UNIVERSITYOF BIRMINGHAM

Ability of verbal autopsy data to detect deaths due to uncontrolled hyperglycaemia

Blackstock, Sarah; Witham, Miles D; Wade, Alisha N; Crampin, Amelia; Beran, David; Ogle, Graham D; Davies, Justine I

Link to publication on Research at Birmingham portal

BMJ Open

Ability of verbal autopsy data to detect deaths due to uncontrolled hyperglycaemia: testing existing methods and development and validation of a novel weighted score

Sarah Blackstock,[1] Miles D Witham,[2,3,4] Alisha N Wade,[3] Amelia Crampin,[5,6] David Beran,[7] Graham D Ogle,[8] Justine I Davies[3,9,10]

For numbered affiliations see end of article.

Correspondence to
Professor Justine I Davies;
j.davies.6@bham.ac.uk

ABSTRACT

Objectives Verbal autopsy (VA) is a useful tool to ascertain cause of death where no other mechanisms exist. We aimed to assess the utility of VA data to ascertain deaths due to uncontrolled hyperglycaemia and to develop a weighted score (WS) to specifically identify cases. Cases were identified by a study or site physician with training in diabetes. These diagnoses were also compared with diagnoses produced by a standard computer algorithm (InterVA-4).

Setting This study was done using VA data from the Health and Demographic Survey sites in Agincourt in rural South Africa. Validation of the WS was done using VA data from Karonga in Malawi.

Participants All deaths from ages 1 to 49 years between 1992 and 2015 and between 2002 and 2016 from Agincourt and Karonga, respectively. There were 8699 relevant deaths in Agincourt and 1663 in Karonga.

Results Of the Agincourt deaths, there were 77 study physician classified cases and 58 computer algorithm classified cases. Agreement between study physician classified cases and computer algorithm classified cases was poor (Cohen's kappa 0.14). Our WS produced a receiver operator curve with area under the curve of 0.952 (95% CI 0.920 to 0.985). However, positive predictive value (PPV) was below 50% when the WS was applied to the development set and the score was dominated by the necessity for a premortem diagnosis of diabetes. Independent validation showed the WS performed reasonably against site physician classified cases with sensitivity of 86%, specificity of 99%, PPV of 60% and negative predictive value of 99%.

Conclusion Our results suggest that widely used VA methodologies may be missing deaths due to uncontrolled hyperglycaemia. Our WS may offer improved ability to detect deaths due to uncontrolled hyperglycaemia in large populations studies where no other means exist.

Strengths and limitations of this study

► It is possible for physicians to interpret verbal autopsy (VA) data to ascertain cases of premature mortality due to uncontrolled hyperglycaemia with reasonable inter-rater reliability.
► A simple, validated algorithm can be used to find cases of uncontrolled hyperglycaemia using VA data.
► Methods can only be used where VA systems exist.
► There was no gold standard laboratory diagnosis to confirm our findings.

Both of these conditions occur in individuals with uncontrolled diabetes mellitus and are usually precipitated by intercurrent illness.[1] DKA is classically seen in patients with type 1 diabetes and has inadequate insulin therapy as a frequent precipitant; HHS is classically seen in type 2 diabetes. Either of these hyperglycaemic emergencies may be seen in either type of diabetes.[1] DKA is the leading cause of mortality in younger people with type 1 diabetes[2] and mortality from HHS ranges from 10% to 20%.[3]

Mortality from hyperglycaemic emergencies has decreased significantly in high-income countries (HICs),[1 4] largely due to improved diagnosis and treatment of these conditions and underlying diabetes. By contrast, deaths from acute complications of uncontrolled hyperglycaemia remain high in lower and middle-income countries (LMICs).[5–8] While delays in presenting to health facilities undoubtedly contribute, deaths from hyperglycaemic emergencies are also an indicator of an unmet need for diabetes care after that care is sought. Long-term glycaemic control as well as rapid diagnosis and treatment of decompensated diabetes would significantly reduce mortality from these conditions.

INTRODUCTION

Hyperglycaemic emergencies, namely diabetic ketoacidosis (DKA) and hyperglycaemic hyperosmolar state (HHS), are preventable causes of premature mortality.

Little data exist on the mortality rates in hyperglycaemic emergencies in LMICs and the data that do exist have often been obtained from hospital records[7 8] and therefore may underestimate true mortality as they do not capture those deaths that occur out of hospital. Verbal autopsy (VA) has been developed to address the deficit of accurate, countrywide, reporting of cause of death in many LMICs[9 10] and may improve mortality estimates in these conditions.

