A comparative analysis of the InterVA model versus physician review in determining causes of neonatal deaths using verbal autopsy data from Nepal

Dinesh Dharel (dineshdharel13@gmail.com)
University of Saskatchewan
Penny Dawson
JSI Research & Training Institute Inc.
Daniel Adeyinka
University of Saskatchewan
Nazeem Muhajarine
University of Saskatchewan
Dinesh Neupane
Johns Hopkins University

Research Article

Keywords: verbal autopsy, neonatal mortality, low- and middle-income countries, inter-rater agreement

Posted Date: February 18th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-151220/v1

A comparative analysis of the InterVA model versus physician review in determining causes of neonatal deaths using verbal autopsy data from Nepal

Dinesh Dharel (ORCID: 0000-0002-9158-8510)[1], Penny Dawson[2], Daniel Adeyinka (ORCID: 0000-0003-1855-6878)[1,3], Nazeem Muhajarine (ORCID: 0000-0001-6781-5421)[1,3], Dinesh Neupane (0000-0002-1501-2990)[2,4,5]

Affiliation:

[1]Department of Community Health and Epidemiology, University of Saskatchewan, Saskatoon, SK, Canada

[2] JSI Research & Training Institute Inc. (JSI) Kathmandu, Nepal

[3]Saskatchewan Population Health and Evaluation Research Unit, Saskatoon, SK, Canada

[4] Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA

5 Nepal Development Society, Bharatpur, Chitwan, Nepal

Corresponding author: Dinesh Dharel, MD, Department of Community Health and Epidemiology, University of Saskatchewan, Saskatoon, SK, Canada. Email: dineshdharel13@gmail.com

Abstract

Background: Verbal autopsy is a common method of ascertaining the cause of neonatal death in low resource settings where majority of causes of deaths remain unregistered. We aimed to compare the causes of neonatal deaths assigned by computer algorithm-based model, InterVA (Interpreting Verbal Autopsy) with the usual standard of Physician Review of Verbal Autopsy (PRVA) using the verbal autopsy data collected by Morang Innovative Neonatal Intervention (MINI) study in Nepal.

Methods: MINI was a prospective community intervention study aimed at managing newborn illnesses at household level. Trained field staff conducted a verbal autopsy of all neonatal deaths during the study period. The cause of death was assigned by two pediatricians, and by using InterVA version 5. Cohen's kappa coefficient was calculated to compare the agreement between InterVA and PRVA assigned proximate cause of death, using STATA™ software version 16.1.

Results: Among 381 verbal autopsies for neonatal deaths, only 311 (81.6%) were assigned one of birth asphyxia, neonatal infection, congenital anomalies or preterm-related complications as the proximate cause of death by both InterVA and PRVA, while the remaining 70 (18.4%) were assigned other or non-specific causes. The overall agreement between InterVA and PRVA-assigned cause of death categories was moderate (66.5% agreement, kappa=0.47). Moderate agreement was observed for neonatal infection (kappa=0.48) and congenital malformations (kappa=0.49), while it was fair for birth asphyxia (kappa=0.39), and preterm-related complications (kappa=0.31); but there was only slight agreement for neonatal sepsis (kappa=0.19) and neonatal pneumonia (kappa=0.16) as specific causes of death within neonatal infections.

Conclusions: We observed moderate overall agreement for major categories of causes of neonatal death assigned by InterVA and PRVA. The moderate agreement was sustained for the

classification of neonatal infection but poor for neonatal sepsis and neonatal pneumonia as distinct categories of neonatal infection. Further studies should investigate the comparative effectiveness of an updated version of InterVA with the current standard of assigning the cause of neonatal death through longitudinal and experimental designs.

Keywords: verbal autopsy, neonatal mortality, low- and middle-income countries, inter-rater agreement

Background

Globally, an estimated four million neonatal deaths occurred annually in low and middle income countries (LMICs), of which the majority occur outside of the formal health care system (1). While the global annual neonatal mortality rate (NMR) decreased by 51%, from 36.6 to 18.0 deaths per 1000 live births between 1990 and 2017, it is estimated that 27.8 million neonatal deaths will occur by 2030, if the current rate of reduction continues in each country (2). Availability of nationally representative vital registration data on neonatal mortality is limited to about 60 countries and neonatal mortality rate is higher in LMICs that do not have a high-quality vital registration data (2). However, information on causes of death is vital to researchers, program planners and policymakers working at local, national and international levels to improve infant survival. In order to ensure the best possible utilization of limited resources available in such a setting, reliable and adequate information about the causes of neonatal deaths is essential. A practical and the most commonly used method to determine probable causes of death at population level in such settings where systems for medical certification of causes of death are weak or non-existent is verbal autopsy in which a series of questions are asked to the primary caregivers for specific signs and

symptoms of the deceased (3, 4). Verbal autopsy tools have been developed over the last few decades (4-10) and used for assigning causes of neonatal death in numerous settings and contexts and using different methods of assigning cause of death (11-24).

