Algorithms for verbal autopsies: a validation study in Kenyan children

M.A. Quigley,[1] J.R.M. Armstrong Schellenberg,[2] & R.W. Snow[3]

The verbal autopsy (VA) questionnaire is a widely used method for collecting information on cause-specific mortality where the medical certification of deaths in childhood is incomplete. This paper discusses review by physicians and expert algorithms as approaches to ascribing cause of deaths from the VA questionnaire and proposes an alternative, data-derived approach.

In this validation study, the relatives of 295 children who had died in hospital were interviewed using a VA questionnaire. The children were assigned causes of death using data-derived algorithms obtained under logistic regression and using expert algorithms. For most causes of death, the data-derived algorithms and expert algorithms yielded similar levels of diagnostic accuracy. However, a data-derived algorithm for malaria gave a sensitivity of 71% (95% CI: 58–84%), which was significantly higher than the sensitivity of 47% obtained under an expert algorithm. The need for exploring this and other ways in which the VA technique can be improved are discussed. The implications of less-than-perfect sensitivity and specificity are explored using numerical examples. Misclassification bias should be taken into consideration when planning and evaluating epidemiological studies.

Introduction

Accurate information on the causes of mortality is necessary for effective planning and evaluation of health care programmes (*1–3*). In recent years, the verbal autopsy (VA) questionnaire has been widely used for collecting such information in situations where the medical certification of deaths in childhood is incomplete. Trained fieldworkers interview bereaved relatives using a structured questionnaire in order to elicit information on the symptoms their child experienced before death. The information from completed questionnaires is then summarized and interpreted to give a likely cause of death for each child. Probably the most common method for ascribing causes of death from VA questionnaires is when the completed questionnaires are reviewed by one or more physicians who ascribe probable causes of death (*4–10*). All parts of the questionnaire, particularly any open-ended sections, are thus incorporated into the diagnosis. However, large-scale surveys can prohibit the use of long, detailed questionnaires; furthermore, open-ended questions, asked by lay interviewers, may prove difficult in establishing case histories.

An alternative means of ascribing causes of death is to follow a set of pre-defined diagnostic criteria in an expert algorithm (*11–14*). In this case, the questionnaire consists of closed questions and yields only pre-coded information. Moreover, as the algorithm uses well-defined diagnostic criteria to ascribe causes of death, changes in cause-specific mortality may be compared over time or between different studies (*2*). Expert algorithms have been the subject of validation studies in the Philippines (170 deaths) (*11*) and in Namibia (135 deaths) (*13*). The accuracy of these algorithms was estimated by comparing the ascribed causes of death with the medically confirmed diagnoses of children who died in hospital. For deaths due to measles and malnutrition, which have signs and symptoms that are readily recognized by lay persons, the algorithms gave a relatively high sensitivity and specificity (Table 1). However, deaths due to other causes, most notably malaria and acute respiratory infection (ARI), were not assigned accurately in either study population.

Poor diagnostic performance of the VA technique may be due to shortcomings of the questionnaire or of the method used to ascribe the cause of death. Information on symptoms leading to a child's death may not be elicited because the symptoms are not easily recognized by lay persons, or are poorly recalled by relatives, or are not included in the questionnaire (*1, 15*). Furthermore, it may not be possible to discriminate between certain causes of

[1] Lecturer, Tropical Health Epidemiology Unit, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, England. Requests for reprints should be sent to this author.

[2] Lecturer, Tropical Health Epidemiology Unit, London School of Hygiene and Tropical Medicine, London, England.

[3] Wellcome Senior Fellow, CRC Research Unit, Kenya Medical Research Institute, Kilifi, Kenya, and Institute of Molecular Medicine, John Radcliffe Hospital, Oxford, England.

