Journal of Perinatology (2010) 30, 616–621

www.nature.com/jp

ORIGINAL ARTICLE

Validity of neonatal jaundice evaluation by primary health-care workers and physicians in Karachi, Pakistan

L Hatzenbuehler[1], AKM Zaidi[2], S Sundar[2], S Sultana[2], F Abbasi[2], A Rizvi[1,2] and GL Darmstadt[1]

[1]Department of International Health, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA and [2]Department of Pediatrics and Child Health, Aga Khan University, Karachi, Pakistan

Objective: The purpose of this study was to validate primary health-care workers' and physicians' visual assessment of neonatal hyperbilirubinemia in Karachi, Pakistan.

Study Design: We compared primary health-care workers' and physicians' clinical identification of jaundice in infants <60 days old.

Result: Primary health-care workers identified 1- to 20-day-old neonates with hyperbilirubinemia ⩾15 mg per 100 ml (260 μmol l^{-1}) with 83.3% sensitivity and 50.5% specificity; neonates aged 1 to 6 days were identified with 76.2% sensitivity and 60.7% specificity. Physicians identified neonates aged 1 to 20 days with hyperbilirubimemia ⩾15 mg per 100 ml (260 μmol l^{-1}) with 51.4% sensitivity and 90.7% specificity, and neonates aged 1 to 6 days with 50% sensitivity and 88.5 % specificity. The primary health-care workers' and physicians' assessments showed fair interobserver agreement (κ statistic 0.29).

Conclusion: Primary health-care workers identified hyperbilirubinemic neonates with adequate sensitivity. With proper training and supervision, their assessment could improve the referral of hyperbilirubinemic neonates in low-resource settings in the developing world.
Journal of Perinatology (2010) **30,** 616–621; doi:10.1038/jp.2010.13; published online 1 April 2010

Keywords: jaundice; neonatal hyperbilirubinemia; Kramer scale

Introduction

Delay in the diagnosis and treatment of neonatal hyperbilirubinemia, due to inadequate health worker recognition of jaundice or lack of proper diagnostic equipment, is common in developing countries and leads to preventable cases of neurodevelopmental impairment, disability and sometimes death.[1] The risk of damage to the neurological system due to hyperbilirubinemia, a condition known as kernicterus, increases with (1) early exposure to high serum bilirubin levels, (2) prematurity or low birth weight, (3) rapid progression of jaundice, (4) high peak levels and (5) prolonged length of exposure to high levels.[2]

Currently, no international protocol exists to aid in primary health-care worker (HCW) identification of infants with clinically significant jaundice in developing country settings. Preliminary results from a study by Goswami *et al.*[3] in 2006, who studied the use of the World Health Organization's Integrated Management of Childhood Illness (IMCI) algorithm for the diagnosis of severe illness in young infants, showed that infants with severe jaundice were commonly missed. The authors concluded: 'the sensitivity of the IMCI algorithm can be further increased if yellowness of lower extremities/palms/soles is included in the algorithm.'[3]

One tool considered for integration into the IMCI algorithm is the clinical jaundice scale, originally introduced by Kramer.[4] The scale is used to predict the severity of hyperbilirubinemia based on the zone of skin discoloration in the neonate extending from the head (zone 1) to the palms and soles (zone 5), and assumes that an increasing zone of staining corresponds to increasing levels of serum bilirubin.[4] Studies by Ebbesen[5] found that neonatal jaundice appearing in Kramer scale zone 4 or higher (that is, icterus progression to the knees (zone 4) and then to the palms and soles (zone 5)) was correlated with clinically significant serum bilirubin levels, on average ⩾11.0 mg per 100 ml (188 μmol l^{-1}), and signaled the need for further evaluation and treatment at a tertiary care center.[5] The Kramer scale, or an adapted version, is widely used by physicians worldwide to predict the necessity of more sophisticated testing to rule out hyperbilirubinemia, and may be particularly advantageous in low-resource settings.

