Published in final edited form as:

Popul Stud (Camb). 2018 March ; 72(1): 1–15. doi:10.1080/00324728.2017.1401654.

Probabilistic Population Projections for Countries with Generalized HIV/AIDS Epidemics

David J. Sharrow^{1,5}, Jessica Godwin², Yanjun He², Samuel J. Clark^{3,4}, and Adrian E. Raftery^{2,3}

¹Center for Statistics and the Social Sciences, University of Washington ²Department of Statistics, University of Washington ³Department of Sociology, University of Washington ⁴MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Sciences, University of the Witwatersrand

Abstract

The United Nations (UN) issued probabilistic population projections for all countries to 2100 in 2015 by simulating future levels of total fertility and life expectancy and combining the results using a standard cohort-component projection method. For the 40 countries with generalized HIV/AIDS epidemics, the mortality projections used the Spectrum/EPP model, a complex, multistate model designed for short-term projections of policy-relevant quantities for the epidemic. We propose a simpler approach that is more compatible with existing UN projection methodology for other countries. Changes in life expectancy are projected probabilistically using a simple time series regression and then converted to age- and sex-specific mortality rates using model life tables designed for countries with HIV/AIDS epidemics. These are then input to the cohort-component method, as for other countries. The method performed well in an out-of-sample cross-validation experiment. It gives similar projections to Spectrum/EPP in the short-run, while being simpler and avoiding multistate modeling.

Keywords

Bayesian hierarchical model; Cohort-component projection method; Estimation and Projection Package; Spectrum; UNAIDS; World Population Prospects

Introduction

Population projections are used by governments at all levels, international organizations, social and health researchers, and by the private sector for policy planning, monitoring development goals and as inputs to economic and environmental models. The United Nations (UN) issues population projections for all the countries of the world by age and sex to 2100, updated every two years, in the World Population Prospects (WPP). It is the only organization to do so, and its projections are the de facto standard at the global level (Lutz and Samir 2010).

⁵Corresponding author: dsharrow@uw.edu.

There has long been great interest in probabilistic population projections, to quantify uncertainty in projections and the risk of adverse demographic events. However, the dominant population projection methods are deterministic, and uncertainty has typically been conveyed by variant scenario projections (such as High and Low projections) using different assumptions about future rates. This approach has been criticized as lacking validity because it has no probabilistic basis, and leads to possible contradictions (Lee and Tuljapurkar 1994; National Research Council 2000).

In a major step forward, the UN issued official probabilistic population projections for all countries in July 2015, available at <http://esa.un.org/unpd/ppp>. These projections were developed using the methodology of Raftery et al. (2012). They take account of uncertainty about future levels of total fertility and life expectancy, using the Bayesian hierarchical models of Alkema et al. (2011) for fertility and Raftery et al. (2013) for life expectancy. Between-country correlation in fertility rates is accounted for by the method of Fosdick and Raftery (2014), and correlation between male and female life expectancy is modeled by the method of Raftery et al. (2014).

These statistical models allow a large number of trajectories of future fertility and mortality for all countries to be simulated from their predictive probability distributions. Each simulated trajectory of future life expectancy is converted to age-specific mortality rates using a modified Lee-Carter method (Li and Gerland 2011; Raftery et al. 2012; Li et al. 2013; Ševčíková et al. 2016a). Future age-specific fertility and mortality rates are then converted to future population by age and sex by the standard cohort-component method (Whelpton 1936; Preston et al. 2001), using the bayesPop R package (Ševčíková et al. 2016b). The result is a sample of possible future trajectories of world population by country, age and sex, to 2100. It can be viewed as a large number of possible versions of the 2100 revision of the UN's World Population Prospects (which won't exist for another 83 years).

The UN's probabilistic World Population Prospects currently treats countries with generalized HIV/AIDS epidemics (defined as having had HIV prevalence greater than 2% at any time since 1980, and hereafter referred to as the "AIDS countries") differently from others. These countries have very different mortality patterns compared to countries without generalized HIV epidemics, with a large mortality hump in middle adulthood due to AIDS, and so they cannot be modeled using, for example, standard model life tables.

The UN's current method for probabilistically projecting mortality for these countries starts with the deterministic projection from the 2015 revision of the World Population Prospects (United Nations, Department of Economic and Social Affairs, Population Division 2015b). This uses the Spectrum model developed for UNAIDS (Avenir Health 2016; Stanecki et al. 2012; Stover et al. 2012), a complex multistate model that divides adults into 15 compartments according to their HIV status, and models transitions among these 15 compartments. In a second stage, the Bayesian hierarchical model of Raftery et al. (2013) for life expectancy and the modified Lee-Carter method are applied to all countries, including the AIDS countries, yielding simulated age- and sex-specific mortality rates. In a third and final stage, the predictive distribution of each future age- and sex-specific mortality

rate is adjusted so that its median coincides with the deterministic projection from the first stage.

This method has several limitations. It relies on the Spectrum multistate model, which was developed primarily to answer short-term policy questions about the AIDS epidemic, such as future need for antiretroviral therapy (ART) drugs or the future number of AIDS orphans. Because of its complexity and reliance on a large number of assumptions about transition rates between the many compartments, it was not designed for medium and longer-term projections. Indeed the UNAIDS Reference Group on Estimates, Modelling and Projection notes that Spectrum projections more than five years into the future are unreliable (UNAIDS 2015, p. 9).

The UN Population Division (UNPD) has very different needs. It provides projections of population that are long-term, to 2100, but output only overall population and vital rates, and so do not need the level of detail in Spectrum. The need to use a different methodology for a subset of countries (about 20% of the world's countries) is also a difficulty.

Here we propose an alternative methodology for probabilistic projection for the AIDS countries that is simpler than the current one, does not require any multistate modeling, and makes use of the bayesPop methodology used by the UNPD for other countries. It requires probabilistic projections of overall HIV prevalence, and these are obtained from the relatively simple non-age-structured Estimation and Projection Package (EPP) (Ghys et al. 2004; Ghys et al. 2008; Brown et al. 2010; Ghys and Garnett 2010). Future life expectancy is projected using a simple time series regression model on HIV prevalence and ART coverage (Godwin and Raftery 2017). The resulting simulated life expectancies and prevalences are converted to age-specific mortality rates using the HIV prevalence-calibrated model life table method of Sharrow et al. (2014), which replicates the middle adult mortality hump characteristic of HIV epidemics. The population projections are then obtained with the same bayesPop methodology as for other countries.

