

Verbal autopsy: current practices and challenges

Nadia Soleman,^a Daniel Chandramohan,^a & Kenji Shibuya^b

Abstract Cause-of-death data derived from verbal autopsy (VA) are increasingly used for health planning, priority setting, monitoring and evaluation in countries with incomplete or no vital registration systems. In some regions of the world it is the only method available to obtain estimates on the distribution of causes of death. Currently, the VA method is routinely used at over 35 sites, mainly in Africa and Asia. In this paper, we present an overview of the VA process and the results of a review of VA tools and operating procedures used at demographic surveillance sites and sample vital registration systems. We asked for information from 36 field sites about field-operating procedures and reviewed 18 verbal autopsy questionnaires and 10 cause-of-death lists used in 13 countries. The format and content of VA questionnaires, field-operating procedures, cause-of-death lists and the procedures to derive causes of death from VA process varied substantially among sites. We discuss the consequences of using varied methods and conclude that the VA tools and procedures must be standardized and reliable in order to make accurate national and international comparisons of VA data. We also highlight further steps needed in the development of a standard VA process.

Keywords Autopsy/methods; Interviews/methods; Questionnaires; Cause of death; Developing countries (*source: MeSH, NLM*).

Mots clés Autopsie/méthodes; Entretien/méthodes; Questionnaires; Cause décès; Pays en développement (*source: MeSH, INSERM*).

Palabras clave Autopsia/métodos; Entrevistas/métodos; Cuestionarios; Causa de muerte; Países en desarrollo (*fuentes: DeCS, BIREME*).

الكلمات المفتاحية: الصفة التشريحية، طرق الصفة التشريحية، المقابلات، طرق المقابلة، استبيانات، أسباب الوفاة، البلدان النامية. (المصدر: رؤوس الموضوعات الطبية، المكتب الإقليمي لشرق المتوسط)

Bulletin of the World Health Organization 2006;84:239-245.

Voir page 244 le résumé en français. En la página 244 figura un resumen en español.

يمكن الاطلاع على الملخص بالعربية في صفحة 244.

Introduction

Knowledge about the distribution of causes of death in populations is essential for public health planning, resource allocation and measuring the impact of interventions. However, particularly in high mortality settings, vital registration data are often missing, incomplete or inaccurate. Medically-certified cause-of-death data are available only for less than one-third of over 57 million deaths occurring worldwide annually.¹ Rapid improvement of poorly performing vital registration systems in many countries is not realistic. Although attaining good quality vital registration data should be a long-term goal, alternative methods of ascertaining and estimating cause-of-death distributions at the population level must be used in the interim.

Verbal autopsy (VA) has been widely used as a method of ascertaining

causes of death in children in places where the majority of deaths occur without medical supervision.² There has been a growing interest in the use of VA in the context of disease surveillance and sample registration systems, particularly for causes of death in adults.³ VA is an indirect method of ascertaining biomedical causes of death from information on symptoms, signs and circumstances preceding death, obtained from the deceased's caretakers. VA has been used not only to gather data on the cause-of-death structure of certain populations, but also in investigations of infectious disease outbreaks and risk factors for certain diseases, and in measuring the effect of public health interventions.⁴⁻⁶ Currently, over 35 Demographic Surveillance Sites (DSS) in 18 countries, the Sample Registration System (SRS) sites in India, and the Disease

Surveillance Points (DSP) system in China regularly use VA on a large scale, primarily to assess the cause-of-death structure of a defined population.⁷⁻⁹

A standard VA tool comprises a VA questionnaire, cause-of-death or mortality classification system, and diagnostic criteria (either expert or data-derived algorithms) for deriving causes of death. The VA process has several stages, and many factors can influence the cause-specific mortality fractions estimated through this process (Fig. 1). Although attempts have been made to standardize VA tools and procedures, the diversity of VA tools currently in use makes it difficult to compare data over time and place.^{10,11} The major objective of this paper is to critically review the current tools and practices in VA and to discuss options for further improvement of the methodology.

^a Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London, England. Correspondence to Nadia Soleman (email: nadia.soleman@lshtm.ac.uk).

^b Measurement and Health Information Systems, Evidence and Information for Policy, World Health Organization, 1211 Geneva, Switzerland.

Ref. No. 05-027003

(Submitted: 26 September 2005 – Final revised version received: 31 January 2006 – Accepted: 1 February 2006)

Methods

We used a number of sources to collect information on VA tools and procedures currently in use, including web-based searches, manual review of archives, and correspondence with VA experts and practitioners working with DSS and SRS sites. Online searches included Medline free-text and MeSH searches for VA literature published in peer-reviewed journals after 1992 as well as Popline database search to identify additional VA tools and validation studies. Manual searches included review of workshop reports and discussion notes archived at the London School of Hygiene and Tropical Medicine, England, and publicly available material from various international and nongovernmental organizations. And finally, we questioned 36 field sites regarding their field operating procedures for deriving diagnoses from questionnaires and training of interviewers. The details of this process are explained elsewhere.¹¹ In order to provide a more current picture of VA tools, we included only the tools that are currently in use. As such, we reviewed 18 VA questionnaires and 10 cause-of-death tabulation lists.