Previous studies have examined the clinical assessment of jaundice by various cadres of HCWs, but none have been conducted, similar to this study, in a developing country to compare HCW and physician (PHY) clinical evaluations of neonatal hyperbilirubinemia (Supplementary Table 1).[5–12] We evaluated the validity of an adapted Kramer scale when used by HCW and PHY to evaluate infants <60 days of age, with a focus

Correspondence: Dr GL Darmstadt, Family Health Division, Global Health Program, Bill & Melinda Gates Foundation, PO Box 23350, Seattle, WA 98102, USA.
E-mail: gary.darmstadt@gatesfoundation.org
Received 5 July 2009; revised 15 December 2009; accepted 22 December 2009; published online 1 April 2010

on neonates aged 1 to 20 days, presenting to community clinics in Karachi, Pakistan.

Methods

Data collection occurred between September 2004 and March 2005 as a prospective surveillance study involving a multiethnic population of infants presenting to three primary health-care clinics (Rehri Goth, Bilal Colony and Ibrahim Hyderi) in Karachi, Pakistan (area population $\sim 170\,000$). The study used the infrastructure of the Young Infant Clinical Signs Study aimed to identify clinical signs of illness signaling need for urgent referral to hospital.[13]

Infants <60 days of age who were brought to the study clinics for evaluation of suspected illness were initially consented and screened for clinical signs of illness, including jaundice, according to the Young Infant Clinical Signs Study protocol.[13] At each of the three primary health-care clinics, irrespective of their presenting clinical complaint, each infant was first evaluated by an HCW ($n = 3$), each of whom were lady health visitors with 2 years of formal health training with additional training in neonatal illness recognition. Each infant was then taken, in a blinded fashion, to a separate room of the clinic, and independently assessed by a PHY ($n = 5$) who had a minimum of 6 months of postgraduate residency work experience in pediatrics in a tertiary hospital. Infants were observed unclothed, under indoor lighting augmented by natural daylight from clinic windows. After visual assessment and application of skin pressure, we independently assigned jaundice scores to each infant (Supplementary Figure 1). Infants determined to have clinically significant signs of jaundice or requiring hospitalization for other reasons were referred to the National Institute of Child Health for appropriate laboratory testing and treatment. Blood for serum bilirubin sampling was collected and then transported in a cool container to the Aga Khan University Laboratory. Samples were processed with a reagent supplied by Beckman Synchron (Brea, CA, USA), and total bilirubin was measured by timed end point using the Diazo method (4.5% coefficient of variation).

Infants were excluded if they were (1) living outside the defined study area, (2) previously enrolled in the study, (3) recently hospitalized (<2 weeks), (4) referred from a health facility, (5) found to have an obvious lethal malformation, (6) in need of immediate cardiopulmonary resuscitation or (7) unable to provide informed consent. In addition, the sensitivity and specificity analysis excluded infants <1 day and >20 days old and those for whom a bilirubin sample was not obtained. The Kappa (κ) statistic analysis excluded those without two independent Kramer scale scores assigned by an HCW and a PHY.

Approval for this study was obtained from the institutional review board at the Johns Hopkins Bloomberg School of Public Health, and the ethical review committee at Aga Khan University.

