

Validation of caregiver interviews to diagnose common causes of severe neonatal illness

Henry D. Kalter^a, Munir Hossain^b, Gilbert Burnham^a, Naila Z. Khan^b, Samir K. Saha^b, Md Anwar Ali^c and Robert E. Black^a

^a*Department of International Health, Johns Hopkins School of Hygiene and Public Health, Baltimore, USA, ^bBangladesh Institute of Child Health, Dhaka Shishu Hospital, Dhaka, and ^cInfectious Disease Hospital, Dhaka, Bangladesh*

Summary. The objective of this study was to validate retrospective caregiver interviews for diagnosing major causes of severe neonatal illness and death. A convenience sample of 149 infants aged < 28 days with one or more suspected diagnoses of interest (low birthweight/severe malnutrition, preterm birth, birth asphyxia, birth trauma, neonatal tetanus, pneumonia, meningitis, septicaemia, diarrhoea, congenital malformation or injury) was taken from patients admitted to two hospitals in Dhaka, Bangladesh. Study paediatricians performed a standardised history and physical examination and ordered laboratory and radiographic tests according to study criteria. With a median interval of 64.5 days after death or hospital discharge, caregivers of 118 (79%) infants were interviewed about their child's illness. Using reference diagnoses based on predefined clinical and laboratory criteria, the sensitivity and specificity of particular combinations of signs (algorithms) reported by the caregivers were ascertained. Sufficient numbers of children with five reference standard diagnoses were studied to validate caregiver reports. Algorithms with sensitivity and specificity > 80% were identified for neonatal tetanus, low birthweight/severe malnutrition and preterm delivery. Algorithms with specificities > 80% for birth asphyxia and pneumonia had sensitivities < 70%, or alternatively had high sensitivity with lower specificity. In settings with limited access to medical care, retrospective caregiver interviews provide a valid means of diagnosing several of the most common causes of severe neonatal illness and death.

Address for correspondence: Dr Henry Kalter, Department of International Health, Johns Hopkins School of Hygiene and Public Health, 615 N. Wolfe St., Suite E8132, Baltimore, MD, 21205, USA.

Introduction

Determining the levels and medical causes of severe illness and death in a locality can help health planners identify service priorities, effectively allocate sparse resources and evaluate the impact of health care programmes. Neonates < 28 days of age presently contribute nearly half of all infant deaths in developing countries.¹ Because intervention programmes aimed at infectious diseases are decreasing mortality among older infants and children,¹ the relative contribution of neonatal mortality is increasing, and it is becoming more important to gather accurate data about neonatal events. Despite its known limitations,² verbal autopsy is the most useful tool available to accomplish this task in developing country settings with limited access to medical care and infrequent certification of cause of death.³

In a verbal autopsy inquiry of a child's death or severe illness, the mother or other main caregiver is retrospectively interviewed about her child's condition and the events surrounding the illness. The cause or causes are determined from predefined combinations of the reported signs and symptoms (algorithms) or by independent classification of the interview findings by a team of physicians. The method has been directly validated against medical reference standard diagnoses for both deaths^{2,4-6} and severe illnesses,^{7,8} and has been found to work best in identifying distinctive and memorable syndromes such as neonatal tetanus, measles and injuries, but also to perform reasonably well for other conditions including diarrhoea and pneumonia.

However, there has been only one validation study that included a variety of neonatal diagnoses,⁶ one other that investigated neonatal tetanus and sepsis,² and a third that examined neonatal tetanus.⁴ All were affected by methodological problems that have limited the utility of verbal autopsy validation studies in general,⁹⁻¹¹ including inadequate control over the quality and reproducibility of reference standards, restricted comparability of findings across settings due to the use of varied methods and instruments, and a possible bias in the application of findings from hospital studies of deaths to population-based surveys, as many children who survive in hospital might have died in a community setting without advanced medical care.

We undertook the present study to validate directly the ability of the verbal autopsy method to diagnose several important causes of severe neonatal illness and death, and to address some of the major methodological problems that have hampered prior validation studies.

Methods

The newborn children reported on in this paper were seen as part of a prospective study of children < 5 years of age who were admitted to the Dhaka Shishu (Children's) Hospital or the Infectious Disease Hospital (IDH) in Dhaka,

Bangladesh, from September 1994 to February 1995. Three experienced Bangladeshi paediatricians at the Shishu Hospital, and one at the IDH, enrolled newborns aged < 28 days with one or more suspected diagnoses of interest (low birthweight [LBW]/severe malnutrition, premature birth, birth asphyxia, birth trauma, neonatal tetanus, pneumonia, meningitis, bacteraemia or septicaemia with no known focus, diarrhoea, congenital malformation or injury), evaluated them using a standardised history and physical format that emphasised problems of neonates, and ordered laboratory and radiographic tests according to study criteria.

