



RESEARCH ARTICLE

REVISED **Does recall time matter in verbal autopsies? Evidence from urban informal settlements in Nairobi, Kenya [version 2; peer review: 2 approved]**

Donnie Mategula ^{1*}, Judy Gichuki ^{2*}

¹Malawi-Liverpool-Wellcome Trust Clinical Research Programme, P.O. Box 30096, Blantyre, Malawi

²Nairobi City County Government, Health Services Department, P.O. Box 34349 -00100, Nairobi, Kenya

* Equal contributors

V2 **First published:** 16 Sep 2020, 5:217
<https://doi.org/10.12688/wellcomeopenres.16243.1>
Latest published: 01 Mar 2021, 5:217
<https://doi.org/10.12688/wellcomeopenres.16243.2>

Abstract

Background: To assign a cause of death to non-medically certified deaths, verbal autopsies (VAs) are widely used to determine the cause of death. The time difference between the death and the VA interview, also referred to as recall time, varies depending on social and operational factors surrounding the death. We investigated the effect of recall time on the assignment of causes of death by VA.

Methods: This is a secondary analysis of 2002-2015 survey data of the Nairobi Urban Health Demographic Surveillance System (NUHDSS). The independent variable recall time was derived from the date of death and the date when the VA was conducted. Univariate and multivariate logistic regression methods were used to calculate odds ratios of assigning a cause of death in defined categories of recall time.

Results: There were 6218 deaths followed up between 2002 and 2016, out of which 5495 (88.3%) had VAs done. Recall time varied from 1-3001 days (median 92 days, IQR 44-169 days). Majority of the VAs (45.7%) were conducted between 1-3 months after death. The effect of recall time varied for different diseases. Compared to VAs conducted between 1-3 months, there was a 24% higher likelihood of identifying HIV/AIDS as the cause of death for VAs conducted 4-6 months after death (AOR 1.24; 95% CI 1.01-1.54; p-value = 0.043) and a 40% increased chance of identifying other infectious diseases as the cause of death for VAs conducted <1 month after death (AOR 1.4; 95% CI 1.02-1.92, p-value = 0.024).

Conclusions: Recall time affected the assignment of VA cause of death for HIV/AIDS, other infectious diseases, maternal/neonatal and indeterminate causes. Our analysis indicates that in the urban informal setting, VAs should be conducted from one month up to 6 months after the death to improve the probability of accurately assigning the cause of death.

Open Peer Review

Reviewer Status

Invited Reviewers

	1	2
version 2 (revision) 01 Mar 2021	 report	 report
version 1 16 Sep 2020	 report	 report

1. **Lucia D'Ambruso** , University of Aberdeen, Aberdeen, UK
 University of the Witwatersrand, Johannesburg, South Africa

2. **Aaron S. Karat** , London School of Hygiene & Tropical Medicine, London, UK

Any reports and responses or comments on the article can be found at the end of the article.

Keywords

Nairobi, Verbal autopsy, Cause of death, Recall, time



This article is included in the [Malawi-Liverpool Wellcome Trust Clinical Research Programme gateway](#).

Corresponding author: Donnie Mategula (dmategula@medcol.mw)

Author roles: **Mategula D:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Gichuki J:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This secondary analysis was not funded by any grant however DM is funded by the Wellcome under the Wellcome Masters Fellowship in Public Health and Tropical Medicine (grant number 205324 Collection of data in the NUHDSS was funded by the Swedish International Development Cooperation Agency – Sida, William and Flora Hewlett Foundation, the Wellcome and Bill and Melinda Gates Foundation.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2021 Mategula D and Gichuki J. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Mategula D and Gichuki J. **Does recall time matter in verbal autopsies? Evidence from urban informal settlements in Nairobi, Kenya [version 2; peer review: 2 approved]** Wellcome Open Research 2021, 5:217 <https://doi.org/10.12688/wellcomeopenres.16243.2>

First published: 16 Sep 2020, 5:217 <https://doi.org/10.12688/wellcomeopenres.16243.1>

REVISED Amendments from Version 1

This manuscript has been updated to address comments from Reviewers 1 and 2. The main changes are:

- Calculation of the cause specific mortality fractions has been redone taking into consideration of all the three probable causes of death assigned by the interVA-4 output.

- We clarify the cut off points for recall time.

- We provide more details for using the 1–3 recall time as reference category in the regression analysis.

- We provide more discussion around the social and economic circumstances affecting the timing of conducting verbal autopsies.

Any further responses from the reviewers can be found at the end of the article

Introduction

Mortality data collected as part of vital registries, disease surveillance systems and epidemiological studies is essential for decision making¹. The need for its accuracy and reliability cannot be overemphasized²⁻³. Currently, the gold standard method for assigning the specific cause of death is through complete diagnostic autopsy⁴. The alternative, where this is not possible, is certification by a medical practitioner using guidelines stipulated in the International classification of diseases and related health problems (ICD), currently available in the eleventh version⁵.

In countries where the vital registration systems are not fully developed, or in situations where there are challenges in medical certification of deaths, such as a home death, verbal autopsies (VAs) are an important means of assigning the cause of death^{6,7}. The VA consists of an interview using standardized questionnaires with the deceased person's close relatives or caregivers who is aware of the circumstances leading to the death⁷. The most widely used VA questionnaires are by the World Health Organization⁷ and the Population Health Metrics Research Consortium⁸.

After completion of the VA interview, data from the VA questionnaire is then interpreted to assign the cause of death. Methods for assigning the cause of death from the VA interview data vary and include the use of either physician certified verbal autopsies (PCVA) or computer coded verbal autopsy (CCVA) systems that utilise algorithms, statistical techniques, machine learning and deep learning approaches^{9,10}. A systematic review comparing PCVAs with various CCVA systems found that although the methods differed in the cause of death output, none of the VA interpretation methods reviewed was superior to the others⁹. Use of CCVA may however decrease the time and cost associated with PCVAs and improve consistency and comparability¹.

A key element in the reliability of cause of death data as determined through VAs is the time between when the verbal autopsy was conducted and when the death occurred, also referred to as the recall time¹¹. There are varying optimal recall time to achieve maximum validity of a VA available in literature ranging from as soon as the death occurs up to 12 months and beyond¹²⁻¹⁴. Data from a study in South Asia indicates that the probability of assigning a correct cause of death by VA methods decreases by 0.55% per increasing month of the recall time¹¹. In contrast, a study in South Africa found that apart from neonatal causes, there was no impact on the validity of VA for recall time as long as one year¹⁵. It is of importance that the timing of the VA not only takes account of the validity but also social and cultural factors¹⁶. As such, it may be imperative to set context-specific optimal timing for VAs. In this paper, we aim to determine the optimal timing of conducting a VA to achieve higher odds of assigning the accurate cause of death in a low-income setting. We use mortality data collected from the Nairobi Urban Health Demographic Surveillance System (NUHDSS). This NUHDSS is run by the African population and research Centre (APHRC) in the informal settlements of Korogocho and Viwandani in Nairobi, Kenya¹⁷.

Methods

Study setting

This paper utilizes mortality data collected from two informal settlements in Nairobi Kenya, that form the NUHDSS, a DSS system run by APHRC. Korogocho is located in Ruaraka Sub-County in Nairobi and covers an area of 0.9 square km with a total population of 36,900 and a density of 42,401 persons per km² as per the 2019 Kenya population and housing census¹⁸. Viwandani is located in Makadara Sub-County, it covers an area of 5 km² with a population of 43,070 and a density of 8554 persons per square km¹⁸. Viwandani forms part of Nairobi's industrial region and its inhabitants are mainly casual labourers within the industries in the area, while Korogocho is inhabited by residents mainly engaged within the informal job sector. Maps for the NUHDSS can be accessed elsewhere¹⁹.

Key challenges in Korogocho and Viwandani, as in many other urban informal settlements, include minimal formal infrastructure such as roads and piped water networks as well as rising cases of communicable and non-communicable diseases, rampant insecurity, environmental pollution, alcohol and drug abuse¹⁷.