Reprint No. **5685**

Table 1: **Sensitivity and specificity of expert algorithms in validation studies in the Philippines (170 deaths) and Namibia (135 deaths)**

	Philippines[a]		Namibia[b]	
Cause of death	Sensitivity	Specificity	Sensitivity	Specificity
Malaria	—	—	45	87
Malnutrition	—	—	73	76
Measles	98	93	71	85
Acute respiratory infection	66	60	72	64
Gastroenteritis[c]	60	85	56	90
Gastroenteritis[c]	78	79	89	61

[a] Source: Kalter et al., 1990 (*11*).
[b] Source: Mobley et al., 1993 (*13*).
[c] Two separate algorithms.

death on the basis of signs and symptoms alone (*16*, *17*). The VA technique may be applied with high diagnostic accuracy in situations where each cause of death is always preceded by a unique set of signs and symptoms (*2*). This set of signs and symptoms may be found using standard statistical methods. For given data, an algorithm is then obtained which results in the highest possible sensitivity and specificity. These techniques have been used in diagnostic methods and screening (*18*, *19*). Clearly, algorithms can only be derived when the medically confirmed causes of death are available, although they may then be used more widely. We report here on algorithms derived using logistic regression and compare their diagnostic accuracy with that from several expert algorithms.

Methods

Study population and prospective surveillance of hospital deaths

A prospective VA study was conducted between May 1989 and April 1993 at Kilifi District Hospital, 60km north of Mombasa on the Kenyan coast. Details of the study population and the VA methods have been described elsewhere (*8*). In brief, all children admitted to the paediatric ward at Kilifi District Hospital were examined on admission and during their stay in hospital. Full clinical examinations and laboratory investigations were carried out on each child and recorded on a standard proforma. If the child died while in the paediatric ward, the detailed clinical notes and laboratory results were reviewed by the examining physicians who ascribed one or two causes of death that were coded according to the Ninth Revision of the International Classification of Diseases (ICD). When there was insufficient evidence to ascribe a cause of death confidently, the cause was classified as undetermined. The ICD codes were used to form the following nine groups: malaria, malnutrition, measles, acute respiratory infection (ARI), gastroenteritis, meningtis, accident, undetermined, and other.

VA questionnaires

All the medically confirmed hospital deaths were followed up by one of two senior field staff, who were fluent in local languages and trained in the sensitive conduct of interviews with the bereaved relatives or guardians using the pre-tested, unambiguous vernacular. Interviews were conducted with the person who had closest contact with the child during the terminal illness, usually the mother. Information was obtained on the child's age and duration of illness and on the following signs and symptoms occurring prior to the child's death: diarrhoea, fever, fits, indrawing of chest, severe coughing, neck stiffness and vomiting. Specific information was obtained on the occurrence of accidents, measles, kwashiorkor and marasmus. Fieldworkers were unaware of the medically confirmed cause of death.

Expert algorithms

Causes of death were assigned from the questionnaires according to particular sets of diagnostic criteria, that is, using expert algorithms. We selected algorithms for malaria, measles, ARI, meningitis and accidents that were used in previous studies (*11*, *13*) or have been recommended for use in VA studies (*20*). We were not able to use algorithms recommended for malnutrition (*13*, *20*) and gastroenteritis (*11*, *13*, *20*) because we had insufficient information on the signs and symptoms to apply them: information on weight loss and the number of stools in a 24-hour period was not collected, because such information collected during pilot studies proved to be unreliable (*8*). A separate and independent algorithm was associated with each cause of death, thus allowing multiple causes of death per child.

Data-derived algorithms

Our objective was to identify, for each medically confirmed cause of death, the set of signs and symptoms as recalled by the bereaved relatives which best discriminated between that cause of death and all other deaths. By ascribing each cause of death separately rather than simultaneously, any child could have had more than one cause of death, as with the expert algorithm process described above. As most of the recalled signs and symptoms formed categorical variables, potential discriminant functions were identified using logistic regression rather than discriminant analysis. Logistic regression was carried out using SAS 6.04 (SAS Institute, Cary, NC, USA).

The completed questionnaires and their associated, medically confirmed causes of death were entered into a database. An outcome variable was created for each medically confirmed cause of death, the variable being equal to 1 for a child with that cause of death and 0 otherwise. Each child was randomly assigned to one of two groups. The children from the first group (sample 1, $n = 145$) were used to derive the algorithms and those from the second group (sample 2, $n = 150$) were used to validate the algorithms. For children in sample 1, each sign and symptom was cross-tabulated with each medically confirmed cause of death, to identify potential discriminating factors. If a sign or symptom yielded a sensitivity and specificity of at least 80% for a particular cause of death, this alone was used in the algorithm for that cause of death. Otherwise, signs and symptoms which were significantly associated with a particular cause of death (χ^2, $P < 0.10$) were assessed in multiple logistic regression models. Any logistic model in which each sign and symptom remained significant (Wald test, $P < 0.10$), after adjusting for all other factors in the model, gave a potential algorithm. For each such model, a score was obtained for every child by summing the coefficients of the model over the recalled signs and symptoms for that child. The children were then assigned to the cause of death for that model if they were below particular cut-off points. Different models and different cut-off points for the same model gave rise to many potential algorithms for any one cause of death. We selected an algorithm for each cause of death which correctly classified the highest number of children and had a reasonably high sensitivity and specificity. We arbitrarily chose 70% as the minimum sensitivity and specificity for an algorithm. If no such algorithm could be found, a lower minimum value, say 60%, for sensitivity and specificity was considered until an algorithm could be found.