The study was conducted at two hospitals because all patients seen in Dhaka with suspected neonatal tetanus are sent to the IDH. At the Shishu Hospital, the study paediatricians were not in charge of the patients' care and were not able to evaluate all admitted patients. During the daytime and evening shifts they assessed patients with any suspected diagnosis of interest as soon as they became aware of the admission. Children admitted overnight were seen the next morning. The IDH study physician was in charge of the patients' care, and all patients admitted with a suspected diagnosis of neonatal tetanus were enrolled in the study.

After obtaining informed consent, the study paediatricians evaluated and enrolled newborn infants and ordered laboratory and radiographic tests based on the following criteria:

- **suspected pneumonia:** a chest radiograph was ordered for **either** severe lower chest wall indrawing **or** > 60 breaths per minute **or** $36^{\circ}\text{C} > \text{rectal temperature} > 38^{\circ}\text{C}$;
- **suspected meningitis:** a lumbar puncture was ordered for $36^{\circ}\text{C} > \text{rectal temperature} > 38^{\circ}\text{C}$ **and either** neck stiffness **or** a Young Infant Observation Scale¹² score ≥ 7 , **or** for convulsions witnessed in the hospital without suspected tetanus;
- **suspected bacteraemia:** a blood culture was ordered for $36^{\circ}\text{C} > \text{rectal temperature} > 38^{\circ}\text{C}$;
- **suspected diarrhoea:** dehydrated newborns were enrolled if the caregiver gave a history of liquid or watery stools ≥ 3 days;
- **suspected tetanus:** a history of normal suck or cry at birth followed by poor suckling or loss of cry starting on or after the second day of life, **or** medically witnessed inability to suck, trismus, muscle spasms or convulsions;
- **low birthweight/severe malnutrition or premature birth:** **either** a weight-for-age Z-score < -3 (calculated by comparison with the National Center for Health Statistics reference population¹³) **or** a medical history of birth at < 37 weeks' gestation **or** a medical examination compatible with a gestational age < 37 weeks;
- **suspected congenital malformation:** a physical examination compatible with malformation;

- **suspected birth asphyxia:** an altered neurological status and $36^{\circ}\text{C} \leq \text{rectal temperature} \leq 38^{\circ}\text{C}$;
- **suspected birth trauma:** a physical examination consistent with birth trauma; and
- **suspected injury:** a history of injury or a physical examination consistent with injury.

The study paediatricians were also asked to formulate up to four discharge diagnoses for each enrolled child based on their clinical judgement and the findings of laboratory and radiographic investigations.

With a minimum interval of 30 days after the child's death or discharge from hospital, caregivers were interviewed with a structured questionnaire about the signs and symptoms of their child's illness that led, respectively, to death or to hospitalisation. We collaborated with colleagues at the London School of Hygiene and Tropical Medicine, the World Health Organisation, and the Kenya Medical Research Center in Kilifi to develop a standardised caregiver interview format for use in a variety of study settings. The questionnaire was translated from English to Bengali by a Bangladeshi physician; the translation was reviewed by a local anthropologist for appropriate use of lay terminology, and the final translation was independently back-translated by a second physician to ascertain that the Bengali version conveyed accurately the original's use of medical concepts. The interviews were conducted by four native Bengali-speaking women of childbearing age who had at a minimum a secondary school education and were supervised by a Bangladeshi female social worker with an extensive knowledge of the local community. In a 1-week course, two of the investigators (HDK and MH) trained the interviewers and the supervisor in the interview procedures and the particulars of the verbal autopsy format.

To measure the sensitivity and specificity of particular combinations of signs and symptoms (algorithms) reported by the caregivers, we compared the caregiver reports to predefined medical reference standard diagnoses based on clinical, laboratory and radiographic criteria (Table 1) developed in collaboration with the above-mentioned colleagues. When determining the validity of the verbal autopsy diagnosis of any particular condition, all newborns without that condition served as the comparison group. Children 1 month to 5 years of age, enrolled in a companion study of older children, were included as part of a comparison group only for the analysis of neonatal tetanus. To gain a better understanding of how the disease mix in a population might affect the verbal autopsy's specificity, we also examined the reference standard diagnoses of patients incorrectly identified by caregiver reports as having the diagnosis of interest. Sensitivities, specificities and 95% exact binomial confidence intervals were calculated using Epi Info computer software.¹⁴ Differences between groups were assessed by the chi-square or Fisher's exact test.