After conducting the verbal autopsy, the collected information is analyzed to assign causes of death. Causes of death assigned by Physician Review of Verbal Autopsy (PRVA) is most commonly used as a reference standard although multiple automated methods using data-driven algorithms have been developed and tested for assigning cause of death using information collected from a VA (25). However, PRVA is labor intensive, and prone to inter-observer variation. InterVA (Interpreting Verbal Autopsy), is one of the computer-based probabilistic model based on Bayes' probability theorem that has been compared with PRVA in some settings (24, 26-28). InterVA offers a promising alternative to expensive and time-consuming physician review in assigning the cause of death in low resource settings (4) (29). InterVA-5 was developed to harmonize with InterVA-4 and WHO 2016 VA standards, which is important for monitoring long-term trends over periods when different VA standards have been used (30). Although InterVA is an affordable and available option to assign causes of death using verbal autopsies, users need to be aware that there is no adequate evidence of equivalence, if not superiority, of its performance over PRVA. A validation study showed suboptimal performance of InterVA in assigning neonatal cause of death and a study from Nepal demonstrated discrepancy and overlap between physician review and algorithm-based assignment the causes of death (24, 31). In the milieu of ongoing search for the best possible method of assigning cause of neonatal death in low resource settings and improvements in InterVA algorithms, there is an ongoing need to compare InterVA with PRVA in different geographical and health system settings, particularly for neonatal mortality. This paper aims to compare the causes of neonatal deaths assigned by the most recent version of

InterVA with PRVA, using the VA data from Morang Innovative Neonatal Intervention (MINI) study.

Methods

Study setting

The current study has utilized the Verbal Autopsy dataset from the MINI program. MINI program focused on community-based management and treatment of illnesses of newborn children in Morang district of southeastern Nepal from 2005 to 2009. The district had 65 Village Development Committees (VDCs), two hospitals, seven primary health care centers, ten health posts, and 49 sub-health posts, and 585 Female Community Health Volunteers at the time of the study (32). In Morang district, 80% of the total population (914,799 at the time of the study) lived in rural areas and the human development and other indices were comparable to the national figures (32). In the MINI model, the trained community health workers (CHWs) followed an algorithm to classify sick young infants with possible severe bacterial infection (PSBI). The CHWs visited homes soon after delivery, recorded the birth, counseled mothers on essential newborn care, and assessed the newborns for danger-signs, thus identifying infants classified as having PSBI. These infants were treated with co-trimoxazole and referred to facility-based CHWs for seven-day treatment with injection gentamicin. A specific monitoring and supervision component were added to the implementation of the model through the existing government infrastructure. Further details of MINI methodology have been published previously (19, 32).

Verbal Autopsy and data collection

MINI field supervisors were trained to collect verbal autopsy data using a pre-tested structured questionnaire which included open-ended narrative and close-ended questions, adapted from the

existing tool from WHO (6).The questionnaire was developed in Nepali language and included demographic data, 39 questions and filters probing into the circumstances and causes of death, and finally the opinion of the caretaker and interviewer about the cause of death.

In 21 of 65 VDCs of Morang district where MINI was implemented, Female Community Health Volunteers (FCHV) notified deaths of neonates to MINI field supervisors. Then, MINI field supervisors visited the house and conducted the interview with the caretaker of the deceased neonate. The verbal autopsy was conducted after a median of 124 days (IQR 57-235 days). The completed questionnaires were checked by senior supervisor and field coordinator for completeness and consistency at the MINI field office.

Interpretation of verbal autopsy data through Physician review and InterVA

Two senior pediatricians in Nepal with similar clinical experience independently assigned the proximate and contributing cause of death using the hard copies of filled questionnaires. For any case in which two different causes of death were assigned by the two pediatricians, only one cause of death was chosen through discussion between the two pediatricians and entered in the MINI database.

We used the latest InterVA model, version 5, for comparison with the proximate cause of death assigned by the PRVA (30). The InterVA algorithm was developed and refined by Peter Byass and his colleagues over many years (33). InterVA accepts a range of symptoms or signs referred to as “indicators” relating to sequence and processes culminating in a death. InterVA processes them in a mathematical model based on Bayes’ theorem, and produces as its output likely cause(s) of death (26, 30). InterVA-5 uses symptom-cause information in the form of conditional probabilities of experiencing a symptom given a specific cause of death, elicited from the physicians as well as from Population Health Matrices Research Consortium gold standard dataset

(9). For the purpose of this study, the proximate (the first most likely) cause of death was assigned by using InterVA-5.

