Reviewer’s report

Title: Improving performance of the Tariff Method for assigning causes of death to verbal autopsies.

Version: 0 Date: 28 Aug 2015

Reviewer: Basia Zaba

Reviewer’s report:

Verbal Autopsy (VA) is a topic of increasing importance as more efforts are made to introduce VA into death registration systems in developing countries. It is important to thoroughly document the workings of proposed automated VA methods in order for their advantages and limitations to be properly understood by users, and to make it possible for software developers to build improved systems based on algorithms developed so far. A strength of this paper is the use of metrics to assess the performance of the automated methods after allowing for chance concordance in allocation of Cause of Death (CoD) and the impact of this on estimated the population Cause Specific Mortality Fractions (CSMF). The authors make reasonably plausible claims for the apparent superiority of the Tariff 2 system over Tariff 1. However, this paper needs further work to meet standards of clarity and accessibility required for VA users and system developers alike.

Major issues

Page 9 lines 26-30: clarify what is meant by Health Care Experience (HCE) data

"diagnoses of chronic illness obtained from health service providers" presumably means reported by respondent as communicated to them by the health service provider (HSP), not obtained directly from HSP through record linkage. Are both items (2) "diagnoses ..." and (4) "details of any interactions ..." considered as part of the HCE data?

Page 10 equation displayed between lines 4-10: explain symbols

The indices k and r have not been defined, or the quantity x. What is the significance of 40 in the summation (this has been explained in the second paper).

Page 10 lines 13-15: add explanation

Explain why some causes have inherently high tariffs (this was done quite well in the second paper).

Page 10 lines 15-20: fundamental clarification required
This is very difficult to follow. "The cause with the highest rank is assigned to the death with the unknown cause" - highest ranked tariff score amongst all the causes in the training data set? This does not make sense to me. Are you trying to say that the unknown cause tariff ranks are compared with the training set cause scores till a matching pattern (of high and low scores) is found? How is the best match chosen? Perhaps you need to show a worked example or a diagram to illustrate the process. Without explaining this properly (and the equation above) the main analysis is impenetrable, and could not be replicated in similar data sets by other VA analysts. When this part of the methods section has been amended the paper should be subject to statistical re-review.

Page 11, lines 57-60 and additional file 3: explain weighting properly

Explain how the Tariff weights and the GBD weights are combined to produce the overall weight.

Page 13 lines 9-20: clarify use of non-gold standard data

How is it possible to identify individual true positives and false negatives, or true Cause Specific Mortality Fraction without using the gold standard hospital data? What does "this analysis" refer to? What does "not linked" mean? The results immediately following, shown in tables 3 and 4, clearly relate to the Gold Standard data set as they refer to the 500 splits of these data (as explained in flow chart figure in additional file 4). Presumably this sentence refers back to page 8 lines 22-31 where some processes are described that use tariff 1 to analyse community-based data without gold-standard causes, but the relevance of the analyses of the non gold-standard data to the "validation" of Tariff 2 are not explained properly.

page 17 lines 56-60: Open Data Kit

If open source code for SmartVA is publically available please give the url for the file in a public software archive. Also state exactly what the open source code actually does: does it calculate individual Cause of Death (or range of causes with probability bounds)? Does it calculate the Cause Specific Mortality Fraction (with confidence limits) for a large set of individual VA records? Or is it simply the code for steering the logic of the VA questionnaire flow and coding the responses?

Discussion: limitations of hospital based gold standard

The analysis is based on a gold standard data set of medically certified deaths, all records should therefore have a non-trivial HCE component - i.e. not all HCE information will be "not stated" or "not applicable". Dropping existing HCE information from the analysis is not the same as including records that have no HCE component to begin with because no health service contact has ever taken place - respondents talking about deceased who had HCE will be more comfortable with modern medical concepts of disease, more knowledgable about the possible CoD and the symptom distributions that they report may be different in range and severity. This is alluded to on page 7 lines 55-57, but it should also be brought up in the discussion as a
limitation of this whole approach and the consequences for the "accuracy" of the system when applied to deaths with no HCE.

Minor problems

Page 7 lines 36-37: ambiguity

The sentence implies that the Bayesian methods (mentioned earlier) cannot scrutinise the relationship between item response and CoD - this is untrue

Page 9 line 1: add detail

How many causes in the original WHO, how many in the reduced list - perhaps add short table to summarise for adults, children and neonates

Page 9 lines 55-59: display problem

the bottom of the equation is truncated so that indices for interquartile range term are not visible. Presumably this is a problem arising from pdf conversion

Page 10 line 1: clarify

"from an unknown cause" is a bit confusing, I think you are not implying that this is done for a subset of deaths with unknown cause, but that you ignore knowledge of the cause based on the gold standard medical certification and use the Tariff algorithm to calculate the score based on its own internal logic.

Page 10 lines 51-60: clarify

Was all the text translated into English before standardized text mining?

Page 13 lines 1-2: clarify

The chance corrected CSMF are not shown in this paper, so leave out this sentence referring to personal communication.

page 14 line 27: jargon

Does "pleomorphic" mean "highly variable"?

page 17 line 1: obscure argument

The VA respondent may not be aware that the deceased was pregnant

Page 17 lines 38-44: further explanation
Unless you explain what Kappa is and how it has been used in ways that you disapprove of this paragraph sounds a bit polemical

Table 2: add more information

Include a panel showing the numbers with gold standard used to train and test. Are the gold standard numbers the same for Tariff 1 and Tariff 2?

Table 5, panel 1: missing line?

Panel 1 is supposed to refer to cause group A including nutritional disorders, but there is no line corresponding to nutritional disorders. Why has this table only been produced for adult deaths?

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

No

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

Yes

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

Unable to assess

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I recommend additional statistical review

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?
2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal