

Strengthening standardised interpretation of verbal autopsy data: the new InterVA-4 tool

Peter Byass^{1,2,3*#}, Daniel Chandramohan⁴, Samuel J. Clark⁵, Lucia D'Ambruoso^{1#}, Edward Fottrell^{1,6}, Wendy J. Graham³, Abraham J. Herbst^{7,8}, Abraham Hodgson⁹, Sennen Hounton¹⁰, Kathleen Kahn^{1,2}, Anand Krishnan¹¹, Jordana Leitao⁴, Frank Odhiambo¹², Osman A. Sankoh^{2,8} and Stephen M. Tollman^{1,2,8†}

¹Umeå Centre for Global Health Research, Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden; ²School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa; ³IMMPACT, The University of Aberdeen, Aberdeen, UK; ⁴London School of Hygiene and Tropical Medicine, Keppel Street, London, UK; ⁵Department of Sociology, University of Washington, Seattle, WA, USA; ⁶Centre for International Health and Development, Institute of Child Health, London, UK; ⁷Africa Centre for Health and Population Studies, University of KwaZulu-Natal, Somkhele, South Africa; ⁸INDEPTH Network, Kanda, Accra, Ghana; ⁹Ghana Health Service, Accra, Ghana; ¹⁰UNFPA, New York, USA; ¹¹Centre for Community Medicine, All India Institute of Medical Sciences, New Delhi, India; ¹²Kenya Medical Research Institute/Centers for Disease Control and Prevention (CDC) Research and Public Health Collaboration, Kisumu, Kenya

Background: Verbal autopsy (VA) is the only available approach for determining the cause of many deaths, where routine certification is not in place. Therefore, it is important to use standards and methods for VA that maximise efficiency, consistency and comparability. The World Health Organization (WHO) has led the development of the 2012 WHO VA instrument as a new standard, intended both as a research tool and for routine registration of deaths.

Objective: A new public-domain probabilistic model for interpreting VA data, InterVA-4, is described, which builds on previous versions and is aligned with the 2012 WHO VA instrument.

Design: The new model has been designed to use the VA input indicators defined in the 2012 WHO VA instrument and to deliver causes of death compatible with the *International Classification of Diseases* version 10 (ICD-10) categorised into 62 groups as defined in the 2012 WHO VA instrument. In addition, known shortcomings of previous InterVA models have been addressed in this revision, as well as integrating other work on maternal and perinatal deaths.

Results: The InterVA-4 model is presented here to facilitate its widespread use and to enable further field evaluation to take place. Results from a demonstration dataset from Agincourt, South Africa, show continuity of interpretation between InterVA-3 and InterVA-4, as well as differences reflecting specific issues addressed in the design and development of InterVA-4.

Conclusions: InterVA-4 is made freely available as a new standard model for interpreting VA data into causes of death. It can be used for determining cause of death both in research settings and for routine registration. Further validation opportunities will be explored. These developments in cause of death registration are likely to substantially increase the global coverage of cause-specific mortality data.

Keywords: *verbal autopsy; cause of death; vital registration; InterVA; World Health Organization*

To access the supplementary material to this article 'The InterVA-4 User Guide' please see Supplementary files under Article Tools online.

† Authors are listed in alphabetical order.

The Deputy Editor Peter Byass and Editor Lucia D'Ambruoso have not participated in the review and decision process for this paper.

Verbal autopsy (VA) covers the entire process of interviewing close caregivers – relatives, friends or witnesses – about the details of a death that occurred and then uses the interview data to arrive at a probable cause of death. This is not a new concept, and although VA might not necessarily be the best or most accurate approach for determining individual causes of death, it is nevertheless a much better option than simply allowing deaths to pass unrecorded. Provided VA is implemented with reasonable rigour and consistency, it leads to information on cause-specific mortality patterns which are otherwise unavailable. The need for more comprehensive death registration, and the potential contribution of VA in this, is described in a recent blog (1).

In recent years increasing attention has been paid to computerised procedures for determining cause of death from VA data (2). VA data have commonly been interpreted by physicians, but this has proved to be a costly, slow and non-reproducible process in many situations, yielding VA cause of death information that cannot readily be compared between settings (3). Proponents of physician interpretation argue that a more nuanced approach to detailed causes of death (including co-morbidities) can be achieved by physicians considering a narrative account of the circumstances leading to death (4), while computerised processes can be comparatively fast, cheap and reproducible over time and place, and also speed up VA interviews by obviating the need for transcribing lengthy narratives (5).

