

RESEARCH

Open Access



# Validation studies of verbal autopsy methods: a systematic review

Buddhika P. K. Mahesh<sup>1†</sup>, John D. Hart<sup>1†</sup>, Ajay Acharya<sup>1</sup>, Hafizur Rahman Chowdhury<sup>1</sup>, Rohina Joshi<sup>2,3</sup>, Tim Adair<sup>1</sup> and Riley H. Hazard<sup>1\*</sup>

## Abstract

**Background:** Verbal autopsy (VA) has emerged as an increasingly popular technique to assign cause of death in parts of the world where the majority of deaths occur without proper medical certification. The purpose of this study was to examine the key characteristics of studies that have attempted to validate VA cause of death against an established cause of death.

**Methods:** A systematic review was conducted by searching the MEDLINE, EMBASE, Cochrane-library, and Scopus electronic databases. Included studies contained 1) a VA component, 2) a validation component, and 3) original analysis or re-analysis. Characteristics of VA studies were extracted. A total of 527 studies were assessed, and 481 studies screened to give 66 studies selected for data extraction.

**Results:** Sixty-six studies were included from multiple countries. Ten studies used an existing database. Sixteen studies used the World Health Organization VA questionnaire and 5 studies used the Population Health Metrics Research Consortium VA questionnaire. Physician certification was used in 36 studies and computer coded methods were used in 14 studies. Thirty-seven studies used high level comparator data with detailed laboratory investigations.

**Conclusion:** Most studies found VA to be an effective cause of death assignment method and compared VA cause of death to a high-quality established cause of death. Nonetheless, there were inconsistencies in the methodologies of the validation studies, and many used poor quality comparison cause of death data. Future VA validation studies should adhere to consistent methodological criteria so that policymakers can easily interpret the findings to select the most appropriate VA method.

**Prospero Registration:** CRD42020186886.

**Keywords:** Verbal autopsy, Verbal autopsy mode, Validation, Cause of death, Civil registration and vital statistics

## Background

Accurate and reliable mortality data are crucial for making informed decisions about public health policy and interventions [1]. However, globally only one in four

countries have well-functioning death registration systems and only one in three deaths are assigned a specific cause with policy utility [2–4]. In resource limited settings, where the majority of deaths occur out-of-hospital, traditional autopsy or cause of death certification by medical doctors is not practical for cause of death determination [5]. Verbal autopsy (VA) has emerged as a solution for collecting information on causes of death in low-income countries [1].

VA involves a trained interviewer administering a structured questionnaire to a carefully selected respondent who was with the deceased during the final illness [1].

<sup>†</sup>Buddhika P. K. Mahesh and John D. Hart are joint first authors.

\*Correspondence:

Riley H. Hazard  
riley.hazard@unimelb.edu.au

<sup>1</sup> Melbourne School of Population and Global Health, The University of Melbourne, Carlton, Victoria, Australia

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

<sup>3</sup> School of Population Health, University of New South Wales, Sydney, Australia



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Historically, physicians evaluated VA responses to assign a cause of death, but more recently there have been several computer-driven automated algorithms that assign cause of death based on VA [6–8]. The relative performance of these different diagnostic algorithms has also been evaluated yet remains controversial [9].

VA has been used increasingly beyond research projects by countries' routine civil registration and vital statistics systems over the past 10 years [1]. The quality of cause of death data from VA will therefore directly impact health policy in many settings. Despite this, to our knowledge there has been no attempt to systematically review the quality of studies validating VA methods.

This review aims to describe the main characteristics of VA validation studies, including the VA questionnaire used, comparator dataset, and metrics to evaluate agreement. This review is a critical knowledge synthesis and discussion of the strengths and weaknesses of the VA validation processes. The findings of the review may facilitate the establishment of guidelines for VA validation processes.

## Methods

This study was registered with PROSPERO (registration number: CRD42020186886) [10]. The review process and reporting were done according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [11].

## Search strategy

A search strategy was developed and refined with librarian support (Additional File 1). MEDLINE, EMBASE, Cochrane-library and Scopus electronic databases were searched from the inception up to the June 2020. The search resulted in 1008 articles (i.e. 294, 307, 69, 338 articles respectively from each database). Those citations were imported into the EndNote-X9 reference manager application for deduplication. Following deduplication, there were 481 articles. A secondary search was conducted from July 2020 to January 2022 to identify new publications. An additional 35 studies were screened, and two were eligible for inclusion. Two reviewers (BPKM, HRC) conducted the search in parallel and independently.

## Eligibility criteria and study selection

Three screening questions were used in the study selection: 1) Does the study/article include a VA component (i.e. interviews with relevant respondents to determine cause of death); 2) Does the study include a validation component (i.e. comparison of cause of death assigned by the VA with another source that is not solely based on the data collected in the VA; and 3) Does the article include

original analysis or a reanalysis (i.e. not a protocol or a comment/correspondence relating to another study).