Results

Comparison of verbal autopsy tools VA questionnaires

The standard, format, wording and question sequence of the 18 VA questionnaires currently used in the field differed considerably. Fourteen implemented separate forms for adults, children and neonates, while four contained questions for all ages mixed on a single form. Fifteen questionnaires had both an open-ended section for recording a verbatim account of symptoms, signs and circumstances leading to death, and a closed section with filter questions on the symptomatology of the disease. However, despite these variations, the key filter questions on symptoms and signs were not substantially different.¹¹ Only the VA adult questionnaires used in the SRS site in India had a different approach emphasizing the narrative section in determining causes of death, while the SRS child and maternal VA questionnaires had a closed section on a few symptoms and signs.⁹

The variation in the structure of VA questionnaires may lead to different sensitivities and specificities of the tools as open questions require the respondent

to recall specifics, whereas closed questions require recognition, with more information likely to be recognized than recalled.¹³ However, one study showed that the sensitivity of VA using physician review for deriving neonatal causes of death from closed sections alone was lower than that from open-ended sections alone or open-ended plus closed sections of the VA questionnaire. Sensitivity of verbatim and the combination of verbatim and closed-section questions were comparable, although the latter was slightly more sensitive for a few causes of death in neonates.¹⁴

Administration of open-ended VA questions generally requires medical training to elicit appropriate symptoms and signs that are not reported spontaneously, whereas this is not necessary for the closed questions. Furthermore, deriving diagnosis by diagnostic algorithms, particularly using automated systems, is more complex from open-ended VA questionnaires than from closed ones. Typically, if a deceased person contacted a health facility during the illness leading to death, the information on the biomedical cause of illness and treatment, recognized and registered by the caretaker, will be captured in the open-ended section of the VA questionnaire. This information is expected to vary between sites depending on the coverage and use of health services. Thus, using this information to derive causes of death introduces variability in the accuracy of VA between sites. However, adding the information from the open-ended section to the closed section for deriving causes of neonatal deaths using a computerized algorithm did not increase the agreement between causes of death reached by the algorithm and by a panel of physicians.¹⁵ This observation may not be applicable to adult deaths as the duration of illness is often longer and the recognition of symptoms, signs and treatment is different than that of neonates.

Causes-of-death classification

The *International statistical classification of diseases and related health problems*, tenth revision (ICD-10) which is the mandatory level of coding for international reporting to the WHO mortality database, has 21 chapters and 2046 categories of diseases, syndromes, external causes or consequences of the external causes.¹⁶ Although all of these categories of causes of death can be diagnosed by clinical judgement and/or laboratory tests, it is impossible to define symptoms

and signs for diagnostic algorithms for the complete list of causes of death. Few currently operating VA systems use the entire list of ICD-10 codes.¹¹

However, most VA systems use a short list for deriving diagnoses from VA. We reviewed 10 short cause-of-death lists currently used in the field. Seven of these have separate sections for children and adults. The number of causes included in the section for children ranged from 4 to 120, and for adults from 53 to 142. Three lists combine all ages, with the number of causes ranging from 32 to 57.