Over the past decade, a series of InterVA models have been developed using Bayesian probabilistic modelling as a means of interpreting VA data to derive causes of death. Starting on a very experimental basis with VA data from Vietnam (6), these models progressed to InterVA-3 which has been widely used in a variety of settings across Africa, Asia and Latin America (7–15) and an associated model, InterVA-M, which dealt separately with deaths among women of reproductive age (16, 17). Additional developments for handling neonatal deaths have also been reported previously (18).

In parallel with these technical developments for ascertaining cause of death, concern has grown about the large proportion of deaths in the world that are not certified by cause and which therefore contribute nothing to global evidence on cause of death and the implications for public health. It is also clear that the chances of a death being registered are strongly determined by geographic location and socio-economic status (19), meaning that widening the scope of death registration (including cause determination) also represents a means of reducing global bias in cause of death information. This has led the World Health Organization (WHO), the Health Metrics Network and the United Nations Statistical Commission

to seek ways forward for extending routine death registration on a much wider scale. Hitherto VA has been largely used in research settings, such as the Health and Demographic Surveillance Sites of the INDEPTH Network (20). Now, particularly with the development of shorter and more automated methods, VA offers new opportunities for wider implementations of routine cause of death registration, not only in research environments.

This paper presents the underlying principles of the latest model in the InterVA suite, InterVA-4, which integrates experience accumulated from previous versions, latest data and research findings, and revisions by an expert panel. The InterVA-4 model is freely available and can be downloaded from www.interva.net. The *InterVA-4 User Guide* is attached to this paper as Supplementary Material. As well as reflecting previous experience, this latest model has been constructed to reflect the structure of the 2012 WHO Verbal Autopsy instrument, developed in an expert consultation in Geneva in December 2011 (21). InterVA-4 aims to provide a consistent and generally applicable means of interpreting VA data, being modelled on (but not restricted to) the 2012 WHO Verbal Autopsy instrument, and hence applicable prospectively and retrospectively. The model is intended for use both within already-enumerated populations and as a stand-alone death registration tool, both in research and in civil registration.

Probabilistic methods

Bayes' theorem (22) links the probability of an event happening given a particular circumstance with the unconditional probability of the same event and the conditional probability of the circumstance given the event. If the event of interest is a particular cause of death, and the circumstance is part of the events leading to death, then Bayes' theorem can be applied in terms of circumstances and causes of death.

Specifically, if there are a predetermined set of possible causes of death $C_1 \dots C_m$ and another set of indicators $I_1 \dots I_n$ representing various signs, symptoms and circumstances leading to death, then Bayes' general theorem for any particular C_i and I_j can be stated as:

$$P(C_i|I_j) = \frac{P(I_j|C_i) \times P(C_i)}{P(I_j|C_i) \times P(C_i) + P(I_j|!C_i) \times P(!C_i)}$$

where $P(!C_i)$ is $(1 - P(C_i))$.

Over the whole set of causes of death $C_1 \dots C_m$ a set of probabilities for each C_i can be calculated using a normalising assumption so that the total conditional probability over all causes totals unity:

$$P(C_i|I_j) = \frac{P(I_j|C_i) \times P(C_i)}{\sum_{i=1}^m P(C_i)}$$

Using an initial set of unconditional probabilities for causes of death $C_1 \dots C_m$ (which can be thought of as $P(C_i|I_0)$) and a matrix of conditional probabilities $P(I_j|C_i)$ for indicators $I_1 \dots I_n$ and causes $C_1 \dots C_m$, it is possible to repeatedly apply the same calculation process for each $I_1 \dots I_n$ that applies to a particular death:

$$P(C_i|I_1 \dots I_n) = \frac{P(I_j|C_i) \times P(C_i|I_0 \dots I_{n-1})}{\sum_{i=1}^m P(C_i|I_0 \dots I_{n-1})}$$

This process typically results in the probabilities of most causes reducing, while a few likely causes are characterised by their increasing probabilities as successive indicators are processed.

The InterVA-4 model

InterVA-4 aims to be a tool for interpreting VA data which can be applied simply, quickly and cheaply to generate cause of death data, compatible with the *International Classification of Diseases* version 10 (ICD-10). It offers the consistency and reproducibility that is characteristic of mathematical models and thus facilitates comparisons of VA cause of death between different places and over time. Furthermore, as a public-domain resource it is freely available and runs on any Windows®-based personal computer, or under the Mac OS® as a virtual Windows® instance. It builds on experience with previous InterVA models but brings a new standard of VA interpretation by conforming to the 2012 WHO VA instrument in terms of cause of death categories and input indicators. Indicators relating to durations in the WHO instrument are split into pre-defined categories for processing in InterVA-4. Tables of cause of death categories and input indicators from the 2012 WHO VA instrument as handled by InterVA-4 are included in the *InterVA-4 User Guide* (Supplementary Material).

