



Global Congress on Verbal Autopsy: State of the Science  
Bali, Indonesia  
Meliá Bali Villas & Spa Resort  
February 15-17, 2011

---

**ABSTRACTS FOR ORAL PRESENTATIONS**

These abstracts represent the oral presentations given at the Global Congress on Verbal Autopsy: State of the Science. For conference purposes only, they have been modified from their original form.

We express our gratitude for the authors for their submission and contributions to the Congress discussions.

## **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.**

S Vergnano, E Fottrell, D Osrin, PN Kazembe, CMwansambo, DS Manandhar, SP Munjanja, S.Lewycka; P Byass, A Costello.

### **Background**

Verbal autopsy is an important method for analysing causes of death in developing countries. We evaluated the InterVA method to identify causes of death in the perinatal (stillbirths and neonatal deaths within the first seven days) and neonatal periods, using data from Malawi, Zimbabwe and Nepal.

### **Methods**

734 stillbirth and neonatal VAs were obtained from recent community studies in rural areas: 169 from Malawi, 385 from Nepal and 180 from Zimbabwe. Initial refinement of the InterVA model was based on 100 physician-reviewed VAs from Malawi. InterVA indicators and matrix probabilities for cause of death were reviewed for clinical and epidemiological coherence by a paediatrician-researcher and an epidemiologist involved in the development of InterVA. The modified InterVA model was evaluated by comparing population-level cause-specific mortality fractions derived from two methods of interpretation (physician review and InterVA) for a further 69 VAs from Malawi, 385 from Nepal, and 180 from Zimbabwe

### **Findings**

Case-by-case agreement between InterVA and physician-reviewed diagnoses, for 69 cases from Malawi, 180 cases from Zimbabwe and 385 cases from Nepal, were 83% (Kappa 0.76 (0.75 - 0.796)), 71% (Kappa 0.41(0.32-0.51)) and 74% (Kappa 0.63 (0.60-0.63)), respectively. The proportion of stillbirths as fresh or macerated identified by the different methods of VA interpretation was similar in all three settings. For cross-country comparison the modified InterVA method showed the proportions of preterm births and deaths due to infection were higher in Zimbabwe than in Malawi or Nepal.

### **Conclusion**

Our modified InterVA method provides reliable information on stillbirths and causes of death of newborns and allows for cross-country comparisons, eliminating potential bias from physician review.

# Assessing the Performance of InterVA on Gold Standard Hospital Deaths for Verbal Autopsy: Using a Probabilistic Model to Assign Adult Causes of Death

Rafael Lozano<sup>1</sup>, Benjamin Campbell<sup>1</sup>, Spencer L James<sup>1</sup>, Alireza Vahdatpour<sup>1</sup>, Sean Green<sup>1</sup>, Abraham D Flaxman<sup>1</sup>, Alan Lopez<sup>2</sup>, Christopher JL Murray<sup>1</sup>

<sup>1</sup>Institute for Health Metrics and Evaluation, University of Washington, Seattle, Washington, USA

<sup>2</sup>University of Queensland School of Population Health, Brisbane, Australia

## Background

Verbal autopsy is an important technique for measuring cause-specific mortality patterns in regions without vital registration systems. In order for verbal autopsy to be useful to researchers, policy-makers, and society at large, it is critical that the results of these interviews can lead to accurate assignments of underlying causes of death. Various methods have been developed to effectively analyse verbal autopsy data. For this study, we assessed the performance of one such method, “InterVA,” (specifically, InterVA version 3.2) a tool used widely to analyse verbal autopsy data. The InterVA model is an expert algorithm based on Bayesian probability theory. The tool has been adapted and used in various settings, including many INDEPTH sites (International Network of field sites with continuous Demographic Evaluation of Populations and Their Health in developing countries), and it has been used to predict cause-specific mortality fractions for important causes of death. However, the method has never been assessed against a “gold standard” of known causes of death to validate diagnoses [1]

## Methods

The Grand Challenges 13 Population Health Metrics Research Consortium (PHMRC) study collected 7296 adult and 4569 child and neonatal verbal autopsy interviews across 6 sites over the course of 3 years. Each interview investigated a death for which a gold-standard underlying cause of death was known. This dataset provides a validation environment for the InterVA method for computer coded verbal autopsy. InterVA requires certain parameters, including “Yes” or “No” responses to 106 symptom indicators, for which we mapped using responses to the study’s VA instrument. InterVA requires HIV and malaria prevalence parameters, which we assigned based on regional estimates. Lastly, to properly match the PHMRC gold standard diagnoses to InterVA outputs, we mapped a 55 cause list from the study to InterVA’s 35 cause list. We generated results using InterVA on adult deaths recorded in the study.

To assess the accuracy of InterVA for computer coded verbal autopsy, we performed 25 replications of the following experiment: we sampled without replacement 25% of the interviews in the PHMRC VA dataset, stratified by underlying cause of death, and used InterVA to obtain the average concordance and the average relative error (ARE) of these predictions. We compared these to the results of physician coded verbal autopsy (PCVA) on the same set of deaths.

## Results

Preliminary results show that InterVA correctly predicts 25.9% of deaths in this dataset, with an average relative error (ARE) of the cause specific mortality fractions of 84.1%. Based on preliminary assessment, the InterVA program performs similarly to PCVA in terms of average concordance.

## Conclusions

This assessment shows that InterVA may be a reasonable substitute for PCVA. The physician review process requires the training and time of physicians, which is often a resource limitation and can be both costly and subjective. Thus, the InterVA method, which is easy to use and freely accessible, is found to be a competitive alternative to PCVA

## References

- [1] Fottrell E, Byass P. Verbal Autopsy: Methods in Transition. *Epidemiol Rev.* 2010; 32:38-55.

## Cause-specific mortality trends in rural KwaZulu-Natal South Africa, 2000-2009.

Kobus Herbst<sup>1</sup>, Tshepiso Mafojane<sup>1</sup>, Marie-Louise Newell<sup>1,2</sup>

<sup>1</sup>Africa Centre for Health and Population Studies, University of KwaZulu-Natal, Somkhele, South Africa

<sup>2</sup>MRC Centre of Epidemiology for Child Health, UCL ICH, London, UK

### Background

The advent of the AIDS pandemic and interventions against it has resulted in extensive and rapid changes in cause-specific mortality rates in sub-Saharan Africa. These changes have happened against the backdrop of a steadily increasing non-communicable burden of disease. In South-Africa these trends have been prominent, and are compounded by a high trauma-related mortality. The objective of this study is to describe cause-specific mortality trends based on verbal autopsies conducted on all deaths in a rural population in KwaZulu-Natal, South Africa over a 10 year period (2000-2009).

### Methods

The study uses population-based mortality data collected by a demographic surveillance system on all resident and non-resident members of 11,000 households. Deaths and person years of observation (pyo) were aggregated for individuals between 01 January 2000 and 31 December 2009. Cause of death was determined by verbal autopsy based on the standard INDEPTH/WHO verbal autopsy questionnaire. Cause of death was assigned by physician review and coded using ICD-10 and further categorised using global burden of disease categories. The verbal autopsy questionnaire data were converted into the input indicators for the InterVA probabilistic model. and the output coded into the same burden of disease categories. Where the InterVA system identified multiple causes of death each cause was weighted based on the likelihood values produced by the system. If an uncertain proportion remained for a case it was assigned to the undefined cause category.

### Results

In 125 432 individuals, representing 782 739 pyo, 11 276 deaths were observed. The cause-specific mortality fractions (CSMF) were (values in brackets refer to the InterVA results): HIV-related (including tuberculosis) 45% (50%); Other communicable diseases 12%(6%); Non-communicable lifestyle related conditions 16%(12%); Other non-communicable diseases 4%(2%); Maternal, perinatal, nutritional and congenital causes 2%(1%); Trauma 8%(7%); Undefined causes 12%(19%). Over the course of the ten years of observation the CSMF of HIV-related causes declined from a high of 56% in 2002 to a low of 38% in 2009 with the largest decline starting in 2004 following the introduction of an anti-retroviral treatment program into the population. The all-cause mortality rate declined over the same period from a high of 17.1 (95% CI 16.2-18.0) deaths per 1000 pyo in 2003 to a low of 11.7 (95% CI 11.0-12.5) in 2009. The decline in the all-cause mortality rate is predominantly due to a decline in the HIV-related mortality rate which declined in the same period from 8.7 (95% CI 8.1-9.4) to 4.7 (95% CI 4.3-5.2) deaths per 1000 pyo.

### Conclusions

Although the standard verbal autopsy questionnaire variables do not map completely to the input indicators required by the InterVA system, the system produced comparable results to physician coding. Minor changes to the standard questionnaire will result in a much better fit to the InterVA input indicators. Automated instruments such as InterVA have the potential to reduce the time and cost associated with physician based allocation to cause of death categories.

Verbal autopsy based methods enable the timely measurement of changing trends in cause specific mortality in order to provide policy makers with the necessary information to allocate resources to appropriate health interventions.

## **Central Medical Evaluation software: A web-based verbal autopsy physician coding system used in the Million Death Study**

Rajeev Kamadod<sup>1</sup>, Varun Malhotra<sup>2</sup>, Peter Rodriguez<sup>2</sup>, Chinthanie Ramasundarahettige<sup>2</sup>, Raju Jotkar<sup>3</sup>, Wilson Suraweera<sup>2</sup>, Stacia Neale<sup>1, 2</sup>, Prabhat Jha<sup>2</sup>

<sup>1</sup> Kentropy Technologies Pvt. Ltd, Bangalore, India

<sup>2</sup> Centre for Global Health Research, Li Ka Shing Knowledge Institute, St. Michael's Hospital, Dalla Lana School of Public Health, University of Toronto, Canada

<sup>3</sup> St. John's Research Institute, St. John's National Academy of Health Sciences, Bangalore, India.

### **Background**

Rapid physician coding of verbal autopsies drawing on narrative and structured questions, paired with simple forms for field data collection by non-medical staff are both central to enabling truly large scale mortality surveys (i.e. in several tens of thousands versus several hundred). The first phase of India's Million Death Study recruited 130 physicians who coded 123,000 verbal autopsy records on paper, based on electronic extracts of simple field forms completed by non-medical staff. We present a novel web-based system – Central Medical Evaluation-2 (CME2) – with rich interface that allows physicians to improve and expedite coding for verbal autopsy using electronic entry.

### **Methods**

A correction algorithm finds errors in verbal autopsy records and then removes them. Physician coding data collected for 2001–2003 deaths is used to construct a medical lexicon with keywords referring to specific diseases and each ICD-10 linked to differential diagnoses. Further, a Delphi method is used to devise clinical guidelines, identify age and gender restrictions related to ICD codes and ensure that keywords and differentials are clinically plausible.

### **Results**

CME2 software provides rich interface with tools for physicians to highlight the verbal autopsy narrative, type keywords to access related ICD codes and subsequently, consider alternate diagnoses. Only clean records are sent to physicians by the correction algorithm thus expediting coding. Physicians' ability to identify problems with records further adds to the quality of coding. Other advantages include permanent archival storage, easy re-sampling and automated generation of analytic and managerial reports through various stages of the coding process.

