Validation of SmartVA using conventional autopsy: A study of adult deaths in Brazil

John D. Hart,[a,*] Paulo Afonso de André,[b] Carmen Diva Saldiva de André,[c] Tim Adair,[a] Lucia Pereira Barroso,[c] Sandra Valongueiro,[d] Ana Luiza Bierrenbach,[e,f] Patrícia Ismael de Carvalho,[q] Maria Bernadete de Cerqueira Antunes,[g] Conceição Maria de Oliveira,[h] Luiz Alberto Amador Pereira,[b] Cátia Martinez Minto,[i] Tânia Maria da Silva Bezerra,[j] Sérgio Parente Costa,[k] Bárbara Araújo de Azevedo,[j] José Ricardo Alves de Lima,[l] Denise Souza de Meira Mota,[l] Ana Maria de Oliveira Ramos,[m,n] Maria de Fátima Marinho de Souza,[f] Luiz Fernando Ferraz da Silva,[b,o] Elisabeth Barboza França,[p] Deirdre McLaughlin,[a] Ian D. Riley,[a] and Paulo Hilário Nascimento Saldiva,[b]

[a]Melbourne School of Population and Global Health, The University of Melbourne, Carlton, Victoria, Australia
[b]University of São Paulo, School of Medicine, São Paulo, São Paulo, Brazil
[c]University of São Paulo, Institute of Mathematics and Statistics, São Paulo, São Paulo, Brazil
[d]Federal University of Pernambuco, Recife, Pernambuco, Brazil
[e]Sírio-Libanês Hospital, São Paulo, São Paulo, Brazil
[f]Vital Strategies, São Paulo, São Paulo, Brazil
[g]University of Pernambuco, Recife, Pernambuco, Brazil
[h]Executive Secretary of the Municipal Health Surveillance of Recife, Pernambuco, Brazil
[i]São Paulo Health State Secretary, São Paulo, São Paulo, Brazil
[j]Secretary of State for Health, Recife, Pernambuco, Brazil
[k]Health Surveillance, Municipal Department of Health, Olinda, Pernambuco, Brazil.
[l]Recife Autopsy Service, Recife, Brazil*
[m]Federal University of Rio Grande do Norte, Health Sciences Center, Natal, Rio Grande do Norte, Brazil
[n]Natal Autopsy Service, Natal, Rio Grande do Norte, Brazil
[o]São Paulo Autopsy Service, University of São Paulo, Sao Paulo, Brazil
[p]Federal University of Minas Gerais, School of Medicine, Belo Horizonte, Minas Gerais, Brazil
[q]Executive Secretary of Health Surveillance of the State of Pernambuco, Brazil

The Lancet Regional Health - Americas 2022;5: 100081
Published online 31 October 2021
https://doi.org/10.1016/j.lana.2021.100081

Summary

Background Accurate cause of death data are essential to guide health policy. However, mortality surveillance is limited in many low-income countries. In such settings, verbal autopsy (VA) is increasingly used to provide population-level cause of death data. VAs are now widely interpreted using the automated algorithms SmartVA and InterVA. Here we use conventional autopsy as the gold standard to validate SmartVA methodology.

Methods This study included adult deaths from natural causes in São Paulo and Recife for which conventional autopsy was indicated. VA was conducted with a relative of the deceased using an amended version of the SmartVA instrument to suit the local context. Causes of death from VA were produced using the SmartVA-Analyze program. Physician coded verbal autopsy (PCVA), conducted on the same questionnaires, and Global Burden of Disease Study data were used as additional comparators. Cause of death data were grouped into 10 broad causes for the validation due to the real-world utility of VA lying in identifying broad population cause of death patterns.

Findings The study included 2,060 deaths in São Paulo and 1,079 in Recife. The cause specific mortality fractions (CSMFs) estimated using SmartVA were broadly similar to conventional autopsy for: cardiovascular diseases (46.8% vs 54.0%, respectively), cancers (10.6% vs 11.4%), infections (7.0% vs 10.4%) and chronic respiratory disease (4.1% vs 3.7%), causes accounting for 76.1% of the autopsy dataset. The SmartVA CSMF estimates were lower than autopsy for "Other NCDs" (7.8% vs 14.6%) and higher for diabetes (13.0% vs 6.6%). CSMF accuracy of SmartVA compared to autopsy was 84.5%. CSMF accuracy for PCVA was 93.0%.