Validation

The two methods were validated by comparing the causes of death assigned by the expert or data-derived algorithm with the medically confirmed diagnoses, while the diagnostic accuracy was assessed using sensitivity and specificity. The data-derived algorithms were validated on sample 2, so as to minimize the bias in the estimates of sensitivity and specificity. The expert algorithms were validated using the whole population and, for purposes of comparison, using sample 2. The diagnostic accuracy of both methods was compared, although the relatively small number of children with each cause of death meant that formal statistical tests, particularly comparisons of sensitivity, had low power. For example, the study would have reasonable power (80% or more) to detect a 15% difference in any two specificities, but would only have sufficient power to detect differences as large as 25% between sensitivities in common causes of death (say, about 50 deaths).

Results

The relatives of 295 post-neonatal children under the age of five years who died at Kilifi District Hospital between May 1989 and April 1993 were interviewed. Forty-eight children (15%) had two medically confirmed causes of death, giving a total of 343 causes of death. Table 2 shows the number of children with each cause of death in the study population in total and in the two random samples. The most common causes of death in the study population were malaria (30%), malnutrition (21%) and measles (17%). Moreover, 190 (64%) of all child deaths were due to one or more of these three causes of death. The proportions of children with these and the other causes of death were similar in the two random samples.

The occurrence of each sign and symptom on the VA questionnaires was examined. Of the 295 children in the study, 259 (88%) had fever, 154 (52%) had diarrhoea, 148 (50%) had vomiting, 147 (50%) had chest indrawing, 95 (32%) had fits, 59 (20%) had kwashiorkor, 59 (20%) had measles, 53 (18%) had a severe cough, 45 (15%) had marasmus, 41 (14%) had neck stiffness, and 12 (4%) had suffered an accident.

Table 3 shows the sensitivity and specificity of various expert algorithms estimated using the whole

Table 2: **Number of children with each cause of death among 295 post-neonatal under-5-year-olds who died in Kilifi District Hospital between May 1989 and April 1993, and among two random samples of these children**

Cause of death	No. in both samples (*n* = 295)	No. in sample 1 (*n* = 145)	No. in sample 2 (*n* = 150)
Malaria	87 (29.5)[a]	39 (26.9)	48 (32.0)
Malnutrition	63 (21.4)	33 (22.8)	30 (20.0)
Measles	49 (16.6)	24 (16.6)	25 (16.7)
Acute respiratory infection	33 (11.2)	19 (13.1)	14 (9.3)
Gastroenteritis	22 (7.4)	10 (6.9)	12 (8.0)
Meningitis	12 (4.1)	6 (4.0)	6 (4.0)
Accidents	9 (3.0)	6 (4.0)	3 (2.0)
Other	47 (15.9)	24 (16.6)	23 (15.3)
Undetermined	21 (7.1)	8 (5.5)	13 (8.7)

[a] Figures in parentheses are percentages.

study population and sample 2. All of the expert algorithms yielded a specificity of at least 73%. The algorithms for measles and accidents had equally high sensitivities. The algorithms for most other causes of death yielded poor sensitivities. In particular, the algorithm for malaria, the most common cause of death, yielded a sensitivity of 47% (95% CI: 37–58).