Data analysis

In those infants who received serum bilirubin testing, sensitivity and specificity calculations compared the HCW and PHY use of the Kramer scale in neonates aged 1 to 20 days, and separately for neonates aged 1 to 6 days and those aged 7 to 20 days, using serum bilirubin as the gold-standard measure at two levels of $\geqslant 15$ mg per 100 ml ($260\ \mu mol\,l^{-1}$) and $\geqslant 20$ mg per 100 ml ($340\ \mu mol\,l^{-1}$). The serum bilirubin thresholds of 15 mg per 100 ml ($260\ \mu mol\,l^{-1}$) and 20 mg per 100 ml ($340\ \mu mol\,l^{-1}$) approximated levels of hyperbilirubinemia, based on a standard American Academy of Pediatrics Clinical Practice Guideline nomogram applicable to normal birth-weight infants, that signal need for referral, the former for consideration for phototherapy and the latter for exchange transfusion.[14] These levels are intermediate or high risk in any infant aged >24 h irrespective of other associated risk factors. Serum bilirubin levels <15 mg per 100 ml ($260\ \mu mol\,l^{-1}$) and $\geqslant 15$ mg per 100 ml ($260\ \mu mol\,l^{-1}$) were compared to assigned Kramer scale scores of 1 to 3 (jaundice of face, chest, abdomen or thighs, and considered to correlate with a serum bilirubin level <15 mg per 100 ml ($260\ \mu mol\,l^{-1}$)) and Kramer scale scores of 4 to 5 (jaundice of the distal extremities, including the legs below the knees (level 4) or the palms and soles (level 5), and indicative of need for referral for evaluation for hyperbilirubinemia, likely $\geqslant 15$ mg per 100 ml ($260\ \mu mol\,l^{-1}$)).[15] Serum bilirubin levels <20 mg per 100 ml ($340\ \mu mol\,l^{-1}$) and $\geqslant 20$ mg per 100 ml ($340\ \mu mol\,l^{-1}$) were compared to assigned Kramer scale scores of both 1 to 3 and 4 to 5, and 1 to 4 and 5 (jaundice extending to the palms and soles, considered as indicative of need for urgent referral and likely hyperbilirubinemia $\geqslant 20$ mg per 100 ml ($340\ \mu mol\,l^{-1}$)) (Supplementary Figure 1).

Interobserver percent agreement and weighted κ statistics compared the clinical assessment of each infant individually by both an HCW and a PHY, first in all infants aged <60 days, and then separately for neonates aged 1 to 6 days and 7 to 20 days. Calculation of agreement and κ statistics was based on two dichotomous categorizations of infants. The first percent agreement calculations examined jaundice of the distal extremities (Kramer scale score 4 or 5) and thus needing referral-level evaluation vs jaundice limited to the face, chest, abdomen and thighs (Kramer scale score 1 to 3). The second percent agreement calculations examined a Kramer scale score of 5 vs a score of 1 to 4. κ Statistics of <0, 0.0 to 0.20, 0.21 to 0.40, 0.41 to 0.60, 0.61 to 0.80, and 0.81 to 1.00 were considered as poor, slight, fair, moderate, substantial and almost perfect agreement, respectively.[16]

Statistical analysis was performed using Stata (version 10) software (StataCorp, College Station, TX, USA).[17]

Results

A total of 976 infants (Rehri Goth ($n = 304$), Bilal ($n = 549$) and Ibrahim Hyderi ($n = 123$)) aged <60 days were enrolled and