Table 1. Criteria for medical reference standard diagnoses

Age at admission < 28 days *plus*:

Pneumonia: (a) chest X-ray diagnosis of infectious pneumonia *or* (b) physician final diagnosis of infectious pneumonia *plus either* crepitations *or* lower chest wall indrawing on examination

Bacterial meningitis: bacteria grown on culture of cerebrospinal fluid *or* bacteria seen on Gram stain of cerebrospinal fluid *or* > 100 leucocytes/mm³ cerebrospinal fluid with > 80% polymorphonuclear leucocytes *or* positive latex agglutination test of cerebrospinal fluid

Acute diarrhoea: caregiver history of liquid or watery stools or diarrhoea for < 14 days *plus either* medically noted liquid, semi-liquid or watery stools *or* medically noted dehydration *plus* no blood observed in stools by medical workers

Acute dysentery: caregiver history of liquid or watery stools or diarrhoea for < 14 days *plus* medically noted liquid, semi-liquid or watery stools *plus* medically noted blood in the stools

Persistent diarrhoea or dysentery: caregiver history of liquid or watery stools or diarrhoea for ≥ 14 days *plus either* medically noted liquid or semi-liquid or watery stools *or* medically noted blood in the stools

Dehydration/severe dehydration: physician diagnosis of dehydration/severe dehydration

Local bacterial infection: red or purulent umbilicus *or* skin pustules on medical examination

Bacteraemia or septicaemia with no known focus of infection: rectal temperature > 38°C *or* < 36°C *plus* positive blood culture *plus either* (for septicaemia) non-consolable irritability, abnormally sleepy or difficult to wake, mottled and cool extremities, *or* pale and shocky on examination (minus all of these signs for bacteraemia) *minus* reference standard pneumonia, bacterial meningitis, acute or persistent diarrhoea or dysentery, and local bacterial infection

Neonatal tetanus: age at admission > 2 days *plus* physician final diagnosis of neonatal tetanus

Low birthweight/severe malnutrition: weight-for-age Z-score < -3 on admission to hospital

Preterm birth: medically documented preterm birth (gestational age < 37 weeks)

Birth asphyxia: medical history of birth asphyxia, evidenced by documented failure to breathe spontaneously at birth *or* 20-minute Apgar score < 4 *plus either* lethargy, coma, hypotonia *or* seizures on examination *plus* 38°C < rectal temperature never < 36°C

Birth trauma: physician final diagnosis of birth trauma

Congenital malformation: congenital abnormality noted on medical examination

The study was approved by the ethical review committees of the Dhaka Shishu Hospital and the Johns Hopkins School of Hygiene and Public Health, and by the medical director of the IDH.

Results

A total of 149 newborn infants aged < 28 days were enrolled in the study. Caregivers of 118 (79%) of the infants were located and interviewed in the home from 31 to 232

days (median = 64.5) after death or hospital discharge; 98 (83%) of the respondents were the children's mothers. The median age of the 118 infants at admission was 5.5 days (range 0–27); 81 (69%) were male. The median age (8.0 days, $P=0.17$) and sex (68% male, $P=0.92$) of those lost to follow-up were similar to the fully studied infants. The two groups differed only in that neonatal tetanus was more common among those lost to follow-up (35% vs. 17%, $P=0.03$). Most of these infants came from an area that was difficult to reach from Dhaka. The remainder of this paper presents the findings for the 118 infants with both a medical examination and a caregiver interview.

The infants had a wide range of medical reference standard diagnoses (Table 2). LBW/severe malnutrition was the most common, present in 46 (39%) of the infants. Neonatal tetanus, birth asphyxia, preterm birth and pneumonia were diagnosed in 16–27% of the infants. Several had multiple diagnoses, most often LBW/severe malnutrition combined with preterm birth or another condition. The mortality rate was high: 11/20 (55%) for neonates with tetanus and 11/98 (11%) for the others; 20 of the deaths occurred in hospital and two more within 24 h of discharge. Another nine infants died before the home interview, six of these within 9 days of hospital discharge and three between 10 and 31 days, but these were not counted as deaths in the analysis.

Table 2. Medical reference standard diagnoses of 22 deceased and 96 surviving neonates (< 28 days old), Dhaka, Bangladesh

	Deaths ^a	Survivors
Pneumonia (2 with birth asphyxia, 1 with birth trauma, one with malformation)	2	30
Birth asphyxia ^b (20; 1 with malformation) and/or birth trauma (6; 1 with local bacterial infection)	5	18
LBW/severe malnutrition ^c (21; 15 preterm) or malnutrition ^d alone	2	21
Neonatal tetanus (1 with malformation, 1 with local bacterial infection)	11	9
Bacteraemia (1) or septicaemia ^e with no known focus	1	9
Diarrhoea (4 acute watery [3 with severe dehydration], 2 persistent with dehydration)	0	6
Bacterial meningitis (1 with birth asphyxia, 1 with birth traum)	0	2
Local bacterial infection	1	1

^aDeath in hospital or within 24 h of discharge.