### **Conclusions**

CME2 software improves quality and expedites physician coding by providing adequate tools. Future work involves creating CME2 application for mobile phones/PDAs, computer desktop icons to provide real-time update on coding process, and integrating support processes such as recruitment, physician training and payment within the web-based system.

# Challenge of Verbal Autopsy in the National Death Registration System in Sri Lanka: the past and the future

Dharmaratne S.D.<sup>1</sup>, Jayasuriya R.L.<sup>1</sup>, Perera M.B.Y.<sup>2</sup>, Gunasekara E.M.<sup>2</sup>, N. Sathasivayyar A.<sup>2</sup>

<sup>1</sup> Department of Community Medicine, Faculty of Medicine, University of Peradeniya, Peradeniya, Sri Lanka.

<sup>2</sup> Registrar General's Department, Colombo, Sri Lanka.

## Background

Analyses of cause-of-death (COD) statistics are fundamental for monitoring the health situation of populations and for planning suitable interventions. Therefore, accuracy of national COD statistics is important for prioritizing health problems to take necessary interventions and for resource allocations. Coverage of the Death Registration System (DRS), which is a major component of the Civil Registration System (CRS) of Sri Lanka is almost 100% and completeness of death registration is above 90%. But, quality of COD statistics has been found to be deficient as 23%-30% of causes of death are categorized into "signs, symptoms and ill defined causes". All deaths coded into these categories are decision making. Verbal autopsy is used worldwide to determine COD and could be used to complement the DRS in improving data quality.

The reasons for low quality of COD statistics have not been comprehensively studied to date. Therefore, we conducted a detailed analysis of the existing DRS in Sri Lanka to identify the causes of the high proportion of R00-R99 (ICD codes for "signs, symptoms and ill defined causes") and to determine the feasibility of introducing verbal autopsy (VA) to improve the quality of COD statistics by supplementing the DRS in Sri Lanka.

## Methods

We searched MEDLINE and Google data bases for published articles from 1900 to 2010, using the following key words; cause of death and mortality statistics for Sri Lanka. Further, the DRS of Sri Lanka was extensively reviewed using documents available at the Registrar General's Department (RGD), Post Graduate Institute of Medicine, University of Colombo and other governmental and non-governmental organizations and institutes. In addition, we conducted in-depth interviews with the Registrar General (RG) of Sri Lanka and with relevant officers in his department and outside who are engaged in collecting, analyzing and coding COD statistics to identify related issues.

## Results

In Sri Lanka, for deaths that occur in hospitals, a COD is declared by the medical officer who attended the diseased, by completing a death declaration form. Excluding 'sudden deaths', for all other deaths that occur outside a hospital the COD is given by the Death Registrar (DR). 'Sudden deaths', (which are a small proportion of total deaths) that occur outside a hospital are attended by an Inquirer into Sudden Death or by Courts of Law.

The majority of the DRs are lay people with minimal or no training on deciding on a probable COD. Deaths that occur outside a hospital, in the majority of instances do not have a death declaration made by a medical officer.

A short questionnaire, named "verbal autopsy" (annexure I), developed by the RG's department has been used by the DRs to arrive at a probable COD since 2006. Subsequent to studying this questionnaire and with information received through interviews with DRs on this instrument, we concluded that this is not adequate to obtain a probable COD (6).

Several studies conducted in Sri Lanka have highlighted the biases that are present in the DRS. They point out that only 30-40% of the registered deaths occur in a government hospital and that 80% of registration of deaths is done by non-medical registrars. A study done in 1996 to assess the quality and coverage of death certification found that 15.5% of the medical officers have misclassified the underlying cause of death and the use of ill-



defined terms (e.g. cardio-vascular arrest) was frequent (76.4%) as was the use of abbreviations leading to misclassification (26.4%).

A post enumeration survey conducted in 1981 to ascertain the extent of coverage of death registration revealed that overall completeness was 94%. Rates for urban, rural and estate sectors were 92.9 per cent, 92.2 per cent and 100.0 per cent respectively. Since 1981, no comprehensive evaluation has been carried out to measure the completeness of death registration in Sri Lanka..

### **Conclusion**

This analysis revealed two important findings. Firstly, around 70% of deaths occur outside healthcare institutions (hospitals) and are registered by lay registrars with insufficient knowledge to decide on a probable COD. This would be an important reason for the high proportions of causes that fall into R00-R99 categories in the ICD – 10 (23% - 30%).

Secondly, although the VAQ used in Sri Lanka has several limitations, VA has been introduced and the DRS have been sensitized to verbal autopsy and has already been accepted by the policy makers.

# Comparative methods in adult verbal autopsy: examining the ability of Symptom Pattern, Machine Learning, Tariff, and InterVA to accurately determine causes of death

Rafael Lozano<sup>1</sup>, Spencer L James<sup>1</sup>, Abraham D Flaxman<sup>1</sup>, Alireza Vahdatpour<sup>1</sup>, Sean Green<sup>1</sup>, Jeanette Birnbaum<sup>1</sup>, Ben Campbell<sup>1</sup>, Charles Atkinson<sup>1</sup>, Henry Kalter<sup>2</sup>, Mohsen Naghavi<sup>1</sup>, Alan Lopez<sup>3</sup>, Christopher JL Murray<sup>1</sup>

<sup>1</sup>Institute for Health Metrics and Evaluation, University of Washington, Seattle, Washington, USA

<sup>2</sup>Johns Hopkins University, Bloomberg School of Public Health

<sup>3</sup>University of Queensland School of Population Health, Brisbane, Australia

## Background

Verbal autopsy is an important technique for measuring cause-specific mortality patterns in populations without adequate vital registration systems. Trained interviewers collect detailed information on signs and symptoms of illness from people familiar with the deceased. For verbal autopsy to be useful, information to individuals or to society, it is essential that the results of these interviews can be mapped correctly to an underlying cause of death. Multiple methods have been developed to determine causes of death from verbal autopsy responses, including Bayesian Symptom Pattern techniques (SP), Machine Learning decision tree algorithms (ML), data-driven symptom-cause tariff computations, and the InterVA verbal autopsy tool. We evaluated each approach's ability to accurately determine both individual causes of death and also cause-specific mortality fractions.

## Methods

The Grand Challenges 13 Population Health Metrics Research Consortium (PHMRC) study collected over 7000 verbal autopsy interviews for adult deaths across 6 sites over the course of 3 years. Each interview investigated a death for which a gold-standard underlying cause of death was known. This dataset provides a validation environment for the four methods of analysing computer coded verbal autopsy. To assess the accuracy of each of the data-driven methods (SP, ML, and tariff) for computer coded verbal autopsy, we performed 25 replications of the following experiment: we extracted a test set of 25% of the interviews in the PHMRC VA dataset (stratified by underlying cause of death) and then applied the each method to the remaining 75% of the interviews, to obtain the method parameters. The InterVA method does not require a training set of data as the algorithm is pre-coded; however, it was used to evaluate the same test sets of data as the other methods to allow for direct comparison. We assessed the quality of each method by predicting for the test set and measuring the concordance and average relative error of cause specific mortality fractions. We compared these to the results of physician coded verbal autopsy (PCVA) on the same set of deaths, and then compared each method's results to each other method to ascertain which method performed best.

## Results

Preliminary results and direct model-to-model comparison show that overall concordance is approximately 38% for the Symptom Pattern, Machine Learning, and Tariff methods. This is 52% higher than preliminary assessments of PCVA, though the currently-available PCVA results did not include the medical record component of the verbal autopsies. By applying a ranking scheme to each model's cause posteriors, a 25% average relative error of the cause-specific mortality fractions predicted by the King-Lu algorithm was retained. The overall concordance from preliminary application of InterVA to the same data set was 25%.

## Conclusions

This validation experiment shows that the Symptom Pattern, Machine Learning, and Tariff methods attain similar overall concordances, outperforming PCVA and InterVA. This provides important evidence for efforts to determine the optimal approach to coding verbal autopsy results in populations with limited resources.

## Effects of different formats of data collection on verbal autopsy findings

Joshi R<sup>1,3</sup>, Praveen D<sup>2</sup>, Chow CK<sup>1,3</sup>, Neal B<sup>1,3</sup>

<sup>1</sup>The George Institute for Global Health, Sydney, Australia

<sup>2</sup>The George Institute for Global Health, Hyderabad, India

<sup>3</sup>University of Sydney, Sydney, Australia

### Aim

To determine the impact of the format of data collection on the pattern of mortality defined by a verbal autopsy-based mortality surveillance system in a rural Indian community.

### Methods

Verbal autopsies were done by primary healthcare workers for all deaths occurring between October 2006 and September 2007 in a community in rural Andhra Pradesh, India (total pop approx 185,629). Each questionnaire had a structured section, comprised of a series of check boxes, and a free text section, in which a narrative description of the events leading to death was recorded. For each death a physician coder was presented first with one section and then the other in random order with a 20 to 40 day interval between. A cause of death was recorded for each data format. Finally, after another 20 to 40 day interval the full combined questionnaire was presented and a gold standard cause of death was assigned.

### Results

1407 verbal autopsies were available for analysis, representing 94% of all deaths in the specified time period. The cause of death assigned using the structured format matched the gold standard in 1012 (72%) of cases at the chapter-heading level in ICD-10 with a kappa statistic of 0.66. For the free text format the corresponding figures were 989 (70%) and 0.64. The patterns of mortality determined by each method and by the combined questionnaire method were similar and the rank order of the ten leading causes of death was the same for the three methods.

### Conclusions

The format of the verbal autopsy data used to assign a cause of death did not substantively influence the pattern of mortality estimated for this community.

## Evaluating symptom recall patterns in verbal autopsy

Ian Riley<sup>1</sup>, Ben Campbell<sup>2</sup>, Spencer L James<sup>2</sup>, Devarsetty Praveen<sup>3</sup>, Veronica Tallo<sup>4</sup>, Rohina Joshi<sup>5,6</sup>, Christopher JL Murray<sup>2</sup>, Alan Lopez<sup>1</sup>

<sup>1</sup>University of Queensland School of Population Health, Brisbane, Australia

<sup>2</sup>Institute for Health Metrics and Evaluation, University of Washington, Seattle, Washington, USA

<sup>3</sup>The George Institute for Global Health, Hyderabad, India

<sup>4</sup>Research Institute for Tropical Medicine, Manila, Philippines

<sup>5</sup>The George Institute for Global Health, Sydney, Australia

<sup>6</sup>University of Sydney, Sydney, Australia

### Background

Verbal autopsy is an important technique for measuring cause-specific mortality patterns in regions without vital registration systems. In order to gather detailed information about signs and symptoms of the deceased, trained interviewers administer verbal autopsies to family members after the event. The effectiveness of the method relies on the premise that symptoms are accurately recalled and reported by people familiar with the deceased. Therefore, the timing of the VA interview can be an important determinant of the overall outcome [1]. Recall patterns may vary significantly based on context and demographics. For example, people interviewed too soon after a death may feel uncomfortable, whereas people interviewed too late after a death may not provide accurate information due to memory degradation [2, 3]. Understanding how timing affects verbal autopsy recall is both an ethical and practical question that needs more attention.