For deaths among women of reproductive age, a separate sub-model is incorporated which uses exactly the same methodology over three pregnancy status outcomes to assess the likelihood of pregnancy being associated with the death (but not necessarily in a causal sense). This was previously implemented in InterVA-M and is important in cases where pregnancy status at death may be ambiguous in the VA data (16).

Apart from the mathematics, the major challenge in building a probabilistic model covering all causes of death to a reasonable level of detail lies in populating the matrix of conditional probabilities $P(I_j|C_i)$. There is no overall source of data available which systematically quantifies probabilities of various signs, symptoms and circumstances leading to death in terms of their associations with particular causes. Therefore, these conditional probabilities have to be estimated from a diversity of incomplete sources (including previous InterVA models) and modulated by expert opinion. In the various versions of InterVA that have been developed, expert panels have

been convened to capture clinical expertise on the relationships between indicators and causes (23). In this case, an expert panel (DC, AH, SH, AK, KK) convened in Geneva in December 2011 and continued to deliberate subsequently, particularly considering issues that built on previous InterVA versions. Experience has shown that gradations in levels of perceived probabilities correspond more to a logarithmic than linear scale (24), and in the expert consultation for InterVA-4, we used a perceived probability scale that was subsequently converted to numbers on a logarithmic scale as shown in Table 1. This was similar to scales used previously but was extended into additional low probability categories. We have also demonstrated by means of sensitivity analyses within the probability matrix that although the conditional probability values are important, the model does not require a high level of precision for these estimates (25).

As was the case in InterVA-3 (5), there are special arrangements for two causes of death (HIV/AIDS and malaria), the occurrences of which vary appreciably in different locations. As described in the *InterVA-4 User Guide*, the model has a facility for setting the unconditional probabilities for each of these causes to reflect their occurrence in a particular population. Conceptually, this is analogous to a physician's background knowledge of local disease profiles, irrespective of the details of a particular case. In the new model, the setting for malaria is also applied to the unconditional probability for deaths due to sickle cell disease, because of the genetically determined geographical overlap between the two diseases (26).

Having experienced difficulties in previous versions of the software with proprietary data file formats, which tend to change in unpredictable ways over time, InterVA-4

Table 1. Qualitative probability scale used as the basis for eliciting expert opinion on probabilities

| Label | Value | Interpretation |
|-------|---------|----------------|
| I | 1.0 | Always |
| A+ | 0.8 | Almost always |
| A | 0.5 | Common |
| A– | 0.2 | |
| B+ | 0.1 | Often |
| B | 0.05 | |
| B– | 0.02 | |
| C+ | 0.01 | Unusual |
| C | 0.005 | |
| C– | 0.002 | |
| D+ | 0.001 | Rare |
| D | 0.0005 | |
| D– | 0.0001 | |
| E | 0.00001 | Hardly ever |
| N | 0 | Never |

has reverted to text with comma-separated values (CSVs) for data input and output, as the lowest common denominator in file formats. CSV data can be relatively easily read and written by most analysis software packages, and examples of Stata scripts for generating InterVA-4 input and processing output from InterVA-4 in a standard way, deriving cause-specific mortality fractions (CSMFs), are included in the user guide.

Having calculated a set of conditional probabilities $P(C_i|I_1 \dots I_n)$ for causes $C_1 \dots C_m$ for each case, methods have to be applied for interpreting the output in ways that make sense from medical cause of death and public health perspectives. In the interests of transparency, InterVA-4 automatically outputs the complete $P(C_i|I_1 \dots I_n)$ data grid for all causes and cases in a batch as a CSV file. However, this is not a very convenient format for routinely analysing cause of death data; hence, the InterVA-4 software also undertakes basic post-processing of the basic cause of death data into a more readily useable format. Although experience shows that the majority of cases arrive at a single overwhelmingly likely cause, some cases, particularly if there is scanty or somewhat contradictory information available, may end up with two or three causes of comparable likelihood. This is analogous to a physician's differential diagnosis in which ambiguity remains between multiple likely causes. In a further minority of cases, there is no high likelihood cause, amounting to no clear conclusion on cause of death. To handle all this in a standardised way, InterVA-4 post-processes likely cause data in the same way that we have previously discussed in relation to InterVA-3 (5). If no single cause has a final likelihood of at least 0.4, the cause of death is considered to be indeterminate. The 0.4 cut-off is considerably higher than the unconditional probability of any cause, and conceptually includes a level approaching a 50:50 possibility for a particular cause, while leaving scope for other likely causes to be included in the overall consideration of the case. Multiple causes are only reported if they reach half of the likelihood of the leading cause. Any residual margin of likelihood not accounted for by the likelihood of the first and, where applicable, second and third causes can then be considered as a partial indeterminate component in analysing overall cause of death and CSMF patterns. We suggest that this is a much more appropriate approach than aggregating the sums of small residual probabilities of unlikely causes, which can lead to misleading results over large numbers of cases.