The resulting method fits observed age-specific mortality data well. In an out-of-sample validation experiment it was reasonably accurate and provided well-calibrated probabilistic projections of aggregate mortality and population quantities. It is simpler than the extant Spectrum method, but still matched its projections closely over the next 15 years.

The rest of the article is organized as follows. In the following section, we describe our new methods and the data we use. In “Results” we give results for four countries with generalized HIV/AIDS epidemics, chosen to represent a range of generalized HIV epidemic experience, namely Botswana, Zimbabwe, Mozambique and Sierra Leone. Following that we show the results of an out-of-sample validation experiment to assess our method, and we also compare the resulting projections with those of Spectrum. We conclude with a discussion of the strengths and limitations of our proposed approach.

Methods and Data

Population projection involves combining future values of age-specific fertility, age- and sex-specific mortality, and international migration rates, in this case using the standard

cohort component model. To make probabilistic projections for high-HIV prevalence countries, we follow the procedure described by Raftery et al. (2012) for making probabilistic projections of fertility in countries without generalized HIV/AIDS epidemics and use the UNPD assumptions about international migration (United Nations, Department of Economic and Social Affairs, Population Division 2015a, pp. 30–31), but we modify the mortality component to account for HIV/AIDS mortality.

Raftery et al. (2012) simulate a large number of trajectories of the Total Fertility Rate (TFR) using the Bayesian hierarchical model of Alkema et al. (2011). The projected TFRs are then converted to age-specific rates using model fertility patterns. For mortality, Raftery et al. (2012) simulate an equal number of trajectories of period life expectancy at birth (e_0) for females using the model of Raftery et al. (2013). Male life expectancies are conditional on the female e_0 and are derived from a model that predicts the gap in male and female e_0 (Raftery et al. 2014). These e_0 projections are converted to age-specific mortality rates using a variant of the Lee-Carter method (Li et al. 2013; Ševčíková et al. 2016a). Each trajectory of fertility, mortality, and migration is then converted to a future trajectory of age- and sex-specific population values using the cohort component method (Preston et al. 2001, ch. 6).

In our application, the projection of fertility remains the same (Fertility is projected using the bayesTFR package (Ševčíková et al. 2011) in the statistical analysis software, R.), along with the use of the UNPD assumptions about international migration and the cohort component method to combine these trajectories. While probabilistic methods for projecting international migration have been developed (Azose and Raftery 2015; Azose et al. 2016), in the interest of isolating any discrepancy between our results and WPP 2015 to our treatment of mortality, we do not incorporate those methods in this paper. In the following sections we outline the methods of Godwin and Raftery (2017) for projecting e_0 accounting for HIV prevalence and Sharrow et al. (2013) for converting those quantities into age-specific mortality rates.

Projecting HIV prevalence

To make probabilistic projections of e_0 for countries with generalized epidemics and to convert those projections into age-specific mortality rates, we first need projections of HIV prevalence. We use the UNAIDS Estimation and Projection Package (EPP) (Alkema et al. 2007; Brown et al. 2010; Ghys and Garnett 2010; Raftery and Bao 2010) in its so-called EPP-Classic non-age structured version for the statistical analysis software R to make probabilistic projections of HIV prevalence up to 2100.

EPP works well to project HIV prevalence into the not too distant future, approximately 5–10 years after the latest surveillance (UNAIDS 2015), but assumptions mirroring those made by UNPD in the WPP 2015 revision were imposed on two of the EPP model parameters to make projections to 2100 (United Nations, Department of Economic and Social Affairs, Population Division 2011). For most countries, the model is fitted assuming that the relevant parameters have remained constant in the past. Beginning in the start year of the projection, the parameter ϕ , which reflects the rate of recruitment of new individuals into the high-risk or susceptible group, is projected to decline by half every 20 years. The parameter r , which represents the force of infection, is projected to decline by half every 30 years. The reduction

in r reflects the assumption that changes in behavior among those subject to the risk of infection and increases in access to treatment for infected individuals will reduce the chances of HIV transmission.

For each country we use a single UNAIDS deterministic trajectory of yearly HIV prevalence to 2100 as a baseline. We construct 1000 trajectories of HIV prevalence using this deterministic trajectory and multipliers created from the output from EPP. Let $z_{c,t}$ be the median of the trajectories of HIV prevalence from EPP for country c and year t . We use

multipliers of the form $\frac{z_{c,t}^k}{z_{c,t}}$ for $k = 1, \dots, 1000$, where $z_{c,t}$ is the projected HIV prevalence for trajectory k . Combining these multipliers with the single deterministic trajectory yields a median determined by the UNAIDS trajectory and uncertainty around that determined by the EPP projections. We generate uncertainty in this manner since Spectrum/EPP cannot be relied on over the long timeframe of our forecasts. From these 1000 yearly trajectories, we construct five-year time period estimates by taking five-year averages. A single trajectory of future ART coverage is then combined with the HIV projections to be used in the projection stage.

Projecting life expectancy at birth, e_0

Because a generalized HIV epidemic can have a considerable depressing effect on life expectancy at birth in a short time (Poit et al. 2001; Ngom and Clark 2003; Blacker 2004; Timaeus and Jasseh 2004; Obermeyer et al. 2010; Bor et al. 2013; Sharrow et al. 2013; Reniers et al. 2014), a model that reflects the impact of HIV prevalence and antiretroviral therapy (ART) coverage is necessary to make appropriate projections of e_0 in generalized epidemics. What follows is a brief description of such a model developed by Godwin and Raftery (2017).

