CSAVR technical review and model development

Report and recommendations from a meeting of the UNAIDS Reference Group on Estimates, Modelling, and Projections
5th June 2020

REPORT & RECOMMENDATIONS

UNAIDS
JOINT UNITED NATIONS PROGRAMME ON HIV/AIDS
UNHCR UNICEF WFP UNDP UNFPA
UNODC ILO UNESCO WHO WORLD BANK

Abbreviations

ADR	AIDS Data Repository
AIC	Akaike information criterion
AP	Asia/Pacific
CAR	Caribbean
CSAVR	Case Surveillance and Vital Registration model
EECA	Eastern Europe and Central Asia
EPP	Estimation and Projection Package
FSW	Female sex workers
FWID	Women who inject drugs
IHME	Institute of Health Metrics and Evaluation
LA	Latin America
MENA	Middle East and North Africa
MSM	Men who have sex with men
MWID	Men who inject drugs
PLHIV	People who live with HIV
TESSy	The European Surveillance System
UNAIDS	Joint United Nations Programme on HIV/AIDS
WCENA	Western/Central Europe and North America

The meeting of the UNAIDS Reference Group on Estimates, Modelling, and Projections was organised for UNAIDS by the Secretariat of the Reference Group (www.epidem.org), managed at Imperial College London, the University of Cape Town, and Stanford University. Participants of the meeting are listed at the end of this document.

Oli Stevens, June 2020

Background

UNAIDS Reference Group on Estimates, Modelling, and Projections

The Joint United Nations Programme on HIV/AIDS (UNAIDS) relies on impartial scientific advice from international experts in relevant subject areas to provide guidance on how to best calculate estimates and projections of the prevalence, incidence, and impact of HIV/AIDS globally. The UNAIDS Reference Group on Estimates, Modelling, and Projections acts as an 'open cohort' of epidemiologists, demographers, statisticians, and public health experts to provide scientific guidance to UNAIDS and partner organisations on the development and use of the tools used by countries to generate annual HIV estimates, which are the source for UNAIDS Global HIV epidemic estimates. The group is coordinated by a secretariat hosted at Imperial College London, the University of Cape Town, and Stanford University.

Meeting Overview

The UNAIDS Reference Group held its virtual technical meeting on *CSAVR technical review and model development* on 5th June 2020. The meeting featured presentations and group discussion to generate consensus recommendations. The programme was divided into the following sessions:

1. CSAVR in the 2020 estimates process
2. Technical review
3. Data inputs and user interface
4. Future model development priorities

Due to time constraints, Session 4 *Future model development priorities*, was omitted from the agenda during the meeting.

This report presents a summary of the meeting presentations and discussions. The presentations are available to meeting participants at www.epidem.org (others, please contact the Secretariat via epidem@imperial.ac.uk). The final recommendations can be found at the end of this report.

The recommendations drafted at these meetings provide UNAIDS with guidance on generating HIV estimates, provide an opportunity to review current approaches, and help to identify the data needed to further improve the estimates. Previous meeting reports are available at www.epidem.org. This transparent process aims to allow the statistics and reports published by UNAIDS and partners to be informed by impartial, scientific peer-review.

The list of participants and meeting agenda are included in Appendix I and Appendix II, respectively.

Meeting objectives

The Case Surveillance and Vital Registration model (CSAVR) is one of three UNAIDS-supported incidence estimation models, alongside the Estimation and Projection Package (EPP) and the AIDS Epidemic Model. CSAVR fits HIV incidence to new diagnosis case surveillance and mortality data, and has been used widely in low HIV prevalence settings.

The objectives of this meeting were to:

1) Review CSAVR's 2020 parametrisation and implementation and validate model performance,
2) Reach recommendations on key population-stratified estimation, capturing uncertainty within input data, and calibrating to CD4 count at diagnosis,
3) Review quality and sources of case surveillance and mortality data.

Session 1: CSAVR in 2020 estimates

Kim Marsh presented an overview of CSAVR and its use in the 2020 UNAIDS estimates round. New HIV diagnoses and estimates of HIV-related deaths for ages 15+ are used as data inputs into the model. HIV diagnosis data must be deduplicated, and HIV diagnoses among migrants where transmission occurred outside the country of interest may be excluded. Countries are recommended to use estimates of HIV-related deaths that account for reclassification of likely AIDS deaths miscoded as other causes or unidentified cause of death ('garbage codes'), such as those produced by the Global Burden of Disease Study from the Institute of Health Metrics and Evaluation (IHME) (see presentation from Hmwe Kyu in Session 2 below).