In the first round, titles and abstracts were reviewed by two authors (BPKM, HRC). In the second round, full texts were reviewed. Two authors (BPKM, HRC) independently selected studies with the DistillerSR application. Discrepancies were reviewed by a third author (RJ). At the end of the second round, 66 articles were selected for data extraction by two authors (BPKM and JDH). The flow diagram of study selection is shown in Fig. 1.

## Risk of bias assessment

Risk of bias assessment was done with a modified tool comprised of 10 domains [12, 13]. Two authors (BPKM, AA) assessed the risks independently and any discrepancy was attended by a third (RHH).

Domain 1: Study's target population/dataset involved was a close representation of the population of a defined geographical area/s (national or sub-national) OR a close representation of the population presenting to a healthcare institution/s in relation to socio-demographic characteristics.

Domain 2: Random selection was used to select the sample OR the total target population/dataset was covered.

Domain 3: Non-response bias was minimal.

Domain 4: Appropriate descriptive statistics have been used in describing the comparison.

Domain 5: Appropriate summary statistics have been used for the comparison.

Domain 6: Comparator had been clearly described.

Domain 7: A validated instrument (e.g. questionnaire) was used for data collection.

Domain 8: Observers were blinded to previous findings.

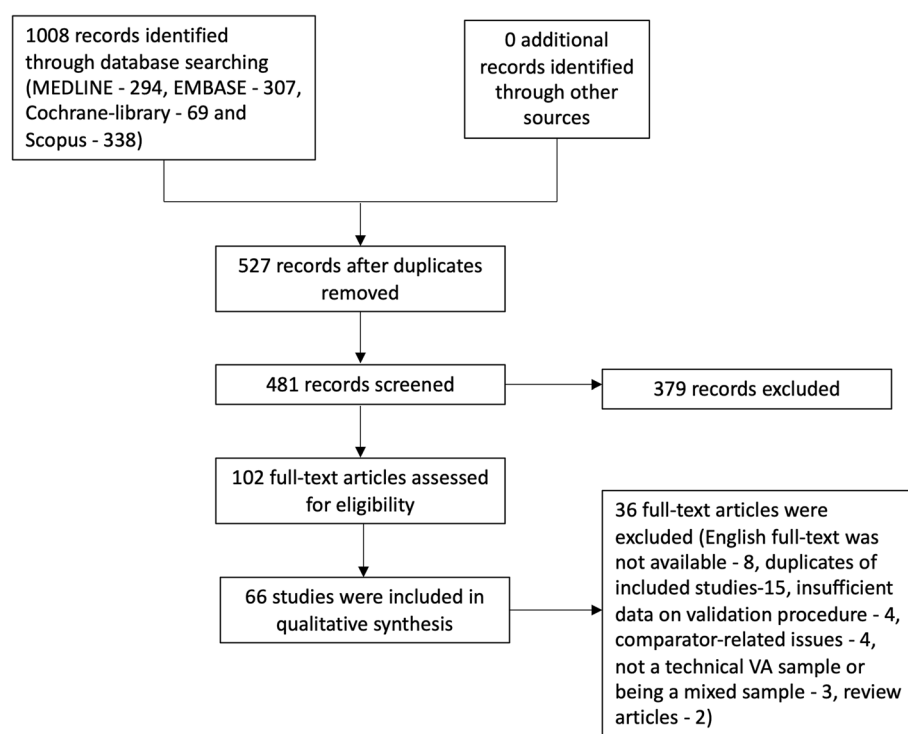
Domain 9: Comparator had been determined before the data collection.

Domain 10: Clear descriptions have been given on the competence of those who were involved in data collection and analysis.

All articles were assessed for each domain and one of the three responses was recorded for each: low-risk, unclear, and high-risk. A risk of bias assessment was summarized using a colour code system.

## Data extraction and narrative synthesis

Seventeen variables were extracted from the selected studies: 1) setting; 2) composition/details of the VA tool used; 3) relation of VA interviewers to the deceased; place of interview; 4) time between death and VA interview; 5) description of the deceased (e.g. general population or



**Fig. 1** Flow diagram of study selection

suspected of a specific condition); 6) age composition of the deceased; 7) gender composition and other relevant characteristics of the deceased; 8) number of death records collected; 9) details of the respondent or source of the details of the deceased; 10) whether comparator was a primary or secondary data source; 11) description of the comparison data, 12) including number of records used; 13) categorization of quality of comparator dataset; 14) whether VA validation was the main study objective; 15) what measures were used in describing the agreement; 16) the findings of the agreement assessment; and 17) whether a computer coded or physician certified VA method was used.

The following categorisation was developed to assess the quality of the comparator dataset.