Since an important part of the objectives behind this model is to extend the use of VA into routine death registration procedures, additional non-medical questions about circumstances of death have been included in the 2012 WHO VA instrument to facilitate the interpretation of cause of death in non-enumerated populations. Although these are not in any way intended to replace the developing role of social autopsy (27), future devel-

opments in describing and modelling non-medical factors associated with cause of death are anticipated and will be reflected in future developments of InterVA.

Comparing InterVA-3 and InterVA-4

As with any software tool, one of our aims in developing this new version of InterVA was to improve the scope of the tool and address known shortcomings of the previous version. The differences between InterVA-3 and InterVA-4 are considerable, even though the same basic mathematical model lies at the heart of both versions. Both the range of causes of death and the scope of input indicators are considerably expanded in the new version and brought into line with the 2012 WHO VA instrument, but this makes a direct and meaningful comparison between the versions difficult. To demonstrate the similarities and differences, we have taken a random sample of 1,000 cases from the Agincourt, South Africa, dataset previously analysed (5). These cases have been run with the InterVA-3 and InterVA-4 models, but without adding data for the additional InterVA-4 indicators. Age-specific CSMFs from both models for these same 1,000 cases are shown in Table 2. This demonstrates the more detailed classification of neoplasms in the new model, as well as the detailed maternal causes brought in from InterVA-M. This illustrative dataset included 44/62 of the 2012 WHO VA instrument possible causes of death.

Discussion

The new InterVA-4 model presented here represents a substantial advance in automated cause of death modelling. It is specifically aligned to the new WHO 2012 Verbal Autopsy instrument (21), although it can be used retrospectively for processing archived VA data, provided a reasonable spread of the necessary indicators are available. It offers the advantages of standardised interpretation over time and place; is freely available as a public-domain resource; and on a typical computer can process around 10,000 cases per hour. These attributes mean that it offers new opportunities for cause of death data to be gathered routinely in settings where hitherto there have been few or no such records and where physicians to assess VAs are few. The speed of processing can yield cause of death information with the potential to feed into local and national health policy development in a timely fashion. Standardised assessment means that trends in causes of death can be tracked over time and can be compared across different settings. We have previously shown that using a standard model across a wide range of settings yields appreciably different, and plausible, cause of death profiles (28).

The potential disadvantage of using any mathematical model for interpreting cause of death is that some subtlety and nuance may be lost as compared to routine

Table 2. Comparison of age- and cause-specific mortality fractions (CSMF) using the InterVA-3 and InterVA-4 models on 1,000 randomly selected VA records from a previous analysis of data in Agincourt HDSS, South Africa (12)