Forty-three countries used CSAVR in the 2020 UNAIDS estimates process, accounting for 6% of global PLHIV and 13% of adult new infections. Aggregating mortality estimates over all 43 countries produced good alignment with both WHO 2016 and IHME Global Burden of Disease 2017 estimates of AIDS mortality in these countries.

Ongoing challenges include:

- How to account for changes to case definitions and testing coverages in a country, leading to jumps in the case notification data time trends;
- Implausibly high estimates of knowledge of HIV status early in the epidemic when HIV testing rates were known to be low
- Modelled estimates of CD4 count at diagnosis are systematically higher than the observed distribution of CD4 at diagnosis from case report data,

particularly in early years of the epidemic when testing coverage was low and diagnoses were largely comprised of symptomatic patients with low CD4 counts

Three points were raised in discussion:

- Model outputs for HIV incidence often decrease rapidly after the end of the data (e.g. Cuba and Chile). Short term projections should be revisited, and additional model constraints considered.
- The case definition and time of recording cases varies across countries. In some settings, cases are record at diagnosis, while in other cases are recorded upon linkage to care. Cases reported at diagnosis may be anonymous and more susceptible to duplication.
- Different subpopulations are expected to have different CD4 count at diagnosis, which is currently not captured by the model.

Session 2: Technical review

Guy Mahiane presented a detailed review of CSAVR including:

- Technical description of the 2020 model implementation
- Model validation
- Proposed methods developments

2.1. Technical description

CSAVR implements four functional forms for incidence trends. These options are the single logistic (incidence increasing), double logistic (incidence formerly increasing, now plateauing/decreasing), spline (flexible incidence), and r-logistic (incidence proportional to prevalence and ART coverage, akin to EPP). Models were compared by fitting all four models to datasets for each country, and in each country the model with the lowest AIC value was identified. In all regions except for WCENA, the r-logistic model had the lowest AIC value for the largest number of countries, followed by the double logistic model (Fig 1),

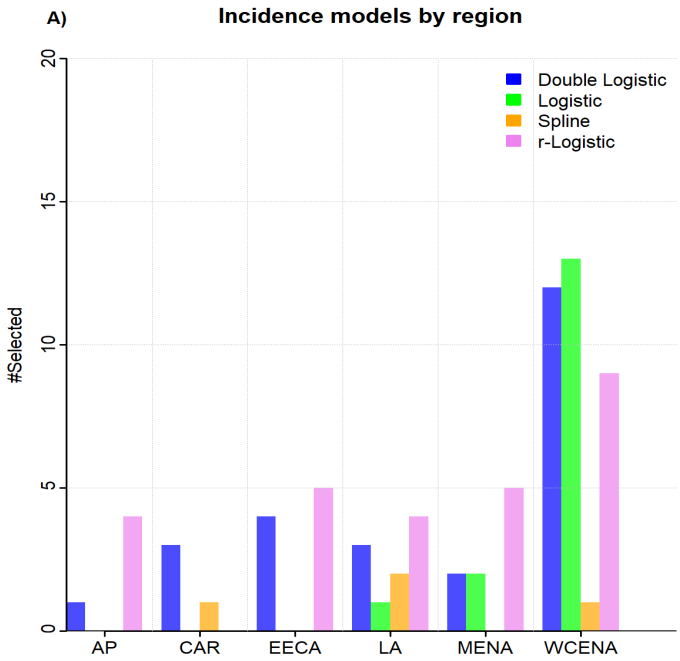

Figure 1: Preferred model by region, as chosen by AIC score

The 2020 revision of CSAVR incorporated a modified version of the HIV diagnosis model implemented in Shiny90 (Maheu-Giroux et al., 2019). The rate of HIV diagnosis is modelled as a function of time, age, sex, and CD4 category. A parametric function is used to represent the change in the rate of HIV diagnosis over time. The diagnosis rate is a mixture of two components which capture the magnitude and rate of the diagnosis trend, and the timing of the inflection point. This provides a flexible form that permits a wide variety of testing trends over time.

