Using vital registration to quantify HIV mortality in South Africa between 1997 and 2009

Msemburi W[1,3], Dorrington R.E[2], Bradshaw D[1] and the SA NBD team[1]
[1]Burden of Disease, Medical Research Council, Cape Town, SOUTH AFRICA.
[2]Centre for Actuarial Research, University of Cape Town, SOUTH AFRICA.
[3]Corresponding author: William Msemburi, email: william.msemburi@mrc.ac.za

Abstract

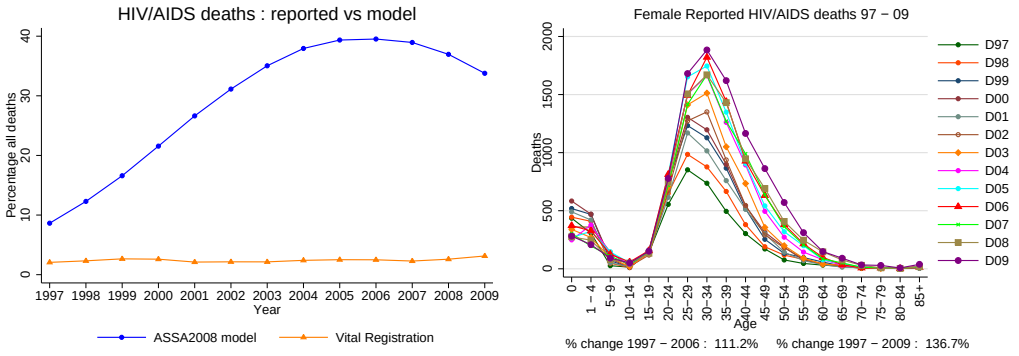
In South Africa the numbers of HIV/AIDS deaths have been under reported due to a greater proportion of AIDS deaths being mis-attributed to other causes. There has been a tendency to record indicator conditions without mentioning HIV or else use a pseudonym such as ``retroviral disease". This research outlines a statistical method used to estimate the numbers of mis-allocated AIDS deaths for South Africa between 1997 and 2009. Cause specific death rates were calculated from the official vital registration data obtained from Statistics SA for the period 1997 to 2009 (corrected for incompleteness of registration) and population estimates from the ASSA2008 population projection model. Causes of death with misclassified AIDS deaths were identified based on a distinct HIV/AIDS age pattern and extreme increase in death-rates akin to the increase in HIV/AIDS death rates. An exponential regression model relating the aggregate mortality rate from identified conditions on lagged antenatal clinic HIV prevalence for the period 1997-2003 provided estimates of the mortality level in the absence of AIDS and the increase in mortality related to prevalence. The background trend in the indicator causes was assessed in the 75-84 and using a stringent significance level ($p \leq 0.01$) was applied to the mortality level in the absence of HIV to estimate the excess AIDS deaths in each cause. The resultant aggregate AIDS proportions were compared to national estimates from various HIV models. Out of 214 causes of death, 19 met the criteria of potentially containing mis-classified AIDS deaths. AIDS deaths were estimated for 18 of these. Overall, the derived estimates of deaths due to AIDS were consistent with existing AIDS models, with results suggesting that approximately 92% of AIDS deaths in South Africa were mis-attributed to other causes for the period 1997 to 2009. Results also suggest peaking of AIDS mortality in 2005 for females and 2006 for males. From the results, we conclude that applying statistical techniques on vital registration data allows one to correct it for some of the systematic and random errors that can exist, in particular misclassification of AIDS deaths. Subtracting the mis-attributed AIDS deaths from each cause provides a more realistic profile of the underlying causes of death and the total number of deaths due to AIDS.

Key words: Cause of death, vital registration, systematic and random errors, HIV/AIDS mortality.

1 Introduction

In South Africa the numbers of HIV/AIDS deaths have been under reported due to a greater proportion of AIDS deaths being mis-attributed to other causes. There has been a tendency to record indicator conditions without mentioning HIV or else use a pseudonym such as ``retroviral disease''. Evaluation of the cause of death data indicates that it cannot be used at face value. When one compares the percentage of all deaths that is due to AIDS according to vital registration data versus various models e.g ASSA2008 (Actuarial Society of South Africa (2012)) shown in Figure 1, it is clear that there has been a systematic under-reporting of AIDS deaths. This research outlines a statistical method used to estimate the numbers of mis-allocated AIDS deaths for South Africa between 1997 and 2009.