^bOne death and three survivors had only a physician final diagnosis of birth asphyxia.

^cLow birthweight/severe malnutrition (weight-for-age Z-score < -3).

^d-3 ≤ weight-for-age Z-score < -2; 25 and 22 additional neonates, respectively, had LBW/severe malnutrition (14 preterm) and malnutrition (1 preterm).

^eOne survivor had only a physician final diagnosis of septicaemia with no reference standard diagnosis of infectious pneumonia, bacterial meningitis, acute or persistent diarrhoea or dysentery, or local bacterial infection.

The sensitivity and specificity of caregiver-reported signs for preterm birth and LBW/severe malnutrition were similar (Table 3). This was because 29/30 preterm infants were also severely malnourished, although the comparison group for

Table 3. Sensitivity and specificity of caregiver-reported signs and symptoms, compared with medical reference standard diagnosis of premature birth or low birthweight/severe malnutrition,^a Dhaka, Bangladesh

	Reference diagnosis		<i>n</i> ^b	Sensitivity % [95% CI]	<i>n</i> ^b	Specificity % [95% CI]
	<i>n</i> with	<i>n</i> without				
Preterm delivery	30	65				
Pregnancy ended early			30	90 [73, 98]	65	78 [67, 88]
Baby was very small or smaller than usual at birth			30	93 [78, 99]	65	68 [55, 79]
Pregnancy ended early <i>and</i> baby was very small or smaller than usual at birth			30	87 [69, 96]	65	85 [74, 92]
Pregnancy ended early <i>or</i> baby was very small or smaller than usual at birth			30	97 [83, 100]	65	62 [49, 73]
Low birthweight/severe malnutrition	46	72				
Pregnancy ended early			45	78 [63, 89]	71	83 [72, 91]
Baby was very small or smaller than usual at birth			45	89 [76, 96]	71	85 [74, 92]
Pregnancy ended early <i>and</i> baby was very small or smaller than usual at birth			45	76 [60, 87]	71	96 [88, 99]
Pregnancy ended early <i>or</i> baby was very small or smaller than usual at birth			45	91 [79, 98]	71	72 [60, 82]
Low birth weight/severe malnutrition <i>or</i> preterm delivery	47	50				
Pregnancy ended early <i>and</i> baby was very small or smaller than usual at birth			46	74 [59, 86]	50	94 [83, 99]
Pregnancy ended early <i>or</i> baby was very small or smaller than usual at birth			46	91 [79, 98]	50	74 [60, 85]

^aWeight-for-age Z-score < -3.

^bThe number of children for whom sufficient caregiver interview data were available to determine the verbal autopsy diagnosis.

preterm infants included 15 newborns with LBW/severe malnutrition that were not preterm. This resulted in several algorithms having nearly identical sensitivity and specificity for LBW/severe malnutrition alone or combined with premature birth. Perhaps the best was 'pregnancy ended early *and* baby was very small or smaller than usual at birth'. These same algorithms were considerably less sensitive for severe or moderate malnutrition (weight-for-age Z-score < -2), with comparable specificity. For example, the sensitivity/specificity of 'pregnancy ended early *and* baby was very small or smaller than usual at birth' was 52%/98%.

Signs and symptoms reported by caregivers also had high sensitivity and specificity for the medical diagnosis of neonatal tetanus (Table 4). The algorithm 'age < 28 days *and* convulsions or spasms' was 100% sensitive and 75% specific. Further restricting the age criteria and adding requirements about the newborn's crying and sucking ability increased the specificity to 89–91%, but at some cost to sensitivity. There was a tendency for these algorithms to be somewhat more specific in survivors than deaths, although none of the differences was statistically significant. For example, the specificity of 'age < 28 days *and* convulsions or spasms' was, respectively, 77% and 56% in the 84 survivors and nine deaths from neonatal tetanus ($P = 0.22$). When children of 1 month to 5 years of age from the companion study of older children were included in the comparison group, the specificity of the algorithms was even greater, 95–99%; this was driven by the specificity of the age criterion, 'age 3–27 days', which by itself had a specificity of 89%.