Table 4 shows the sensitivity and specificity of the data-derived algorithms estimated using sample 2. The algorithms for malnutrition, measles and accidents were based on the answers to specific questions using the local terms for kwashiorkor, measles, and accidents respectively, as these questions alone yielded high (at least 80%) sensitivity and specificity. The VA questionnaire included a specific question on marasmus, but this had a lower sensitivity (46%) than the question on kwashiorkor. The algorithms for malnutrition, measles and accidents were highly sensitive and specific. No logistic regression model for ARI or meningitis yielded an algorithm with sensitivity above 50%. However, the data-derived algorithm for malaria yielded a reasonably high sensitivity (71%, 95% CI: 58–84) and specificity (80%, 95% CI: 73–88). The sensitivity of this data-derived algorithm for malaria was significantly higher ($P = 0.013$) than the sensitivity (47%) under the expert algorithm proposed by Mobley (*13*).

Discussion

In this hospital-based validation study, we have shown that data-derived diagnostic algorithms can give comparable, or better, diagnostic accuracy than expert algorithms. Both approaches gave consistently high specificity (usually >80%), and for deaths due to malnutrition, measles and accidents, the sensitivity was equally high. For malaria, however, our data-derived algorithm gave a sensitivity

Table 3: **Sensitivity and specificity of expert algorithms**

Cause of death	Expert algorithm	Both samples (*n* = 295) Sensitivity	Both samples (*n* = 295) Specificity	Sample 2 (*n* = 150) Sensitivity	Sample 2 (*n* = 150) Specificity
Malaria	Fever *and* fits (*13*)[a]	47	79	52	83
Measles	Age > 4 months *and* duration > 3 days *and* rash *and* fever (*11*)[b,c], (*13*)[b,c]	84	96	88	96
Measles	Age > 3 months *and* duration of 4–89 days *and* rash *and* fever (*20*)[c,d]	84	96	88	97
Measles	Age > 4 months *and* rash (*13*)[c]	94	95	96	95
Acute respiratory infection	Duration > 1 day *and* cough *and* indrawing (*20*)[e]	24	91	36	92
Meningitis	Fever *and* fits (*20*)	67	73	67	74
Accident	Accident (*20*)	89	99	100	98

[a] Figures in italics, within parentheses, are the source references.
[b] Where the source used "rash & fever >3 days", we used "duration of illness >3 days & rash & fever".
[c] Where the source used "rash", we used the local term for measles.
[d] Where the source used "rash & fever for >3 days within 90 days of death", we used "duration of illness of 4–89 days & rash & fever".
[e] Where the source used "cough >1 day", we used "duration of illness >1 day & cough".

Table 4: **Sensitivity and specificity of data-derived algorithms**

Cause of death	Logistic regression algorithm derived using sample 1	Algorithm validated using sample 2	
		Sensitivity	Specificity
Malaria[a]	Duration < 4 days *and* no measles or Duration = 4–7 days *and* no measles *and* no diarrhoea	71	80
Malnutrition	Kwashiorkor	80	97
Measles	Measles	96	94
Acute respiratory infection	Age < 12 months *and* no measles	36	78
Meningitis	Age < 12 months *and* fits	50	90
Accident	Accident	100	98

[a] The following is a *single* algorithm for malaria:
Duration < 4 days *and* no measles
or
Duration = 4–7 days *and* no measles *and* no diarrhoea

of 71%, significantly higher than the expert algorithm. Both methods had low sensitivity (<40%) for ARI.

The diagnostic accuracy of the VA technique in ascribing causes of death due to malnutrition, measles and accidents has been demonstrated in several studies (*8*, *9*, *11*). Deaths due to these causes are preceded by symptoms or events, which are easily recognized and recalled by lay persons and which are specific to that cause of death. Several studies have also demonstrated the poor diagnostic performance of the VA technique for ascribing causes of death due to ARI (*8*, *9*, *11–13*).

Few studies have validated the use of the VA technique in ascribing deaths due to malaria. In Namibia (*13*), an algorithm based on the occurrence of fever and convulsions or loss of consciousness was reasonably accurate for cerebral malaria (72% sensitivity and 85% specificity) but the same algorithm, used on the same population, had poor sensitivity (45%) for all deaths associated with malaria parasitaemia. In our study, the expert algorithm used by Mobley et al. (*13*) yielded a sensitivity of 47% for malaria deaths. However, a significantly higher sensitivity (71%) was found using a logistic regression algorithm which classified a cause of death as malaria if the duration of illness was less than 4 days with no measles, or if the duration was between 4 and 7 days with no measles and no diarrhoea. This algorithm was derived using those signs and symptoms which best discriminated malaria deaths from all other deaths in our study, and these are not necessarily the signs and symptoms deemed by physicians as essential, confirmatory or supportive in diagnosing malaria. For example, fever may be regarded as *essential* in the diagnosis of malaria, but in our study, it had *poor discriminating power* because 93% of all malaria deaths and 86% of non-malaria deaths had fever. Diarrhoea had high discriminating power in our study because 74% of all malaria deaths and only 37% of non-malaria deaths had no diarrhoea. Measles also had high discriminating power in our study because measles was a common cause of death which rarely occurred with malaria. Clearly, our algorithm for malaria may result in lower diagnostic accuracy if applied in a setting where deaths due to measles are less common.