	WCENA	**EECA**	**MENA**	**LA**	**AP**	**CAR**
New diagnoses						
Availability over time	Increasing					
Disaggregated by age	Yes	Yes	Limited	Limited	Limited	Limited
Disaggregated by sex	~80% of countries	~80%	~30%	~20%	~40%	No
AIDS deaths						
Availability over time	Constant 1990-2019					
Disaggregated by sex	~90%	~60%	~65%	~60%	~100%	~25%

Table 1: Availability of new diagnosis and HIV-related death data by age

An inverse Gaussian distribution was specified as a likelihood for the number of new diagnoses and AIDS deaths, summed across all ages and both sexes. These are then disaggregated by sex (binomial distribution) and by age (multinomial distribution). This distribution, $IG(\mu, \mu^{\zeta})$, has a heuristically derived variance, with $\zeta =$

$-3/4$. A formal justification should be developed for this distribution and choice of variance specification.

The model produces outputs for:

- New HIV diagnoses by 5 year age groups, sex, and CD4 category,
- AIDS deaths by 5 year age groups and sex,
- Proportion of PLHIV aware of their status by 5 year age groups, sex, and CD4 category,
- Mean CD4 at diagnosis by 5 year age groups and sex.

Estimates of incidence 1990-present were stable and low in WCENA, AP, and MENA, increasing in LA and EECA, and decreasing in CAR having peaked in the 1990s (Fig 2). Mean CD4 count at diagnosis was broadly stable, with most countries >400 cells/ml and higher in women than in men (Fig 3a). Knowledge of HIV status has seen large increases over the estimation period, though the timing of increase varies between region (Fig 3b). Women are more aware of their status in all regions, particularly in MENA and LA. High estimates of knowledge of status early in the epidemic are implausible, particularly those years before the development of an HIV diagnostic test.

