Tropical Medicine and International Health doi:10.1111/j.1365-3156.2006.01603.x

VOLUME 11 NO 5 PP 681–696 MAY 2006

# Validity of verbal autopsy procedures for determining cause of death in Tanzania

**Philip W. Setel[1], David R. Whiting[2], Yusuf Hemed[3], Daniel Chandramohan[4], Lara J. Wolfson[5], K. G. M. M. Alberti[6] and Alan D. Lopez[7]**

1 *Department of Epidemiology and MEASURE Evaluation, Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA*
2 *University of Newcastle upon Tyne School of Population and Health Sciences; Faculty of Medicine; Newcastle upon Tyne, UK*
3 *SAVVY Regional Coordinator, MEASURE Evaluation and Tanzanian Ministry of Health, Dar es Salaam, Tanzania*
4 *Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, UK*
5 *Department of Immunization, Vaccines, and Biologicals, World Health Organization, Geneva, Switzerland*
6 *Department of Endocrinology and Metabolism, Imperial College, London, UK*
7 *School of Population Health, University of Queensland, Herston, Australia*

**Summary**

OBJECTIVES To validate verbal autopsy (VA) procedures for use in sample vital registration. Verbal autopsy is an important method for deriving cause-specific mortality estimates where disease burdens are greatest and routine cause-specific mortality data do not exist.

METHODS Verbal autopsies and medical records (MR) were collected for 3123 deaths in the perinatal/neonatal period, post-neonatal <5 age group, and for ages of 5 years and over in Tanzania. Causes of death were assigned by physician panels using the International Classification of Disease, revision 10. Validity was measured by: cause-specific mortality fractions (CSMF); sensitivity; specificity and positive predictive value. Medical record diagnoses were scored for degree of uncertainty, and sensitivity and specificity adjusted. Criteria for evaluating VA performance in generating true proportional mortality were applied.

RESULTS Verbal autopsy produced accurate CSMFs for nine causes in different age groups: birth asphyxia; intrauterine complications; pneumonia; HIV/AIDS; malaria (adults); tuberculosis; cerebrovascular diseases; injuries and direct maternal causes. Results for 20 other causes approached the threshold for good performance.

CONCLUSIONS Verbal autopsy reliably estimated CSMFs for diseases of public health importance in all age groups. Further validation is needed to assess reasons for lack of positive results for some conditions.

**keywords** cause of death, autopsy/methods, epidemiological measurements, epidemiological methods, vital statistics, child mortality

## Background

Many countries with the highest burdens of poverty and disease continue to lack routine, representative and high quality information on the levels and causes of death (Mathers *et al.* 2005). Mortality surveillance systems using validated verbal autopsy (VA) procedures and focused demographic surveillance in a nationally representative sample of district clusters represent a cost-effective and sustainable medium-term solution to this problem (Setel *et al.* 2005). VA-derived mortality estimates, usually from studies conducted in small and non-representative samples, are growing in importance as a basis for setting and evaluating international health priorities, policies and interventions. Recent meta-analyses of VA-based data to estimate the burden of diarrhoea, malaria and acute respiratory infection among children (Williams *et al.* 2002; Korenromp *et al.* 2003; Morris *et al.* 2003) and the evaluation of the integrated management of childhood illnesses (IMCI) reflect this growing influence (Bryce *et al.* 2004). To ensure internationally comparable data of known quality, standard VA procedures should be promoted and rigorously validated.

The majority of VA validation studies published since 1992 focus on neonates and children under 5 (Benara & Singh 1999; Snow *et al.* 1992; Dowell *et al.* 1993; Kamali *et al.* 1996; Mobley *et al.* 1996; Quigley *et al.* 1996; Chandramohan *et al.* 1998a,b; Rodriguez *et al.* 1998;

© 2006 Blackwell Publishing Ltd

Kalter *et al.* 1999; Coldham *et al.* 2000; Kahn *et al.* 2000; Marsh *et al.* 2003). Four studies focus on causes of death in adults (Chandramohan *et al.* 1998a,b; Kamali *et al.* 1996; Kahn *et al.* 2000) of which only two examine more than one condition (Chandramohan *et al.* 1998a; Kahn *et al.* 2000). Measures of sensitivity and specificity alone are most commonly used to assess performance of VA in these studies; accuracy of VA-derived cause-specific mortality fractions (CSMF) is rarely included in VA validation. (The CSMF is the proportion of deaths due to cause X divided by the total number of deaths in that age group; in the case of maternal causes, the denominator is limited to the deaths among women aged 15–59 years.) Finally, none of the validation studies that rely on existing medical records (MR) as a 'gold standard' adjust the measure of validity for uncertainty in the reference diagnoses or otherwise systematically account for the poor quality of documentary evidence from health facilities.

We tested the validity of VA procedures developed in Tanzania for use in sentinel community-based surveillance (Ministry of Health 2004a; Setel & Hemed 2004). The VA forms, along with coding and tabulation procedures are intended for application in nationally representative sample registration systems, and research studies to evaluate impact of the efforts aimed at reducing mortality due to specific causes. The procedures developed in Tanzania were adapted for use in China as part of a national sample vital registration programme and were concurrently validated in China using a similar protocol (Yang *et al.* in press). Here, we present the main results from the Tanzania validation study.

## Methods

### Data collection and cause of death attribution

Data were collected and processed prospectively from 2000 to 2003 according to the methods depicted in Figure 1. We used two recruitment strategies (community-based and facility-based) to obtain a sample consisting of deaths that:

- occurred at a health facility; or
- occurred at home following treatment at a health facility (and hence, for which MRs could be obtained) and
- for which a VA could be conducted.

Community recruitment was embedded in a routine demographic surveillance system in three areas of Tanzania as described elsewhere (Mswia *et al.* 2003; Ministry of Health 2004b). The inclusion criteria for this recruitment strategy also included all deaths occurring 18 months prior to the beginning of the study. We obtained informed consent from an appropriate surviving family member to access the decedent's MRs.

Facility-based case recruitment was conducted in several tertiary hospitals and some health centres located in or near the demographic surveillance areas. This included all deaths in these facilities that occurred to individuals residing within a specified geographic radius of the facilities, and that occurred during the course of the study. The geographic restriction was used to maintain similar community characteristics within the sample. We obtained informed consent from bereaved families to visit their household 8–12 weeks after the death to conduct a VA interview. VA interviewers did not have knowledge of the cause of death according to hospital records, and visited families within the prescribed time. VAs were obtained using the same procedures under both recruitment strategies. MRs for all decedents were abstracted, photocopied and filed for coding causes of death. All VA interviews also included questions about whether the respondents were informed about the cause of death by health workers.

Standard three-line death certificates used in Tanzanian hospitals were produced for each VA and each MR by independent two-doctor panels. For the review of the MRs, panel members arrived at their own diagnosis based on all evidence available from the time of admission to the time of discharge. They did not simply assess the validity of diagnoses already contained in the MR based on an independent review of the evidence. For both MR and VA, doctors used their clinical judgment augmented by objective guidelines for malaria, TB, HIV, cholera and stillbirth (details in Annex 1). In addition, training on ICD coding rules aided in determining the correct relationship among underlying, intermediate and immediate causes. While stillbirth is not a cause of death *per se*, it represents a major area of confusion and misclassification with regard to early neonatal mortality and, as a public health burden, deserves quantification in demographic statistics. Therefore, we elected to include an assessment of the validity of VA for ascertaining stillbirth.

The method of relying on the clinical judgement of doctors may introduce some bias or preference in areas of particular interest or specialization on the part of the coding doctor. However, we felt that relying on generalist rather than specialist doctors, providing some criteria for common and potentially ambiguous causes, and the insistence on two independent attributions would minimize such bias.

Death certificates were coded according to ICD-10 coding rules and to the core code and, where possible, four-digit levels. Perinatal events were also coded using the three-line certificate. After the first doctor produced a