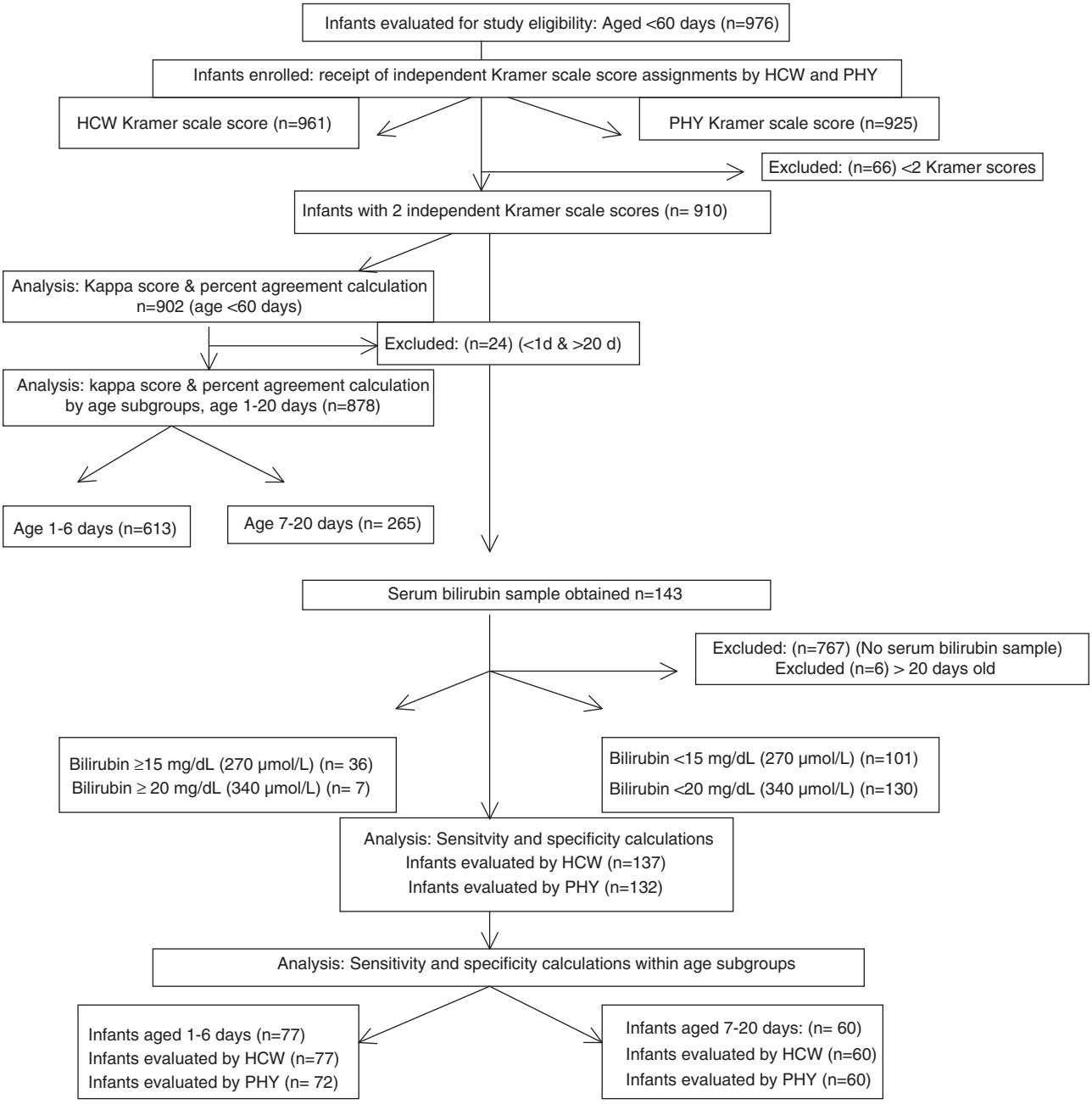

Figure 1 Sample flowchart.

evaluated for clinical signs of jaundice; 961 were evaluated and assigned a Kramer scale score by an HCW, and 925 were independently evaluated by a PHY (Figure 1). A total of 910 infants were assigned two independent Kramer scale scores. Among those in whom serum bilirubin sampling was obtained, over 98% were exclusively breastfed, and 90% were considered aseptic, acyanotic, non-tachypneic, non-hypoxic and had normal temperature (Supplementary Tables 2 and 3).

Among 193 infants (20.9%, 193/925) identified with jaundice, a serum bilirubin sample was collected in 143 infants; the remaining infants refused referral or did not provide consent for obtaining blood. Of these 143 infants, 137 were 1 to 20 days of age and were referred from the following study sites: Rehri Goth ($n = 14$), Bilal ($n = 119$) and Ibrahim Hyderi ($n = 4$). Of the three study sites, Bilal is considered the best equipped for laboratory blood draws, is located closest to the referral hospital and has a higher rate of referral acceptance. Among infants aged 1 to 20 days, the median and mean bilirubin levels were 11.8 mg per 100 ml (201 μmol l^{-1}) and 12.8 mg per 100 ml (216 μmol l^{-1}), respectively, and the range was 1 to 28.6 mg per 100 ml (17 to 489 μmol l^{-1}).

Thirty-six infants 1 to 20 days old (26.3%) were found to have a serum bilirubin level $\geqslant$15 mg per 100 ml (260 μmol l^{-1}), and 7 (5.1%) were found to have a serum bilirubin level $\geqslant$20 mg per 100 ml (340 μmol l^{-1}), which were considered high bilirubin levels necessitating referral irrespective of age of the infant (Supplementary Figures 2 and 3).