Figure 1: Reported vs modeled AIDS and HIV/AIDS deaths age-pattern



2 Methods

The first step taken towards estimating the numbers of mis-allocated HIV/AIDS deaths was to identify the causes to which AIDS deaths are being mis-attributed. This was a 3 step process

1. Identify causes displaying characteristic HIV/AIDS age pattern. See Figure 1.
2. Identify causes with significant percentage increases over the period
3. Compare list of identified causes with causes identified by other researchers (Birnbaum et al. (2011); Groenewald et al. (2005))

Following the identification of causes a statistical model is applied to the data to estimate the true deaths for each cause. The model for cause-age-time patterns of AIDS mortality as a function of prevalence[1] is formulated as follows. Observed numbers of death attributed to indicator cause h aged x at time t are corrected for completeness of reporting and aggregated across indicator causes to give $D(x,t) = \sum_h D(h,x,t)$. Next, for ages younger than 74, key parameters are estimated by fitting the following regression model to data for the period 1997 to 2003

$$\frac{D(x,t)}{N(x,t)} = \alpha(x)e^{\beta(x)p(t)} \tag{1}$$

[1]Based on Dorrington's UCT Actuarial and Demography seminar, 2/9/2010: ``Ins and Outs: Estimation of the HIV incidence on all women from the change in antenatal prevalence and the number of AIDS deaths from the CoD vital registration data''.

where $p(t)$ is lagged antenatal prevalence depending on age-group, t is measured in years and is equal to 1 when $p(t)$ is first greater than zero, $N(x,t)$ is mid-year population, $\beta(x)$ is an exponential scaling factor relating death rate to prevalence at time t and $\alpha(x)$ is the aggregate non-AIDS mortality rate at time $t = 0$.

In order to determine a trend in non-AIDS mortality the following regression model is fitted to deaths in the age group 75-84 on the assumption that the proportion AIDS deaths are negligible in this age range. γ_h is set to zero if not statistically significant i.e. $p > 0.01$.

$$\frac{D(h,x,t)}{N(x,t)} = \mu_h + \gamma_h \ln(1+t) \tag{2}$$

The cause h non-AIDS rate is estimated by applying the cause proportions for 1997 to the aggregate non-AIDS rate

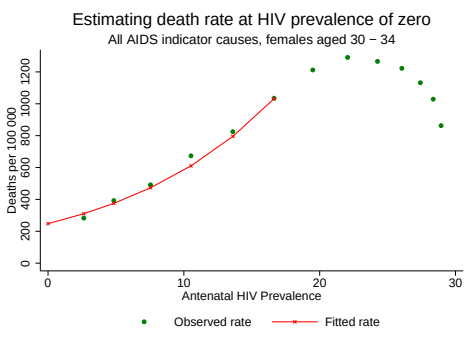
$$\alpha(x,h) = \alpha(x) \times \frac{D(h,x,1997)}{D(x,1997)} \tag{3}$$

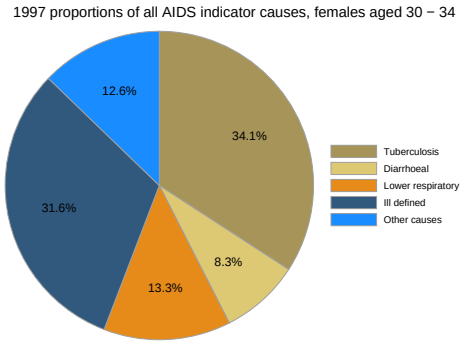
Finally AIDS deaths misclassified as cause h aged x at time t are estimated as follows:

$$d_h(a,x,t) = D(h,x,t) - \alpha(x,h)(1 + \frac{\gamma_h}{\mu_h}\ln(1+t))N(x,t) \tag{4}$$

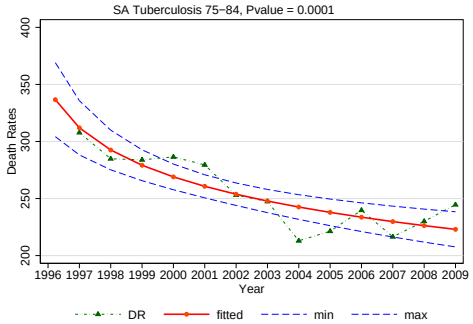
2.1 Example of application of model

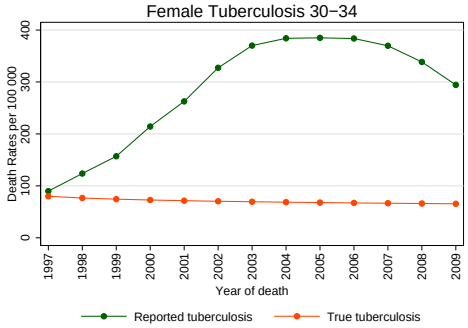
1. Estimate aggregate mortality at HIV prevalence of zero. Estimate cause specific proportion of aggregate for zero HIV rate based on 1997





2. Estimate cause specific death rate for each year in the absence of HIV. Calculate predicted cause specific death rate for each year and age in the absence of HIV





3 Results

There were 19 causes identified as potentially containing misclassified AIDS deaths. These causes are given in Table 1.