Design

This secondary analysis uses VA data collected through a series of surveys by the APHRC in the NUHDSS. The general methodology of the NUHDSS is summarized elsewhere²⁰. In summary, these were a series of surveys conducted from 2002 to 2015 using methodologies described by The International Network for the Demographic Evaluation of Populations and their Health (INDEPTH) network²¹. Surveys were done every four months. Information on the events surrounding the death within the NUHDSS was collected by field

interviewers and later independently validated by three physicians to assign a probable cause of death. Data from the questionnaire was then used as inputs in the [InterVA-4 software](#), a tool that uses probabilistic models based on Bayes' theorem to interpret symptom data and determine possible causes of death²². The output from the InterVA-4 model consists of up to three likelihoods per case attributed to different causes. An indeterminate residual portion is assigned when the sum of the probability of the three likelihoods is less than 100%²². For this study, we utilized the output from the InterVA-4 model to carry out analysis of optimal recall time.

Data analysis

We used STATA version 15.1 in the analysis and included all records where a successful VA was conducted. The main independent variable of interest in the analysis, recall time, is derived from the date of death and the date when the VA was conducted. Based on consideration of the average mourning time, and the acceptable time to conduct a VA as per the Kenya Verbal Autopsy Guidelines²³, we categorized recall time into less than 1 month, 1 to 3 months, 4 to 6 months, 7 to 12 months and greater than 12 months. The upper limit integer month for each category included all the possible decimals before the next category (for example, 1–3 months included all possible days from 1 month up to 3.99 months). In the descriptive analysis, we tabulate the recall time against other background characteristics to calculate their frequencies in the categories of recall time.

For each death, there were three possible causes as an output from the InterVA-4 model. The interVA-4 output was converted into cause-specific mortality fractions considering all the three probable causes of death assigned. The dependent variables were derived from each specific probable cause of death, transformed into a binary variable of yes/no. There were 22 most probable causes of death as an output in the interVA-4 model. To reduce data sparsity in our modelling, we collapsed these 22 probable causes of death into 12 categories guided by their counts and “reasonable relatedness”. Frequently occurring diseases such as tuberculosis and HIV/AIDS were retained in their original categories. Asthma, diabetes mellitus, chronic obstructive pulmonary disease (COPD) and other chronic diseases were reclassified as other chronic/non-communicable diseases. Injuries were grouped together with other external causes of death. Malaria and anaemia were grouped together given their limited numbers and that malaria is a major cause of anaemia in this region²⁴. Direct obstetric and neonatal causes were combined into one category as maternal/neonatal causes. Details of this re-categorization of the probable causes of death are found in [Table 1](#). This approach to re-categorization has also been applied in other studies^{15,20}.

To investigate the effect of recall time on determining the cause of death outcome of the VA, logistic regression methods were used to calculate odds ratios of assigning a probable cause of death using the categories of recall time defined above.

The 1 to 3 months recall time was used as the reference category for the logistic regression as the highest number of VAs were conducted within this period and this falls within the Kenya VA guidelines on appropriate time to conduct a VA. In the univariate logistic regression analysis, association of recall time and probable cause of death outcome of the VA was tested using chi square tests while in the multivariate analysis, likelihood ratio tests were carried out. Wald tests were used to determine the effect of various categories of recall time. The backward modelling approach was used in the multivariate analysis to adjust for the background characteristics.

Ethical considerations

Permission to use the dataset was obtained from APHRC. The original NUHDSS study obtained ethical clearance from Kenya Medical Research Institute (KEMRI) Ethical Review Committee (Ref number KEMRI/RES/7/3/1)

Table 1. Cause of death re-categorization criteria.

Cause of death as per interVA-4	Re-classified cause of death groups
HIV/AIDS	HIV/AIDS
Asthma	Other chronic/NCDs
Anaemia	Malaria/Anaemia
Cardiovascular	Cardiovascular
COPD	Other chronic/NCDs
Diabetes mellitus	Other chronic/NCDs
Diarrhoeal diseases	Other infectious diseases
Direct obstetric	Neonatal/Maternal
Infectious diseases	Other infectious diseases
Injuries	Injuries/Other external
Liver disease	Other chronic/ NCDs
Malaria	Malaria/Anemia
Malnutrition	Other chronic/NCDs
Meningitis	Meningitis
Neonatal causes	Neonatal/Maternal
Malignancies	Malignancies
Respiratory tract infections	Respiratory tract infections
Tuberculosis	Tuberculosis
Other communicable disease	Other infectious diseases
Other non-communicable disease	Other chronic/NCDs
Other external cause	Injuries/Other external
Indeterminate	Indeterminate

Results

Distribution of deaths in the NUHDSS 2002–2016

There were a total of 6218 deaths followed up between 2002 and 2016, of which the highest number were in 2011(533; 8.6%), out of which 5495 (88.3%) had VAs conducted. Most of the deceased were between the ages of 15–49 (3147, 50.6%). Distribution of deaths by year and age is shown in [Figure 1](#). The most common cause of death in the period of 2002 to 2016 was tuberculosis and the least commonly identified cause of death was anaemia. The cause-specific mortality fractions are shown in [Table 2](#).

Description of VA recall time

The range of recall time for deaths with VAs done was from 1 day to 3001 days, with a median recall time of 92 days (Interquartile range 44–169 days). Majority of the verbal autopsies (45.7%) were conducted between one to 3 months of death. Distribution of socio-demographic characteristics varied based on the various verbal autopsy recall periods as shown in [Table 3](#). There were more VAs conducted within the first month of death between 2010 and 2013 (26%) as compared to the other years under study.

The effect of recall time on VA cause of death assignment

The effect of recall time varied for different causes of death ([Table 4](#)). We found strong evidence (p-value = 0.0124) that identifying HIV/AIDS as the likely cause of death varied based on

the timing of the VA. Conducting the VA 4 to 6 months after the death increased the likelihood of identifying HIV/AIDS as the cause of death by 24% compared to conducting it between 1 to 3 months (AOR 1.24; 95% CI 1.01-1.54; p-value = 0.043).

Both the crude and the adjusted analysis showed that an assignment of malaria/anaemia, meningitis, malignancies and other chronic diseases as the probable cause of death did not depend on the timelines within which the VA was conducted. We observed that assignment of tuberculosis as the probable cause of death was dependent on the recall time in the unadjusted analysis (p value= 0.0442) but this effect was lost in the adjusted analysis.

In the crude analysis, recall time did not affect the identification of other infectious diseases as the probable cause of death but we found an effect of recall time in the adjusted analysis (p-value = 0.024). Compared to doing the VA one to three months after the death, conducting it in less than one month increased the chance of identifying other infectious diseases as the probable cause of death by 40 per cent (AOR 1.4; 95% CI 1.02-1.92). Similarly, identifying maternal /neonatal causes of death did not depend on the recall time in the crude analysis but in the adjusted analysis, there was an effect of recall time with a 43% lower chance of assigning probable cause of death to maternal/neonatal causes in VAs done less than one month compared to those done between

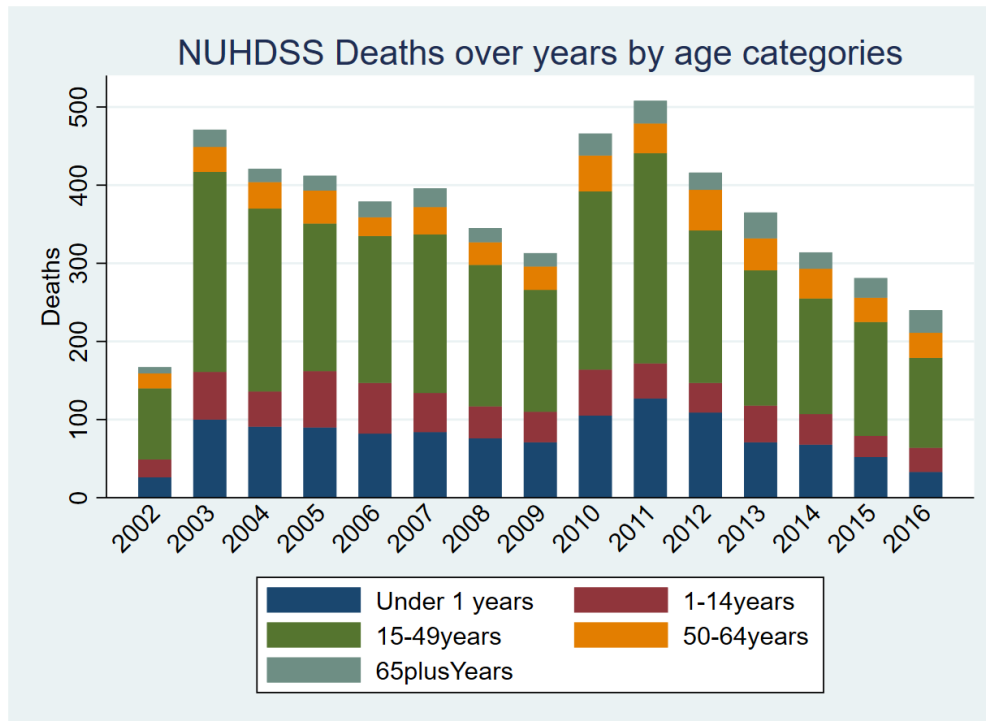


Figure 1. Number of deaths over years by age categories in the NUHDSS.

Table 2. Cause-specific mortality fractions in the NUHDSS between 2002–2016.

Cause of death	CSMF*
HIV/AIDS	15.60
Anaemia	0.12
Asthma	0.35
Cardiovascular	6.01
COPD**	0.46
Diabetes mellitus	0.15
Diarrheal diseases	1.64
Direct obstetric	0.86
Infectious diseases	2.25
Injuries	14.07
Liver disease	0.27
Malaria	1.32
Malnutrition	0.40
Meningitis	4.22
Neonatal causes	6.30
Malignancies	3.18
Respiratory tract infections	10.24
Tuberculosis	19.14
Other communicable disease	3.14
Other non-communicable disease	2.93
Other external cause	0.61
Indeterminate	6.73
Total	100

*= Cause specific mortality fractions

**= Chronic Obstructive Pulmonary Disease

1 to 3 months (AOR 0.57; 95%CI 0.40-0.82; p-value=0.002). Compared to a recall period of 1–3 months, cases with recall periods of >12 months were 41% less likely to be classified as indeterminate (AOR 0.59 ; 95%CI 0.39-0.88; p value=0.01).

Discussion

In this secondary analysis, we investigate the effect of recall time on the cause of death assigned through the InterVA-4 software. There was a notable variation in the timing of VAs with a higher percentage of VAs being conducted within three months of death in the latter years of follow up (2010 – 2016) as compared to the earlier years (2002–2009) suggesting an improved death notification and follow up system in the latter years.

The recall time in a VA is crucial as it determines if the respondent(s) can accurately recount the prevailing symptoms, signs and probable diagnosis before the deceased person's death²⁵. The interVA-4, as well as other VA systems, rely on the accuracy of this information to assign the possible cause of death. In our adjusted analysis, we found a significant difference in the odds of the assignment of a cause of death as HIV/AIDS, other infectious disease and maternal/neonatal causes by recall time. We found that there was a 40% higher chance of identifying other infectious diseases as the cause of death in our analysis for VAs conducted less than one month of death as compared to those conducted between one to three months.

The timing of VAs should consider prevailing circumstances to the death and ensure the appropriate mourning period be observed based on the accepted cultural norms and religion. In a study²⁶ looking at effects of culture and ethics on VAs in Ghana, it was noted that the timing should take into consideration the respondent's emotional distress, for example from the death of an only child or a maternal death. The mourning period may also vary based on the relationship between the interviewee(s) and the deceased as well as the nature of the death and role of the deceased person in the household²⁶. The Kenya 2019 verbal autopsy guidelines provide for a 30–40-day period of mourning before conducting a VA with a maximum allowable period of one year after the death²³.

In the NUHDSS, the VA was preceded by the field supervisor's visit to the deceased home to determine the appropriate time to conduct the VA, based on an assessment of the situation and the availability of a credible respondent²⁰. A maximum of up to five visits were conducted to contact the deceased household members and where it was established that they were no longer residents of the area, a willing credible neighbour was interviewed.

Some of the factors that could affected the timing of the VA are migration of the deceased family following the death due to changes in the economic situation and other factors. Mourning period's in Kenya also vary based on tribe and cultures, with some cultures having extensive periods of mourning as compared to others, while in some cultures, the length of the mourning period will be dependent on marital status, age, gender, role in community, and birth order²⁷. Data on tribe/ethnicity was, however, not available for further analysis in this study.

Our results, in general, find no effect of recall time on identifying the cause of death by VA for most chronic diseases (except HIV/AIDS) and indeterminate causes as compared to acute conditions. We observe that VA respondents in situations where the deceased had a chronic disease are more likely to remember the symptoms, signs and the probable cause of death overtime.

Delay of VA for periods beyond 12 months has been hypothesised to lead to a less accurate recall of symptoms¹⁴. Our analysis did not show any significant differences for any

Table 3. Distribution of background characteristics by recall time.

Recall time	< 1 month	1–3 Months	4–6 Months	7–12 Months	>12 Months	Total
	N= 5,494(100%)					
	936 (17.04)	2,509 (45.67)	1,012 (18.42)	579 (10.54)	458 (8.34)	5,494 (100%)
Variable						
Gender						
Female	371 (15.85)	1,071 (45.77)	457 (19.53)	247 (10.56)	194 (8.29)	2340 (100)
Male	565 (17.91)	1438 (45.59)	555 (17.60)	332 (10.53)	264 (8.37)	3154 (100)
Age						
<1 year	167 (14.09)	551 (46.50)	223 (18.82)	144 (12.15)	100 (8.44)	1185 (100)
1–14years	108 (15.84)	294 (43.11)	126 (18.48)	89 (13.05)	65 (9.53)	682 (100)
15–49 years	509 (18.36)	1234 (44.52)	508 (18.33)	268 (9.67)	253 (9.13)	2772 (100)
50–64 years	95 (18.16)	264 (50.48)	93 (17.78)	47 (8.99)	24 (4.59)	523 (100)
65 plus	57 (17.17)	166 (50.00)	62 (18.67)	31 (9.34)	16 (4.82)	332 (100)
Residence						
Korogocho	523 (16.24)	1491 (46.29)	614 (19.06)	346 (10.74)	247 (7.67)	3221 (100)
Viwandani	413 (18.18)	1018 (44.81)	397 (17.47)	233 (10.26)	211 (9.29)	2272 (100)
Year of death						
2002–2005	148 (10.06)	636 (43.24)	286 (19.44)	161 (10.94)	240 (16.32)	1471 (100)
2006–2009	142 (9.91)	578 (40.33)	331 (23.10)	231 (16.12)	151 (10.54)	1433 (100)
2010–2013	462 (26.32)	917 (52.25)	245 (13.96)	80 (4.56)	51 (2.91)	1755 (100)
2014–2016	184 (17.04)	378 (45.27)	150 (17.96)	107 (12.81)	16 (8.34)	835 (100)
Place of death						
House	286 (15.04)	878 (46.19)	364 (19.15)	209 (10.99)	164 (8.63)	1901 (100)
Health facility	432 (16.88)	1162 (45.39)	479 (18.71)	271 (10.59)	216 (8.44)	2560 (100)
Enroute to health facility	65 (18.62)	150 (42.98)	63 (18.05)	44 (12.61)	27 (7.74)	349 (100)
Other	151 (22.47)	315 (46.88)	102 (15.18)	54 (8.04)	50 (7.44)	672 (100)
Missing	2 (16.67)	4 (33.33)	4 (33.33)	1 (8.33)	1 (8.33)	12 (100)
Respondent lives in the HH						
Yes	470 (19.00)	1118 (45.19)	441 (17.83)	272 (10.99)	173 (6.99)	2474 (100)
No	360 (18.58)	873 (45.05)	362 (18.68)	198 (10.22)	145 (7.48)	1938 (100)
Not in the universe*	67 (10.29)	312 (47.93)	110 (16.90)	52 (7.99)	110 (16.90)	651 (100)
Missing	39 (9.05)	206 (47.80)	99 (22.97)	57 (13.23)	30 (6.96)	431 (100)
Sought health care						
Yes	662 (15.43)	1957 (45.62)	814 (18.97)	485 (11.31)	372 (8.67)	4290 (100)
No	263 (23.11)	524 (46.05)	186 (16.34)	87 (7.64)	78 (6.85)	1138 (100)
Don't know	4 (10.53)	16 (42.11)	8 (21.05)	3 (7.89)	7 (18.42)	38 (100)
Missing	7 (25.00)	12 (42.86)	4 (14.29)	4 (14.29)	1 (8.34)	28 (100)

*Implies that the response given did not apply to the context of the deceased individual

Table 4. Crude and adjusted logistic regression analysis to examine the effect of recall time on the VA cause of death assignment groups.

Disease/Recall time	Number (n)	OR (95% CI)	P-Value ^a	P-value ^b	AOR (95%CI) ^c	P- Value ^a	P-Value ^b
HIV/AIDS							
<1 Month	96	0.7 (0.55-0.89)	0.004	<0.0001	0.85 (0.66-1.09)	0.198	0.0124
1-3 Months	352	Ref			Ref		
4-6 Months	178	1.31 (1.07-1.60)	0.008		1.24 (1.01-1.54)	0.043	
7-12months	99	1.26 (0.99-1.61)	0.060		1.16 (1.89-1.52)	0.256	
>12 months	63	0.98 (0.73-1.30)	0.876		0.77 (0.57-1.05)	0.106	
Malaria -Anaemia							
<1 Month	10	0.67 (0.33-1.34)	0.254	0.3285	0.72 (0.35-1.46)	0.359	0.3606
1-3 Months	40	Ref			Ref		
4-6 Months	15	0.93 (0.51-1.69)	0.808		0.91 (0.50-1.67)	0.767	
7-12months	8	0.86 (0.40-1.86)	0.710		0.72 (0.33-1.57)	0.408	
>12 months	12	1.66 (0.86-3.19)	0.128		1.65 (0.82-3.32)	0.158	
Tuberculosis							
<1 Month	139	0.85 (0.70-1.06)	0.148	0.0442	0.90 (0.71-1.13)	0.381	0.7997
1-3 Months	424	Ref			Ref		
4-6 Months	182	1.08 (0.89-1.31)	0.440		1.01 (0.81-1.24)	0.963	
7-12months	89	0.89 (0.70-1.15)	0.373		0.89 (0.67-1.18)	0.437	
>12 months	96	1.30 (1.02-1.67)	0.036		1.05 (0.79-1.41)	0.731	
Respiratory tract infections							
<1 Month	84	1.05 (0.81-1.37)	0.707	0.0714	1.17 (0.88-1.54)	0.280	0.0585
1-3 Months	215	Ref			Ref		
4-6 Months	95	1.11 (0.86-1.42)	0.438		1.12 (0.86-1.47)	0.380	
7-12months	62	1.28 (0.95-1.72)	0.105		1.12 (0.81-1.54)	0.493	
>12 months	58	1.54 (1.13-2.11)	0.006		1.70 (1.21-2.41)	0.002	
Meningitis							
<1 Month	31	0.82 (0.54-1.23)	0.333	0.4481	0.83 (0.55-1.26)	0.378	0.4409
1-3 Months	101	Ref			Ref		
4-6 Months	38	0.93 (0.64-1.36)	0.709		0.89 (0.60-1.31)	0.553	
7-12months	22	0.94 (0.59-1.51)	0.802		0.83 (0.51- 1.34)	0.447	
>12 months	25	1.38 (0.88-2.16)	0.164		1.35 (0.08-2.19)	0.217	
Other infectious diseases							
<1 Month	68	1.16 (0.86-1.55)	0.329	0.0953	1.40 (1.02-1.92)	0.038	0.0214
1-3 Months	159	Ref			Ref		
4-6 Months	52	0.80 (0.58-1.10)	0.176		0.75 (0.54-1.06)	0.102	
7-12months	48	1.34 (0.95-1.87)	0.091		1.11 (0.78-1.60)	0.560	
>12 months	25	0.85 (0.55-1.32)	0.474		0.75 (0.47-1.20)	0.227	

Disease/Recall time	Number (n)	OR (95% CI)	P-Value ^a	P-value ^b	AOR (95%CI) ^y	P- Value ^a	P-Value ⁵
Maternal/neonatal							
<1 Month	42	0.61 (0.43-0.86)	0.005	0.0618	0.57 (0.40-0.82)	0.002	0.024
1-3 Months	179	Ref			Ref		
4-6 Months	70	0.97 (0.73-1.29)	0.820		1.01 (0.75-1.36)	0.962	
7-12months	37	0.97 (0.62-1.28)	0.527		0.98 (0.67-1.43)	0.91	
>12 months	28	0.85 (0.56-1.28)	0.431		1.02 (0.66-1.58)	0.938	
Malignancies							
<1 Month	30	1.14 (0.74-1.75)	0.562	0.7464	1.08 (0.69-1.71)	0.726	0.5190
1-3 Months	71	Ref			Ref		
4-6 Months	24	0.83 (0.52-1.33)	0.448		0.79 (0.48-1.29)	0.340	
7-12months	17	1.04 (0.61-1.78)	0.890		1.05 (0.59-1.86)	0.871	
>12 months	16	1.24 (0.72-2.16)	0.440		1.44 (0.79-2.65)	0.228	
Injuries/other external causes							
<1 Month	221	1.62 (1.34-1.94)	<0.001	>0.001	1.28 (0.96-1.70)	0.094	0.2626
1-3 Months	403	Ref			Ref		
4-6 Months	143	0.86 (0.70-1.06)	0.152		1.06 (0.78-1.44)	0.701	
7-12months	80	0.84 (0.65-1.09)	0.180		1.37 (0.93-2.011)	0.108	
>12 months	64	0.85 (0.64-1.13)	0.260		0.91 (0.58-1.42)	0.673	
Cardiovascular							
<1 Month	66	1.34 (0.99-1.82)	0.057	0.0202	1.36 (0.98-1.89)	0.064	0.0858
1-3 Months	134	Ref			Ref		
4-6 Months	64	1.19 (0.88-1.62)	0.252		1.15 (0.83-1.60)	0.394	
7-12months	24	0.77 (0.49-1.20)	0.240		0.78 (0.48-1.25)	0.299	
>12 months	16	0.64 (0.38-1.09)	0.100		0.70 (0.40-1.24)	0.226	
Other Chronic NCDs							
<1 Month	42	1.03 (0.72-1.49)	0.180	0.911	0.97 (0.67-1.41)	0.880	0.4877
1-3 Months	109	Ref			Ref		
4-6 Months	46	1.05 (0.74-1.49)	0.792		1.11 (0.78-1.60)	0.563	
7-12months	30	1.20 (0.79-1.82)	0.382		1.36 (0.88-2.09)	0.162	
>12 months	23	1.16 (0.73-1.85)	0.518		1.40 (0.86-2.28)	0.175	
Indeterminate							
<1 Month	107	0.88 (0.69-1.11)	0.268	0.0027	0.92 (0.71-1.17)	0.490	0.0339
1-3 Months	322	Ref			Ref		
4-6 Months	105	0.79 (0.62-0.99)	0.044		0.77 (0.60-0.98)	0.036	
7-12months	63	0.82 (0.62-1.10)	0.200		0.80 (0.59-1.09)	0.158	
>12 months	32	0.51 (0.35-0.74)	0.000		0.59 (0.39-0.88)	0.011	

^aWald p value. ^bchi square p value ^yVariables adjusted in the model: Gender, Age, Slum area, Year of death, Place of death, Respondent lives in household, Sought health care ⁵Likelihood ratio p value. OR = odds ratio, CI = confidence interval, AOR = adjusted odds ratio, Ref = reference group, HIV = human immunodeficiency virus, AIDS = acquired immunodeficiency syndrome, NCDs = non-communicable diseases.

of the identified causes of deaths for recall periods above 12 months apart from the indeterminate causes of death. Compared to a recall period of 1–3 months, cases with recall periods of >12 months were 41% less likely to be classified as indeterminate. In comparison, in a study in South Africa, the only significant difference for recall periods over 12 months were for neonatal deaths¹⁵.

In the regression analyses, we do not adjust for seasonality as seasons and weather patterns in Kenya have varied considerably over the period of analysis in the paper²⁸, it would be less prudent to include seasonality in the model. Additionally, the NUHDSS is located in an urban informal settlement where disruptions of participant availability for reasons like farming are less as compared to a similar study done in South Africa¹⁵.

The interVA-4 outputs consist of three likelihoods, other approaches to the analysis include inclusion of all the possible causes of death and their likelihoods as weights to the regression models. However, where there is need to recategorize the causes of the death, assumptions will have to be made on how the likelihoods combine before and after the categorisation, whether additive, multiplicative or other complex form of combination. We avoid these assumptions by conducting the regression analysis with the most probable cause of death as assigned by the interVA-4. We acknowledge that there might be alternative preferences to this.

The ideal method of identifying the effect of recall time would have been a comparison of the cause of death generated from gold standard death certification to the VA assigned cause of death for the different recall times. However, in the absence of the gold standard comparator, as in our case, the alternative

approach used in this paper was to compare probabilities or odds of making a specific VA cause of death assignment for different recall periods. To our best knowledge this is the first paper within the East African region to expound on the effects of recall time in VAs.

Conclusion

Recall time affected the assignment of VA cause of death for HIV/AIDS, other infectious diseases, maternal /neonatal and indeterminate causes. Our analysis indicates that in the urban informal setting, VAs should be conducted from one month up to 6 months after the death to improve the probability of accurately assigning the cause of death.

Data availability

The KENYA - NUHDSS - Verbal Autopsy, Causes of deaths 2002–2015 dataset was used in this secondary analysis. The dataset is available upon request from the APHRC Microdata Portal <http://microdataportal.aphrc.org/index.php/catalog>. Access to the dataset can be obtained through submission of a written request in the APHRC portal following creation of an account. Further information on how to apply for access to the data can be found [here](#).

Author contributions

Conceptualization: DM, JG; Data curation: DM, JG Formal Analysis: DM, JG; Writing-original draft preparation: DM, JG; Writing-Review & Editing: DM, JG

Acknowledgements

The authors wish to acknowledge APHRC as the source of the data utilized for the secondary analysis.

References

- Byass P: **Who Needs Cause-of-Death Data?** *PLoS Med.* 2007 [cited 2020 Jun 26]; 4(11): e333. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- AbouZahr C, De Savigny D, Mikkelsen L, *et al.*: **Civil registration and vital statistics: Progress in the data revolution for counting and accountability.** *Lancet.* Lancet Publishing Group; 2015 [cited 2020 Jun 26]; 386(10001): 1373–85. [PubMed Abstract](#) | [Publisher Full Text](#)
- Thomas LM, D'Ambruoso L, Balabanova D: **Verbal autopsy in health policy and systems: A literature review.** *BMJ Global Health.* BMJ Publishing Group; 2018 [cited 2020 Jun 26]; 3(2): e000639. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Menéndez C, Quintó L, Castillo P, *et al.*: **Limitations to current methods to estimate cause of death: A validation study of a verbal autopsy model. [version 1; peer review: 2 approved, 1 approved with reservations]** *Gates Open Res.* 2020 [cited 2021 Feb 2]; 4: 55. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- WHO: **International classification of diseases for mortality and morbidity statistics (11th Revision).** 2018 [cited 2020 Jun 26]. [Reference Source](#)
- Murray CJL, Lozano R, Flaxman AD, *et al.*: **Using verbal autopsy to measure causes of death: The comparative performance of existing methods.** *BMC Med.* 2014; 12(1): 5. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- WHO: **Verbal Autopsy Standards: The 2016 WHO verbal autopsy instrument.** 2017; 1–143. [Reference Source](#)
- 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; 9: 27. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Leitao J, Desai N, Aleksandrowicz L, *et al.*: **Comparison of physician-certified verbal autopsy with computer-coded verbal autopsy for cause of death assignment in hospitalized patients in low- and middle-income countries: Systematic review.** *BMC Med.* 2014; 12(1): 22. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Mapundu M, Kabudula C: **Overview of Statistical and Machine Learning techniques for Determining Causes of Death from Verbal Autopsies: A Systematic Literature Review.** 2020; [cited 2021 Jan 3]. [Publisher Full Text](#)
- Serina P, Riley I, Hernandez B, *et al.*: **What is the optimal recall period for verbal autopsies? Validation study based on repeat interviews in three populations.** *Popul Health Metr.* 2016 [cited 2020 May 20]; 14(1): 40. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Byass P, D'Ambruoso L, Ouédraogo M, *et al.*: **Assessing the repeatability of verbal autopsy for determining cause of death: two case studies among**

- women of reproductive age in Burkina Faso and Indonesia. *Popul Health Metr.* 2009 [cited 2020 Jun 27]; **7**: 6.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
13. 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 [cited 2020 Jun 27]; **10**(1): 1272882.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 14. McLaughlin D: **CRVS best-practice and advocacy Introducing verbal autopsies into CRVS systems: Guiding principles.** 2020 [cited 2020 Jun 27].
[Reference Source](#)
 15. Hussain-Alkhateeb L, Petzold M, Collinson M, *et al.*: **Effects of recall time on cause-of-death findings using verbal autopsy: empirical evidence from rural South Africa.** *Emerg Themes Epidemiol.* 2016; **13**(1): 10.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 16. Huong DL, Van Minh H, Byass P: **Applying verbal autopsy to determine cause of death in rural Vietnam.** *Scand J Public Health.* 2003 [cited 2020 Jun 26]; **62**: 19–25.
[PubMed Abstract](#) | [Publisher Full Text](#)
 17. Beguy D, Elung'ata P, Mberu B, *et al.*: **Health & Demographic Surveillance System Profile: The Nairobi Urban Health and Demographic Surveillance System (NUHDSS).** *Int J Epidemiol.* 2015 [cited 2020 Jun 27]; **44**(2): 462–71.
[PubMed Abstract](#) | [Publisher Full Text](#)
 18. KNBS: **2019 Kenya Population and Housing Census: Volume II i.**2019.
[Reference Source](#)
 19. Wamukoya M, Kadengye DT, Iddi S, *et al.*: **The Nairobi Urban Health and Demographic Surveillance of slum dwellers, 2002– 2019: Value, processes, and challenges.** *Glob Epidemiol.* 2020; **2**: 100024.
[Publisher Full Text](#)
 20. Oti SO, Kyobutungi C: **Verbal autopsy interpretation: A comparative analysis of the InterVA model versus physician review in determining causes of death in the Nairobi DSS.** *Popul Health Metr.* 2010; **8**(1): 21.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 21. **Founding documents | INDEPTH Network.** [cited 2020 Jun 26].
[Reference Source](#)
 22. Byass P, Fottrell E, Huong DL, *et al.*: **Refining a probabilistic model for interpreting verbal autopsy data.** *Scand J Public Health.* 2006 [cited 2020 Aug 5]; **34**(1): 26–31.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 23. Ministry of Health; Kenya: **Verbal Autopsy Standards and Guidelines.** 2019.
[Reference Source](#)
 24. White NJ: **Anaemia and malaria.** *Malar J.* BioMed Central Ltd.; 2018 [cited 2021 Jan 4]; **17**(1): 371.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 25. Joshi R, Kengne AP, Neal B: **Methodological trends in studies based on verbal autopsies before and after published guidelines.** *Bull World Heal Organ.* 2009 [cited 2020 Jul 4]; **87**(9): 678–82.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 26. Aborigo RA, Allotey P, Tindana P, *et al.*: **Cultural imperatives and the ethics of verbal autopsies in rural Ghana.** *Glob Health Action.* 2013 [cited 2020 Jul 4]; **6**(1): 18570.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 27. Njue JRM, Rombo D, Lutomia AN, *et al.*: **Death, Grief and Culture in Kenya: Experiential Strengths-Based Research.**2015; 3– 23.
[Publisher Full Text](#)
 28. Warner K, van de LP: **Climate Change Profile, Kenya.** Minist Foreign Aff Netherlands. 2018; 14.
[Reference Source](#)

Open Peer Review

Current Peer Review Status:  

Version 2

Reviewer Report 06 April 2021

<https://doi.org/10.21956/wellcomeopenres.18326.r42886>

© 2021 D'Ambruoso L. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Lucia D'Ambruoso 

¹ University of Aberdeen, Aberdeen, UK

² University of the Witwatersrand, Johannesburg, South Africa

Overall comments:

Thank you for the opportunity to review the revised paper, which appears to have responded to comments on the earlier version. This is an interesting paper considering recall time as a key a methodological aspect of the VA method. The final, published piece might usefully be checked for grammar.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Global public health; qualitative methods; health policy and systems research; South Africa

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 22 March 2021

<https://doi.org/10.21956/wellcomeopenres.18326.r42887>

© 2021 Karat A. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Aaron S. Karat 

TB Centre, London School of Hygiene & Tropical Medicine, London, UK

My thanks to the authors for their careful and detailed responses and my apologies for the delay in re-reviewing. I am happy with the changes made and have no further concerns - congratulations on the paper!