### Methods

The Grand Challenges 13 Population Health Metrics Research Consortium (PHMRC) study collected 11865 verbal autopsy interviews across 6 sites over the course of 3 years. In a subsequent study, 1021 households in Andhra Pradesh, India and 908 households in Bohol, Philippines were revisited to collect a second verbal autopsy on the same death. In this study, we compared the verbal autopsy responses from the initial stage of interviews with the follow-up interviews to assess the accuracy of symptom recall and its effect on accuracy of diagnosis. The study also worked to ascertain what impact, if any, different respondents or any other relevant variables might have on symptom recall.

### Results

In preliminary analyses, we found that symptom recall can vary as a function of time. We are now exploring how symptom recall varies by relationship to the decedent, site, cause of death and other factors.

### Conclusions

Accurate symptom recall is crucial to determining cause of death from verbal autopsy analysis. Therefore, the factors identified in this study should be considered in the design and implementation of verbal autopsy systems. This will help lead to more successful use of verbal autopsy as a means to determine individual causes of death and population-level cause-specific mortality fractions.

### References

[1] Soleman N, Chandramohan D, Shibuya K. Verbal autopsy: current practices and challenges. *Bull World Health Organ.* 2006; 84(3): 239-245.

[2] Gajalakshmi V, Peto R, Kanaka S, et al. Verbal autopsy of 48,000 adult deaths attributable to medical causes in Chennai (formerly Madras), India. *BMC Public Health.* 2002; 2:7.

[3] Mobley CC, Boerma JT, Titus S, et al. Validation study of a verbal autopsy method for causes of childhood mortality in Namibia. *J Trop Pediatr*. 1996; 42(6): 365-369.

# Evaluation of causes of Childhood Mortality in the Pakistan Demographic and Health Survey 2006-2007(PDHS) by Verbal Autopsy

Farrukh Raza, Zamir Suhag, Arjumand Rizvi, Zulfiqar A Bhutta

Department of Paediatrics and Child Health, Division of Women and Child Health, Aga Khan University, Karachi, Pakistan

## Background

In 2006-2007, the National Institute of Population Studies (NIPS) carried out its fifth Pakistan Demographic and Health Survey (PDHS). The PDHS is the largest household-based survey ever concluded in Pakistan. The survey adopted a two-stage stratified random sample design. The survey covered 972 out of 1000 sample points., 102,037 households (HHs) were selected for the sample. However at the time of field work only 97,687 HHs were occupied and amongst them 95,441 (98%) were successfully interviewed.

## Method

A standard and validated VA instrument was used to ascertain the causes of child death, to evaluate the causes of stillbirths a further modified instrument was used. The VA questionnaires had both close-ended and verbatim parts. For the allocation of cause of death, two methods were applied as mentioned below:

Allocation of cause of death using computerized algorithm: A hierarchal-based computerized algorithm was applied to stillbirths, neonatal and post maternal deaths by using conditions from closed ended questions of the VA questionnaire. Each algorithm was based on a biological, plausible set of conditions drawn from closed ended questions. For example, if a child with asphyxia was also preterm and had a secondary infection, asphyxia was taken as the primary cause of death. Similarly, if death occurred early, prematurity was chosen as the primary cause of death.

Allocation of cause of death using questionnaire and verbatim reviews: To review the narrative part of the questionnaire, a review process was undertaken. The expert reviewers' team was divided into two groups; A & B, and the two groups were completely blinded to each other. An ICD-10 based list of diagnoses classification was used. In the case of agreement between groups A and group B (concordance), the consensus diagnosis was recorded as a final diagnosis and if there was disagreement (discordance) between the two groups then the case was referred to an expert panel for review and final diagnosis.

Based on the two systems above, the expert panel assigned the final cause of death using both the computer algorithm as well as final manual assessments performed by the teams. The same system was used for classification of stillbirths, but it was found that there was no agreed system for the classification of stillbirths using the ICD10 coding system.

## Result

A total of 4,438 verbal autopsies were completed, of which 1,337 were stillbirths and 3,101 were children under five years of age. The current infant mortality rate was 78 deaths per 1000 live births, and the under-five mortality rate is 94 deaths per 1000 live births. There is a very close link between maternal and child mortality with 40% of all stillbirths occurring during childbirth overall. The major causes of death among children under five were birth asphyxia (22%), sepsis (14%), pneumonia (13%), diarrhea (11%), prematurity(9%) and unclassified cases (one in nine deaths). Pneumonia and diarrhoea deaths were responsible for over half of all infant and child deaths. The data support a strong focus on addressing newborn deaths and reducing deaths from diarrhoea and pneumonia which are well-known causes of death among children. The data show that girls are more likely die with pneumonia and diarrhea compared with boys who die mainly due to birth asphyxia, sepsis and prematurity in neonatal period. The data indicate that over half of the stillbirths occur in antepartum period (mostly due to

APH and pre-eclampsia/eclampsia) and 42% in intrapartum period due to asphyxia. Childhood mortality is highest among children whose mothers have no education. The patterns and location of deaths suggest that poor care seeking as well as quality of care in the health system are important determinants of newborn and child deaths. We found concordant (1022) and discordant (1974) cases in under five children on cross tabulation of two methods (i.e. computer algorithm and reviewers diagnosis). Results indicate that asphyxia (37.5%), sepsis (14.4%), prematurity (17.1%), pneumonia (7.8%) and diarrhoea (7.7%) were concordant or matched cases among the two methods of ascertaining cause of death. For example in stillbirths, intrapartum asphyxial SB (35.9%) and unexplained antemartum SB (33.2%) were concordant cases. The PDHS data highlights that most of the deaths are preventable, such as diarrhoea, pneumonia and asphyxia, in early newborn and asphyxia-related stillbirths. These findings strengthen the argument that by integrating and scaling up existing evidence-based interventions within community and outreach programs, most of the deaths are avoidable at national level.

# Experimental validation of ranking-rounding method for verbal autopsy: imposing mortality fraction constraints on individual cause-of-death assignments

Spencer L James<sup>1</sup>, Abraham D Flaxman<sup>1</sup>, Alireza Vahdatpour<sup>1</sup>, Rafael Lozano<sup>1</sup>, Alan Lopez<sup>2</sup>, Christopher J.L. Murray<sup>1</sup>

<sup>1</sup>Institute for Health Metrics and Evaluation, University of Washington, Seattle, WA, USA

<sup>2</sup>University of Queensland School of Population Health, Brisbane, Australia

## Background

Verbal autopsy is a method of gathering cause-specific mortality information in areas without a vital registration system. Past research in verbal autopsy techniques has mostly focused on making individual cause-of-death assignments, or on ascertaining population-level mortality patterns such as cause-specific mortality fractions. Verbal autopsy methods can predict individual cause-of-death assignments using data-driven approaches [1] or expert algorithms [2]. Other techniques have been developed to use verbal autopsy responses in both hospital and community deaths to determine the cause-specific mortality fractions in the community deaths [3]. Previous work in verbal autopsy has focused on developing and validating the population-level and individual-level cause-of-death assessment. These techniques, however, have not been used in tandem to predict individual cause assignments constrained by the predicted cause distribution.

## Methods

The Grand Challenge 13 Population Health Metrics Research Consortium Project (PHMRC) study collected 7289 adult verbal autopsy interviews across 6 sites over the course of 3 years. This data set was used as a validation environment for the ranking-rounding method for computer coded verbal autopsy. The ranking-rounding method (RR) makes individual cause assignments based on cause-ranked posterior probabilities from the Bayesian Symptom Pattern method, the Machine Learning Random Forest method, and the Tariff method, but cause assignments are only made within a prior cause-specific mortality fraction distribution determined by the King-Liu method. To assess the accuracy of the ranking-rounding method for computer coded verbal autopsy, we performed the 25 replications of the following experiment: we extracted a test set of 25% of the interviews in the PHMRC VA dataset (stratified by underlying cause of death) and then applied the Symptom Pattern, Tariff, and Machine Learning algorithms to the remaining 75% of the interviews, to obtain the posterior probability ranks of individual deaths for each cause, where the rank indicates the probabilistic hierarchy of individual deaths within each cause. The individual deaths were then assigned to causes based on their ranking within each cause and the previously-determined CSMF for that cause. The order in which causes receive their assigned deaths is based on each cause's overall rank ROC. We assessed the quality of the RR method by predicting using the test set and measuring the concordance and average median error of cause specific mortality fractions. We compared these to the results of physician coded verbal autopsy (PCVA) on the same set of deaths.

## Results

In preliminary assessments, we applied the ranking-rounding method to Tariff Model predictions in the PHMRC data. In a head-to-head comparison, applying the RR algorithm resulted in a 10% better performance in terms of prediction concordance with true cause, while maintaining an overall 25% average relative error of CSMF predictions.

## Conclusions

This validation experiment shows that the RR method is a valuable approach to computer coded verbal autopsy, particular for environments in which both the CSMF distribution and individual-level cause assignments need to be reliably predicted. The flexibility of the algorithm also allows for easy implementation of more accurate CSMF estimates, or novel individual-level cause assignments.



## References

- [1] Murray, C. J. L., Lopez, A. D., Feehan, D. M., Peter, S. T., & Yang, G. (January 01, 2007). Validation of the Symptom Pattern Method for Analyzing Verbal Autopsy Data. *Plos Medicine*, 4, 11, 1739-1753.
- [2] Assessing a new approach to verbal autopsy interpretation in a rural Ethiopian community: the InterVA model. (March 01, 2006). *Bulletin of the World Health Organization*, 84, 3, 204-10.
- [3] King, G., & Lu, Y. (January 01, 2008). Verbal Autopsy Methods with Multiple Causes of Death. *Statistical Science : a Review Journal of the Institute of Mathematical Statistics*, 23, 1, 78.

## **Experimental validation of the Physician Coded Verbal Autopsy (PCVA) method for verbal autopsy: establishing a baseline for performance**

Charles Atkinson<sup>1</sup>, Rafael Lozano<sup>1</sup>, Abraham D Flaxman<sup>1</sup>, Spencer L James<sup>1</sup>, Alan Lopez<sup>2</sup>, Christopher JL Murray<sup>1</sup>

<sup>1</sup>Institute for Health Metrics and Evaluation, University of Washington, Seattle, Washington, USA

<sup>2</sup> University of Queensland School of Population Health, Brisbane, Australia

### **Background**

Physician Coded Verbal Autopsy (PCVA) is the most common approach for evaluating Verbal Autopsy (VA) data. With increasing use of VA in developing countries, it is important to validate the method to ensure comprehensiveness and consistency of the interview instrument. Higher accuracy of cause assignment by physicians leads to more precise estimation of cause-specific mortality fractions. However, it is challenging to validate the accuracy of PCVA, since the true underlying causes of deaths are difficult to know. In the Grand Challenges 13 Population Health Metrics Research Consortium (PHMRC) VA study, blinded interviews were collected on deaths for which a gold-standard underlying cause of death was known., providing a validation environment for the PCVA approach.

