Published in final edited form as:
Trop Med Int Health. 2011 January ; 16(1): 18–29. doi:10.1111/j.1365-3156.2010.02679.x.

An alternative strategy for perinatal verbal autopsy coding: single versus multiple coders

C. Engmann[1], **I. Jehan**[2], **J. Ditekemena**[3], **A. Garces**[4], **M. Phiri**[5], **M. Mazariegos**[6], **E. Chomba**[5], **O. Pasha**[2], **A. Tshefu**[3], **E. M. McClure**[7], **V. Thorsten**[7], **H Chakraborty**[7], **R. L. Goldenberg**[8], **C. Bose**[1], **W. A. Carlo**[9], and **L. L. Wright**[10]

[1]University of North Carolina at Chapel Hill NC, USA

[2]The Aga Khan University, Karachi, Pakistan

[3]Kinshasa School of Public Health, Kinshasa, Democratic Republic of Congo

[4]IMSALUD/San Carlos University, Guatemala City, Guatemala

[5]University Teaching Hospital, Lusaka, Zambia

[6]Institute of Nutrition for Central America and Panama, Guatemala City, Guatemala

[7]Research Triangle Institute, Durham NC, USA

[8]Drexel University, Philadelphia PA, USA

[9]University of Alabama at Birmingham AL, USA

[10]Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda MD, USA

Summary

Objective—To determine the comparability between cause of death by a single physician coder and a two-physician panel, using verbal autopsy.

Methods—The study was conducted between May 2007 and June 2008. Within a week of a perinatal death in 38 rural remote communities in Guatemala, the Democratic Republic of Congo, Zambia and Pakistan, VA questionnaires were completed. Two independent physicians, unaware of the others decisions, assigned an underlying cause of death, in accordance with the causes listed in the chapter headings of the *International classification diseases and related health problems, 10th revision* (ICD-10). Cohen's kappa statistic was used to assess level of agreement between physician coders.

Results—There were 9461 births during the study period; 252 deaths met study enrollment criteria and underwent verbal autopsy. Physicians assigned the same COD for 75% of stillbirths (K=0.69; 95% confidence interval: 0.61–0.78) and 82% early neonatal deaths (K=0.75; 95% confidence interval: 0.65–0.84). The patterns and proportion of stillbirths and early neonatal deaths determined by the physician coders were very similar compared to causes individually assigned by each physician. Similarly, rank order of the top 5 causes of stillbirth and early neonatal death were identical for each physician.

Conclusion—This study raises important questions about the utility of a system of multiple coders that is currently widely accepted, and speculates that a single physician coder may be an

Corresponding Author: Cyril Engmann, Division of Neonatal-Perinatal Medicine, Department of Pediatrics, University of North Carolina School of Medicine, CB # 7596, 4th Floor, UNC Hospitals, UNC-Chapel Hill, Chapel Hill, NC 27599-7596. cengmann@med.unc.edu.

effective and economical alternative to VA programs that use traditional two-physician panels to assign COD.

Keywords

verbal autopsy; perinatal death; comparing coders

Introduction

Understanding population-based causes of perinatal death (stillbirths and newborn deaths in the first 7 days of life) is critical to the development of an effective perinatal health policy (Lopez & Mathers 2006). Because there will always be competing demands for healthcare resources, a well-established system for identifying all perinatal deaths and assigning a medically determined cause of death (COD) for each death is highly desirable (Engmann et al. 2009). In many high-income countries, there is complete recording of deaths and for over 90% of these, medical certification is provided (Mathers et al. 2005). By contrast, fewer than 3% of all perinatal deaths in low and middle-income countries (LMIC) have medical certification of COD (Lawn Shibuya & Stein 2005). Many of these countries have the highest burden of poverty and disease and continue to lack routine, representative and high quality information on the levels and causes of death (Setel et al 2007). Part of the explanation for this may be that over half of all births and perinatal deaths occur in the home, and are frequently unrecorded in vital registration systems (Lawn et al. 2008).