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Tuberculosis, epidemiology, infectious diseases, verbal autopsy.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 30 Mar 2021

Donnie Mategula, Malawi-Liverpool-Wellcome Trust Clinical Research Programme , P.O Box 30096, Blantyre, Malawi

Thank you for the review and the kind words.

Competing Interests: No competing interests were disclosed.

Version 1

Reviewer Report 08 December 2020

<https://doi.org/10.21956/wellcomeopenres.17841.r41360>

© 2020 Karat A. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Aaron S. Karat

TB Centre, London School of Hygiene & Tropical Medicine, London, UK

Thank you for the opportunity to review this paper, which is generally well written and deals with an important aspect of verbal autopsy implementation. This is a secondary analysis of VA data from more than 5,000 deaths and, though the findings are interesting, I have some concerns around some of the methods and around the interpretation of results.

Major comments:

1. Calculation of CSMFs - The method used to calculate CSMFs is potentially problematic. InterVA output (up to three causes with likelihoods +/- residual likelihood) is intended to be converted into CSMFs in a way that takes account of all causes of death assigned (see Appendix 4 of the InterVA-4 user guide), not only the most likely cause, as described by the authors. The 'collapsing' process described seems perfectly justifiable but may have been better applied after the calculation of CSMFs to reflect better the true output of InterVA.

Adopting this approach would likely have produced different CSMFs for each recall period, and may therefore have affected the results of the regression.

2. Logistic regression - Although most factors likely to affect the relationship between recall time and cause of death were adjusted for, time of year/seasonality was not. This is potentially very important, as it may have affected mortality patterns but also affected the ability of VA teams to access respondents (e.g. this may have been harder in the rainy season) and therefore affected time between death and VA. This may be less relevant in the urban setting where this study was conducted (compared with the rural setting in the study by Hussain-Alkhateeb *et al.*) but this should at least have been discussed and the omission justified by the authors.
3. Interpretation of results - My apologies if I have misunderstood, but I do not entirely follow some of the statements made in the discussion and I am not sure the conclusions flow from the results presented. The authors suggest that respondents may be more likely to remember the symptoms of an acute illness if the VA is done in a shorter time after death, yet in this analysis the odds of being assigned a respiratory tract infection CoD were higher for VAs done after >12 months compared with 1–3 months.
4. I also do not entirely follow the possible explanation for why maternal deaths were less likely to be assigned <1 months after death than 1–3 months. Are the authors suggesting that fewer maternal deaths were assigned because VAs may have been delayed when a particular cause of death was suspected? It would seem important to know if this was done when any other cause of death was suspected, or if it was only for maternal deaths. The authors also state that the Kenyan guidelines recommend 30–40 days mourning prior to conducting a VA, but there is no description of how often these guidelines were followed (and/or if they were followed more closely in one district than another) – this is a potentially a major source of bias and should ideally have been explored in more detail in the analysis.

Minor comments:

Introduction:

1. **Paragraph 1:** The ‘gold standard’ for cause of death is one assigned by complete diagnostic autopsy, not by a medical practitioner, though of course for the majority of deaths a practitioner-assigned CoD is often the best that can be hoped for. It may be worth making this point a little more clearly in the first paragraph.
2. **Paragraph 2:** This is a very minor point, but the multiple plurals used here (‘interviews’, ‘questionnaires’, ‘relatives’, ‘caregivers’) suggests that a standard VA involves interviewing more than one person, when this is rarely the case.
3. **Paragraph 2:** I believe the full name of the organisation is the Population Health Metrics Research Consortium. A better reference when discussing the VA instrument (instead of the paper currently cited in Ref 7) may be the ‘VA tools’ page on the [PHMRC website](#).
4. **Paragraph 3:** the list of CCVA methods is slightly dated (the ref provided is a 2014 review). [This pre-print](#) may provide a more up-to-date list (InterVA is also misspelled in this paragraph, as “inter-VA”).

Methods:

1. If possible, please could the authors provide the exact cut-offs (in days) used for the recall time categories. At present it is not entirely clear if a VA conducted after ~3.5 months would be categorised as "1–3 months" or "4–6 months" (same for recall times between 6 and 7 months).
2. There may be a good reason for using 1–3 months as a reference instead of <1 month but the reason for this approach is not described in the methods. In my opinion this makes it more difficult to interpret the results and to look at trends across recall periods (regardless of statistical 'significance').

Results

1. **Page 4:** it may be more interesting to show the median + IQR for recall time, given that it is unlikely to be normally distributed?
2. **Table 2:** It is not clear what 'Not in the universe' means (under 'Respondent lives in the HH').
3. **Table 2:** Typo under 'year of death ("2104").
4. **Table 3:** It would make the results more transparent if the CSMF proportion for each cause was shown in this table, very ideally for each recall period for each cause, so readers are clear on the 'number' of outcomes associated with each OR/aOR.
5. **Table 3:** Typo in the 95% CI for the aOR for HIV/AIDS deaths with a recall time of 7–12 months ("10.89").

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?

Partly

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Partly

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Tuberculosis, epidemiology, infectious diseases, verbal autopsy.

I confirm that I have read this submission and believe that I have an appropriate level of

expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 03 Feb 2021

Donnie Mategula, Malawi-Liverpool-Wellcome Trust Clinical Research Programme , P.O Box 30096, Blantyre, Malawi

Dr Aaron S.Karat Many thanks for your helpful and thoughtful comments. We have re-submitted an updated manuscript which addresses your suggestions and concerns. Please find our detailed responses below.

Major comments:

1) Calculation of CSMFs - The method used to calculate CSMFs is potentially problematic. Inter-VA output (up to three causes with likelihoods +/- residual likelihood) is intended to be converted into CSMFs in a way that takes account of all causes of death assigned (see Appendix 4 of the InterVA-4 user guide), not only the most likely cause, as described by the authors. The 'collapsing' process described seems perfectly justifiable but may have been better applied after the calculation of CSMFs to reflect better the true output of InterVA. Adopting this approach would likely have produced different CSMFs for each recall period, and may therefore have affected the results of the regression.

Response: Thank you for the comment. The cause specific mortality fractions have been re-calculated as suggested and are presented in table 2 of the revised paper. However, we have not used the CSMFs for the regression analysis as adding the cause of death specific weights/likelihoods to a model where the cause of death have been collapsed makes assumptions on how the cause of death weights/likelihoods combine before and after the re-categorization. To avoid making these assumptions, we prefer to use the most likely cause of death in the logistic regression analysis. However, we have included this limitation in the discussion.

2) Logistic regression - Although most factors likely to affect the relationship between recall time and cause of death were adjusted for, time of year/seasonality was not. This is potentially very important, as it may have affected mortality patterns but also affected the ability of VA teams to access respondents (e.g. this may have been harder in the rainy season) and therefore affected time between death and VA. This may be less relevant in the urban setting where this study was conducted (compared with the rural setting in the study by Hussain-Alkhateeb *et al.*) but this should at least have been discussed and the omission justified by the authors.

Response: Seasons and weather patterns in Nairobi have varied considerably over the period of analysis in the paper, it would have been less prudent to include seasonality in the model. Additionally, we agree with the reviewer that this variable is less relevant in this context unlike in the Hussain Alkhateeb *et al* paper. We have included text in the discussion

section to reflect our rationale for the lack of inclusion of this variable in the model.