VA from population samples is a useful tool to ascertain causes of mortality and trends thereof.[11–13] During VA, a respondent—usually a relative—who cared for the deceased during his or her last illness is asked a set of standard questions by a trained data collector about the illness.[14 15] VA reports are either reviewed by physicians who assign a cause of death or, increasingly, processed automatically by computer models to derive likely causes of death. Such models (eg, InterVA-4) are derived using a mixture of data and expert opinion.[16] They have been shown to be reasonably reliable in determining causes of death which are commonly seen, but data on reliability of these methods for estimating causes of death for less prevalent diseases are lacking.[11 16–18]

Where there are classic features of an acute death from hyperglycaemia (symptom combinations including diagnosis of diabetes, increased thirst, increased urine output, coma and so on), it is relatively straightforward for either a reviewing physician or an automated model like InterVA-4 to arrive at a high likelihood of diabetes as a cause of death. However, for many people living with diabetes, the symptoms occurring around the time of death may be less obvious, especially in contexts where symptoms of hyperglycaemia can be attributable to other, more common conditions. In these cases, increased index of reviewing physician suspicion, for example due to greater exposure to or knowledge of the disease, may improve detection of the condition.

We therefore aimed to assess ability of the VA methodology, using InterVA-4 algorithm determined cause of death, to detect deaths attributable to uncontrolled hyperglycaemia, as compared with diagnoses made by a study physician with experience of diabetes care in a LMIC setting. A further aim was to derive and test a weighted score (WS)—which could later be applied to other settings using VA reports—for detecting deaths due to uncontrolled hyperglycaemia and validate this WS in VA data from an independent data set. Other aims were to compare our WS and the study physician diagnoses. We limited our age range between 1 and 49 years with an aim of focusing particularly on premature mortality due to uncontrolled hyperglycaemia.

METHODS
Setting and VA methodology
Our study was done using VA data from the Agincourt Health and Socio-Demographic Surveillance System (HDSS) and validated using data from the Karonga HDSS in Malawi.[19 20] The Agincourt HDSS is based in the Agincourt sub-district of rural, northeast South Africa, near the Mozambique border; the Karonga HDSS is based in the south of Karonga district, in rural northern Malawi. From Agincourt, we used VA data collected on annual census visits between 1992 and 2015; from Karonga, we used VA data collected at household visits initiated after reporting of a death by a community informant at monthly reporting session between 2002 and 2016. VA methodology is described in detail elsewhere.[16] In brief, for any death, household members are approached and asked to take part in an interview based on standard WHO questionnaires and administered by a local, trained, data collector or medical assistant. VA questionnaires consist of responses (which are converted to binary for processing with InterVA-4) to a range of questions on signs, symptoms and diagnoses during the terminal illness. Some VAs (eg, those in Agincourt and Karonga) also have a 'free text' section where respondents are able to freely describe circumstances leading up to the death. In Agincourt, prior to 2010, cause of death was defined by physician review of each VA. After 2010, all causes of death have been determined using InterVA computer algorithms; all VAs have also been retrospectively formatted for input into InterVA-4 with binary variables (presence of symptom or absence/unknown). In Karonga, both physician review and computer models are used; two physicians review—blind to each other's coding—and allocate underlying cause of death as well as direct and contributory causes.[20] Where there is discrepancy between physicians, a third reviewer considers the VA and the responses of the first two physicians and decides on the cause(s) of death to be coded. Questionnaire and free text responses are de-identified and stored in electronic databases.

Participant selection and creation of study physician coded data set
Given the rarity of presenting with uncontrolled hyperglycaemia in infants under 1 year old, the increased likelihood of deaths being due to other competing causes in those of older years, and our focus on premature mortality, we agreed a priori to restrict our sample age range to deaths occurring between 1 and 49 years of age, inclusive. To enable later application of the WS to other VA data sets, we aligned our age range selection with the WHO 2012 standard VA age groups and therefore included those in VA age groups 'under 5' (1–4), 'child' (5–14) and 'adult' (15–49).

Data sets that use clinical data including hospital diagnoses and laboratory results to provide a gold standard cause of death have been developed and enable testing of standard VA methodologies. However, the VA input parameters collected in these data sets are not complete,[12 21] and although fields of use for diagnosing diabetes—for example polyuria and polydipsia—are present in standard VA questionnaires, they are not captured in these gold standard sets. We therefore created a study physician coded data set that was to act as our 'gold standard'.

A clinician with experience in diabetes management in HICs and LMICs (SB) reviewed all data from VA records at Agincourt. Study physician classified cases were determined using responses to the answers to VA questions and examination of the free text. Cases were defined as those for whom uncontrolled hyperglycaemia would be acceptable as the main cause of death on a standard UK death certificate. Any cases where the reviewing physician was unsure were discussed with clinical colleagues with expertise in adult internal medicine, diabetes and endocrinology (AW, JD and MDW) until consensus was reached.