Increasing numbers of LMIC are using verbal autopsy (VA) as a cost-effective and sustainable alternative to a thorough medical diagnostic evaluation as a source of data to inform mortality surveillance systems (Hill et al 2007).To determine the cause of fetal or infant mortality, the VA method relies on information obtained from a standardized interview with the primary caregiver (usually the mother) of the deceased. During this process, the symptoms, signs and behaviors during the illness of the deceased, or of the mother in the case of fetal death, are recorded. Trained coders review these data and apply diagnostic algorithms to determine COD. Typically, 2 or 3 trained physician coders review the data and independently assign a COD (Soleman et al. 2006). Any discrepancies between the COD assigned by each physician member of the panel are resolved by discussion and review of the VA data, and a final consensus COD is agreed upon by the physician panel (Setel et al 2005). The use of multiple physician coders in VA has been used to prevent random and systematic errors. Some researchers have urged that physicians should be encouraged to assign more than just a single cause of death, and that discussion of discrepant cases among a panel of physicians to reach consensus be considered less appropriate than allowing all physician diagnoses to contribute to the cause of death profile, whether or not individual physician diagnoses agree (Joshi et al 2009). Other authors have suggested methods for simultaneous analysis of cause of death (King and Yu 2008). Alternatively, COD can be assigned by the use of predetermined criteria/algorithms or computer simulations, a method that does not require the presence of a physician (Soleman 2006).

A recent report from a general population in India suggests that one trained physician determining COD facilitated by a series of algorithms developed for the Sample Registration System may be as effective as a physician panel in coding COD (Joshi et al. 2009). We sought to determine the potential effectiveness of using a single physician coder to assign cause of perinatal deaths by comparing COD assigned by two members of a physician panel. Each panel was based in rural districts in one of four low income countries.

Methods

Setting, Subjects and Study Design

This prospective observational study was nested within an ongoing, cluster randomized, controlled trial, the FIRST BREATH Trial, conducted by the Global Network McClure et al). The FIRST BREATH trial investigated the effects of implementing a package of newborn care practices, using the WHO Essential Newborn Care (ENC) program, and a neonatal resuscitation training program, a simplified version of the American Academy of Pediatrics Neonatal Resuscitation Program, in community settings. As part of this study, birth attendants were trained to collect basic maternal, fetal and neonatal outcomes data, which included demographics, mode of delivery, birthweight, gestational age, receipt of resuscitation, and adverse events. All birth attendants were trained to check for fetal and neonatal vital signs on every baby by auscultating the abdomen of every pregnant woman before delivery, and after delivery by feeling the umbilical cord of the neonate for a pulse, auscultating lungs for breath sounds, and assessing for any movement (Engmann et al. 2009). Birth weights were measured within 48 hours of delivery using UNICEF spring Salter Scales (UNICEF model 145555) provided for the study.

This study included sites in Guatemala (Chimaltenango province), the Democratic Republic of Congo (Equateur province), Zambia (Kafue district) and Pakistan (Thatta district). Within these sites, 38 communities participated in this study. Each community comprised a cluster of villages with approximately 300 deliveries per year. Data describing births was collected by birth attendants and reviewed by trained nurses or health workers assigned to each community and designated as Community Coordinators.

Within one week of an early neonatal death (END) or stillbirth (SB), birth attendants notified Community Coordinators who then visited the family, determined eligibility for the study and requested consent from eligible mothers. Perinatal deaths were excluded if they occurred in a hospital, if a birth attendant was absent at delivery, if the mother was unavailable for any reason (including peripartum death), or attempts to enroll the mother did not occur within 7 days of death. A 7-day window within a perinatal death was chosen to reduce the variability in the quality of reporting introduced by recall bias (Soleman 2006), Lee 2008, Fottrell 2010). Since the conventional perinatal verbal autopsy respondents are mothers, we elected to enroll only those subjects whose mothers were available for interview. Informed consent was obtained from mothers in a private and confidential setting. The consent form was read to all mothers who then provided their signatures or, if they were illiterate, thumbprints.

Training and VA methodology

All Community Coordinators and physicians participating in this study received standardized training in VA methodology (Engmann et al. year??). Community Coordinators were trained to interview mothers using the VA questionnaire. To assign COD, physicians were trained in ICD-10 classification, rules and guidelines (WHO 2005).