 © 2006 Blackwell Publishing Ltd

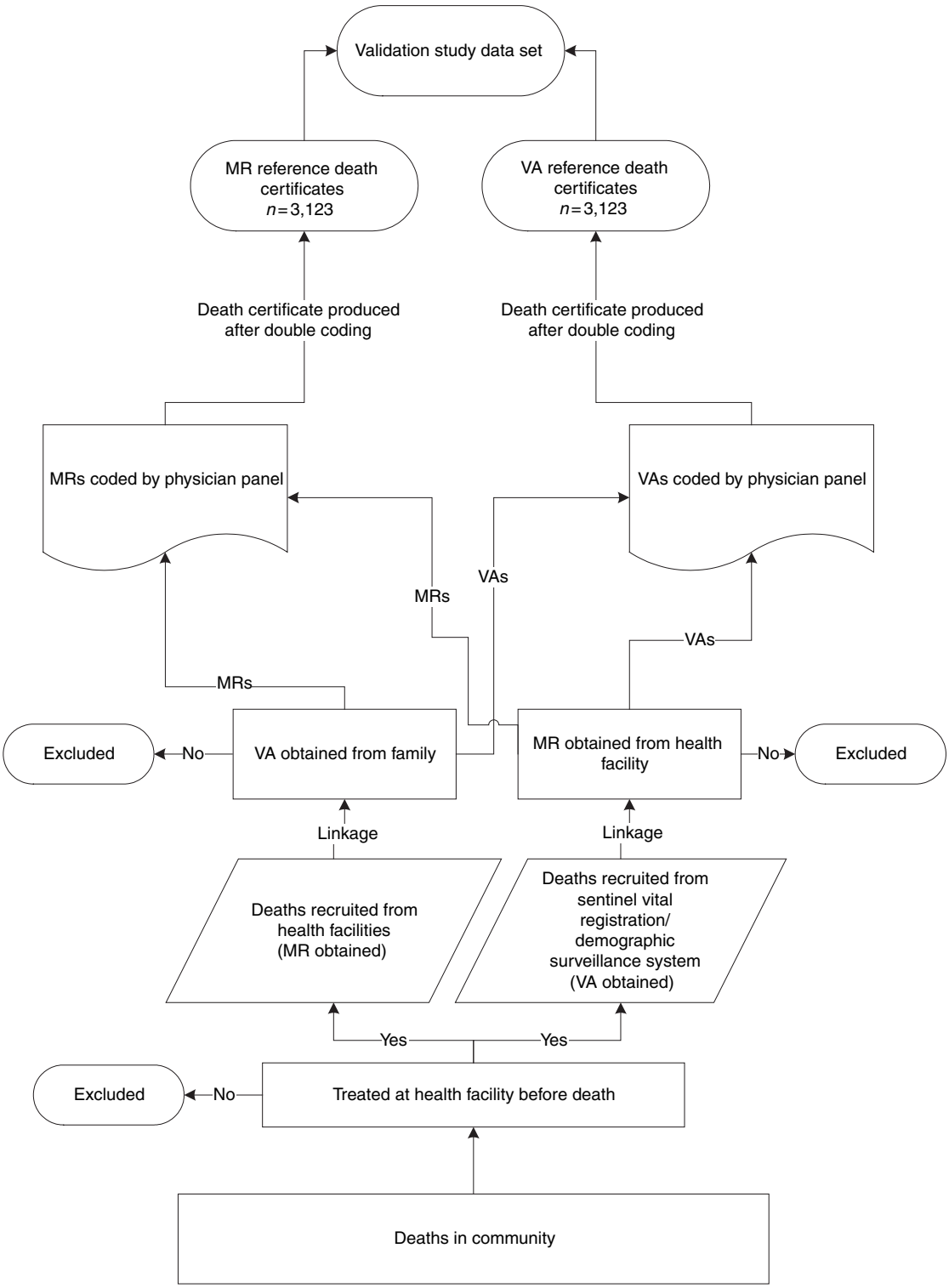


**Figure 1** Recruitment strategies and data flow; VA, verbal autopsy; MR, medicial record.

© 2006 Blackwell Publishing Ltd

death certificate for the VA or the MR, a second doctor, blinded to the results of the first review, coded the same materials to produce a comparison death certificate. In cases of discrepancy, the two coding doctors discussed the case and reached consensus. No doctor coded the MR and VA for the same individual. Following double data entry, all ICD-10 codes were processed by PC-ACME/TRANSAX software to ensure that valid codes for underlying causes of death had been used.

### Adjustment for strength of evidence

Medical records were rated for the strength of evidence in support of the underlying cause of death. The five domains used in this scoring were: appropriate diagnostic tests; appropriate treatment; documented signs; reported presenting symptoms and consistent past medical history. The strength of evidence was assessed in each of these domains for each MR as being 'absent', 'weak', 'strong' and (for appropriate diagnostic tests, appropriate treatment and documented signs) 'confirmatory.' The five scores for each MR death certificate were then combined into a single weight using the formula:

$$W_i = \frac{w_i N}{\sum w_i}$$

These weights were developed using the 'analytic hierarchy process' (Saaty 1990; Expert Choice 2003). This process involved a series of pairwise comparisons made by the study team of the relative importance of each domain. For example, the team made a series of collective value judgments about whether 'appropriate diagnostic tests' were more or less important than 'appropriate treatment' to the overall strength of the MR evidence. This process was repeated with each paired combination of domains, and summary weights were derived and scaled to range from 0 to 1. Missing values were imputed using the mean value for all MR with available weights.

The weights were then scaled to the data and summed to calculate the four cells of adjusted two-by-two tables for each cause, which in turn were used to calculate sensitivity and specificity. Overall, 975 (31.4%) of MRs were given the maximum score (representing greatest confidence in the quality of evidence for the MR diagnosis), 1174 (37.8%) were considered to be fair to very good and the remaining 960 (30.9%) of MRs were assessed as having limited quality and quantity of evidence in support of the reference diagnosis. The weights were then applied in the analysis of sensitivity, specificity and positive predictive value (PPV) reported in Tables 2–4. In order to adjust the two-by-two table parameters, we applied the appropriate weighting function to each MR case before summing across all MR to derive the appropriate value for each table cell and the margin totals.

### Analysis

We developed criteria to evaluate the overall performance of VA procedures in producing correct CSMFs that did not contain a high degree of misclassification. These criteria were as follows: sensitivity >50%; specificity > $(1 - \mathrm{CSMF_{MR}})$; and relative difference between $\mathrm{CSMF_{VA}}$ and $\mathrm{CSMF_{MR}}$ within 20%. These criteria further developed those used by Quigley *et al.* (1999) to assess the performance of data-derived algorithms for VA. They are also based on the mathematical relationship between sensitivity, specificity and CSMF, which implies that to improve the accuracy of the VA it is important to maximize specificity, especially when the CSMF is low, even at the expense of decreasing sensitivity (Anker 1997).

The VA was validated against a 51-item mortality tabulation list comprised of groupings of ICD-10 codes. The cause groupings, ICD-10 codes they contain, and counts for the cause groupings (for both MR and VA) are presented in Table 1.

No ICD-10 code occurred in more than one tabulation category. In order to assess the performance of the VA in each age group, we calculated and compared CSMFs and 95% confidence interval (CI) for the VA and MR along with the coefficient of variation (CV). The CV measures the relative accuracy of the CSMF and provides a qualitative guide to the likely sensitivity of the CSMFs to sample sizes. In general, the higher the CV, the less reliable are the results; the lower the CSMF, the higher the CV. Analysis of sensitivity, specificity and PPV of the VA was carried out adjusting for degree of uncertainty in the reference diagnosis. A separate analysis of sensitivity, specificity and PPV using conventional two-by-two table analysis was carried out, restricted to those cases for which the MR were scored as having the highest levels of evidence in support of the underlying cause of death on the MR death certificate. Confidence intervals were calculated using the delta method to estimate the variance of the ratio of two proportions (DeGroot 1986). No funding source played an active role in the conduct of this research.

### Ethical approval

Ethical Approval for this study was obtained from Harvard University Human Subjects Protection Committee; and the University of Newcastle upon Tyne Ethics Committee. The study was also implemented with the full knowledge and support of the Ministry of Health of Tanzania as part of

 © 2006 Blackwell Publishing Ltd

**Table 1** Tabulation groupings, corresponding ICD-10 core code ranges, and counts

| | | Perinatal/ neonatal | | 1 month to 5 years | | 5 years and above | |
|---|---|---|---|---|---|---|---|
| Tabulation grouping | ICD code range | VA | MR | VA | MR | VA | MR |
| Abdominal pain | R10 | 0 | 0 | 0 | 0 | 1 | 4 |
| All other conditions originating in the perinatal period | P04, P08, P25:P29, P35, P37:P94, P96 | 12 | 19 | 0 | 0 | 0 | 0 |
| All other diseases | D00:D48, D55:D63, D65:D89, E00:E07, E15:E34, E50:E88, G04:G98, H00:H95, I01:I09, I26:I49, I51:I52, I70:I99, J00:J11, J20:J39, J45:J99, K00:K22, K28:K73, K75:K92, L00:L98, M00:M99, N30:N98, Q10:Q88, Q90:Q99 | 6 | 3 | 15 | 14 | 172 | 138 |
| Bacterial sepsis of newborn | P36 | 7 | 18 | 0 | 0 | 0 | 0 |
| Birth asphyxia, and other respiratory disorders specific to the perinatal period | P20:P24 | 91 | 86 | 0 | 0 | 0 | 0 |
| Birth trauma | P10:P15 | 0 | 1 | 0 | 1 | 0 | 0 |
| Cancers of blood | C83:C92 | 0 | 0 | 0 | 0 | 2 | 8 |
| Cancers of the GI tract | C00:C22, C25 | 0 | 0 | 0 | 0 | 32 | 41 |
| Cerebrovascular diseases | I60:I69 | 0 | 0 | 0 | 0 | 109 | 95 |
| Chronic obstructive pulmonary diseases | J40:J44 | 0 | 0 | 0 | 0 | 0 | 1 |
| Cirrhosis liver | K74 | 0 | 0 | 0 | 0 | 2 | 5 |
| Congenital malformations of the central nervous system | Q00:Q07 | 1 | 2 | 0 | 0 | 0 | 0 |
| Diabetes mellitus | E10:E14 | 0 | 0 | 0 | 0 | 30 | 37 |
| Direct maternal | O00:O08, O10:O16, O44:O46, O70:O72, O74:O75 | 0 | 0 | 0 | 0 | 20 | 19 |
| Disorders of the kidney | N00:N29 | 0 | 0 | 1 | 0 | 26 | 15 |
| Fetus and newborn affected by complications of labor and delivery | P03 | 28 | 26 | 0 | 0 | 0 | 0 |
| Fetus and newborn affected by complications of placenta, cord and membranes | P02 | 64 | 71 | 0 | 0 | 0 | 0 |
| Fetus and newborn affected by maternal complications of pregnancy | P01 | 13 | 19 | 0 | 0 | 0 | 0 |
| Fetus and newborn affected by maternal conditions that may be unrelated to present pregnancy | P00 | 110 | 28 | 0 | 0 | 0 | 0 |
| Gastric and duodenal ulcer | K25:K27 | 0 | 0 | 0 | 0 | 4 | 2 |
| Heart failure | I50 | 0 | 0 | 0 | 0 | 26 | 30 |
| HIV disease | B20:B24 | 0 | 0 | 43 | 66 | 563 | 634 |
| Hypertensive diseases | I10:I13 | 0 | 0 | 0 | 1 | 64 | 63 |
| Indirect maternal | O85:O92, O98 | 0 | 0 | 0 | 0 | 7 | 11 |