Let $HIV_{c,t}$ and $ART_{c,t}$ be the HIV prevalence and adult ART coverage in country c at time t , respectively. We define $HnA_{c,t} = HIV_{c,t} \times (100 - ART_{c,t})$, so that $HnA_{c,t}$ represents the percentage of infected people not receiving ART in country c at time t . The model for projecting e_0 , Eq. (1), predicts the five-year increase in female life expectancy at birth from time $t-5$ to t as

$$\Delta e_{0,c,(t-5:t)} = g(e_{0,c,t-5} | \theta^{(c)}) + \beta_{HnA} \Delta HnA_{c,t-5} + \varepsilon_{c,t-5}, \quad (1)$$

where $e_{0,c,(t-5:t)}$ is the change in life expectancy for country c at time $t-5$ to t , $g(e_{0,c,t-5} | \theta^{(c)})$ is the double logistic fitted change in life expectancy at time $t-5$ to t , given life expectancy at time $t-5$, and $\varepsilon_{c,t-5}$ is the error term.

The double logistic term is the same as the model used for countries not substantially impacted by HIV/AIDS (Raftery et al. 2013) and reflects the transition from high to low mortality, which can be broken down into two processes, each represented by a single logistic function. The first process describes initial slow growth in e_0 with small improvements in mortality at low levels of e_0 resulting from gains in hygiene and nutrition

followed by a quicker pace of improvement, and the second represents continuing gains from combating non-communicable diseases (Omran 2005; United Nations, Department of Economic and Social Affairs, Population Division 2015a).

The Bayesian hierarchical model (Eq. 1) is constructed by placing priors on all country-specific parameters, which can be interpreted as world parameters. Proper and diffuse hyperpriors are placed on all world parameters, and the model is estimated via Markov chain Monte Carlo simulation. See Raftery et al. (2013) and Godwin and Raftery (2017) for further details.

For our observed data we use five-year female life expectancy estimates from 1950–2015 for all 201 countries with population over 100,000 from the WPP 2015 (United Nations, Department of Economic and Social Affairs, Population Division 2015b). We use the single UNAIDS trajectory of HIV prevalence to construct five-year averages for 1950–2015. We use a single trajectory of ART coverage (percentage of seropositive individuals receiving ART) for the periods 1950–2100 for each country obtained from UNPD internal tabulations (United Nations, Department of Economic and Social Affairs, Population Division 2011). Note that countries not experiencing a generalized epidemic have HnA values of 0 for all time periods.

To project female life expectancy to 2100, we input each trajectory of HnA and life expectancy period by period into the model starting from 2015, drawing random perturbations, $\varepsilon_{c,b}$ from a zero-mean Gaussian distribution whose variance is a function of life expectancy at time t for each five-year time period. Countries with a generalized HIV/AIDS epidemic have more intrinsic variability than those without. Therefore, HIV/AIDS countries draw random perturbations from a distribution with larger variance than those not experiencing an epidemic. The process for deriving sex- and age-specific mortality rates from the projected female e_0 is described in the following section.

Converting e_0 and HIV prevalence projections to age-specific mortality rates

Once we have obtained probabilistic projections of female life expectancy and HIV prevalence, we need to map those quantities onto a set of sex- and age-specific mortality rates that can be combined with age-specific fertility rates and net migration using the cohort component method. In the WPP 2015 Revision for countries without high HIV prevalence, age-specific mortality rates are either extrapolated from recent data if they are available or by converting e_0 projections to age-specific mortality rates using model mortality patterns (United Nations, Department of Economic and Social Affairs, Population Division 2015a, pp. 27–28). However, recent data are typically unavailable in this region of the world and model mortality patterns are unable to replicate the particular age pattern of mortality resulting from large scale HIV epidemics (United Nations, Department of Economic and Social Affairs, Population Division 2015a, p. 28). Available model mortality patterns also have no relationship to HIV prevalence.

To convert the female life expectancy and HIV prevalence projections to sex- and age-specific mortality rates we use the model of Sharrow et al. (2014), shown in Eq. (2). This model can reproduce the characteristic age pattern of mortality associated with generalized

epidemics, i.e. an accentuated adult mortality hump concentrated between ages 30 and 45. The model represents the age pattern of mortality rates as a weighted sum of three age-varying components. The components, $b_{i,x}$ in Eq. (2), are derived from a Singular Value Decomposition (SVD) of the matrix of observed historical mortality rates and the weights, $w_{i,l}$ are modeled as a function of HIV prevalence and female life expectancy at birth.

The SVD is performed on a matrix that keeps the male and female mortality rates for a given country and period together such that the $b_{i,x}$ are vectors of length 44 (two times the number of age groups). This step allows us to generate male and female mortality rates simultaneously as a function of female e_0 and HIV prevalence. Because the age-varying components, $b_{i,x}$ are fixed, the component weights, $w_{i,x}$, which are modeled as a function of HIV prevalence and e_0 , are the effective parameters in this model. We refer to this model as “HIV MLT,” for “HIV-calibrated model life table.” The model is defined as follows:

$$\ln(m_{x,l}) = c_l + \sum_{i=1}^3 \omega_{i,l} b_{i,x} + \varepsilon_{x,l}, \quad (2)$$

where $m_{x,l}$ is the period age-specific mortality rate for age x in life table l , c is a constant specific to life table l , $b_{i,x}$ is the value of the i^{th} component for age x , $w_{i,l}$ is the weight of the i^{th} component for life table l , and $\varepsilon_{x,l}$ is the error term.

Figure 1(a) plots the fit from the HIV MLT model and four existing model life table systems for South Africa females for 2010–15. HIV prevalence has remained high in South Africa and was roughly 17.4% during this period, resulting in a large adult mortality hump. Figure 1(a) demonstrates how this pattern was well captured by the HIV MLT model. The other systems considered tended to produce patterns of mortality rates that matched the overall level of mortality as measured by period life expectancy at birth but missed the age-specific rates, which are critical for accurate population projection.

Note that as the epidemic matures and availability of ART increases, a paradox emerges where HIV prevalence stays elevated or even increases while life expectancy also increases as greater numbers of infected individuals live to older ages. During this period in the epidemic, the “hump” in middle age will eventually merge into the natural increase in mortality rates at older ages. HIV MLT is designed to handle this dynamic in mature epidemics with high life expectancy mitigating the effect of high prevalence producing a flat pattern of mortality rates in middle age (Sharrow et al. 2014).

This model takes e_0 and HIV prevalence as inputs and produces a set of age-specific mortality rates that reflect those two inputs, i.e. HIV MLT is designed to produce a set of age-specific mortality rates that yield an output life expectancy matching the input life expectancy. The HIV MLT model was originally calibrated with sex-specific e_0 (Sharrow et al. 2014), but for the present purpose we have recalibrated the model with the female e_0 since that is what is projected by the model described in the previous section.