Uniform data describing the circumstances surrounding a perinatal death were collected from each mother using a standardized VA questionnaire developed specifically for this study from a validated VA tool (Engmann 2009), Mswia 2006). The questionnaire was administered by the Community Coordinators who then sent these data separately to two local physicians who independently assigned a COD. All physicians were provided with demographic and other descriptive data collected as part of the FIRST BREATH Trial. Each physician assigned one underlying COD, a final cause of death and contributing causes of death. Underlying COD was defined as the single most important disease or condition which initiated the train of morbid events leading directly to fetal or neonatal death. Underlying

COD was assigned in a non-hierarchical manner by physicians familiar with prevailing local diseases and health conditions/patterns (Thatte et al 2009). After the COD was assigned and entered independently, any discrepancy in assignment of COD between physicians was discussed and a consensus underlying COD assigned. In all cases, the two physicians were able to reach consensus after discussion.

Data collection and analysis

Data were collected between May 2007 and June 2008. Data were entered and transmitted electronically to the data coordinating center (RTI: Research Triangle Institute International, Research Triangle Park, NC, USA) where data edits, including inter- and intra-form consistency checks were performed. The study was reviewed and approved by the institutional ethics review committees of the Research Triangle International, the University of North Carolina at Chapel Hill and local institutional review boards.

The level of agreement between physician coders for underlying COD was calculated using Cohen's kappa statistic (K). Levels of agreement based on ranges of kappa values were defined as follows: 0.81–0.99 almost perfect agreement, 0.61–0.80, substantial agreement, 0.41–0.6 moderate agreement and less than 0.4 slight to fair agreement. (Viera & Garrett 2005). Data were analyzed using SAS (SAS/STAT® Software version 9.0. Descriptive statistics were generated for participant demographics and circumstances surrounding the deaths. Relationships between categorical variables were evaluated by examining cross-tabulations. Relationships between continuous variables were evaluated by examining means, standard deviations, medians and ranges.

Results

There were 9461 infants born in the designated communities during the study period (Figure 1). Among these, there were 518 SB and END. The SB, END, and perinatal mortality rates were 30/1000 births, 25/1000 live births, and 55/1000 births, respectively. Of the 518 deaths, 81 were ineligible for the study because the delivery occurred in a hospital (79) or the birth attendant was absent at the time of delivery (2). Among eligible deaths, 185 were not enrolled because the mother was not available for interview within seven days after the death (145) or did not provide consent (40). This study includes data describing deliveries of 241 women which resulted in 252 perinatal deaths (134 SBs and 118 ENDs).

The 5 major causes of early neonatal death were attributable to infections (45%), birth asphyxia (26%), prematurity (17%), tetanus (4%) congenital malformations (3%) and other/unknown causes (5 %). Major causes of stillbirth were attributable to infections (37%), obstructed/prolonged labor (11%), antepartum hemorrhage (10%), prematurity (7%) and cord complications such as prolapse (6%). For 12% of stillbirths, a COD could not be determined.

Agreement among coders

Physician coders assigned the same COD for 82% of END and 75% of SB. The kappa statistic for overall inter-coder agreement was 0.75 (0.65, 0.84) for END and 0.69 (0.61, 0.78) for SB.

Early Neonatal Death

Table 1 compares physician coder 1 and physician coder 2 responses for END. Overall, physicians agreed on the same COD for 97 out of 118 (82%) END. Table 2 is a comparison of physician coder 1 versus physician coder 2 responses for specific causes of END. As an example, physicians agreed 109 times out of 118 (92%) on prematurity as a COD. They

agreed that prematurity was a COD 13 times, and that prematurity was not the COD 96 times. The kappa value (level of agreement) between physicians was 0.7 (95% CI 0.51–0.88).

Stillbirth

Table 3 compares physician 1 and physician coder 2 responses for SB. Overall, physicians agreed on the same underlying cause of death at the chapter-heading level of ICD-10 for 101 out of 134 (75%) of stillbirths.

Table 4 compares physician coder 1 versus physician coder 2 responses for specific causes of stillbirth. Using maternal infection as an example, physicians agreed 120 times out of 134 (90%). They agreed that maternal infection was the cause of stillbirth 42 times and that maternal infection was not the cause of stillbirth 78 times. Physicians disagreed 14 times on the designation of infection as a cause of stillbirth. The kappa level of agreement was 0.78 (0.67, 0.89)

Individual assignment of COD and consensus COD

The proportion of stillbirths and early neonatal deaths determined by the two physicians were very similar, as were the patterns derived from the consensus process, compared to causes individually assigned by each physician. Similarly, the rank order of the top 5 leading causes of stillbirth and early neonatal death were identical for each physician (Figures 1 and 2).

