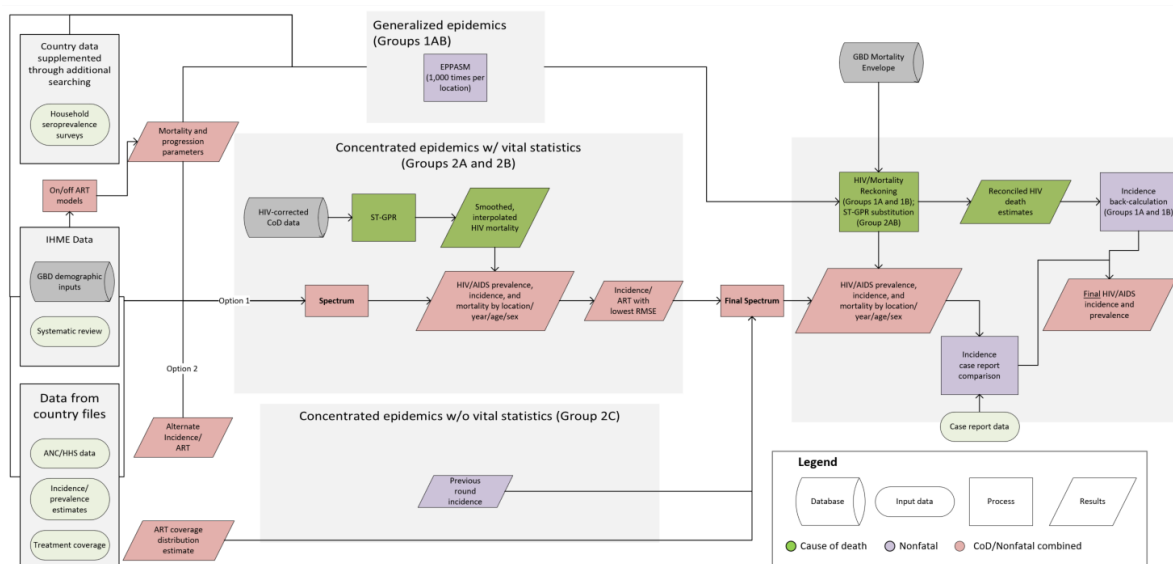


HIV/AIDS



Case definition

Infection with the human immunodeficiency virus (HIV) causes influenza-like symptoms during the acute period following infection and can lead to acquired immunodeficiency syndrome (AIDS) if untreated. HIV attacks the immune system of its host, leaving infected individuals more susceptible to opportunistic infections like tuberculosis. Although there are two different subtypes of HIV, HIV-1 and HIV-2, no distinction is made in our estimation process or presentation of results. For HIV, ICD 10 codes are B20-B24, C46-C469, D84.9; ICD 9 codes are 042-044, 112-118 (after 1980), 130 (after 1980), 136.3-136.8 (after 1980), 176.0-176.9 (after 1980), 279 (after 1980); and ICD9 BTL codes are B184-B185.

Input data

Household seroprevalence surveys

Geographically representative HIV seroprevalence survey results were used as inputs to the model for countries with generalised HIV epidemics where available.

GBD demographic inputs

Location-specific population, fertility, migration and HIV-free survival rates from GBD 2020 were used as inputs in modelling all locations.

Data from countries

The files compiled by UNAIDS for their HIV/AIDS estimation process were one of our sources of data for producing estimates of HIV burden. The files are often built by within-country experts with the support of UNAIDS, which publishes estimates annually on behalf of countries and only shares their files when permission is granted. The files contain the HIV-specific information which is needed to run the Spectrum¹, and Estimation and Projection Package-Age Sex² (EPP-ASM) models.

Spectrum and EPP-ASM require the following input data: AIDS mortality among people living with HIV with and without ART, CD4 progression among people living with HIV not on ART, ART coverage among adults and children, cotrimoxazole coverage among children, coverage of breastfeeding among women living with HIV, prevention of mother-to-child transmission coverage, and CD4 thresholds for treatment eligibility. EPP-ASM additionally uses HIV prevalence data from surveillance sites and representative surveys. Antenatal care (ANC), incidence, prevalence, and treatment coverage data from UNAIDS were used in modelling for all locations. We extracted all of these data from the proprietary format used by UNAIDS.