Given the likely rarity of study physician classified cases, we produced an enriched sample of cases for study physician review by searching the VA database for cases with features suggestive of uncontrolled hyperglycaemia as individual symptoms, symptom combinations or terms (chosen to reflect both chronic and acute symptoms of diabetes) as follows:

► Ante-mortem (AM) diagnosis of diabetes.
► Polyuria.
► Polydipsia.
► Weight loss combined with polyuria or polydipsia.
► Weight loss combined with polyuria or polydipsia and in combination with acute rapid breathing, abdominal pain, confusion or coma.
► 'Sugar' or 'diabetes' in the free text search.
► Site-physician review indication of deaths due to diabetes (in Agincourt VA data prior to 2010).

Any cases that did not have any of the above features were thought clinically unlikely to have died from uncontrolled hyperglycaemia and were therefore categorised as deaths from other causes (hereafter termed 'negative cases').

Comparison of study physician classified cases with computer algorithm classified cases

We compared the predictive value of InterVA-4 (with positive cases termed computer algorithm classified cases) against study physician classified cases in the Agincourt data set using χ^2 and used Cohen's kappa as a measure of inter-rater correlation, with the recognition that the 'raters' in this context included algorithms. As there is no InterVA-4 category of cause of death due to hyperglycaemia, computer algorithm classified cases were defined as determined by InterVA-4 as being greater than 50% likely to be due to diabetes.

Development of predictive score

After producing a set of study physician classified cases and negative cases, we tested the predictive value of VA-recorded variables. These variables were chosen after consensus was reached by study clinicians on which were likely to be seen in uncontrolled hyperglycaemia in clinical practice (as either subacute or chronic features).[22] Additionally, we identified symptom variables which we agreed would reduce likelihood of a death being due to uncontrolled hyperglycaemia and which were captured by the VA questions. We performed univariable testing of each individual symptom for its ability to predict study physician classified cases using the χ^2 test. Symptoms that were significant univariable predictors at p<0.1 (Pearson's χ^2) were entered into a multivariable binary logistic regression model using stepwise entry. A WS was then developed based on the relative beta-weights in the final multivariable model.

We constructed a receiver–operator characteristic (ROC) curve based on the relationship between the WS for each individual VA entry and study physician classification. We then determined the sensitivity, specificity, positive and negative predictive values (PPV and NPV) of the WS at various cut points.

Score validation

We used VA data from Karonga HDSS in Malawi to externally validate our WS. Cases above a cut point determined based on analysis of Agincourt data were extracted for review and these 'WS classified cases' were compared with Karonga site-physician classified cases.

In addition, and to allow for differences between Karonga site-physician diagnosis and one made by a physician with expertise in diabetes and experience in working in an LMIC setting, in a sample of 100 cases we compared WS classified cases with classification by an independent endocrinologist (AW, who determined independent physician classified cases (investigator physician classified cases)). These 100 cases were made up of all WS classified cases plus a random selection of cases not determined to be deaths due to uncontrolled hyperglycaemia .

We compared computer algorithm classified cases and study physician classified cases with WS classified cases using χ^2. We also describe the predictive value of symptoms identified by multivariable analysis in determining computer algorithm classified cases.

Determining timing of diagnosis of diabetes

For study physician classified cases or WS classified cases, we ascertained which had been diagnosed with diabetes prior to or during the final illness by examining responses to the VA question, 'did the deceased have diabetes' and examining the VA free texts. Cases where it was stated in the free text that diagnoses of diabetes were given, or patients were told their sugar levels were high in the final illness, with no noted history of diabetes on VA question response were assumed to have been diagnosed in the final illness. We assumed that diabetes was diagnosed prior to the final illness in cases where it was stated in the free text that patients were known to have diabetes. Cases where there was no mention of diabetes in the free text were classified as unknown.

Statistical analysis

SPSS V.22 was used for all analyses.

Patient and public involvement

No patients were involved in this study.

Figure 1 Study flow chart.

RESULTS

Determining study physician classified cases

There were 15 261 deaths occurring in the Agincourt HDSSs which had a VA report completed between 1992 and 2015, of which 8699 were between the ages of 1 and 49 years. After limiting cases to those with symptoms suggestive of uncontrolled hyperglycaemia, there were 3708 cases which were reviewed by the study physician. There were two cases with missing data. Of all VA reports reviewed, 77 cases were determined as positive (study physician classified cases), 3626 were negative and 3 were deemed indeterminate even after discussion among investigators. Figure 1 shows the flow of classification of VA cases at Agincourt into study physician classified cases and negative cases.