Discussion

After preparation using a standardized VA training program, two physicians were in substantial agreement when assigning the major causes of early neonatal death. There was almost perfect agreement when tetanus and congenital malformations were the causes of early neonatal death. Any discrepancies noted in the COD assigned to individual cases had little impact on their rank order or the overall pattern of reported mortality. Substantial agreement between physicians was observed in the assignment of the major causes of stillbirth (antepartum hemorrhage, maternal infection, prolonged labor). There was only moderate agreement on the assignment of cord prolapse and prematurity as a cause of stillbirth, while there was almost perfect agreement when maternal accident was assigned as the COD.

Few studies have evaluated the impact of different methods for assigning cause of neonatal death using VA. In a recent paper, Joshi et al compared the assignment of the COD in 45 villages in Southern India by single versus multiple coders (Joshi et al. year???). This was a study of mortality in a general population of all ages, and fewer than 1% of the deaths occurred in children 0–28 days of age. They reported that physician coders agreed on the same diagnosis 94% of the time, with overall kappa values of 0.93 suggesting almost perfect agreement among physician coders. Among deaths in children aged 0–28 days, they reported kappa values of 1.0, although there were only 11 cases. Our study examined the comparability of the assignment of causes by two physicians for perinatal deaths only. For the three most important causes of early neonatal death (infections, birth asphyxia and prematurity), physicians agreed on the same cause of death approximately 90% of the time, suggesting substantial agreement. For two other causes of early neonatal death, congenital malformations and tetanus, physician agreement was nearly 100%. Similar results were reported byEdmond et al. (2008) on levels of agreement among three physicians determining cause of 590 neonatal deaths from verbal autopsies in rural Ghana. There was substantial agreement among three physicians for prematurity, birth asphyxia, infections (kappa values

0.8, 0.77, 0.72). In contrast to our study, they reported a kappa value of 0.63 for congenital abnormalities as a COD.

In our study, physicians showed substantial agreement for certain causes of stillbirths (antepartum hemorrhage, maternal infection, prolonged labor), and almost perfect agreement for maternal accidents. There was only moderate agreement for prematurity and cord prolapse. When the diagnostic accuracy of VA as determined by three experienced community pediatricians to determine cause of stillbirths from rural Ghana was compared to a hospital reference standard, VA performed poorly for causes of stillbirth diagnosis such as congenital abnormalities and maternal hemorrhage, while accuracy was higher for intrapartum obstetric complications and antepartum maternal disease (Edmond et al. 2008).

Even in settings where placental examinations, autopsies, cultures, karyotypes, x-rays, MRIs and other imaging are available, up to 60% of stillbirths are unexplained, highlighting the inherent difficulties that understanding and obtaining agreement over cause of stillbirth can pose (McClure et al. 2006; Silver et al 2007). In our study, the low rate of “unknown” COD may be an artifact of the study during which coders may have perceived some pressure to assign a cause of stillbirth. Nonetheless, the rate of concurrence between coders also suggests that VA may be a useful tool in determining population-based causes of stillbirth

There are economic and resource implications of the results of this study. The cost of programs using VA to assign COD could be substantially reduced by switching to a system of single coding. Joshi et al estimate that with deaths coded only once, the cause-of-death assignment costs can be halved and project management costs reduced by one third. They also suggest that funds currently used for duplicate coding could be reassigned to conduct validation studies that compare cause of death assignments from single coders against cause of death derived from reliable medical records, diagnosis by autopsy, or physician-diagnosed deaths in the community. Since VA is most typically used within weak health systems which suffer a shortage of physicians, utilizing fewer physicians and provide standardized training to them to code VA and redeploying them to other clinical tasks could be a more appropriate use of scarce human resources.