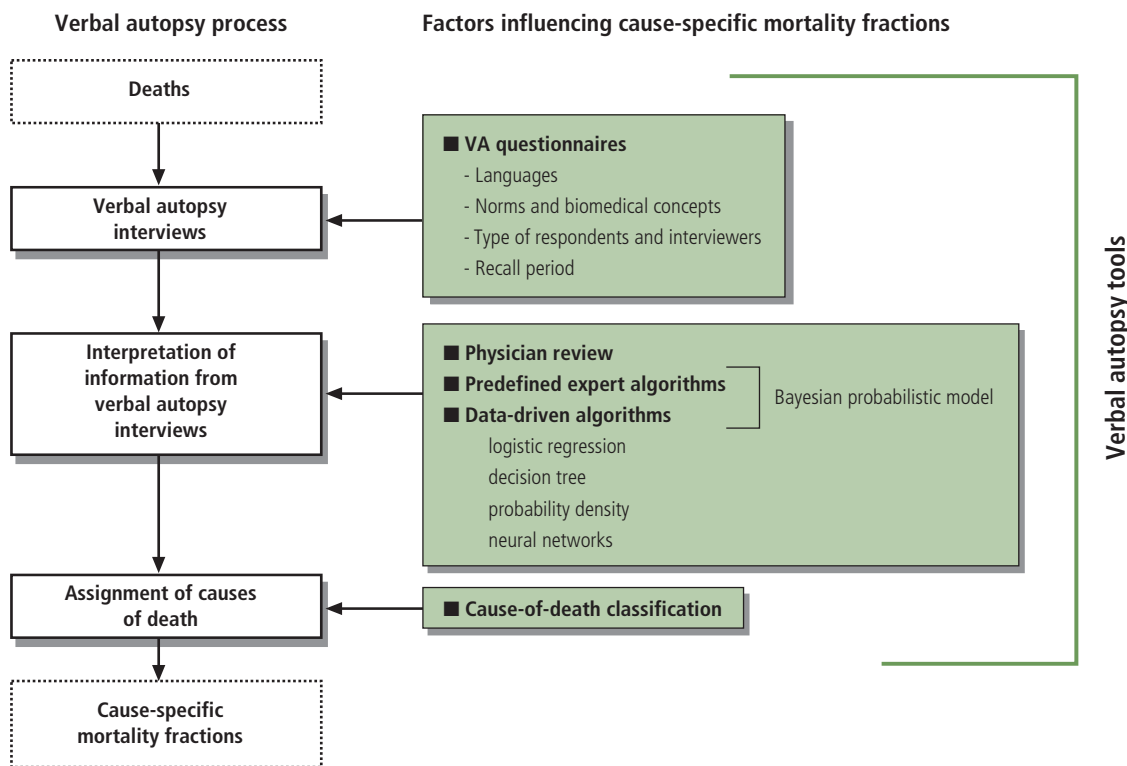
The structure of the cause-of-death lists currently used in the field varied: three had free listing of causes of death without any subgroupings; four grouped causes of death by organ system, consistent with ICD-10; and six grouped causes of death by pathophysiological mechanisms. Diagnosis of subgroups of diseases by VA is likely to be more accurate than diagnosis of specific diseases. For example, a diagnosis of infectious and parasitic diseases derived from VA is likely to be more accurate than a specific diagnosis of malaria. Furthermore, misclassification of malaria as pneumonia will not affect the accuracy of estimates of mortality due to infectious and parasitic diseases if all infectious diseases are included in a subgroup.

In ICD-10 some infections of specific organs are grouped under their respective organ systems. For example pneumonia is included under respiratory disorders and meningitis under nervous system disorders. Although causes of death can be regrouped for estimating subgroup level mortality, diagnostic algorithms to derive subgroup level causes of death such as infectious and parasitic diseases are difficult to derive unless all infectious diseases are grouped under this subgroup.

Algorithms to derive causes of death

Algorithms map diagnostic criteria in order to provide a systematic means of deriving cause of death from VA, and also aid in the development of the VA questionnaire. They increase the reliability of VA tools and allow automation of the cause-of-death coding process. However, standardized algorithms, validated in some epidemiological settings, are available only for neonatal and childhood deaths.² Attempts have been made to develop algorithms for selected adult causes of death.^{17,18}

Fig. 1. Verbal autopsy process and factors influencing cause-specific mortality fractions



WHO 06.18

There is, however, a substantial gap between VA theory and practice as most field sites do not currently employ diagnostic criteria for deriving causes of death from VA. Diagnostic algorithms are available at only two sites, but their use in deriving causes of death is obligatory at only one site, and the algorithms are not fully developed at the other. A third site is developing algorithms for implementation using handheld personal digital assistants (PDAs).

The most widely-used approach to derive causes of death from VA is physician review, in which a panel of physicians assigns cause of death based on clinical judgement of the information from the VA questionnaire. Typically a cause of death is reached if two physicians agree on an underlying cause. Thirteen sites report using this approach (23 sites did not specify the method used to derive causes of death from VA questionnaires).

Field assessment of the application of verbal autopsy tools

Many factors influence the validity of a VA tool — not just those inherent in the tool or affected by prevalence of diseases and causes of death, but also issues with the operational process of collecting and

coding VA data. This section summarizes the findings of the review of the practices of the VA application in the field.

Interviewers

The educational background of VA interviewers varied between sites. Six sites reported using medical professionals (medical officers, nurses, medical assistants), and five sites employed people with secondary education to conduct VA interviews. The remaining 25 sites did not report the characteristics of VA interviewers. All VA interviewers, particularly the non-medical ones, underwent training in interview techniques. VA interviews may cause emotional stress to some bereaved relatives, making counselling techniques a valuable training component, though not yet a common one.¹⁹

Some experts believe that medically-trained interviewers more accurately determine signs and symptoms of the deceased from VA interviews.²⁰ Others believe that medical knowledge may bias the result towards certain causes of death familiar to the interviewer. Several studies suggest that well-trained lay people can obtain accurate information when using culturally and linguistically appropriate questionnaires.^{18,21,22} Large

numbers of interviewers may produce diverse results, and the interviewer's gender and ethnic background in relation to respondents can also influence the outcome.²⁰ The choice of interviewer should be adapted to the local community, but the characteristics and training of interviewers used in various sites should be standardized.

Respondents

Most sites reported identifying as the respondent a relative who had taken care of the deceased during the final illness. However, the process of identifying an appropriate respondent is not formalized. Few sites reported interviewing friends or neighbours if a caretaker was not available.