Changes for GBD 2020

We supplemented the antenatal care and treatment coverage data available through processing done by the Local Burden of Disease team,³ and retrieving data on adult antiretroviral (ART) treatment coverage rates from country reports, respectively. The addition of ANC sites affected 33 countries, while ART data was added in 45 countries. During the Local Burden of Disease alignment process the antenatal care clinic prevalence estimates were corrected in a number of facets. There were 17 estimates with placeholder sample sizes that were corrected, duplicate observations in Togo were removed, 123 additional observations were added, 1491 non-ANC observations were removed, and 232 points were outliered based on comparison reports of HIV burden in a given area.

We did not have country UNAIDS files for 40 locations, many of them countries with small populations and/or low HIV prevalence. As in previous rounds, we generated regional averages of all needed inputs in these locations. This enabled us to run Spectrum for every GBD location.

Vital registration data

We used all available sources of vital registration and sample registration data from the GBD Causes of Death database after garbage code redistribution and HIV/AIDS mis-coding correction⁴ in Group 2 countries and India. There are two different cause of death data sources for HIV/AIDS in China: the Disease Surveillance Point (DSP) system and the Notifiable Infectious Disease Reporting (NIDR) system. Both systems are administered by the Chinese Center for Disease Control and Prevention, but the reported number of deaths due to HIV is significantly lower in DSP. Therefore, we have used the provincial-level ratio of deaths due to HIV/AIDS from NIDR to those from DSP, choosing the larger ratio between years 2013 and 2014, and scaled the reported deaths in the DSP system, which is in turn used in the spatiotemporal Gaussian process regression (ST-GPR). ST-GPR deaths were used as final deaths in group 2A and group 2B.

Case notifications data

We searched for case notifications data using the ECDC database and country reports series in countries with four- and five-star vital registration data. We identified 59 countries with available information.

On-ART literature data

Data were identified by using search string: "hiv"[MeSH Terms] OR "hiv"[All Fields]) AND ("mortality"[Subheading] OR "mortality"[All Fields] OR "mortality"[MeSH Terms]) AND antiretroviral[All Fields] AND ("therapy"[Subheading] OR "therapy"[All Fields] OR "therapeutics"[MeSH Terms] OR "therapeutics"[All Fields]) in PubMed.

To be included, studies must include only HIV-positive people over the age of 15, who receive antiretroviral therapy (ART) but who were ART-naïve prior to the study. In addition, studies must report either a duration-specific (time since initiation of ART) mortality proportion or a hazard ratio across age or sex, and must not include children.

For duration-specific survival data, studies must report uncertainty on mortality estimates or provide stratum-specific sample sizes and must include duration-specific data to allow for calculation of 0-6, 7-12, or 13-24 month conditional mortality. In addition, studies must either report separate mortality and loss-to-follow-up (LTFU) curves, be corrected for LTFU using vital registration data or double sampling or be conducted in a high-income setting. Finally, studies must report the percentage of participants who are male and the median age of participants.

Hazard ratio data for ages or sexes can only be used if the hazard ratios are controlled for other variables of interest (age, sex, and CD4 category). In GBD 2020, we included 61 studies, thirteen of which were new this cycle. Of these studies, we added 10 to inform the estimation age-sex hazard ratios, and three studies informed LTFU curves.

Off-ART literature data

In GBD 2013, we systematically reviewed the literature on mortality without ART to characterise uncertainty in the progression and death rates. We searched terms related to pre-ART or ART-naive survival since seroconversion.⁵ After screening, we identified 13 cohort studies that included the cohorts used by UNAIDS, from which we extracted survival at each one-year point after infection. Screening for additional, recently published studies in GBD 2015, GBD 2016 and GBD 2017 identified no new cohort studies for inclusion in this analysis. We did not search for new studies in GBD 2019 or GBD 2020.