A major strength of this study is the standardized VA training and tools program which we have reported on previously. After initial training, a train-the-trainer model was used to spread it in the different countries within the GN. This strategy increases knowledge, promotes ownership, builds capacity and enables sustainability of programs (Enweronu-Laryea et al. 2009). In contrast to other studies which delay interview, we performed them within one week and found mothers eager to discuss their baby’s death. Early interviews may also yield more accurate diagnoses. There are also limitations to our study. It is possible that the duplicate coding process may be a poor method for detecting systematic errors in the assignment of causes of death. Also, poor training of coders could also result in a bias towards a particular cause of death assignment, which could be repeated by subsequent coders. Each coder was tested after training in the VA program, making these potential biases less likely. Another potential limitation may be bias towards certain diagnoses resulting from prior knowledge of the coders of disease patterns in their community. Therefore, use of coders from the community in which the deaths occur would be expected to result in a high level of agreement, but with less certainty of the correct assignment of COD. Although the VA tool has been validated previously in hospital settings, it application in a community setting where deaths occur outside of hospitals and the formal health care system has not been validated. Therefore, we cannot be certain of the accurate assignment of COD. However, even if tests of validity discovered incorrect assignment of COD for particular causes, it is unlikely that such a problem would affect agreement between coders.

Although inter-observer and intra-observer variations have been recognized over the years, the impact of these variations has not been studied in detail (Garenne & Fauveau 2006; Fauveau 2006). The findings from our study suggest that a single physician coder may be as effective as two coders in determining cause of stillbirth and early neonatal death when trained in a standardized VA program. This study also raises important questions about the utility of a system of multiple coders that is currently widely accepted, and speculates that a single physician coder may be an effective and economical alternative to VA programs that use traditional two-physician panels to assign COD.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding was provided by grants from the National Institutes of Child Health and Human Development and the Bill and Melinda Gates Foundation.

References

Edmond KM, Quigley MA, Zandoh C, et al. Diagnostic accuracy of verbal autopsies in ascertaining the causes of stillbirths and neonatal deaths in rural Ghana. Paediatr Perinat Epidemiol. 2008; 22:417–429. [PubMed: 18782250]

Engmann C, Matendo R, Kinoshita R, et al. Stillbirth and early neonatal mortality in rural Central Africa. Int J Gynaecol Obstet. 2009; 105:112–117. [PubMed: 19201402]

Engmann C, Jehan I, Ditekemena J, et al. Using verbal autopsy to ascertain perinatal cause of death: are trained non-physicians adequate? Trop Med Int Health. 2009; 14:1496–1504. [PubMed: 19799757]

Enweronu-Laryea C, Engmann C, Osafo A, Bose C. Evaluating the effectiveness of a strategy for teaching neonatal resuscitation in West Africa. Resuscitation. 2009; 80:1308–1311. [PubMed: 19720439]

Fauveau V. Assessing probable causes of death without death registration or certificates: a new science? Bull World Health Organ. 2006; 84:246–247. [PubMed: 16583085]

Fottrell E, Byass P. Verbal autopsy: methods in transition. Epidemiol Rev. 2010; 32:38–55. [PubMed: 20203105]

Garenne M, Fauveau V. Potential and limits of verbal autopsies. Bull World Health Organ. 2006; 84:164. [PubMed: 16583068]

Hill K, Lopez AD, Shibuya K, Jha P. Interim measures for meeting needs for health sector data: births, deaths, and causes of death. Lancet. 2007

Joshi R, Lopez AD, MacMahon S, et al. Verbal autopsy coding: are multiple coders better than one? Bull World Health Organ. 2009; 87:51–57. [PubMed: 19197404]

Joshi R, Kengne A, Neal B. Methodological trends in studies based on verbal autopsies before and after published guidelines. Bull World Health Organ. 2009; 87:678–682. [PubMed: 19784447]

King G, Lu Y. Verbal autopsy methods with multiple causes of death. Stat Sci. 2008; 23(1):78–91.

Lawn J, Shibuya K, Stein C. No cry at birth: global estimates of intrapartum stillbirths and intrapartum-related neonatal deaths. Bull World Health Organ. 2005; 83:409–417. [PubMed: 15976891]

Lawn JE, Osrin D, Adler A, Cousens S. Four million neonatal deaths: counting and attribution of cause of death. Paediatr Perinat Epidemiol. 2008; 22:410–416. [PubMed: 18782248]

Lee AC, Mullany LC, Tielsch JM, et al. Verbal autopsy methods to ascertain birth asphyxia deaths in a community-based setting in southern Nepal. Pediatrics. 2008; 121(5) e1372-e1380sign.

