid triangles)

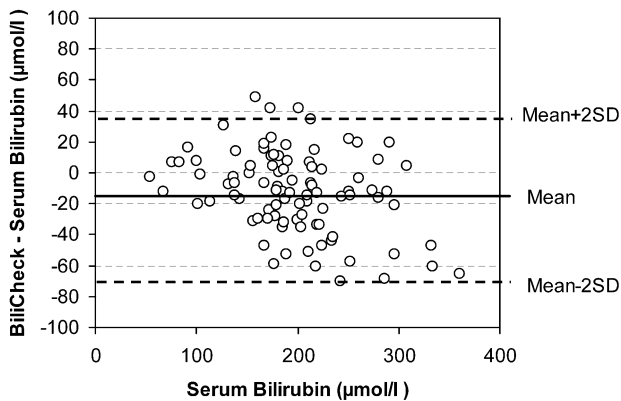
Table 2 Linear correlation coefficients (R^2) and ICC for non-invasive methods compared with serum bilirubin values for white and non-white infants

Method	R^2		ICC	
	White	Non-white	White	Non-white
Kramer (nurse)	0.74	0.71	0.67	0.70
Kramer (paediatrician)	0.70	0.65	0.63	0.44
Minolta sternum	0.90	0.91	0.93	0.92
BiliCheck forehead	0.90	0.87	0.88	0.80
BiliCheck sternum	0.89	0.81	0.89	0.78

performing better than the clinical assessment. This result was also confirmed by the intraclass correlation (Table 2).

The scattering for the Minolta instrument, quantified as ± 2 SDs or 95% CI, was 4 units (equivalent to 56 $\mu\text{mol/l}$) for both white and non-white infants. The BiliCheck showed more variability ($P=0.02$) in non-white infants ($\pm 68 \mu\text{mol/l}$ CI) than in white infants ($\pm 52 \mu\text{mol/l}$ CI) (Fig. 3). Assessments made using the

White infants



Non-white infants

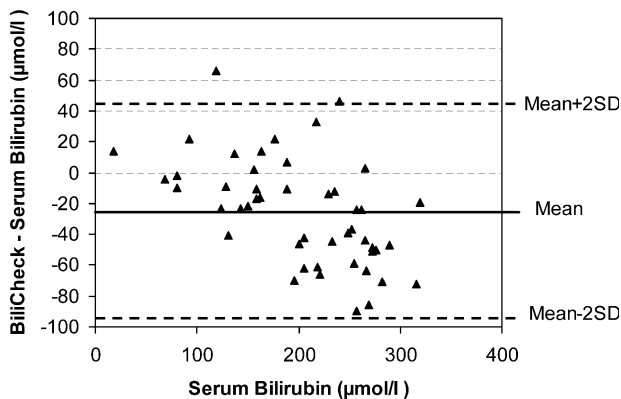


Fig. 3 Bland-Altman-Plot: BiliCheck (forehead) in white (open circles) and non-white infants (solid triangles) versus serum bilirubin

Kramer method had a 95% CI of ± 1.5 zones (corresponding to $\pm 75 \mu\text{mol/l}$). The inter-observer agreement between clinical assessment by nurses and by the physician was significantly ($P < 0.05$) better for white infants (kappa 0.56) than for non-white infants (kappa 0.36). Table 3 shows the sensitivity and specificity for all three non-invasive methods at various cut-offs for serum bilirubin values $> 250 \mu\text{mol/l}$. The area under the ROC-curve was 0.98 for Minolta, 0.92 for BiliCheck over forehead, 0.88 for BiliCheck over sternum, and 0.84 for the Kramer method.

Influence of the ambient light

The Kramer method and the BiliCheck measurements over the sternum showed significantly less accuracy in daylight than in white fluorescent light ($P < 0.001$ for Kramer, $P < 0.02$ for BiliCheck). Measurements performed with the Minolta JM-102 instrument were not influenced by ambient light.

Discussion

Our study confirms the previously reported highly significant correlation between BiliCheck and Minolta-102 readings with total serum bilirubin values [3, 11, 16, 21]. The correlation of clinical assessment with serum bilirubin values for both nurses and the primary investigator is lower but still statistically significant. These results, based on linear regression analyses, were also confirmed by ICC, thereby ruling out systematic deviations.

BiliCheck applied to the forehead and clinical assessment by the physician were significantly less

Table 3 Sensitivity and specificity for detection of serum bilirubin $> 250 \mu\text{mol/l}$ by the method of Kramer (nurses) and the two transcutaneous methods (Minolta JM-102) and BiliCheck (on the forehead)

		Sensitivity	Specificity
Kramer nurse (zone)	2	1	0.36
	3	0.93	0.46
	4	0.84	0.71
	5	0.38	0.95
	Minolta (index)	19	1
	20	1	0.58
	21	1	0.74
	22	0.97	0.83
	23	0.94	0.94
	24	0.66	0.99
BiliCheck forehead ($\mu\text{mol/l}$)	180	0.97	0.60
	190	0.94	0.72
	200	0.88	0.80
	210	0.84	0.86
	220	0.78	0.89
	230	0.63	0.94
	240	0.59	0.94
	250	0.44	0.97

accurate in non-white infants than in white infants. The BiliCheck systematically underestimated serum bilirubin values by 25 $\mu\text{mol/l}$ on average in non-white ($P < 0.001$) infants (Fig. 3). This finding is in contrast with previous reports and should be investigated further. In the meantime, caution is indicated with transcutaneous bilirubinometers in infants with dark skin.