Severity splits and disability weights

The basis of the GBD disability weight survey assessments are lay descriptions of sequelae highlighting major functional consequences and symptoms. The lay descriptions and disability weights for HIV/AIDS severity levels are shown below.

Severity level	Lay description	DW (95% CI)
Symptomatic HIV	Has weight loss, fatigue, and frequent infections.	0.274 (0.184–0.377)
AIDS with antiretroviral treatment	Has occasional fevers and infections. The person takes daily medication that sometimes causes diarrhoea.	0.078 (0.052–0.111)
AIDS without antiretroviral treatment	Has severe weight loss, weakness, fatigue, cough and fever, and frequent infections, skin rashes, and diarrhoea.	0.582 (0.406–0.743)

Modelling strategy

Countries were divided into groups: Group 1A and 1B and 2A, 2B and 2C.

In group 1 countries we used EPP-ASM; these were mostly countries in Sub-Saharan Africa with HIV peak prevalence of greater than 0.5%. Group 1A consisted of countries with available antenatal care clinic

data and household surveys Similarly, Group 1B was similar to 1A but had HIV peak prevalence of at least 0.25%. Mauritania, Sao Tome and Principe, and Comoros were modelled as group 1A countries in GBD2020.

The remaining countries comprised group 2, where we continued to use the Spectrum program rewritten in Python for GBD 2013 to facilitate faster and more flexible execution necessary for our more intensive computational needs. Group 2A consisted of countries with high quality vital registration data, Group 2B consisted of countries with any available vital registration data which was generally lower quality, and Group 2C countries were those without any vital registration data. Quality was measured based on a star rating system as described elsewhere.⁶

On-ART

First, we corrected reported probabilities of death for loss to follow-up using an approach developed by Verguet and colleagues.⁶ Verguet and colleagues used tracing and follow-up studies to empirically estimate the relationship between death in loss to follow-up (LTFU) and the rate of LTFU.

To create estimates of age-specific hazard ratios, we synthesised hazard ratio data in five broad age groups: 15-25, 25-35, 35-45, 45-55, 55-100, and modelled the data using DisMod-MR 2.1.

To create estimates of sex-specific hazard ratios, we use the *metan* function in Stata to create estimates of relative risks separately by region, using female as the reference group.

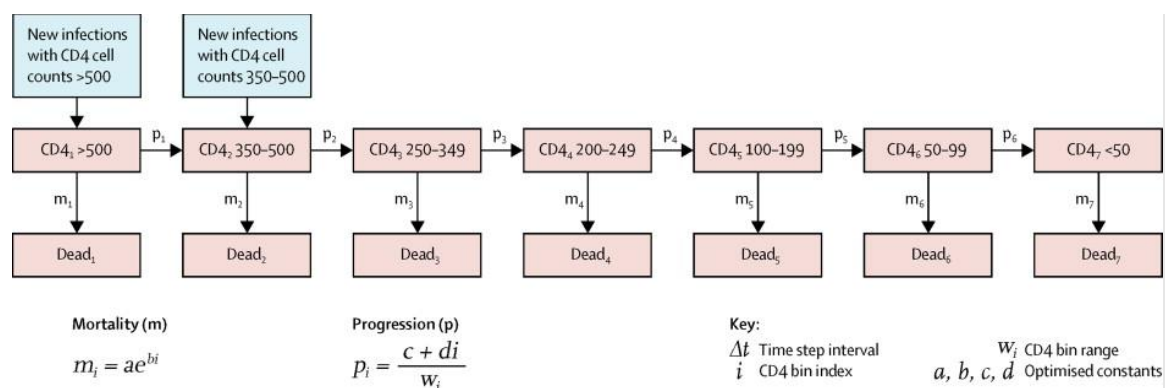
The age and sex hazard ratios were applied to the study-level mortality rates, accounting for the distribution of ages and sexes in the mortality data. We then subtracted HIV-free mortality from the model life table process to calculate study-level age-sex HIV-specific mortality.