To maintain the gap between male and female e_0 , the recalibrated HIV MLT model produces complete sets of male and female age-specific mortality rates simultaneously and matches

the input e_0 to the female life expectancy derived from the output female mortality rates by adjusting the intercept, c_7 in Eq. (2). The male rates are adjusted using the same adjusted intercept used to match female e_0 . Figure 1(b) plots the sex-specific output e_0 from the HIV MLT model while varying the two input parameters, female e_0 and HIV prevalence. The sex gap in life expectancy is maintained over all combinations of the two input parameters. The female output e_0 equals the input e_0 , and the male output e_0 is always below the output female e_0 .

As described above, HIV MLT itself is deterministic (i.e. for a given e_0 and prevalence, HIV MLT will produce a specific set of mortality rates), but in validation experiments we found that prediction intervals (derived after running the probabilistic e_0 and prevalence projections through HIV MLT) with the deterministic form occasionally did not cover mortality rates at the oldest ages. To increase coverage, we added noise from the error term by randomly sampling from a normal distribution for each age with mean zero and standard deviation estimated from the distribution of residuals at each age.

The HIV MLT model is calibrated with five-year age-specific mortality rates obtained from WPP 2015 from 1970–2015 for the 40 countries experiencing a generalized epidemic (defined as having >1% prevalence at any time 1980–2015). HIV prevalence for calibrating this model is the same as for the model for projecting e_0 . To produce a draw from the predictive distribution of the vector of age-specific mortality rates, a sample from each of the predictive distributions of e_0 and prevalence projections is first produced. Then a sample vector of age-specific mortality rates is produced conditionally on the simulated values of e_0 and HIV prevalence using Eq. (2).

Making full probabilistic population projections

We modified the bayesPop software (Ševčíková et al. 2016b), which combines the fertility and mortality projections using the cohort component method, to make full population projections. The software uses the method described by Raftery et al. (2012) to produce mortality projections, so the package functions were altered to include the mortality methodology described above.

Results

We discuss results here for four countries: Botswana, Zimbabwe, Mozambique and Sierra Leone. These countries were chosen to represent different levels of current HIV prevalence. Botswana and Zimbabwe represent the largest historical epidemics with peak HIV prevalence of about 25%; Mozambique represents a smaller but still substantial epidemic (peak HIV prevalence $\approx 11\%$); while Sierra Leone has a small generalized epidemic (peak HIV prevalence $\approx 1.6\%$). Results for Botswana are shown in Figure 2, while results for the other three countries are in Figures S1–S3 in Supplemental Materials.

Figure 2(a) (top left panel) plots the total population projection for Botswana. We project an increase in total population until about 2065 when the total population begins to decline (median trajectory). Also in Figure 2(a), note the increasing width of the prediction intervals as the projection reaches farther into the future reflecting the increase in uncertainty, a

feature of the projection for all countries. WPP 2015 (solid blue line) also shows sustained population growth followed by a mild reversal in that trend toward the end of the projection period, but our median projection predicts fewer people in the total population over the entire projection horizon compared to WPP 2015.

The differences between the WPP 2015 projections and ours result from our treatment of mortality, which depends on projections of HIV prevalence and life expectancy. Botswana has one of the largest HIV epidemics in the world in terms of prevalence. Figure 3 shows the probabilistic projection of HIV prevalence and life expectancy at birth for Botswana 2015–2100. HIV prevalence in Botswana is projected to remain high, dropping from a current level of about 25% to roughly 13% by 2050 in the median projection. Although declining, these high prevalence rates result in relatively low female life expectancies as modeled with Eq. (1), reaching just above 70 years by 2050, up from less than 50 years in 2005. Compared to the WPP 2015 projected life expectancy, our median projection shows sustained lower life expectancy with an increasing gap over the entire projection period.

Figure 4 shows the difference between our median projection of the total population and the WPP 2015 projection of total population as a proportion of the WPP 2015 estimate for each of the 40 countries with generalized epidemics from 2015–2100. By 2050, Botswana is projected to have one of the largest differences from WPP 2015 with about 7% fewer people in the total population compared to the WPP 2015 Revision. In addition to projecting lower female life expectancy than WPP 2015 (see Figure 3(a)), our method also produces relatively high age-specific mortality rates at ages 25–45, consistent with the mortality generated under high HIV prevalence. This decline in projected total population is evidence of the crucial role age-specific mortality rates play in shaping future population structure in countries with generalized HIV epidemics. Sustained high mortality during the reproductive years for women compared to WPP 2015 results in fewer women alive during the reproductive years and thus fewer births compared to WPP 2015.

These effects can be seen in Figs. 2(b)–2(d). Figure 2(c) shows the probabilistic population projection for women aged 15–49. We project a decline in the number of women in this age category beginning around 2055, and Figure 2(b) shows a shrinking population under age five over the entire projection period. The effect of high mortality in the reproductive adult years on future population size, especially for women, reverberates for a number of years as smaller cohorts are born in each projected period, resulting in smaller total population compared to WPP 2015. Combined with declining fertility, the effect of high adult mortality yields an eventual reversal in population growth for Botswana.

Similar conditions exist for some of the other countries with large negative proportional differences at 2050 (e.g. Zimbabwe, Namibia, Central African Republic—all with differences of greater than five percent at 2050) shown in Figure 4. These countries typically have large-scale HIV epidemics (> 10% prevalence) and our median life expectancy projection is lower than that for WPP 2015, reducing the number of women of reproductive age resulting in smaller birth cohorts. Zimbabwe, for example, is projected to have about 13.7% fewer people in the total population in 2050 compared to the WPP 2015 Revision. HIV prevalence is also projected to decline from about 14% to 10% between 2015 and 2050. We project

higher probabilities of death for women of reproductive age compared to WPP 2015 (Figure S1(d)) over the entire projection period resulting in a likely decreasing number of women of reproductive age past 2050 (Figure S1(c)) and consequently smaller birth cohorts over the projection horizon (Figure S1(b)).