© 2006 Blackwell Publishing Ltd

**Table 1** Continued

| Tabulation grouping | ICD code range | Perinatal/ neonatal | | 1 month to 5 years | | 5 years and above | |
|---|---|---|---|---|---|---|---|
| | | VA | MR | VA | MR | VA | MR |
| Injuries | V01:V99, W00:W19, W65:W74, X00:X09, X40:X49, X60:X84, X85:Y09, W20:W64, W75:W99, X10:X39, X50:X59, Y10:Y89 | 0 | 0 | 3 | 3 | 53 | 50 |
| Intestinal infectious diseases (including diarrhoeal diseases) | A00:A09 | 0 | 0 | 56 | 64 | 29 | 35 |
| Ischaemic heart diseases | I20:I25 | 0 | 0 | 0 | 0 | 1 | 0 |
| Malaria | B50:B54 | 4 | 2 | 262 | 213 | 392 | 345 |
| Malignant neoplasm of cervix, other and unspecified, parts of uterus | C53:C55 | 0 | 0 | 0 | 0 | 7 | 12 |
| Malignant neoplasm of trachea, bronchus and lung | C33:C34 | 0 | 0 | 0 | 0 | 3 | 2 |
| Malnutrition | E40:E46 | 0 | 0 | 4 | 12 | 0 | 1 |
| Meningitis | G00, G03 | 0 | 0 | 6 | 10 | 13 | 9 |
| Mental and behavioural disorders | F00:F99 | 0 | 0 | 0 | 0 | 9 | 7 |
| Other anaemias | D64 | 0 | 0 | 4 | 7 | 29 | 35 |
| Other classified malformations, not elsewhere classified | Q89 | 3 | 7 | 0 | 0 | 0 | 0 |
| Other maternal causes | O20:O43, O47:O63, O67:O69, O73, O76:O84, O95:O97, O99 | 0 | 0 | 0 | 0 | 6 | 4 |
| Pneumonia | J12:J18 | 4 | 5 | 172 | 165 | 70 | 83 |
| Prematurity and low birth weight | P05:P07 | 41 | 23 | 0 | 0 | 0 | 0 |
| Remainder of infectious and parasitic diseases | A20:A28, A30:A32, A36:A38, A40:A49, A50:A64, A65:A79, A80:A89, A90:A99, B00:B04, B06:B09, B25:B49, B56:B64, B65:B99 | 1 | 0 | 0 | 12 | 7 | 7 |
| Remainder of malignant neoplasms | C23:C24, C26:C32, C37:C50, C51:C52, C56:C82, C93:C97 | 0 | 0 | 1 | 2 | 39 | 25 |
| Stillbirths | P95 | 243 | 318 | 0 | 0 | 0 | 0 |
| Tetanus | A33:A35 | 0 | 0 | 0 | 0 | 3 | 6 |
| Tuberculosis | A15:A19 | 0 | 1 | 15 | 11 | 188 | 163 |
| Unspecified causes of mortality | R00:R09, R11:R49, R51:R55, R57:R99 | 1 | 0 | 0 | 1 | 7 | 27 |
| Total | | 629 | 629 | 582 | 582 | 1912 | 1912 |

GI, gastrointestinal; VA, verbal autopsy; MR, medical record.



© 2006 Blackwell Publishing Ltd

**Table 2** Sensitivity, specificity and positive predictive value for VA in perinatal and neonatal mortality adjusted for degree of uncertainty in reference diagnosis ($n = 629$)

| | Unadjusted | | | | | | | Adjusted | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *N* | | | | | | | | | |
| Cause | MR | VA | Sensitivity | Specificity | PPV | $1 - CSMF_{MR}$ | $CSMF_{MR} - CSMF_{VA}$ | Sensitivity (95% CI) | Specificity (95% CI) | PPV (95% CI) |
| Stillbirths | 318 | 243 | 0.61 | 0.84 | 0.79 | 0.50 | 0.12 | 0.61 (0.57–0.65) | 0.84 (0.80–0.88) | 0.86 (0.83–0.90) |
| Maternal conditions unrelated to pregnancy | 28 | 110 | 0.71 | 0.85 | 0.18 | 0.96 | 0.13 | 0.65 (0.51–0.80) | 0.82 (0.80–0.85) | 0.15 (0.10–0.21) |
| **Birth asphyxia/respiratory disorders*,†** | **86** | **91** | **0.43** | **0.90** | **0.41** | **0.86** | **0.01** | **0.54 (0.41–0.66)** | **0.93 (0.92–0.95)** | **0.38 (0.28–0.48)** |
| **Intrauterine complications** | **71** | **64** | **0.59** | **0.96** | **0.66** | **0.89** | **0.01** | **0.63 (0.54–0.72)** | **0.96 (0.94–0.97)** | **0.67 (0.58–0.77)** |
| Prematurity and low birth weight* | 23 | 41 | 0.48 | 0.95 | 0.27 | 0.96 | 0.03 | 0.64 (0.43–0.86) | 0.97 (0.95–0.98) | 0.31 (0.17–0.45) |
| Complications of labour and delivery | 26 | 28 | 0.23 | 0.96 | 0.21 | 0.96 | 0.00 | 0.22 (0.09–0.36) | 0.97 (0.95–0.98) | 0.22 (0.09–0.36) |
| Maternal complications of pregnancy | 19 | 13 | 0.11 | 0.98 | 0.15 | 0.97 | 0.01 | 0.15 (0.00–0.35) | 0.98 (0.98–0.99) | 0.14 (0.00–0.32) |
| Bacterial sepsis* | 18 | 7 | 0.11 | 0.99 | 0.29 | 0.97 | 0.02 | 0.06 (0.00–0.18) | 1.00 (1.00–1.00) | 0.35 (0.00–0.97) |
| **Pneumonia** | **6** | **5** | **0.50** | **1.00** | **0.60** | **0.99** | **0.00** | **0.51 (0.02–0.99)** | **1.00 (1.00–1.00)** | **0.82 (0.34–1.00)** |
| Malaria | 2 | 4 | 0.50 | 1.00 | 0.25 | 1.00 | 0.00 | 0.50 (0.00–1.00) | 0.99 (0.99–1.00) | 0.29 (0.00–0.65) |

Asterisk (*) indicates unadjusted value of sensitivity, specificity, or PPV outside 95% CI of adjusted value.
†Causes in boldface meet criteria for likelihood of robustness of VA cause-specific mortality fractions (see text).
VA, verbal autopsy; MR, medical record; PPV, positive predictive value; CSMF, cause-specific mortality fraction; CI, confidence interval.

**Table 3** Sensitivity, specificity and positive predictive value for VA in deaths among children 1 month to 5 years adjusted for degree of uncertainty in reference diagnosis ($n = 582$)

| | Unadjusted | | | | | | | Adjusted | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *N* | | | | | | | | | |
| Cause | MR | VA | Sensitivity | Specificity | PPV | $1 - CSMF_{MR}$ | $CSMF_{MR} - CSMF_{VA}$ | Sensitivity (95% CI) | Specificity (95% CI) | PPV (95% CI) |
| Malaria* | 213 | 262 | 0.69 | 0.69 | 0.56 | 0.63 | 0.08 | 0.67 (0.61–0.73) | 0.67 (0.62–0.73) | 0.65 (0.60–0.71) |
| **Pneumonia†** | **165** | **171** | **0.63** | **0.84** | **0.61** | **0.72** | **0.01** | **0.62 (0.54–0.70)** | **0.84 (0.81–0.88)** | **0.58 (0.50–0.66)** |
| HIV | 65 | 43 | 0.31 | 0.96 | 0.47 | 0.89 | 0.04 | 0.37 (0.20–0.54) | 0.96 (0.94–0.97) | 0.34 (0.18–0.50) |
| Intestinal infections/ diarrhoea | 64 | 56 | 0.41 | 0.94 | 0.46 | 0.89 | 0.01 | 0.42 (0.27–0.56) | 0.94 (0.92–0.96) | 0.39 (0.25–0.52) |
| Malnutrition | 12 | 4 | 0.00 | 0.99 | 0.00 | 0.98 | 0.01 | 0.00 (0.00–0.00) | 0.99 (0.99–1.00) | 0.00 (0.00–0.00) |
| Tuberculosis | 11 | 15 | 0.18 | 0.98 | 0.13 | 0.98 | 0.01 | 0.19 (0.00–0.50) | 0.98 (0.97–0.99) | 0.12 (0.00–0.32) |
| Meningitis | 10 | 6 | 0.30 | 0.99 | 0.50 | 0.98 | 0.01 | 0.31 (0.04–0.58) | 0.99 (0.98–1.00) | 0.44 (0.10–0.79) |
| Other anaemias | 7 | 4 | 0.00 | 0.99 | 0.00 | 0.99 | 0.01 | 0.00 (0.00–0.00) | 0.99 (0.99–1.00) | 0.00 (0.00–0.00) |
| **Injuries** | **3** | **3** | **0.67** | **1.00** | **0.67** | **0.99** | **0.00** | **0.91 (0.53–1.00)** | **1.00 (1.00–1.00)** | **0.80 (0.32–1.00)** |