Interpretation The results suggest that SmartVA can, with reasonable accuracy, predict the broad cause of death groups important to assess a population's epidemiological transition. VA remains a useful tool for understanding causes of death where medical certification is not possible.

*Corresponding author
E-mail address: john.hart@unimelb.edu.au (J.D. Hart).

Funding: This study was funded by: Brazil Ministry of Health (grant number: 815781/2014); FAPESP São Paulo Research Foundation (grant number: 2013/21728-2); and FAPESP SPRINT University of Melbourne (grant number: 2016/502215).

Keywords: Cause of death; Verbal autopsy; Autopsy; Brazil; Validation; Vital statistics

Research in context

Evidence before this study

Verbal autopsy (VA) is used increasingly in resource-limited settings to improve cause of death data. There are three automated VA algorithms that are endorsed by the World Health Organisation for interpreting VA data: SmartVA, InterVA and InSilicoVA. Prior to this study, to our knowledge, no research had been conducted to validate any of these tools using conventional autopsy as the gold standard.

Added value of this study

This study provides researchers and countries implementing VA methods with a robust comparison of SmartVA and conventional autopsy diagnosis for the first time. We show that SmartVA can predict broad cause of death groups, that are important for health policy in resource-limited settings, with reasonable accuracy.

Implications of all the available evidence

Despite improvements in mortality surveillance and civil registration and vital statistics globally, verbal autopsy will remain an important tool for poorer countries to understand the causes of death in their populations for many years to come. This research, further to the initial validation of SmartVA using hospital cause of death data, should provide confidence in the accuracy of SmartVA methodology for countries in which medical certification of cause of death is inadequate. Further validation studies using different VA methodologies would be useful for countries deciding whether to implement VA and what methods to use.

Introduction

Accurate cause of death data are essential for understanding countries' major health problems and to guide health policy debate. These data are also required for evaluation and optimisation of health programs. Ideally, cause of death data are derived from complete and accurate medical certification of cause of death by trained physicians using the standard international form to identify the underlying cause of death.[1,2] However, many populations, particularly in low-income settings, lack effective mortality surveillance due to limited health and civil registration services. The required systems for recording complete and accurate cause of death data are not available in some areas of Brazil, which results in a high proportion of ill-defined causes of death.

Where medical certification is not a feasible option in the short term, verbal autopsy (VA) methodology is increasingly used to estimate causes of death at the population level.[3] VA involves a structured interview, usually with a relative of the deceased, to identify key factors in the history and circumstances leading to death that will enable prediction of most likely cause of death. Until recently, physician review of VA interviews was required for prediction of causes of death from the interviews, however computer algorithms that cost less, save time and are more consistent are now generally preferred by countries implementing VA.[4] There are three VA algorithms that are endorsed by the World Health Organisation: SmartVA, InterVA and InSilicoVA.[5–7] SmartVA analyses data from the shortened PHMRC questionnaire, which contains approximately 50% fewer questions than the World Health Organisation 2016 questionnaire that is required to run InterVA and InSilicoVA, which may represent an important time saving when introduced into routine health worker activities.[8]

SmartVA has been validated using gold standard causes of death diagnosed through robust criteria, including clinical endpoints, laboratory findings, medical imaging, and pathology.[9,10] However, SmartVA and, to our knowledge, no other VA methodology has been validated using conventional autopsy. Although conventional autopsy rates have plummeted from as high as 60% of hospital deaths in high income settings to lower than 10% now, autopsy is generally still considered the most accurate gold standard diagnosis for comparing other methods of diagnosing cause of death, such as medical certification.[11] Reasons why conventional autopsy has not been used to validate VA likely include the expense and logistical difficulty of conducting conventional autopsy in the remote areas where VA is required, the potential for a biased sample as not all deaths are submitted to autopsy procedures, and responses to VA questions may differ between remote areas and the urban areas in which most autopsies are performed.

Here we present the validation of SmartVA methodology using conventional autopsy for deaths occurring in São Paulo and Recife. Whilst we may expect some discrepancy in the causes identified from these two very different methods that are generally applied in different

settings, the development of VA algorithms is likely to benefit through validation with the highest quality gold standard causes. We discuss potential reasons for discrepancies in cause predictions and their implications for the more widespread use of SmartVA.