There is limited information on the effect of respondents' characteristics on the accuracy of VA tools. One study that examined the effect of age, sex, relationship and language of the respondents found no significant effect of these variables.¹⁸ However, the accuracy of the verbal autopsy tool improved if the respondents had taken care of the deceased during the final illness. Cultural and societal factors must be taken into account when choosing the most appropriate respondent. For example, female respondents may be preferred for

a maternal mortality survey.²³ However, one study reported societal constraints on women reporting intimate details of the deceased's gynaecological history to a male interviewer without permission from the family head.²³ The process of identifying appropriate respondents for VA interview needs to be standardized taking cultural factors into account.

Recall period

Currently, a wide range of recall periods from the time of death is used in VA. Some perform interviews as soon as possible after death, while others visit the household of the deceased after a minimum of four weeks to allow an adequate mourning period. The maximum recall period varied between the sites from six months to an indefinite amount of time.

A long recall period is likely to impair a respondent's ability to recollect and report relevant information. However, inadequate time for mourning may cause distress and influence a respondent's willingness and ability to engage in a VA interview.²⁰ A recall period ranging from 1 to 12 months is generally thought to be acceptable.^{18,24} One validation study showed no significant effect on sensitivity or specificity using differences in recall period length of one to 21 months.¹⁸ The effects of recall period may differ between VAs concerning children and adults, and those investigating sudden or unexpected deaths. The differences in the recall period may influence the validity of the VA tool, and thus affect comparisons of VA data between sites. Further work is needed to define the acceptable recall period and to harmonize the recall period used between sites.

Language

In order to minimize misreporting, the questionnaire should ideally be written in the local language. In one validation study, using a language other than the mother tongue did not affect the sensitivity and specificity of VA.¹⁸ However, the agreement between VA and the gold standard estimates of cause-specific mortality fractions was stronger if respondents spoke the same language used in the questionnaire.¹⁸

Multiple translators should be involved in the translation of VA questionnaires, ideally medically-trained persons familiar with both health terms and the local language (Yoder S, Macro ORC, personal communication). Sometimes no written form exists for indigenous

languages, or several languages are spoken in a small area.^{18,23} In these circumstances, interviewers must be able to translate freely and incorporate local phraseology. In addition to the nuances of language, local concepts of health and disease may differ considerably between cultures. Questionnaires should also be field-tested in order to gain information to optimize layout, language and local biomedical concepts.

Analytical challenges Deriving causes of death from verbal autopsy results

There are several approaches to derive cause of death from VA: physician review, predefined expert algorithms, and data-driven algorithms. The most widely-used approach is physician review. The validity of physician review of VA has been tested in children and adults, and shown to have reasonable sensitivity and specificity for selected causes of death.^{2,11,14,17,21,22,25–27} However, the repeatability of causes of death derived by physician review is low.²⁸ Although inter-observer agreement is shown to be high in some studies, this may simply reflect consistency in physicians' prior knowledge of the local epidemiology.²⁸ This technique tends to reach a single cause even if a death is very likely due to multiple causes. Only a minority of VAs (13%) assigned more than one cause of death by physicians, while 25% of these deaths had two causes recorded in hospital.²⁹ In a study comparing physician review to algorithm-based cause-of-death assignment, only 11% were assigned more than one cause of death by physician review, while 58% of deaths were assigned multiple causes by algorithms.¹⁵ Ascertaining causes of large numbers of deaths by physician review is both time-consuming and cost-ineffective.³⁰

Algorithms would increase the reliability of VA tools and allow automation of the cause-of-death coding process. Expert algorithms are predefined diagnostic criteria agreed to by a panel of physicians. This approach overcomes the inconsistencies of physician review, and can reduce the cost and time needed for deriving diagnoses from VA. The validity of expert algorithms varies, and for many adult causes of death they do not perform as well as physician review.³¹

A Bayesian approach to defining the probability of a given cause of death in the presence of a particular symptom or sign could improve the performance

of expert algorithms.^{30,32} A model for selected causes of death using this approach showed a 90% concordance with causes of death derived by physician review.³² However, this approach has not been validated against a gold standard such as diagnoses from hospital records. (In some cases medical records are subject to bias and the quality is disputable — therefore some VA experts prefer to call medical records "reference standard" instead of "gold standard".¹¹)

Data-derived algorithms are another potential alternative to expert algorithms. The choice between the various analytical techniques to derive causes of death based on linear and other discriminant techniques (logistic regression), probability density estimation, and decision tree and rule-based methods (including artificial neural networks) depends on the intended use.³³ The arguments in favour of data-derived methods include their relatively low cost and potentially high reliability and consistency over time and between sites. Although the reported validity of data-derived algorithms was comparable to physician review in some studies, the same datasets were used both to generate and validate the algorithms.³⁴ Arguably, the validity of data-derived algorithms is underestimated because these use only the closed-ended sections of the VA questionnaire, while physicians usually use both closed- and open-ended responses. Physicians may also use information on drugs and reported hospital diagnoses when deriving causes of death. If data-derived algorithms are to be applied on a large scale, further research and validation against external data sets are needed.