Table 1: Causes identified as containing mis-classified AIDS deaths

1. Tuberculosis	11. Protein energy malnutrition
2. Sexually transmitted excl HIV	12. Iron deficiency anemia
3. Diarrhoeal diseases	13. Endocrine, nutritional, blood and immune
4. Selected vaccine preventable	14. Other interstitial lung
5. Meningitis and encephalitis	15. Other respiratory diseases
6. Septicaemia	16. Other digestive diseases
7. Other infectious diseases	17. Other neurological conditions
8. Lower respiratory infections	18. Adult respiratory distress
9. Maternal conditions	19. Ill defined
10. Other nutritional deficiencies	

Tuberculosis, lower respiratory infections (pnuemonia), diarrhoeal diseases and ill-defined conditions as well as HIV psuedonyms accounted for most of the misclassified deaths.

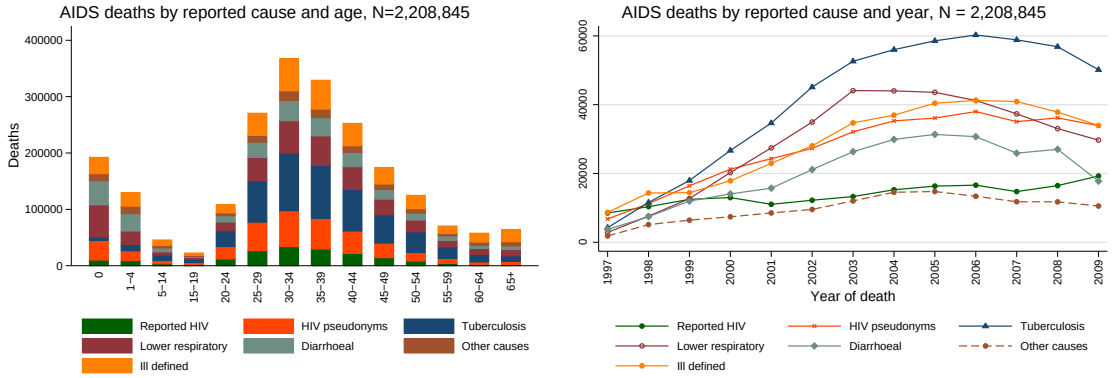


Table 2 gives the estimated AIDS deaths and percentage of all deaths attributable to AIDS for males, females and all persons. The minimum number of AIDS deaths is estimated for 1997 (36,695 deaths), rising to a maximum of 241,373 deaths in 2006 and falling to 195,278 deaths in 2009. Approximately 92% of the estimated 2,208,845 AIDS deaths were mis-attributed to other causes.

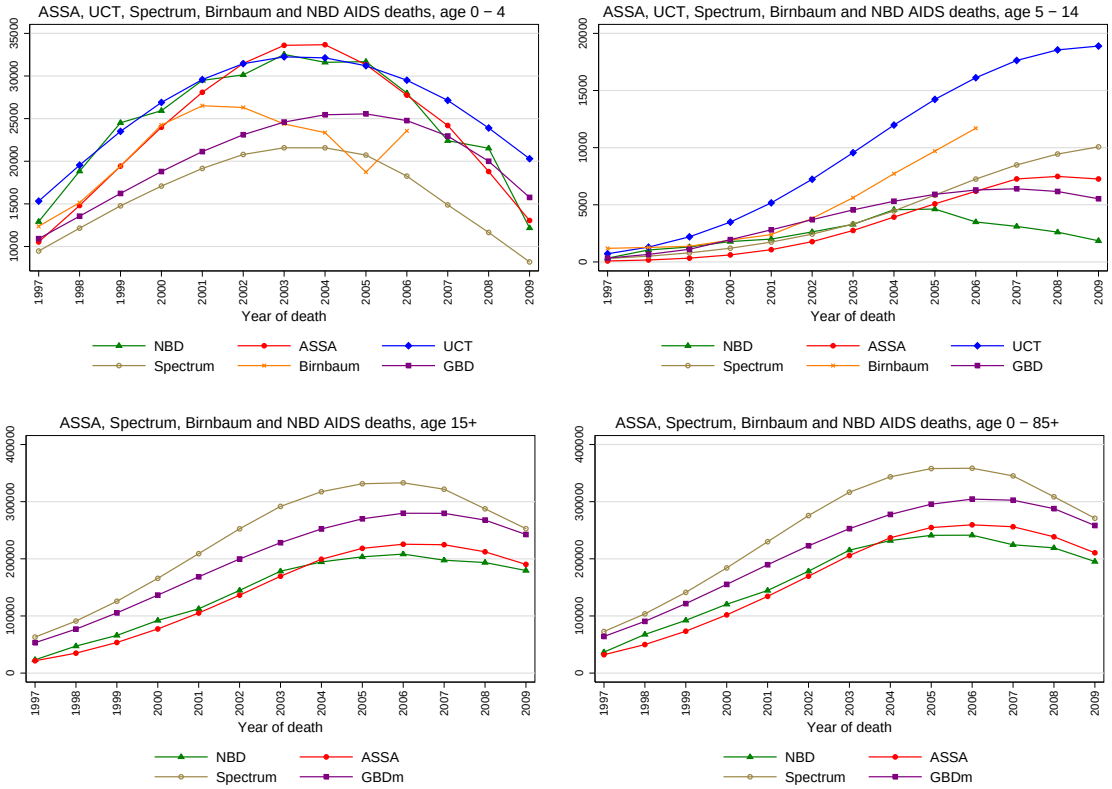
Table 2: AIDS deaths and % of all deaths, 1997 to 2009