Methods

Study Setting

Data for this study were collected in São Paulo and Recife. The Faculty of Medicine of the University of São Paulo houses the Post-Mortem Verification Service (PMVS), which performs autopsies for all natural deaths for which a physician is unable to complete the medical certificate of cause of death. Approximately 14,000 autopsies are conducted per year, which corresponds to approximately 15-20% of all natural deaths. Similar procedures are followed in Recife.

Verbal autopsies were conducted for deaths at age 18 years and over that were due to natural causes and for which the cause of death was undetermined clinically, requiring them to be sent for the local Post Mortem Verification Service in São Paulo or Recife. At the study sites, approximately 80% of autopsies conducted are for natural deaths occurring in the community, or within 48 hours of arrival at hospital, when a cause of death cannot be determined. The remaining 20% of autopsies are for cases that were hospitalised for a longer period of time but for which there is no clear diagnosis or the case is of academic interest — these cases were excluded from this study. VAs were conducted from May 2016 to June 2018 in São Paulo and from December 2018 to April 2019 in Recife. A convenience sample of cases that underwent autopsy during these periods was included in the study.

Verbal autopsy interview

The verbal autopsy instrument was amended from the standard SmartVA tool. The questionnaire was translated into Portuguese and questions were added relating to Chagas disease, alcohol use and dementia, as well as to measure the length of the interview. Questions relating to injuries and some of the female-specific questions were excluded. Initially a paper form with subsequent electronic data entry was used prior to the development of an electronic tool amended from the SmartVA questionnaire. An online system that included the questionnaire was then implemented to schedule all study tasks, providing the required information for each member of the study, including interviewers, supervisors, pathologists, clinicians, and coders.

A single VA supervisor in São Paulo and one in Recife oversaw the data collection processes and supported the interviewers. Due to the challenging nature of conducting VAs, that requires empathy and psychological resilience, as well as technical skills, several interviewers were trained on all aspects of VA, including consent, ethics and the study rationale, and their roles were changed over the course of the study. A study interviewer explained the research and if the relative consented, the interview was conducted in a private room. VA interviews were conducted with family members when they visited the PMVS to await the release of their relatives' bodies and for the medical certificate of cause of death (MCCOD). In Recife, home interviews were conducted when the person taking the body to the PMVS was not appropriate for the VA interview.

Gold standard criteria

Autopsies at PMVS involve macroscopic examination plus histological examination of samples typically taken from the heart, pancreas, kidney, lungs, liver, spleen and brain, plus other organs as indicated. Sample blocks are prepared, sliced, and stained with haematoxylin and eosin; other staining techniques may be used if required. Information on medical conditions and risk factors is routinely available for the pathologists from a short interview conducted with the next of kin by administrative staff, social workers and sometimes the pathologist, depending on the site. For this purpose, in São Paulo, the open narrative section of the VA interview was made available to the pathologists. The pathologists may be able to list the underlying, intermediate, and immediate causes of death from the macroscopic findings alone, in which case the MCCOD is completed without delay. Others are completed after review of the histopathological findings.

The autopsy diagnoses for this study were produced by senior pathologists who conducted the histopathology slide reading and had access to the macroscopic findings at autopsy. The sequence of causes that led to death, from underlying to immediate, were defined according to International Statistical Classification of Diseases Tenth Revision (ICD-10) criteria.[12] For cases where the pathologist was unsure of the diagnosis, a panel discussion with two or three pathologists was used to determine the causal sequence, which was recorded in an Excel sheet. The causal sequence was then compiled and coded by a single senior medical coder trained in ICD-coding.

Data analysis

The VA interview data were cleaned to remove the additional questions that were asked, and blank columns were added for variables that were removed. Interviews were then analysed using the SmartVA-Analyze program to diagnose most likely causes of death. The SmartVA program redistributes cases for which there is not enough information to assign a most likely cause of death to other VA causes based on: a) the likelihood of each cause being assigned as undetermined using a database of 12,542 VAs for which the true COD was

known; and b) the Brazil COD estimates reported in the Global Burden of Disease Study (GBD).[10,13] As redistribution is an integral part of the SmartVA algorithm, which is designed for predicting population cause of death patterns, the validation was conducted of the redistributed cause of death patterns as opposed to individual causes of death. SmartVA can predict 33 adult causes of death, listed in Table 1 with ICD-10 codes. This cause-list includes nine external causes, including a residual "other injuries" category, which we would not expect to be predicted for many deaths as known non-natural deaths were excluded from the study.