Asterisk (*) indicates unadjusted value of sensitivity, specificity, or PPV outside 95% CI of adjusted value.
†Causes in boldface meet criteria for likelihood of robustness of VA cause-specific mortality fractions (see text).
VA, verbal autopsy; MR, medical record; PPV, positive predictive value; CSMF, cause-specific mortality fraction; CI, confidence interval.

© 2006 Blackwell Publishing Ltd

the Adult Morbidity and Mortality Project (AMMP). AMMP was a project of the Tanzanian Ministry of Health, funded by the Department for International Development (DFID), UK, and implemented in partnership with the University of Newcastle upon Tyne.

## Results

We analysed 4014 deaths. Of these 3123 had both MR and VA available for cause of death determination, including 582 perinatal and neonatal; 629 post-neonatal children and 1912 over ages of 5 years. Community-based recruitment yielded a total of 2386 potentially linkable deaths, of which 51% ($n = 1211$) were linked to the health facility level, and 26.7% ($n = 637$) were ultimately linked to useable MRs. Final linkage rates for deaths recruited from the community were not significantly different between rural and urban areas (59.3% *vs.* 57.5%; $P = 0.61$). Data were obtained from 14 facilities, with four municipal and one zonal referral hospital contributing 73% of all deaths analysed ($N = 2281$). Twenty percent of respondents (421 of 2121) reported that they had discussed the cause of death of their relative with a health worker.

Figures 2–4b present CSMFs (with 95% CI) and CVs for VA and MR for important causes of death in each age group. Among 629 deaths recorded in the perinatal/neonatal period (Figure 2), $CSMF_{VA}$ and $CSMF_{MR}$ differed markedly for stillbirths and maternal conditions unrelated to the current pregnancy, with VA underestimating the former and overestimating the latter relative to MR. There were no other statistically significant differences between MR and VA CSMFs for other causes of death. CV ranged from 0.00 to 0.01 for CSMFs based on 60 or more events per cause; 0.03 to 0.05 for CSMFs based on 18–28 events per cause and 0.20 to 0.33 for causes of death based on seven or fewer events. All deaths were included in the denominator for CSMF calculations; however, we have not included results from 30 deaths as a result of 12 causes with small numbers of observations. One death was coded as 'ill-defined' in the MR; there were no perinatal or neonatal deaths in the VA due to ill-defined causes. Although the use of the ICD-10 codes P00–P04 'Foetus and newborn affected by maternal factors and by complications of pregnancy, labour and delivery' are not valid for coding underlying causes of perinatal mortality (World Health Organization 1993) we used these codes to tabulate and validate such causes. This was performed in order to investigate and highlight the contribution of maternal conditions to perinatal death as a possible target for intervention.

In 582 post-neonatal child deaths (Figure 3), the $CSMF_{VA}$ (45.2%) and $CSMF_{MR}$ (36.7%) for malaria differed significantly. There were no statistically significant differences in VA and MR CSMFs for any other causes. CV ranged from 0.00 to 0.02 for CSMFs based on 40 or more events per cause; 0.07 to 0.25 for CSMFs based on 10–12 events per cause and 0.14 to 0.33 for causes of death based on seven or fewer events. All deaths were included in the denominator for CSMF calculations. We do not show results from 31 MR and 17 VA deaths due to seven causes with small samples. One post-neonatal child death was coded as 'undetermined' in the MR; there were no ill-defined deaths in the VA group.

Figure 4a,b are based on the 1912 deaths in older children and adults. There were no statistically significant differences in the CSMFS of the MR and VA at the 95%

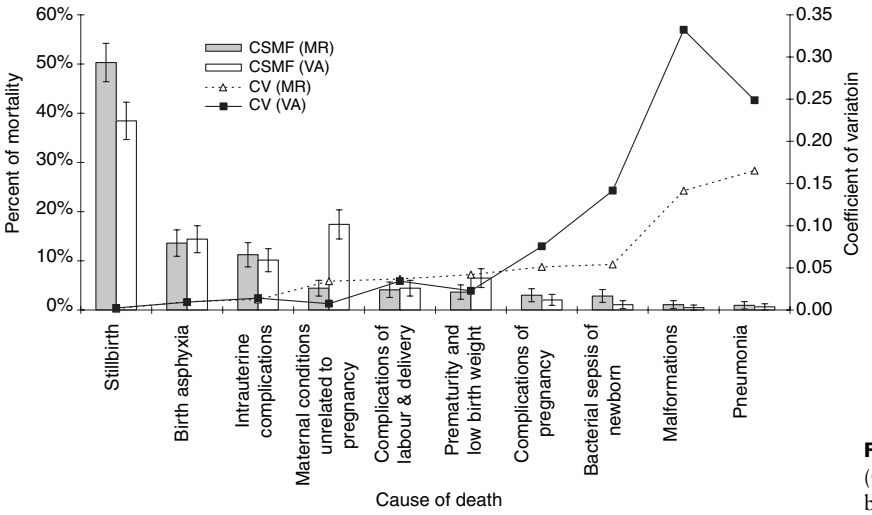


**Figure 2** Cause-specific mortality fractions (CSMF) and coefficients of variation (CV), birth to 29 days ($N = 629$).


© 2006 Blackwell Publishing Ltd

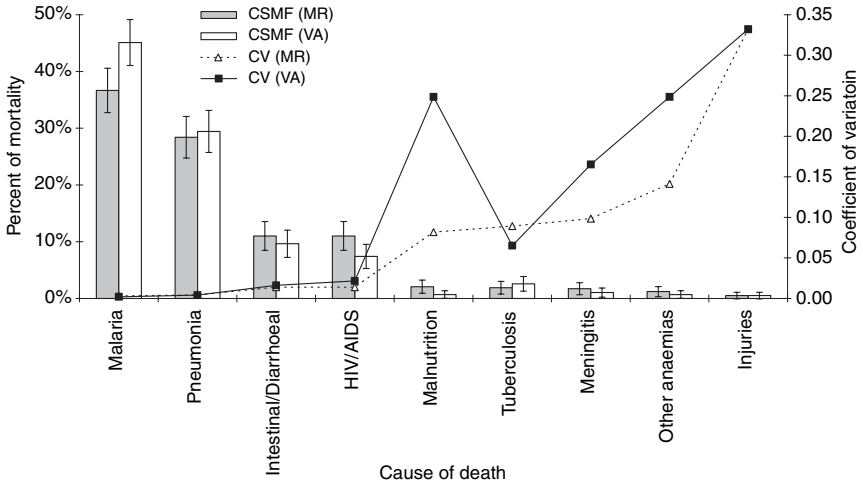


**Figure 3** Cause-specific mortality fractions (CSMF) and coefficients of variation (CV), 1 month to 5 years ($N = 582$).

level for any of 20 causes of death in this age group. Coefficients of variation ranged from 0.00 to 0.02 for CSMFs based on 50 or more events per cause; 0.02 to 0.07 for CSMFs based on 15–50 events per cause and 0.08 to 0.50 for causes of death based on 10 or fewer events. All deaths were included in the denominator for CSMF calculations. We did not display results from 138 MR ($CSMF_{MR}$ 7.2%) and 172 VA ($CSMF_{VA}$ 9.0%) deaths due to other specified conditions; 27 MR ($CSMF_{MR}$ 1.4%) and seven VA ($CSMF_{VA}$ 0.4%) deaths from unspecified causes; 33 MR deaths due to nine minor causes; and 33 VA deaths from eight minor causes in the VA.