Changes for GBD 2020

To synthesise the age-sex-split study-level data into estimates of conditional probability of death over initial CD4 count for GBD 2020, we replaced the use of Dismod⁶ in favour of the Meta Regression; Bayesian; Regularized; Trimmed (MR-BRT) model,⁷ This model is a mixed effects meta-regression that accounts for between-study heterogeneity. We ran MR-BRT models for each age group (15-25, 25-35, 35-45, 45-55, or 55-100), sex (male or female), duration since ART initiation (0-6, 7-12 or 13-24 months) and super region (sub-Saharan Africa, High-income, or Other) strata.

Off-ART

Following UNAIDS assumptions, no-ART mortality is modelled as shown in the figure below.



The death and progression rates between CD4 categories vary by age according to four age groups: 15–24 years, 25–34 years, 35–44 years, and 45 years or older. We modelled the logit of the conditional probability of death between years in these studies using the following formula:

$$\text{logit}(m_{ijk}) = \beta_0 + \sum_{i=1}^4 \beta_{1i} a_i + \sum_{j=1}^{12} \beta_{2j} t_j + u_k + \epsilon_{ijk}$$

In the formula, m is conditional probability of death from year t_j to t_{j+1} , a_i is an indicator variable for age group at seroconversion (15–24 years, 25–34 years, 35–44 years, and 45 years or older), t_j is an indicator variable of year since seroconversion, and u_k is a study-level random effect.

By sampling the variance-covariance matrix of the regression coefficients and the study-level random effect, we generated 1,000 survival curves for each age group that capture the systematic variation in survival across the available studies. For each of the 1,000 survival curves, we used a framework modelled after the UNAIDS optimisation framework in which we find a set of progression and death rates that minimises the sum of the squared errors for the fit to the survival curve.^{8,9}

Finally, in cases where estimated on-ART mortality rates were higher than off-ART mortality rates, we replaced our estimated on-ART mortality rates by the corresponding off-ART mortality rates to account for progression to lower CD4 categories. This ensured individuals would not experience higher mortality when they entered treatment in Spectrum or EPP-ASM.

GBD 2020 HIV burden estimation overview

We used two different components to derive year-, age- and sex-specific estimates of HIV incidence, prevalence, and mortality depending on locations' availability of data and extent of HIV burden, as described below:

1. EPP-ASM was used to estimate incidence, prevalence, and mortality that are consistent with serosurveillance data from antenatal care clinics and/or prevalence surveys.

Spectrum is a compartmental HIV progression model used to generate age-sex-specific incidence, prevalence, and death rates from input incidence and prevalence curves and assumptions about intervention scale-up and local variation in epidemiology. This model was used for all Group 2 countries, and in conjunction with EPPASMEPP-ASM for India.

EPP-ASM model

For GBD 2020, we continued to use our modified version of EPP-ASM both to improve the fit to age-sex specific prevalence survey data among adults and to generate paediatric estimates. We built a paediatric module in EPP-ASM that mirrored early updates to the paediatric module in Spectrum.¹⁰ This child module included CD4 progression and CD4-specific mortality rates taken from a model fit to survival data from leDEA and child initiation of ART based on ART distribution data from leDEA. Perinatal and breastfeeding transmission was calculated as a function of prevalence among pregnant women and PMTCT programme data. We were thus able to utilise EPP-ASM to produce HIV incidence, prevalence, and mortality estimates for all ages. Additionally, we improved fit to prevalence data through allowing flexibility in the age distribution of incidence over time. We parameterised the ratio of incidence among ages 15-24:25+ as a constant before year 2000 and a linear regression thereafter. This allowed for the shifts in the age distribution of incidence observed over the course of the HIV epidemic to be reflected in our results. Finally, we utilised GBD demographic inputs and substituted in our own assumptions about HIV progression rates and on/off-ART mortality.