### **Methods**

The PHMRC study collected 11,865 verbal autopsy interviews across 6 sites over the course of 3 years. Physicians at the collection sites reviewed verbal autopsies to fill out death certificates using their expert judgement. All verbal autopsies were reviewed without access to accompanying medical record information, and 50% were additionally reviewed by a second physician with medical records included. It was previously found that duplicate verbal autopsy coding offers little advantage over single reviews [1]. To further investigate this finding, a 10% sample of the total VAs was distributed to another physician at the same site for double-physician review. A separate 10% sample was sent for cross-site review to examine site-specific effects in physician coding. The death certificates were input into the National Center for Health Statistics' Mortality Medical Data System (MMDS) for automated ICD-10 coding. The ICD-10 codes were mapped to the original PHMRC cause list to compare to the gold-standard underlying cause of death.

### **Results**

Preliminary results show that the Physician Coded Verbal Autopsy (PCVA) method produces an estimated overall concordance of less than 30%, with significant variation by cause of death.

### **Conclusions**

The Physician Coded Verbal Autopsy (PCVA) method for analysis of the PHMRC gold-standard dataset establishes a baseline against which the performance of other verbal autopsy measurement techniques can be assessed. PCVA is widely used to assess of verbal autopsy questionnaires. Our study provides an extensive validation environment for assessing the accuracy of this approach.

### **References**

[1] Joshi, R., Lopez, A. D., MacMahon, S., Reddy, S., Dandona, R., Dandona, L., & Neal, B. (January 01, 2009). Verbal autopsy coding: are multiple coders better than one?. *Bulletin of the World Health Organization*. *Bulletin De L'organisation Mondiale De La Santé*, 87, 1, 51.

# Experimental validation of the Tariff Method for verbal autopsy: using empirical cause-symptom associations to levy cause-of-death assignments

Spencer L James, Abraham D Flaxman, Alireza Vahdatpour, Christopher JL Murray

Institute for Health Metrics and Evaluation, University of Washington, Seattle, WA, USA

## Background

Verbal autopsy is a method of ascertaining cause-of-death information from the signs and symptoms exhibited by a person prior to death. The development and refinement of data-driven verbal autopsy methods requires mapping the results of experimental verbal autopsies to true causes of death. This information is also required to rigorously evaluate the accuracy of different verbal autopsy algorithms. Previous efforts at developing successful verbal autopsy algorithms have lacked such validation data, which has inhibited adoption of this important health metrics technology.

## Methods

The Grand Challenges 13 Population Health Metrics Research Consortium (PHMRC) study collected 7289 adult verbal autopsy interviews across 6 sites (Bohol, Philippines; Andhra Pradesh, India; Uttar Pradesh, India; Dar es Salaam, Tanzania; Pemba Island, Tanzania; and Mexico City, Mexico) over the course of 3 years. Each interview investigated a death for which a gold-standard underlying cause of death was known. This dataset provides a development and validation environment for the Tariff Method for computer coded verbal autopsy. The Tariff Method uses empirical symptom-cause relationships (tariffs) calculated in a 75% sample of the data (stratified by cause). A symptom-cause tariff is computed in the following way for each cause  $j$  (46 different causes were used), for each symptom  $i$ :

1.  $\text{freq}_{i,j}$  = proportion of times symptom  $i$  was endorsed for cause  $j$
2.  $m_i = \text{Median}(\text{freq}_{i,1}, \dots, \text{freq}_{i,46})$
3.  $s_i = \text{StDev}(\text{freq}_{i,1}, \dots, \text{freq}_{i,46})$
4.  $\text{tariff}_{i,j} = (\text{freq}_{i,j} - m_i) / s_i$

In the remaining 25% of the data, an additive score based on the symptoms reported and not reported for each death and for each cause is then calculated, and the scores for each cause are then ranked within that cause. The rankings are then processed to determine the most likely cause of death for each death, as well as the most likely deaths for each cause of death. We assessed the quality of the Tariff Method by repeating this process 25 times and then measuring the average concordance by cause and average relative error of cause specific mortality fractions. We also compared these to the results of physician coded verbal autopsy (PCVA) on the same set of deaths.

## Results

Based on preliminary assessments, we found that the Tariff Method performs approximately 50% better than PCVA in terms of concordance. The average concordance for using all symptoms for each cause prediction was 39%, although with more refined methods for choosing the number of symptoms to include for each cause prediction, the average concordance is expected to increase.

## Conclusions

This validation experiment shows that the Tariff Method is a valuable approach to computer coded verbal autopsy, particular for causes which have symptoms with high tariff values, as deaths reporting these symptoms are quickly identified by the model. This provides important evidence in the efforts to determine the optimal approach to coding verbal autopsy results in any given situation.

## **Improving verbal autopsy design: Multi-site gold standard validation of 11,865 deaths in selected countries**

The Bill & Melinda Gates Foundation Grand Challenges in Global Health #13 Population Health Metrics Research Consortium

Executive Team: Christopher J.L. Murray, Robert Black, Alan Lopez

Core Team: Rafael Lozano, Charles Atkinson, Jeanette Birnbaum, Ben Campbell, Abraham D. Flaxman, Sean Green, Spencer L. James, Henry Kalter, Summer Lockett Ohno, Alireza Vahdatpour

Andhra Pradesh Site: Bruce Neal, Rohina Joshi, Devarsetty Praveen

Bohol Site: Ian Riley, Marilla Lucero, Diozele Sanvictores, Veronica Tallo

Dar es Salaam Site: Zulfiqarali Premji, Wafaie Fawzi, Saurabh Mehta, Mwanaidi Said

Mexico Site: Bernardo Hernandez Prado, Minerva Romero Alvarez, Sara Gomez, Dolores Ramirez Villalobos

Pemba Island Site: Sunil Sazawal, Usha Dhingra, Arup Dutta

Uttar Pradesh Site: Abdullah Baqui, Ramesh Ahuja, Aarti Kumar, Vishwajeet Kumar

### **Background**

To best guide policy making, priority setting, and program planning, it is essential to understand the leading causes of death for adults and children. For areas without sufficient vital statistics, verbal autopsy (VA) instruments provide an inexpensive and easily implemented tool for measuring a profile of many important causes of death. The instrument collects key information about demographic characteristics, symptoms of the deceased and possible risk factors from the next of kin. However, there is significant variation in the design and quality of verbal autopsies being used by different audiences in the field. Recognizing the need for improved strategies to measure population health, the Grand Challenges 13 (GC13) Population Health Metrics Research Consortium (PHMRC) brought together researchers at the Institute for Health Metrics and Evaluation at the University of Washington, the University of Queensland, Johns Hopkins University, and Harvard University to develop a science-based, standardized tool that is widely applicable. To achieve this goal, the PHMRC implemented a rigorous validation study using gold standard records with known causes of death.

### **Methods**

To conduct the validation study, a stringent list of diagnostic criteria was developed for a set of targeted causes of death. Researchers reviewed medical records in local facilities to identify deaths that met the requisite criteria to be enrolled as a gold standard cause of death case. For these cases, blinded interviewer teams implemented a verbal autopsy survey of the next of kin. Both the medical record information and the blinded verbal autopsy interviews were entered via double data entry and are being currently analyzed. The study was implemented in 6 research sites: Andhra Pradesh, India; Bohol, Philippines; Dar es Salaam, Tanzania; Mexico City, Mexico; Pemba Island, Tanzania; and Uttar Pradesh, India. The results were reviewed both by the sites and by researchers at the Institute for Health Metrics and Evaluation, to ensure the highest quality.

### **Results**

Through this study, we amassed what is, to our knowledge, the largest known verbal autopsy validation dataset. The Andhra Pradesh site collected 2,400 validated verbal autopsies: 1554 adult, 468 child, and 378 neonatal cases. The Bohol site collected 1898 total: 1260 adult, 264 child, and 374 neonatal cases. The Dar es Salaam site collected 3249 total: 1716 adult, 480 child, and 1053 neonatal. The Pemba Island site collected 730 total: 233 adult, 228 child, and 269 neonatal. The Uttar Pradesh site collected 2170: 1419 adult, 499 child, and 252 neonatal cases. Finally, the Mexico site is still in the midst of in the process of data collecting data, but has collected 1418 total to date: 1114 adult, 105 child, and 199 neonatal, with an anticipated total of 2100. The collective dataset is being used to develop new methods for analyzing verbal autopsy data and to improve the instrument itself in order to ultimately produce a user-friendly package for collecting and interpreting verbal autopsy data.

**Conclusions**

Together, the robust dataset collected through this study acts as a basis for testing the effectiveness of specific questions within the instrument, and will allow us to reduce the survey length to include only the most effective questions. In addition, we are able to evaluate the performance of the instrument in multiple locations and cultural contexts over the six sites. This study will result in a streamlined instrument that can be easily implemented in any various settings. By developing an objective and a rigorously validated survey tool, this study can act as a foundation for making recommendations on how to best collect cause of death information in settings where the vital registration is insufficient, inadequate, or non-existent.

## **Item reduction analysis in verbal autopsy research: the effects of symptom exclusion on metrics of success**

Spencer L James<sup>1</sup>, Abraham D. Flaxman<sup>1</sup>, Alireza Vahdatpour<sup>1</sup>, Rafael Lozano<sup>1</sup>, Alan Lopez<sup>2</sup>, Christopher J.L. Murray<sup>1</sup>

<sup>1</sup>Institute for Health Metrics and Evaluation, University of Washington, Seattle, WA, USA

<sup>2</sup>University of Queensland School of Population Health, Brisbane, Australia

### **Background**

Verbal autopsy methods attempt to accurately predict causes of death based on questions about the signs and symptoms experienced by a deceased person. Responses to these questions can be used with several analytical techniques to estimate the true cause of death. Previous work has investigated the success of these methods in estimating the true cause of death in addition to accurately assessing cause-specific mortality fractions. However, the relative importance of the different questions asked in the verbal autopsy survey instrument has not previously been examined when extensive validation data (including both verbal autopsy responses and true cause of death) is available. It would be preferable to remove questions from the instrument that do not cause any loss of performance, as this makes the survey process more efficient for the interviewer and less burdensome for the interviewee. These improvements will become increasingly important as implementation of verbal autopsy becomes more common in survey and census projects, or routine application in vital registration.

### **Methods**

Several analytical verbal autopsy methods currently under development were applied to the Grand Challenges 13 Population Health Metrics Research Consortium Project (PHMRC) study dataset. To analyse the specific effects of item reduction, the Tariff Model [1] was used to predict causes of death using all symptoms in the dataset. Symptoms with low cross-cause variance were then removed from the dataset, and the model was rerun. This process was repeated with an increasing number of symptoms removed in order to assess the extent to which additional symptoms help inform the model.

### **Results**

Preliminary results show that a 10% sample of symptoms can be excluded from the model without adversely affecting predictive validity. The initial model concordance was 39% with the full symptom set, and this remained constant after removing the 10% sample of symptoms. Removing 25% of the symptoms resulted in less than a 5% decrease in overall concordance.

### **Conclusions**

This experiment shows that full symptom inclusion is not required in order to achieve maximum concordance, and that the survey instrument can hypothetically be shortened without adversely affecting the subsequent model's predictive validity. Further analysis will focus on identifying the specific symptoms which can be removed, how many symptoms can be excluded without affecting performance, and on exploring whether qualitatively-similar symptoms can be identified and removed from the model.