Table 4. Sensitivity and specificity of caregiver-reported signs and symptoms, compared with medical reference standard diagnosis of neonatal tetanus, Dhaka, Bangladesh

	Reference diagnosis			Sensitivity % [95% CI]	n ^a	Specificity % [95% CI]
	n with	n without	n ^a			
Neonatal tetanus	20	98				
Age < 28 days and 'tetanus' (local term)			17	88 [64, 99]	75	95 [87, 99]
Age < 28 days and convulsions or spasms			18	100 [81, 100]	93	75 [65, 84]
Age 3–27 days and convulsions or spasms and able to suckle <i>and</i> cry normally after birth and stopped suckling or crying			18	67 [41, 87]	94	91 [84, 96]
Age 3–27 days and convulsions or spasms and able to suckle <i>or</i> cry normally after birth and stopped suckling or crying			18	83 [59, 96]	95	89 [81, 95]

^aThe number of children for whom sufficient caregiver interview data were available to determine the verbal autopsy diagnosis.

Table 5. Sensitivity and specificity of caregiver-reported signs and symptoms, compared with medical reference standard diagnosis of pneumonia, Dhaka, Bangladesh

	Reference diagnosis			Sensitivity % [95% CI]	n ^a	Specificity % [95% CI]
	n with	n without	n ^a			
Pneumonia	32	86				
Cough > 1 day and difficult breathing > 1 day			32	53 [35, 71]	79	90 [79, 95]
Fast breathing and chest indrawing			29	69 [49, 85]	64	81 [70, 90]
Fast breathing > 1 day			28	71 [51, 87]	64	75 [63, 85]
Child had 'pneumonia' (local term)			27	74 [54, 89]	52	87 [74, 94]
Cough <i>or</i> difficult breathing and fast breathing <i>or</i> chest indrawing			30	80 [61, 92]	64	59 [46, 71]

^aThe number of children for whom sufficient caregiver interview data were available to determine the verbal autopsy diagnosis.

'Convulsions or spasms' alone presented great difficulty in differentiating neonatal tetanus from other reference standard diagnoses of newborn infants. Of the 23 newborns without tetanus who were reported by their caregivers to have had convulsions or spasms during the illness, 12 in fact had convulsions on admission to the hospital. Seven of the 12 were febrile. The reference diagnoses of these infants were pneumonia (4), septicaemia with no known focus (3), LBW/severe malnutrition with preterm delivery (2), meningitis (1), local bacterial infection (1) and birth trauma (1). The 11 neonates without medically observed convulsions had reference standard birth asphyxia (4), pneumonia (4), acute diarrhoea with severe dehydration (1), persistent diarrhoea (1) and LBW/severe malnutrition with premature birth (1). None was febrile on admission.

'Stopped suckling after the second day of life' was also problematic in the interview-based diagnosis of tetanus. The study physicians found 'no suckling' at the time of hospital admission in 21 of the 26 newborns without reference standard tetanus who were reported by their caregivers to have this sign. Taken together, the physicians confirmed the presence of convulsions or no suckling in 29/39 (74%) of the neonates without tetanus whose caregivers reported one or both of these signs.

Pneumonia was more difficult to diagnose by verbal autopsy, generally requiring a sacrifice in sensitivity to achieve higher specificity, or vice versa (Table 5). Still, it was possible to realise a sensitivity/specificity mix of 69%/81% with the

Table 6. Sensitivity and specificity of caregiver-reported signs and symptoms, compared with medical reference standard diagnosis of birth asphyxia, Dhaka, Bangladesh

	Reference diagnosis			Sensitivity % [95% CI]	n ^a	Specificity % [95% CI]
	n with	n without	n ^a			
Birth asphyxia	19	86				
No fever and not able to cry after birth and either not able to breathe after birth or not able to suckle normally after birth or convulsions/spasms and no tetanus by caregiver report (Table 4, algorithm 3)			15	53 [27, 79]	74	82 [72, 90]
No fever and either convulsions/spasms or not able to cry after birth or not able to suckle normally after birth			15	60 [32, 84]	71	63 [51, 75]
Not able to cry after birth and either not able to breathe after birth or not able to suckle normally after birth			15	73 [45, 92]	75	72 [60, 82]
and no tetanus by caregiver report (Table 4, algorithm 4)			15	73 [45, 92]	76	75 [64, 84]
Not able to cry after birth and either convulsions/spasms or not able to breathe after birth or not able to suckle normally after birth			15	87 [60, 98]	75	69 [58, 79]
and no tetanus by caregiver report (Table 4, algorithm 4)			15	87 [60, 98]	76	72 [61, 82]

^aThe number of children for whom sufficient caregiver interview data were available to determine the verbal autopsy diagnosis.

algorithm 'fast breathing *and* chest indrawing'. While 25 caregivers could not provide the information needed to determine this verbal autopsy diagnosis, only three of their infants had reference standard pneumonia. Algorithms consisting of 'cough and difficult breathing', including the duration of these signs in various iterations, had high specificity but very low sensitivity. Algorithms based on the duration of difficult breathing alone, or on the duration of fast breathing alone, did not approach the performance of the algorithms shown in Table 5.