Comparison of study physician classified cases with computer algorithm classified cases

There were 58 computer algorithm classified cases. χ^2 testing showed that there was dependence between study physician classified cases and computer algorithm classified cases (Pearson's χ^2 of 176, 1 degree of freedom, p<0.001); however although they were associated on the χ^2 test (online supplementary appendix table 1), kappa showed poor concordance, with the VA algorithm not finding the majority of cases identified by the physician; Cohen's kappa for inter-rater agreement was low at 0.14.

Development of the WS

On discussion between investigators, we identified AM diagnosis of diabetes, polyuria, polydipsia, sunken eyes, weight loss, wasting, acute rapid breathing, abdominal pain or acute abdominal pain (which are two separate responses on VA), coma and confusion as being variables

collected on VA which were likely to be seen in cases of deaths due to uncontrolled hyperglycaemia. In environments where tuberculosis and other respiratory conditions are common causes of death, we did not consider that other symptoms of breathlessness captured on VA were likely to be discriminating enough for our purposes. Of our potential predictive variables, nine which were significantly associated with study physician classified cases on univariable testing—and thus entered into the multivariable regression analysis—were AM diagnosis of diabetes (p<0.001), polyuria (p=0.001), polydipsia (p<0.001), confusion (p<0.001), weight loss (p=0.001), chronic abdominal pain (p<0.001), abdominal pain (p<0.001), acute rapid breathing (p=0.095) and wasting (p=0.098) (online supplementary appendix table 2).

We identified 27 variables that were captured on the VA questions that were considered to decrease the clinical likelihood of the death being due to uncontrolled hyperglycaemia; these were HIV, tuberculosis (TB), chronic cough, cough, productive cough, bloody cough, TB combined with chronic cough, chronic fever, whooping cough, wheeze, night sweats, chronic diarrhoea, bloody diarrhoea, jaundice, haematemesis, haematuria, abdominal mass, swollen abdomen, swollen legs, injury, died in labour, died 24 hours after labour, vaginal bleeding after menopause, kidney disease, liver disease, cancer or measles. Of these variables, on univariable testing, we found injury (p=0.021), TB combined with chronic cough (p=0.097), night sweats (0.032) and chronic diarrhoea (0.032) were significant negative predictors of positive cases. Haematuria (p=0.06) and measles (p=0.002) were significant predictors of study physician classified cases (table 1 and online supplementary appendix table 2). The positive association between cases and measles lacked face validity. However, haematuria could be a misunderstanding of urinary frequency also seen in uncontrolled hyperglycaemia. We therefore entered TB combined with chronic cough, chronic diarrhoea, injury, night sweats and haematuria into the regression analysis.

The WS was produced using variables that were significantly associated with study physician classified cases on binary logistic regression analysis, scaled to the lowest positive beta weight and rounded up (or down) to the nearest whole number (table 1). This produced a ROC (figure 2) with an area under the curve of 0.952 (95% CI 0.920 to 0.985). Sensitivity, specificity, NPV and PPV of various cut points for the WS applied to the Agincourt data set and compared with study physician cases are shown in table 2.

We chose a cut point of 8 to identify cases with reasonable specificity while maintaining sensitivity. Applying this cut point in the Agincourt data set gave 134 WS classified cases, thus 1.54% (134/8699) of all deaths were estimated due to uncontrolled hyperglycaemia. Characteristics of these deaths are shown in table 3. In particular, all had an AM diagnosis of diabetes recorded; 62 cases had diabetes diagnosed in the final illness and 24 cases had symptoms on the free text which were determined to be suggestive of

Table 1 Binary logistic regression showing variables entered into the weighted score and derived score weights

	Number of cases with symptom	Beta	SE	P value	Weighting (beta/0.751)	Rounded weight
Ante-mortem diagnosis of diabetes	140	6.462	0.357	<0.001	8.6045273	9
Polyuria	265	1.542	0.583	0.008	2.05326232	2
Polydipsia	2539	1.406	0.353	<0.001	1.87217044	2
Confusion	1569	0.751	0.352	0.033	1	1
TB and chronic cough	944	−1.627	0.716	0.023	−2.1664447	−2
Chronic diarrhoea	992	−2.058	0.795	0.01	−2.7403462	−3
Constant	–	−7.114	0.357	<0.001	–	–

TB, tuberculosis.

of uncontrolled hyperglycaemia. Of note, 138 of all VA deaths in the age range of interest from Agincourt had an AM diagnosis of diabetes—4 more than those detected by the algorithm.

Comparison of WS classified cases with computer algorithm classified cases

Online supplementary appendix table 1 shows the performance of our WS in comparison with InterVA-4. χ^2 showed dependence (Pearson's χ^2 of 146, 1 degree of freedom, p<0.001), however, inter-rater agreement was poor (Cohen's kappa).