The VA data were also assessed by one physician in São Paulo, experienced in family health, to assign a physician coded verbal autopsy (PCVA) diagnosis. All VA output data, and the corresponding conventional autopsy diagnoses, were grouped into 10 broad causes for the validation of SmartVA (Table 1). The rationale for this is that the real-world utility of VA generally lies in the broad assessment of population data to assess the stage of the epidemiological transition that a population has reached by measuring the leading causes of mortality broadly, to guide policy debates.

Cause specific mortality fraction (CSMF) accuracy was calculated for ten broad cause categories, both from the application of SmartVA and PCVA, in each case compared to conventional autopsy. The CSMF for each cause is calculated as the number of deaths from that cause divided by all deaths. CSMF accuracy is defined as one minus the sum of all absolute CSMF errors across causes divided by the maximum total error.[14] The metric varies between zero and one; the higher the value, the more accurate are the cause of death diagnoses compared with the reference (autopsy) standard,

Broad cause	SmartVA Cause	ICD-10 codes
Cardiovascular diseases	Ischaemic heart disease	I20-I25
	Stroke	I60-I69
	Other cardiovascular diseases	I00-I19 I26-I59, I70-I99
Other NCDs	Other NCDs	All other ICD-10 codes if greater than 12 years of age
Cancers	Breast cancer	C50
	Cervical cancer	C53
	Colorectal cancer	C18-C21
	Oesophageal cancer	C15
	Leukaemia/lymphoma	C81-C85
	Lung cancer	C34
	Prostate cancer	C61
	Stomach cancer	C16
	Other cancers	C00-C14, C17, C22-C33, C35-C49, C51-C52, C54-C60, C62-C80, C86-C90, C97-D48
Infections	AIDS	B20-B24
	Diarrhoea/dysentery	A00-A09
	Malaria	B50-B54
	Other infectious diseases	A10-A14, A20-B19, B25-B49, B55-B99
	Pneumonia	J10-J22, J85
	TB	A15-A19
Diabetes	Diabetes	E10-E14
Chronic respiratory disease	Chronic respiratory disease	J40-J46
Cirrhosis	Cirrhosis	K70-K76
Renal failure	Chronic kidney disease	N17-N19
Maternal	Maternal	O00-O99
External	Homicide	X85-Y09
	Falls	W00-W19
	Road traffic	V01-V89
	Drowning	W65-W74
	Fires	X00-X19
	Bite of venomous animal	X20-X29
	Poisonings	X40-X49
	Suicide	X60-X84
	Other injuries	S00-T98, V90-V99, W20-W64, W75-W99, X30-X39, X50-X59, Y10-Y98

***Table 1*: Adult causes of death from SmartVA mapped to 10 broad causes with corresponding International Statistical Classification of Diseases Tenth Revision (ICD-10) codes**

	Age group				
	18-44 years	45-64 years	65-84 years	85+ years	Total
Study deaths					
Male *n (%)*	135 (4.3)	626 (19.9)	745 (23.7)	164 (5.2)	1,670 (53.2)
Female *n (%)*	85 (2.7)	372 (11.9)	672 (21.4)	340 (10.8)	1,469 (46.8)
Total *n (%)*	220 (7.0)	998 (31.8)	1,417 (45.1)	504 (16.1)	3,139 (100.0)
GBD estimates					
Male %	4.1	14.9	24.9	8.2	52.0
Female %	3.0	10.1	22.1	12.9	48.0
Total %	7.0	25.0	46.9	21.0	100.0

***Table 2*: Mortality by sex and age group in the study sample compared to GBD estimates**

regardless of the number of causes. GBD 2019 data were used as an additional comparator to assess the plausibility of the cause of death distributions from autopsy and SmartVA.[15] This is an important consideration as SmartVA is primarily designed to understand population cause of death patterns, not individual causes of death.[4]

Ethical considerations

This study was approved by the Research Ethics Committee of Hospital das Clínicas, University of São Paulo School of Medicine; reference number 17261814.8.0000.0068. Relatives of all decedents included in the study provided written informed consent.