For clinical practice, the specificity and sensitivity of a test are considered to be the most useful parameters for identification or exclusion of patients with a disease (Table 3). As both can be varied—in opposite directions—by changing the cut-offs, a combined analysis such as the ROC area is best suited to compare the performance of a test. In the present study, carried out in term newborn infants, Minolta had the best performance (ROC = 0.98), closely followed by BiliCheck over forehead (ROC = 0.92) and over sternum (ROC = 0.88), with clinical assessment providing a weaker, though still acceptable, performance (ROC = 0.84). It should be noted that the presented results are valid only for term newborns. We have previously shown that all three non-invasive test performances were lower in preterm infants in the range of 34 0/7 to 36 6/7 gestational weeks [19].

Clinical assessment

Our study confirms the findings of Kramer [13] who reported a mean serum bilirubin increase of $50 \pm 37 \mu\text{mol/l}$ (± 1 SD) for each dermal zone of jaundice in white and non-white infants. However, a closer analysis of the relation between dermal zones and serum bilirubin shows only a small increase of 13 $\mu\text{mol/l}$ from zone 2 to zone 3. This means that the relation between these two variables is not linear but S-shaped (Fig. 2).

In our study all infants with jaundice progression below zone 3 (only head and upper trunk), assessed by nurses and by the physician, had serum bilirubin below 250 $\mu\text{mol/l}$. Therefore, for practical purposes this cut-off can be used to rule out hyperbilirubinaemia with very high confidence, eliminating the need to perform further investigations. In infants with jaundice progression to zones 3 and 4, the risk for hyperbilirubinaemia is 14% and 25% respectively. As a consequence, transcutaneous bilirubinometry or a serum bilirubin measurement should be performed in these infants.

Minolta JM-102

The Minolta (Airshield) JM-102 showed the best performance of all three non-invasive methods. However, the inconvenience of having to derive institutional based calibration factors, adapted to both the population and the reference method, remains a problem. Our action guide values were established with white infants (Table 1) and are approximately the same as those reported by Wick [20]. However, our reference range is

about 20 $\mu\text{mol/l}$ higher than that published by Harish and Sharma [11] who investigated a non-white population of 60 term newborns in India. The practical consequence from this difference is that skin pigmentation affects transcutaneous jaundice assessment.

With its high sensitivity, the Minolta JM-102 bilirubinometer is ideal for excluding infants without hyperbilirubinaemia ($< 250 \mu\text{mol/l}$). This means that infants with a Minolta index below 22 need not endure a heel prick for serum bilirubin measurement. Of all infants with Minolta readings of 23 units or lower, only 6% had serum bilirubin levels $> 250 \mu\text{mol/l}$ and none had a serum bilirubin $> 300 \mu\text{mol/l}$. Therefore, this cut-off can be considered safe provided that transcutaneous measurement is repeated every 4 to 6 h. It should be noted, however, that this is valid only for healthy term infants more than 36 h old and not for younger or for sick infants.

BiliCheck

Our findings confirm the previously reported good correlation of BiliCheck values with serum bilirubin values in healthy white term infants [8, 16, 17]. However, in infants with non-white skin, the BiliCheck instrument showed a significantly lower performance (Fig. 3), with wider CI and underestimation of serum bilirubin by $25 \pm 35 \mu\text{mol/l}$, a finding similar to that already reported by Engle et al. [9]. Until this finding has been confirmed or refuted in a larger patient sample, BiliCheck values should be interpreted with caution in non-white infants.

In the daily clinical situation, a heel prick can only be avoided if the BiliCheck shows values below 190 $\mu\text{mol/l}$ (sensitivity 94%, Table 3).

Another unexpected finding is the fact that BiliCheck measurements were affected by ambient light. For BiliCheck measurements over the sternum, the correlation with serum bilirubin was significantly less accurate with daylight than with fluorescent room light ($P < 0.018$). The same tendency was observed over the forehead ($P < 0.184$). We therefore suggest that transcutaneous bilirubinometer readings are dependent on the light conditions under which the measurements are performed. This hypothesis should be tested in a larger population.

Conclusion

All three non-invasive methods are well suited for estimation of serum bilirubin but have relatively large 95% CI. Non-invasive measurements may be affected by skin pigmentation and ambient light. In healthy term newborns, hyperbilirubinaemia ($> 250 \mu\text{mol/l}$) can be safely ruled out by eye if jaundice does not reach the abdomen or the extremities (Kramer zones 1 and 2), with < 22 units ($< 230 \mu\text{mol/l}$) for the Minolta instrument or with a cut-off of 190 $\mu\text{mol/l}$ for the BiliCheck system.