Lopez AD, Mathers CD. Measuring the global burden of disease and epidemiological transitions: 2002–2030. Ann Trop Med Parasitol. 2006; 100:481–499. [PubMed: 16899150]

Mathers CD, Fat DM, Inoue M, Rao C, Lopez AD. Counting the dead and what they died from: an assessment of the global status of cause of death data. Bull World Health Organ. 2005; 83:171–177. [PubMed: 15798840]

McClure EM, Wright LL, Goldenberg RL, et al. The global network: a prospective study of stillbirths in developing countries. Am J Obstet Gynecol. 2007; 197:247 e1–245 e1. [PubMed: 17826406]

Mswia RDC, Setel P, Hemed Y, Whiting D, Williams D. Sample Vital Registration with Verbal Autopsy : Key informants manual; Verbal autopsy interviewers manual. 2006

McClure EM, Nalubamba-Phiri M, Goldenberg RL. Stillbirth in developing countries. Int J Gynaecol Obstet. 2006; 94:82–90. [PubMed: 16730726]

Setel PW, Macfarlane SB, Szreter S, et al. A scandal of invisibility: making everyone count by counting everyone. Lancet. 2007

Setel PW, Sankoh O, Rao C, et al. Sample registration of vital events with verbal autopsy: a renewed commitment to measuring and monitoring vital statistics. Bull World Health Organ. 2005; 83:611–617. [PubMed: 16184280]

Silver RM, Varner MW, Reddy U, et al. Work-up of stillbirth: a review of the evidence. Am J Obstet Gynecol. 2007; 196:433–444. [PubMed: 17466694]

Soleman N, Chandramohan D, Shibuya K. Verbal autopsy: current practices and challenges. Bull World Health Organ. 2006; 84:239–245. [PubMed: 16583084]

Thatte N, Kalter HD, Baqui AH, Williams EM, Darmstadt GL. Ascertaining causes of neonatal deaths using verbal autopsy: current methods and challenges. J Perinatol. 2009; 29:187–194. [PubMed: 19110535]

Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. Fam Med. 2005; 37:360–363. [PubMed: 15883903]

WHO. International Statistical Classification of Diseases. 10th Revision, Second Edition2005.

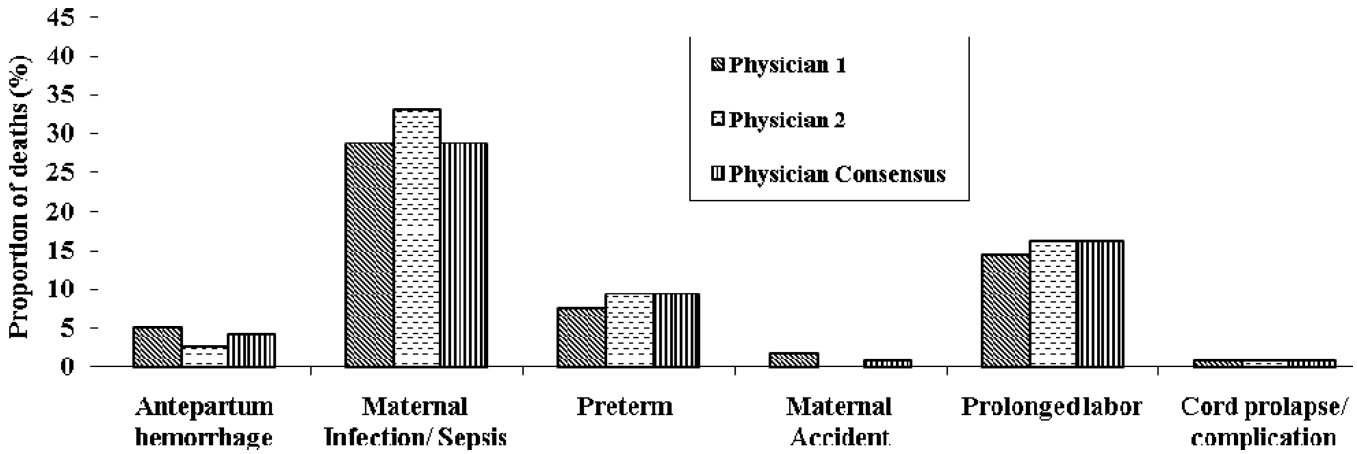

Figure 1.
Proportion of stillbirths assigned an underlying COD by each of two physician coders and by physician consensus