| InterVA-4 cause (WHO 2012 categories*) | Cause-specific mortality fractions by age group (%) | | | | | | | | | | | | | | | | InterVA-3 cause |
|--|---|----|---------------------|----|-----------------------|----|---------------------|----|----------------------|----|------------------------|----|------------------------|----|----------------------|----|------------------------------------|
| | All-age n = 1000 | | ≤ 28 days n = 15 | | 1–11 months n = 50 | | 1–4 years n = 79 | | 5–14 years n = 35 | | 15–49 years n = 437 | | 50–64 years n = 139 | | 65+ years n = 245 | | |
| | v4 | v3 | v4 | v3 | v4 | v3 | v4 | v3 | v4 | v3 | v4 | v3 | v4 | v3 | v4 | v3 | |
| 01.02 ARI, including pneumonia | 5 | 4 | | | 35 | 25 | 11 | 6 | 3 | | 1 | 1 | 5 | 5 | 4 | 6 | <i>Pneumonia/sepsis</i> |
| 01.03 HIV/AIDS related death | 19 | 17 | | | 16 | 22 | 39 | 49 | 15 | 6 | 25 | 25 | 16 | 8 | 4 | 0 | <i>HIV/AIDS-related death</i> |
| 01.04 diarrhoeal diseases | 3 | 2 | | | 10 | 21 | 13 | 5 | | 4 | 3 | 0 | 2 | | 1 | 1 | <i>Diarrhoea</i> |
| 01.05 malaria | 1 | 0 | | | 5 | | 4 | | 4 | | 1 | 0 | 2 | | 0 | | <i>Malaria</i> |
| 01.07 meningitis | 0 | 2 | | | 2 | 5 | 1 | 3 | | | 15 | | 1 | | | 2 | <i>Meningitis</i> |
| 01.08 tetanus | | 0 | | | | | | | | | | | | | | 1 | <i>Tetanus</i> |
| 01.09 pulmonary tuberculosis | 19 | 20 | | | 1 | | 5 | 4 | 10 | 5 | 25 | 26 | 21 | 20 | 19 | 22 | <i>Pulmonary tuberculosis</i> |
| 01.10 pertussis | 0 | | | | 2 | | 0 | | | | | | | | | | |
| 01.99 other infectious diseases | 3 | | | | 2 | | | | 8 | | 3 | | 4 | | 4 | | |
| 02.01 oral neoplasms | 0 | 1 | | | | | | | | | | 0 | | 2 | 0 | 3 | <i>Malignancy</i> |
| 02.02 digestive neoplasms | 4 | | | | | | | | | | 2 | | 9 | | 8 | | |
| 02.03 respiratory neoplasms | 3 | | | | | | | | | | 2 | | 4 | | 5 | | |
| 02.05, 02.06 reproductive neoplasms | 2 | | | | | | | | | | 1 | | 0 | | 5 | | |
| 02.99 other neoplasms | 1 | | | | | | | | | | 1 | | 0 | | 2 | | |
| 03.02 severe malnutrition | 0 | 0 | | | | | | 3 | | | 0 | | | | 0 | 1 | <i>Malnutrition</i> |
| 03.03 diabetes mellitus | 2 | 3 | | | | | 2 | 1 | 5 | | 1 | 1 | 2 | 4 | 5 | 5 | <i>Diabetes</i> |
| 04.01 acute cardiac disease | 1 | 0 | | | | | | | | | 1 | 0 | 1 | 1 | 2 | 0 | <i>Acute cardiac death</i> |
| 04.02 stroke | 1 | | | | | | | | | | 1 | 0 | 2 | 1 | 1 | 2 | <i>Stroke</i> |
| 04.03 sickle cell with crisis | | 1 | | | | | | 1 | | | | | | | | | <i>Haemoglobinopathy</i> |
| 04.99 other cardiac disease | 2 | 4 | | | | | | | 3 | 4 | 1 | 2 | 2 | 8 | 4 | 9 | <i>Chronic cardiac death</i> |
| 05.01 COPD | 1 | 2 | | | | | | | | | 0 | 1 | 1 | 5 | 3 | 3 | <i>Chronic respiratory disease</i> |
| 05.02 asthma | 1 | | | | | | | | 3 | | 1 | | | | 1 | | |
| 06.01 acute abdomen | 2 | | | | 1 | | 2 | | | | 2 | | 2 | | 4 | | |
| 06.02 liver cirrhosis | 0 | 4 | | | | 1 | | | | | 0 | 2 | 1 | 6 | 1 | 8 | <i>Liver disease</i> |
| 07.01 renal failure | 1 | 1 | | | 2 | | | | | | 1 | 0 | 2 | 2 | 1 | 4 | <i>Kidney disease</i> |
| 08.01 epilepsy | 1 | 0 | | | | | 1 | | 4 | | 1 | 0 | | | 1 | | <i>Disease of nervous system</i> |
| 98 other NCD | 0 | 0 | | | | | | | | | 0 | | | 0 | 0 | 0 | |
| 09.01 ectopic pregnancy | 0 | 0 | | | | | | | | | 0 | 1 | | | | | <i>Maternal-related death</i> |
| 09.02 abortion-related death | 0 | | | | | | | | | | 0 | | | | | | |
| 09.06 pregnancy-related sepsis | 0 | | | | | | | | | | 0 | | | | | | |
| 09.07 anaemia of pregnancy | 0 | | | | | | | | | | 0 | | | | | | |

Table 2 (Continued)