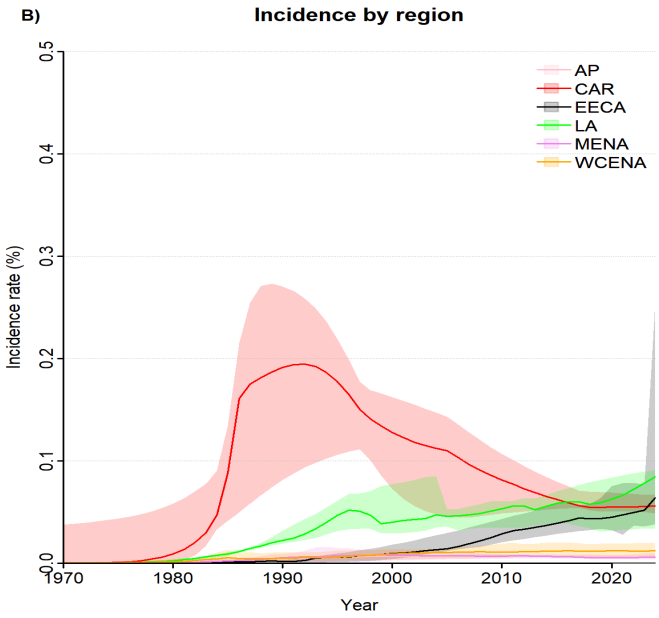

Figure 2: HIV incidence by UNAIDS region

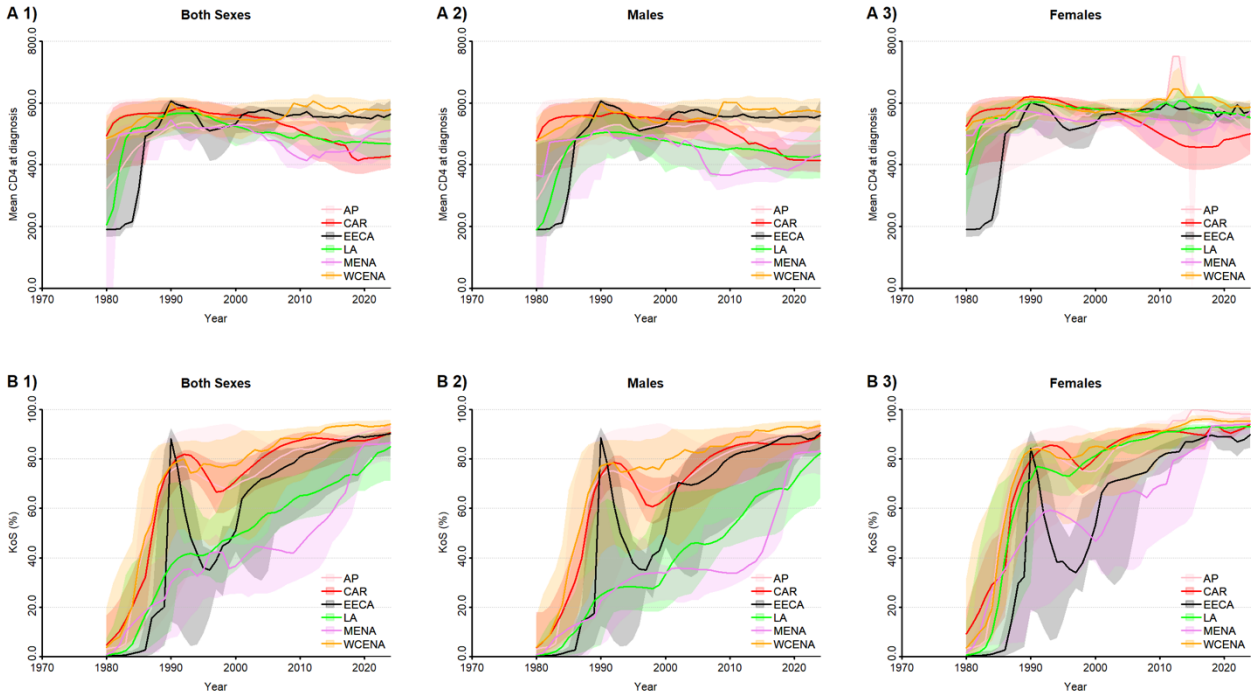

Figure 3: CD4 at diagnosis (top) and knowledge of HIV status (bottom), by UNAIDS region

It was noted that the mean CD4 at diagnosis in WCENA is higher than observed in the surveillance data from TESSy. The quality and completeness of CD4 count at diagnosis data is further addressed below.

2.2. Model performance and validation

Two analyses were undertaken to evaluate model performance:

1. Out-of-sample validation: To assess the predictive accuracy of the model with observed data, the most recent three years of data were withheld from model fitting and used as test data for model predictions.
2. Simulation study: Simulated datasets were generated from the model with configurations similar to existing data to which the model was fit.

Out-of-sample validation

Mahiane assessed projection performance by:

- Excluding the last three years of new HIV diagnoses and/or HIV-related deaths
- Fitting all four incidence options with uncertainty analysis

- Quantifying the forecast performance of the model with the continuous ranked probability score (CRPS)
- Estimating the proportion of data points that fall within the 95% confidence intervals, denoted “coverage”.

Akin to fitting CSAVR to all available data, the r-logistic remained the preferred model, followed by double logistic when omitting the final three years’ data. In the out-of-sample validation analysis, the double logistic model had the lowest CRPS (indicating best prediction) in 36% of countries, followed by the spline model (29%), and the logistic model (18%).

Detailed results were presented for Barbados, Bulgaria, and Chile. In Barbados, all four models produced similar estimates and projections for the number of new diagnoses and were consistent with the withheld observations. Data were not stratified by sex. In Bulgaria and Chile, CSAVR struggled to capture observed trends in the sex-specific model and the withheld data points fell outside the 95% prediction intervals. Mahiane hypothesised that model estimates for women may be worse than for men due to lower number of cases reported and should be investigated further.

Out-of-sample coverage for the 71 test countries was calculated by region for each incidence model. Between 45-80% of omitted points lay within the 95% prediction interval, substantially lower than the target coverage of 95% (Fig 4). CRPS and AIC often identify different models as the best fitting model, raising concerns regarding the use of AIC in model selection. The reported coverage estimates were for both withheld data about new diagnoses and AIDS deaths. It was suggested to review out-of-sample predictive coverage for stratified by data type to assess whether goodness-of-fit differed by data type.