Role of the funding source

This study was funded by: Brazil Ministry of Health (grant number: 815781/2014); FAPESP São Paulo Research Foundation (grant number: 2013/21728-2); and FAPESP SPRINT University of Melbourne (grant number: 2016/502215). The funder of the study played no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

The study included 2,060 deaths in adults aged 18 years or older in São Paulo from 2,286 next of kin approached to participate (90.1% included); and 1,079 adult deaths in Recife. Conventional autopsy was performed and verbal autopsy interviews were conducted with relatives for all deaths. All VA interviews were analysed using the SmartVA-Analyze software. PCVA was only available for the São Paulo dataset.

The age and sex structure of the autopsy dataset compared to GBD estimates for Brazil is shown in Table 2. The greatest difference was a higher proportion of male deaths in the 45-64-year age range compared to GBD (19.9% vs 14.9%).

The cause specific mortality fractions estimated using SmartVA were broadly similar to autopsy for: cardiovascular diseases, cancers, infections and chronic respiratory disease, a group of causes that account for 76.1% of the autopsy dataset (Table 3). Compared to autopsy, the SmartVA estimates were lower for "Other NCDs" (7.8% vs 14.6%) and higher for diabetes (13.0% vs 6.6%).

CSMF accuracy of SmartVA for 10 broad causes, using conventional autopsy as the gold standard, was 84.5%. CSMF accuracy for PCVA on the subset of data for which this was available was higher, at 93.0%.

Cardiovascular disease mortality, identified using conventional autopsy, was higher at the São Paulo site than Recife (56.8% vs 48.8%). Infectious mortality, identified using conventional autopsy, was lower at the São Paulo site compared to Recife (5.5% vs 9.7%). The data are also presented by study site in Table 3.

Details of cause specific mortality fractions, before and after redistribution of undetermined cases, are provided in Supplementary table 1. Estimates of stroke mortality were higher from application of both VA methods than autopsy (SmartVA: 20.9%; PCVA: 15.5%; autopsy: 6.5%). The converse was true for ischaemic heart disease, where the VA and PCVA estimates were similar, but lower than predicted from autopsy (SmartVA: 25.4%; PCVA: 26.5%; autopsy: 31.2%). Overall, the total CSMFs for all cardiovascuular diseases were similar across the three data sources.

Discussion

Our study has demonstrated that SmartVA can, with reasonable accuracy, predict the broad cause of death groups that are important for identifying a population's stage of epidemiological transition, which in turn is broadly useful for guiding policy. Differences in CSMF estimates between conventional autopsy and VA for some non-communicable diseases, such as ischaemic disease and diabetes, are to be expected, and are discussed in more detail below. More accurate

Broad cause	Cause specific mortality fractions (%)							
	Recife *N*=1,079		São Paulo *N*=2,060			Combined *N*=3,139		GBD 2019
	Autopsy	SmartVA	Autopsy	SmartVA	PCVA	Autopsy	SmartVA	
Cardiovascular diseases	48.8	44.8	56.8	47.9	53.9	54.0	46.8	33.6
Other NCDs	18.0	8.6	12.8	7.3	13.5	14.6	7.8	14.1
Cancers	9.8	10.6	12.0	10.5	8.9	11.4	10.6	22.4
Infections	9.7	10.2	5.5	10.6	10.5	7.0	10.4	10.9
Diabetes	7.0	13.2	6.3	12.9	3.2	6.6	13.0	5.6
Chronic respiratory disease	2.5	3.7	4.3	4.3	4.9	3.7	4.1	6.5
Cirrhosis	3.7	2.1	2.0	1.5	3.1	2.6	1.7	3.3
Renal failure	0.2	3.5	0.1	2.4	0.8	0.2	2.8	3.5
Maternal	0.2	0.9	0.0	0.8	0.0	0.1	0.8	0.1
External	0.0	2.3	0.0	1.9	0.0	0.0	1.9	-
Undetermined	0.0	0.0	0.0	0.0	1.2	0.0	0.0	-
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

***Table 3*: Cause specific mortality fractions for deaths in Recife and São Paulo by broad cause categories from conventional autopsy, SmartVA, and physician-certified VA; GBD estimates of CSMFs for Brazil are shown as an additional comparator**

identification of specific causes, as might be required for some research purposes, or to more precisely guide specific disease control programs, may require more detailed follow-up than is possible through VA methods. PCVA more accurately predicted the conventional autopsy causes of death in this study than SmartVA. However, the study was designed to validate the performance of SmartVA against other methods and the use of the narrative section of the PCVA by the pathologists may have led to increased concordance between those methods.