Level 1: Autopsy diagnosis.

Level 2: Hospital diagnosis with evidence suggestive that investigation findings (except autopsy) were used in the diagnosis, in addition to the signs and symptoms.

Level 3: Hospital diagnosis without further details on the criteria used/ Hospital diagnosis only based on signs or symptoms.

Level 4: Externally made diagnosis (e.g. by a physician not involved in the management of the patient) using available clinical records.

Level 5: Civil registry or other non-medical records.

The identified domains were summarised using narrative synthesis under the domains; objective and settings, risk of bias, VA instruments and data collectors, study populations and analysis techniques including comparator dataset.

## Results

The characteristics of the selected studies are shown in Additional File 2.

### Objective and settings

Among the 66 studies, 59 included VA validation as the primary objective, whereas in 7 studies it was either a secondary objective, the comparator, or done on a sub-sample [14–20]. 10 validation studies were conducted on existing datasets [8, 9, 15, 21–27]. The other studies involved a data collection component in diverse countries including Ethiopia, South Africa, India, China, Philippines, Pakistan, Bangladesh, Georgia, Haiti, Ghana,

Tanzania, Indonesia, Iran, Kenya, Zimbabwe, Liberia, Malawi, Malaysia, Mexico, Uganda and Vietnam.

### Risk of bias

The risk of bias assessment is summarized in Additional File 3. Most studies were low risk across the individual domains assessed, however, 9 studies [15, 28–35] did not collect a representative sample or cover the entire population (Domain 2), 17 studies [15, 17, 29, 30, 32, 35–46] had a high risk of non-response (Domain 3), and 15 studies [16, 17, 20, 31, 32, 40, 43, 46–53] did not use a validated instrument (Domain 7).

### VA instruments and data collectors

Except for the 10 studies [8, 9, 15, 21–27] that used existing databases, the most used VA instruments were the World Health Organization questionnaire ( $n=16$ ) [29, 39, 41, 51, 54–63] and Population Health Metrics Research Consortium (PHMRC) questionnaire ( $n=5$ ) [8, 42, 64–66]. In 7 studies [18, 40, 43, 48, 53, 67, 68], the tool was not mentioned. In other studies, locally designed VA instruments were used. In 5 studies [30, 32, 49, 50, 67], the data collectors were nurses, in another 3 [20, 51, 59] they were doctors, in 10 [14, 29, 38, 43, 60, 61, 64, 66, 69, 70] the type of data collectors was not mentioned. Other studies included several other categories such as field workers and non-medical graduates.

In the majority ( $n=38$ ) of validation studies, the data collection was conducted at the household setting of the deceased. In 4 studies [30, 35, 58, 71], a combination of

settings was used. In 10 studies [28, 29, 40, 44, 45, 52, 64, 68–70], the data collection setting was not mentioned. Other data collection settings included during pilgrimage, in hospital, and at the mortuary. The time between death and the interview varied from immediately on collection of the body up to 42 months. In 20 studies [16–19, 28, 34, 35, 42, 45, 46, 48, 52, 53, 57, 59, 64, 67–70], the time between the death and the interview was not mentioned. The interviewees were not specified in 18 studies [20, 28, 29, 40, 43–46, 56, 58, 62, 64, 65, 67–71], however in all others they were clearly described as family members, relatives or principal caregivers at the time of death.

### Study populations

In 29 studies, the study population included the general adult population and in 15 studies [14, 30, 31, 33, 38, 39, 52, 54–56, 60, 61, 63, 70, 72] it included general neonatal or child deaths. Three studies included only female deaths [34, 47, 73] and others were conducted on selected groups such as deaths from HIV/TB, pilgrims, and respiratory infections. In general, the studies with general adult deaths had an approximately equal male to female ratio. In three studies there were fewer than 50 interviews [30, 53, 74], and in the others it ranged from 100 to over 26,000.

### Analysis techniques and comparator dataset

Physician certification was used in 36 studies and computer coded methods were used in 14 studies [8, 15, 18, 23, 24, 27, 29, 32, 42, 44, 50, 64, 66, 75]. Both techniques

**Table 1** Categorization of the quality of the comparator dataset

Level	Description	Frequency (%)	References
1	Autopsy diagnosis	1 (1.5%)	[53]
2	Hospital diagnosis with evidence suggestive that investigation findings (i.e. except autopsy) were used in the diagnosis in addition to the signs and symptoms	36 (54.5%)	[8, 9, 15, 21–29, 31–34, 37–40, 44, 47–50, 52, 54, 55, 58, 59, 63–65, 70, 72, 76]
3	Hospital diagnosis without further details of the criteria used/Hospital diagnosis only based on signs and symptoms	7 (10.6%)	[16, 17, 30, 36, 51, 56, 67]
4	Externally made diagnosis (e.g. by a physician not involved in the management of the patient) using available clinical records	11 (16.7%)	[35, 41, 45, 46, 60–62, 68, 69, 74, 77]
5	Civil registry or other non-medical records	8 (12.1%)	[14, 18–20, 42, 57, 66, 73]
Mixed	More than 1 level	3 (4.5%)	[43, 71, 75]
Total		66 (100.0%)	

were used in 11 studies [21, 22, 28, 40, 47, 52, 58, 65, 70, 71, 76] and the method was not mentioned clearly in 6 studies [9, 26, 38, 45, 67, 68]. The comparator data was a primary source in the majority of studies ( $n=53$ ), while in 13 it was a secondary source.