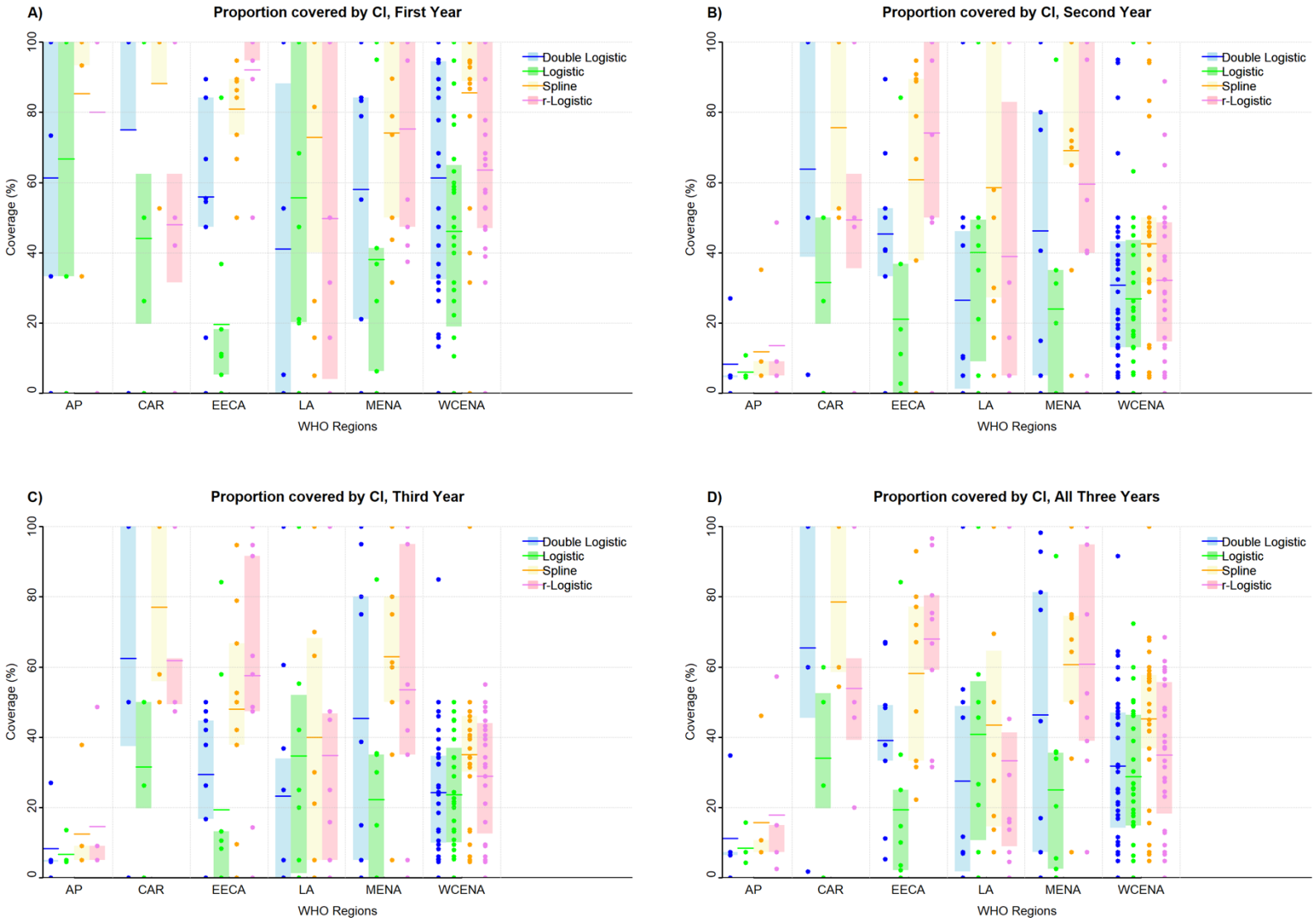

Figure 4: Regional and incidence model estimates of out-of-sample validation coverage by each year of omitted data (A-C), and for all three years' omitted data (D)

Model fitting to simulated data

In many countries using CSAVR, available data about HIV diagnoses and AIDS deaths are fragmentary and only available several years after the start of the HIV epidemic. Model fits to simulated data assessed whether the CSAVR models for HIV incidence and diagnosis rate are able to reliably reconstruct historical HIV incidence trends from available input data. For each country, fitted model results were used to generate simulated observations with the same configuration as the observed data (that is with the same years and stratifications as reported data). Simulated observations for HIV diagnoses and AIDS deaths were simulated from a Poisson distribution with equal mean and variance, differing from the $IG(\mu, \mu^{\zeta})$ distribution used in regular model fitting. Future simulation studies could consider the use of the negative binomial to address overdispersed data.