Validation of the WS

We validated our WS using VA data from Karonga in Malawi. There were 3614 VA reports between 2001 and 2016 and 1663 of these were from people between 1 and 49 years old. Application of our WS to Karonga VA data

Figure 2 Receiver–operator characteristic curve for weighted score applied to study physician coded data set.

and use of a cut point of 8 or above identified 20 cases of deaths due to uncontrolled hyperglycaemia (table 3). Comparison of these 20 WS classified cases with site physician classified cases (either as underlying or as contributory cause) showed 12 WS classified cases which were true positives; the Karonga site physician also identified an additional two case of deaths from diabetes which the algorithm did not detect; thus, giving a sensitivity of 86%, specificity of 99%, PPV of 60% and NPV of 99%.

Of the 20 WS classified cases, all were in the adult age group: 9 between 20 and 29, 5 between 30 and 39 and 6 between 40 and 49 years of age. All cases had a premortem diagnosis of diabetes, and we determined that 4 of these diagnoses were made in the final illness (of note, 22 deaths in the whole Karonga VA data set for this age range had premortem diagnosis of diabetes). Ten relatives reported symptoms on the free text that were suggestive of uncontrolled hyperglycaemia, and for 7 of the 20 WS detected cases, relatives recalled being told by treating healthcare workers that the death was caused by diabetes (table 3).

Out of the 100 cases assessed, the external investigator physician (AW) determined that a total of 22 were due to uncontrolled hyperglycaemia, 74 deaths were not due to uncontrolled hyperglycaemia and 4 were unclassifiable due to missing data. Compared with the external investigator physician, the WS gave 15 true positives and 7 false negatives. All site physician classified cases were deemed positive by the independent physician investigator. However, compared with the site physician, the independent physician investigator determined that there were an additional nine deaths likely to be due to uncontrolled hyperglycaemia (kappa for inter-rater agreement=0.69).

Performance of clinically determined variables in detecting computer algorithm classified cases

Variables which we determined were clinically likely to increase or decrease chance of death due to uncontrolled hyperglycaemia were tested for association with computer algorithm classified cases. In univariable testing, we found AM diagnosis of diabetes (p<0.001), polyuria (p<0.001) and polydipsia (p<0.001) to be
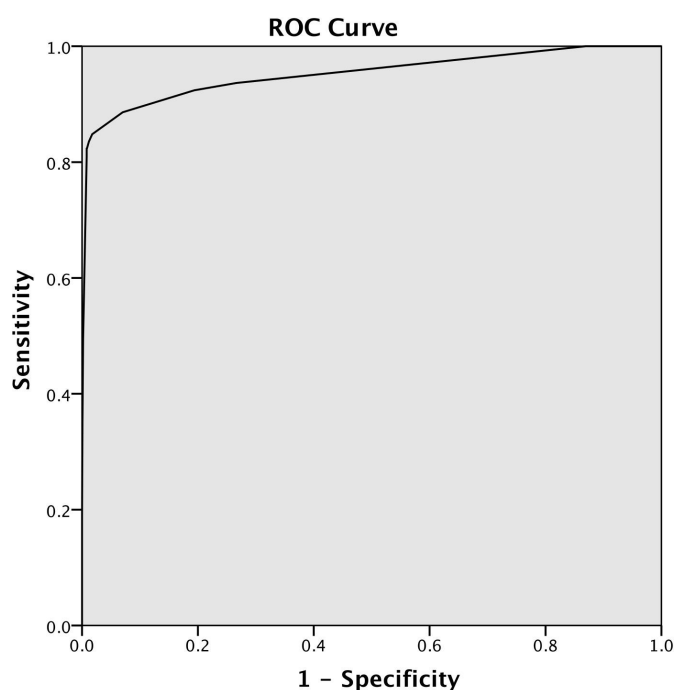
Table 2 Sensitivity and specificity of the weighted score above different score cut points (summed weightings) as applied to Agincourt data and tested against study physician classification of cases

Cut point	Number of deaths above cut point	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
5	169	83.54	98.91	39.05	99.85
6	138	82.27	99.15	47.10	99.84
7	136	82.28	99.17	47.79	99.84
8	134	82.28	99.20	48.01	99.84
9	133	81.01	99.20	48.12	99.82

NPV, negative predictive value; PPV, positive predictive value.

positive predictors in these cases. TB and chronic cough (p=0.025), wasting (p<0.001), sunken eyes (p=0.012), chronic diarrhoea (p=0.006), chronic cough (p<0.001), chronic fever (p=0.003), wheeze (p<0.001), productive cough (p=0.003), night sweats (p<0.001), abdominal swelling (p=0.048) and injury (p=0.03) were significant negative predictors of diabetes. Results of multivariable testing are shown in online supplementary appendix table 3. Online supplementary appendix table 1 shows the performance of InterVA-4 in comparison with the study physician categorisation.