Categorisation of quality of the comparator dataset is shown in Table 1. Most studies ( $n=37$ ) used a comparator dataset of high quality (level 1 or 2) that included investigation findings.

The measures used in the comparison included percentage agreement, chance-corrected concordance (CCC), CSMF accuracy, Kappa coefficient, concordance correlation coefficient, sensitivity, specificity, predictive value and statistical tests for exploring significant differences. Eleven studies [8, 21, 23–25, 27, 58, 64, 65, 71, 75] used CCC, 14 studies [8, 21–27, 58, 64, 65, 68, 71, 75] used CSMF accuracy, 16 studies [8, 18–20, 23, 28, 34, 39, 55, 57–60, 66, 67, 73] used kappa coefficient, and 38 studies [9, 16, 17, 20, 28, 29, 31–35, 37, 39, 40, 43–46, 48–50, 52–57, 59, 60, 62, 63, 66, 67, 69, 72, 73, 75, 77] used sensitivity, specificity, or positive predictive value. In general, the studies showed overall acceptable agreements. Specificity was observed to be higher than sensitivity in most of the studies that measured those two parameters. Additionally, the negative predictive value was higher than the positive predictive value.

## Discussion

This is the first systematic review conducted on the characteristics of VA validation studies. This review revealed that while most studies have confirmed the validity of VA methods, the quality of the comparator datasets varied considerably, impacting the interpretation of VA methods. Additionally, most studies used VA certified by physicians and a variety of VA instruments and agreement metrics.

This review highlights several factors related to the quality of VA validation methods. First, even though some studies were designated “validation studies”, on closer inspection, key criteria were missing. For example, in 7 studies, the main objective was not to validate the VA method. Further, in 13 studies the validation was conducted using a secondary data collection method in which the comparator did not include the same set of deaths. Essential details that should have been mentioned were missing for some studies. For example, the type of data collectors was not mentioned in 8 studies; the data collection setting was not mentioned in 10 studies; the time between death and interview was not mentioned in 20 studies; and the interviewees were not specified in 18 studies.

The second implication of this review is that the quality of the comparator data used needs to be evaluated carefully. We developed a new classification to assess the quality of comparator data assessment. Twenty-six

studies used a lower quality comparator than level 2, meaning the comparator dataset did not include laboratory or imaging findings. For some causes of death, such as injuries or external causes of death, further investigations beyond the clinical record may not be necessary to determine cause of death. But for other causes, such as specific cancers, laboratory findings are essential to accurately assigning cause of death. The availability of laboratory and imaging investigations can be limited, especially in rural areas where VA is most often used, which presents a challenge for VA validation studies. Furthermore, most studies were not focused on external causes of death, which suggests higher quality comparison cause of death data is needed to reliably interpret the validity of VA methods.

Third, VA validation studies should adhere to more consistent use of agreement metrics. Most studies used sensitivity, specificity, or positive predictive value to evaluate the agreement for cause of death assignment. While these metrics are appropriate for some comparisons, they do not adjust for assignment by chance as in CCC and chance-corrected cause-specific mortality fraction [78, 79]. Additionally, since VA is a blunt tool, its primary purpose remains at population, rather than individual level cause of death assignment [80]. As such, validation studies should include a metric that assesses the population level agreement, such as CSMF accuracy.

There were several limitations of this review. Firstly, only articles in English were included in the review. In the risk of bias assessment, although the aim was to conduct objective assessments, the possibility of some subjectivity could not be excluded. As an example, one criterion was whether a validated tool was used. In reporting the studies, the exact phrase “validated tool” was used in some studies whereas in others, less specific descriptions were used, such as “modifications which were made to validated tools”.

## Conclusions

This review highlights that while the majority of VA validation studies have reported favourable findings for the VA method under assessment, major differences were observed in the methodologies. Many used poor quality comparison cause of death data that affects the interpretation of the validation assessment. Future VA validation studies should adhere to consistent methodological criteria. Robust validation studies will help health policy planners and those involved in VA implementation make informed decisions before selecting a particular VA method.