To incorporate uncertainty in our demographic and progression parameters, we run EPP-ASM with separate draws of CD4 progression, on- and off-ART mortality rates, fertility, and HIV-free mortality. This

process produced 1000 posterior distributions for each of the locations that make up Group 1. For every location in the group, we sampled one draw from each of the sets of EPP-ASM results in order to create a final distribution of 1000 draws. By sampling one draw from each set, we ensured that the distribution of mortality parameters dictating the relationship between incidence and prevalence aligned with those used in the GBD demographics estimates.

We also continued to use the approach implemented in GBD 2019 to address selection bias resulting from temporal and geographical variation in ANC reporting. The ANC data which EPP-ASM uses cannot be assumed as representative of HIV prevalence in the full population. This is especially the case when there are minimal or no nationally representative prevalence surveys to anchor estimates, as in the early epidemic.¹¹

EPP-ASM has embedded approaches to adjust for the bias associated with using prevalence among ANC-site-attending pregnant women to estimate prevalence among the both-sexes population. For the bias between pregnant women and the national both-sexes population, it makes assumptions around the difference in total fertility rate among HIV positive and HIV negative women, and the difference in prevalence between men and women. For the bias associated with the data coming from ANC sites, the specification of the likelihood of observed ANC data includes random intercepts for each clinic. The random intercepts allow each site's baseline prevalence to vary randomly around the overall mean prevalence. In other words, factors that could drive differences between sites' HIV prevalence levels are "adjusted" for.

However, the embedded approach does not explicitly account for the fact that the location of the clinic in space may also drive its HIV prevalence level. For example, we might expect rural sites to be more correlated than urban sites. Thus, to further adjust for this bias, we used an offset term that represents the difference in the prevalence among the national, both-sexes population and the prevalence among the female, pregnant population associated with an ANC site location. The offset term was derived for each location as the difference between the adjusted prevalence in a given site-year and the adjusted national prevalence in that year. These estimates are adjusted for covariates that are thought to influence prevalence, for example, access to health-care facilities, malaria incidence, and male circumcision.

Thus, our final strategy for estimating the likelihood of the observed ANC data was:

$$W_{st} = \varphi^{-1}(\rho_t) + \vartheta_{st} + u_s + e_{st}$$

$$e_{st} \sim N(0, \sigma_{st}^2)$$

$$u_s \sim N(0, \sigma_s^2)$$

Where:

W_{st} = the probit transformed prevalence among ANC-attending pregnant women at site s and time t

ρ_t = The national prevalence adjusted to represent prevalence among pregnant women from the model simulation

ϑ_{st} = The offset term representing the difference between the adjusted prevalence in a given site-year and the adjusted national prevalence in that year

φ^{-1} = probit transformation

e_{st} = Site-year-specific error term

u_s = Site-specific intercept

Spectrum

For GBD 2013, we created an exact replica of Spectrum in Python. This enabled us to run thousands of iterations of the model at once on our computing cluster and allowed for more flexible input data structures. Additionally, we scaled all input values by a uniformly sampled factor between 0.9 and 1.1 to generate estimates with realistic ranges of uncertainty. For example, if treatment retention rates across CD4 categories were 0.906, 0.759, 0.787, 0.795, 0.785, 0.756, 0.813, and 0.700, we multiplied each number by an array of equivalent size that contained factors ranging from 0.9 to 1.1. At each draw, the array would contain different, randomly selected factors in the same range. Further, we previously improved our sex-specific modelling strategy in Spectrum by sex-splitting incidence based on a model fit to the sex ratio of prevalence observed in countries with representative surveys and updated the Spectrum paediatric module to reflect changes made by UNAIDS.¹⁰ Our child module was revised to include CD4 progression and CD4-specific mortality rates taken from a model fit to survival data from leDEA. Finally, we updated child initiation of ART to include data on ART distribution from leDEA. These changes were retained in GBD 2020.

ART coverage distribution

Spectrum determines the number of people initiating ART treatment across each CD4 category based on eligibility criteria, and the number of expected deaths and untreated people. In other words, groups with a large proportion of people living with HIV and high numbers of expected deaths initiated the most individuals into treatment.

