Reviewer's report

Title: Using verbal autopsy to track epidemic dynamics: the case of HIV-related mortality in South Africa

Version: 1 Date: 15 April 2011

Reviewer: varun malhotra

Reviewer's report:

Review summary

This paper started with three objectives: investigating disease dynamics using Physician coded Verbal autopsy (PCVA) and InterVA methods, 2. looking at approaches for handling common co-morbidities with HIV such as diarrhea, malnutrition, TB; and 3. to compare single coded VAs between two physicians coding for HIV.

Objective 1 is interesting because mostly VA has been used for point estimates of disease prevalence. Therefore, studying disease dynamics over ten years makes for an interesting read. Objective 2 is not particularly original as TB for instance has been repeatedly considered as a co-morbidity/opportunistic infection with HIV/AIDS. Although malnutrition and diarrhea are introduced, much has been written on them, especially diarrhea. Objective 3 is interesting because not much work until recently has been done on factors contributing to physician agreement in VA.

The paper, however, has certain areas where it can be especially improved. These areas are mentioned below and labeled according to Major compulsory revisions (C), Minor essential revisions (M), Discretionary revision (D).

1. Is the data sound and well controlled?

No comments because do not have access to many tables or appendices, etc. detailing methods. Some issues are:
- 373 deaths is 6.4% of 5828 but no where in the paper 5828 deaths are mentioned. Instead only 5794 of the total 6153 deaths in Agincourt are VA coded (C).

Note that I could not verify HIV related mortality rates as did not have access to Aigncourt population denominators.

2. Is the interpretation (discussion and conclusion) well balanced and supported by data?

Some comments are:
- Agreement between physicians (single-coded) is low. (M)
b. Why the difference in proportion of VA deaths attributable to B20 ICD code are different for InterVA and Physician coded VA? (C)

c. “The proportional differences in rates between the various approaches were greatest during the first three years.” – This has not been shown clearly. (C)

d. Do not understand and therefore cannot evaluate this line in discussion section of paper – “…Initial work on the InterVA model suggested that only causes likely to vary by an order of magnitude in terms of overall proportion needed to have an adjustment [23], with the crossover between “low” and “high” being at around 1% of total mortality. The “high” setting was therefore appropriate overall here…” (C)

3. Are methods appropriate and well described, and are sufficient details provided to allow others to evaluate and/or replicate the work?

a. Case definition of HIV – Authors mention use of B20-B24 ICD codes but these codes provide categories into which HIV is classified. There should be an explicit mention of disease guidelines/case definition that is used. (C)

4. What are strengths and weaknesses of methods?

No comments. Overall, methods are simple and believe sufficient to meet the objective of the paper. As far as statistical analysis is concerned, they do not need to be assessed specifically by an additional reviewer with statistical expertise.

5. Can writing, organization, tables and figures be improved?

- Authors mention that HIV/AIDS age-wise distribution is bimodal. Here, the age range of 20-64 is too wide. A narrower division of 5 year intervals can provide specific information as to where between 20 and 64 years of age, there lies high HIV mortality rate. (M)

- Context for change in disease trends over time – It is important to characterize the diseases prevalent in Agincourt area. If like most South Africa, the transition is from communicable to non-communicable diseases, then over time the burden of communicable diseases would decrease and that of non-communicable diseases will increase. The exception to this could be rising trend of HIV which can be associated with communicable diseases (i.e. infections, TB, etc.) and also non-communicable diseases (e.g. neoplasms, etc.). It would be worth mentioning how the disease landscape has evolved over time. I believe one of the authors – KK – has written extensively on this. Perhaps, referring to her papers would be helpful. (M)

- Table 1 - Mean age groups between InterVA and physician coding for 1995-97 (30.2 in physician vs. 25.7 in InterVA) and 2002-2005 (31.0 in physician vs 26.4 in InterVA) difference is more than or equal to 5 years. This difference is substantial. (C)

- The 2nd part of Results 1st sub-section – The evolving epidemics of HIV related mortality – seems misplaced and do not understand what point is being made
# Authors can do sensitivity analysis where different levels of strictness criteria for HIV case definition can be used. (D)

# Authors do mention that they looked for records explicitly mentioning HIV positivity. But overall, this aspect can be worded better. (C)

# Malnutrition is a risk factor and is not well placed in this section. It would instead be useful to look at 2-3 risk factors separately in another paragraph. Risk factors and/or other causes of death (where multiple causes of death are allowed with HIV being one of the underlying causes) can change over time, and this element would add to the epidemic transition element of the study. (C)

- Overall, the Results section needs work. It needs structure. For example, talk about:

# Physician coding vs. InterVA (C)
• Comparing trends over time
• Compare trends age-wise
• Then compare overall physician coding to InterVA

# Talk about difference criteria used for HIV – so different case definitions (C)

# Risk factors (C)

# B20-B24 – split the ICD codes and talk about what percentage HIV attributable each specific ICD code (C)

- Figure 2 – I do not understand the grey lines. They represent physician coding but it would be easier to read visually if two separate lines (physician 1 and physician 2) were shown and labeled. Also, note that the paper mentioned that one physician could be enough but this figure tells us that relying on one physician can produce huge variation. For instance, see the grey lines from 1997/98 till 2001 - they don’t even follow the general trend and show that HIV rates are decreasing. Also, variation peaks seen in 2003 and 2004 which as mentioned in discussion section are due to physician’s perception that HIV is widely prevalent in the area. Overall, Figure 2 needs work as is confusing at present. (C)

- Table 2 – the table presents a very general view and does not seem to tell much. What could improve this analysis is to see physician agreement by age, gender and ICD codes (B20-B24). Such analysis will provide more insight into which areas are where physicians are agreeing more or less. Perhaps they agree less for those above 64 years than for those less than 1 year. For instance, HIV seems more prevalent in above 65 as time progresses. Would be worth looking at physician agreement particularly for later time periods. (C)

6. Miscellaneous observations
- What do you mean by “aggregated findings” (last line) in the results section of Abstract?

- Abstract – Last line of conclusion – “The ability of the InterVA…are not necessary.” The study compares the two methods of InterVA and physician VA. It
does not necessarily show that one is better than the other. For this reason, InterVA alone cannot be regarded as the method to analyze VA data. 
- Interpretation in support of InterVA is overly supportive

Level of interest: An article of limited interest

Quality of written English: Needs some language corrections before being published

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

I declare I have no competing interests.