3) Interpretation of results - My apologies if I have misunderstood, but I do not entirely follow some of the statements made in the discussion and I am not sure the conclusions flow from the results presented. The authors suggest that respondents may be more likely to remember the symptoms of an acute illness if the VA is done in a shorter time after death, yet in this analysis the odds of being assigned a respiratory tract infection CoD were higher for VAs done after >12 months compared with 1–3 months.

Response: In our analysis, there is no overall association between recall time and RTI based on the likelihood ratio test (P value 0.0585) and that is why we did not focus on RTIs. Our general observation that respondents are more likely to remember acute illness if the VA is done in a shorter time still holds.

4) I also do not entirely follow the possible explanation for why maternal deaths were less likely to be assigned <1 months after death than 1–3 months. Are the authors suggesting that fewer maternal deaths were assigned because VAs may have been delayed when a particular cause of death was suspected? It would seem important to know if this was done when any other cause of death was suspected, or if it was only for maternal deaths. The authors also state that the Kenyan guidelines recommend 30–40 days mourning prior to conducting a VA, but there is no description of how often these guidelines were followed (and/or if they were followed more closely in one district than another) – this is a potentially a major source of bias and should ideally have been explored in more detail in the analysis.

Response: Thank you for the comment. We have updated the discussion section to further explain the potential sources of delay to conducting a VA in the NUHDSS. This was not specific to maternal deaths but to the prevailing circumstances. We have edited the discussion to focus less on the maternal/neonatal finding. While the numbers for combined direct obstetric and neonatal causes were low, the majority of these were neonatal deaths. The effect observed could therefore have been a chance finding due to the few numbers or as a result of the higher proportion of neonatal deaths to direct obstetric causes (6 to 1) especially in cases where the neonatal deaths were not reviewed promptly.

Minor comments:

Introduction:

5) Paragraph 1: The 'gold standard' for cause of death is one assigned by complete diagnostic autopsy, not by a medical practitioner, though of course for the majority of deaths a practitioner-assigned CoD is often the best that can be hoped for. It may be worth making this point a little more clearly in the first paragraph.

Response: This has been updated

6) Paragraph 2: This is a very minor point, but the multiple plurals used here ('interviews', 'questionnaires', 'relatives', 'caregivers') suggests that a standard VA involves interviewing

more than one person, when this is rarely the case.

Response: Multiple plurals have been reviewed and revised

7) Paragraph 2: I believe the full name of the organisation is the [Population Health Metrics Research Consortium](#). A better reference when discussing the VA instrument (instead of the paper currently cited in Ref 7) may be the 'VA tools' page on the [PHMRC website](#).

Response: The changes have been made and the reference updated

8) Paragraph 3: the list of CCVA methods is slightly dated (the ref provided is a 2014 review). [This pre-print](#) may provide a more up-to-date list (InterVA is also misspelled in this paragraph, as "inter-VA").

Response: We have updated the list of CCVA methods based on the updated reference, however we have retained our original citation as well since Mapundo et al paper is in a preprint.

Methods:

9) If possible, please could the authors provide the exact cut-offs (in days) used for the recall time categories. At present it is not entirely clear if a VA conducted after ~3.5 months would be categorised as "1–3 months" or "4–6 months" (same for recall times between 6 and 7 months).

Response: Thank you for the comment, the highest cutoff integer month for our categories included all the possible decimals before the next category, for example 1-3 months includes all possible days up to 3.99. We have updated this in the methods section.

10) There may be a good reason for using 1–3 months as a reference instead of <1 month but the reason for this approach is not described in the methods. In my opinion this makes it more difficult to interpret the results and to look at trends across recall periods (regardless of statistical 'significance').

Response: The 1 to 3 months recall time was used as the reference category for the logistic regression for two reasons; the highest number of VAs were conducted within this period and secondly this falls within the Kenya VA guidelines on appropriate time to conduct a VA thus we set to compare all the results against this recall period. We have updated this in the methods section.

Results

11) Page 4: it may be more interesting to show the median + IQR for recall time, given that it is unlikely to be normally distributed?

Response: Thanks, we concur. The Median and IQR for recall time have been included.

12) Table 2: It is not clear what 'Not in the universe' means (under 'Respondent lives in the HH').

Response: Implies that the response given did not apply to the context of the deceased under question. We have added a footnote in the table to explain

13) Table 2: Typo under 'year of death ("2104").

Response: This has been amended, thank you

14) Table 3: It would make the results more transparent if the CSMF proportion for each cause was shown in this table, very ideally for each recall period for each cause, so readers are clear on the 'number' of outcomes associated with each OR/aOR.

Response: We have added the number of outcomes associated with each OR/aOR. Please also refer to our response to comment number 1 to justify our choice.

14 Table 3: Typo in the 95% CI for the aOR for HIV/AIDS deaths with a recall time of 7–12 months ("10.89").

Response: This has been amended, thank you

Competing Interests: No competing interests were disclosed.

Reviewer Report 02 December 2020

<https://doi.org/10.21956/wellcomeopenres.17841.r41357>

© 2020 D'Ambruoso L. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Lucia D'Ambruoso 

¹ University of Aberdeen, Aberdeen, UK

² University of the Witwatersrand, Johannesburg, South Africa

Overall comments:

This is an interesting paper examining recall times on the outcomes of VA surveys and contributes to the evidence base on this methodology. The paper is generally well written and clearly presented. Overall comments are as follows: first, the paper would be improved with fuller justification of key assumptions and aspects of the analytical approach around ideal recall times and CoD categories (see specific comments). Second, while I am not well-qualified to comment on statistical aspects, it would appear that not all the statistically significant findings are considered in the narrative. Third, the discussion could be further developed by considering the plausibility of the main findings, offering reasons as to why the recall time effects on some cause of death categories were found. Fourth, further attention to the social and cultural aspects of recall time,

and in the context of Covid-19/other infectious outbreaks would add to the current relevance of the paper. Finally, a statistician should review the paper. As a qualitative researcher, I have not commented on the statistical approach in anything more than general terms. The paper is likely to be worthy of indexing, with attention to these issues.

Specific comments:

Introduction:

1. Paragraph 3: the authors may wish to acknowledge that computer-coded VA addresses observed inconsistencies within and between physician coders, as well as obviating the time implications of physician coding: in the settings where VA is needed, using physician time for VA interpretation is ethically problematic.
2. Paragraph 4: the need to account for social and cultural factors as a key factor determining recall period is convincingly introduced – how the authors deal with this issue is not clear however and could usefully be expanded on in the paper.

Methods:

1. Paragraph 1, Study setting: A map would be good to include, if possible.
2. Paragraph 3, Design: The term demographic health system may be misleading, suggest replace and use HDSS consistently.
3. Paragraph 3, Design: InterVA5 model is currently available, it may be worth considering applying InterVA5 for these data.
4. Page 4, Table 1: the title of this table might be rephrased - strictly speaking, criteria are not presented. Recategorization of InterVA CoDs to constructed CoD categories might be more accurate.
5. Page 4, Table 1: The collapsed CoD categories are not quite clear. The authors state that category classifications were guided by counts and reasonable relatedness - this requires some clarification, especially around amnesia and malaria. The authors also state that the collapsed categories are based on classifications used in other work, however there are 8 categories in the Hussain *et al.* study and 15 in the Oti and Kyobutungi paper. The relationship of these to the 12 categories used in this analysis similarly would benefit from some clarification.
6. Page 4, Table 1: It is not customary to cite AIDS/HIV as a category, rather HIV/AIDS, as is done later in the paper – suggest consistency with commonly used term.
7. Page 4, paragraph 1: CSMFs typically account for all probable CoD assigned per case. Suggest recalculating the CSMF using all probable CoDs from InterVA.
8. It is not entirely clear why 1-3 months recall time is used as the reference period. The paper indeed presents research from the same region where 12 months recall is found to be suitable. The authors may wish to clarify this element of the analysis.

Results:

- While I am not well-qualified to comment on the statistical analyses, it would appear that not all significant findings from the adjusted regression analysis are highlighted in the narrative, please see (d) and (e) below. The main findings would appear to be as follows: Compared to a recall period of 1-3 months.
 - (a) Cases with recall periods of 4-6 months are 24% more likely to be classified as due to HIV;
 - (b) Cases with recall periods of <1 month 40% more likely to be classified as due to other infectious causes.
 - (c) Cases with recall periods of <1 month 43% less likely to be classified as due to maternal causes.
 - (d) Cases with recall periods of >12 months 70% more likely to be classified as due to RTI.
 - (e) Cases with recall periods of >12 months 41% less likely to be classified as indeterminate.

Discussion:

1. The discussion could be further developed by considering the plausibility of the main findings offering reasons as to why the recall time effects on some cause of death categories were found. For example, that cases with recall periods of <1 month are 43% less likely to be classified as due to maternal causes is counter-intuitive – maternal deaths are relatively rare, generally avoidable and therefore likely to be highly memorable.
2. Page 9, paragraph 5: the statement 'our analysis did not show any significant differences for any of the identified causes of deaths for recall periods above 12 months' needs to be checked. Again, from a general perspective, Table 3 appears to suggest that deaths investigated >12 months are (a) 70% more likely to be classified as due to RTI (AOR 1.70 (CI 1.21-2.41) P=0.002) and (b) 41% less likely to be diagnosed by InterVA as indeterminate (AOR 0.59 (CI 0.39-0.88) P=0.011).
3. Further attention to social and cultural influences on recall times, and how these vary by country/region would be useful in this section. Thorough assessment of social and cultural norms prior to VA surveys is of critical importance, whether these factors have an impact on the findings presented in this paper would be useful to consider.
4. Some attention to how the analysis and recommendations on recall periods related to the Covid-19 pandemic and other infectious outbreaks, e.g., Ebola, is likely to be of relevance to readers.
5. The authors may wish to include strengths and limitations section, where a sensitivity analysis with e.g., up to 12 months as an ideal recall period (as is indicated in other research) is applied.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Partly

Are sufficient details of methods and analysis provided to allow replication by others?

Partly

If applicable, is the statistical analysis and its interpretation appropriate?

I cannot comment. A qualified statistician is required.

Are all the source data underlying the results available to ensure full reproducibility?

Partly

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Global public health; qualitative methods; health policy and systems research; South Africa

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 03 Feb 2021

Donnie Mategula, Malawi-Liverpool-Wellcome Trust Clinical Research Programme , P.O Box 30096, Blantyre, Malawi

Dr Lucia D'Ambruoso, Many thanks for your helpful and thoughtful comments. We have re-submitted an updated manuscript which addresses your suggestions and concerns. Please find our detailed responses below:

Introduction:

Paragraph 3: the authors may wish to acknowledge that computer-coded VA addresses observed inconsistencies within and between physician coders, as well as obviating the time implications of physician coding: in the settings where VA is needed, using physician time for VA interpretation is ethically problematic.

Response: Thank you for the suggestion. This has been included.

Paragraph 4: the need to account for social and cultural factors as a key factor determining recall period is convincingly introduced – how the authors deal with this issue is not clear however and could usefully be expanded on in the paper.

Response: We have considered the social factors that were available in the dataset in the

regression analysis. In the discussion, we have expounded on some of the potential social-cultural factors around recall period not available in the dataset.

Methods:

Paragraph 1, Study setting: A map would be good to include, if possible.

A reference to where the NUHDSS map can be accessed has been provided in the revised version.

Paragraph 3, Design: The term demographic health system may be misleading, suggest replace and use HDSS consistently.

Response: This has been updated.

Paragraph 3, Design: InterVA5 model is currently available, it may be worth considering applying InterVA5 for these data.

Thank you for the comment. The data set was provided following the Inter-VA4 analysis therefore it would not be possible to analyze it using Inter-VA5. InterVA-5 was developed to accommodate the WHO 2016 standard VA tool that was not utilized for this data set.

Page 4, Table 1: the title of this table might be rephrased - strictly speaking, criteria are not presented. Recategorization of InterVA CoDs to constructed CoD categories might be more accurate.

Response: Thank you. This has been updated

Page 4, Table 1: The collapsed CoD categories are not quite clear. The authors state that category classifications were guided by counts and reasonable relatedness - this requires some clarification, especially around amnesia and malaria. The authors also state that the collapsed categories are based on classifications used in other work, however there are 8 categories in the Hussain *et al.* study and 15 in the Oti and Kyobutungi paper. The relationship of these to the 12 categories used in this analysis similarly would benefit from some clarification.

Response: We have updated the text to further detail how we did the recategorization. The approach to this classification has been used by other studies as cited, however the final number of categories varies depending on data and specific circumstances in the study areas. The collapsing of malaria and anemia is based on the precinct that malaria is a major cause of anemia in tropical areas.

Page 4, Table 1: It is not customary to cite AIDS/HIV as a category, rather HIV/AIDS, as is done later in the paper – suggest consistency with commonly used term. This has been updated. Thank you.

Page 4, paragraph 1: CSMFs typically account for all probable CoD assigned per case. Suggest recalculating the CSMF using all probable CoDs from InterVA. CSMFs have been recalculated as suggested. It is not entirely clear why 1-3 months recall time is used as the reference period. The paper indeed presents research from the same region where 12 months recall is found to be suitable. The authors may wish to clarify this

element of the analysis.

Response: The 1 to 3 months recall time was used as the reference category for the logistic regression for two reasons; the highest number of VAs were conducted within this period and secondly this falls within the Kenya VA guidelines on appropriate time to conduct a VA thus we set to compare all the results against this recall period. We have updated this in the methods section.

Results:

While I am not well-qualified to comment on the statistical analyses, it would appear that not all significant findings from the adjusted regression analysis are highlighted in the narrative, please see (d) and (e) below. The main findings would appear to be as follows: Compared to a recall period of 1-3 months.

- (a) Cases with recall periods of 4-6 months are 24% more likely to be classified as due to HIV;
- (b) Cases with recall periods of <1 month 40% more likely to be classified as due to other infectious causes.
- (c) Cases with recall periods of <1 month 43% less likely to be classified as due to maternal causes.
- (d) Cases with recall periods of >12 months 70% more likely to be classified as due to RTI.
- (e) Cases with recall periods of >12 months 41% less likely to be classified as indeterminate.

Response: Thank you. We have updated the missing significant results for the indeterminate, however we have left out the RTIs as the association with recall time is not significant on the overall likelihood ratio test of association.

Discussion:

The discussion could be further developed by considering the plausibility of the main findings offering reasons as to why the recall time effects on some cause of death categories were found. For example, that cases with recall periods of <1 month are 43% less likely to be classified as due to maternal causes is counter-intuitive – maternal deaths are relatively rare, generally avoidable and therefore likely to be highly memorable.

Response: While the numbers for combined direct obstetric and neonatal causes were low, the majority of these were neonatal deaths. The effect observed could therefore have been a chance finding due to the few numbers or as a result of the higher proportion of neonatal deaths to direct obstetric causes (6 to 1) especially in cases where the neonatal deaths were not reviewed promptly.

Page 9, paragraph 5: the statement 'our analysis did not show any significant differences for any of the identified causes of deaths for recall periods above 12 months' needs to be checked. Again, from a general perspective, Table 3 appears to suggest that deaths investigated >12 months are (a) 70% more likely to be classified as due to RTI (AOR 1.70 (CI 1.21-2.41) P=0.002) and (b) 41% less likely to be diagnosed by InterVA as indeterminate (AOR 0.59 (CI 0.39-0.88) P=0.011).

Response: Thank you for the observation. We have revised the statement and included the

findings on the indeterminate causes of death. In our analysis, there is no evidence of an overall association between recall time and RTI based on the likelihood ratio test (P value 0.0585) and that is why we did not focus on RTIs in the results section

Further attention to social and cultural influences on recall times, and how these vary by country/region would be useful in this section. Thorough assessment of social and cultural norms prior to VA surveys is of critical importance, whether these factors have an impact on the findings presented in this paper would be useful to consider.

Response: Thank you for the suggestion. We have included an assessment of the potential social and cultural circumstances prior to a VA.

Some attention to how the analysis and recommendations on recall periods related to the Covid-19 pandemic and other infectious outbreaks, e.g., Ebola, is likely to be of relevance to readers.

Response: Thank you for the suggestion, however, given the period within which the VA were conducted (2002-2016) it might be difficult to extrapolate the findings to the current situation of Covid-19.

Competing Interests: There no competing interests