| InterVA-4 cause (WHO 2012 categories*) | Cause-specific mortality fractions by age group (%) | | | | | | | | | | | | | | | | InterVA-3 cause |
|--|---|-----|---------------------|-----|-----------------------|-----|---------------------|-----|----------------------|-----|------------------------|-----|------------------------|-----|----------------------|-----|-----------------------------------|
| | All-age n = 1000 | | ≤ 28 days n = 15 | | 1–11 months n = 50 | | 1–4 years n = 79 | | 5–14 years n = 35 | | 15–49 years n = 437 | | 50–64 years n = 139 | | 65+ years n = 245 | | |
| | v4 | v3 | v4 | v3 | v4 | v3 | v4 | v3 | v4 | v3 | v4 | v3 | v4 | v3 | v4 | v3 | |
| 10.01 prematurity | 0 | 0 | 24 | 13 | | 2 | | | | | | | | | | | <i>Pre-term/small baby</i> |
| 10.02 birth asphyxia | 0 | 1 | 19 | 43 | | | | | | | | | | | | | <i>Perinatal asphyxia</i> |
| 10.03 neonatal pneumonia | 1 | 0 | 40 | 26 | | | | | | | | | | | | | <i>Pneumonia/sepsis</i> |
| 10.06 congenital malformation | 0 | | | | 4 | | | | | | | | | | | | |
| 10.99 other neonatal conditions | 0 | | 5 | | | | | | | | | | | | | | |
| 12.01 road traffic accident | 3 | 4 | | | | 0 | 1 | 4 | 6 | 10 | 5 | 6 | 2 | 2 | 1 | 1 | <i>Transport-related accident</i> |
| 12.03 accidental fall | 0 | | | | | | 1 | | 3 | | | | | | | | |
| 12.04 accidental drowning | 0 | 0 | | | | | 1 | 1 | 6 | 5 | | | | | | | <i>Accidental drowning</i> |
| 12.05 smoke, fire and flames | 0 | | | | | | 1 | | | | | | | | | 1 | |
| 12.07 accidental poisoning | 1 | 1 | | | | | 1 | 3 | | | 0 | 1 | | | 1 | 0 | <i>Accidental poisoning</i> |
| 12.08 intentional self-harm | 1 | 1 | | | | | | | 3 | 3 | 2 | 1 | 1 | | 1 | 1 | <i>Suicide</i> |
| 12.09 assault | 4 | 3 | | | 2 | | | | | | 7 | 6 | 4 | 4 | 0 | 0 | <i>Homicide</i> |
| 12.99 other external causes | 1 | 0 | | | | 1 | 3 | 1 | 4 | 1 | 1 | 0 | 2 | | 2 | 0 | <i>Other fatal accident</i> |
| 99 indeterminate | 17 | 29 | 12 | 18 | 18 | 23 | 11 | 22 | 23 | 36 | 12 | 26 | 15 | 31 | 19 | 31 | <i>Indeterminate</i> |
| Total over all causes | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | |

*These demonstration data included 44/62 of the cause of death categories in the 2012 WHO VA instrument and InterVA-4. CSMFs are rounded to nearest 1%, 0 representing a finite value <0.5%.

physician certification of cause of death (either from attending physicians or from physicians coding individual VA cases). It is certainly true that mathematical models will fail to identify very unusual or complex causes of death. However, such cases tend to be of more local or esoteric interest rather than contributing significantly to the public health understanding of population mortality patterns.

Considering the validity of any VA approach raises difficult issues. VA is normally used in settings where other sources of information on circumstances of death are lacking, precluding many direct comparisons. The VA interview stage carries inherent uncertainty as the quality of information obtained depends on a variety of factors including relationship of respondent to deceased, knowledge of signs and symptoms during terminal illness, recall and willingness to disclose information, especially for conditions that are stigmatised or have culturally sensitive connotations (29). The absolute validity of deaths certified by attending physicians is by no means a given, and studies have shown considerable inconsistencies between hospital causes of death and pathologists' findings (30, 31). The validity of physician interpretation of VA material has also been shown to be questionable (3) and subject to inter-observer variation (12). It is also clear that any approach to VA data must involve some degree of capturing expert opinion, alongside the use of established information (32). Previous InterVA models have been extensively evaluated against local physician interpretations, with generally concordant findings (5, 7, 10, 12, 14, 15). One evaluation suggested that InterVA-3 did not perform well against VA data from tertiary health facilities (33), but those data were not available for further study. Nevertheless, certain shortcomings have been identified in InterVA-3, including over-diagnosis of meningitis, under-diagnosis and lack of differentiation between various cancers, and fairly high levels of indeterminate cases. InterVA-4 is designed to handle these more effectively, and we will continue to explore realistic opportunities for further validations.

The comparison between InterVA-3 and InterVA-4 demonstrated in Table 2 says nothing about the absolute validity of either model or about the mortality profile in the Agincourt population. It does demonstrate, however, the general continuity in results between the two versions of the model, when using exactly the same input data, as well as evidence of deliberately introduced changes in response to some of the shortcomings identified in InterVA-3. This comparison does not necessarily make the most of some of the new aspects of InterVA-4, since data for new indicators were not included in order to demonstrate continuity. We will continue to make InterVA-3 and InterVA-M available as legacy downloads, but we recommend the use of InterVA-4 wherever possible.