If respective thresholds are exceeded, serum bilirubin concentration should be measured.

Acknowledgements The authors thank the parents of the infants studied and the nurses for making this study possible. We also thank Konrad Gasser for statistical advice and acknowledge the editorial assistance of Helen Keller-Kurzman and Derek Brown.

References

1. Beck M, Kau N, Schlebusch (2003) Transcutaneous bilirubin measurement in newborn infants: evaluation of a new spectrophotometric method. *Arch Dis Child Fetal Neonatal Ed* 88: F350–F351
2. Bhutani VK, Gourley GR, Adler S, Kreamer B, Dalin C, Johnson LH (2000) Noninvasive measurement of total serum bilirubin in a multiracial predischarge newborn population to assess the risk of severe hyperbilirubinemia. *Pediatrics* 106: E17
3. Bilgen H, Ince Z, Ozek E, Bekiroglu N, Ors R (1998) Transcutaneous measurement of hyperbilirubinaemia: comparison of the Minolta jaundice meter and the Ingram icterometer. *Ann Trop Paediatr* 18: 325–328
4. Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1: 307–310
5. Briscoe L, Clark S, Yoxall CW (2002) Can transcutaneous bilirubinometry reduce the need for blood tests in jaundiced full term babies? *Arch Dis Child Fetal Neonatal Ed* 86: F190–F192
6. Bucher HU, Baumgartner R, Bucher N, Seiler M, Fauchère JC (2000) Artificial sweetener reduces nociceptive reaction in term newborn infants. *Early Hum Dev* 59: 51–60
7. Carbonell X, Botet F, Figueras J, Riu-Godo A (2001) Prediction of hyperbilirubinaemia in the healthy term newborn. *Acta Paediatr* 90: 166–170
8. Ebbesen F, Rasmussen LM, Wimberley PD (2002) A new transcutaneous bilirubinometer, BiliCheck, used in the neonatal intensive care unit and the maternity ward. *Acta Paediatr* 91: 203–211
9. Engle WD, Jackson GL, Sendelbach D, Manning D, Frawley WH (2002) Assessment of a transcutaneous device in the evaluation of neonatal hyperbilirubinemia in a primarily Hispanic population. *Pediatrics* 110: 61–67
10. Hanley JA, McNeil BJ (1982) The managing and use of the area under the Receiver Operating Characteristic(ROC) curve. *Radiology* 143: 29–36
11. Harish R, Sharma DB (1998) Transcutaneous bilirubinometry in neonates: evaluation of Minolta Airshields jaundicemeter. *Indian Pediatr* 35: 264–267
12. Heick C, Mieth D, Fallenstein F, Schubiger G, Nars PW, Amato M (1982) Transcutaneous bilirubin determination in the newborn infant. *Helv Paediatr Acta* 37: 589–597
13. Kramer LI (1969) Advancement of dermal icterus in the jaundiced newborn. *Am J Dis Child* 118: 454–458
14. Lee J (1992) Evaluating agreement between two methods for measuring the same quantity: a response. *Comput Biol Med* 22: 369–371
15. Maisels MJ, Kring E (1997) Transcutaneous bilirubinometry decreases the need for serum bilirubin measurements and saves money. *Pediatrics* 99: 599–601
16. Robertson A, Kazmierczak S, Vos P (2002) Improved transcutaneous bilirubinometry: comparison of SpectR(X) Bili-Check and Minolta Jaundice Meter JM-102 for estimating total serum bilirubin in a normal newborn population. *J Perinatol* 22: 12–14
17. Rubaltelli FF, Gourley GR, Loskamp N, Modi N, Roth-Kleiner M, Sender A, Vert P (2001) Transcutaneous bilirubin measurement: a multicenter evaluation of a new device. *Pediatrics* 107: 1264–1271
18. Shrout PE and JL F (1979) Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 86: 420–427
19. Szabo P, Wolf M, Bucher HU, Haensse D, Fauchère JC, Arlettaz R (2004) Assessment of jaundice in preterm neonates: comparison between clinical assessment, two transcutaneous bilirubinometers and serum bilirubin values. *Acta Paediatr* 93: 1–5
20. Wick RS (1998) Anwendung des Bilirubinometers während der Phototherapie bei ikterischen Neugeborenen, Thesis, Basel
21. Wong CM, van Dijk PJ, Laing IA (2002) A comparison of transcutaneous bilirubinometers: SpectRx BiliCheck versus Minolta AirShields. *Arch Dis Child Fetal Neonatal Ed* 87: F137–140
22. Yamanouchi I (1980) Transcutaneous bilirubinometry: Preliminary studies of non-invasive transcutaneous bilirubin meter in the Okayma National Hospital. *Pediatrics* 65: 195–202
23. Yamauchi Y, Yamanouchi I (1991) Factors affecting transcutaneous bilirubin measurement: effect of daylight. *Acta Paediatr Jpn* 33: 658–662