DISCUSSION
There are several key findings from our study. The first is that compared with a physician classification of VA data, the widely used InterVA-4 algorithm performs poorly in detecting cases of deaths thought clinically likely due to uncontrolled hyperglycaemia. Reports from InterVA-4 on numbers of deaths due to diabetes should therefore be interpreted with caution. Second, we found that a large number of deaths due to uncontrolled hyperglycaemia

received only a diagnosis of diabetes in their final illness, especially in the Agincourt sample. VA captures data from carers, rather than health records, and thus this finding may be overly negative. It is nevertheless troubling, and suggests that further investigation of health system's ability to diagnose and manage diabetes is needed. Third, it was possible to develop a WS to detect cases of uncontrolled hyperglycaemia, that, at a cut point of greater than 8, had reasonable sensitivity and specificity. Our WS also had better agreement with study physician classified cases than the InterVA-4 model. On validation in an independent data set, the score also showed good sensitivity and specificity for predicting deaths due to uncontrolled hyperglycaemia both when compared with site physician classified cases and investigator physician classified cases. However, the score was dominated by a premortem diagnosis of diabetes and it could be argued that inclusion of the other factors provided minimal further discriminatory value. Apart from premortem diagnosis of diabetes, the predominant symptoms that our WS detected were related to uncontrolled hyperglycaemia (polydipsia and

Table 3 Characteristics of people identified by the study physician and the algorithm as having died of diabetes

	Agincourt	Agincourt	Karonga
	study physician classified cases	weighted score classified cases (cut point >8)	weighted score classified cases (cut point >8)
Total	77	134	20
Infant	0	3	0
Under 5	6	6	0
Child	3	3	0
Adult	68	122	20
Female	38	71	12
Male	39	63	8
Any recorded AM diagnosis of diabetes	64	134	20
No AM diagnosis of diabetes	13	0	0
AM diagnosis of diabetes made in final illness	37	62	4
AM diagnosis of diabetes made prior to final illness	3	10	10
Unknown when AM diagnosis of diabetes made	24	62	6

AM, ante-mortem.

polyuria). Our score, or indeed, VA cannot discriminate between deaths due to hyperglycaemia-related complications of type 1 or type 2 diabetes. Additionally, given recent evidence from HIC which suggests the fall in incidence of type 1 diabetes with age may not be as steep as previously thought[23] and, from sub-Saharan Africa, where there is some evidence that the peak age of presentation may be older than in other countries,[24 25] limiting the age range of cases is unlikely to be a strategy to enable detection of cases of uncontrolled hyperglycaemia due to type 1 diabetes.

As mentioned, we found that AM recording of a diagnosis of diabetes was the strongest predictor of a death being due to uncontrolled hyperglycaemia, in fact, using our score cut point of 8, it is not possible to assign diabetes as a cause of death without diabetes having been recorded on the VA report. This is a further limitation of our score, making it only applicable in settings where health systems are advanced enough to diagnose diabetes or report hyperglycaemia; unfortunately, in many LMICs, laboratory services are focused on detecting infectious rather than non-communicable diseases.[26] This lack of diagnostic capacity also impacts on the ability of health systems to detect and treat diabetes to prevent untimely deaths. Even in countries with reasonable diagnostic ability, it may not be deployed early enough in the disease course to avert death.[27 28] Such delayed diagnosis is reflected in our finding that diagnosis of diabetes was often made in the final illness, suggesting that the death could have been averted if diagnosis had been made earlier in the illness. Access to diagnostic testing has to be paired with an increased index of suspicion of the diagnosis early in the disease course. We also acknowledge that even if diagnoses are made, hurdles of access to treatment still need to be overcome.[29 30] Unfortunately, given the small numbers of deaths found in this study, we were not able to reliably look at temporal trends in access to care.

Although we found that our WS performance was substantially more reliable in ascertaining cases of deaths due to uncontrolled hyperglycaemia than InterVA-4 when applied to Agincourt, our study aim was to look specifically for this condition, and we produced an algorithm that was optimised to find cases. In contrast, InterVA-4 takes into consideration numerous competing diseases to deliver an adequate performance to determine population-level causes of death across a wide range of diseases[16] and was not developed to detect single diseases. Furthermore, while physicians use both presence and absence of symptoms to determine diagnoses, InterVA-4 predominately relies on presence of symptoms.[31] It is also interesting that there were differences between physicians in ascribing uncontrolled hyperglycaemia as a cause of death. These differences are likely to result from different exposure to disease prevalence and be influenced by the reason for examining the data; it would be expected that physicians who were used to dealing with a condition and who were specifically looking for that condition would

find a greater prevalence of that condition.[32–34] However, that there was reasonable agreement between the investigator physician (AW) and the algorithm with the Karonga site physician diagnoses is reassuring.

