Modeling Age-Specific Mortality for Countries with Generalized HIV Epidemics*

David J. Sharrow, Samuel J. Clark, Adrian E. Raftery
University of Washington

November 5, 2012

Abstract

Population projections that forecast the future size and age-composition of a country are crucial tools for appropriately planning the future allocation of societal resources. A projection model for countries with generalized HIV epidemics should take into account the future trajectory of the epidemic given the severe effect a generalized epidemic can have on the mortality conditions and composition of a population. We present a model of age-specific mortality as a function of life expectancy, HIV prevalence, and anti-retroviral therapy coverage for the 39 countries of the world experiencing a generalized HIV epidemic. We perform an in-sample validation where results show slight errors for several mortality indicators. Combined with the outputs of existing epidemiological and demographic models, this model makes it possible to estimate future mortality profiles for countries with generalized HIV epidemics.

*This work was supported by grants R01 HD054511, R01 HD070936, and K01 HD057246 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD).

Motivation

Population projections that forecast the future size and age-composition of a country or region are crucial tools for appropriately planning the future allocation of societal resources such as medical, educational, or government services. These kinds of projections for countries with generalized HIV epidemics ($> 1\%$ prevalence in the general population and no concentration of the epidemic in high-risk subgroups such as IV drug users, men who have sex with men, or sex workers) are especially useful given the dramatic effect a generalized epidemic can have on life expectancy and the age-composition of a population (Ngom and Clark, 2003; Blacker, 2004; Timaues and Jasseh, 2004; Porter and Żaba, 2004).

Mortality conditions play a key role in future population composition thus the mortality component of projections for countries with generalized epidemics should include the influence of the future trajectory of the epidemic. Using tools like the United Nations EPP model (Brown et al., 2010; Ghys and Garnett, 2010) one can estimate or project features of an HIV epidemic such as prevalence and coverage of anti-retroviral therapy (ART) among the infected population, which can then be mapped onto a set of age-specific mortality rates. In this paper, we develop a model of age-specific mortality as a function of life expectancy, HIV prevalence, and ART for the 39 countries of the world experiencing a generalized HIV epidemic.

Data

We calibrate this model with a set of 312 5-year sex-specific life tables from the World Population Prospects 2010 revision (United Nations, Department of Economic and Social Affairs, Population Division, 2011). This dataset contains mortality information from 39 countries with generalized HIV epidemics with eight 5-year life tables from 1970-2010 for each country. All life tables have uniform 5-year age intervals except for the youngest age groups up to an open interval of 100 (0, 1-4, 5-9, 10-14,...,100+). We model the logged ${}_nm_x$ column of these life tables. Finally, we obtain the mid-period (1973, 1978, 1983,...,2008) HIV prevalence as well as adult and child ART coverage for each of the 312 country-periods from UNAIDS estimates.

Model

Our objective is a parsimonious model that can represent the age-pattern of mortality rates for countries with generalized HIV epidemics as a function of life expectancy at birth, HIV prevalence, and ART coverage. The general form of the model represents a set of age-specific mortality rates as the weighted sum of three independent, age-varying components that represent the age-varying nature of the mortality schedule. A step-by-step discussion of

the modeling strategy is below.

We first derive the age-varying components from a Singular Value Decomposition (SVD) of the matrix of observed mortality rate schedules. The resulting 'left-singular vectors' are the independent components we need, and they have the convenient property of encoding the bulk of the variation among the observed mortality schedules in a small number of these vectors.[1]

1. SVD components of life tables L: $X_{i,x}$

We next regress (OLS) each of the 312 mortality rate schedules on the first few left-singular vectors from the SVD yielding a set of weights ($\omega_{l,i}$) for each country-period life table that relate the individual mortality rate schedule to the components.

2. Project logged mortality rates onto SVD components $X_{i,x}$:

$$\lambda_{\ell,x} = c_\ell + \sum_{i=1}^{3} \omega_{\ell,i} \cdot X_{i,x} \tag{1}$$

where

- $\lambda_{\ell,x}$ are logged age-specific mortality rates for life table ℓ: $\lambda_{\ell,x} = \log({}_n m_x)_\ell$
- X are age-specific SVD components
- ω are weights specific to SVD components
- c is a constant
- ℓ indexes life tables: $\ell \in \{1, \ldots, L\}$
- i indexes age-specific SVD components: $i \in \{1, 2, 3\}$
- x indexes age: $x \in \{0, 1, 5, 10, \ldots, 100\}$

Once we have obtained a set of ω's for each country-period life table, we model each ω_i as some combination of life expectancy at birth ($e0$), HIV prevalence, and ART coverage for both adults and children. We use Bayesian Model Averaging (Raftery, 1995) to adjudicate amongst the various combinations of these variables to find the combination that best predicts each ω_i.[2]

[1] We tested the model using a varying number of components and found little improvement in fit when using more than three components. The lack of improvement with inclusion of higher order components likely reflects the fact that, similar to a Principal Components Analysis, each successive component accounts for successively smaller proportions of the overall variance.