Tables 2–4 report the sensitivity, specificity and PPV of VA for each age group. Both unadjusted and adjusted (for degree of uncertainty in the reference diagnosis) are provided, with 95% CI for adjusted values only. In the 582 perinatal and neonatal deaths (Table 2), the sensitivity of VA ranged from nearly 0 for bacterial sepsis to 0.65 for maternal conditions affecting the foetus or newborn that may be unrelated to the current pregnancy. Other causes of death with adjusted sensitivity above 0.50 included stillbirths, birth asphyxia, intrauterine complications, prematurity/low birth weight and pneumonia. Specificity was generally high, ranging from 0.82 (0.80–0.85) for maternal conditions unrelated to present pregnancy, to 1.00 (0.99–1.00) for pneumonia and bacterial sepsis. Adjustment for quality of evidence in the MR led to significant differences in parameter estimates (i.e. unadjusted values outside the 95% CI range of adjusted values) for the specificity of birth asphyxia, prematurity/low birth weight and bacterial sepsis, and for the PPV of stillbirth. In all cases, where significant differences were found, the adjusted estimate and CIs were higher than the unadjusted values.

In post-neonatal children under 5 years (Table 3), sensitivity ranged from 0 for anaemia to 0.67 (0.61–0.73) for malaria. Sensitivity also exceeded 0.50 for pneumonia and injuries. Specificity was above 0.95 for all causes except malaria (0.67; 0.62–0.72); pneumonia (0.84; 0.81–0.88) and intestinal infections/diarrhoea (0.94; 0.92–0.96). Adjusting for the strength of evidence in support of the underlying cause of death from MRs, led to significant improvement in the PPV for malaria, but did not change significantly parameter estimates for any other cause.

For deaths of those aged 5 years and older (Table 4), the sensitivity of VA ranged from 0.10 (0.00–0.30) for disorders of the kidney to 0.73 (0.57–0.89) for injuries. The sensitivity of VA also exceeded 0.50 for HIV/AIDS; malaria; tuberculosis; cerebrovascular diseases; cancers of the gastrointestinal tract; cancers of the female genital tract; direct maternal causes and tetanus. The lowest specificities were 0.86 (0.84–0.88) for HIV/AIDS, 0.90 (0.98–0.92) for malaria and 0.94 (0.95–0.95) for tuberculosis; values of specificity for all other causes were >0.95. After adjusting for the strength of evidence, unadjusted estimates of PPV for HIV/AIDS, malaria, tuberculosis and cancers of the female genital tract were outside the 95% CI range of the adjusted point estimate. In all cases except HIV/AIDS, PPV increased with the adjustment procedure.

Although we felt that the truest reflection of VA performance was to be gleaned from an analysis of all available data with systematic adjustments as described, we also performed an analysis of summary measures limited only to cases where the MR was scored as having the highest level of evidence in support of the underlying cause of death. This subset analysis generally resulted in significant improvements to sensitivity, some improvement to specificity and PPV. However, there was an increase in

© 2006 Blackwell Publishing Ltd

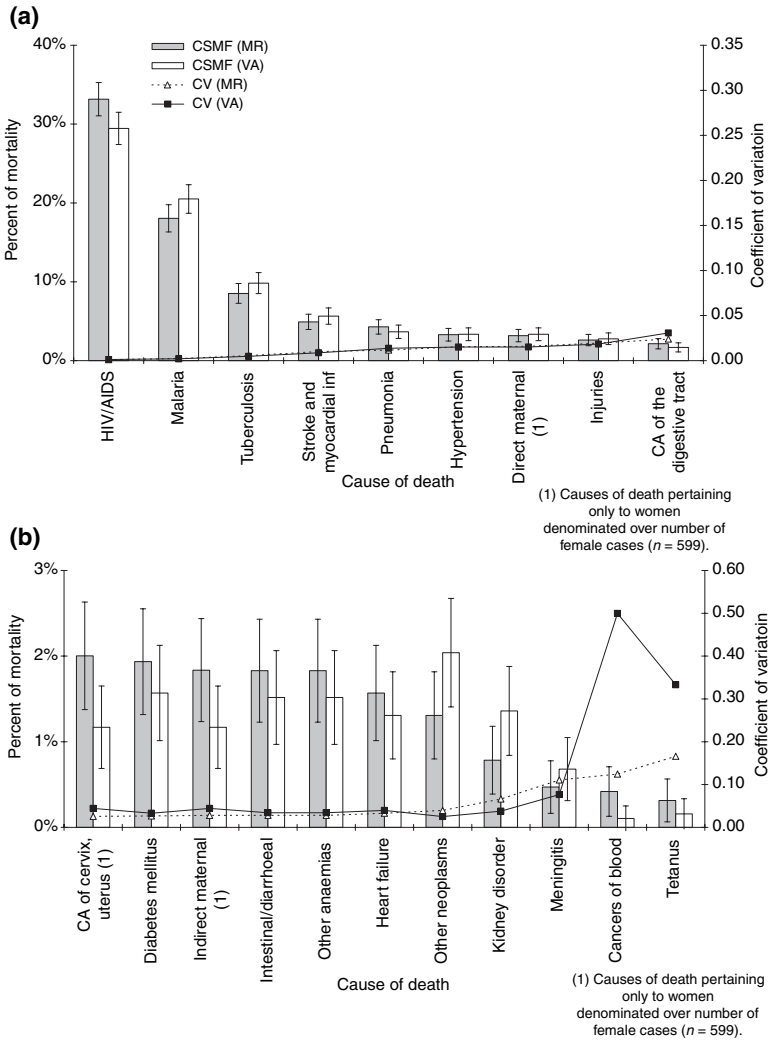


**Figure 4** (a) Cause-specific mortality fractions (CSMF) and coefficients of variation (CV), 5 years and older for causes >2% of all mortality ($N = 1912$). (b) CSMF and CV, 5 years and older for causes <1.9% of all mortality ($N = 1912$).

the differences between $CSMF_{VA}$ compared with $CSMF_{MR}$ for some causes of death (data not shown).

Evaluating the performance of VA indicates that accurate results were obtained for nine different causes of death of public health importance across all age groups. Causes of perinatal and neonatal mortality included: birth asphyxia/respiratory disorders; intrauterine complications and pneumonia. Causes of post-neonatal child mortality were pneumonia and injuries. Causes in those age over 5 years included: HIV/AIDS; malaria; tuberculosis; cerebrovascular diseases; injuries and direct maternal causes.

## Discussion

Some previous validation studies differ in methods of cause of death attribution and cause of death categorization to an extent that makes direct comparison problematic (i.e. used strict rules for cause of death attribution based on symptoms and signs; Kalter *et al.* 1990, 1999; Anker *et al.* 1999; Coldham *et al.* 2000; Marsh *et al.* 2003). Validation studies that are more directly comparable with the present research (i.e. used the clinical judgement of doctors to assign causes, and/or are coded to ICD or ICD-derived cause groups) include those by Snow *et al.* (1992), Dowell *et al.* (1993), Kamali *et al.* (1996), Chandramohan *et al.* (1998a,b) and Kahn *et al.* (2000).

Verbal autopsy appeared to underestimate the proportion of stillbirths, and overestimate the number of deaths due to maternal conditions unrelated to the pregnancy. With regard to the underestimate of stillbirths by the VA, anecdotally it has been reported that a tendency may exist at health facilities to report early neonatal deaths as

 © 2006 Blackwell Publishing Ltd

**Table 4** Sensitivity, specificity and positive predictive value for VA in deaths in ages 5 years and older adjusted for degree of uncertainty in reference diagnosis ($n = 1912$)