Single or multiple causes of death

Using multiple rather than single causes of death probably more accurately reflects the interaction of different diseases that lead to death.¹⁰ For instance, if a fatally ill child suffered from diarrhoea and an acute lower respiratory infection, it is likely that it was the combination of the two that ultimately led to death — curing or preventing one may have prevented the death.² To count only one cause of death would distort mortality estimates and hence underestimate potential gains from health interventions.¹⁰ This is particularly true among children and older age groups in which co-morbidity is common.

On the other hand, the definition of the underlying cause of death and rules for classifying causes of death into under-

lying and contributory causes are defined in ICD-10.¹⁶ The cause-specific mortality fraction should be primarily based on the underlying cause of death as defined in ICD-10. Assessment of multiple causes of deaths based on underlying and contributing causes of death would still be possible. However, the cause-of-death categories used by VA studies have been inconsistent. While some classify according to the ICD-10 definition, others use classes such as primary or secondary, main or underlying cause of death. The method of classifying cause of death should be harmonized between VA systems according to the ICD-10 rules.¹⁶

Accuracy of the mortality data from VA systems

Several studies have attempted to assess the validity of VA tools.^{14,17,21,22,25,27,35} All validation studies except one have used causes of death based on medical records as the “gold standard”.²⁵ Most studies have estimated the sensitivity and specificity of VA, but few studies have assessed the agreement in the cause-specific mortality fraction (CSMF) between the gold standard and VA diagnoses. The sensitivity and specificity of VA varied by cause of death and between sites for the same causes.

VA is considered to have an acceptable level of diagnostic accuracy at the individual level if the sensitivity and specificity are at least 90%. At the population level, the VA is deemed to have reasonable diagnostic accuracy if sensitivity is at least 50%, specificity at least 90%, and the CSMF is within 20% of the true value.³¹ However these criteria of diagnostic accuracy are not uniformly regarded as acceptable.¹⁸

Low sensitivity and specificity does not necessarily imply that VA estimates of CSMF are over- or underestimates, as false positives and false negatives may cancel each other out, thereby not affecting the VA estimate.^{17,36} On the other hand, even in the presence of relatively high sensitivity and specificity, misclassification can result in serious over- or underestimates of CSMFs.^{36,37} This is because the accuracy of VA estimates depends not only on sensitivity and specificity, but also on the true underlying CSMF itself.³⁶ If the true specificity and sensitivity are known, the difference between the true CSMF and that estimated by VA can be calculated and the effect of misclassification on the VA estimates corrected. However, reported sensitivity and specificity

levels are estimated from hospital-based validation studies where the underlying causes-of-death structure is likely to be different from that of the general population.^{36,37} As VA is primarily needed in communities with restricted access to secondary or even primary care, applying values from hospital-based patient populations is inappropriate for correcting misclassification in areas with an unknown causes-of-death structure.¹⁷ Thus, at present it is not recommended to adjust for misclassification using sensitivities and specificities from validation studies.²

Measuring trends

Data from different geographical areas may not be comparable due to the use of heterogeneous VA tools, though this problem could be overcome by standardizing the tools and field procedures. Misclassification, and varying misclassification patterns across time and location, can mask or exaggerate geographical differences in cause-specific mortality and modify trends over time.^{37,38} The complex relationship between sensitivity, specificity and the underlying CSMF must be taken into account if trends in CSMF are to be measured by VA.^{36,37}

Maude & Ross introduced two models to correct the effect of misclassification on differences in cause-specific mortality estimates across time and space.³⁷ However, these models require values of true sensitivity and specificity for correcting the effects of misclassification. One study attempted to overcome the lack of data from the community level by using the average sensitivity and a logical constraint factor to the specificity of VA from seven validation studies to adjust VA estimates of malaria mortality from 28 different sites in sub-Saharan Africa.³⁸ This method is based on the following assumptions: 1) variation in the sensitivity of VA due to differences in the VA tool and procedures, malaria endemicity and distribution of causes of death is negligible. 2) the relationship between the specificity of VA and CSMF of malaria is linear and approaches 100% as malaria mortality reaches 0%. However, only seven data points were available to validate these assumptions. More validation studies in different epidemiological settings are needed.

Sample size is another challenge in measuring trends with VA. Detecting trends in CSMF requires large sample sizes, depending on the sensitivity and specificity of VA, CSMFs, and changes

in the CSMF of the prevailing causes of death. Many DSS do not cover populations that are large enough to detect significant differences in CSMF within reasonable time frames.^{37,38}

The way forward

Over the past decade the VA process has become widely used as a method to determine causes of death among various age groups in places where the majority of deaths occur without medical supervision. However, the various limitations of this method, discussed in this paper, must be overcome in order for VA data to be used for international comparisons. The introduction of a uniform and reliable method to derive causes of death and standardization of the VA questionnaires and field-operating procedures are important steps towards further improvement of the VA process.