The poor diagnostic performance of the VA technique may be due to shortcomings of the questionnaire. A more likely explanation for the poor diagnostic performance of the VA is that it is difficult to discriminate between many causes of death on the basis of recalled signs and symptoms alone (*16*, *17*). For example, the consistently poor diagnostic accuracy for ARI may be due to the overlap in ARI symptoms with those from other diseases, most notably malaria. Similarly, the poor diagnostic accuracy for meningitis may in part be attributed to the overlap in meningitis symptoms with those of cerebral malaria.

For the data-derived algorithms, a further explanation for poor diagnostic accuracy might be a lack of discriminating power of logistic regression.

Table 5: **Effects of misclassification bias on the estimated values of the proportion of deaths due to a particular cause and the protective efficacy (PE) of an intervention**

Cause of death	True deaths (%)	Sensitivity (%)	Specificity (%)	Estimated deaths (%)	Estimated protective efficacy
Accident	3	90	90	12	22% of true PE
Malnutrition	21	90	90	27	70% of true PE
Meningitis	4	70	80	22	13% of true PE
Malaria	30	70	80	35	60% of true PE

Other statistical techniques such as decision trees, Bayesian classification methods and expert systems might yield more powerful discriminant functions. However, these techniques would result in a *single*, complex algorithm for all causes of death; they are more suitable when every cause of death category has large numbers, which is not usually the case, particularly with multiple causes of death. We allowed for multiple causes of death by using logistic regression to derive *separate* algorithms for each cause of death. Furthermore, by ensuring that each algorithm had a reasonably high specificity, there was no tendency for children to be assigned too many causes of death: the percentage of children in sample 2 with 0, 1 and 2 causes of death were 9%, 75% and 16% respectively, whereas the percentage assigned 0, 1, 2 and 3 causes of death by the data-derived algorithms were 12%, 51%, 33% and 4% respectively. Logistic regression has the further advantage that it is widely used by epidemiologists and available on most statistical software.

The most obvious limitation of a data-derived algorithm is that it can only be derived if the true or medically confirmed causes of death are known: this is not usually the case in areas where the VA is most needed. Furthermore, results obtained from hospital-based validation studies may not be applicable when the VA is used to ascertain causes of death in the surrounding communities (*1*). It would seem likely, however, in a study population such as ours, with good access to hospital and where about a third of deaths in children occur in hospital, that such results are indeed relevant in the surrounding communities. Any data-derived algorithm with sufficient diagnostic accuracy could be used to ascribe causes of deaths in future studies in the same population or in other populations. Clearly, such algorithms would need to be validated periodically and modified, in order to take account of changing mortality patterns associated with seasonality, epidemics, interventions and the emergence of new diseases.

Accurate information on causes of mortality is essential when estimating cause-specific mortality rates. Misclassification of the causes of mortality may lead to inaccurate estimates of mortality rates (*8*).[a] If the true proportion of deaths due to a particular cause is denoted by m, then m is estimated under the VA technique as follows:[a]

$$m = m \text{ sens} + (1 - m)(1 - \text{spec})$$

The effect of less than 100% sensitivity and specificity on estimates of cause-specific mortality rates may vary, but usually the proportion of deaths due to a particular cause will be overestimated. Another potential use of the VA is to measure the protective efficacy of community-based interventions using the ratio of two mortality rates. If a cause of death is ascribed with a specificity of less than 100%, the efficacy will generally be underestimated. If sensitivity is also less than 100%, the power of the study will be reduced. Using the VA, efficacy is estimated as:

$$\text{efficacy} = m \text{ sens efficacy}/[m \text{ sens} + (1 - m)(1 - \text{spec})]$$