| | Unadjusted | | | | | | | Adjusted | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *N* | | | | | | | | | |
| Cause | MR | VA | Sensitivity | Specificity | PPV | $1 - CSMF_{MR}$ | $CSMF_{MR} - CSMF_{VA}$ | Sensitivity (95% CI) | Specificity (95% CI) | PPV (95% CI) |
| **HIV/AIDS*†** | **634** | **563** | **0.59** | **0.85** | **0.67** | **0.67** | **0.04** | **0.60 (0.55–0.65)** | **0.86 (0.84–0.88)** | **0.57 (0.53–0.62)** |
| **Malaria*** | **345** | **392** | **0.64** | **0.89** | **0.56** | **0.82** | **0.02** | **0.63 (0.59–0.68)** | **0.90 (0.89–0.92)** | **0.70 (0.65–0.75)** |
| **Tuberculosis*** | **163** | **188** | **0.45** | **0.93** | **0.39** | **0.91** | **0.01** | **0.51 (0.43–0.58)** | **0.94 (0.93–0.95)** | **0.49 (0.41–0.56)** |
| Cerebrovascular diseases | 94 | 108 | 0.63 | 0.97 | 0.55 | 0.95 | 0.01 | 0.59 (0.47–0.70) | 0.97 (0.96–0.98) | 0.47 (0.36–0.57) |
| Pneumonia | 82 | 70 | 0.18 | 0.97 | 0.21 | 0.96 | 0.01 | 0.17 (0.08–0.26) | 0.97 (0.96–0.98) | 0.19 (0.09–0.29) |
| Hypertensive diseases | 63 | 64 | 0.33 | 0.98 | 0.33 | 0.97 | 0.00 | 0.40 (0.27–0.52) | 0.97 (0.96–0.98) | 0.33 (0.22–0.45) |
| **Injuries** | **50** | **53** | **0.74** | **0.99** | **0.70** | **0.97** | **0.00** | **0.73 (0.57–0.89)** | **0.99 (0.99–1.00)** | **0.66 (0.50–0.83)** |
| Cancers of the GI tract | 41 | 32 | 0.56 | 1.00 | 0.72 | 0.98 | 0.00 | 0.55 (0.40–0.70) | 0.99 (0.99–1.00) | 0.69 (0.53–0.84) |
| Diabetes mellitus | 37 | 30 | 0.43 | 0.99 | 0.53 | 0.98 | 0.00 | 0.41 (0.28–0.55) | 0.99 (0.99–1.00) | 0.61 (0.45–0.78) |
| Intestinal infections/diarrhoea | 35 | 29 | 0.34 | 0.99 | 0.41 | 0.98 | 0.00 | 0.45 (0.27–0.64) | 0.99 (0.99–1.00) | 0.51 (0.31–0.71) |
| Other anaemias | 35 | 29 | 0.26 | 0.99 | 0.31 | 0.98 | 0.00 | 0.31 (0.15–0.46) | 0.99 (0.98–0.99) | 0.38 (0.21–0.56) |
| Heart failure | 30 | 25 | 0.27 | 0.99 | 0.32 | 0.98 | 0.00 | 0.25 (0.08–0.41) | 0.99 (0.99–1.00) | 0.31 (0.11–0.50) |
| Other cancers | 25 | 39 | 0.36 | 0.98 | 0.23 | 0.99 | 0.01 | 0.37 (0.21–0.53) | 0.98 (0.97–0.99) | 0.28 (0.15–0.41) |
| **Direct maternal** | **19** | **20** | **0.63** | **1.00** | **0.60** | **0.99** | **0.00** | **0.63 (0.40–0.86)** | **1.00 (0.99–1.00)** | **0.57 (0.34–0.80)** |
| Disorders of the kidney | 15 | 26 | 0.13 | 0.99 | 0.08 | 0.99 | 0.01 | 0.10 (0.00–0.30) | 0.99 (0.98–0.99) | 0.04 (0.00–0.12) |
| Cancer female genital tract* | 12 | 7 | 0.50 | | 1.00 | 0.86 | 0.99 | 0.00 | 0.58 (0.34–0.82) | 1.00 (1.00–1.00) |
| Indirect maternal | 11 | 7 | 0.18 | 1.00 | 0.29 | 0.99 | 0.00 | 0.22 (0.00–0.45) | 1.00 (1.00–1.00) | 0.47 (0.05–0.88) |
| Meningitis | 9 | 13 | 0.22 | 0.99 | 0.15 | 1.00 | 0.00 | 0.19 (0.00–0.53) | 0.99 (0.99–1.00) | 0.09 (0.00–0.27) |
| Cancers of blood | 8 | 2 | 0.25 | 1.00 | 1.00 | 1.00 | 0.00 | 0.21 (0.00–0.45) | 1.00 (1.00–1.00) | 1.00 (1.00–1.00) |
| Tetanus | 6 | 3 | 0.50 | 1.00 | 1.00 | 1.00 | 0.00 | 0.52 (0.01–1.00) | 1.00 (1.00–1.00) | 1.00 (1.00–1.00) |

Asterisk (*) indicates unadjusted value of sensitivity, specificity, or PPV outside 95% CI of adjusted value.
†Causes in boldface meet criteria for likelihood of robustness of VA cause-specific mortality fractions (see text).
VA, verbal autopsy; MR, medical record; PPV, positive predictive value; CSMF, cause-specific mortality fraction; CI, confidence interval; GI, gastrointestinal.

© 2006 Blackwell Publishing Ltd

stillbirths, and that stillbirth reporting rates may vary drastically among settings that would ordinarily be expected to have comparable rates. Reasons for intentional misattribution may include a concern that representing a death as a stillbirth might be less distressing to bereaved parents, or that reporting a neonatal death may trigger a cumbersome additional reporting or investigation procedure. Very few deaths were assigned cause of death 'unknown' by VA in our study compared with previous validation studies because: (i) the doctors discussed the cause of death and reached a diagnosis by consensus when there was a discordant diagnosis at the time of independent coding; and (ii) the ICD-10 coding scheme allows assignment of syndromes within an anatomical organ system as a valid cause of death.

During the period of observation, we observed only one fatality because of congenital defects. Although the MR and VA agreed on the diagnosis in this case, it was clearly not possible to validate the VA against this cause group. The scarcity of congenital defects as a cause of death in both the reference and VA data sets may be due to a combination of factors. First, some defects likely to be more prevalent, such as septal/cardiac defects, may result in a stillbirth or in an early neonatal death not coded to congenital abnormality. This would occur because congenital cardiac defects cannot be ascertained without an autopsy that entails dissection – and such procedures are very uncommon in Tanzania – and would be unlikely to be picked up by VA. In cases of congenital defects that are fatal beyond 29 days after birth, the proportion of deaths due to this cause would likely be small in relation to all other causes identified in the 1-month to 5-year-old age group.

Our results for sensitivity, specificity and PPV add to and tend to support some of the previous conclusions about the use of VA in Africa. In particular, PPV for malaria (0.65) is similar to that reported previously from Kenya (0.57; Snow *et al.* 1992) as well as for diarrhoeal diseases and childhood injuries. While our estimate of sensitivity for malaria in post-neonatal children (0.67) was somewhat higher than that found in previous validation studies (range: 0.45–0.75), our estimate of specificity (0.67) was somewhat lower (range: 0.77–1.00; Korenromp *et al.* 2003). It may be further noted that sensitivity of VA appears to be inversely related to the proportion of deaths due to malaria (Korenromp *et al.* 2003), and that the proportion of deaths in the MR data set (37%) is considerably higher than that found in previous validation studies.

Conversely, our results suggest that, in Tanzania at least, these VA methods had substantially higher PPV for childhood pneumonia and meningitis than those validated elsewhere, including Kenya (0.58, 0.29, respectively, for pneumonia; 0.44, 0.20 for meningitis). Our methods also suggest a similar level of sensitivity and PPV for VA in diagnosing HIV/AIDS deaths in children as other studies (Snow *et al.* 1992; Dowell *et al.* 1993; Kahn *et al.* 2000). Specificities, which are central to the accuracy of VA (Anker 1997), are above 0.90 for most causes of death.

Among adults our findings are broadly consistent with results of previous research, with wide ranges in sensitivity and PPV, and specificities typically ≥0.90. In particular, for HIV, sensitivity, specificity and PPV were all similar to a previous study in Tanzania (Chandramohan *et al.* 1998a) and within the range reported from Uganda (Kamali *et al.* 1996). Similar results were observed for tuberculosis, cerebrovascular disease and diarrhoea. VA performed better for malaria deaths among adults than in previous research, but less well for pneumonia, injuries, meningitis and maternal causes (Chandramohan *et al.* 1998a,b; Kahn *et al.* 2000). It is not possible to determine the degree to which variation in performance across countries is due to differences in VA procedures and epidemiological context, or even due to the varying quality of MRs against which VA is validated.

The VA procedures validated in Tanzania provide robust measurement of nine causes of death of public health importance across all age groups. Additionally, VA data for other important causes, such as malaria in children, may still be relevant for priority setting and monitoring trends in cause-specific mortality, even if these causes did not meet all three accuracy criteria. For causes with samples that should be adequate for validation and that did not reach the threshold for good performance, it is uncertain that modifying VA procedures would necessarily improve matters; for some important causes of death it may not be feasible, using VA, to attain a degree of reliability to be desired in a policy-making or impact evaluation tool.

The generalizability of results from a VA validation study depends on a number of factors. Chief among these are the: (i) degree to which the cause structure of the validation sample resembles that of the general population to which the results are to be applied and (ii) degree to which responses to VA questions have been influenced by contact with the health system (and hence may systematically differ from the responses of caregivers and family members of those who die at home and without medical attention in the period before death). These conditions pose what may be insuperable conundrums for VA validation. In the former case, the entire enterprise would be unnecessary were the mortality cause structure of the general population known. Comparisons of community-based surveillance (based on unvalidated CSMFs) with the data set for this study are the subject of a forthcoming

 © 2006 Blackwell Publishing Ltd

publication, but do suggest that selection biases related to whether causes of death are acute or chronic affect the probability of inclusion in the validation set.

In the latter case, there is no practical alternative to carrying out VA validation studies with respondents whose knowledge about cause of death may have been influenced by contact with health workers, although we note that in this study only about 20% of respondents reported discussing the cause of death with a healthcare worker. Under these circumstances a qualitative assessment of the robustness of the results by experts may be the most that can be hoped for. Using the estimates of sensitivity and specificity to adjust VA estimates from other sources must be approached with caution as these values also vary according to the underlying mortality cause structure; the cross-application of sensitivity and specificity to situations with significantly different cause structures can lead to spurious results (Chandramohan *et al.* 2001).

Verbal autopsy procedures are growing in importance as a source of data for populations lacking other reliable sources of mortality information. Application and refinement of existing VA methods holds out the possibility of obtaining replicable, sustainable and internationally comparable mortality statistics of known quality for the majority of the world's population for whom such knowledge has, to date, been unavailable. Where its validity is known, VA has the potential to provide cost-effective information to guide policy, set priorities and track impact, particularly in countries undergoing rapid epidemiological and mortality transitions.