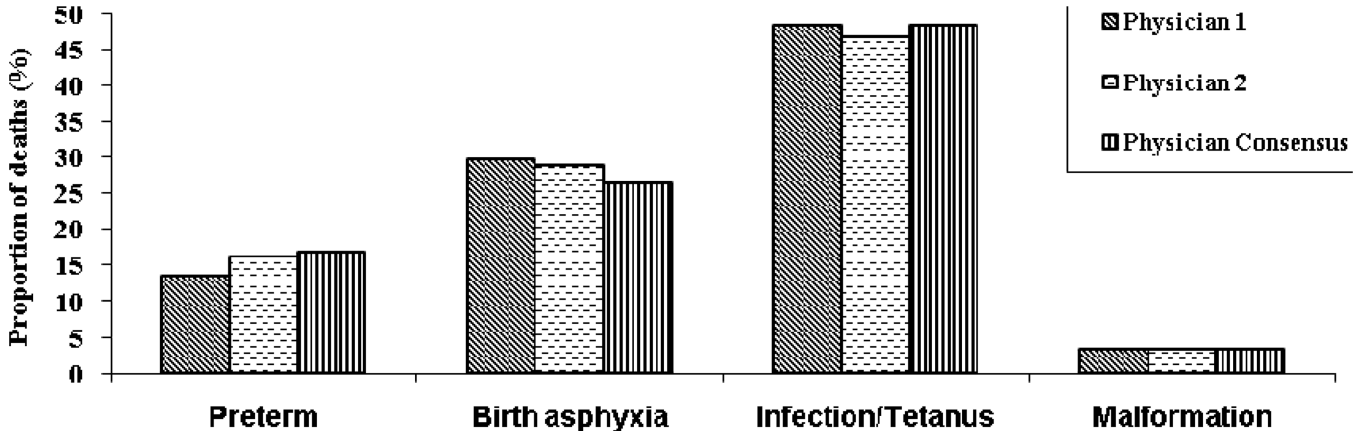

Figure 2.
Proportion of early neonatal deaths assigned an underlying COD by each of two physician coders and by physician consensus

Table 1

Comparison of Physician coder 1 and Physician coder 2 assigned cause of early neonatal death (n=118)

Physician 1 Responses	Physician 2 Responses							Physician 1 Total n (%)	Physician Consensus n (%)
	Preterm	Infection	Birth As-phyxia	C. Malform	Tetanus	Unknown / No Cause	Other		
Preterm	**13**	1	1	0	0	0	1	**16 (14)**	20 (17)
Neonatal Infection	2	**44**	4	0	1	1	0	**52 (44)**	52 (44)
Birth Asphyxia	3	3	**28**	0	0	1	0	**35 (30)**	31 (26)
C. Malformation	0	0	0	**4**	0	0	0	**4 (3)**	4 (3)
Tetanus	0	0	0	0	**5**	0	0	**5 (4)**	5 (4)
Unknown/no cause	0	0	0	0	0	**2**	0	**2 (2)**	4 (3)
Other	1	1	1	0	0	0	**1**	**4 (3)**	2 (2)
Physician 2 Total n (%)	**19 (16)**	**49 (42)**	**34 (28)**	**4 (3)**	**6 (5)**	**4 (3)**	**2 (2)**	**118**	

The bold numbers along the diagonal indicate agreement reached independently by the two physicians. Percent agreement = 97/118=82.2%. The percentages in parenthesis, provided along the Physician 1 total column, indicate how often Physician 1 reported the cause of death out of the total. Similarly, the percentages in parenthesis, provided along the bottom Physician 2 total row indicate how often Physician 2 reported the cause of death out of the total.

The percentages provided in parenthesis in the extreme right Physician Consensus column, refer to how often the Physician Consensus reported the cause of death out of the total. Thus taking infection as an example, Physician 1 and Physician 2 concluded neonatal infection was the cause of death 52 times. Physician 1 and Physician 2 initially agreed that infection was the underlying cause of death for 44 of the 52 cases. After discussing the 13 discrepant cases where one but not both attributed the underlying cause of death to infection, the physicians came to final consensus that 8 of the 13 cases had an underlying cause of death of neonatal infection. Other causes of early neonatal death were hypothermia, low birth weight and birth trauma.