Year	Male deaths	%	Female deaths	%	Person deaths	%
1997	18,329	7.6	18,366	9.5	36,695	8.4
1998	33,377	12.9	34,269	16.0	67,646	14.3
1999	44,663	16.7	47,700	20.9	92,363	18.6
2000	55,943	20.2	64,574	26.0	120,516	23.0
2001	66,749	23.4	77,879	29.7	144,628	26.4
2002	80,653	26.6	97,671	34.0	178,324	30.2
2003	98,461	29.9	116,719	37.2	215,180	33.5
2004	104,111	31.4	127,846	39.7	231,957	35.5
2005	108,690	32.3	132,432	40.0	241,122	36.1
2006	110,943	32.5	130,430	39.2	241,373	35.8
2007	104,262	31.2	120,341	37.4	224,603	34.2
2008	103,140	31.2	116,021	36.7	219,161	33.9
2009	93,662	29.2	101,616	33.5	195,278	31.3
All	1,022,982	25.8	1,185,862	32.3	2,208,845	28.9

4 Comparisons with models

Figure 2 shows the estimated deaths that are attributable to HIV, comparing the estimates derived using this regression method with those derived using the UCT paediatric HIV model by Johnson et al. (2012), the estimates from Birnbaum et al. (2011), the Spectrum model estimates by UNAIDS (2012), the GBD2010 results by the Institue for Health Metrics and Evaluation (2013) and the ASSA2008 Actuarial Society of South Africa (2012) models.

Figure 2: ASSA and NBD proportions HIV deaths by broad age group, 1997 to 2009



Although the estimates from the ASSA2008 model and NBD approach do not exhibit much variation for the 15+ age group, for the younger age groups there is a significant amount of between model variation as well as variation relative to the NBD approach. While one could propose that the number of AIDS deaths in the 5-14 age group is small and the model differences are thus insignificant, this proposition does not hold for the youngest age group (0 to 4) for which HIV/AIDS deaths is a leading cause of death. The differences between the NBD approach and the models suggests there were more AIDS related child deaths before ART than is being estimated by most models. Additionally, the differences suggest that for children the ART roll out has not had an effect of the magnitude reflected in the model estimates. The similarities in the death trends for the NBD and model AIDS estimates do confirm that ART rollout made a positive difference in reducing AIDS mortality in all age groups. The NBD as well as the models investigated generally suggest that AIDS mortality has been falling following its peaking around 2005/2006.

5 Conclusion

Applying statistical techniques on vital registration data allows one to correct it for some of the systematic and random errors that can exist, in particular misclassification of AIDS deaths. Subtracting the mis-attributed AIDS deaths from each cause provides a more realistic profile of the underlying causes of death and the total number of deaths due to AIDS.

References

Actuarial Society of South Africa (2012). ASSA2008 model, aids.actuarialsociety.org.za/ASSA2008-Model-3480.htm.

Birnbaum, J., Murray, C. and Lozano, R. (2011). Estimating misclassified HIV/AIDS deaths in South Africa, *Bulletin of the World Health Organization Type : Research* (BLT10.086280).

Groenewald, P., Nannan, N., Bourne, D., Laubscher, R. and Bradshaw, D. (2005). Identifying deaths from AIDS in South Africa, *AIDS* **19**(2): 193--201.

Institue for Health Metrics and Evaluation (2013). GBD Compare visualization tool, Accessed June 2013 at http://viz.healthmetricsandevaluation.org/gbd-compare/.

Johnson, L., Davies, M., Moultrie, H., Sherman, G., Bland, R., Rehle, T., Dorrington, R. and Newell, M. (2012). The effect of early initiation of antiretroviral treatment in infants on pediatric AIDS mortality in South Africa: A model-based analysis., *Pediatric Infectious Disease Journal.* **31**(5): 474--480.

UNAIDS (2012). Spectrum model, www.unaids.org/en/dataanalysis/tools/spectrumepp2011/.