Model fits to simulated data reliably recovered the true epidemic well with a low mean absolute error during the period in which data were available. Reconstruction of epidemic trends during the early epidemic period before data were available were generally robust, but the percentage absolute error was slightly larger than during the data period. During the projection period, the spline model had a larger mean percentage error than other models (Fig 5).

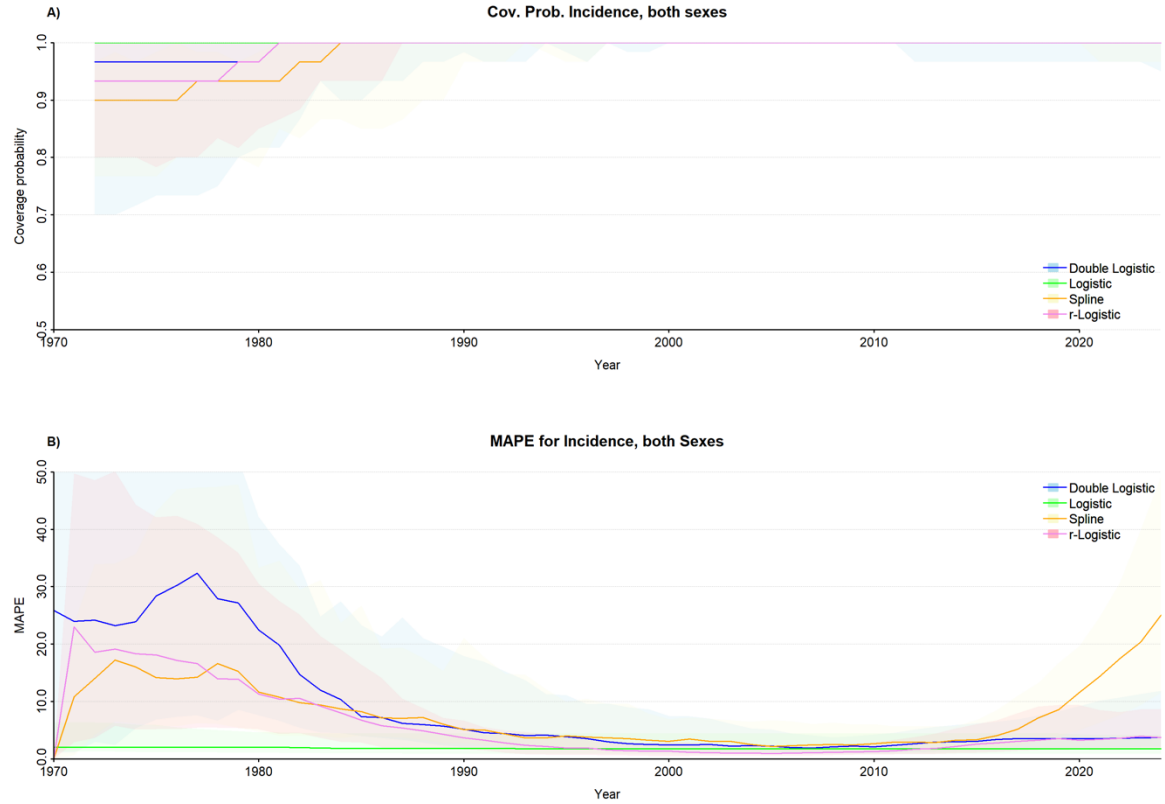

Figure 5: In-sample coverage probability (B) and mean absolute percentage error (B) for HIV incidence in Bulgaria

2.3. Future model development

Key population-stratified estimation

CSAVR is used in countries with low prevalence, concentrated epidemics. The present implementation of CSAVR does not produce estimates disaggregated by key population. New diagnoses data stratified by key population are available from the ECDC TESSy surveillance database for men who have sex with men (MSM), men who injected drugs (MWID), and women who inject drugs (FWID). Female sex worker-stratified data are absent from the TESSy database, but included within the proposed key population-stratified CSAVR. Key population stratification is not available for AIDS deaths data in any setting.