There are several limitations of our study. The lack of a gold standard data set which contained both confirmed clinical diagnosis and relevant VA parameters necessitated our use of a physician to ascertain cases, and without laboratory results, the diagnosis can never be certain. To enrich the data set for clinical review with likely cases, we also preselected cases that had one or more responses on VA which could suggest the diagnosis; although it is unlikely that deaths with no symptoms of uncontrolled hyperglycaemia noted would have died of this condition, it is not impossible. The enrichment of the review data set with these cases could also have led to overfitting of the WS. For this WS development study, we limited the age range of cases to between 1 and 49 years to ensure that we detected premature mortality and to avoid confounding from competing symptoms that may be seen in older people who likely have multiple comorbidities. We may have missed cases in older deaths, and how this WS performs in older age groups needs to be the subject of separate study. As in the development of InterVA-4, we, a priori, decided to use clinical knowledge to guide choice of our input variables of our model. While we argue that this is a reasonable method for model development, an alternative approach could have been to assess all VA variables for association and include all those that are statistically significant regardless of clinical validity. Lastly, VA tools to ascertain cause of death are not as accurate as vital statistics reporting which is based on clinical diagnoses. However, such reporting is lacking in many populations, especially in LMICs. In these situations, VA is proven to be a reliable alternative method of ascertaining cause of death.

In summary, we have found that the InterVA-4 algorithm performs poorly in detecting cases of deaths due to uncontrolled hyperglycaemia. Our algorithm improves detection, however is dominated by necessity for a premortem diagnosis of diabetes with other variables adding little discrimination. We also found that a high proportion of deaths due to uncontrolled hyperglycaemia received a diagnosis in their final illness. In countries where information on numbers of deaths due to uncontrolled hyperglycaemia is lacking and where VA reports exist, our algorithm can be used to give an indication of the numbers of deaths due to the condition, hence expose health system gaps in the provision of care which would result in earlier diagnosis and treatment.

Author affiliations
[1]Department of Paediatric Rheumatology, Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK
[2]Department of Aging and Health, School of Medicine, University of Dundee, Dundee, UK
[3]MRC/Wits Rural Public Health and Health Transitions Research Unit, University of the Witwatersrand School of Public Health, Johannesburg, South Africa

[4]AGE Research Group, NIHR Newcastle Biomedical Research Centre, Newcastle University and Newcastle upon Tyne Hospitals Trust, Newcastle, UK

[5]Department of Population Health, London School of Hygiene and Tropical Medicine, London, UK

[6]Malawi Epidemiology and Intervention Research Unit, Malawi Epidemiology and Intervention Research Unit, Lilongwe, Malawi

[7]Division of Tropical and Humanitarian Medicine, University of Geneva and Geneva University Hospitals, Geneva, Switzerland