### **References**

[1] James, S.L., Flaxman, A., Murray, C.J.L. (December 2010). Experimental validation of the Tariff Method for verbal autopsy. Research publication under submission to 2011 Verbal Autopsy conference.

## Machine Learning for Verbal Autopsy Analysis: Validation Study of Random Forest

Alireza Vahdatpour<sup>1</sup>, Sean Green<sup>1</sup>, Spencer L James<sup>1</sup>, Abraham D Flaxman<sup>1</sup>, Rafael Lozano<sup>1</sup>, Mohsen Naghavi<sup>1</sup>, Alan Lopez<sup>2</sup>, Christopher JL Murray<sup>1</sup>

<sup>1</sup>Institute for Health Metrics and Evaluation, University of Washington, Seattle, Washington, USA

<sup>2</sup>University of Queensland School of Population Health, Brisbane, Australia

### Background

This study investigated the use of Machine Learning techniques for Verbal Autopsy (VA) data analysis. A machine learning method for VA analyzes example (or training) data to capture the characteristics of causes of deaths and builds classifiers to find the underlying cause for interviews which lack a known cause of death. The use of verbal autopsy for cause-of-death assignment is becoming more frequently used in countries with incomplete vital registration systems. It is important that the results of these interviews be mapped correctly to the underlying cause of death in order for verbal autopsy to be useful to inform health policy. The Grand Challenges 13 (GC13) Population Health Metrics Research Consortium (PHMRC) collected 11,865 VA interviews across six sites over the course of three years. Each interview investigated a death for which a gold-standard underlying cause of death was known. As a result, it is now possible to validate the accuracy of various verbal autopsy coding techniques.

### Methods

In this study, we conduct an extensive investigation using RandomForest classifiers [1] for verbal autopsy analysis. Random Forest is an ensemble classifier that consists of many decision trees. It has been proven to be an efficient technique for classification especially where there is a large sparse variable set. However, due to high uncertainty and the sparse nature of signs and symptoms in real world verbal autopsy settings, none of the machine learning techniques perform as desired (they achieve less than 25% concordance with Gold Standard causes). We improved the classification accuracy of Random Forest by applying a two-phase enhancement technique.

In the first phase, we constructed single cause Random Forest classifiers to predict the likelihood of each cause for each death. In the second phase, a novel Ranking/Rounding method was used to select the most likely cause for each death. We also investigated several variations of these enhancements (including pair-wise cause classifiers, high posterior Random Forest, etc.), which did not improve performance substantially.

### Conclusions

As a result of these improvements, the average concordance rate increased from ~20% to ~38%. Compared to preliminary physician coded verbal autopsy, on average Random Forest performs ~50% better. There is substantial variance between the accuracy of the method for different causes of death. Preliminary results show that while Random Forest can detect causes such as “bite of venomous animal” and “drowning” with high accuracy (> 80% concordance), less common causes such as “Lymphomas” are more difficult to detect (< 20% concordance).

### References

[1] Leo Breiman. Random Forests. *Machine Learning*, 45(1):5–32, October 2001.

## **Quality of death certification in Mexican hospitals: preliminary results from the validation of cause of death assigned through death certificates vs. gold standard medical records review.**

Bernardo Hernández<sup>1</sup>, Dolores Ramirez<sup>1</sup>, Minerva Romero<sup>1</sup>, Sara Gómez<sup>1</sup>, Christopher Murray<sup>2</sup>, Rafael Lozano<sup>2</sup>.

<sup>1</sup>National Institute of Public Health, Mexico.

<sup>2</sup>Institute for Health Metrics and Evaluation, University of Washington, Seattle, Washington, USA

### **Introduction**

As part of the GC13 project, validation of a verbal autopsy instrument is being conducted in several sites around the world, including Mexico. Besides assessing the validity of the verbal autopsy instrument, it is important to assess the concordance between the causes of death (COD) assigned on death certificates and the ones obtained through a review of medical records, to provide a benchmark for the validity of the verbal autopsy.

### **Objective**

To assess the validity of the COD registered on death certificates compared with causes obtained from a review of medical records, considered gold standard, in a sample of hospitals in Mexico City and the State of Morelos, Mexico.

### **Methods**

Adult, child and neonatal deaths that occurred between 2009 and 2010 were identified in public hospitals. Trained physicians reviewed medical records (including autopsy reports when available) associated with those deaths and assigned them a COD, considered the gold standard. In addition, information from the death certificates of those cases was obtained and coded by three independent coders. For each specific COD, we estimated the proportion of cases that would reveal concordance between COD assigned on the death certificate and COD assigned by physician review, and calculated Cohen's Kappa.

### **Results**

446 adult and 41 neonatal/child deaths have been reviewed to date. Among adults, 50% of cases were assigned the same cause of death on the death certificate as that assigned as the gold standard. The causes of death with higher concordance were AIDS, leukemia, postpartum hemorrhage, as well as breast, cervical, prostate and gastric cancer. The causes with lower concordance were pneumonia, tuberculosis, lymphomas, colon and ovarian cancers, diabetes with diabetic coma or with renal failure, ischemic heart disease with acute myocardial infarction, eclampsia, and renal failure. Among children and neonates, the concordance was low for all causes of death in the study (22%).

### **Conclusion**

These preliminary results indicate that the concordance between the assignment of causes of death in death certificates vs. medical records is low, suggesting actions be taken to improve the quality of routine death certification. This information provides an allocation basis for understanding the likely true distribution of causes of death using verbal autopsies.

\*We acknowledge the support of the General Hospital of Mexico and the Dr. Manuel Gea González Hospital, the National Institutes of Respiratory Diseases and Oncology of Mexico, as well as Hospitals from the Secretary of Health from the Federal District and the State of Morelos in the conduction of this study



## Sample registration of vital events with verbal autopsy in Zambia

Sheila S. MUDENDA<sup>1</sup>, Stanley KAMOCHA<sup>2</sup>, Robert MSWIA<sup>3</sup>, Palver SIKANYITI<sup>1</sup>, William C. MAYAKA<sup>1</sup>, Brivine SIKAPANDE<sup>4</sup>, Dara POTTER<sup>2</sup>, Melissa MARX<sup>2</sup>

<sup>1</sup>Central Statistical Office, Zambia

<sup>2</sup>Centers for Disease Control and Prevention, Zambia

<sup>3</sup>Futures Group/MEASURE Evaluation

<sup>4</sup>Ministry of Health, Zambia.

### Background

A well-functioning vital events registration system that produces accurate statistics on births, deaths, and the causes of death is fundamental to evidence-based health policy and evaluation. Most developing countries do not have functioning vital statistics registration systems (1). We estimate that fewer than ten countries in sub-Saharan Africa have systems that produce usable data. In spite of having the necessary regulatory framework that supports the maintenance of a vital statistics system in Zambia, the current system does not generate usable vital statistics to inform health policy. Data on the number and causes of death in Zambia are lacking. Currently reported data are based on numbers of deaths recorded in health facilities. The majority of deaths occur at home in Zambia, like many other developing countries (1). Thus, coverage of the current system is estimated to be low. Mortality estimates have not been sufficiently reliable for setting health sector priorities, or for assessing program progress and impact. To address this gap, the Zambian Central Statistical Office (CSO) and the Zambian Ministry of Health are implementing a pilot sample registration of vital events system with verbal autopsy in order to collect age, sex and cause-specific mortality data.

### Implementation Approach

Sample Vital Registration with Verbal Autopsy (SAVVY) is community-based and implemented in a nationally representative cluster sample. In Zambia, SAVVY collects data continuously on births and deaths. It also conducts periodic independent re-enumeration of populations to verify resident populations and assess birth and death registration completeness (2). The primary outcomes of the SAVVY in Zambia (SAVVY-Z) are causes of mortality and fertility rates. SAVVY-Z is being implemented in two phases. The first phase is being implemented in four provinces. The second expands coverage to all nine provinces by 2011. SAVVY-Z is using the standard World Health Organization (WHO) (3) verbal autopsy (VA) questionnaires for the collection of data on neonatal, child and adult deaths, and the causes of death. Demographic characteristics including sex, age, marital status, education, place of death, health services utilization, rural or urban residence, and province are collected enabling assessment by these covariates. Nurses and other medical personnel have been trained by Measure Evaluation as verbal autopsy interviewers, and are assisted by key informants whose duty is to inform the data collectors about births, deaths and vital events that occur in a cluster. These key informants were chosen from the communities/clusters where the SAVVY-Z is being implemented. Nine physicians were trained by Measure Evaluation on VA questionnaire review, how to assign an immediate and underlying cause of death based on the International Classification of Diseases tenth revision (ICD-10) coding principles and guidelines and death certificate production.

The baseline census was implemented in the selected clusters and was followed by verbal autopsy interviews for all deaths in households reported to have occurred in the 12 months preceding the baseline census. After the VA interviews were conducted, two physicians independently reviewed each VA questionnaire to determine a probable cause of death. Then, each physician completed a death certificate for the VA death and assigned an ICD-10 code. The death certificates and ICD-10 codes were then compared. If the two coders agreed on a cause of death, the assigned cause is considered final. If the two coders disagreed on the cause of death, then the two physicians review the VA questionnaire together to reach an agreement on the most likely cause of death. After

12 months of continuous data collection, an updated census will be conducted to ensure all people residing in the selected clusters are included.

### **Data Analysis and Outcomes**

SAVVY and VA data will be analyzed to produce mortality rates by age and sex, as well as cause-specific mortality fractions (CSMF) for the leading causes of death.. Infant and child mortality rates, maternal mortality ratios, HIV/AIDS-related mortality, and intervention-addressable conditions and summaries of place of death, health services utilization in the period before death - by type of service, treatment and by leading causes of death, will also be presented.

### **Results**

Final results from SAVVY implementation from January 2010 to January 2011 in the four provinces are expected by June 2011. Preliminary results from 4 provinces (*n=388 deaths*) indicate that overall, the leading causes of death in Zambia include HIV/AIDS (28% of all reported deaths, *n=107 deaths*), injuries and accidents (10%, *n=40 deaths*), malaria (8%, *n=32 deaths*), malnutrition (7%, *n=27 deaths*), hypertensive/ischemic heart diseases (6%, *n=22 deaths*) and perinatal and early neonatal conditions (5%, *n=19 deaths*). The remaining 141 deaths were caused by hypertensive/ischemic heart diseases, tuberculosis, pneumonia, neoplasms, intestinal infectious diseases, meningitis, maternal conditions, cerebrovascular diseases, anaemia and diabetes mellitus each causing less than 5% of the total. The CSMF vary by province. Fifty-two percent (*n=202 deaths*) of all reported deaths occurred at home, and about 40% (*n=155 deaths*) of deaths occurred in hospitals or other health facilities. Place of death varies considerably among the provinces. For example in Luapula province, close to 70% (*n=65 deaths*) of all reported deaths occurred at home while in Lusaka, less than 37% of deaths occurred at home, with or without contact with formal health care in the period leading to death.