InterVA-5 supersedes InterVA-4, which was designed to be compatible with the WHO 2012 Verbal Autopsy Instrument and incorporated previous versions of the model for maternal and neonatal deaths, building on experience from InterVA-3 and preceding models. Thus, InterVA-5 brings the model in line with the WHO-2016 VA standard. With subsequent versions of WHO VA instruments there were corresponding changes in the indicators used in InterVA models in subsequent years.

Data management and Statistical analysis

Out of 498 verbal autopsies collected in the MINI study, 381 (76.5%) verbal autopsies for neonatal deaths were included in the analysis. For each VA, the cause of death was assigned by PR and InterVA. The reasons for exclusion were: 36 permanently transferred from the study sites, 2 refusals, 9 still births, 62 older than 28 days at the time of death, and 8 had incomplete data. The cause of death was assigned for each of 381 deceased neonates using InterVA-5 version 5.1 (released on 9th April 2020), using the executable software, code and user documentation available in the GitHub repository (34). The causes of neonatal death were assigned by PRVA using the WHO verbal autopsy guide and coded using ICD 10 codes. For this analysis, we combined neonatal pneumonia, neonatal sepsis and neonatal tetanus in PRVA to make a single category neonatal infection and combined preterm, low birth weight and respiratory distress syndrome as preterm-related complication. The other categories included hypothermia, hemorrhagic disease of newborn, sudden infant death syndrome and accidental injury. We used the proximate (first) cause of death output of InterVA-5, which consisted of one of the six WHO VA cause of death categories and other and unspecified perinatal cause of death category. InterVA assigned the second cause of

death for 34 of 381 verbal autopsies that had lower percentage of likelihood of the proximate diagnosis. For this analysis, we combined neonatal pneumonia, and meningitis/encephalitis in InterVA in a single category neonatal infection. InterVA-5 assigned prematurity as a cause of death in the category of preterm related complications.

Cohen's kappa coefficient was calculated to compare the agreement between InterVA and PRVA causes of death. For calculating kappa value for cause-specific mortality, for example birth asphyxia, data were assigned 1 (yes) for that cause and 0 (no) for all other causes. We compared the categories of cause of neonatal death common to PRVA and InterVA-5, including two subcategories of neonatal infection, namely neonatal sepsis and neonatal pneumonia. Analysis was done using STATA™ software version 16.1. We used Landis and Koch classification of Kappa scores, in which kappa values < 0 is taken as no agreement, 0-0.20 as slight agreement, 0.21 – 0.40 as fair agreement, 0.41-0.60 as moderate agreement, 0.61-0.80 as substantial agreement and 0.81-1 as perfect agreement (35).

Ethics

Ethical approval for the study was obtained from the Western Institutional Review Board (WIRB, Pro No. 20031870) in the USA. The Ministry of Health and Population, Nepal approved the intervention. The study was conducted by JSI Research & Training Institute, Inc. (JSI). Informed verbal consent was obtained from infant's legal guardian. All the procedures were followed in accordance with the Declaration of Helsinki.

Results

Background characteristics of neonates

182 Out of 381 neonatal deaths studied, the majority (78.5%) were early neonatal deaths, occurring
183 within 7 days of birth. The median survival time was 3 days from birth. Circumstances Of
184 Mortality CATegories (COMCAT) assigned by InterVA-5 showed that one third of neonatal
185 deaths involved circumstances other than emergencies, which would preclude life-saving actions
186 (29). The background characteristics are shown in Table 1.

187 **Table 1: Background characteristics of Neonates (N=381)**

Characteristics (N=381)	N (%)	95%CI (Lower limit, upper limit)
Timing of neonatal deaths		
Within 24hrs	133 (34.9)	30.3, 39.8
1-7 days	166 (43.6)	38.7, 48.6
>8 days	82 (21.5)	17.7, 25.9
Sex		
Male	219 (57.5)	52.4, 62.4
Female	162 (42.5)	37.63, 47.6
Birth weight (g)		
Birth weight not available	169 (44.4)	39.4,49.4
Very low birth weight (<1500 grams)	36 (9.5)	6.9, 12.8
Low birth weight (1500-2499 grams)	78 (20.5)	16.7, 24.8
Normal birth weight (>2500 grams)	98 (25.7)	21.6, 30.4
Gestation type		
Single	332 (87.1)	83.4, 90.2
Multiple	49 (12.9)	9.9, 16.6

Type of delivery		
Normal vaginal	327 (85.8)	81.9, 89.0
Instrumental or Caesarian	54 (14.2)	11.0, 18.1
Place of delivery		
Home	261 (68.5)	63.7, 73.0
Health facilities	120 (31.5)	27.0, 36.4
Circumstances of mortality categories		
Emergencies	250 (65.6)	60.7, 70.2
Health System	86 (22.6)	18.6, 27.1
Traditions	24 (6.3)	4.3, 9.2
Multiple	21 (5.5)	3.6, 8.3

188

189 The top three cause categories of neonatal deaths by physician review were neonatal infection
190 (40.7%), birth asphyxia (34.7%) and preterm-related complications (14.4%). The same three
191 categories were the three major causes of neonatal death by InterVA-5, with slightly different
192 percentages: neonatal infection (48.3%) birth asphyxia (23.8%), preterm-related complications
193 (14.5%), as shown in Table 2. Interestingly, we observed remarkably different proportions of
194 neonatal sepsis and neonatal pneumonia assigned by InterVA and PRVA. Both methods assigned
195 congenital anomalies in 2.6% of neonatal deaths (Table 2).