Table 2

Comparison of Physician coder 1 and Physician coder 2 responses for specific causes of early neonatal deaths (n=118)

Underlying Cause of Death	Physician Response - n (%)				Kappa with corresponding 95% CI
	Physicians agreed			Physicians disagreed (one physician noted the condition positive and other noted the condition negative)	
	Condition noted as underlying cause	Condition not noted as underlying cause	Total		
Preterm	13	96	109 (92%)	9 (8%)	0.70 (0.51, 0.88)
Neonatal Infection	44	61	105 (89%)	13 (11%)	0.78 (0.66, 0.89)
Birth Asphyxia	28	77	105 (89%)	13 (11%)	0.73 (0.60, 0.87)
C. Malformation	4	114	118 (100%)	0	1.00 (1.00, 1.00)
Tetanus	5	112	117 (99%)	1 (1%)	0.90 (0.72, 1.00)
Unknown/no cause	2	114	116 (98%)	2 (2%)	0.66 (0.22, 1.00)

Table 3

Comparison of Physician coder 1 and Physician coder 2 responses for cause of stillbirth (n=134)

Physician 1 Responses	Physician 2 Responses								Physician 1 Total n (%)	Physician Consen-sus n (%)
	Antepar-tum He-morrhage	Maternal Infection	Preterm	Maternal Accident	Prolonged labor	Cord pro-lapse/ comp-lication	Unknown / No Cause	Other		
Antepartum Hemorrhage	**8**	3	0	0	0	0	0	1	**12 (8)**	13 (10)
Maternal Infection	0	**42**	0	0	0	0	0	1	**43 (32)**	50 (37)
Preterm	1	3	**5**	0	0	0	4	0	**13 (10)**	9 (7)
Maternal Accident	0	0	0	**6**	1	0	0	0	**7 (5)**	7 (5)
Prolonged labor	0	0	0	0	**11**	1	0	0	**12(9)**	15 (11)
Cord prolapse/ complication	0	0	0	0	0	**5**	3	0	**8 (6)**	8 (6)
Unknown/ no cause	1	2	1	0	0	0	**11**	0	**15 (11)**	16 (12)
Other	0	5	1	0	3	2	0	**13**	**24 (18)**	16 (12)
Physician 2 Total n(%)	**10 (7)**	**55 (41)**	**7 (5)**	**6 (4)**	**15 (11)**	**8 (6)**	**18 (13)**	**15 (11)**	**134**	

The bold numbers along the diagonal indicate agreement reached independently by the two physicians. Percent agreement = 101/134=75.4%. The percentages provided in parenthesis along the Physician 1 total column indicate how often Physician 1 reported the cause of death out of the total. Similarly, the percentages provided in parenthesis, along the total row indicate how often Physician 2 reported the cause of death out of the total.

The percentages provided in parenthesis in the extreme right Physician Consensus column, refers to how often the Physician Consensus reported the cause of death out of the total. Other causes of stillbirth were identified as malpresentation, folic acid deficiency, hypertension, post-term delivery, multiple birth, polihydramnios and multipara.

Table 4

Comparison of Physician coder 1 and Physician coder 2 responses for specific causes of stillbirth (n=134)

Underlying Cause of SB	Physician Response - n (%)				Kappa with corresponding 95% CI
	Physicians agreed			Physicians disagreed (one physician noted condition positive and other noted condition negative)	
	Condition noted as underlying cause	Condition not noted as underlying cause	Total		
Antepartum Hemorrhage	8	120	128 (96%)	6 (4%)	0.70 (0.48, 0.93)
Maternal Infection	42	78	120 (90%)	14 (10%)	0.78 (0.67, 0.89)
Preterm	5	119	124 (93%)	10 (7%)	0.46 (0.19, 0.74)
Maternal Accident	6	127	133 (99%)	1 (1%)	0.92 (0.76. 1.00)
Prolonged labor	11	118	129 (96%)	5 (4%)	0.79 (0.62, 0.97)
Cord prolapse/ complication	5	123	128 (96%)	6 (4%)	0.60 (0.31, 0.89)
Unknown/ no cause	11	112	123 (92%)	11 (8%)	0.62 (0.42, 0.83)