[8]Life for a Child Program, Diabetes NSW, Glebe, New South Wales, Australia

[9]Centre for Global Health, King's College London, London, UK

[10]Institute for Applied Health Research, Birmingham University, Birmingham, UK

Twitter Justine I Davies @drjackoids

REFERENCES

1. Maletkovic J, Drexler A. Diabetic ketoacidosis and hyperglycemic hyperosmolar state. *Endocrinol Metab Clin North Am* 2013;42:677–95.
2. Umpierrez G, Korytkowski M. Diabetic emergencies—ketoacidosis, hyperglycaemic hyperosmolar state and hypoglycaemia. *Nat Rev Endocrinol* 2016;12:222–32.
3. Pasquel FJ, Umpierrez GE. Hyperosmolar hyperglycemic state: a historic review of the clinical presentation, diagnosis, and treatment. *Diabetes Care* 2014;37:3124–31.
4. Gregg EW, Li Y, Wang J, et al. Changes in diabetes-related complications in the United States, 1990–2010. *N Engl J Med Overseas Ed* 2014;370:1514–23.
5. Anthanont P, Khawcharoenporn T, Tharavanij T. Incidences and outcomes of hyperglycemic crises: a 5-year study in a tertiary care center in Thailand. *J Med Assoc Thai* 2012;95:995–1002.
6. Chung ST, Perue GG, Johnson A, et al. Predictors of hyperglycaemic crises and their associated mortality in Jamaica. *Diabetes Res Clin Pract* 2006;73:184–90.
7. Desse TA, Eshetie TC, Gudina EK. Predictors and treatment outcome of hyperglycemic emergencies at Jimma University Specialized Hospital, southwest Ethiopia. *BMC Res Notes* 2015;8:553.
8. Ekpebegh CO, Longo-Mbenza B, Akinrinmade A, et al. Hyperglycaemic crisis in the Eastern Cape province of South Africa: high mortality and association of hyperosmolar ketoacidosis with a new diagnosis of diabetes. *S Afr Med J* 2010;100:822–6.
9. Colin Mathers D D, Inoue M, Rao C, et al. Counting the dead and what they died from: an assessment of the global status of cause of death data. *Bulletin of the World Health Organization* 2005;83:6.
10. Byass P. Who needs cause-of-death data? *PLoS Med* 2007;4:e333.
11. Byass P, Calvert C, Miiro-Nakiyingi J, et al. InterVA-4 as a public health tool for measuring HIV/AIDS mortality: a validation study from five African countries. *Glob Health Action* 2013;6:22448.
12. Byass P. Usefulness of the Population Health Metrics Research Consortium gold standard verbal autopsy data for general verbal autopsy methods. *BMC Med* 2014;12.
13. Thomas L-M, D'Ambruoso L, Balabanova D. Verbal autopsy in health policy and systems: a literature review. *BMJ Global Health* 2018;3.
14. Leitao J, Chandramohan D, Byass P, et al. Revising the WHO verbal autopsy instrument to facilitate routine cause-of-death monitoring. *Glob Health Action* 2013;6:21518.
15. Nichols EK, Byass P, Chandramohan D, et al. The WHO 2016 verbal autopsy instrument: an international standard suitable for automated analysis by InterVA, InSilicoVA, and Tariff 2.0. *PLoS Med* 2018;15:e1002486.
16. Soleman N, Chandramohan D, Shibuya K. Verbal autopsy: current practices and challenges. *Bull World Health Organ* 2006;84:239–45.
17. Herrera S, Enuameh Y, Adjei G, et al. A systematic review and synthesis of the strengths and limitations of measuring malaria mortality through verbal autopsy. *Malar J* 2017;16.
18. Karat AS, Maraba N, Tlali M, et al. Performance of verbal autopsy methods in estimating HIV-associated mortality among adults in South Africa. *BMJ Global Health* 2018;3.
19. Kahn K, Collinson MA, Gomez-Olive FX, et al. Profile: Agincourt Health and Socio-Demographic Surveillance System. *Int J Epidemiol* 2012;41:988–1001.
20. Crampin AC, Dube A, Mboma S, et al. Profile: The Karonga Health and Demographic Surveillance System. *Int J Epidemiol* 2012;41:676–85.
21. 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.
22. Atkinson MA, Eisenbarth GS, Michels AW. Type 1 diabetes. *The Lancet* 2014;383:69–82.
23. Thomas NJ, Jones SE, Weedon MN, et al. Frequency and phenotype of type 1 diabetes in the first six decades of life: a cross-sectional, genetically stratified survival analysis from UK Biobank. *The Lancet Diabetes & Endocrinology* 2018;6:122–9.
24. Alemu S, Dessie A, Seid E, et al. Insulin-requiring diabetes in rural Ethiopia: should we reopen the case for malnutrition-related diabetes? *Diabetologia* 2009;52:1842–5.
25. Marshall SL, Edidin D, Arena VC, et al. Prevalence and incidence of clinically recognized cases of type 1 diabetes in children and adolescents in Rwanda. *Diabet. Med.* 2015;32:1186–92.
26. Davies J, Abimiku Alash'le, Alobo M, et al. Sustainable clinical laboratory capacity for health in Africa. *The Lancet Global Health* 2017;5:e248–9.
27. Atun R, Davies JI, Gale EAM, et al. Diabetes in sub-Saharan Africa: from clinical care to health policy. *Lancet Diabetes Endocrinol* 2017;5:622–67.
28. Beran D, Yudkin JS, de Courten M. Access to care for patients with insulin-requiring diabetes in developing countries: case studies of Mozambique and Zambia. *Diabetes Care* 2005;28:2136–40.
29. Beran D, Ewen M, Laing R. Constraints and challenges in access to insulin: a global perspective. *Lancet Diabetes Endocrinol* 2016;4:275–85.
30. Ogle GD, Middlehurst AC, Silink M. The IDF Life for a Child Program Index of diabetes care for children and youth. *Pediatr Diabetes* 2016;17:374–84.
31. Gill CJ, Sabin L, Schmid CH. Why clinicians are natural bayesians. *BMJ* 2005;330:1080–3.
32. Sirota M, Kostopoulou O, Round T, et al. Prevalence and alternative explanations influence cancer diagnosis: an experimental study with physicians. *Health Psychol* 2017;36:477–85.
33. Hertwig R, Barron G, Weber EU, et al. Decisions from experience and the effect of rare events in risky choice. *Psychol Sci* 2004;15:534–9.
34. Hamm RM. Physicians neglect base rates, and it matters. *Behavioral and Brain Sciences* 1996;19:25–6.