196 **Table 2 Comparison of proximate cause of neonatal death by InterVA and PRVA**

Proximate cause of death	**Physician review**		**InterVA-5**	
	n(%)	95%(CI)	n(%)	95%CI

Birth asphyxia	132 (34.7)	30.0, 39.6	90 (23.6)	19.6, 28.2
Neonatal infection	155 (40.7)	35.8, 45.7	184 (48.3)	43.3,53.3
• Sepsis	141 (37.0)	32.3, 42.0	91 (23.9)	19.9, 28.4
• Pneumonia	11 (2.9)	1.6, 5.2	89 (23.4)	19.4, 27.9
• Neonatal tetanus	3 (0.8)	0.3, 2.4	0	0
• Meningitis and encephalitis	0	0	4 (1.1)	0.4, 2.8
Congenital malformations	10 (2.6)	1.4, 4.8	10 (2.6)	1.4, 4.8
Preterm-related complications	57 (15.0)	11.7, 18.9	55 (14.4)	11.2, 18.4
• Prematurity	34 (8.9)	6.4, 12.2	55 (14.4)	11.2, 18.4
• Low birth weight	21 (5.5)	3.6, 8.3	0	0
• Respiratory distress syndrome	2 (0.5)	0.1, 2.1	0	0
Others causes	27 (7.1)	4.9,10.2	0	0
• SIDS	13 (3.4)	2.0, 5.8	0	0
• Hypothermia	9 (2.4)	1.23, 4.49	0	0

• Hemorrhagic disease of newborn	4 (1.1)	0.39, 2.77	0	0
• Accidental injury	1 (0.3)	0.037,1.85	0	0
Unspecified cause of death	0	0	42 (11.0)	8.2, 14.6
Total	381		381	

197

198 **Agreement between PRVA and InterVA**

199 The inter-rater reliability between InterVA and PRVA was assessed using Cohen's kappa statistic.
200 In our study, overall Kappa value was 0.47 corresponding to 66.5% agreement, which shows a
201 moderate agreement between InterVA and PRVA (Table 3). The agreement for neonatal infection,
202 as a category, was moderate (kappa=0.48) but it was only slight for neonatal sepsis (kappa=0.19)
203 and neonatal pneumonia (kappa=0.16), assigned by both methods. The agreement was fair for
204 prematurity as well as for category of preterm related complications. The agreement was fair for
205 birth asphyxia moderate for congenital malformations (Table 3).

206 **Table 3 Agreement between InterVA and PRVA for cause of neonatal death**

Proximate cause of death	Agreement (%), 95%CI	Expected agreement (%)	Kappa 95%CI	p-value	Degree of agreement (35)
Overall	66.5 (60.5,72.5)	36.5	0.47 (0.39, 0.55)	<0.001	Moderate

Neonatal infections*	74.0 (0.7,0.8)	50.3	0.48 (0.39,0.57)	<0.001	Moderate
Neonatal Sepsis#	64.8 (60.0,69.7)	56.8	0.19 (0.09,0.29)	0.0001	Slight
Neonatal Pneumonia#	79.0 (74.9,83.1)	75.1	0.16 (0.07,0.25)	<0.001	Slight
Birth asphyxia	74.3 (69.9,78.7)	58.1	0.39 (0.29,0.48)	<0.001	Fair
Preterm-related complications**	82.7 (78.9,86.5)	74.9	0.31 (0.18.0.44)	<0.001	Fair
Prematurity#	84.5 (80.9,88.2)	79.2	0.26 (0.12,0.39)	<0.001	Fair
Congenital malformations	97.4 (95.8,99.0)	94.9	0.49 (0.21,0.76)	<0.001	Moderate

* Neonatal infections included sepsis, pneumonia, neonatal tetanus and meningitis/encephalitis;

**preterm-related complications included prematurity, LBW and RDS

Sub-category compared as neonatal sepsis, neonatal pneumonia and prematurity were specific proximate diagnoses in both PRVA and InterVA

Discussion

The causes of neonatal deaths, observed in descending order of frequency in our study, both by InterVA and PRVA, included neonatal infection, birth asphyxia, preterm-related complications and congenital malformations which is in agreement with those described in literature (17, 24). Although reliable data on causes of neonatal deaths and the timing around neonatal deaths are often sparse, available evidence from low and middle income countries suggests that the major causes of death have not changed significantly despite a decrease in overall neonatal mortality rate and continue to occur most frequently in the first day and week of life (1, 2, 36). The focus of this paper is to compare assignment of cause of neonatal death using InterVA compared to PRVA from available VA data from Nepal and is not intended to determine the causes and timing of neonatal death.