## Abbreviations

CSMF	Cause-specific mortality fraction
CCC	Chance-corrected concordance
PHMRC	Population Health Metrics Research Consortium
VA	Verbal autopsy



# Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12889-022-14628-1>.

**Additional file 1.** Search Strategies used in MEDLINE and EMBASE.

**Additional file 2.** Characteristics of all studies selected from the literature search.

**Additional file 3.** Colour coded risk of bias assessment across ten domains.

**Additional file 4.**

## Acknowledgements

Not applicable.

## Authors' contributions

RHH conceptualised the study. BPKM, AA, and TA designed and conducted the search. BPKM, HRM, and RJ screened the studies. BPKM and JDH extracted the data. BPKM, AA and RHH conducted the risk of bias assessment. BPKM and RHH wrote the first draft. All authors contributed to study design and revisions. The author(s) read and approved the final manuscript.

## Funding

This work was funded by Bloomberg Philanthropies. Bloomberg Philanthropies had no role in decision to publish, or preparation of the manuscript.

## Availability of data and materials

All data generated or analysed during this study are included in this published article and its supplementary information files.

## Declarations

## Ethics approval and consent to participate

Not applicable.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

Received: 29 June 2021 Accepted: 14 November 2022

Published online: 29 November 2022

## References

- Hazard RH, Buddhika MPK, Hart JD, Chowdhury HR, Firth S, Joshi R, et al. Automated verbal autopsy: from research to routine use in civil registration and vital statistics systems. *BMC Med*. 2020;18(1):60.
- Mikkelsen L, Phillips DE, AbouZahr C, Setel PW, de Savigny D, Lozano R, et al. A global assessment of civil registration and vital statistics systems: monitoring data quality and progress. *Lancet*. 2015;386(10001):1395–406.
- Collaborators GBDCoD. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2018;392(10159):1736–88.
- Mikkelsen L, de Alwis S, Sathasivam S, Kumarapeli V, Tennakoon A, Karunapema P, et al. Improving the Policy Utility of Cause of Death Statistics in Sri Lanka: An Empirical Investigation of Causes of Out-of-Hospital Deaths Using Automated Verbal Autopsy Methods. *Frontiers in Public Health*. 2021;9(587).
- AbouZahr C, de Savigny D, Mikkelsen L, Setel PW, Lozano R, Nichols E, et al. Civil registration and vital statistics: progress in the data revolution for counting and accountability. *Lancet*. 2015;386(10001):1373–85.
- Joshi R, Praveen D, Jan S, Raju K, Maulik P, Jha V, et al. How much does a verbal autopsy based mortality surveillance system cost in rural India? *PLoS ONE*. 2015;10(5):e0126410.
- Byass P, Huong DL, Minh HV. A probabilistic approach to interpreting verbal autopsies: methodology and preliminary validation in Vietnam. *Scand J Public Health Suppl*. 2003;62:32–7.
- Serina P, Riley I, Stewart A, James SL, Flaxman AD, Lozano R, et al. Improving performance of the Tariff Method for assigning causes of death to verbal autopsies. *BMC Med*. 2015;13:291.
- Murray CJ, Lozano R, Flaxman AD, Serina P, Phillips D, Stewart A, et al. Using verbal autopsy to measure causes of death: the comparative performance of existing methods. *BMC Med*. 2014;12:5.
- International prospective register of systematic reviews (PROSPERO). Validation studies of verbal autopsy methods: A systematic review: National Institute for Health Research; 2020 [Available from: [https://www.crd.york.ac.uk/prospero/display\\_record.php?RecordID=186886](https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=186886)].
- Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ*. 2009;339:b2700.
- Hoy D, Brooks P, Woolf A, Blyth F, March L, Bain C, et al. Assessing risk of bias in prevalence studies: modification of an existing tool and evidence of interrater agreement. *J Clin Epidemiol*. 2012;65(9):934–9.
- Brink Y, Louw QA. Clinical instruments: reliability and validity critical appraisal. *J Eval Clin Pract*. 2012;18(6):1126–32.
- Arudo J, Gimnig JE, ter Kuile FO, Kachur SP, Slutsker L, Kolczak MS, et al. Comparison of government statistics and demographic surveillance to monitor mortality in children less than five years old in rural western Kenya. *Am J Trop Med Hyg*. 2003;68(4 Suppl):30–7.
- Byass P. Usefulness of the Population Health Metrics Research Consortium gold standard verbal autopsy data for general verbal autopsy methods. *BMC Med*. 2014;12:23.
- Gajalakshmi V, Peto R. Verbal autopsy of 80,000 adult deaths in Tamilnadu. South India *BMC Public Health*. 2004;4:47.
- Gajalakshmi V, Peto R, Kanaka S, Balasubramanian S. Verbal autopsy of 48 000 adult deaths attributable to medical causes in Chennai (formerly Madras). India *BMC Public Health*. 2002;2:7.
- Hussain-Alkhateeb L, Fottrell E, Petzold M, Kahn K, Byass P. Local perceptions of causes of death in rural South Africa: a comparison of perceived and verbal autopsy causes of death. *Glob Health Action*. 2015;8:28302.
- Joubert J, Bradshaw D, Kabudula C, Rao C, Kahn K, Mee P, et al. Record-linkage comparison of verbal autopsy and routine civil registration death certification in rural north-east South Africa: 2006–09. *Int J Epidemiol*. 2014;43(6):1945–58.
- Wang L, Yang G, Jiemin M, Rao C, Wan X, Dubrovsky G, et al. Evaluation of the quality of cause of death statistics in rural China using verbal autopsies. *J Epidemiol Community Health*. 2007;61(6):519–26.
- Flaxman AD, Vahdatpour A, Green S, James SL, Murray CJ. Population Health Metrics Research C. Random forests for verbal autopsy analysis: multisite validation study using clinical diagnostic gold standards. *Popul Health Metr*. 2011;9:29.
- Flaxman AD, Vahdatpour A, James SL, Birnbaum JK, Murray CJ. Population Health Metrics Research C. Direct estimation of cause-specific mortality fractions from verbal autopsies: multisite validation study using clinical diagnostic gold standards. *Popul Health Metr*. 2011;9:35.
- Kalter HD, Perin J, Black RE. Validating hierarchical verbal autopsy expert algorithms in a large data set with known causes of death. *J Glob Health*. 2016;6(1):010601.
- Lozano R, Freeman MK, James SL, Campbell B, Lopez AD, Flaxman AD, et al. Performance of InterVA for assigning causes of death to verbal autopsies: multisite validation study using clinical diagnostic gold standards. *Popul Health Metr*. 2011;9:50.
- Lozano R, Lopez AD, Atkinson C, Naghavi M, Flaxman AD, Murray CJ, et al. Performance of physician-certified verbal autopsies: multisite validation study using clinical diagnostic gold standards. *Popul Health Metr*. 2011;9:32.
- Murray CJ, James SL, Birnbaum JK, Freeman MK, Lozano R, Lopez AD, et al. Simplified Symptom Pattern Method for verbal autopsy analysis: multisite validation study using clinical diagnostic gold standards. *Popul Health Metr*. 2011;9:30.
- Serina P, Riley I, Stewart A, Flaxman AD, Lozano R, Mooney MD, et al. A shortened verbal autopsy instrument for use in routine mortality surveillance systems. *BMC Med*. 2015;13:302.