Table 5 demonstrates the effect of misclassification bias on *estimated* cause-specific mortality and efficacy of a new intervention for given values of true cause-specific mortality, sensitivity and specificity. For causes of death due to measles, malnutrition and accidents, an estimate of 90% for the sensitivity and specificity is representative of the results of this study and others (*8, 9, 11*). However, for diseases such as malaria, ARI, gastroenteritis and meningitis, a sensitivity of 70% and a specificity of 80% is the best that we can currently expect of the VA technique. Under the scenarios given in Table 5, cause-specific mortality is overestimated and efficacy is underestimated. The misclassification bias is greatest for relatively rare causes of death with low sensitivity and specificity. Even for a common cause of death such as malnutrition, with a high sensitivity and specificity (90%), the true proportion of deaths due to this cause (21%) would be overestimated

[a] **Maude GH, Ross DA.** *The effect of different cause-of-death structures on the operating characteristics and sample size requirements of studies using verbal autopsies to determine cause-specific mortality rates in children.* Unpublished report for WHO, 1994.

(27%) and the efficacy would be estimated as only 70% of its true value.

The VA remains the only method of estimating cause-specific mortality in the absence of routinely and reliably collected mortality data. Consequently, there is an urgent need to explore the ways in which the VA questionnaire can be improved, particularly for ascertaining causes of death due to malaria and ARI. Given the improved diagnostic accuracy obtained using our data-derived algorithm for malaria, logistic regression and other methods for deriving algorithms should be explored in a wider setting. But until the diagnostic accuracy of the VA technique is improved, the implications of misclassification bias need to be taken into consideration when planning and evaluating epidemiological studies.

Acknowledgements

This paper is published with the permission of the Director of the Kenya Medical Research Institute (KEMRI). We thank Dr F. Pereira, Mr C. Frost, the staff of the KEMRI unit at Kilifi, especially Dr K. Marsh and Dr N. Peshu, and the staff of the Kilifi District Hospital. This study was supported by funds from the Wellcome Trust, the Overseas Development Administration (UK) Research Scheme R4672, the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases, and KEMRI. M.A. Quigley is funded by the Overseas Development Administration (UK) and the Wellcome Trust. JRM Armstrong Schellenberg is funded by the U.K. Medical Research Council. R.W. Snow is a Wellcome Senior Fellow in basic biomedical sciences.

Résumé

Algorithmes d'autopsie verbale: étude de validation sur des enfants du Kénya

Les questionnaires d'autopsie verbale sont largement utilisés pour recueillir des informations sur la mortalité par causes lorsque le certificat de décès des enfants est incomplet. Cette méthode permet un diagnostic précis dans le cas où chaque cause de décès est toujours précédée d'un ensemble caractéristique de signes et de symptômes. Dans la présente étude de validation, on a appliqué une méthode de régression logistique pour tâcher d'identifier un tel ensemble de signes et symptômes et comparé l'exactitude du diagnostic de ces algorithmes fondés sur les données à celles de différents algorithmes établis par des experts.

On a interviewé les parents de 295 enfants morts à l'hôpital en leur demandant de répondre à un questionnaire d'autopsie verbale. La cause du décès a été déterminée à l'aide d'un algorithme fondé sur les données et de différents algorithmes d'experts. Dans la plupart des cas, la précision du diagnostic a été la même avec les deux types d'algorithmes. Avec un algorithme d'expert attribuant la cause du décès au paludisme en cas de fièvre et de convulsions, la sensibilité a été de 47%. Par contre, on a obtenu une sensibilité nettement plus élevée (75%, IC à 95%: 58–84%) avec un algorithme fondé sur les données, qui attribuait le décès au paludisme lorsque la durée de la maladie était inférieure à 4 jours en l'absence de rougeole, ou comprise entre 4 et 7 jours en l'absence de rougeole et de diarrhée. Dans la présente étude, 30% des décès étaient dus au paludisme. Si une cause aussi fréquente de décès était établie avec une sensibilité de 70% et une spécificité de 80%, le taux de mortalité attribuable à cette cause serait estimé à 35% au lieu de 30% et l'efficacité protectrice d'une intervention serait sous-estimée à 60% seulement de sa vraie valeur.