## Acknowledgements

This study is, in part an output of AMMP. Research funding and publication support was also provided by grant P10462-109/9903GLOB-2, The Global Burden of Disease 2000 in Aging Populations, US National Institutes of Health (NIH), National Institute on Ageing (NIA). Funding for the preparation of this study was provided through MEASURE Evaluation, Phase 2, a USAID Cooperative Agreement (No. GPO-A-00-03-00003-00) implemented by the Carolina Population Center, University of North Carolina at Chapel Hill. The views expressed are not necessarily those of DFID, NIH, NIA, or USAID. The authors also wish to express their appreciation to Chalapati Rao for input into the study.

## References

Anker M (1997) The effect of misclassification error on reported cause-specific mortality fractions from verbal autopsy. *International Journal of Epidemiology* **26**, 1090–1096.

Anker M, Black RE, Coldham C *et al.* (1999) *A Standard Verbal Autopsy Method for Investigating Causes of Death in Infants and Children (WHO/CDS/CSR/ISR/99.4)*. World Health Organization, Geneva, Switzerland.

Benara SK & Singh P (1999) Validity of causes of infant death by verbal autopsy. *Indian Journal of Pediatrics* **66**, 647–650.

Bryce J, Victora CG, Habicht J-P, Vaughan JP & Black RE (2004) The multi-country evaluation of the integrated management of childhood illness strategy: lessons for the evaluation of public health interventions. *American Journal of Public Health* **94**, 406–415.

Chandramohan D, Maude GH, Rodrigues LC & Hayes RJ (1998a) Verbal autopsies for adult deaths: their development and validation in a multi-centre study. *Tropical Medicine and International Health* **3**, 436–446.

Chandramohan D, Rodrigues LC, Maude GH & Hayes RJ (1998b) The validity of verbal autopsies for assessing the causes of institutional maternal death. *Studies in Family Planning* **29**, 414–422.

Chandramohan D, Setel P & Quigley M (2001) Effect of misclassification of causes of death in verbal autopsy: can it be adjusted? *International Journal of Epidemiology* **30**, 509–514.

Coldham C, Ross D, Quigley M, Segura Z & Chandramohan D (2000) Prospective validation of a standardized questionnaire for estimating childhood mortality and morbidity due to pneumonia and diarrhoea. *Tropical Medicine and International Health* **5**, 134–144.

DeGroot MH (1986) *Probability and Statistics*, 2nd edn. Addison Wesley, Reading, UK.

Dowell SF, Davis HL, Holt EA *et al.* (1993) The utility of verbal autopsies for identifying HIV-1-related deaths in Haitian children. *AIDS* **7**, 1255–1259.

Expert Choice I (2003) *Expert Choice*. Expert Choice, Inc., Arlington, VA, USA.

Kahn K, Tollman SM, Garenne M & Gear JS (2000) Validation and application of verbal autopsies in a rural area of South Africa. *Tropical Medicine and International Health* **5**, 824–831.

Kalter HD, Gray R, Black R & Gultiano SA (1990) Validation of postmortem interviews to ascertain selected causes of death in children. *International Journal of Epidemiology* **19**, 380–386.

Kalter HD, Hossain M, Burnham G *et al.* (1999) Validation of caregiver interviews to diagnose common causes of severe neonatal illness. *Paediatric and Perinatal Epidemiology* **13**, 99–113.

Kamali A, Wagner HU, Nakiyingi J, Sabiiti I, Kengeya-Kayondo JF & Mulder DW (1996) Verbal autopsy as a tool for diagnosing HIV-related adult deaths in rural Uganda. *International Journal of Epidemiology* **25**, 679–684.

Korenromp EL, Williams BG, Gouws E, Dye C & Snow RW (2003) Measurement of trends in childhood malaria mortality in Africa: an assessment of progress toward targets based on verbal autopsy. *The Lancet Infectious Diseases* **3**, 349–358.

Marsh DR, Sadruddin S, Fikree FF, Krishnan C & Darmstadt GL (2003) Validation of verbal autopsy to determine the cause of 137 neonatal deaths in Karachi, Pakistan. *Paediatric and Perinatal Epidemiology* **17**, 132–142.

© 2006 Blackwell Publishing Ltd

Mathers C, Ma Fat D, Inoue M, Rao C & Lopez AD (2005) Counting the dead and what they died of: an assessment of the global status of cause of death data. *Bulletin of the World Health Organization* **83**, 171–177.

Ministry of Health (2004a) *The Policy Implications of Tanzania's Mortality Burden: A Ten-Year Community-Based Perspective*. AMMP-2, Final Report, Vol. 1. Dar es Salaam, Adult Morbidity and Mortality Project, Tanzania Ministry of Health, University of Newcastle upon Tyne, UK Department for International Development, Newcastle upon Tyne, UK.

Ministry of Health (2004b) *The Policy Implications of Tanzania's Mortality Burden: Field Operations and Validation Studies*. AMMP-2, Final Report, Vol. 3. Dar es Salaam, Adult Morbidity and Mortality Project, Tanzania Ministry of Health, University of Newcastle upon Tyne, UK Department for International Development, Newcastle upon Tyne, UK.

Mobley CC, Boerma JT, Titus S, Lohrke B, Shangula K & Black RE (1996) Validation study of a verbal autopsy method for causes of childhood mortality in Namibia. *Journal of Tropical Pediatrics* **42**, 365–369.

Morris SS, Black RE & Tomaskovic L (2003) Predicting the distribution of under-five deaths by cause in countries without adequate vital registration systems. *International Journal of Epidemiology* **32**, 1041–1051.

Mswia R, Lewanga M, Moshiro C *et al.* (2003) Community-based monitoring of safe motherhood in the United Republic of Tanzania. *Bulletin of the World Health Organization* **81**, 87–94.

Quigley MA, Armstrong Schellenberg JR & Snow RW (1996) Algorithms for verbal autopsies: a validation study in Kenyan children. *Bulletin of the World Health Organization* **74**, 147–154.

Quigley MA, Chandramohan D & Rodrigues LC (1999) Diagnostic accuracy of physician review, expert algorithms and data-derived algorithms in adult verbal autopsies. *International Journal of Epidemiology* **28**, 1081–1087.

Rodriguez L, Reyes H, Tome P, Ridaura C, Flores S & Guiscafre H (1998) Validation of the verbal autopsy method to ascertain acute respiratory infection as cause of death. *Indian Journal of Pediatrics* **65**, 579–584.

Saaty T (1990) *Multicriteria Decision Making: The Analytic Hierarchy Process*. RWS Publications, Pittsburgh, PA, USA.

Setel P & Hemed Y (2004) Box 1.1: Sentinel vital registration in the United Republic of Tanzania. In: *World Health Report 2003. Shaping the Future* (ed. World Health Organization) World Health Organization, Geneva, Switzerland.

Setel PW, Sankoh O, Mathers C *et al.* (2005) Sample registration of vital events with verbal autopsy: innovative approaches to measuring and monitoring vital statistics. *Bulletin of the World Health Organization* **83**, 611–617.

Snow RW, Armstrong JRM, Forster D *et al.* (1992) Childhood deaths in Africa: uses and limitations of verbal autopsies. *Lancet* **340**, 351–355.

Williams BG, Gouws E, Boschi-Pinto C, Bryce J & Dye C (2002) Estimates of world-wide distribution of child deaths from acute respiratory infections. *The Lancet Infectious Diseases* **2**, 25–32.

World Health Organization (1993) *International Statistical Classification of Diseases and Related Health Problems*, Tenth Revision, Vol. 2. World Health Organization, Geneva, Switzerland.

Yang G, Rao C, Ma J *et al.* (in press) Validation of verbal autopsy procedures for adult deaths in China. *International Journal of Epidemiology*.

**Corresponding Author Philip W. Setel**, Department of Epidemiology and MEASURE Evaluation, Carolina Population Center, University of North Carolina at Chapel Hill, 206 W Franklin St, 2nd Floor, Chapel Hill, NC 27516, USA. Tel.: +1 919 966 7541; Fax: +1 919 966 2391; E-mail: psetel@unc.edu

**Validité des procédures de l'autopsie verbale pour la détermination des causes de décès en Tanzanie**

OBJECTIFS Valider les procédures d'autopsie verbale pour leur utilisation dans l'enregistrement d'échantillon vital. L'autopsie verbale est une méthode importante pour déterminer la mortalité cause-spécifique dans les endroits où les taux de décès sont les plus élevés et les données sur la mortalité cause-spécifique sont inexistantes.

MÉTHODES Les autopsies verbales et les records médicaux ont été collectés sur 3123 décès dans la période péri/néonatale, post-néonatale pour les groupes d'âge <5 ans et ≥5 ans, en Tanzanie. Les causes de décès ont été assignées par un groupe de médecins en utilisant la Classification Internationale des Maladies, version numero 10. La mesure de la validité a été basée sur: les fractions de mortalité cause-spécifique, la sensibilité, la spécificité et la valeur prédictive positive. Les records de diagnostics médicaux ont été classifiés par degré d'incertitude et la sensibilité et spécificité ont été ajustées. Nous avons appliqué des critères d'évaluation de la performance de l'autopsie verbale à générer une vraie mortalité proportionnelle.