The estimate of diabetes mortality from SmartVA was roughly twice as high as autopsy (13.0% vs 6.6%). The estimate from PCVA was considerably lower (3.2%). These discrepancies may be related to the significant challenges associated with identification of the underlying cause for some clusters of conditions, for example between diabetes and cardiovascular diseases in patients with evidence of both conditions, and how the data available using the different diagnostic methods may be systematically interpreted differently. Diabetes and cardiovascular disorders are part of the same pathophysiological cascade and have overlapping characteristics, which may often pose a challenge in selecting the most relevant factor as underlying cause of death. Prevalence of diabetes is known to have increased considerably in Brazil, as well as being underdiagnosed. A study using data from 13 primary health clinics in São Paulo reported diabetes prevalence between 6.7% and 10.7% in women; and between 5.1% and 7.1% in men.[16] Cross sectional studies in São Paulo state have reported diabetes prevalence of 12.1% in 30-69 year olds in Ribeirao Preto, 46% of which were undiagnosed; and 13.5% in 30-79 year olds in São Carlos.[17,18]

In 9.6% of VAs in this study, a previous diagnosis of diabetes was reported so the true prevalence of diabetes may be considerably higher than this. An urban-weighting of the study sample may be expected to increase the SmartVA estimate of diabetes mortality above the GBD estimate of 5.6%. However, the underlying cause that is selected in such cases by the different methods will inevitably vary. Medical certification of cause of death data from Recife suggested cardiovascular complications were the most common recorded underlying cause of death in those aged >50 years when diabetes was reported on the MCCOD.[19] This highlights in the study setting that these conditions are common comorbidities and that cardiovascular causes are usually interpreted as the underlying cause. This would support the higher diabetes mortality estimated by SmartVA, and the low diabetes mortality estimated from PCVA may be explained by the trained physician usually interpreting cardiovascular causes as underlying cause.

The relatively low estimate of "Other NCDs" (close to half that of conventional autopsy: 7.8% vs 14.6%) may be explained by the relatively blunt nature of VA diagnosis. "Other NCDs" is a residual category that includes all ICD codes not specified in the rest of the SmartVA cause list — it consists of multiple different causes. If a small proportion of each of these causes (with symptoms that overlap with those of specific causes identified by SmartVA) were misdiagnosed to the specific SmartVA causes, it would contribute to this underrepresentation of "Other NCDs" by SmartVA.

There are particular challenges related to conventional autopsy diagnosis of some conditions such as myocardial infarction and stroke. Myocardial infarction becomes evident macroscopically about 18 hours after the ischaemic injury in cases that survive that long; and microscopically after about four hours when loss of cross striations and contraction bands on the myocardial fibres may appear as evidence of myocardial infarction.

Identification of myocardial infarction in patients that die within four hours of infarction is therefore particularly challenging, and it is estimated that approximately 50% of deaths in acute myocardial infarction occur in the first hour and the majority before arrival at hospital.[20] Due to these difficulties, in this study, the autopsy diagnosis for myocardial infarction included cases presenting with either: a) sudden death exhibiting classical histopathological signs of necrosis; or b) a family report of acute death preceded by chest pain, with evidence of chronic ischaemic cardiomyopathy, acute pulmonary oedema, and acute vascular congestion on histology, and without evidence of chronic congestive heart failure. It is feasible that these criteria could lead to overestimates of myocardial infarction deaths by the pathologist, particularly in place of other cardiovascular diseases, such as stroke. Indeed, stroke patients are likely to have widespread atherosclerotic changes and acute neurogenic pulmonary oedema is a common complication.

The main difference in the age structure of the study dataset compared to GBD was the higher proportion of male deaths in the 45-64 year age group (19.9% vs 14.9%). The reasons for this are unclear although may reflect constraints on recruitment for this study. The exclusion of injury deaths, which the GBD estimates as 12.1% of adult deaths in Brazil, may be expected to increase the proportion of deaths occurring at older ages, which is not apparent in this study. Injury deaths are generally well recalled by VA respondents and accurately predicted by VA. Their exclusion will inevitably decrease the predictive accuracy of SmartVA. It is worth noting that the (very low) injury mortality predicted by SmartVA is due to the redistribution of undetermined cases.