Sensitivity and specificity of clinical assessments of jaundice
Primary health-care workers. Among the 36 neonates 1 to 20 days of age with serum bilirubin level $\geqslant$15 mg per 100 ml (260 μmol l^{-1}), the HCWs assigned 30 neonates a Kramer scale score of 4 or 5, signaling need for referral (sensitivity 83.3%) (Table 1). HCW clinical assessments were less specific, as among 101 neonates with a serum bilirubin level <15 mg per 100 ml (260 μmol l^{-1}), 51 were assigned a Kramer scale score of 1 to 3 (specificity 50.5%). Similarly, the HCW assessments corresponded with sensitivity of 76.2 and 93.3%, but specificity of 60.7 and 37.8 % in the separate age groups, 1 to 6 days and 7 to 20 days, respectively.

Among the seven neonates who were found to have a serum bilirubin level $\geqslant$20 mg per 100 ml (340 μmol l^{-1}), the HCWs were 71.4% sensitive and 42.3% specific in assigning a Kramer scale score of 4 or 5 (Table 1). Similar results were found among the separate age categories of neonates. The HCWs were 57.1% sensitive and 81.7% specific in assigning these seven neonates a Kramer score of 5 (Table 2). Within the age category 1 to 6 days, HCWs were 60.0% sensitive and 84.7% specific in assigning a Kramer scale score of 5 to neonates who had a serum bilirubin level $\geqslant$20 mg per 100 ml, and 50% sensitive and 77.6% specific for neonates aged 7 to 20 days.

Physicians. Among the 36 neonates aged 1 to 20 days who had a serum bilirubin value $\geqslant$15 mg per 100 ml (260 μmol l^{-1}), PHYs evaluated 35 of the neonates and assigned 18 (sensitivity 51.4%) a Kramer score of 4 or 5 (Table 1). Conversely, among 97 neonates with a serum bilirubin level <15 mg per 100 ml (260 μmol l^{-1}), PHYs assigned a Kramer scale score of 1 to 3 in 88 (specificity 90.7%). Similarly, PHY assessments corresponded with moderate sensitivity (50.0 and 53.3%, respectively) and high specificity (88.5 and 93.3%, respectively) in the separate age groups, 0 to 6 days and 7 to 20 days.

Among the seven neonates who were found to have a serum bilirubin level $\geqslant$20 mg per 100 ml (340 μmol l^{-1}), PHYs were 57.1% sensitive in assigning them a Kramer scale score of 4 to 5, and 81.6% specific. Among those neonates aged 1 to 6 days, PHYs were 80.0% sensitive and 82.1% specific. Neither of the two infants aged 7 to 20 days was correctly assigned a Kramer scale score of 4 to 5, but PHYs correctly ruled out clinical signs of jaundice in 81.0%.

At a Kramer score cutoff of 5 (jaundice in the palms and soles), PHYs were 57.1% sensitive in assigning these seven neonates a

Table 2 Sensitivity and specificity of clinical assessments of jaundice reaching a Kramer score of 5 by HCWs and PHYs at serum bilirubin threshold $\geqslant$20 mg per 100 ml (340 μmol l^{-1})

Evaluator	*Sensitivity (%)*	*Specificity (%)*
All infants 1–20 days (n = 137)		
HCW (*n* = 137)	(4/7) 57.1	(106/130) 81.6
PHY (*n* = 132)	(4/7) 57.1	(119/125) 95.2
Infants 1–6 days (n = 77)		
HCW (*n* = 77)	(3/5) 60.0	(61/72) 84.7
PHY (*n* = 72)	(4/5) 80.0	(64/67) 95.5
Infants 7–20 days (n = 60)		
HCW (*n* = 60)	(1/2) 50.0	(45/58) 77.6
PHY (*n* = 60)	(0/2) 0	(55/58) 94.8

Abbreviations: HCW, primary health-care worker; PHY, physician.