We observed an overall moderate agreement (kappa 0.47, percentage agreement 66.5%) between InterVA and PRVA assigned categories of causes of neonatal death, with moderate agreement for neonatal infection and congenital anomalies and fair agreement for birth asphyxia and preterm-related complications. Interestingly, the agreement was weaker (only fair) for neonatal sepsis and neonatal pneumonia when compared as distinct causes of death, rather than neonatal infection. There is a considerable overlap in the clinical presentation of pneumonia and sepsis in a neonate and the management is the same for both conditions (36). In fact, pneumonia is a form of neonatal sepsis. In the MINI, PSBI was defined as presence of any one of the following: unable to feed, lethargic or unconscious, weak cry, fast breathing (≥ 60 breaths per minute), severe chest indrawing, grunting, fever or hypothermia, redness around umbilicus, >10 pustules or 1 large abscess and there were 1051 episodes (9%) of PSBI identified by the CHWs among 0-28 days infants, and they were not differentiated as neonatal sepsis or neonatal pneumonia. Identifying neonatal infection as a cause of neonatal death in the community setting serves the public health purpose of knowing the relative proportion of one of the main reasons for preventable neonatal death. However, distinguishing neonatal pneumonia from neonatal sepsis is neither accurate nor does it offer any benefit. Our finding of better agreement between InterVA-5 and PRVA for neonatal infection as a category, but not for neonatal pneumonia or sepsis suggests that InterVA-5 is a reasonable alternative approach for informing population level programs and policies targeting the reduction of preventable causes of neonatal deaths in resource limited settings, where specifying the exact etiology for neonatal deaths is not practically feasible and cause of death remain otherwise uncertain. However, the findings also speak to the need to continue upgrading this automated VA model to improve the accuracy of assigning more specific etiology of neonatal

mortality, which may be more meaningful in distinguishing death from neonatal tetanus, which demands more specific preventive practices.

Preterm-related complication, as a category of cause of neonatal death, only had fair agreement between InterVA and PRVA, along with prematurity as a common proximate cause of death by both methods. Low birth weight appeared as a proximate cause of neonatal death only in PRVA. Preterm-related complications could include several issues specific to preterm babies including respiratory distress syndrome. Our findings do not offer such comparisons, reflecting the limitation to accurately report the gestational age and/or birth weight by the caretaker during verbal autopsy interview, particularly when the birth is unattended and occurs at home. Still, capturing the broad cause of death as being related to prematurity and/or low birth weight would provide valuable information to the maternal and newborn health programs in reducing the neonatal deaths from preterm-related complications.

Assigning cause to neonatal death remains challenging due to non-specific signs and symptoms in sick newborns (3). Specifying a cause of death in general is itself challenging (37) and attributing a single cause could be an oversimplification (38) with the possibility of synergistic relationship among the causes such as infection, asphyxia and prematurity/intrauterine growth restriction. Therefore, the limitations of attributing a cause for neonatal deaths are well documented and recognized in the scientific community (2). The verbal autopsy dataset we used, which was collected using questionnaire based on an older iteration of WHO verbal autopsy standards, did not allow input for all relevant indicators in the InterVA-5 algorithm. The accuracy of cause of death assigned by InterVA would heavily depend on accuracy and completeness of the data available from verbal autopsy interview. Still, it is possible, as well as programmatically useful, to review broad ‘best guesses’ for cause of death in a reasonable sample of neonatal deaths. Our

findings, along with findings from other similar studies, demonstrate how assessing and assigning the causes of neonatal deaths is complex but can still be done with reasonable accuracy using InterVA as an acceptable substitute for PRVA.

The verbal autopsy is carried out by medically untrained enumerators from usually illiterate or just literate parents in rural settings. So, defining the cause of death by strict criteria using verbal autopsy data may lead to under- or over-estimation of neonatal deaths due to inability to get the right information, which is sensitive and specific enough to make a diagnosis. In particular, early neonatal death may have been overestimated due to misclassification of stillbirth as neonatal death by the respondent. Assigning the specific cause of neonatal death is difficult even in a hospital setting by trained health professionals owing to non-specific signs and symptoms of common neonatal conditions. However, there are limited possible diagnoses or disease categories leading to preventable neonatal death, which potentially allows making reasonable estimate of the proximate cause of death using algorithm-based methods as well.