The estimates of cardiovascular disease mortality were significantly higher in the study sample compared to GBD estimates, which may reflect the study population. The fact that the study was conducted at sites where about 80% of autopsies conducted are for individuals who die of natural causes at home, on the streets, or on arrival at hospital, was expected to provide a relatively representative sample of mortality in these populations, which in turn was expected to be similar to the general Brazil population.[15] However, the discrepancies between GBD estimates and the study sample, particularly the cardiovascular disease mortality, indicate the sample may not be representative of the broader population in Brazil. Whilst this would not affect the validation of a technique purely for the assessment of individual causes, for VA, which is designed for use at the population level, significant differences in the cause patterns may affect the CSMF accuracy.

The SmartVA questionnaire is generally implemented in a highly standardised manner, with exact wording for each question. This is because the wording influences the responses and the whole basis of SmartVA relies on the symptom-cause associations that were identified in the large Population Health Metrics Research Consortium study.[10] It is possible that the changes and additions to the questionnaire in this study could negatively affect the performance of SmartVA. An additional consideration is that only one physician in São Paulo was used in this study to diagnose the PCVA causes, which may affect the reliability of the PCVA output and comparisons compared to the usual practice of using two or three physicians. Due to the subjectivity of physician coding of VA, it is common to use two physicians to diagnose the causes with a third physician to resolve discrepancies, in order to reduce bias.[21]

This study is the first to use conventional autopsy to validate SmartVA methodology. We have discussed expected differences when comparing these vastly different methods, where conventional autopsy aims to accurately diagnose each death, whereas VA is a tool solely aimed at describing population cause of death patterns. However, the results suggest that SmartVA can predict the broad cause of death groups to identify the stage of a population's epidemiological transition even when comparing to a mostly urban cohort eligible for conventional autopsy. VA remains a useful tool for improving our understanding of causes of death in remote areas where medical certification of cause of death by a trained physician is not possible.

Contributors

The study was designed by FM, PHNS, EBF, ALB, CDSdeA, PAdeA, LFFS, LPB, SV, PIdeC, MBdeCA and CMdeO. Control of data acquisition was by PAdeA, LPB, CDSdeA, ALB, EBF, FM, SV, MBdeCA and BAdeA. Autopsies were conducted by PHNS, LFFdaS, AMOR, JRAdeL and DSdeMM. JDH and TA analysed the data and FM, EBF, PHNS, LFFS, ALB, PAdeA, CDSdeA, LPB, LAAP, CMM, TMdaSB, SP, DM and IDR were involved in data interpretation. The manuscript was written by JDH; all authors revised it and approved the final version.

Data Sharing Statement

The autopsy data used in this study are available from the Post-Mortem Verification Services (PMVS) in São Paulo and Recife. However, restrictions apply to the availability of these data, which were used with permission for the current study. Applications for use of the data should be made directly with PMVS in São Paulo and Recife.

Declaration of interests

The authors declare no conflicts of interest.

Acknowledgments

We thank Mauro T. Taniguchi for the ICD10 attribution and Pedro Losco Takecian and Leonardo Tadashi Kamaura for the development of the online version of the questionnaire and of the data storage system. We are grateful to all interviewers who applied VA during the study period and to Lisie Tocci Justo for the supervision of the interviews in São Paulo and Raquel Aquino de Souza for the supervision in Recife. We thank Cândida Pereira, Maria Lígia Leite Teixeira de Araújo, Flávio Santos de Azevedo and Carolina Cândida da Cunha for the support given to the Project development in Recife.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.lana.2021.100081.