[2] This step is performed with the function `bicreg()` from the package `BMA` (Raftery et al., 2012) in the statistical analysis software `R`.

3. Model ω's :

$$\omega_{\ell,i} = \beta_{0,i} + \beta_{1,i} \cdot e_{0,\ell} + \beta_{2,i} \cdot P_\ell + \beta_{3,i} \cdot A_\ell + \epsilon_{\ell,i} \tag{2}$$

where

- P is HIV prevalence
- A is ART coverage
- e_0 is the expectation of life at birth

Equation 2 presents a model predicting a ω as a function of all predictor variables, but BMA actually selects more parsimonious models. We also fit separate models for each ω_i by sex as well as for African countries and non-African countries (Bahamas, Belize, Guyana, Haiti, Jamaica, Thailand). The sex-region-specific models for each ω_i are presented below. The same set of predictor variables is found through BMA for both regions and sexes.[3]

$$\omega_{1,s,r} = \beta_{0,1,s,r} + \beta_{1,1,s,r} \cdot e_{0,s,r} \tag{3}$$
$$\omega_{2,s,r} = \beta_{0,2,s,r} + \beta_{1,2,s,r} \cdot e_{0,s,r} + \beta_{2,2,s,r} \cdot P_r \tag{4}$$
$$\omega_{3,s,r} = \beta_{0,3,s,r} + \beta_{1,3,s,r} \cdot e_{0,s,r} + \beta_{2,3,s,r} \cdot P_r \tag{5}$$

where

- s indexes sex
- r indexes region

Using equations 3 through 5 we can predict $\hat{\omega}$'s, which when substituted in equation 1 produce a set of predicted mortality rates. The effective parameters in this model are the weights ($\hat{\omega}$) as the components are age-specific and fixed.

4. Use models from step 3 to predict $\hat{\omega}$'s both in and out of sample.

5. Use predicted $\hat{\omega}$'s to predict $\hat{\lambda}$'s:

$$\hat{\lambda}_{\ell,x} = c_m + \sum_{i=1}^{3} \hat{\omega}_{\ell,i} \cdot X_{i,x} \tag{6}$$

where

- c_m is the median of the distribution of c_ℓ values
- The 'hat' symbol over a variable, as in $\hat{\omega}$, indicates *predicted* value

[3]We make a further refinement to the modeling of the ω's by testing for any non-linear transformations that would improve fit using alternating conditional expectation. This method suggested a transformation to the outcome variable in equation 5 which did improve overall fit and is incorporated into our final model.

Model Validation

To assess the model's ability to replicate mortality rate schedules for populations with generalized HIV epidemics, we use the model outlined above to predict a set of age-specific mortality rates for each of the in-sample country-period life tables as a function of the life expectancy and mid-period prevalence for each country-period life table. We then compare the predicted mortality rates to those in the observed dataset. For comparison with existing models of all-age mortality, we fit the same set of tables using the WHO modified logit system (Murray et al., 2003), the UN model life tables for developing countries (United Nations, 1982), and the Coale-Demeny regional model life tables (Coale and Demeny, 1966; Coale and Guo, 1989). Table 1 presents the mean absolute error amongst all age groups and all country-periods along with the mean absolute error for life expectancy at birth (e_0), under-5 mortality ($_5q_0$), and adult mortality ($_{45}q_{15}$). Because the system described in this paper is essentially an HIV-calibrated model life table system we refer to it as "HIV MLT" in the table below. For all metrics, smaller numbers are better.