### **Policy Impact of SAVVY implementation in Zambia**

The findings from SAVVY-Z provide national and sub-national mortality burden estimates for the Ministry of Health and other stakeholders to inform decision-making, policy-formulation, planning, and for implementation, monitoring and evaluation of health intervention programs.

### **References**

1. MEASURE Evaluation. Sample Vital Registration with Verbal Autopsy (SAVVY); A Resource Kit for implementing the SAVVY system 2007.
2. Chandrasekar C, Deming W. On a method of estimating birth and death rates and the extent of registration. *Journal of American Statistics Association* 1949;44:101-15.
3. World Health Organization. International Standard Verbal Autopsy Questionnaires. Available from [www.who.int/whosis/mort/verbal/\\_autopsy\\_standards2.pdf](http://www.who.int/whosis/mort/verbal/_autopsy_standards2.pdf)

## Social Autopsy: Concept and Development

Henry D. Kalter<sup>1</sup>, Marzio Babilie<sup>2</sup>, Robert E. Black<sup>1</sup>

<sup>1</sup>Johns Hopkins Bloomberg School of Public Health, Baltimore, USA;

<sup>2</sup>UNICEF/Chad Country Office

Health policy makers and health programmers require data on causes of death to identify health priorities, allocate sparse resources, and evaluate health programs. Such data on causes of death often come from verbal autopsy. Effective delivery of child survival interventions that can overcome these causes of death depends on improved understanding of modifiable social, cultural and health systems factors affecting health care access and utilization. This was shown by the Integrated Management of Childhood Illness (IMCI) multi-country assessment, which found that although IMCI health facilities provided the gold standard of child illness care in developing countries, the strategy failed to decrease child mortality. In part this was due to inadequate implementation of IMCI's household and community component, and assuming that quality health care services alone would lead to increased care seeking and appropriate home care practices.

The "Pathway to Survival" conceptual framework was designed to support IMCI implementation, with an emphasis on the household and community component, by identifying the steps that families, communities and health systems must take to prevent illness and return sick children to health. The "Pathway Analysis" social autopsy instrument was developed to assess these factors, together with a verbal autopsy, in children who suffered a fatal illness. As shown with data from Guinea, this tool has been employed in several countries to collect data that supported the formulation of health interventions with increased access, coverage, and utilization. Community participation in these studies seemed to increase the use of the data, the implementation of effective programs, and appropriate health care seeking.

Recently, the Child Health Epidemiology Reference Group (CHERG) undertook a review and update of the Pathway Analysis social autopsy instrument to improve its assessment of neonatal deaths; and of preventive, social and behavioral factors related to the deaths of children of all ages. The CHERG social autopsy tool has been integrated with the GC-13 verbal autopsy tool, and this integrated verbal/social autopsy (VASA) instrument will be used in several upcoming studies of neonatal and child mortality at country or sub-national level in sub-Saharan Africa.

Social autopsy is also an important part of maternal death review, which has been conducted in Britain for more than 50 years and more recently in several developing countries. The rationale is similar to that for child social autopsy. Simple, effective technologies exist to overcome maternal mortality, but health care access, coverage, and utilization are hampered by social and cultural factors as well as inadequate up scaling of effective health programs. Since 2005, in 10 states of India with technical assistance from UNICEF, Maternal Death Inquiry and Response (MADIR) has emphasized the importance of community and health system participation in the MADIR process as a means of increasing the visibility and awareness of maternal death as a preventable problem. This has empowered communities and engaged health programs to take effective action. MADIR data are presented that illustrate how the process has increased maternal death reporting, led to the collection of actionable data on the biological causes and social, cultural and health system factors related to maternal death, and supported the development of community- and health program-initiated interventions. In 2010 the Government of India implemented maternal death audits at the community and facility levels as a national program.

In conclusion, social autopsy findings have proven capable of raising awareness and the visibility of child and maternal mortality as preventable problems, and of supporting the development of appropriate, data driven interventions that increased access, coverage and utilization. The death inquiry and response process has built institutional awareness and political commitment to overcome these problems; and community participation in

the process may in itself act as an intervention. New, representative, neonatal and child VASA studies are planned for several African countries.

## **Strengthening of mortality registration system and causes of death statistics: Experience of Indonesia**

Soewarta Kosen<sup>1</sup>, Chalapati Rao<sup>2</sup> and Tati Suryati<sup>1</sup>

<sup>1</sup>Center for Health Systems and Policy Research and Development, NIHRD, Jakarta

<sup>2</sup> University of Queensland School of Population, Brisbane, Australia

The Indonesian mortality registration system and cause of death statistics in Indonesia are characterized poor coordination among different responsible institutions, use of non-standardized data collection nomenclature, limitations in the understanding and skills of personnel responsible for registration, poor coverage and the absence of a basic legal framework for population data collection. With the majority of deaths occurring at home, reliable data on the underlying causes of death are difficult to obtain. Even for deaths occurring in health facilities, the data on causes of death has not been compiled in a systematic manner to yield proper statistical tabulations.

With the enactment of Law No. 23/2006 on Population Administration and its related regulations, there is an opportunity to improve the quality of registration as well as the application of multiple causes of death based on the International Classification of Diseases and Related Health Problems (ICD – 10, 1992). Health systems research to strengthen mortality registration and cause of death statistics has been implemented by the National Institute of Health Research & Development in Surakarta City and Pekalongan District, Central Java, with technical support from the Ministry of Home Affairs and the University of Queensland, Australia, and financial support from WHO Indonesia and AusAID. The objectives are to institutionalize mechanisms for collection and compilation of mortality statistics within the broader process of civil registration, to support institutional strengthening of civil registration offices, community health centers, hospitals, and other institutions at various levels and to generate routine and timely evidence as input for health policy formulation and planning.

The pilot registration system determines causes of death using two methods. For deaths occurring in a health facility, a Medical Certificate of Cause of Death (MCCD) is completed by the attending physician. Information on deaths occurring at home is collected by a nurse or a midwife using a structured verbal autopsy instrument. The completed instruments are reviewed by health center physicians who then complete a certificate of cause of death. All death certificates are sent to the District/City Health Office, where a team of trained medical record staff codes causes of death according to the ICD-10 and apply prescribed conventions to select the underlying cause of death. The results are promising and showed the increased coverage of mortality registration in the pilot area and the identification of major underlying causes of death from non-communicable diseases, especially cerebrovascular diseases (stroke). The findings have implication for policy and further action, such as control of risk factors. Expansion of this cause of death registration system throughout Indonesia would be invaluable for assessing progress with health development and monitoring the achievements of health programs and MDGs.

# Trends in causes of death among children under-five in Bangladesh: An exercise applying standard computer algorithm to assign cause of death using verbal autopsy data

Li Liu<sup>1</sup>, Qingfeng Li<sup>2</sup>, Rose Lee<sup>1</sup>, Neff Walker<sup>1</sup>, Robert E. Black<sup>1</sup>

<sup>1</sup>Department of International Health, Johns Hopkins Bloomberg School of Public Health

<sup>2</sup>Department of Population Family and Reproductive Health, Johns Hopkins Bloomberg School of Public Health

## Background

Child causes of death serve as important public health information to facilitate policy making, resource allocation, and program evaluation. With child mortality continuing to drop at the global and national levels, distribution of child cause of death change accordingly. It is hence of great interest to understand trends in child cause of death, especially using empirical data. The three rounds of Bangladesh Demographic and Health Survey (BDHS) with verbal autopsy (VA) module provide a unique opportunity to study such a topic.

## Methods

The 1993/94, 1996/97, and 2004 BDHS+VA questionnaires and datasets were obtained through ICF Macro. The 93/94 and 96/97 VA studies are similar in general, but are different from the 2004 study in many ways. In particular, more detailed symptom data were collected in the 2004 survey. To generate comparable cause of death distribution, a standard algorithm needs to be generated to assign cause of death.

In this study, we choose to use a computer-based algorithm. A computer algorithm is usually composed of two parts, the cause-of-death assigning hierarchy and the case definition for each cause. The hierarchy and the case definitions are standardized by first reviewing and comparing relevant information from each of the three surveys. A standard hierarchy is then determined by harmonizing the hierarchy across the three surveys and referencing hierarchies applied in other DHS+VA surveys. A similar review process is done to determine the standard case definitions. In surveys where the case definitions are not available, those from another BDHS or the 1999 WHO validation study<sup>1</sup> are used instead. To standardize the entire algorithm, the cause of death categories are adapted from the Child Health Epidemiology Reference Group (CHERG) causes.<sup>2</sup> The standard computer algorithm is then applied to all three DHS+VA datasets.

## Results

The hierarchies applied in the 93/94, 96/97 and 2004 BDHS+VA surveys, as well as the standard hierarchy are presented in Figure 1. Compared to the 93/94 and 96/97 surveys, more causes can be assigned when additional symptoms data were collected in the 2004 survey. Neonatal causes were also better assigned in 2004. Three important causes, measles, diarrhea, and ARI, were assigned simultaneously in all three BDHSs. In the standard hierarchy, however, we choose to assign the three causes in the order of measles followed by diarrhea and then by ARI instead. Similar hierarchy has been applied in more recent DHS+VA studies, such as the Nepal 2006 survey.<sup>3</sup> Additionally, in the standard hierarchy, meningitis/encephalitis is separated from other possible serious infections and treated as a separate cause. The case definitions for meningitis are borrowed from the WHO validation study.<sup>1</sup> The detailed comparison of case definitions across the three surveys and the standard case definitions are presented in Table 1.

During the study period, the under-five mortality rate dropped from 133.1 per 1,000 live births in the 1993/94 survey to 115.7 in 1996/97, and then to 88 in 2004 in the five years preceding the survey.<sup>4-6</sup> In the meantime, neonatal mortality reduced from 52.3 in 1993/94 to 48.4 in 1996/97, and then to 41 in 2004 during the same referencing period. Corresponding to the decline in under-five and neonatal mortality, proportion of children dying in the neonatal period increased from 38% in 1993/94 to 42% in 1996/97, and then to 56% in 2004 (Figure 2).

With the neonatal slice increasing across the three surveys, the top ranking neonatal cause of death changed from tetanus in 93/94 and 96/97 (6% and 7%, respectively) to birth asphyxia (11%) in 2004 (Figure 2). Other possible serious infections and prematurity/low birth weights also rank higher proportionately among neonatal causes. Among children ages 1-59 months, however, pneumonia remains the top killer across time, with 12%-15% of deaths attributable to this cause. The contribution of diarrhea reduced from 13% and 12% in the first two surveys to only 6% in the 2004 survey. Injury also poses a major threat to child survival, contributing 4% to 8% to the postneonatal deaths. Among all children under five years of age, pneumonia is responsible for 15% to 19% of deaths. Diarrhea was the second important cause in the 93/94 and 96/97 surveys, accounting for 14% of deaths. However, in the 2004 survey, birth asphyxia took over as the second important cause, claiming 11% of the lives of children under-five in Bangladesh.