Physicians had access to additional remarks from the respondent (caretaker) and the interviewer about their impression on the possible cause of death, but this information was not coded for analysis by InterVA. When there is more than one possible cause of death, assigning a proximate cause of death by InterVA depends on whether a hierarchical or nonhierarchical approach is used and how the cause of death is assigned. However, physicians may assign cause of death as per their discretion, influenced by their medical knowledge and contextual information. However, InterVA-5 offers the COMCAT functionality which categorizes the circumstances of death (30).

There may be recall bias on late interviewing after the neonatal deaths. However, verbal autopsy taken between 3 to 12 months resulted in comparable results with those taken within 3 months of death (39). Some interviews in our cohort were repeated due to lack of meaningful information on

290 the first interview, which could have introduced bias. InterVA and PRVA used the same set of VA
291 data, thus the effect of recall bias would apply to both methods.

292 Considering PRVA as the reference standard for assigning cause of death from verbal autopsy, the
293 InterVA needs to be updated and standardized further to make it closer to the actual cause of
294 neonatal death. This may demand a more complex computer-based algorithm and more rigorous
295 data collection method to obtain more sensitive information about the signs and symptoms that led
296 to death of the neonate. Whether certain signs and symptoms predict the diagnosis which is used
297 as a cause of death of the neonate should be validated by well-designed clinical studies. As the
298 accuracy of verbal autopsy based assignment of cause of neonatal death is sensitive to the setting
299 and the process in which verbal autopsy is conducted and the cause of death is assigned, further
300 reports on comparability of PRVA and InterVA from different settings is necessary to advance
301 this field. We utilized the verbal autopsy data from a community-based intervention research
302 model in a rural setting of an emerging district health system of a developing country to add to the
303 ongoing search for the optimum method of assigning cause of neonatal death.

304 **Conclusions**

305 This study revealed a moderate overall agreement between InterVA and PRVA for assigning
306 broad category of cause of neonatal death from verbal autopsy data from Nepal, with moderate
307 agreement for neonatal infection and congenital malformation and fair agreement for birth
308 asphyxia and preterm-related complications. Only slight agreement for neonatal pneumonia and
309 neonatal sepsis between InterVA and PRVA, compared to moderate agreement for neonatal
310 infection as a category suggests the need for revision in classifying causes of deaths based on
311 similar clinical and programmatic implications of preventable causes of neonatal death. Further

studies should look at the comparative effectiveness of updated version of InterVA with the current standard of assigning the cause of death at population-level through longitudinal and experimental designs.

List of Abbreviations

CHW	Community Health Workers
COMCAT	Circumstances Of Mortality CATegories
FCHV	Female Community Health Volunteers
ICD	International Classification of Disease
InterVA	Interpreting Verbal Autopsy
LMIC	Low and Middle Income Countries
MINI	Morang Innovative Neonatal Intervention
NMR	Neonatal Mortality Rate
PRVA	Physician Review of Verbal Autopsy
PSBI	Possible Severe Bacterial Infection
VA	Verbal Autopsy
VDC	Village Development Committee
WHO	World Health Organization

Declarations:

Ethics approval and consent to participate:

Ethical approval for the study was obtained from the Western Institutional Review Board (WIRB, Pro No. 20031870) in the USA. The Ministry of Health and Population, Nepal approved the intervention. The study was conducted by JSI Research & Training Institute, Inc. (JSI). Informed

verbal consent was obtained from infant's legal guardian. All the procedures were followed in accordance with the Declaration of Helsinki.

Consent for publication:

Not applicable

Availability of data and materials:

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests:

The authors (DD, PD, DA, NM, DN) declare that they have no competing interests.

Funding:

Funding for MINI was provided through Saving Newborn Lives/Save the Children-USA to JSI Research and Training Institute, Inc. (JSI). Additional financial, administrative and technical support was provided by the United States Agency for International Development, Nepal, through the Nepal Family Health Program/JSI Research & Training Institute, Inc. The authors have not entered into any agreement with the funding agencies that may have limited their ability to complete the research and have had full control over the primary data.

Authors contribution:

DD contributed to the conceptualization, design, analysis and writing the initial draft and revising the manuscript. PD contributed to the conceptualization, design and revision of subsequent drafts of the manuscript. DA contributed to the analysis and writing of the methods and results section of the manuscript. NM contributed to the revision of the subsequent drafts of the manuscript and provided guidance to analysis and interpretation of data. DN contributed to

the conceptualization, design, analysis and writing and revision of subsequent draft of the manuscript.

All authors read and approved the final manuscript and agree to be accountable for the contents in the manuscript.

Acknowledgement and dedication:

The authors acknowledge Dr. Steve Wall (SC/SNL), and Dr. Jaganath Sharma (USAID/Nepal) for their review of earlier versions of this paper. During implementation of the MINI program, many colleagues contributed to the success of the initiative, for which the authors wish to express their thanks: Dr. Prakash Sundar Shrestha, Dr. Ranendra Prakash Bahadur Shrestha, Dr. Stephen Hodgins, Dr. Robin Houston, Dr. Sudhir Khanal, Dr. Gargi KC and Dr. Vijay Singh GC. The authors also wish to thank the entire team in the District Public Health Office, Morang for their continuous support and vigilance over the program during the implementation stage and to the Female Community Health Volunteers who served their communities with great commitment.