RÉSULTATS L'autopsie verbale a procuré des fractions précises de mortalité cause-spécifique dans différents groupes d'âge, pour 9 causes de décès: asphyxie à la naissance, complications intra-utérines, pneumonie, VIH/SIDA, malaria (adultes), tuberculose, maladies cérébro-vasculaires, blessures et causes maternelles directes. Les résultats pour 20 autres causes de décès approchaient le seuil de bonne performance.

CONCLUSION L'autopsie verbale a estimé de façon valable les fractions de mortalité cause-spécifiques pour des maladies importantes en terme de santé publique dans tous les groupes d'âge. Des validations supplémentaires sont nécessaires pour évaluer les raisons du manque de résultats positifs dans certaines conditions.

**mots clés** cause de décès, autopsie/*méthodes, mesures épidémiologiques, méthodes épidémiologiques, statistiques vitales, mortalité infantile

 © 2006 Blackwell Publishing Ltd

**Validez de las autopsias verbales en la determinación de las causas de muerte en Tanzania**

OBJETIVOS Validar el procedimiento de las autopsias verbales para su uso en el muestreo de registros vitales. La autopsia verbal es un método importante para obtener datos sobre mortalidad causa específica en lugares en donde la carga de enfermedad es mayor y los datos de mortalidad causa-específico no se recogen de forma rutinaria.

MÉTODO Se recogieron autopsias verbales e informes médicos de 3123 muertes, en Tanzania, para las siguientes edades: período peri -/neonatal, post-neonatal, menores de 5 años y mayores de 5 años. Las causas de muerte fueron asignadas por paneles médicos utilizando la Clasificación Internacional de Enfermedades, décima revisión. La validez se midió mediante fracciones de mortalidad causa-específica; sensibilidad, especificidad y valor predictivo positivo. Los diagnósticos de informes médicos fueron puntuados según el grado de incertidumbre y se ajustaron la sensibilidad y la especificidad. Se aplicaron criterios para evaluar el desempeño de la autopsia verbal en generar una mortalidad proporcional real.

RESULTADOS Las autopsias verbales produjeron fracciones acertadas de mortalidad causa-específica para nueve causas en diferentes grupos de edad: asfixia durante el parto; complicaciones intrauterinas; neumonía; VIH/SIDA; malaria (en adultos); tuberculosis; enfermedad cerebrovascular; lesiones; y causas maternas directas. Los resultados para otras 20 causas estuvieron cerca del umbral del buen desempeño.

CONCLUSIONES Las autopsias verbales estimaron, con fiabilidad, las fracciones de mortalidad causa específica para enfermedades con importancia para la salud pública en todos los grupos de edad. Se requiere una validación adicional para evaluar las razones por las cuales no se obtienen resultados positivos para ciertas condiciones.

**palabras clave** causa de muerte, autopsia/*métodos, mediciones epidemiológicas, métodos epidemiológicos, estadísticas vitales, mortalidad infantil

## Annex 1 ICD coding guidelines and criteria provided to doctor panel members

A formal course of instruction based on volume 2 of ICD-10 was provided to doctor panel members. The training provided practical instructions for cause of death attribution. The instruction was augmented with extensive practice on sample records and discussion of the problems.

In addition to the theoretical work, participants were given adequate opportunity to practice the skills obtained through extensive exercises both in cause of death assignment and clinical cause of death coding using ICD-10. These skills were applied to sample VA and MR data. These were then subjected to a quality analysis, after which common errors and mistakes were identified and reviewed.

In the context of the VA validation study, and after thorough review of the general principles, selection rules, modification rules and other ICD rules (vol. 2; pp. 32–68), the research team found it necessary to introduce certain modifications to the ICD rules and instructions. In particular, stillbirths with known/unknown underlying maternal cause as well as circumstances surrounding the external causes of injuries led to discrepancies in assigning a cause of death and reaching a final, valid underlying cause of death. The modifications and special rules applied in this validation study are noted here.

**Special cases and exceptions to ICD-10 cause of death coding:** *Stillbirths, fetal deaths, intrauterine fetal deaths*

- The terms 'stillbirth' or 'fetal death' will be used (not 'intra-uterine fetal death') as the death is recorded after birth of the dead fetus and not while *in utero*.
- There is no difference in cause of death between stillbirths recorded as 'fresh' or 'macerated'.
- When a fetal death or stillbirth can be attributed to a particular cause (e.g. antepartum haemorrhage; maternal infection; eclampsia or pre-eclampsia), the cause of death is recorded as 'stillbirth' (fetal death) due to the appropriate obstetric cause taken from the 'O' series of blocks in Chapter XV.
- P95 – 'Fetal death of unspecified cause/dead born fetus not otherwise specified/stillbirth not otherwise specified' is to be used where there is no possible cause for the stillbirth (from the history).

*Malaria (B50–B54)*: VA is unable to support the inclusion of some forms of malaria known to be of public health importance, and clinical notes also frequently lack adequate confirmatory information. For example, classification of '*Plasmodium falciparum* malaria with cerebral complications' (B50.0) requires microscopic confirmation of *P. falciparum*, which is unlikely to be found in verbal autopsy or MRs. Thus, strict adherence to ICD criteria would almost always preclude the use of B50.0 as this diagnosis cannot be made based on the symptoms and signs alone. Similarly, the use of B51, B52 and B53.0 and B53.1 all require specification of *Plasmodium* species, which was not encountered in over 3000 MRs examined for the study. Therefore, for the purposes of the validation study, the possible causes of death due to malaria were restricted to the following:

- B53.8: Parasitologically confirmed malaria, not otherwise specified; and
- B54: Unspecified malaria/clinically diagnosed malaria without parasitological confirmation.

*HIV disease (B20–B24)*: HIV disease can present with many complications and infections, each having its own

© 2006 Blackwell Publishing Ltd

unique cause of death. ICD-10 explicitly employs the combined categories of B20–B23 for optional use where it is not possible or desired to use multiple causes of death. This includes B20.0 'HIV disease resulting in mycobacterial infection/HIV disease resulting in tuberculosis'. However, considering the public health importance of tuberculosis, and in order to maintain uniformity of the cause of deaths assigned, the following guidelines were used for diagnosing HIV disease in the study:

- B20.0 – 'HIV disease resulting in mycobacterial infection/HIV disease resulting in tuberculosis' was given priority as the underlying cause of death if the history or findings indicate evidence of tuberculosis;
- B20.7 – 'HIV disease resulting in multiple infections' was used when there was evidence of more than a single infection in HIV (e.g. candidiasis, mycoses, parasitic diseases, etc.). The use of this code when there was evidence of multiple infections avoided assigning several causes of death for each type of associated infection; and
- B22.0 – 'HIV disease resulting in encephalopathy/HIV dementia' was used when there was history of confusion, dementia and loss of consciousness of more than 1 day or other central nervous system (CNS) manifestations such as stroke associated with HIV. This is a common presentation of terminal HIV disease; however, where there was evidence of tuberculosis infection/disease, the cause of death B20.0 has been used.
- In cases of HIV disease with only one infection identified (e.g. candidiasis only), then the appropriate four-digit ICD code was assigned (e.g. B20.4 – 'HIV disease resulting in candidiasis').
- When a case of HIV disease with TB presented with CNS manifestations, the cause of death 'HIV with encephalopathy' (B22.0) has been used, as the immediate cause of death followed by B20.0 – 'HIV disease resulting in tuberculosis' as the underlying cause.
- Where HIV presents with Kaposi sarcoma, this complication was not coded separately but was included under B20.7 in the multiple infection category, unless it appeared to be the sole complication, in which case B21.0 – 'HIV disease resulting in Kaposi's sarcoma' was used. This is because Kaposi sarcoma is multicentric and is regarded as a malignancy with a viral infectious origin.

*Paediatric HIV disease*: ICD-10 does not provide specific classification and cause of death of HIV disease in children. Because of difficulty in diagnosing HIV in children, the following guidelines have been used to assign cause of death in children who presented with HIV disease:

- Clinical symptoms suggesting HIV disease in the absence of other obvious causes of immune suppression (e.g. malnutrition); or
- Clinical symptoms suggesting HIV disease and a family and social history suggestive of HIV (e.g. parental death due to suspected HIV disease including cases where the child's mother was sick at the time of death of the child); or
- Clinical symptoms suggesting HIV disease and the attending doctor had requested an HIV test to confirm the diagnosis, regardless of whether the MR contained the results of the serology.

*Respiratory tuberculosis (A15–A16)*: Definitive diagnosis of tuberculosis can only be made where acid-fast bacilli (AFB) can be microscopically identified (typically from sputum). However, such information was not available in most of the MRs available for the study, and cannot be considered a reliable datum to elicit in a VA. Therefore, diagnosis of respiratory tuberculosis was made based on any of the following criteria:

- Sputum-positive for AFB; or
- Current history of taking antituberculosis drugs; or
- A chest X-ray suggestive of pulmonary tuberculosis and symptoms consistent with tuberculosis (e.g. chronic cough >1 month with or without blood; prolonged fever and weight loss); or
- History and symptoms suggestive of tuberculosis, but in the absence of other signs and symptoms suggestive of HIV disease resulting in either mycobacterial infection or other infectious process (ICD codes B20.0–B20.9).


© 2006 Blackwell Publishing Ltd