Table 1 Sensitivity and specificity of clinical assessments of jaundice reaching a Kramer score of 4 or 5 by HCWs and PHYs at serum bilirubin thresholds $\geqslant$15 mg per 100 ml (260 μmol l^{-1}) and $\geqslant$20 mg per 100 ml (340 μmol l^{-1})

Evaluator	*Serum bilirubin threshold $\geqslant$15 mg per 100 ml (260 μmol l^{-1})*		*Serum bilirubin threshold $\geqslant$20 mg per 100 ml (340 μmol l^{-1})*	
	Sensitivity (%)	*Specificity (%)*	*Sensitivity (%)*	*Specificity (%)*
All infants 1–20 days (n = 137)				
HCW (*n* = 137)	(30/36) 83.3	(51/101) 50.5	(5/7) 71.4	(55/130) 42.3
PHY (*n* = 132)	(18/35) 51.4	(88/97) 90.7	(4/7) 57.1	(102/125) 81.6
Infants 1–6 days (n = 77)				
HCW (*n* = 77)	(16/21) 76.2	(34/56) 60.7	(3/5) 60.0	(37/72) 51.4
PHY (*n* = 72)	(10/20) 50.0	(46/52) 88.5	(4/5) 80.0	(55/67) 82.1
Infants 7–20 days (n = 60)				
HCW (*n* = 60)	(14/15) 93.3	(17/45) 37.8	(2/2) 100.0	(18/58) 31.0
PHY (*n* = 60)	(8/15) 53.3	(42/45) 93.3	(0/2) 0.0	(47/58) 81.0

Abbreviations: HCW, primary health-care worker; PHY, physician.

Table 3 Percent agreement and κ statistic calculations to compare the clinical use of the Kramer scale by HCW to PHY

	Percent agreement (%)	*Expected agreement (%)*	κ *statistic*	*Standard error*
Kramer score grouping: (1–3), (4/5)				
All infants <60 days ($n=902$)	67.7	54.4	0.29	0.029
Infants 1–20 days ($n=878$)	67.6	54.2	0.29	0.029
Infants 1–6 days ($n=613$)	68.9	55.3	0.30	0.037
Infants 7–20 days ($n=265$)	65.3	51.1	0.29	0.048

Kramer score of 5 and 95.2% specific (Table 2). Among the five neonates aged 1 to 6 days, PHYs correctly assigned 80.0% a Kramer score of 5; specificity was 95.5%. Only two infants aged 7 to 20 days had serum bilirubin levels $\geqslant 20$ mg per 100 ml (340 μmol l^{-1}), neither of whom was identified by PHYs, but 94.5% of those with serum bilirubin levels <20 mg per 100 ml (340 μmol l^{-1}) were correctly assigned a Kramer scale <5.

Intraobserver agreement and kappa statistics

There was fair interobserver agreement between HCWs and PHYs in the use of the Kramer scale when assessing jaundice in infants <60 days of age (Table 3). Agreement of 67.7% corresponded with Kramer scale assignments in zones 1, 2 or 3 and Kramer scale assignments 4 or 5, which resulted in a κ statistic of 0.29.[15] Similar results were seen among neonates 1 to 20 days, and in the two age groups, 1 to 6 days and 7 to 20 days.

Discussion

The primary limitations of this study are related to its conduct in a low-resource setting in a developing country community setting, where socioeconomic and cultural barriers impede neonatal care. Only 16% of the infants independently evaluated by HCW and PHY for any illness received serum bilirubin measurements (143/910). Ideally, all infants should have received a serum bilirubin measurement, regardless of whether they had evidence of clinical jaundice, to validate HCW and PHY visual assessments of neonatal hyperbilirubinemia in our population sample. However, blood was only obtained if a PHY identified the presence of any level of jaundice (Kramer score $\geqslant 1$) and the family provided consent for obtaining blood. Of the 910 newborns with two independent assessments by an HCW and a PHY, 17% (153/910) were judged by a PHY to have no jaundice (Kramer score 0); none of these patients had a blood sample taken for bilirubin determination. Of the 196 infants with a PHY Kramer score assignment of 4 to 5, only 13% (26/196) obtained a blood draw. Despite this limitation, our data provide some useful information on the performance of the Kramer scale as a screening tool in a low-resource setting.