### **Discussion**

This study applies a standard computer-based algorithm to assign cause of death using symptom data collected through nationally representative verbal autopsy studies. The results provide insights into the trends in distribution of child cause of death for one decade. Our results corroborate the previous finding that pneumonia remains the top ranking cause of death among children below five years of age.<sup>2</sup> It also demonstrates changes in the distribution of child causes of death when child mortality decreases, highlighting the increasing significance of neonatal causes.

Some methodological limitations are acknowledged to inform future research. First, despite the comprehensive review process, certain cause cannot be assigned due to lack of symptom data in some surveys (e.g. meningitis in the first two surveys). The computer algorithm has been standardized to the degree that we possibly can. But because information is not consistently available across surveys, some cause fractions may not be directly comparable. For example, data on a number of signs of serious infections were collected in the 2004 survey which were not available from the two previous surveys. With the same case definition, i.e., the child had 2+ signs of serious infection, it is more likely to assign cause to other possible serious infections in 2004. So 10% of other possible serious infections in 2004 cannot be simply compared with the 4% in 96/97 and the 5% in 93/94. In addition, several important causes, such as neonatal sepsis and pertussis, still cannot be assigned because lack of established case definitions and relevant symptom data.

Despite these limitations, verbal autopsy studies combined with standard computer-based algorithms remain a promising tool to generate information on trends in child causes of death in places where vital registration is far from adequate.

### **Reference List**

1. Anker M, Black RE, Coldham C et al. A Standard Verbal Autopsy Method for Investigating Causes of Death in Infants and Children. Geneva, Switzerland: WHO; 1999.
2. Black RE, Cousens S, Johnson H et al. Global, Regional and National Causes of Child Mortality, 2008. *Lancet* 2010;375(9730):1969-87.
3. Ministry of Health and Population (MOHP) [Nepal], New ERA, Macro International Inc. Nepal Demographic and Health Survey 2006. Kathmandu, Nepal: Ministry of Health and Population, New ERA, and Macro International Inc.; 2007.
4. Mitra SN, Ali MN, Islam S, Cross AR, Saha T. Bangladesh Demographic and Health Survey 1993-1994. Calverton, Maryland: National Institute of Population Research and Training (NIPORT), Mitra and Associates, and Macro International Inc.; 1994.

5. Mitra SN, Al-Sabir A, Cross AR, Jamil K. Bangladesh Demographic and Health Survey, 1996-1997. Dhaka and Calverton, Maryland: National Institute of Population Research and Training (NIPORT), Mitra and Associates, and Macro International Inc.; 1997.
6. National Institute of Population Research and Training (NIPORT), Mitra and Associates, ORC Macro. Bangladesh Demographic and Health Survey 2004. Dhaka, Bangladesh and Calverton, Maryland [USA]: National Institute of Population Research and Training (NIPORT), Mitra and Associates, and ORC Macro; 2005.



## Using statistical models to predict malaria deaths in children under 5

Robert Mswia, Stirling Cummings, Siân Curtis

### Background

In 2005, it was estimated that somewhere between 300 and 500 million illnesses and nearly one million deaths resulted from malaria. The vast majority afflicted by this disease were children under five years of age. Currently, vital registration of deaths is limited or nonexistent in many of the sub-Saharan Africa countries where the risk of malaria morbidity and mortality are greatest. The collection of verbal autopsy (VA) data is one way the contribution of malaria to child mortality can be estimated in many of these countries, where there are inadequate vital registration systems and many deaths occur outside of the health care system and, therefore, are not recorded.

Currently, the most widely used approach for classifying VA forms is by physician review. This method is not without problems: it can be expensive, time consuming and biased by the physicians' knowledge of the local epidemiology. Overcoming these deficiencies has led many researchers to investigate other methods such as expert and data-driven algorithms, and probabilistic techniques to classify VA results. An algorithm or probability-based method for coding cause of death (COD) from VA data has the potential to lower costs, save time, and reduce coder biases in low income countries.

Though many researchers have found that these methods compare favorably to physician coding, there is little information published on how to go about developing these models and applying them to other VA data sets, in the absence of gold-standard cause of death evidence, such as a death certificate. In this paper, such a model is created using the physician-designated malaria deaths in children under 5 years of age and their symptoms from one data set, and then applying the model to a second data set. The model's performance is evaluated by how well it predicts physician-coded malaria deaths in the second data set. In order to minimize the effects, that different study protocols could have on the two data sets, data were selected from studies that used the most current VA questionnaires recognized by WHO, as well as the same enumerator and physician training protocols.

### Methods

The first VA data set, from which the model was derived, consists of 3,133 child deaths from the 2007 Mozambique Post-Census Mortality Survey (INCAM). A probability sample was selected from deaths reported by 2007 Mozambique General Population and Housing Census. A Verbal Autopsy questionnaire was completed for each eligible death enumerated by the census within the sampled areas. The International Standard Verbal Autopsy Questionnaire 2 (WHO) was used for the collection of the child mortality data.

The second data set, used to validate the model, is based on 283 child deaths from the 2008 Rwanda Verbal Autopsy study, a follow-up survey to the 2008 Rwanda Interim Demographic and Health Survey. The study used the same questionnaires and training materials for enumerators and medical coders as above, but adapted to the Rwandan context.

For each record in the Mozambique data, the underlying cause of death was dichotomized as either a 'Malaria' or 'Non-Malaria' death. The duration information for symptoms experienced by the child prior to death were dichotomized to reflect either a 'Yes' or 'No' response. Logistic regression analysis, using a forward selection process, was performed on these data in order to produce parameter estimates of a much reduced number of symptoms. The statistical model was then applied to the Rwanda data set in order to calculate a model-predicted probability of malaria death. The Rwanda model-based result was then compared to the Rwandan physician-coded result.

## **Results**

According to physician coded deaths of children aged 28 days to 4 years from the data, 49 percent of child deaths in Mozambique and 16 percent in Rwanda were due to malaria. Goodness-of-fit statistics for logistic regression analysis of the Mozambique data resulted in a model that fitted the data adequately. Twenty-eight symptom variables from the model were found to be statistically significant in predicting malaria deaths in Mozambique.

The parameter estimates from the Mozambique model were used to estimate the probability of malaria deaths in the Rwanda dataset, based on the symptoms of each of the deceased. Since the proportion of child deaths from malaria in Rwanda was approximately 16 percent, the upper 16th percentile of the model-based estimated probabilities of malaria death were coded as malaria deaths. Using McNemar's Test we concluded that there was a significant level of agreement ( $p$ -value=0.043) between the physician-coded cause of death and the model-predicted cause of death from malaria. The sensitivity of the model was low (0.29), however, the specificity was high (0.93).

## **Conclusions**

Statistical models show promise in predicting the cause of death from malaria in settings where there is little or no infrastructure to monitor vital events. Development of better data-driven models has the potential to reduce the cost of conducting mortality research as well as reducing the burden on health systems where medical coding of VA questionnaires requires trained medical personnel. Using data from two Sub-Saharan African countries that shared the same standard VA questionnaires and common data collection and ICD-10 cause of death certification protocols has minimized sources of bias from the analysis. This approach, with the added contributions of many researchers in this field, could serve as a means of monitoring progress with international malaria initiatives, such as Roll Back Malaria (RBM) and the President's Malaria Initiative (PMI).

# Using Verbal Autopsy in a Post-Census Mortality Survey to Capture Causes of Death in Mozambique, 2006-2007

Francisco Mbofana<sup>1</sup>, Renee Lewis<sup>2</sup>, Loraine West<sup>2</sup>, Elisio Mazive<sup>3</sup>, Stirling Cummings<sup>4</sup>, Robert Mswia<sup>4</sup>

<sup>1</sup>Mozambique Ministry of Health

<sup>2</sup>U.S. Census Bureau

<sup>3</sup>Mozambique National Institute of Statistics

<sup>4</sup>MEASURE Evaluation

## Background

Most developing countries and their development partners are committed to reducing the burden of disease, such as HIV/AIDS, with a number of initiatives focused on this goal. However, these actors often do not have adequate national and sub-national mortality estimates to inform their decision-making processes.

Implementing a post-census mortality survey is one way to supplement data from weak or nonexistent vital registration systems. Information gained from such a survey can provide a baseline to measure the impact of intervention programs so that funds can be allocated effectively and specific populations can be targeted.

This paper examines findings from a nationally representative post-census mortality survey utilizing verbal autopsy (VA) implemented in Mozambique following the 2007 census.

## Methods

A stratified cluster sample design was used to select rural and urban areas in all 11 provinces. Interviewers trained to administer VAs visited households with census-reported deaths. The World Health Organization standardized VA questionnaires were used. Trained physicians examined the completed verbal autopsy questionnaires, assigned a cause of death and coded the underlying cause of death using standards set by the International Classification of Disease, 10th Revision (ICD-10).

## Results

There were 10,080 completed VAs in our sample. The post-census mortality survey was successful in determining the leading causes of death in Mozambique and cause-specific mortality fractions (CSMF). Newborns and infant deaths accounted for the largest number of deaths, and approximately one quarter of all deaths occurred among children less than one year of age. Deaths of children between the ages of 1 to 4 years comprised the next largest age group, accounting for 19 percent of all deaths. Cumulatively, children under 5 years of age accounted for 43 percent of all deaths and half of all deaths were in children under the age of fifteen.

Malaria and HIV/AIDS were the leading causes of death and accounted for over 50% of all deaths. Furthermore, malaria deaths in children under 5 accounted for 18 percent of all enumerated deaths. Other leading causes of death included perinatal conditions, diarrhoeal diseases, pneumonia, accidents and external causes, diseases of the circulatory system, tuberculosis, and malignant neoplasms.

The post-census mortality survey also shed light on maternal mortality. There were 213 sample cases of maternal deaths in the survey; 4,803 weighted maternal deaths. These deaths accounted for 14 percent of deaths in women of reproductive age (15-49). Fifty-five percent of maternal deaths were due to direct obstetric causes. About 18 percent of maternal deaths were HIV/AIDS-related indirect obstetric deaths and the remaining 27 percent were due to indirect causes other than HIV/AIDS (these two shares are not significantly different). Furthermore, only two percent of maternal deaths occurred in the 24 hours post-delivery.

The survey was also useful for collecting information on use of health services prior to death and place of death, by cause, sex, age group, and area of residence.

### **Conclusions and Recommendations**

A post-census mortality survey using VA is a practicable community-based approach to collect mortality data at national and sub-national levels. Cooperation among countries and their development partners to institute the VA process can make up for weaknesses in, or lack of mortality data. However, linking the mortality survey (VA) to a census is dependent on the quality of the census, in terms of cartography, data collection, and data processing. The Mozambique Ministry of Health now has representative mortality data to guide planning and priority setting. The survey also provided indicators for international initiatives, such as PEPFAR, UNAIDS, and the Global Fund to Fight AIDS, TB and Malaria.