There are many ongoing attempts to harmonize and collaborate. In a recent WHO consultation on VA,¹¹ the majority of experts agreed on the need for a standardized questionnaire with separate components for deaths of neonates, children and adults. The adult VA questionnaire should include all closed-ended questions needed to ascertain maternal deaths. The panel also agreed that a VA questionnaire with both an open-ended section and closed sections with filter questions is preferable, and information from both sections should be used to maximize the accuracy of VA. For comparisons across locations and over time, there is also a need for a standardized cause-of-death classification that lists globally important causes and relates them to ICD-10 codes.¹¹ A list of proposed cause-of-death classifications for VA can be obtained from the authors. Further consensus and agreement on VA tools, in particular algorithms to derive causes of death and the content of VA questionnaires, is urgently needed.

Conclusions

The major focus of cause-of-death analysis has shifted from global and regional estimates to national and subnational estimates for monitoring the progress towards the Millennium Development Goals and providing evidence for backing national policies. In many middle- and especially low-income countries, VA is the only method currently available to obtain estimates of the distribution of causes of death.

As our review of the currently used tools shows, the methodologies and standards in practice vary substantially. Although progress has been achieved in harmonizing the VA methodologies internationally, collaboration and harmonized efforts are needed to realize the full potential of this methodology in obtaining internationally comparable

causes-of-death data. ■

Acknowledgements

The authors would like to thank the participants of the WHO meeting on verbal autopsy tools in Talloires, France, 2-3 November 2004, for valuable inputs. Special thanks to Sue Piccolo and Suzanne Scheele for editing manu-

scripts and to Robert Jakob and Doris Ma Fat for their comments on the cause-of-death list.

Funding: This study was partly supported by the grant from the Ministry of Health, Labour and Welfare of Japan.

Competing interests: none declared.

Résumé

Autopsie verbale : pratiques actuelles et défis à surmonter

Pour la planification, la définition de priorités, la surveillance et les évaluations dans le domaine sanitaire, on fait de plus en plus appel aux données relatives aux causes de décès provenant des autopsies verbales (AV) dans les pays ne tenant pas de registres d'état civil ou disposant de registres incomplets. Dans certaines régions du monde, c'est même la seule méthode disponible pour obtenir des estimations de la distribution par causes des décès. Actuellement, l'autopsie verbale est couramment utilisée sur plus de 35 sites, principalement en Afrique et en Asie. Le présent article expose brièvement les procédures d'autopsie verbale et les résultats d'un bilan des outils d'autopsie verbale et des procédures opératoires utilisés sur les sites de surveillance démographique et par les systèmes d'enregistrement par sondage des faits d'état civil. Il a été demandé à 36 sites de terrain de fournir des informations

sur les procédures opératoires qu'ils appliquent : 18 questionnaires d'autopsie verbale et 10 listes de causes de décès utilisés dans 13 pays ont ainsi été examinés. Le format et le contenu des questionnaires pour AV, des procédures opératoires de terrain, des listes de causes de décès et des procédures employées pour extraire les causes de décès des AV varient de manière importante d'un site à l'autre. L'article analyse les conséquences de l'application de méthodes différentes et conclut à la nécessité de définir des outils et des procédures d'autopsie verbale standards et fiables pour permettre des comparaisons précises aux niveaux national et international des données d'AV. L'étude attire aussi l'attention sur les étapes nécessaires dans l'avenir au développement d'une procédure d'autopsie verbale standard.

Resumen

Autopsias verbales: práctica y retos

Los datos sobre causas de defunción obtenidos a partir de autopsias verbales (AV) son usados con creciente frecuencia con fines de planificación de la salud, establecimiento de prioridades, seguimiento y evaluación en los países con sistemas de registro civil incompletos o inexistentes. En algunas regiones del mundo es el único método disponible para poder estimar la distribución de las causas de mortalidad. Hoy día el método de las AV se utiliza sistemáticamente en más de 35 lugares, sobre todo en África y Asia. En este artículo presentamos un panorama general del sistema de las AV y los resultados de un análisis de los instrumentos de AV y los procedimientos operativos utilizados en los sitios de vigilancia demográfica y los sistemas de registro de estadísticas vitales por muestreo. Solicitamos información a 36 sitios sobre el

terreno acerca de los procedimientos operativos y examinamos 18 cuestionarios de autopsia verbal y 10 listas de causas de defunción usadas en 13 países. El formato y el contenido de los cuestionarios de AV, los procedimientos operativos sobre el terreno, las listas de las causas de defunción y los procedimientos empleados para calcular las causas de mortalidad a partir de las AV diferían sustancialmente de un sitio a otro. Analizamos las consecuencias de utilizar distintos métodos y llegamos a la conclusión de que es necesario normalizar los instrumentos y los procedimientos de AV y hacerlos más fiables si se desea hacer comparaciones más precisas de los datos de AV en los planos nacional e internacional. Ponemos de relieve, además, las medidas adicionales que habría que adoptar para desarrollar un procedimiento de AV normalizado.