28. Bauni E, Ndila C, Mochamah G, Nyutu G, Matata L, Ondieki C, et al. Validating physician-certified verbal autopsy and probabilistic modeling (InterVA) approaches to verbal autopsy interpretation using hospital causes of adult deaths. *Popul Health Metr*. 2011;9:49.
29. Byass P, Calvert C, Miuro-Nakiyingi J, Lutalo T, Michael D, Crampin A, et al. InterVA-4 as a public health tool for measuring HIV/AIDS mortality: a validation study from five African countries. *Glob Health Action*. 2013;6:22448.
30. Dowell SF, Davis HL, Holt EA, Ruff AJ, Kissinger PJ, Bijoux J, et al. The utility of verbal autopsies for identifying HIV-1-related deaths in Haitian children. *AIDS*. 1993;7(9):1255–9.
31. 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(2):380–6.
32. Lopman BA, Barnabas RV, Boerma JT, Chawira G, Gaitskill K, Harrop T, et al. Creating and validating an algorithm to measure AIDS mortality in the adult population using verbal autopsy. *PLoS Med*. 2006;3(8):e312.
33. 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(2):132–42.
34. Midhet F. Validating the verbal autopsy questionnaire for maternal mortality in Pakistan. *Int J Health Sci (Qassim)*. 2008;2(1):91–6.
35. Mondal RN, Alam S, Haque A, Azad SJAQ, Rani MRM, Shaheeduzzaman BA, et al. Validity of Verbal Autopsy for Assessment of Causes of Death in Bangladesh. *European Academic Research*. 2015;3(8):8725–40.
36. Ganapathy SS, Yi Yi K, Omar MA, Anuar MFM, Jeevananthan C, Rao C. Validation of verbal autopsy: determination of cause of deaths in Malaysia 2013. *BMC Public Health*. 2017;17(1):653.
37. Mayanja BN, Baisley K, Nalweyiso N, Kibengo FM, Mugisha JO, Van der Paal L, et al. Using verbal autopsy to assess the prevalence of HIV infection among deaths in the ART period in rural Uganda: a prospective cohort study, 2006–2008. *Popul Health Metr*. 2011;9:36.
38. 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(6):365–9.
39. Nausheen S, Soofi SB, Sadiq K, Habib A, Turab A, Memon Z, et al. Validation of verbal autopsy tool for ascertaining the causes of stillbirth. *PLoS ONE*. 2013;8(10):e76933.
40. 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(6):1081–7.
41. Qureshi JS, Samuel JC, Mulima G, Kakoulides S, Cairns B, Charles AG. Validating a verbal autopsy tool to assess pre-hospital trauma mortality burden in a resource-poor setting. *Trop Med Int Health*. 2014;19(4):407–12.
42. Ramirez-Villalobos D, Stewart AL, Romero M, Gomez S, Flaxman AD, Hernandez B. Analysis of causes of death using verbal autopsies and vital registration in Hidalgo, Mexico. *PLoS ONE*. 2019;14(7):e0218438.
43. Setel PW, Whiting DR, Hemed Y, Chandramohan D, Wolfson LJ, Alberti KG, et al. Validity of verbal autopsy procedures for determining cause of death in Tanzania. *Trop Med Int Health*. 2006;11(5):681–96.
44. Tensou B, Araya T, Telake DS, Byass P, Berhane Y, Kebebew T, et al. Evaluating the InterVA model for determining AIDS mortality from verbal autopsies in the adult population of Addis Ababa. *Trop Med Int Health*. 2010;15(5):547–53.
45. Tran HT, Nguyen HP, Walker SM, Hill PS, Rao C. Validation of verbal autopsy methods using hospital medical records: a case study in Vietnam. *BMC Med Res Methodol*. 2018;18(1):43.
46. Yang G, Rao C, Ma J, Wang L, Wan X, Dubrovsky G, et al. Validation of verbal autopsy procedures for adult deaths in China. *Int J Epidemiol*. 2006;35(3):741–8.
47. 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(4):414–22.
48. 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(11):824–31.
49. Kamali A, Wagner HU, Nakiyingi J, Sibiiti I, Kengeya-Kayondo JF, Mulder DW. Verbal autopsy as a tool for diagnosing HIV-related adult deaths in rural Uganda. *Int J Epidemiol*. 1996;25(3):679–84.
50. Lopman B, Cook A, Smith J, Chawira G, Urassa M, Kumogola Y, et al. Verbal autopsy can consistently measure AIDS mortality: a validation study in Tanzania and Zimbabwe. *J Epidemiol Community Health*. 2010;64(4):330–4.
51. Pane M, Imari S, Alwi Q, NyomanKandun I, Cook AR, Samaan G. Causes of mortality for Indonesian Hajj Pilgrims: comparison between routine death certificate and verbal autopsy findings. *PLoS ONE*. 2013;8(8):e73243.