Table 1: Mean absolute error for three mortality indicators and amongst all-ages and life tables after fitting WPP life tables with various all-age mortality models[†]

	Female				Male			
Model	All-ages[‡]	e_0	$_5q_0$	$_{45}q_{15}$	All-ages[‡]	e_0	$_5q_0$	$_{45}q_{15}$
HIV MLT	0.002	0.81	0.012	0.020	0.003	0.86	0.014	0.020
WHO[§]	0.005	0.93	—	—	0.006	1.18	—	—
CD[§]	0.003	—	0.017	0.033	0.004	—	0.015	0.036
UN	0.004	1.43	0.027	0.037	0.006	1.18	0.012	0.038

[†] All errors in this table are calculated from life tables with ages 0, 1-4, 5-9, 10-14, ..., 75+.
[‡] 'All-ages' refers to the mean absolute error for the non-logged mortality rates across all age groups ($_nm_x$) and amongst all life tables ($1, ..., L = 312$).
[§] 'WHO' and 'CD' contain blank spaces as these quantities are inputs to these systems and thus have no error.

One can see from table 1 that our model is able to outperform these three existing model life table systems. This result is no surprise not because the method used to generate those models is flawed but rather because the data sets used to calibrate those models do not contain data from countries and periods where the HIV epidemic was generalized. We show smaller errors on nearly every metric for both sexes. The 'HIV MLT' model shows mean absolute errors for life expectancy of less than one year for both sexes along with slight errors in predicting the probability of childhood death of 0.014 and 0.012 for the male and female models respectively. The error in predicting adult death is slightly higher than for childhood death at about 0.02 for both sexes.

An advantage of this model and what tends to explain the overall better fit is it's ability to produce the accentuated adult mortality humps associated with a generalized HIV epidemic that other models often miss. Figure 1 plots the predicted ${}_nm_x$ schedule from the 'HIV MLT' model along with the WHO, Coale and Demeny, and UN fits for Swaziland females 2005-10. This figure makes clear that when HIV prevalence is high and the HIV hump is present, the 'HIV MLT' model is able to closely estimate age-specific mortality. Additional selected fits can be found in Appendix 1.

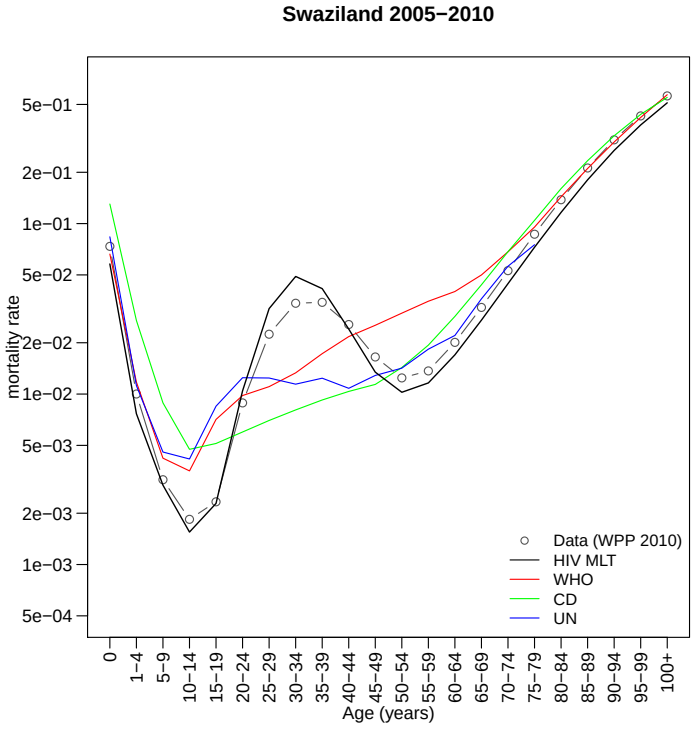


Figure 1: Fits of 'HIV MLT' model and other model life table systems for Swaziland females 2005-2010.

Discussion

We present a flexible, parsimonious model of age-specific mortality for countries with generalized HIV epidemics. First, a set of age-specific mortality rates is represented as the weighted combination of a set of age-varying components. On a second level, the weights are modeled as a function of life expectancy and a set of epidemic characteristics. This structure allows us to map life expectancy and prevalence onto a set of mortality rates reflecting the mortality impact of a generalized HIV epidemic.

We perform an in-sample validation method by predicting age-specific mortality for the 39 generalized HIV epidemic countries based on life expectancy and prevalence for each country-period. Results suggest a three-component model fits best with modest errors for several mortality indicators.

Combined with the outputs of existing epidemiological and demographic models, this model makes it possible to estimate future mortality profiles for countries with generalized HIV epidemics.