InterVA-4 is launched here as a global resource. It offers substantial benefits over the default situation of not routinely capturing cause of death data (which is the *de facto* standard in many resource-limited parts of the world), considerable improvements over previous versions and the ability to more reliably compare causes of death across different settings and over time. This is important for research, for example, as INDEPTH cross-site research accelerates (20), and also offers new opportunities for cause-specific civil registration of deaths on a much wider scale. Work is in progress on a portable electronic version of the model which will be particularly useful for routine registration. In the longer term, there will undoubtedly be further evaluations and developments as the science and methods of attributing cause of death improve. We will continue to work with old and new collaborators on further refinements in this process.

Acknowledgements

The core development work for InterVA-4 was undertaken within The Umeå Centre for Global Health Research which is supported from FAS, the Swedish Council for Working Life and Social Research (www.fas.se) (grant no. 2006-1512). The Agincourt health and socio-demographic surveillance system, including verbal autopsies, has been funded by The Wellcome Trust, UK (grants 058893/Z/99/A; 069683/Z/02/Z; 085477/Z/08/Z) with contributions from the Medical Research Council and University of the Witwatersrand, South Africa, and the William and Flora Hewlett Foundation, National Institute on Aging (NIA) of the NIH, and Andrew W Mellon Foundation, USA. Contributions to this work of collaborators at the World Health Organization, the INDEPTH Network, University College London, the University of Aberdeen and other institutions are appreciated. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest and funding

The authors have not received any funding or benefits from industry or elsewhere to conduct this study.

References

1. Byass P. A new era dawns for death registration. *PLoS Speaking of Medicine*; 22 August 2012. Available from: <http://blogs.plos.org/speakingofmedicine/2012/08/22/a-new-era-dawns-for-death-registration/> [cited 23 August 2012].
2. Fottrell E, Byass P. Verbal Autopsy – methods in transition. *Epidemiol Rev* 2010; 32: 38–55.
3. Lozano R, Lopez AD, Atkinson C, Naghavi M, Flaxman AD, Murray CJ; Population Health Metrics Research Consortium (PHMRC). Performance of physician-certified verbal autopsies: multisite validation study using clinical diagnostic gold standards. *Popul Health Metr* 2011; 9: 32.
4. Montgomery AL, Morris SK, Bassani DG, Kumar R, Jotkar R, Jha P. Factors associated with physician agreement and coding choices of cause of death using verbal autopsies for 1130 maternal deaths in India. *PLoS One* 2012; 7: e33075.
5. Byass P, Kahn K, Fottrell E, Collinson MA, Tollman SM. Moving from data on deaths to public health policy in