We improved the basis for this distribution using survey microdata and country-level wealth information. Three relevant surveys were available: Uganda AIS 2011 and Kenya AIS 2007 and 2012. These surveys conducted CD4 count measurements and include a question regarding the amount of time that an individual receiving ART had been enrolled in treatment. Survey data provide cross-sectional CD4 count information; however, the Spectrum modelling framework tracks individuals by categorical CD4 count at the initiation of treatment. In order to cross-walk the cross-sectional survey data into estimates of CD4 count at treatment initiation, we built a model using relevant cohort data which tracked changes in CD4 count after initiation of treatment to translate an individual's current CD4 count and duration on treatment into CD4 count at initiation of treatment. The functional form for changes in CD4 count as a function of duration on treatment was a natural spline on duration with knots at 3, 12, 24, and 36 months, and an interaction between initial CD4 count and duration.

After cross-walking, we predicted the probability of being on treatment as a function of individual income (measured through an asset-based index), stratified by CD4 count, age, and sex. The results of this prediction were translated into country-specific age-sex-year-CD4 count probabilities of coverage using a conversion factor between individual income and lag-distributed GDP per capita. We used stochastic frontier analysis to constrain the maximum possible coverage for a given degree of income and CD4 count.

Predicted probabilities of coverage were input to Spectrum to inform the distribution, and not the overall level, of ART treatment by CD4 count. Within Spectrum, the probabilities of coverage are converted to counts of expected individuals on treatment in each CD4 count group. These are scaled to the distribution across CD4 count groups to match the input data on the number of people on ART coming from UNAIDS country files. In cases where the predicted number of individuals initiating treatment exceeds the total number of untreated individuals in a CD4 count group, we reallocate treatment evenly to other CD4 count groups.

Countries with seroprevalence surveys and antenatal clinic data (Groups 1ABC)

53 countries – as well as subnational locations in India, Kenya, Ethiopia, Nigeria and South Africa – were included in group one with available antenatal care clinic (ANC) data and/or least one geographically representative HIV seroprevalence survey. For all these locations we used EPP-ASM, which was updated to incorporate the new ANC bias adjustment.

In EPP-ASM, the transmission rate, $r(t)$, is a simple transmission model applied at each time step (1/10 of a year) to the population. 'r' represents the number of new cases expected to emanate from a single case. Over 3000 iterations, a new $r(t)$ is drawn, the full epidemic is determined and compared to the observed prevalence data to determine its likelihood. Beyond the end of the data, a prior distribution on $r(t)$ helps to determine how we should expect the epidemic to behave. This assumption was different in EPP-ASM versus EPP. In EPP-ASM in most countries we extended a random walk into the future based on the 'r-hybrid' $r(t)$. The r-hybrid assumes a logistic decay until the year 2003, a linear interpolation until year 2008 and a random walk form after this.

Changes for GBD 2020

For India, Comoros, Sao Tome and Principe and Mauritania (new Group 1B countries), we used EPP-ASM to model HIV burden for GBD 2020. For India, we used EPPASMEPP-ASM in combination with Spectrum, to be able to capitalize on SRS data. The SRS data was used to inform age and sex distribution. In addition, we used an 'equilibrium prior,' for $r(t)$ rather than 'r-hybrid' for India, which provided a better fit to the comparatively lower magnitude of the epidemic.¹² The equilibrium prior extends into the future with a rate of change following a normal distribution with a mean equal to the value of r expected when the proportion of the population infected is saturated, ie, the epidemic has stabilised.

When age-, sex-specific prevalence data included a zero proportion (no observed positive HIV tests), a binomial likelihood was used in place of the normal likelihood. Prior, imputation of a half positive observation was used to allow for probit transformation. This improved the fit to the zero proportion data while minimally impacting fits in non-zero prevalence age-, sex- strata. South Africa, India, Kenya, Gambia, Niger, Burundi, Ethiopia, Rwanda, Ghana, Sao Tome and Principe, Senegal, and Sierra Leone were affected by this change.