Although previous studies have examined the clinical assessment of jaundice by HCWs, trained nurses and PHYs, none have been conducted in a developing country to compare HCW and PHY clinical evaluations of neonatal hyperbilirubinemia. Most previous studies on the clinical assessment of neonatal hyperbilirubinemia were conducted in the United States or Europe, were limited to a single or bi-ethnic study population and evaluated the clinical assessment of jaundice by a single observer compared to a serum bilirubin measurement taken in all the enrolled subjects (Supplementary Table 1).[5–12]

Riskin *et al*.[6] showed that neonatologists' identification of jaundice, and predicted serum bilirubin levels based on the Kramer scale, was highly correlated with serum total bilirubin levels ($r=0.682$, $P<0.001$). This study concluded: 'the clinical impression of jaundice by an…experienced clinician is a reliable method to assess newborns for … (neonatal hyperbilirubinemia) and serum bilirubin levels are necessary only in severely jaundiced neonates.' Similarly, Madlon-Kay *et al*.[7,8] found that nurses trained in clinical assessment of jaundice adequately ($r=0.61$, $P<0.01$) identified newborns with total serum bilirubin levels <12 mg per 100 ml (205 μmol l^{-1}), and all three infants with serum bilirubin levels $\geqslant 17$ mg per 100 ml (290 μmol l^{-1}) as having jaundice requiring referral to hospital. Szabo *et al*.[9,10] reported that trained nurses and PHYs were able to identify neonates with jaundice that did not reach the abdomen or extremities in whom hyperbilirubinemia did not exceed 250 μmol l^{-1}.

Alternatively, multiple studies have contested the validity of HCW diagnosis of hyperbilirubinemia using a jaundice scale. Knudsen and Brodersen[11] concluded that the yellow color of the skin is influenced by basal skin color, albumin binding capacity and serum pH, and is therefore an unreliable surrogate of plasma bilirubin levels. Studies by Moyer *et al*.[12] and Knudsen and Brodersen[11] have challenged the accuracy and reliability of Kramer's scale and found a poor correlation between PHY observations and total serum bilirubin levels ($r=0.16$ to 0.24). Despite these limitations, they concluded that PHY clinical observations were 97% specific in identifying infants who are not likely to have bilirubin levels >12 mg per 100 ml (205 μmol l^{-1}), based on verification that jaundice did not extend below the nipple line. On the basis of these and other studies, it is standard practice in the United States and abroad to clinically evaluate every infant for jaundice before both additional testing and discharge from the hospital. Many studies support the scale as a valid preliminary screening tool (Supplementary Table 1).[5–12]

This study showed that a clinical assessment tool used to identify infants with jaundice extending to the distal extremities shows some promise for guiding referral of infants with moderate to severe hyperbilirubinemia in a developing country setting, where sophisticated and expensive equipment is not widely available.[7–11] HCWs identified both infants aged 1 to 20 days (sensitivity 83.3%) and those <1 week old (aged 1 to 6 days) (sensitivity 76.2%) with

a serum bilirubin level $\geqslant 15$ mg per 100 ml ($260\ \mu mol\ l^{-1}$), based on identification of jaundice of the distal extremities, including the legs below the knees (Kramer level 4) or the palms and soles (Kramer level 5). Thus, few infants with clinical signs of moderate to severe hyperbilirubinemia would be missed, which in turn, prevents long-term health-care costs due to the associated morbidity and mortality resulting from untreated hyperbilirubinemia.[1]

PHYs were better than HCWs at correctly ruling out moderate to severe jaundice, thereby limiting the number of laboratory evaluations or serum bilirubin measures performed at the tertiary center. However, PHYs were also less sensitive in correctly identifying infants 1 to 20 days with serum bilirubin levels $\geqslant 15$ mg per 100 ml ($260\ \mu mol\ l^{-1}$). Once patients are identified as having jaundice, referred by HCWs from peripheral primary care clinics and overcome care seeking barriers to arrive at the tertiary center for evaluation, a lower threshold of suspicion on the part of the PHY, and thus for serum sampling for hyperbilirubinemia, may be indicated.

This study suggests that HCWs trained in the clinical assessment of jaundice have potential for improving the identification of young infants who are at risk of severe consequences of hyperbilirubinemia.[13] Infants aged 1 to 20 days old with yellow discoloration extending below the knees or to the palms on the upper extremities should be considered at high risk of moderate to severe hyperbilirubinemia and require referral to a tertiary setting for further evaluation.

Conflict of interest

The investigators have no competing interests and received no financial gain from the study results.