References

1 The University of Melbourne CRVS Gateway. The International Form of Medical Certificate of Cause of Death. https://crvsgateway.info/The-International-Form-of-Medical-Certificate-of-Cause-of-Death~356.
2 Hart JD, Sorchik R, Bo KS, et al. Improving medical certification of cause of death: Effective strategies and approaches based on experiences from the Data for Health Initiative. *BMC Med* 2020. https://doi.org/10.1186/s12916-020-01519-8.
3 de Savigny D, Riley I, Chandramohan D, et al. Integrating community-based verbal autopsy into civil registration and vital statistics (CRVS): System-level considerations. *Glob Health Action* 2017. https://doi.org/10.1080/16549716.2017.1272882.
4 Hazard RH, Buddhika MPK, Hart JD, et al. Automated verbal autopsy: from research to routine use in civil registration and vital statistics systems. *BMC Med* 2020;**18**(1):60. https://doi.org/10.1186/s12916-020-01520-1.
5 Fantahun M, Fottrell E, Berhane Y, Wall S, Högberg U, Byass P. Assessing a new approach to verbal autopsy interpretation in a rural Ethiopian community: The InterVA model. *Bull World Health Organ* 2006;**84**(3):204–10. https://doi.org/10.2471/BLT.05.028712.
6 McCormick TH, Li ZR, Calvert C, Crampin AC, Kahn K, Clark SJ. Probabilistic Cause-of-Death Assignment Using Verbal Autopsies. *J Am Stat Assoc* 2016. https://doi.org/10.1080/01621459.2016.1152191.
7 Serina P, Riley I, Stewart A, et al. Improving performance of the Tariff Method for assigning causes of death to verbal autopsies. *BMC Med* 2015. https://doi.org/10.1186/s12916-015-0527-9.
8 Serina P, Riley I, Stewart A, et al. A shortened verbal autopsy instrument for use in routine mortality surveillance systems. *BMC Med* 2015. https://doi.org/10.1186/s12916-015-0528-8.
9 James SL, Flaxman AD, Murray CJL. Performance of the Tariff Method: Validation of a simple additive algorithm for analysis of verbal autopsies. *Popul Health Metr* 2011. https://doi.org/10.1186/1478-7954-9-31.
10 Murray CJL, Lopez AD, Black R, et al. Population Health Metrics Research Consortium gold standard verbal autopsy validation study: Design, implementation, and development of analysis datasets. *Popul Health Metr* 2011. https://doi.org/10.1186/1478-7954-9-27.
11 Aase S. Autopsy – still the gold standard? *Tidsskr Den Nor legeforening* 2013;**133**(7). https://doi.org/10.4045/tidsskr.13.0293. 730-730.
12 World Health Organization. International Statistical Classification of Diseases and Related Health Problems, 10th Revision ICD-10. *World Heal Organ* 2016.
13 Roth GA, Abate D, Abate KH, et al. 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. https://doi.org/10.1016/S0140-6736(18)32203-7.
14 Murray CJL, 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. https://doi.org/10.1186/1478-7954-9-28.
15 Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2019 (GBD 2019) Results. Seattle, United States: Institute for Health Metrics and Evaluation (IHME). http://ghdx.healthdata.org/gbd-results-tool. Published 2020. Accessed December 11, 2020.
16 Miraglia JL, Mafra ACCN, Monteiro CN, Borges LM. The variation of the burden of hypertension and diabetes in two large districts of the city of São Paulo, Brazil, based on primary health care routinely-collected data. *PLoS One* 2019. https://doi.org/10.1371/journal.pone.0213998.
17 Bosi PL, Carvalho AM, Contrera D, et al. Prevalence of diabetes and impaired glucose tolerance in the urban population of 30 to 79 years of the city of São Carlos, São Paulo. *Arq Bras Endocrinol Metabol* 2009. https://doi.org/10.1590/s0004-27302009000600006.
18 de Almeida-Pititto B, Dias ML, de Moraes ACF, Ferreira SRG, Franco DR, Eliaschewitz FG. Type 2 diabetes in Brazil: Epidemiology and management. *Diabetes, Metab Syndr Obes Targets Ther*. 2015. doi:10.2147/DMSO.S72542
19 Melo MS, de Lolio CA, Lucena MA, Kirzner CF, Martins SM, Barros MN. Multiple causes of death in diabetics in the municipality of Recife, 1987. *Rev Saude Publica* 1991. https://doi.org/10.1590/s0034-89101991000600004.
20 Mechanic OJ, Gavin M, Grossman SA. Acute Myocardial Infarction. StatPearls. Treasure Island (FL): StatPearls Publishing; 2021 Jan-. [Updated 2021 Mar 9]Available from: https://www.ncbi.nlm.nih.gov/books/NBK459269/.
21 Joshi R, Praveen D, Jan S, et al. How much does a verbal autopsy based mortality surveillance system cost in rural India? *PLoS One* 2015. https://doi.org/10.1371/journal.pone.0126410.