The opinions expressed herein are those of the authors and do not necessarily reflect the views of any concerned agency.

The authors wish to dedicate this article to the late Professor Peter Byass, for his relentless efforts in developing the InterVA model and providing support for analyzing our data using InterVA-5.

References

1. Lawn JE, Cousens S, Zupan J, Team LNSS. 4 million neonatal deaths: when? Where? Why? The lancet. 2005;365(9462):891-900.
2. Hug L, Alexander M, You D, Alkema L, for Child UI-aG. National, regional, and global levels and trends in neonatal mortality between 1990 and 2017, with scenario-based projections to 2030: a systematic analysis. The Lancet Global Health. 2019;7(6):e710-e20.
3. Thatte N, Kalter HD, Baqui A, Williams E, Darmstadt G. Ascertaining causes of neonatal deaths using verbal autopsy: current methods and challenges. Journal of Perinatology. 2009;29(3):187-94.
4. Nichols EK, Byass P, Chandramohan D, Clark SJ, Flaxman AD, Jakob R, et al. The WHO 2016 verbal autopsy instrument: An international standard suitable for automated analysis by InterVA, InSilicoVA, and Tariff 2.0. PLoS medicine. 2018;15(1):e1002486.
5. Nykanen M, Tamaona W, Cullinan T, Van VO, Ashorn P. Verbal autopsy as a technique to establish causes of infant and child mortality. East African medical journal. 1995;72(11):731-4.
6. World Health Organization. A standard verbal autopsy method for investigating causes of death in infants and children. World Health Organization; 1999.
7. World Health Organization. WHO technical consultation on verbal autopsy tools. Geneva: WHO. 2005.
8. Soleman N, Chandramohan D, Shibuya K. Verbal autopsy: current practices and challenges. Bulletin of the World Health Organization. 2006;84:239-45.
9. Byass P. Usefulness of the Population Health Metrics Research Consortium gold standard verbal autopsy data for general verbal autopsy methods. BMC medicine. 2014;12(1):23.
10. Soofi SB, Ariff S, Khan U, Turab A, Khan GN, Habib A, et al. Diagnostic accuracy of WHO verbal autopsy tool for ascertaining causes of neonatal deaths in the urban setting of Pakistan: a hospital-based prospective study. BMC pediatrics. 2015;15(1):144.
11. Duby J, Pell LG, Ariff S, Khan A, Bhutta A, Farrar DS, et al. Effect of an integrated neonatal care kit on cause-specific neonatal mortality in rural Pakistan. Global Health Action. 2020;13(1):1802952.
12. Kalter HD, Perin J, Amouzou A, Kwamdera G, Adewemimo WA, Nguefack F, et al. Using health facility deaths to estimate population causes of neonatal and child mortality in four African countries. BMC Medicine. 2020;18(1):1-14.
13. Koffi AK, Perin J, Kalter HD, Monehin J, Adewemimo A, Black RE. How fast did newborns die in Nigeria from 2009-2013: a time-to-death analysis using Verbal/Social Autopsy data. Journal of global health. 2019;9(2).
14. Dandona R, Kumar GA, Bhattacharya D, Akbar M, Atmavilas Y, Nanda P, et al. Distinct mortality patterns at 0–2 days versus the remaining neonatal period: results from population-based assessment in the Indian state of Bihar. BMC medicine. 2019;17(1):140.
15. Ahmed I, Ali SM, Amenga-Etego S, Ariff S, Bahl R, Baqui AH, et al. Population-based rates, timing, and causes of maternal deaths, stillbirths, and neonatal deaths in south Asia and sub-Saharan Africa: a multi-country prospective cohort study. The Lancet Global Health. 2018;6(12):e1297-e308.
16. Halim A, Dewez JE, Biswas A, Rahman F, White S, van den Broek N. When, where, and why are babies dying? Neonatal death surveillance and review in Bangladesh. PloS one. 2016;11(8):e0159388.
17. Fottrell E, Osrin D, Alcock G, Azad K, Bapat U, Beard J, et al. Cause-specific neonatal mortality: analysis of 3772 neonatal deaths in Nepal, Bangladesh, Malawi and India. Archives of Disease in Childhood-Fetal and Neonatal Edition. 2015;100(5):F439-F47.
18. Vergnano S, Fottrell E, Osrin D, Kazembe PN, Mwansambo C, Manandhar DS, et al. Adaptation of a probabilistic method (InterVA) of verbal autopsy to improve the interpretation of cause of stillbirth and neonatal death in Malawi, Nepal, and Zimbabwe. Population health metrics. 2011;9(1):48.