For countries with smaller HIV epidemics, the reduction in the number of women of reproductive age is less severe and the difference between our projections and WPP 2015 tends to be smaller. Mozambique is projected to have approximately 7% prevalence (median projection) by 2050, down from around 11% in 2015. Figure S2(a) shows that our median projection of the total population is similar to the WPP 2015 projection over the entire projection period. These smaller differences between our projections and the WPP 2015 projections can also be seen in Figs. S2(b) and S2(c) showing the population projections for under age five and women aged 15–49, respectively.

Some countries see larger populations at 2050 in our projection compared to WPP 2015. For Sierra Leone, where HIV prevalence is projected to decrease from about 1.5% in 2015 to less than 1% by 2050, the difference from WPP 2015 is in the opposite direction. We project about 11.5% more people in the total population by 2050 compared to WPP 2015. The lower rates of HIV prevalence have a far less extreme depressing effect on total population in Sierra Leone in the long run as evidenced by Figs. S3(b) and S3(c). In addition to relatively little effect from HIV on the age-specific mortality rates, compared to WPP 2015, we project consistently higher life expectancy over the projection period for Sierra Leone, which is consistent with the life expectancy projection for other countries for which our method projects higher populations than does WPP 2015 (e.g. Djibouti, Ethiopia, Guinea Bissau, Jamaica, Nigeria; see Figure 4).

Figure 5 plots the probabilistic projections of female life expectancy for six countries with positive differences in projected population at 2050 as shown in Figure 4. Figure 5 also includes the WPP 2015 female life expectancy projections for these countries. As this figure shows, a large portion of the difference between the WPP 2015 total population projections and our median projections for these countries arises from differences in the projections of life expectancy, since these countries have comparatively small HIV epidemics, which will limit the influence of prevalence on the age pattern of mortality rates.

Validation

Out-of-Sample Validation

To validate our method, we calibrated the e_0 projection and HIV MLT models with five-year data from the WPP 2015 revision up to 2010 and used those models to generate a population projection for the 2010–15 period for each of the 40 countries. We then compare the resulting mortality and population distributions with the observations from WPP 2015 for the 2010–15 period.

Because our method addresses the mortality component of the projection, we first assessed the accuracy of the mortality predictions for 2010–15 by calculating the mean absolute error (MAE) of our median projection for 2010–15, treating the WPP 2015 estimate as the

observed value. We did this for four mortality indicators: e_0 (life expectancy at birth), ${}_5q_0$ (the probability a newborn will die before reaching age five), ${}_{45}q_{15}$ (the probability that a 15-year-old will die before reaching age 60), and ${}_{35}q_{10}$ (the probability that a 10-year-old will die before age 45). Table 1 presents the MAE by sex for the four mortality indicators. For males, the mean absolute error for life expectancy is about 2.3 years, while the MAE for e_0 for females is slightly higher at 2.5 years, suggesting a good fit for the level of mortality. For the other three indicators, the MAE is less than 4 per 1000 for both males and females. Overall, our method predicted the WPP 2015 estimate for 2010–15 well for most countries (Figures S4–S9 plot the predicted distribution of these three indicators for 2010–15 along with the WPP 2015 estimate for the same period for all 40 countries.).

We further assess the accuracy of our method for projecting age-specific mortality rates by calculating the proportion of age-specific mortality rates for 2010–15 (observed data that were left out for calibration) captured in the 80%, 90%, and 95% predictive intervals (PI) for 2010–15 produced with our method. If the method is doing well, we should capture about the same percentage of the out-of-sample mortality rates as defines the predictive intervals. The top two rows of Table 2 show the coverage (percentage of WPP 2015 mortality rates across all ages and all 40 countries captured by the various interval widths) for the 2010–15 period. For each PI, our method captures just under the expected proportion, indicating that the method provides reasonably well-calibrated projections for the out-of-sample period. The third row of Table 2 gives the coverage for the single quantity of female life expectancy in the out-of-sample period (i.e. the percentage of countries whose PIs captured the WPP 2015 estimate for female life expectancy). Coverage for female life expectancy is also good with our method again capturing approximately the expected proportion of predicted out-of-sample life expectancies.

Finally, we calculate coverage similar to that of mortality and life expectancy but for the total population prediction at 2010–15 (Calculation of total population includes running fertility projections for the out-of-sample period 2010–15 with the bayesTFR software.). The bottom row of Table 2 shows the percentage of WPP 2015 total population estimates for 2010–15 captured in the 80%, 90%, and 95% predictive intervals. While the coverages for the mortality rates were not markedly different from the nominal coverages of the PIs, coverage of the single total population estimate is near universal (i.e. the observed total population estimate for the out-of-sample period 2010–15 was within the 80% and 90% PI for all but one country, Swaziland, and within the 95% PI for all countries). This near universal coverage of the total population estimate is not surprising since total population is unlikely to be greatly impacted by the changes we make to the mortality component of the projection in only one five-year period. The change in mortality from the method used in WPP 2015 markedly affects only certain ages and even then may not show those effects for some time on total population. Further mitigating the short-term effect of mortality change is the treatment of fertility, which remains the same in our projections and WPP 2015.

Comparison to Spectrum/EPP

Spectrum/EPP (Stanecki et al. 2012; Stover et al. 2012; Avenir Health 2016) provides useful information to UNAIDS about high-HIV prevalence countries including estimates of HIV

prevalence, total population, life expectancy, and a host of other demographic and epidemiological variables in the short term (five years). However, it is, by necessity, quite complex. We propose a simpler method to project over a longer projection horizon, but it is instructive to compare our projection results to those obtained from Spectrum in the near term.

Table 3 shows the mean difference (MD) between our median projection using the full dataset for calibration (i.e. no out-of-sample period was removed) and Spectrum among all 40 countries for prevalence and life expectancy at birth for the first three projection periods. All Spectrum results were obtained with Spectrum version 5.441, downloaded July 2016 (Avenir Health 2016). For prevalence, our projections differed from Spectrum by less than one percentage point on average across the 40 countries during each of the first three projection periods, which is substantially less than the prediction margin of error. The mean differences for life expectancy show that on average across the 40 countries our projection is between one and two years lower than the Spectrum projection in the very near term. Thus, overall our projections are similar to those of Spectrum for these quantities, but using a much simpler model.

The right three columns of Table 3 show the mean proportional difference (MPD) between our median projection and Spectrum for three population quantities: the total population across all ages, the total population aged 0–4, and the female population aged 15–49. The MPD is the average difference between our median estimate and Spectrum as a proportion of the Spectrum estimate for a given population indicator. Again, we present a simpler method for long-term projection, but it should approximate Spectrum at least in the short run. For total population, our median estimates of total population were about one percent lower than the Spectrum estimate for the most recent projection period, 2015–20. The mean proportional difference is between one and two percent for each of the next two periods. Our projections of the population aged 0–4 are also lower than the Spectrum result by roughly one percent. Finally, our short-term projection results for the female reproductive age population are also close to the Spectrum result with less than one percent average proportional difference over all three periods. In sum, results from Table 3 suggest our less complex method approximates the short-term projections of Spectrum for these demographic quantities reasonably closely.

Discussion

We have presented a method for making probabilistic population projections for countries with generalized HIV epidemics. We accomplish this by following the Bayesian probabilistic projection method described by Raftery et al. (2012) for fertility and the international migration assumptions of the UNPD, but because of the singular nature of mortality in generalized HIV epidemics, we modify the mortality component of the projection to incorporate the future trajectory of the epidemic in terms of HIV prevalence and ART coverage. The probabilistic fertility and mortality projections and the UNPD's assumptions about future migration are combined using the cohort component method of projection.

These projections are potentially useful to researchers and policy makers as this method provides a predictive distribution for population quantities of interest such as total population, life expectancy, and support ratios into the future. Our method takes into account uncertainty about future levels of mortality and fertility, the major drivers of population change, as well as uncertainty about the trajectory of HIV prevalence. Our approach is less complex than the UNPD's current method for projecting mortality in high-HIV prevalence settings and better captures the age pattern of mortality rates during a generalized HIV epidemic.

Results from the projections described here show that by 2050 and beyond, we project smaller total populations for about half of the 40 countries under study here compared to WPP 2015. Many of the countries with the largest negative differences in projected total population compared to WPP 2015 have large-scale HIV epidemics. For these countries, we tend to project lower total life expectancy over the course of the projection period. Combined with projected high HIV prevalence, the lower life expectancies result in high age-specific mortality rates during the younger adult years, and thus fewer women of reproductive age and consequently smaller birth cohorts. Projected into the future, these trends lead to smaller total population projections compared to WPP 2015. Coupled with declining fertility, high mortality rates resulting from HIV/AIDS-related deaths produce a reversal in population growth by 2100 for some countries with very large epidemics. Overall, these trends amount to a -1.7% difference in total population amongst all 40 countries by 2100 compared to WPP 2015, a difference of approximately 52.5 million people.

Although the method presented here for mortality and elsewhere for fertility takes into account uncertainty about future levels of fertility and mortality, it does not include uncertainty about international migration in the future, which can be an important source of forecast errors. While this paper focuses on the treatment of mortality in population forecasts for countries with generalized HIV/AIDS epidemics, future projections should incorporate recent advances in probabilistic forecasting of migration (Azose and Raftery 2015; Azose et al. 2016) to address this potential source of forecast error. Likewise, the life expectancy projection model and the model used to convert e_0 projections to age-specific mortality rates are calibrated with results from WPP 2015, some of which are themselves modeled, so they reproduce only the variability in the quantities of interest contained in the WPP results. To the extent that the WPP 2015 data and results used to calibrate these models reflect the empirical reality, the models we present here should as well. Finally, as in Raftery et al. (2012), this method does not take into account random variation in the number of births or deaths, given the fertility and mortality rates.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development under Grant K01HD057246, Grant R01HD054511, and Grant R01 HD070936.

References

- Alkema, Leontine, Raftery, Adrian E., Clark, Samuel J. Probabilistic projections of HIV prevalence using Bayesian melding. *Annals of Applied Statistics*. 2007; 1(1):229–248.
- Alkema, Leontine, Raftery, Adrian E., Gerland, Patrick, Clark, Samuel J., Pelletier, François, Buettner, Thomas, Heilig, Gerhard K. Probabilistic projections of the total fertility rate for all countries. *Demography*. 2011; 48(3):815–839. [PubMed: 21748544]
- Avenir Health. Spectrum/EPP version 5.441. 2016. (downloaded July 2016). <http://www.unaids.org/en/dataanalysis/datatools/spectrumepp>
- Azose, Jonathan J., Raftery, Adrian E. Bayesian probabilistic projection of international migration. *Demography*. 2015; 52(5):1627–1650. [PubMed: 26358699]
- Azose, Jonathan J., Ševčíková, Hana, Raftery, Adrian E. Probabilistic population projections with migration uncertainty. *Proceedings of the National Academy of Sciences*. 2016; 113(23):6460–6465.
- Blacker J. The impact of AIDS on adult mortality: Evidence from national and regional statistics. *AIDS*. 2004; 18(suppl. 2):S19–S26.
- Bor, Jacob, Herbst, Abraham J., Newell, Marie-Louise, Barnighausen, Till. Increases in adult life expectancy in rural South Africa: Valuing the scale-up of HIV treatment. *Science*. 2013; 339(6122):961. [PubMed: 23430655]
- Brown, Tim, Bao, Le, Raftery, Adrian E., Salomon, Joshua A., Baggaley, Rebecca F., Stover, John, Gerland, Patrick. Modeling HIV epidemics in the antiretroviral era: The UNAIDS Estimation and Projection Package 2009. *Sexually Transmitted Infections*. 2010; 86:3–10.
- Coale, Ansley J., Demeny, Paul. *Regional Model Life Tables and Stable Populations*. Princeton University Press; 1966.
- Coale, Ansley J., Demeny, Paul, Vaughan, Barbara. *Regional Model Life Tables and Stable Populations*. 2. Academic Press; New York: 1983.
- Fosdick, Bailey K., Raftery, Adrian E. Regional probabilistic fertility forecasting by modeling between-country correlations. *Demographic Research*. 2014; 30:1011–1034. [PubMed: 25242889]
- Ghys PD, Brown T, Grassly NC, Garnett G, Stanecki KA, Stover J, Walker N. The UNAIDS Estimation and Projection Package: A software package to estimate and project national HIV epidemics. *Sexual Transmitted Infections*. 2004; 80:i5–i9.
- Ghys, Peter D., Garnett, Geoff P. The 2009 HIV and AIDS estimates and projections: Methods, tools and analyses. *Sexually Transmitted Infections*. 2010; (Suppl. 2):86. ii1–ii2.
- Ghys PD, Walker N, McFarland W, Miller R, Garnett GP. Improved data, methods and tools for the 2007 HIV and AIDS estimates and projections. *Sexually Transmitted Infections*. 2008; 84:i1–i4. [PubMed: 18647859]
- Godwin, Jessica, Raftery, Adrian E. Bayesian projection of life expectancy accounting for the HIV/AIDS epidemic. *Demographic Research*. 2017 forthcoming.
- Lee, Ronald D., Tuljapurkar, Shripad. Stochastic population forecasts for the United States: Beyond high, medium, and low. *Journal of the American Statistical Association*. 1994; 89:1175–1189.
- Li, Nan, Gerland, Patrick. Paper presented at the Population Association of America annual meeting. Washington, D.C: 2011. Modifying the lee-carter method to project mortality changes up to 2100.
- Li, Nan, Lee, Ronald, Gerland, Patrick. Extending the Lee-Carter method to model the rotation of age patterns of mortality decline for long-term projections. *Demography*. 2013; 50(6):2037–2051. [PubMed: 23904392]
- Lutz, Wolfgang, KC, Samir. Dimensions of global population projections: What do we know about future population trends and structures? *Philosophical Transactions of the Royal Society B*. 2010; 365:2779–2791.
- Murray, Christopher JL., Ferguson, Brodie D., Lopez, Alan D., Guillot, Michel, Salomon, Joshua A., Ahmad, Omar. Modified logit life table system: Principles, empirical validation, and application. *Population Studies*. 2003; 57:165–182.
- National Research Council. *Beyond Six Billion: Forecasting the World's Population*. National Academy Press; Washington, D.C: 2000.