52. Quigley MA, Armstrong Schellenberg JR, Snow RW. Algorithms for verbal autopsies: a validation study in Kenyan children. *Bull World Health Organ*. 1996;74(2):147–54.
53. Rodriguez L, Reyes H, Tome P, Ridaura C, Flores S, Guiscafie H. Validation of the verbal autopsy method to ascertain acute respiratory infection as cause of death. *Indian J Pediatr*. 1998;65(4):579–84.
54. Aggarwal AK, Jain V, Kumar R. Validity of verbal autopsy for ascertaining the causes of stillbirth. *Bull World Health Organ*. 2011;89(1):31–40.
55. Aggarwal AK, Kumar P, Pandit S, Kumar R. Accuracy of WHO verbal autopsy tool in determining major causes of neonatal deaths in India. *PLoS ONE*. 2013;8(1):e54865.
56. Edmond KM, Quigley MA, Zandoh C, Danso S, Hurt C, OwusuAgyei S, et al. Diagnostic accuracy of verbal autopsies in ascertaining the causes of stillbirths and neonatal deaths in rural Ghana. *Paediatr Perinat Epidemiol*. 2008;22(5):417–29.
57. Hong TT, Phuon Hoa N, Walker SM, Hill PS, Rao C. Completeness and reliability of mortality data in Viet Nam: Implications for the national routine health management information system. *PLoS ONE*. 2018;13(1):e0190755.
58. Karat AS, Maraba N, Tlali M, Charalambous S, Chihota VN, Churchyard GJ, et al. Performance of verbal autopsy methods in estimating HIV-associated mortality among adults in South Africa. *BMJ Glob Health*. 2018;3(4):e000833.
59. Khademi H, Etemadi A, Kamangar F, Nouraei M, Shakeri R, Abaie B, et al. Verbal autopsy: reliability and validity estimates for causes of death in the Golestan Cohort Study in Iran. *PLoS ONE*. 2010;5(6):e11183.
60. Mpimbaza A, Filler S, Katureebe A, Quick L, Chandramohan D, SG S. Verbal Autopsy: Evaluation of Methods to Certify Causes of Death in Uganda. *PLoS One*. 2015;10(1):e0128801.
61. Mpimbaza A, Filler S, Katureebe A, Kinara SO, Nzabandora E, Quick L, et al. Validity of verbal autopsy procedures for determining malaria deaths in different epidemiological settings in Uganda. *PLoS ONE*. 2011;6(10):e26892.
62. Polprasert W, Rao C, Adair T, Pattaraarchachai J, Porapakkham Y, Lopez AD. Cause-of-death ascertainment for deaths that occur outside hospitals in Thailand: application of verbal autopsy methods. *Popul Health Metr*. 2010;8:13.
63. Soofi SB, Ariff S, Khan U, Turab A, Khan GN, Habib A, et al. Diagnostic accuracy of WHO verbal autopsy tool for ascertaining causes of neonatal deaths in the urban setting of Pakistan: a hospital-based prospective study. *BMC Pediatr*. 2015;15:144.
64. Chowdhury HR, Flaxman AD, Joseph JC, Hazard RH, Alam N, Riley ID, et al. Robustness of the Tariff method for diagnosing verbal autopsies: impact of additional site data on the relationship between symptom and cause. *BMC Med Res Methodol*. 2019;19(1):232.
65. James SL, Flaxman AD, Murray CJ. Population Health Metrics Research C. Performance of the Tariff Method: validation of a simple additive algorithm for analysis of verbal autopsies. *Popul Health Metr*. 2011;9:31.
66. Yokobori Y, Matsuura J, Sugiura Y, Mutemba C, Nyahoda M, Mwango C, et al. Analysis of causes of death among brought-in-dead cases in a third-level Hospital in Lusaka, Republic of Zambia, using the tariff method 2.0 for verbal autopsy: a cross-sectional study. *BMC Public Health*. 2020;20(1):473.
67. Dehghan A, Nasirian M, Haghdoust AA, Bahramali E, Sharifi H. Validation of the verbal autopsy questionnaire for adult deaths in Iran. *Med J Islam Repub Iran*. 2018;32:7.
68. Murray CJ, Lopez AD, Feehan DM, Peter ST, Yang G. Validation of the symptom pattern method for analyzing verbal autopsy data. *PLoS Med*. 2007;4(11):e327.
69. Akgun S, Colak M, Bakar C. Identifying and verifying causes of death in Turkey: National verbal autopsy survey. *Public Health*. 2012;126(2):150–8.
70. Ndila C, Bauni E, Nyirongo V, Mochamah G, Makazi A, Kosgei P, et al. Verbal autopsy as a tool for identifying children dying of sickle cell disease: a validation study conducted in Kilifi district. *Kenya BMC Med*. 2014;12:65.
71. Karat AS, Tlali M, Fielding KL, Charalambous S, Chihota VN, Churchyard GJ, et al. Measuring mortality due to HIV-associated tuberculosis among adults in South Africa: Comparing verbal autopsy, minimally-invasive autopsy, and research data. *PLoS ONE*. 2017;12(3):e0174097.