The HIV/mortality reckoning process is intended as a method of reconciling separate estimates of HIV mortality (and its resulting effect on estimates of HIV-free and all-cause mortality) in Group 1 countries by averaging estimates of HIV mortality from the model life table process and EPP-ASM. Additional details on the reckoning can be found elsewhere.¹³

Since EPP-ASM produces HIV incidence, prevalence, and deaths that are consistent with one another over time, the reckoning process results in death numbers that are no longer consistent with the incidence and prevalence produced in Spectrum. In order to recreate this consistency, we recalculated incidence for all Group 1 locations using reckoned deaths and prevalence produced by EPP-ASM. The updated incidence is calculated by aggregating counts of new infections, HIV deaths from EPP-ASM, and HIV deaths after reckoning at the year-sex level. The difference between reckoned HIV deaths and HIV deaths from EPP-ASM is added to EPP-ASM incidence, and we calculate the ratio between updated incidence and EPP-ASM incidence. Age-specific counts of new infections are then scaled by their corresponding sex-year ratios.

Countries with vital registration data (Group 2A and 2B)

Vital registration is one of the highest-quality sources of data on HIV burden in many countries, so generating estimates that are consistent with these data with necessary adjustment to account for any potential underreporting is critical. We identified 121 countries – as well as 760 subnational locations from China, Japan, Indonesia, India, Mexico, Sweden, Philippines, Poland, Italy, the UK, Ukraine, Russia, New Zealand, Iran, Norway and the USA – with usable points of vital registration data, verbal autopsy (VA) data, or sample registration system (SRS) data. In India, Vietnam, and Indonesia, we used SRS and VA data, respectively, as input mortality for CIBA. For India, we extracted the CIBA derived age-sex distribution of incidence but scaled the level to match the adult incidence rate estimated from EPP for each state.

We imputed missing years of data to generate a complete time series for HIV from the estimated start year of the epidemic using ST-GPR. We analysed mortality trends using ST-GPR starting in 1981, the year that HIV was first identified in the USA.¹³ For ST-GPR, we adjusted the lambda (time weight) and GPR scale according to the completeness of vital registration data, with 4- and 5-star quality VR using parameters designed to follow the data more closely. Separately by country/age group, we fit a piecewise linear spline with a single knot located at the empirical peak year of death rate using robust regression. The model includes fixed effects on region, age, and sex. Following this, we ran space-time residual smoothing, in which time, age, and space weights are used to inform smoothing of the residuals between datapoints and the linear regression estimate. From this process, we generated space-time estimates with the applied weights, along with the median absolute deviation (MAD) of the space-time estimates from the data. The MAD was calculated at various levels of the geographical hierarchy (eg, subnational and national), and was added into the data variance term. The data variance and space-time estimates were then analysed using Gaussian process regression to return a final estimate of mortality along with uncertainty.

Although Spectrum produces HIV mortality estimates that are within the realm of possibility in most countries using the incidence curves provided in the UNAIDS country files, it is a deterministic model that has not yet been integrated into an optimisable framework. Therefore, in order to “fit” it to vital registration data, we need to adjust input incidence. In contrast to GBD 2019 and previous cycles, in addition to adjusting input incidence, we determined the most plausible best treatment input based on fit to vital registration as well.

Changes for GBD 2020

For GBD 2020, we then created a grid of incidence and treatment options and reran Spectrum for each using each of these options, rather than using the CIBA-adjusted incidence for our final Spectrum run in all locations. The incidence options included the CIBA-adjusted incidence and the non-CIBA adjusted incidence from the initial Spectrum run, both using the most recent data and the last cycle, in addition to incidence data available from public use UNAIDS files. The CIBA process is described in more detail in the GBD 2019 write up.⁶ The adult ART options included the data available from public use UNAIDS files. Where these data were provided in terms of the number of people on treatment, we created additional treatment options by dividing the number on treatment by prevalence, as estimated by the current and previous GBD cycles. We ran Spectrum on every combination of incidence and treatment options, and then determined the root mean squared error of the resulting mortality relative to the vital registration data.