## Using verbal autopsy to track epidemic dynamics: the case of HIV-related mortality in South Africa

Peter Byass<sup>1,2</sup>, Kathleen Kahn<sup>1,2</sup>, Edward Fottrell<sup>2</sup>, Paul Mee<sup>1</sup>, Mark A. Collinson<sup>1,2</sup>, Stephen M. Tollman<sup>1,2</sup>

<sup>1</sup>MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

<sup>2</sup>Umeå Centre for Global Health Research, Department of Public Health and Clinical Medicine, Umeå University, Sweden

### Background

Verbal autopsy (VA) has been used extensively for estimating cause-specific mortality at particular times, but less for characterizing long-term changes in epidemic patterns. Monitoring emerging causes of death over time raises questions about practitioners' developing perceptions of new disease patterns and demands consistent methods and practices. In this paper we retrospectively analyse the emergence of HIV-related mortality in a rural South African population from 1992 to 2005. VA data have been interpreted both by physicians and using the InterVA probabilistic model. The development of the HIV/AIDS mortality epidemic is described, together with a discussion of VA methods and variations between individual physician coders as the epidemic progressed.

### Methods

Between 1992 and 2005, 6,153 deaths were registered in a rural South African population (living in the Agincourt sub-district, under longitudinal health and socio-demographic surveillance). VAs were carried out for 94% of deaths, and coded independently by two physicians as well as being processed by the InterVA model. The physician causes of death were consolidated into a single consensus underlying cause for each case, with an additional physician arbitrating where different diagnoses persisted. HIV-related mortality rates were calculated using population denominators from the Agincourt site. Proportions of deaths coded as HIV-related by individual physicians, physician consensus and the InterVA model were compared over time.

### Results

Approximately 20% of deaths were rated as HIV-related over the whole time period, ranging from very low levels in the early years to population rates of 2.5 per 1,000 person-years latterly, a more than ten-fold increase. The age distribution was consistently bi-modal, with higher rates among children under 5 years, and among adults aged 20 to 64 years. Adult mortality shifted to older ages as the epidemic progressed, with a small but noticeable number of HIV-related deaths in the over-65 year age group latterly. InterVA results from the early years suggested that there may have been slightly higher early HIV-related mortality than recorded in physician consensus results, even though both rates were low compared with later years. Otherwise physician consensus and InterVA results characterised the epidemic in much the same way. Analysing rates of HIV-related deaths as coded by individual physicians showed a marked degree of inter-observer variation, with the physician consensus findings generally reflecting slightly lower proportions of HIV-related deaths than individual physicians. Aggregated findings for first versus second physician did not differ appreciably.

### Conclusions

VA was clearly able to detect and characterise the emergence of a very significant epidemic of HIV-related mortality in this population. Using either physician interpretation or the InterVA probabilistic model gave closely comparable findings on the overall development of the epidemic. The high degree of aggregate consistency between first and second physician coders suggests that double coding of VA data may not be worthwhile. However, consistency within and between individual physician coders, individual perceptions of changing epidemiology in a dynamic situation and the inherent consistency over time of probabilistic models are important considerations in these kinds of long-term analyses.

# Validating physician review and probabilistic modelling (InterVA) approaches to verbal autopsy interpretation using hospital causes of death

E. Bauni<sup>1</sup>, C. Ndila<sup>1</sup>, G. Mochamah<sup>1</sup>, G. Nyutu<sup>1</sup>, T. N. Williams<sup>1,2,3</sup>

<sup>1</sup> KEMRI-Wellcome Trust Research Programme , Kilifi Kenya

<sup>2</sup> Nuffield Department of Medicine, John Radcliffe Hospital, Oxford, United Kingdom

<sup>3</sup> Department of Paediatrics, John Radcliffe Hospital, Oxford, United Kingdom

## Background

The common method to determine cause of death is certification by trained physicians, based on available medical records, or through verbal autopsy (VA) (by physician review). However, physician review approach is costly and inconvenient because two physicians must review, code and agree on the cause of death. However, recent work shows the potential of a computer-based probabilistic model (InterVA) to interpret verbal autopsy data in a more convenient, consistent and rapid way. This study validated both physician review (PR) and the InterVA probabilistic model, using hospital cause of death (COD) as a gold standard.

## Methods

Between March 2007 and June 2010, VA interviews were conducted for 145 adult deaths that occurred in hospital. The VA data were reviewed by physicians and the cause of death established. Agreement between the two physicians was reasonable (67%). Various indicators (including age, gender, signs and symptoms, pregnancy status, medical history, circumstances) from the VA data were entered into the InterVA probabilistic model for interpretation. Cause-specific mortality fractions (CSMF), Cohen's kappa (k) statistic, receiver operating characteristic (ROC) curve, sensitivity, specificity, and positive predictive values were applied to compare agreement between the physician review (PR), InterVA and hospital COD.

## Results

Hospital-based diagnosis, InterVA model and Physician review yielded the same top five underlying causes of adult deaths. The probabilistic InterVA model overestimated TB cases as compared to the hospital deaths. Conversely, physicians overestimated diabetes as the cause of death. Overall multi-rater agreement across the three methods was 0.41. The areas under the receiver operator characteristic curve (ROC) were, 0.82 and 0.88 for InterVA model and Physician review respectively. The observed sensitivities and specificities across the five major cause of death varied from 43%-100% and 87%-99%, respectively, between InterVA model/physician and hospital based diagnosis.

## Conclusion

Both the probabilistic InterVA model and PR compared reasonably well with the hospital causes of death and determined accurately the top five underlying causes of death in the rural community of Kilifi.

# The Verbal Autopsy Method as an Important Tool in the Investigation of Ill-defined Conditions Declared as the Underlying Cause of Death in the State of Minas Gerais, Brazil

Elisabeth França<sup>1</sup>, Deise Campos<sup>2</sup>, Maria de Fátima Marinho de Souza<sup>3</sup>

<sup>1</sup>Graduate Program in Public Health, Federal University of Minas Gerais, Brazil

<sup>2</sup>Research Group in Epidemiology and Health Evaluation, Federal University of Minas Gerais, Brazil

<sup>3</sup>PAHO

## Background

The Mortality Information System (MIS) in Brazil records mortality data in hospitals and civil registries, with responsibility for compiling and analyzing cause of death data. Recently, the Ministry of Health has implemented a series of efforts focused on helping states to reduce ill-defined causes of death. Despite continuous improvements in the MIS, Minas Gerais maintains more than 10 per cent of deaths assigned to symptoms, signs, and ill-defined conditions, especially in areas with lower socioeconomic levels. Deaths coded to this category are in fact most likely miscoded deaths from communicable and noncommunicable diseases. More complete registration of deaths from injury is expected. In Brazil however, some local studies have provided evidence of underreporting of injury in small municipalities. The aim of this study was to investigate ill-defined causes of death using the verbal autopsy method to estimate injury specific mortality fractions in a poor region of Minas Gerais state, Brazil.

## Methods

Minas Gerais is in the Southeast region of Brazil, one of the most developed in the country, with an estimated total population in 2007 of about 19 million. Because of the geographically dispersed population, all 63 municipalities of the state were sorted by population size into three groups: a) municipalities with less than 10,000 inhabitants (49,2%); b) municipalities of 10,000 to 19,999 inhabitants (31,8%) and c) municipalities with 20,000 inhabitants or more (19,0%). A probability proportionate to population group size random sample of 10 municipalities was obtained, and then verbal autopsy interviews were conducted for all ill-defined deaths reported for the previous year to the State Health Department up to April 2008. Trained interviewers questioned relatives using a standardized verbal autopsy (VA) questionnaire to elicit information on symptoms experienced by the deceased before death. Important attempts have been made to collect all existing information about the disease or death using hospitals records, health department records, autopsy records, registry office records or by interviewing the Family Health Program professional. Three forms were applied: (i) for neonatal deaths, (ii) for deaths from 28 days old to less than 10 years old and, (iii) for people 10 years or older. Probable causes of death were assigned by physician review of the completed questionnaires and the underlying cause selected according to ICD-10 rules.

## Results

The random sample of municipalities yielded a total of 573 deaths for 202 of which the underlying cause was an ill-defined condition. Of these cases, 151 were investigated using verbal autopsy and 12.6% (n=19) had injury as the underlying cause of death. The proportional mortality fraction from injury among defined causes increased from 6.8% to 12.9% after investigation. Different sex distribution, age and specific injury category causes were observed between recorded injury causes and those detected by verbal autopsy. Falls and drowning were the top specific injury causes detected after investigation.

## Conclusions

This study provides evidence that approximately 13% of deaths certified as ill-defined causes in small counties in Minas Gerais, Brazil, were in fact due to injuries. The use of VA for diagnosis of causes of death among ill-defined conditions in resource poor communities can provide information on the relevance of injury as a priority health

problem in poor areas of developing countries. Verbal autopsy is an important tool in the investigation of ill-defined conditions and more validation studies in specific populations should be encouraged to resolve the uncertainty surrounding the true causes of these deaths. From another perspective, local research with VA should be brought to the attention of regional health policy makers to improve the quality of data for their planning.



## **The Verbal Autopsy success rate and factors associated with undetermined cause of death in poor resource settings in rural Tanzania**

Mathew Alexander Mwanyangala<sup>1</sup>, Honarath Urassa<sup>1</sup>, Jensen Charles<sup>1</sup>, Chrisotom Mahutanga<sup>1</sup>, Angelina Lutambi<sup>1,2</sup>, Honorati Masanja<sup>1,2</sup>, Salim Abdulla<sup>1</sup>, Rose Nathan<sup>1</sup>

<sup>1</sup>Ifakara Health Institute, Ifakara, Kilombero. Morogoro Tanzania

<sup>2</sup>Swiss Tropical Public Health School

### **Background**

Verbal Autopsy (VA) is a widely-used tool for ascertaining probable cause of death in areas with limited or incomplete vital registration systems. Its uses in priority setting, and health planning is well documented in sub-Saharan Africa (SSA) and Asia. Issues related to the VA instrument, including design, analysis and use in the health information system, are also widely discussed. However, discussion relating to completeness and success rate, from the stage of identification of death in the community to assigning cause of death, and factors associated with undetermined causes is rare in SSA. Such information is needed for better estimation of the burden of disease and understanding the limitations of VA.

### **Objective**

To determine the success rate and investigate factors associated with undetermined cause of deaths in rural Tanzania.

### **Methods**

A database of deaths reported from the Ifakara Health and Demographic Surveillance System (HDSS) from 2002 to 2007 was used. The success rates were determined at various stages, based on a) the number of deaths identified, b) VA interviews conducted, c) VA forms submitted to physicians, d) coding and e) assigning cause of death. Logistic regression was used to investigate factors associated with deaths that were coded as “undetermined”.

### **Results**

VA was successfully applied to 82 to 88 percent of deaths between 2002 and 2006. Ninety four percent of deaths submitted to physician review ended with a specific cause of death, and among those 27% ended with an undetermined cause. The undetermined cause of death is associated with the types of relationship the respondent had with the diseased, the place where the death occurred, and the level of education of the respondent. Furthermore, neonatal deaths are less likely to be associated with an undetermined cause if the respondent was the mother.

### **Conclusion**

Despite high success rates, there are still a number of deaths lost during the processing of VA especially between identification of death and VA interview. The reasons for this phenomenon need to be explored. Results suggest that the choice of the respondent has a substantial effect on assigning cause of death. The current study suggests that further work is needed to improve coding and minimize the steps and logistics for determining a specific cause of death, so that there are fewer undetermined causes leading to better disease burden estimates.