72. 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(2):134–44.
73. Lomia N, Berdzuli N, Sturua L, Kereselidze M, Topuridze M, Pestvenidze E, et al. Leading causes of death of women of reproductive age in the Republic of Georgia: findings from the National Reproductive Age Mortality Survey (2014). *Int J Womens Health*. 2018;10:437–52.
74. Pacque-Margolis S, Pacque M, Dukuly Z, Boateng J, Taylor HR. Application of the verbal autopsy during a clinical trial. *Soc Sci Med*. 1990;31(5):585–91.
75. Murtaza SS, Kolpak P, Bener A, Jha P. Automated verbal autopsy classification: using one-against-all ensemble method and Naive Bayes classifier. *Gates Open Res*. 2018;2:63.
76. 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(6):436–46.
77. Misganaw A, Mariam DH, Araya T, Aneneh A. Validity of verbal autopsy method to determine causes of death among adults in the urban setting of Ethiopia. *BMC Med Res Methodol*. 2012;12:130.
78. Flaxman AD, Serina PT, Hernandez B, Murray CJ, Riley I, Lopez AD. Measuring causes of death in populations: a new metric that corrects cause-specific mortality fractions for chance. *Popul Health Metr*. 2015;13:28.
79. Murray CJ, Lozano R, Flaxman AD, Vahdatpour A, Lopez AD. Robust metrics for assessing the performance of different verbal autopsy cause assignment methods in validation studies. *Popul Health Metr*. 2011;9:28.
80. de Savigny D, Riley I, Chandramohan D, Odhiambo F, Nichols E, Notzon S, et al. Integrating community-based verbal autopsy into civil registration and vital statistics (CRVS): system-level considerations. *Glob Health Action*. 2017;10(1):1272882.

# Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