References

Blacker, John. 2004. "The impact of AIDS on adult mortality: evidence from national and regional statistics." *AIDS* 18:S19–S26.

Brown, T, L Bao, AE Raftery, JA Salomon, RF Baggaley, J Stover, and P Gerland. 2010. "Modeling HIV epidemics in the antiretroviral era: the UNAIDS Estimation and Projection package 2009." *Sexually transmitted infections* 86:3–10.

Coale, Ansley J. and Paul Demeny. 1966. *Regional Model Life Tables and Stable Populations*. Princeton University Press.

Coale, Ansley J. and G. Guo. 1989. "Revised regional model life tables at very low levels of mortality." *Population Index* 55:613–643.

Ghys, Peter D and Geoff P Garnett. 2010. "The 2009 HIV and AIDS estimates and projections: methods, tools and analyses." *Sexually transmitted infections.* 86:ii1.

Murray, C.J.L, B.D. Ferguson, A.D. Lopez, M. Guillot, J.A. Salomon, and O. Ahmad. 2003. "Modified Logit Life Table System: Principles, Empirical Validation, and Application." *Population Studies* 57:165–182.

Ngom, Pierre and Samual Clark. 2003. "Adult Mortality in the Era of HIV/AIDS: sub-Saharan Africa." Technical report, Population Division, United Nations, New York.

Porter, Kholoud and Basia Żaba. 2004. "The empirical evidence for the impact of HIV on adult mortality in the developing world: data from serological studies." *AIDS* 18:S9–S17.

Raftery, Adrian, Jennifer Hoeting, Chris Volinsky, Ian Painter, and Ka Yee Yeung. 2012. *BMA: Bayesian Model Averaging*. R package version 3.15.1.

Raftery, Adrian E. 1995. "Bayesian Model Selection in Social Research." *Sociological Methodology* 25:111–163.

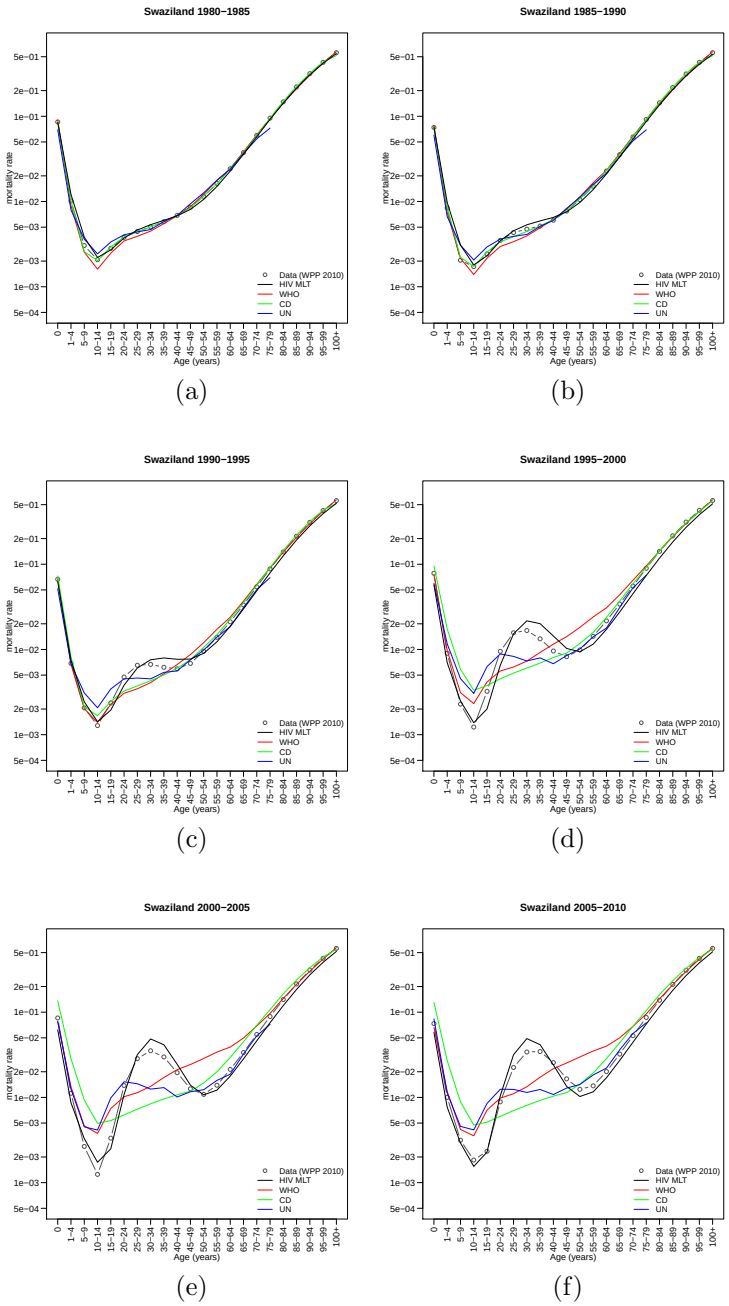
Timaues, Ian M. and Momodou Jasseh. 2004. "Adult Mortality in sub-Saharan Africa: Evidence from Demographic and Health Surveys." *Demography* 41:757–772.