Of the 12 infants without pneumonia who were reported by their caregivers to have had 'fast breathing *and* chest indrawing' during the illness, two had respiratory rates ≥ 60 per minute and had severe lower chest wall indrawing on admission to the hospital. Their reference diagnoses were neonatal tetanus and

LBW/severe malnutrition with preterm delivery. None of the other 10 newborns in this group had tachypnoea or indrawing that was observed in the hospital. Their reference diagnoses included birth asphyxia and birth trauma (3), acute diarrhoea with severe dehydration (2), persistent diarrhoea (1), neonatal tetanus (1), LBW/severe malnutrition with preterm delivery (1), septicaemia with no known focus (1) and local bacterial infection (1).

Birth asphyxia could be detected by verbal autopsy with equal or higher sensitivity than pneumonia but somewhat lower specificity (Table 6). Several algorithms had both sensitivity and specificity in the 70–80% range. The algorithm 'not able to cry after birth *and either* convulsions or spasms *or* not able to breathe after birth *or* not able to suckle normally after birth' had the highest sensitivity (87%), while maintaining specificity of 69%. Adding the exclusion of a caregiver report of neonatal tetanus (algorithm 4 in Table 4) boosted the specificity to 72%. Of the 14 infants for whom the caregivers could not provide sufficient data to determine the verbal autopsy diagnosis of birth asphyxia, four actually had this reference standard diagnosis.

Of the 21 neonates with a verbal autopsy diagnosis of birth asphyxia (algorithm 6 in Table 6), but without reference standard birth asphyxia, four were reported by their caregivers to have had 'convulsions or spasms' during their illness and three of these four were convulsing on admission to the hospital. All three had reference standard septicaemia with no known focus. Of the same 21 newborns, four were said to be 'not able to suckle normally after birth', and three of the four were actually unable to suck when seen at the hospital; their reference diagnoses were meningitis (1), LBW/severe malnutrition with preterm delivery (1) and pneumonia (1). Seventeen of the infants were said to 'not be able to breathe after birth', but only three of these were assessed by a study paediatrician to be in respiratory failure; their reference diagnoses were LBW/severe malnutrition (2, 1 born preterm) and acute diarrhoea with severe dehydration (1). The reference diagnoses of the remaining 12 infants in this group included LBW/severe malnutrition born preterm (2), pneumonia (2), birth asphyxia (2), septicaemia (2) or bacteraemia with no known focus (1), acute diarrhoea with severe dehydration (1), birth trauma (1) and malnutrition (1).

Discussion

We undertook this study to validate the verbal autopsy method of diagnosing several causes of severe neonatal illness and death in developing countries. While we took a convenience sample of newborn infants from two paediatric hospitals in a capital city, the study children suffered from a broad array of reference standard diagnoses whose distribution was similar to that found in several community investigations.³ Therefore, our sensitivity and specificity estimates for the verbal autopsy diagnoses are likely to be as unbiased as possible in a hospital validation study.

Retrospective caregiver reports of their newborn infants' signs and symptoms during a severe or fatal illness were able to identify LBW/severe malnutrition and preterm birth with high sensitivity and specificity. We assumed that infants < 28 days old with a weight-for-age Z-score < -3 were severely malnourished from birth, and demonstrated that algorithms based on the child's appearance at birth were highly valid when compared with the current weight. Caregiver reports were more sensitive for preterm birth than for LBW because these newborns had both problems, but were less specific for this diagnosis because the comparison group included the LBW infants who were not born preterm.

Marsh *et al.*⁶ in Pakistan used birthweight < 2500 g as their LBW reference standard, and found that the sensitivity of 'baby smaller than usual' (includes 'very small' in their study) was 72%, compared with our finding of 89%. This discrepancy is probably because of our use of a more severe reference standard; sensitivity for reference standard weight-for-age Z-score < -2 was also lower (67%) in our study. This suggests that caregiver reports of body size may be insensitive to less than very LBW or severe malnutrition in settings where many or most of the children in their experience have been malnourished.

Algorithms for the diagnosis of neonatal tetanus were highly sensitive and specific when all young children < 5 years of age were included in the comparison group – most of these comparison children were successfully excluded from the verbal autopsy diagnosis by age criteria alone – but a greater trade-off between sensitivity and specificity was required to identify and distinguish tetanus among the newborns. These values more accurately reflect the performance of verbal autopsy in the community for the diagnosis of neonatal tetanus.