- Agincourt, South Africa: approaches to analysing and understanding verbal autopsy findings. *PLoS Med* 2010; 7: e1000325.
6. Byass P, Huong DL, Minh HV. A probabilistic approach to interpreting verbal autopsies: methodology and preliminary validation in Vietnam. *Scand J Public Health* 2003; 31(Suppl 62): 32–7.
 7. Fantahun M, Fottrell E, Berhane Y, Wall S, Högberg U, Byass P. Assessing a new approach to verbal autopsy interpretation in a rural Ethiopian community: the InterVA model. *B World Health Organ* 2006; 84: 204–10.
 8. WHO (2009). Global Malaria Programme. Deployment at community level of artemether-lumefantrine and rapid diagnostic tests, Raya Valley, Tigray, Ethiopia. Geneva: World Health Organization.
 9. Tensou B, Araya T, Telake DS, Byass P, Berhane Y, Kebebew T, et al. Evaluating the InterVA model for determining AIDS mortality from verbal autopsies in the adult population of Addis Ababa. *Trop Med Int Health* 2010; 15: 547–53.
 10. Bauni E, Ndila C, Mochamah G, Nyutu G, Matata L, Ondieki C, et al. Validating physician-certified verbal autopsy and probabilistic modeling (InterVA) approaches to verbal autopsy interpretation using hospital causes of adult deaths. *Popul Health Metr* 2011; 9: 49.
 11. Herbst AJ, Mafojane T, Newell M-L. Verbal autopsy-based cause-specific mortality trends in rural KwaZulu-Natal, South Africa, 2000–2009. *Popul Health Metr* 2011; 9: 47.
 12. Byass P, Kahn K, Fottrell E, Mee P, Collinson MA, Tollman SM. Using verbal autopsy to track epidemic dynamics: the case of HIV-related mortality in South Africa. *Popul Health Metr* 2011; 9: 46.
 13. Ferri CP, Acosta D, Guerra M, Huang Y, Llibre-Rodriguez JJ, Salas A, et al. Socioeconomic factors and all cause and cause-specific mortality among older people in Latin America, India and China: a population-based cohort study. *PLoS Med* 2012; 9: e1001179.
 14. Ramroth H, Lorenz E, Rankin JC, Fottrell E, Yé M, Neuhann F, et al. Cause of death distribution with InterVA and physician coding in a rural area of Burkina Faso. *Trop Med Int Health* 2012; 17: 904–13.
 15. Oti SO, Kyobutungi C. Verbal autopsy interpretation: a comparative analysis of the InterVA model versus physician review in determining causes of death in the Nairobi DSS. *Popul Health Metr* 2010; 8: 21.
 16. Fottrell E, Byass P, Ouedraogo T, Tamini C, Gbangou A, Sombie I, et al. Revealing the burden of maternal mortality: a probabilistic model for determining pregnancy-related causes of death from verbal autopsies. *Popul Health Metr* 2007; 5: 1.
 17. Bell JS, Ouedraogo M, Ganaba R, Sombie I, Byass P, Baggaley RF, et al. The epidemiology of pregnancy outcomes in rural Burkina Faso. *Trop Med Int Health* 2008; 13(Suppl 1): 31–43.
 18. Vergnano S, Fottrell E, Osrin D, Kazembe PN, Mwansambo C, Manandhar DS, et al. Adaptation of a probabilistic method (InterVA) of verbal autopsy to improve the interpretation of cause of stillbirth and neonatal death in Malawi, Nepal, and Zimbabwe. *Popul Health Metr* 2011; 9: 48.
 19. Byass P. The unequal world of health data. *PLoS Med* 2009; 6: e1000155.
 20. Sankoh O, Byass P. The INDEPTH Network: filling some international gaps in epidemiology. *Int J Epidemiol* 2012; 41: 579–88.
 21. WHO (2012). Verbal autopsy standards: the 2012 WHO verbal autopsy instrument. Geneva: World Health Organization. Available from: http://www.who.int/healthinfo/statistics/WHO_VA_2012_RC1_Instrument.pdf [cited 23 August 2012].
 22. Bayes T. A problem in the doctrine of chances. *Philos Trans R Soc Lond* 1763; 53: 370–418.
 23. Byass P, Fottrell E, Huong DL, Berhane Y, Corrah T, Kahn K, et al. Refining a probabilistic model for interpreting verbal autopsy data. *Scan J Public Health* 2006; 34: 26–31.
 24. Kong A, Barnett GO, Mosteller F, Youtz C. How medical professionals evaluate expressions of probability. *New Engl J Med* 1986; 315: 740–4.
 25. Fottrell E, Kahn K, Tollman S, Byass P. Probabilistic methods for verbal autopsy interpretation: InterVA robustness in relation to variations in *A Priori* probabilities. *PLoS ONE* 2011; 6: e27200.
 26. Grosse SD, Odame I, Atrash HK, Amendah DD, Piel FB, Williams TN. Sickle cell disease in Africa: a neglected cause of early childhood mortality. *Am J Prev Med* 2011; 41(Suppl 4): 398–405.
 27. Waiswa P, Kalter HD, Jakob R, Black RE. Increased use of social autopsy is needed to improve maternal, neonatal and child health programmes in low-income countries. *Bull World Health Organ* 2012; 90: 403.
 28. Fottrell E, Kahn K, Ng N, Sartorius B, Huong DL, Minh HV, et al. Mortality measurement in transition: proof of principle for standardised multi-country comparisons. *Trop Med Int Health* 2010; 15: 1256–65.
 29. Byass P, D'Ambruoso L, Ouedraogo M, Qomariyah SN. Assessing the repeatability of verbal autopsy for determining cause of death: two case studies among women of reproductive age in Burkina Faso and Indonesia. *Popul Health Metr* 2009; 7: 6.
 30. Cox JA, Lukande RL, Lucas S, Nelson AM, Van Marck E, Colebunders R. Autopsy causes of death in HIV-positive individuals in sub-Saharan Africa and correlation with clinical diagnoses. *AIDS Rev* 2010; 12: 183–94.
 31. Shojania K, Burton E, McDonald K, Goldman L (2002). The autopsy as an outcome and performance measure. Evidence report/technology assessment No. 58. Rockville, MD: Agency for Healthcare Research and Quality.
 32. Byass P. Whither verbal autopsy? *Popul Health Metr* 2011; 9: 23.
 33. Lozano R, Freeman MK, James SL, Campbell B, Lopez AD, Flaxman AD, et al. Performance of InterVA for assigning causes of death to verbal autopsies: multisite validation study using clinical diagnostic gold standards. *Popul Health Metr* 2011; 9: 50.

***Peter Byass**

Umeå Centre for Global Health Research
 Department of Public Health and Clinical Medicine
 Umeå University
 SE-90185 Umeå, Sweden
 Email: peter.byass@epiph.umu.se