19. Khanal S, Dawson P, Houston R. Verbal autopsy to ascertain causes of neonatal deaths in a community setting: a study from Morang, Nepal. Journal of the Nepal Medical Association. 2011;51(181).
20. Manandhar S, Ojha A, Manandhar D, Shrestha B, Shrestha D, Saville N, et al. Causes of stillbirths and neonatal deaths in Dhanusha district, Nepal: a verbal autopsy study. Kathmandu University medical journal (KUMJ). 2010;8(29):62.
21. Baqui AH, Darmstadt G, Williams E, Kumar V, Kiran T, Panwar D, et al. Rates, timing and causes of neonatal deaths in rural India: implications for neonatal health programmes. Bulletin of the World Health Organization. 2006;84:706-13.
22. Lee AC, Mullany LC, Tielsch JM, Katz J, Khatry SK, LeClerq SC, et al. Verbal autopsy methods to ascertain birth asphyxia deaths in a community-based setting in southern Nepal. Pediatrics. 2008;121(5):e1372-e80.
23. Setel PW, Rao C, Hemed Y, Whiting DR, Yang G, Chandramohan D, et al. Core verbal autopsy procedures with comparative validation results from two countries. PLoS Med. 2006;3(8):e268.
24. Freeman JV, Christian P, Khatry SK, Adhikari RK, LeClerq SC, Katz J, et al. Evaluation of neonatal verbal autopsy using physician review versus algorithm-based cause-of-death assignment in rural Nepal. Paediatric and Perinatal Epidemiology. 2005;19(4):323-31.
25. Murray CJ, Lozano R, Flaxman AD, Serina P, Phillips D, Stewart A, et al. Using verbal autopsy to measure causes of death: the comparative performance of existing methods. BMC medicine. 2014;12(1):5.
26. Byass P, Huong DL, Van Minh H. A probabilistic approach to interpreting verbal autopsies: methodology and preliminary validation in Vietnam. Scandinavian Journal of Public Health. 2003;31(62_suppl):32-7.
27. Fantahun M, Fottrell E, Berhane Y, Wall S, Högberg U, Byass P. Assessing a new approach to verbal autopsy interpretation in a rural Ethiopian community: the InterVA model. Bulletin of the World Health Organization. 2006;84:204-10.
28. Oti SO, Kyobutungi C. Verbal autopsy interpretation: a comparative analysis of the InterVA model versus physician review in determining causes of death in the Nairobi DSS. Population health metrics. 2010;8(1):21.
29. Hussain-Alkhateeb L, D'Ambruoso L, Tollman S, Kahn K, Van Der Merwe M, Twine R, et al. Enhancing the value of mortality data for health systems: adding circumstances of mortality categories (COMCATs) to deaths investigated by verbal autopsy. Global health action. 2019;12(1):1680068.
30. Byass P, Hussain-Alkhateeb L, D'Ambruoso L, Clark S, Davies J, Fottrell E, et al. An integrated approach to processing WHO-2016 verbal autopsy data: the InterVA-5 model. BMC medicine. 2019;17(1):102.
31. Lozano R, Freeman MK, James SL, Campbell B, Lopez AD, Flaxman AD, et al. Performance of InterVA for assigning causes of death to verbal autopsies: multisite validation study using clinical diagnostic gold standards. Population health metrics. 2011;9(1):50.
32. Khanal S, Jaganath Sharma VSG, Dawson P, Houston R, Khadka N, Yengden B. Community health workers can identify and manage possible infections in neonates and young infants: MINI—a model from Nepal. Journal of health, population, and nutrition. 2011;29(3):255.
33. Clark SJ. A Guide to Comparing the Performance of VA Algorithms. arXiv preprint arXiv:180207807. 2018.
34. Byass P. InterVA-5 Verbal Autopsy Model 2020 [updated 9th April 2020. 5.1:[Available from: http://www.interva.net.
35. Landis JR, Koch GG. The measurement of observer agreement for categorical data. biometrics. 1977:159-74.

36. Duke T. Neonatal pneumonia in developing countries. Archives of Disease in Childhood-Fetal and Neonatal Edition. 2005;90(3):F211-FF9.
37. Marsh D, Majid N, Rasmussen Z, Mateen K, Khan AA. Cause-specific child mortality in a mountainous community in Pakistan by verbal autopsy. Journal of Pakistan Medical Association. 1993;43(11):226.
38. Peebles DM, Wyatt JS. Synergy between antenatal exposure to infection and intrapartum events in causation of perinatal brain injury at term. BJOG: An International Journal of Obstetrics & Gynaecology. 2002;109(7):737-9.
39. Serina P, Riley I, Hernandez B, Flaxman AD, Praveen D, Tallo V, et al. What is the optimal recall period for verbal autopsies? Validation study based on repeat interviews in three populations. Population health metrics. 2016;14(1):40.