Many other neonatal medical conditions, including premature birth, birth asphyxia and septicaemia, are associated with a cardinal sign of neonatal tetanus, convulsions. Indeed, the study paediatricians observed convulsions in more than half of the newborns without tetanus whose caregivers reported that they had this sign. Another critical sign in diagnosing neonatal tetanus, the cessation of suckling after the second day of life, was commonly reported by caregivers and confirmed by the study physicians in neonates without tetanus. These false-positive verbal autopsy diagnoses therefore appear to have been largely caused by the overlap in the clinical presentations of several medical conditions of newborn infants, not by faulty recognition or recall of signs and symptoms by the caregivers.

Several validation studies have shown that pneumonia is difficult to diagnose by verbal autopsy,^{2,4,5,7} although none of these studies specifically investigated this diagnosis in neonates. Only three of the newborns in Pakistan⁶ had a reference standard diagnosis of acute respiratory tract infection. Our findings therefore are the first reported data specifically on the verbal autopsy diagnosis of pneumonia in neonates. Algorithms to detect pneumonia performed favourably compared with the studies in older infants and children, although, as expected, cough was a less sensitive sign in the neonates. Many caregivers could not provide the

information about fast breathing or chest indrawing needed to make the verbal autopsy diagnosis using these signs, but almost all of their infants did not have pneumonia. This suggests that the presence of these clinical signs is more memorable than their absence, and that a 'don't know' response might be considered equivalent to a 'no' response. This possibility should be further tested in other data sets.

Birth asphyxia was also less easily detected than the other neonatal diagnoses, although perhaps somewhat more so than might have been predicted. Relevant signs that caregivers might recognise, including difficulty in breathing, convulsions and other signs of neurological dysfunction, also occur with pneumonia, tetanus, septicaemia and other neonatal conditions. However, the age distributions of infants suffering from these conditions are distinct (in our study asphyxia = 2.1 days vs. pneumonia = 13.2 days, $P < 0.001$; asphyxia = 2.1 days vs. neonatal tetanus = 8.7 days, $P < 0.001$), suggesting, as we found, that caregivers might perceive other differences as well.

Marsh *et al.*⁶ found the sensitivity (54%) of caregiver-reported signs similar to those that we assessed for birth asphyxia to be lower than in our study (87%), with comparable specificity (80% vs. 72%). This might have been partially due to some misclassification of reference standard diagnoses in Pakistan as, although Marsh *et al.*⁶ provided their physician coders with clinical case definition guidelines, these may not have been applied consistently. This points to the importance of greater standardisation of procedures both within and between verbal autopsy validation studies, if their results are to be more comparable across settings.

We refined our methods over those of prior validation studies, including those of neonatal illnesses,^{2,4,6} to decrease the subjectivity and misclassification of reference standard diagnoses and to increase the generalisability of our findings and their applicability to population-based studies. Snow *et al.*² determined reference standard and verbal autopsy diagnoses through consensus review of medical and interview data by physician teams that used unspecified diagnostic criteria. The use by Marsh *et al.*⁶ of an interview format with screening questions followed by disease-specific modules appears to have limited their sensitivity estimates, because some screening questions proved to be insensitive. Kalter *et al.*⁴ relied on retrospectively collected reference standard data, and could not determine the specificity of the verbal autopsy diagnosis of neonatal tetanus because they did not include a comparison group of newborns without tetanus. In addition, most prior studies have used different interview formats, decreasing the ability to compare findings across sites.

In the present study, the reference diagnoses were based on stringent, predefined criteria, applied through use of a computer program to prospectively collected medical data, and interview data were collected via a standardised questionnaire that is being used in other disease and cultural settings. We studied a wide variety of neonatal illnesses, and included survivors as well as deaths, in

order to decrease the possible biases resulting from applying the findings of a hospital-based study to community settings with different diseases and less advanced medical care. Although our study sample included a good disease mix, the specificity estimates may still be biased in settings where the sample does not represent illnesses present in the community whose clinical syndromes are easily confused with those of the study diagnoses.

In summary, LBW/severe malnutrition, premature birth and tetanus can be detected in newborn infants by caregiver interview with high sensitivity and specificity, whereas the diagnoses of pneumonia and birth asphyxia are more difficult but still feasible. In settings with limited access to medical care, retrospective caregiver interviews provide a valid means of diagnosing several of the most common causes of severe neonatal illness and death.