Finally, to produce location-, year-, age-, and sex-specific estimates of HIV incidence, prevalence, and mortality, we ran a final spectrum run using the incidence and treatment option that resulted in the best fit to VR data, or, the lowest RMSE.

Adjustment to case notification data

To estimate final incidence, we scaled the Spectrum output incidence up to the level of observed case notifications data, with a five-year lag to account for the difference between infection and detection. This was done in countries with 4 and 5 start vital registration systems, with available case notification data.

Countries without prevalence or deaths data (Group 2C)

32 countries had neither geographically representative seroprevalence surveys nor reliable vital registration systems. To produce estimates of HIV burden in these countries, we used Spectrum to produce estimates of burden. As above, the estimates of incidence, prevalence, and mortality were incorporated into the rest of the machinery via the reckoning process.

Changes for GBD 2020

Group 2C countries no longer sampled bias adjustment ratios from other Group 2 countries within the same super-region.

References

- 1 Stover J, Glaubius R, Mofenson L, *et al.* Updates to the Spectrum/AIM model for estimating key HIV indicators at national and subnational levels. *AIDS Lond Engl* 2019; **33**: S227–34.
- 2 Eaton JW, Brown T, Puckett R, *et al.* The Estimation and Projection Package Age-Sex Model and the r-hybrid model: new tools for estimating HIV incidence trends in sub-Saharan Africa. *AIDS Lond Engl* 2019; **33 Suppl 3**: S235–44.
- 3 Dwyer-Lindgren L, Cork MA, Sligar A, *et al.* Mapping HIV prevalence in sub-Saharan Africa between 2000 and 2017. *Nature* 2019; **570**: 189–93.
- 4 Kyu HH, Jahagirdar D, Cunningham M, *et al.* Accounting for misclassified and unknown cause of death data in vital registration systems for estimating trends in HIV mortality. *J Int AIDS Soc* 2021; **24**: e25791.
- 5 Murray CJL, Ortblad KF, Guinovart C, *et al.* Global, regional, and national incidence and mortality for HIV, tuberculosis, and malaria during 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet* 2014; **384**: 1005–70.
- 6 Vos T, Lim SS, Abbafati C, *et al.* Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet* 2020; **396**: 1204–22.
- 7 Zheng P, Barber R, Sorensen RJD, Murray CJL, Aravkin AY. Trimmed Constrained Mixed Effects Models: Formulations and Algorithms. *J Comput Graph Stat* 2021; **0**: 1–13.
- 8 Ghys PD, Zaba B, Prins M. Survival and mortality of people infected with HIV in low and middle income countries: results from the extended ALPHA network. *AIDS Lond Engl* 2007; **21 Suppl 6**: S1-4.
- 9 Hallett TB, Zaba B, Todd J, *et al.* Estimating incidence from prevalence in generalised HIV epidemics: methods and validation. *PLoS Med* 2008; **5**: e80.

- 10 MAHY M, PENAZZATO M, CIARANELLO A, *et al.* Improving estimates of children living with HIV from the Spectrum AIDS Impact Model. *AIDS Lond Engl* 2017; **31**: S13–22.
- 11 Ng M, Gakidou E, Murray CJ, Lim SS. A comparison of missing data procedures for addressing selection bias in HIV sentinel surveillance data. *Popul Health Metr* 2013; **11**: 12.
- 12 Eaton JW, Brown T, Puckett R, *et al.* The Estimation and Projection Package Age-Sex Model and the r-hybrid model: new tools for estimating HIV incidence trends in sub-Saharan Africa. *AIDS Lond Engl* 2019; **33**: S235–44.
- 13 Pneumocystis Pneumonia --- Los Angeles.
https://www.cdc.gov/mmwr/preview/mmwrhtml/june_5.htm (accessed Feb 4, 2021).