L'avantage d'un algorithme fondé sur les données est de reconnaître les signes et symptômes qui distinguent le mieux une cause particulière de décès de toutes les autres, et qui ne sont pas nécessairement ceux que les médecins associent à cette cause. La méthode souffre cependant d'une limite évidente: elle suppose que la cause véritable ou médicalement confirmée du décès soit connue, au moins au départ, ce qui est rarement le cas dans les régions où l'autopsie verbale serait le plus utile. Compte tenu de la plus grande précision du diagnostic obtenue dans le cas du paludisme à l'aide de notre algorithme fondé sur les données, la technique de régression logistique et d'autres méthodes d'établissement d'algorithmes devraient être évaluées dans un contexte plus large. Mais tant que la précision du diagnostic de l'autopsie verbale ne sera pas meilleure, il faudra tenir compte du biais introduit par les erreurs de classification dans la planification et l'évaluation des études épidémiologiques.

References

1. **Snow RW, Marsh K.** How useful are verbal autopsies to estimate childhood causes of death? *Health policy and planning*, 1992, **7**(1): 22–29.
2. Measurement of overall and cause-specific mortality in infants and children: Memorandum from a WHO/UNICEF meeting. *Bulletin of the World Health Organization*, 1994, **72**: 707–713.
3. **The World Bank**. *World development report 1993: investing in health.* New York, Oxford University Press, 1993.

4. **Greenwood BM et al.** Deaths in infancy and early childhood in a well-vaccinated rural West African population. *Annals of tropical paediatrics*, 1987, **7**: 91–99.
5. **Alonso PL et al.** The effect of insecticide-treated bed nets on mortality of Gambian children. *Lancet*, 1991, **337**: 1499–1502.
6. **Ghana VAST Study Team.** Vitamin A supplementation in northern Ghana: effects on clinic attendances, hospital admissions and child mortality. *Lancet*, 1993, **342**: 7–12.
7. **Bang AT et al.** Reduction in pneumonia mortality and total childhood mortality by means of community-based intervention trials in Gadchiroli, India. *Lancet*, 1990, **336**: 201–206.
8. **Snow RW et al.** Childhood deaths in Africa: uses and limitations of verbal autopsies. *Lancet*, 1992, **340**: 351–355.
9. **Sachdev HPS et al.**, cited in: Measurement of overall and cause-specific mortality in infants and children: Memorandum from a WHO/UNICEF meeting. *Bulletin of the World Health Organization*, 1994, **72**: 707–713
10. **Mirza NM et al.** Verbal autopsy: a tool for determining cause of death in a community. *East African medical journal*, 1990, **67**: 693–698.
11. **Kalter HD et al.** Validation of postmortem interviews to ascertain selected causes of deaths in children. *International journal of epidemiology*, 1990, **19**: 380–386.
12. **Osinski P**, cited in: Measurement of overall and cause-specific mortality in infants and children: Memorandum from a WHO/UNICEF meeting. *Bulletin of the World Health Organization*, 1994, **72**: 707–713.
13. **Mobley CC**, cited in: Ewbank DC, Gribble JN, eds. *Effects of health programs on child mortality in sub-Saharan Africa*. Washington, National Academic Press, 1993: 20.
14. **Pacqué-Margolis S et al.** Application of the verbal autopsy during a clinical trial. *Social science and medicine*, 1990, **31**(5): 585–591.
15. **Snow RW et al.** Maternal recall of symptoms associated with childhood deaths in rural East Africa. *International journal of epidemiology*, 1993, **22**: 667–683.
16. **O'Dempsey TJD et al.** Overlap in the clinical features of pneumonia and malaria in African children. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 1993, **87**: 662–665.
17. **Todd JE et al.** The limitations of verbal autopsy in a malaria-endemic region. *Annals of tropical paediatrics*, 1994, **14**: 31–36.
18. **Richens J et al.** An algorithm for the clinical differentiation of malaria and typhoid: a preliminary communication. *Papua New Guinea medical journal*, 1992, **35**: 298–302.
19. **Goldman LG et al.** A computer-derived protocol to aid in the diagnosis of emergency room patients with acute chest pain. *New England journal of medicine*, 1982, **307**: 588–596.
20. **Bang AT et al.** Diagnosis of causes of childhood deaths in developing countries by verbal autopsy: suggested criteria. *Bulletin of the World Health Organization*, 1992, **70**: 499–507.