Acknowledgements

The authors acknowledge the contributions of Dr David Ross and Mr Christopher Coldham of the London School of Hygiene and Tropical Medicine; Dr Robert Snow of the Kenya Medical Research Center in Kilifi; Ms Martha Anker of the World Health Organisation; Dr Mohammad Hanif of the Dhaka Shishu Hospital; the study paediatricians, Dr Pervez and Dr Shameem at Dhaka Shishu Hospital, and Dr Mohiuddin at the Infectious Disease Hospital; and Johns Hopkins graduate student, Mr Randy Kolstad. The study materials were developed in cooperation with the World Health Organisation, and financial support was provided by the United States Agency for International Development through the Johns Hopkins Family Health and Child Survival Cooperative Agreement, and by the United Kingdom Department for International Development through the World Health Organisation.

References

- 1 World Health Organisation. *The World Health Report 1995. Bridging the Gaps*. Geneva: WHO, 1995.
- 2 Snow RW, Armstrong JR, Forster D, Winstanley MT, Marsh VM, Newton CRJC, *et al*. Childhood deaths in Africa: uses and limitations of verbal autopsies. *Lancet* 1992; **340**:351–355.
- 3 Gray RH. *Verbal autopsy. Using interviews to determine causes of death in children*. Johns Hopkins University Institute for International Programs, Occasional Paper Series no. 14, March 1991.
- 4 Kalter HD, Gray RH, Gultiano C, Black RE. Validation of postmortem interviews to ascertain selected causes of death in children. *International Journal of Epidemiology* 1990; **19**:380–386.
- 5 Mobley CC, Boerma JT, Titus S, Lohrke B, Shangula K, Black RE. Validation study of a verbal autopsy method for causes of childhood mortality in Namibia. *Journal of Tropical Pediatrics* 1996; **42**:365–369.
- 6 Marsh DR, Sadruddin S, Fikree FF, Mehnaz A, Kisat H, Billoo AG, *et al*. *Validation of Diagnostic Criteria for Cause of Child Death by Delayed Interview of Bereaved Parent: a Refined Verbal Autopsy Method*. Karachi, Pakistan: Department of Community Health Sciences, The Aga Khan University, 1995.

- 7 Kalter HD, Gray RH, Gultiano C, Black RE. Validation of the diagnosis of childhood morbidity using maternal health interviews. *International Journal of Epidemiology* 1990; **20**:1–6.
- 8 Alonso PL, Bowman A, Marsh K, Greenwood BM. The accuracy of the clinical histories given by mothers of seriously ill African children. *Annals of Tropical Pediatrics* 1987; **7**:187–189.
- 9 Measurement of overall and cause-specific mortality in infants and children: memorandum from a WHO/UNICEF meeting. *Bulletin of the World Health Organisation* 1994; **72**:707–713.
- 10 Ross DA. *Monitoring Cause-specific Infant and Child Mortality Rates in Areas where Death Certification Systems are Weak*. WHO/UNICEF consultation on the measurement of overall and cause-specific mortality in infants and children. Geneva: WHO, December 1992.
- 11 Gray RH, Smith G, Barss P. *The Use of Verbal Autopsy Methods to Determine Selected Causes of Death in Children*. Johns Hopkins University Institute for International Programs, Occasional Paper Series no. 10, February 1990.
- 12 Bonadio WA, Hennes H, Smith D, Ruffing R, Melzer-Lange M, Lye P, *et al.* Reliability of observation variables in distinguishing infectious outcome of febrile young infants. *Pediatric Infectious Disease Journal* 1993; **12**:111–114.
- 13 United States Public Health Service, Health Resources Administration. *NCHS Growth Charts, Rockville, MD, 1976*. (HRA 76-1120, 25, 3).
- 14 Dean AG, Dean JA, Coulombier D, Burton AH, Brendel KA, Smith DC, *et al.* *Epi Info Version 6: a Word Processing, Database, and Statistics Program for Epidemiology on Microcomputers*. Atlanta, GA: Centers for Disease Control and Prevention, 1994.

Book review

Modern Epidemiology, 2nd Edition. Edited by K.J. Rothman and S. Greenland. Philadelphia: Lippincott-Raven, 1998, pp. 750, £49.

Unlike the first single-authored volume, this has a new section with articles that will be of particular interest to our readers. These include: reproductive epidemiology (Weinberg and Wilcox), genetic epidemiology (Khoury), nutritional epidemiology (Willett), ecological studies (Morgenstern), infectious diseases (Halloran), environmental epidemiology (Hertz-Picciotto) and the analysis of vital statistics data (Moolgavkar, Lee and Stevens). The major part of the work is rightly concentrated on study design and data analysis with excellent expositions by the two main authors.

For every epidemiologist who was a devotee of the first edition, and to those who were not, this book is a must. Given its size and detail it is very reasonably priced and therefore a tome that should be in every epidemiologist's private library.

JEAN GOLDING