ملخص

الصفة التشريحية اللفظية: الممارسات والتحديات الراهنة

موقعاً من المواقع الميدانية حول الإجراءات العملية واستعرضنا (18) استبياناً للصفات التشريحية اللفظية و(10) قوائم لأسباب الوفيات تستخدم في 13 بلداً. وتختلف أشكال ومضمون الاستبيانات والإجراءات الميدانية العملية، وقوائم أسباب الوفيات والإجراءات المتبعة لاستخلاص أسباب الوفيات من عملية الصفة التشريحية اللفظية اختلافاً ملحوظاً من موقع لآخر. وقد ناقشنا العواقب التي تترتب على استخدام طرق مختلفة وتوصلنا إلى استنتاج أن الأدوات والإجراءات المستخدمة في الصفة التشريحية اللفظية ينبغي أن تكون معيارية وموثوقة لكي توصلنا إلى مقارنات وطنية ودولية دقيقة لمعطيات الصفة التشريحية اللفظية. وقد استعرضنا أيضاً الخطوات الأخرى اللازمة لإعداد عملية معيارية للصفة التشريحية اللفظية.

يتزايد استخدام المعطيات المستخلصة من الصفة التشريحية اللفظية حول أسباب الوفيات في التخطيط الصحي وفي تقدير الأولويات والمراقبة والتقييم في البلدان التي تغيب فيها نظم السجلات المدنية أو تكون فيها غير مستكملة. وقد تكون الصفة التشريحية اللفظية هي الطريقة المتاحة الوحيدة في بعض مناطق العالم للحصول على تقديرات لأسباب الوفيات وعلى توزيعها. وتستخدم الصفة التشريحية اللفظية في الوقت الراهن بشكل روتيني في ما يزيد على 35 موقعاً، ولاسيما في آسيا وأفريقيا. ونستعرض في هذه الورقة نبذة عن الصفة التشريحية اللفظية ونتائج لمراجعة الأدوات المستخدمة فيها والإجراءات العملية المستخدمة في مواقع الرصد الديموغرافي وفي نظم السجلات المدنية المأخوذة كعينة. وقد جمعنا المعلومات من 36