Acknowledgments

We thank the study participants from the three study communities for their willingness to participate in the study; the lady health visitors (Razia Awaldad, Nasira A Jabbar, Mussarat Omer); Save the Children-US (through a grant from the Bill & Melinda Gates Foundation) who provided funding for the data collection and analysis; Martin Weber and Rajiv Bahl from the Child and Adolescent Health and Development Department at the World Health Organization, David Hamer at Boston University and John Carlin at the University of Melbourne, for their technical help in study design, and Dr Marie Deiner-West and Dr John McGready of the Biostatistics Department, and Dr Rosa Crum of the Epidemiology Department at the Johns Hopkins Bloomberg School of Public Health for their professional advice in the data analysis and reporting of results.

References

1 Narayanan I, Banwalikar J, Mehta R, Ghorpade M, Peesay MR, Nanda S *et al*. A simple method of evaluation of jaundice in the newborn. *Ann Trop Paediatr* 1990; **10**: 31–34.
2 Connolly AM, Volpe JJ. Clinical features of bilirubin encephalopathy. *Clin Perinatol* 1990; **17**: 371–379.
3 Goswami V, Kumar DA, Singh V, Chandra J. Evaluation of simple clinical signs of illness in young infants (0–2 months) and its correlation with WHO IMCI algorithm (7 days to 2 months). *Indian Pediatr* 2006; **43**: 1042–1049.
4 Kramer LI. Advancement of dermal icterus in the jaundiced newborn. *Am J Dis Child* 1969; **118**: 454–458.
5 Ebbesen F. Relationship between the cepahlo-pedal progress of clinical icterus and the serum bilirubin concentration in newborn infants without blood type sensitization. *Acta Obstet Gynecol Scand* 1975; **54**: 329–332.
6 Riskin A, Kuglman A, Abend-Weinger M, Green M, Hemo M, Bader D. In the eye of the beholder: how accurate is clinical estimation of jaundice in newborns? *Acta Paediatr* 2003; **92**: 574–576.
7 Madlon-Kay DJ. Recognition of the presence and severity of newborn jaundice by parents, nurses, physicians, and icterometer. *Pediatrics* 1997; **100**: E3.
8 Madlon-Kay DJ. Home health nurse clinical assessment of neonatal jaundice: comparison of 3 methods. *Arch Pediatr Adolesc Med* 2001; **155**: 583–586.
9 Szabo P, Wolf M, Ulrich Bucher H, FauchŸre JC, Haenesse D, Arlettaz R. Detection of hyperbilirubinaemia in jaundiced full-term neonates by eye or by bilirubinometer. *Eur J Pediatr* 2004; **163**: 722–727.
10 Szabo P, Wolf M, Ulrich Bucher HU, Haenesse D, FauchŸre JC, Arlettaz R. Assessment of jaundice in preterm neontates: comparison between clinical assessment, two transcutaneous bilirubinometers and serum bilirubin values. *Acta Paediatr* 2004; **93**: 1491–1495.
11 Knudsen A, Brodersen R. Skin colour and bilirubinemia in neonates. *Arch Dis Child* 1989; **64**: 605–609.
12 Moyer VA, Ahn C, Sneed S. Accuracy of clinical judgement in neonatal jaundice. *Arch Pediatr Adolesc Med* 2000; **154**: 391–394.
13 Young Infants Clinical Signs Study Group. Clinical signs predicting severe illness in young infants: a multicentre study. *Lancet* 2008; **12**: 135–142.
14 Subcommittee on Hyperbilirubinemia. Management of hyperbilirbubinemia in the newborn infant 35 or more weeks of gestation. *Pediatrics* 2004; **114**: 297–316.
15 Kaplan M, Hammerman C. Understanding and preventing severe neonatal hyperbilirubinemia: is bilirubin neurotoxicity really a concern in the developed world?. *Clin Perinatol* 2004; **31**: 555–575, x.
16 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**: 159–174.
17 StataCorp. *Stata Statistical Software: Release 10.0*. Stata Corporation, College Station: TX, 2005.

Supplementary Information accompanies the paper on the Journal of Perinatology website (http://www.nature.com/jp)