- Ngom, Pierre, Clark, Samuel J. Adult mortality in the era of HIV/AIDS: Sub-Saharan Africa. Technical report, Population Division, United Nations; New York: 2003.
- Obermeyer, Ziad, Rajaratnam, Julie Knoll, Park, Chang H., Gakidou, Emmanuela, Hogan, Margaret C., Lopez, Alan D., Murray, Christopher JL. Measuring adult mortality using sibling survival: A new analytical method and new results for 44 countries, 1974–2006. *PLoS Medicine*. 2010; 7(4)
- Omran, Abdel R. The epidemiologic transition: A theory of the epidemiology of population change. *Milbank Quarterly*. 2005; 83(4):731–757. [PubMed: 16279965]
- Poit, Peter, Bartos, Michael, Ghys, Peter D., Walker, Neff, Schwartländer, Bernhard. The global impact of HIV/AIDS. *Nature*. 2001; 410:968–973. [PubMed: 11309626]
- Preston, Samuel H., Heuveline, Patrick, Guillot, Michel. *Demography: Measuring and Modeling Population Processes*. Blackwell; Malden, MA: 2001.
- Raftery, Adrian E., Bao, Le. Estimating and projecting trends in HIV/AIDS generalized epidemics using incremental mixture importance sampling. *Biometrics*. 2010; 66(4):1162–1173. [PubMed: 20222935]
- Raftery, Adrian E., Chunn, Jennifer L., Gerland, Patrick, Šev íková, Hana. Bayesian probabilistic projections of life expectancy for all countries. *Demography*. 2013; 50(3):777–801. [PubMed: 23494599]
- Raftery, Adrian E., Lalic, Nevena, Gerland, Patrick. Joint probabilistic projection of female and male life expectancy. *Demographic Research*. 2014; 30(4):795–82. [PubMed: 25580082]
- Raftery, Adrian E., Li, Nan, Šev íková, Hana, Gerland, Patrick, Heilig, Gerhard K. Bayesian probabilistic population projections for all countries. *Proceedings of the National Academy of Sciences of the United States of America*. 2012; 109(35):13915–13921. [PubMed: 22908249]
- Reniers, Georges, Slaymaker, Emma, Nakiyingi-Miir, Jessica, Nyamukapa, Constance, Crampin, Amelia Catharine, Herbst, Kobus, Urassa, Mark, Otieno, Fred, Gregson, Simon, Sewe, Maquins, Michael, Denna, Lutalo, Tom, Hosegood, Victoria, Kasamba, Ivan, Price, Alison, Nabukalu, Doreen, Mclean, Estelle, Zaba, Basia, on behalf of the ALPHA Network. Mortality trends in the era of antiretroviral therapy: Evidence from the network for Analysing Longitudinal Population based HIV/AIDS data on Africa (ALPHA). *AIDS*. 2014; 28(Suppl 4):S533. [PubMed: 25406756]
- Šev íková, Hana, Alkema, Leontine, Raftery, Adrian E. bayesTFR: An R package for probabilistic projections of the total fertility rate. *Journal of Statistical Software*. 2011; 43(1):1–29.
- Šev íková, Hana, Li, Nan, Kantorová, Vladimíra, Gerland, Patrick, Raftery, Adrian E. Age-specific mortality and fertility rates for probabilistic population projections. In: Schoen, R., editor. *Dynamic Demographic Analysis*. Cham: Springer; 2016a. p. 285–310.
- Šev íková, Hana, Raftery, Adrian E., Buettner, T. bayesPop: Probabilistic Population Projection. R package version 6.0-1. 2016b
- Sharrow, David J., Clark, Samuel J., Collinson, Mark, Kahn, Kathleen, Tollman, Stephen. The age pattern of increases in mortality affected by HIV: Bayesian fit of the Heligman-Pollard model to data from the Agincourt HDSS field site in rural northeast South Africa. *Demographic Research*. 2013; 29(39):1039–1096. [PubMed: 24453696]
- Sharrow, David J., Clark, Samuel J., Raftery, Adrian E. Modeling age-specific mortality for countries with generalized HIV epidemics. *PLoS ONE*. 2014; (5):9.
- Stanecki, Karen, Garnett, Geoff P., Ghys, Peter D. Developments in the field of HIV estimates: Methods, parameters and trends. *Sexually Transmitted Infections*. 2012; 88:i1–i2. [PubMed: 23172340]
- Stover, John, Brown, Tim, Marston, Milly. Updates to the Spectrum/Estimation and Projection Package (EPP) model to estimate HIV trends for adults and children. *Sexually Transmitted Infections*. 2012; 88:11–16.
- Timæus, Ian M., Jasseh, Momodou. Adult mortality in Sub-Saharan Africa: Evidence from Demographic and Health Surveys. *Demography*. 2004; 41(4):757–772. [PubMed: 15622953]
- UNAIDS. *Quick Start Guide for Spectrum 2015*. UNAIDS. 2015
- United Nations, Department of Economic and Social Affairs, Population Division. *World Population Prospects: The 2010 Revision, Special Tabulations*. 2011