United Nations. 1982. "Model Life Tables for Developing Countries." In *Population Studies*, 77. New York: United Nations.

United Nations, Department of Economic and Social Affairs, Population Division. 2011. *World Population Prospects: the 2010 Revision*, volume 3. New York: United Nations.

Appendix 1: Selected Fits

The following series of plots show example fits with this model. In this section we show a mix of regions, sexes, and prevalence levels.



(a) (b) (c) (d) (e) (f)

Figure 2: Swaziland, Females 1980-2010

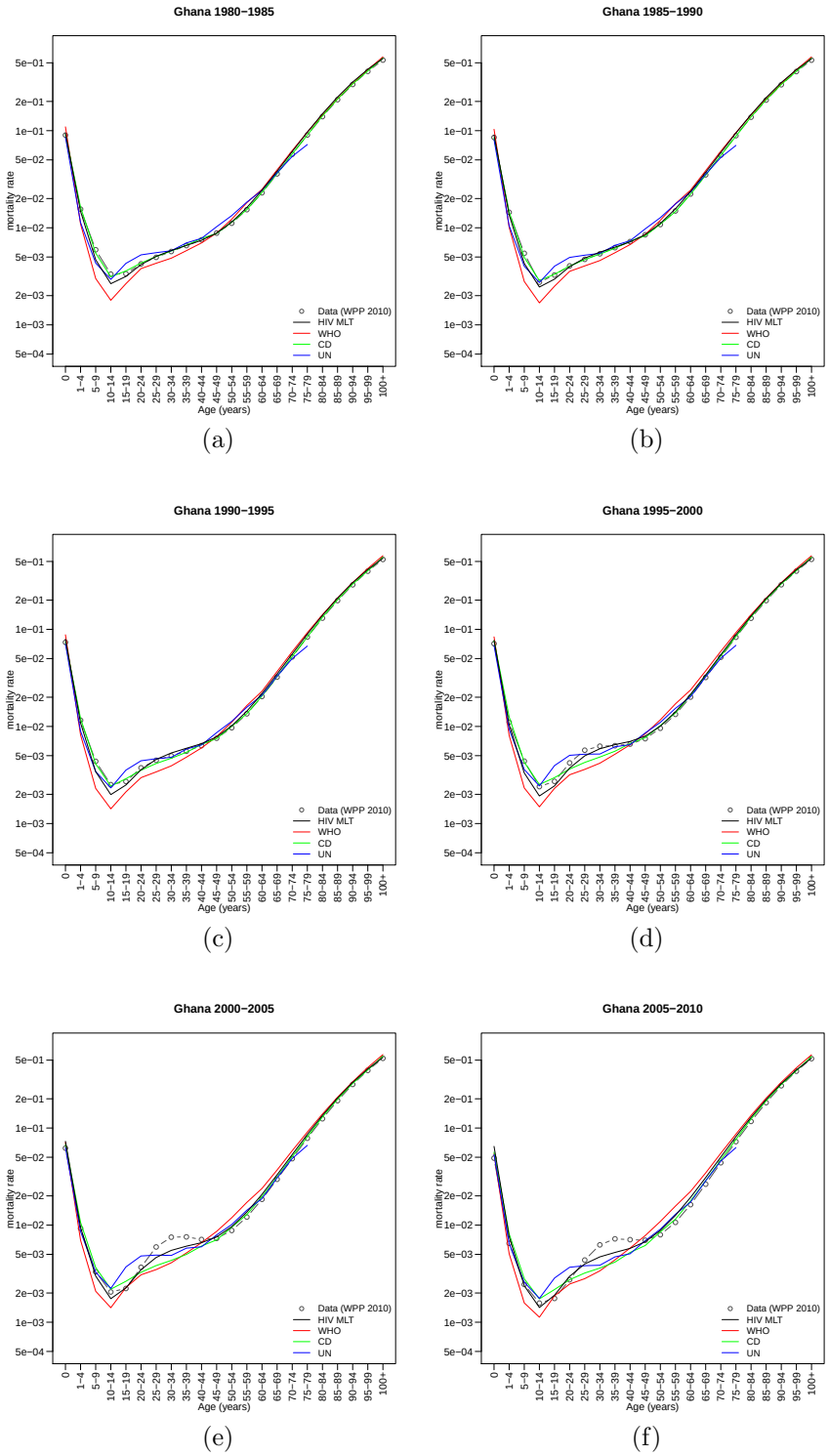


Figure 3: Ghana, Females 1980-2010

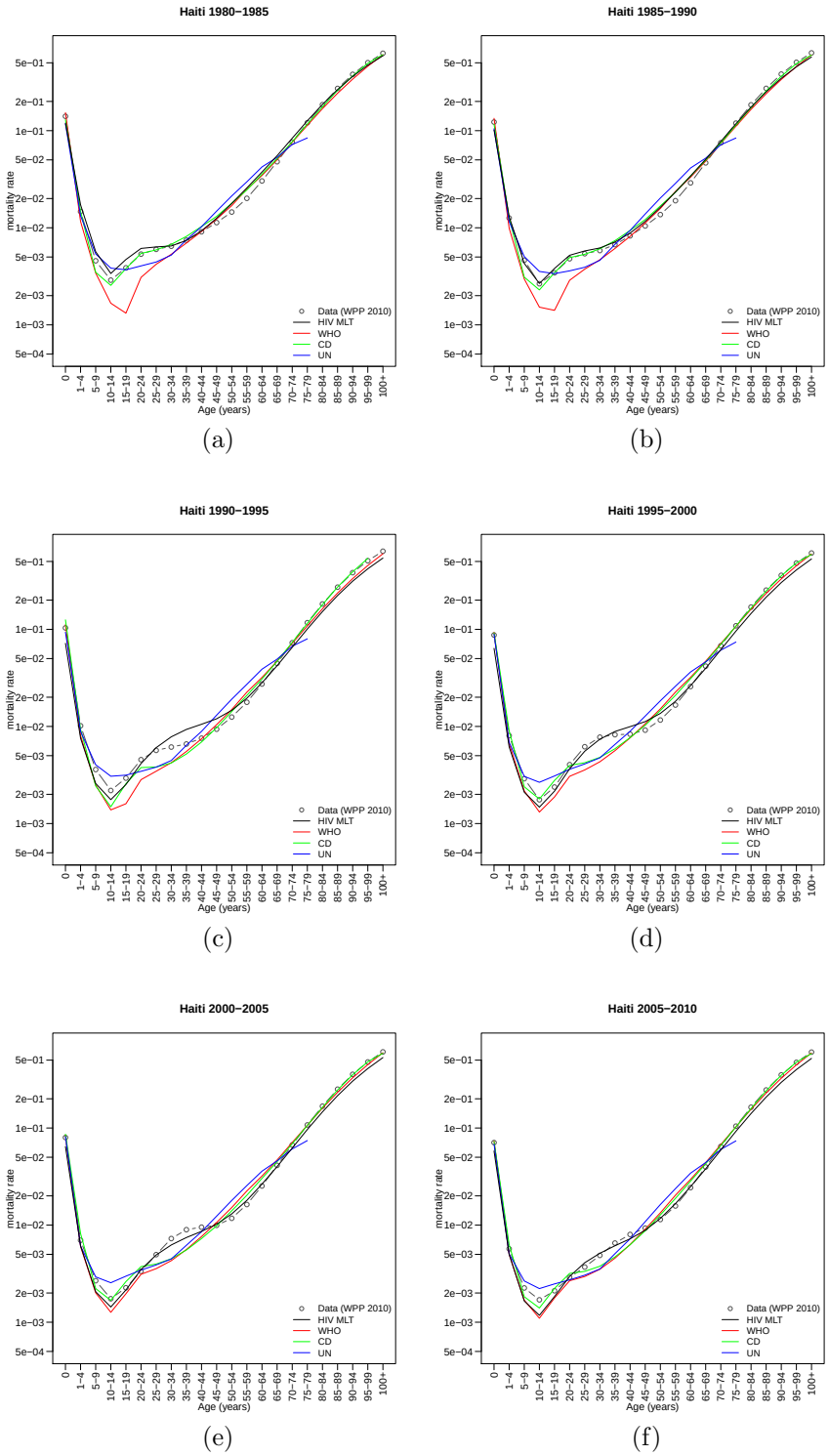


Figure 4: Haiti, Males 1980-2010

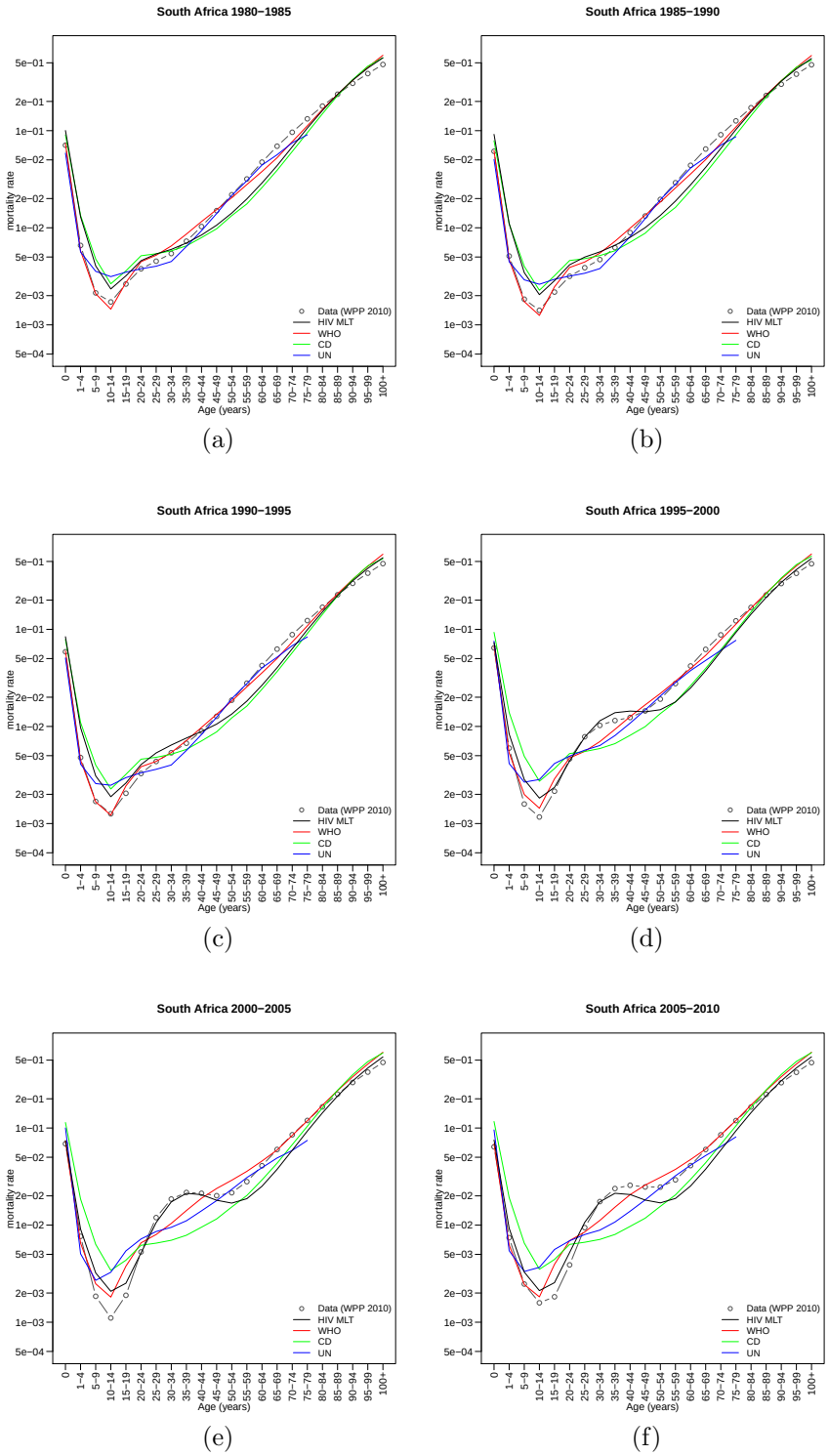


Figure 5: South Africa, Males 1980-2010