References

1. World Health Organization. *World health report 2004 — changing history*. Geneva: WHO; 2004.
2. Anker M, Black RE, Coldham C, Kalter HD, Quigley M, Ross D, et al. *A standard verbal autopsy method for investigating causes of death in infants and children*. Geneva: World Health Organization; 1999.
3. Setel PW, Sankoh O, Rao C, Velkoff VA, Mathers C, Gonghuan Y, et al. Sample registration of vital events with verbal autopsy: a renewed commitment to measuring and monitoring vital statistics. *Bull World Health Organ* 2005; 83:611-7.
4. Andraghetti R, Bausch D, Formenty P, Lamunu M, Leitmeyer K, Mardel S, et al. *Investigating causes of death during an outbreak of Ebola virus haemorrhagic fever: draft verbal autopsy instrument*. Geneva: World Health Organization; 2003.
5. Pacqué-Margolis S, Pacqué M, Dukuly Z, Boateng J, Taylor HR. Application of the verbal autopsy during a clinical trial. *Soc Sci Med* 1990;31:585-91.
6. Telishevka M, Chenet L, McKee M. Towards an understanding of the high death rate among young people with diabetes in Ukraine. *Diabet Med* 2001; 18:3-9.
7. Yang G, Hu J, Rao KQ, Ma J, Rao C, Lopez AD. Mortality registration and surveillance in China: History, current situation and challenges. *Popul Health Metr* 2005;3:3.
8. INDEPTH Network. *An International Network of field sites with continuous Demographic Evaluation of Populations and Their Health in developing countries*. INDEPTH Network; 2005.
9. Centre for Global Health Research. *India sample registration system prospective study*. Toronto: University of Toronto, Centre for Global Health Research; 2005.
10. Bang AT, Bang RA. Diagnosis of causes of childhood deaths in developing countries by verbal autopsy: suggested criteria. *Bull World Health Organ* 1992;70:499-507.
11. World Health Organization. *WHO technical consultation on verbal autopsy tools*. Geneva: WHO; 2005.
12. Tanzania Ministry of Health. *The Policy Implications of Tanzania's Mortality Burden: Field Operations and Validation Studies. AMMP-2 Final Report Volume 3*. 2004, Dar es Salaam: Adult Morbidity and Mortality Project, Tanzania Ministry of Health, University of Newcastle upon Tyne, UK Department for International Development. Available from: http://www.ncl.ac.uk/ammp/site_files/public_html/finrep/index.html
13. Bennett AE, Ritchie K. *Questionnaires in medicine. A guide to their design and use*. Oxford: Oxford University Press; 1975.
14. Marsh DR, Sadruddin S, Fikree FF, Krishnan C, Darmstadt GL. Validation of verbal autopsy to determine the cause of 137 neonatal deaths in Karachi, Pakistan. *Paediatr Perinat Epidemiol* 2003;17:132-42.
15. Freeman JV, Christian P, Khatri SK, Adhikari RK, LeClerq SC, Katz J, et al. Evaluation of neonatal verbal autopsy using physician review versus algorithm-based cause-of-death assignment in rural Nepal. *Paediatr Perinat Epidemiol* 2005;19:323-31.
16. World Health Organization. *International Statistical classification of diseases and related health problems*, tenth revision, 2nd ed. Geneva: WHO; 1992.
17. Chandramohan D, Maude GH, Rodrigues LC, Hayes RJ. Verbal autopsies for adult deaths: their development and validation in a multicentre study. *Trop Med Int Health* 1998;3:436-46.
18. Chandramohan D. *Verbal autopsy tools for adult deaths*. [PhD Thesis]. London School of Hygiene and Tropical Medicine; 2001.
19. Chandramohan D, Soleman N, Shibuya K, Porter J. Ethical issues in the application of verbal autopsies in mortality surveillance systems. *Trop Med Int Health* 2005;10:1087-9.
20. Huong DL, Minh HV, Byass P. Applying verbal autopsy to determine cause of death in rural Vietnam. *Scand J Public Health Suppl* 2003;62:19-25.
21. Kahn K, Tollman SM, Garenne M, Gear JS. Validation and application of verbal autopsies in a rural area of South Africa. *Trop Med Int Health* 2000;5:824-31.
22. Kalter HD, Gray RH, Black RE, Gultiano SA. Validation of postmortem interviews to ascertain selected causes of death in children. *Int J Epidemiol*. 1990;19:380-6.
23. Hoj L, Stensballe J, Aaby P. Maternal mortality in Guinea-Bissau: the use of verbal autopsy in a multi-ethnic population. *Int J Epidemiol* 1999;28:70-6.
24. Mirza NM, Macharia WM, Wafula EM, Agwanda RO, Onyango FE. Verbal autopsy: a tool for determining cause of death in a community. *East Afr Med J* 1990;67:693-8.
25. Rodriguez L, Reyes H, Tome P, Ridaura C, Flores S, Guiscafre H. Validation of the verbal autopsy method to ascertain acute respiratory infection as cause of death. *Indian J Pediatr* 1998;65:579-84.
26. Chandramohan D, Rodrigues LC, Maude GH, Hayes RJ. The Validity of Verbal Autopsies for Assessing the Causes of Institutional Maternal Death. *Stud Fam Plann* 1998;29:414-22.
27. Mobley CC, Boerma JT, Titus S, Lohrke B, Shangula K, Black RE. Validation study of a verbal autopsy method for causes of childhood mortality in Namibia. *J Trop Pediatr* 1996;42:365-9.
28. Todd JE, De Francisco A, O'Dempsey TJ, Greenwood BM. The limitations of verbal autopsy in a malaria-endemic region. *Ann Trop Paediatr* 1994;14:31-6.
29. Snow RW, Armstrong JR, Forster D, Winstanley MT, Marsh VM, Newton CR, et al. Childhood deaths in Africa: uses and limitations of verbal autopsies. *Lancet* 1992;340:351-5.
30. Byass P, Huong DL, Minh HV. A probabilistic approach to interpreting verbal autopsies: methodology and preliminary validation in Vietnam. *Scand J Public Health*, 2003; *Suppl* 62:32-7.
31. Quigley MA, Chandramohan D, Rodrigues LC. Diagnostic accuracy of physician review, expert algorithms and data-derived algorithms in adult verbal autopsies. *Int J Epidemiol* 1999;28:1081-7.
32. Byass P, Fottrell E, Huong, DL, Berhane Y, Corrah T, Kahn K, et al. Refining a probabilistic model for interpreting verbal autopsy data. *Scand J Public Health*. In press.
33. Reeves BC, Quigley M. A review of data-derived methods for assigning causes of death from verbal autopsy data. *Int J Epidemiol* 1997;26:1080-9.
34. Quigley MA, Chandramohan D, Setel P, Binka F, Rodrigues LC. Validity of data-derived algorithms for ascertaining causes of adult death in two African sites using verbal autopsy. *Trop Med Int Health* 2000;5:33-9.
35. Coldham C, Ross D, Quigley M, Segura Z, Chandramohan D. Prospective validation of a standardized questionnaire for estimating childhood mortality and morbidity due to pneumonia and diarrhoea. *Trop Med Int Health* 2000;5:134-44.
36. Anker M. The effect of misclassification error on reported cause-specific mortality fractions from verbal autopsy. *Int J Epidemiol* 1997;26:1090-6.
37. Maude GH, Ross DA. The effect of different sensitivity, specificity and cause-specific mortality fractions on the estimation of differences in cause-specific mortality rates in children from studies using verbal autopsies. *Int J Epidemiol* 1997;26:1097-106.
38. Korenromp EL, Williams BG, Gouws E, Dye C, Snow RW. Measurement of trends in childhood malaria mortality in Africa: an assessment of progress toward targets based on verbal autopsy. *Lancet Infect Dis* 2003;3:349-58.