- United Nations, Department of Economic and Social Affairs, Population Division. World population prospects: The 2015 revision, methodology of the United Nations population estimates and projections. 2015a Paper No. ESA/P/WP.242.
- United Nations, Department of Economic and Social Affairs, Population Division. World population prospects: The 2015 revision, volume i: Comprehensive tables. 2015b Paper No. ST/ESA/SER.A/379.
- United Nations. Department of International Economic and Social Affairs. Model life tables for developing countries. United Nations; New York: 1982.
- Whelpton PK. An empirical method of calculating future population. Journal of the American Statistical Association. 1936; 31(195):457–473.
- Wilmoth, John, Zureick, Sarah, Canudas-Romo, Vladimir, Inoue, Mie, Sawyer, Cheryl. A flexible two-dimensional mortality model for use in indirect estimation. Population Studies. 2012; 66(1):1–28. [PubMed: 22150635]

(a) Fit of HIV MLT model for South Africa females 2010–2015 (b) Sex-specific e_0 output for HIV MLT model at varying female e_0 and prevalence

Figure 1.

Output from HIV MLT model. (a) Fit of HIV MLT model for South Africa females 2010–2015. HIV MLT model shown with black line. For comparison, fits from the WHO modified logit model (Murray et al. 2003) [red solid line], Coale and Demeny model life tables (Coale and Demeny 1966; Coale et al. 1983) [green solid line], UN model life tables for developing countries (United Nations. Department of International Economic and Social Affairs 1982) [teal solid line], and the Log-Quad model (Wilmoth et al. 2012) [purple solid line] are also shown. (b) Sex-specific output e_0 , varying the two HIV MLT model inputs: female e_0 and HIV prevalence.

(a) Population Projection: total population

(b) Population Projection: age 0-4

(c) Population Projection: female age 15-49

(d) Projection of female ${}_{35}q_{10}$ **Figure 2.**

Probabilistic population projections for Botswana 2015–2100. Past, observed data: black circles; median probabilistic projection: solid red line, 80% predictive interval: dashed red lines; 95% predictive interval: dotted red lines; WPP 2015 projection: solid blue line.

(a) Probabilistic Life Expectancy Projection

(b) Probabilistic HIV Prevalence Projection

Figure 3.

Probabilistic life expectancy and HIV prevalence projections for Botswana 2015–2100.

Median of probabilistic projection: solid, red line; 80% predictive interval: dashed, red line; 95% predictive interval: dotted, red line; WPP 2015 projection: solid, blue line; observed: black circles. The gray lines in these figures are a random sample of ten trajectories from the final sample of 1,000 trajectories from the posterior distribution.

Figure 4.

Comparison between our median total population projection and WPP 2015 median projection for total population. Each line shows the percentage difference from the WPP 2015 total population estimate for each country and the solid black line shows the difference at each period for all countries combined.

Figure 5.

Probabilistic life expectancy projections for six countries 2015–2100. Median of probabilistic projection: solid, red line; 80% predictive interval: dashed, red line; 95% predictive interval: dotted, red line; WPP 2015 projection: solid, blue line; observed: black circles. The gray lines in these figures are a random sample of ten trajectories from the final sample of 1,000 trajectories from the posterior distribution.

Table 1

Mean absolute error for four mortality indicators for the 2010–15 out-of-sample period ($5q_0$, $45q_{15}$, and $35q_{10}$, all $\times 1000$) using our method for projecting HIV prevalence and life expectancy and using the HIV MLT model for converting to age-specific mortality rates. All models have been calibrated with data up to 2010 and used to predict the out-of-sample period 2010–15.

	Mean Absolute Error (MAE)			
	e_0	$5q_0$	$45q_{15}$	$35q_{10}$
Male	2.3	16.3	31.9	22.9
Female	2.5	13.9	33.4	30.6

Table 2

Coverage of WPP 2015 sex-age-specific mortality rates, female life expectancy at birth, and total population in the 80%, 90%, and 95% PIs for the 2010–15 out-of-sample period using our method for projecting HIV prevalence and life expectancy and using the HIV MLT model for converting to age-specific mortality rates. Fertility and migration were projected using the bayesPop software. Data from 1970–2010 were used for calibration, and the resulting estimated model was used to predict the out-of-sample period 2010–15.

	80% PI	90% PI	95%PI
m_x , male	75	82	88
m_x , female	69	81	86
e_0 , female	75	88	90
Total Population	97	97	100

Table 3

Mean difference (MD) for prevalence and life expectancy and mean proportional difference (MPD) for total population, population aged 0–4, and female population aged 15–49, between our method and the Spectrum/EPP software for the first three projection periods. All results for Spectrum derived from the Spectrum software version 5.441.

	MD ¹		MPD ²	
	Prevalence	e0	Total Population	Female Population 15–49
2015–20	0.7	–1.76	–0.9	–0.8
2020–25	0.8	–1.10	–1.3	–1.0
2025–30	0.9	–0.97	–1.6	–1.2

¹Mean Difference between our estimate and that of the Spectrum software for all 40 countries. These numbers are on the scale of prevalence (percentage points) and life expectancy (years).

²Mean Proportional Difference is the difference between our median estimate and the Spectrum estimate as a proportion of the Spectrum estimate. Varying absolute population sizes among the 40 countries necessitate taking the proportional error. MPD is in units of per 100.