Key population stratification for MSM, MWID, FWID, and FSW was proposed to be incorporated into CSAVR using the follow assumptions:

1. The size of each key population is specified as a fixed proportion of the total population for each sex. This proportion is constant over time. For FSW and PWID populations, population 'turnover' is applied to the entire population group at a fixed rate over time.
2. The HIV incidence rate among key population groups is modelled relative to the general population incidence for the same sex (incidence in MSM and MWID is relative to male general population incidence, and incidence in FWID and FSW is relative to female general population incidence). The incidence rate ratio for each key population group is estimated relative to the general population and varies over time modelled by a logistic function.
3. The diagnosis rate for each key population is proportional to diagnosis rate the general population of the same sex. The relative diagnosis rate is estimated for each key population and is constant over time.

Example results were presented for Bulgaria (Fig 6). Though the incidence rate was estimated to be higher in key populations than in the general population, the majority of new infections occur outside key populations.

It was noted in discussion that substantial misclassification of source of infection occurs within case surveillance systems, and the true number of new diagnoses within key populations is likely significantly higher than is recorded in many settings. This likely varies across settings. It was recommended to review existing research literature to guide potential adjustments for misclassified source of infection.

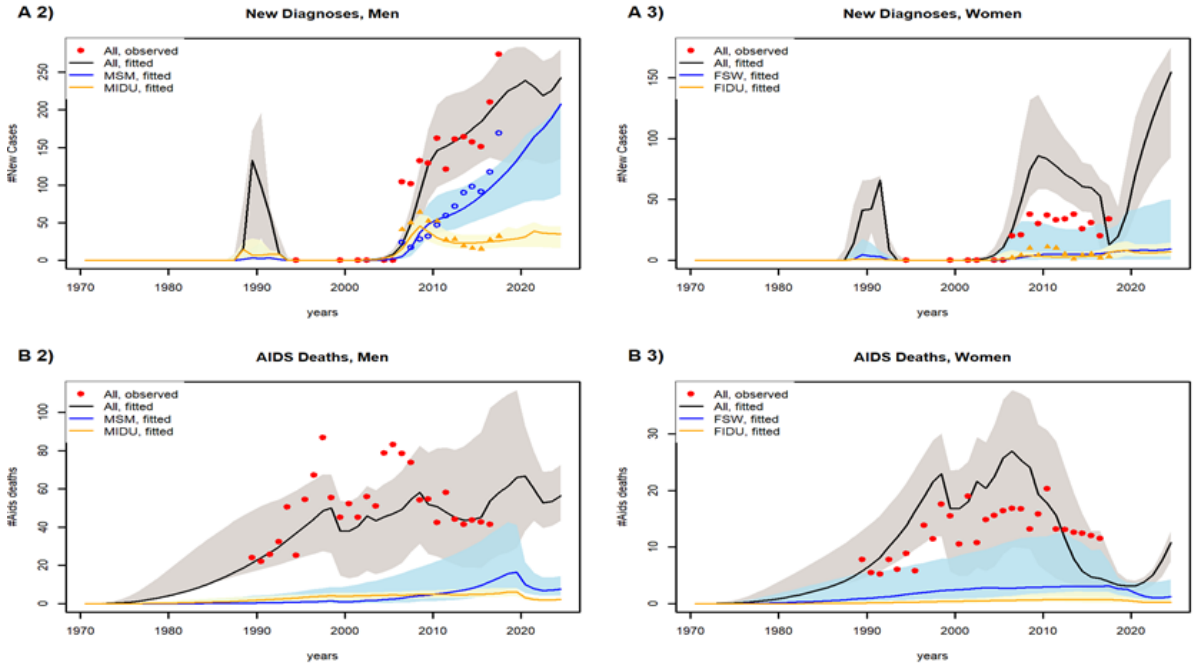

Figure 6: Key population and sex stratified estimates of new diagnoses (A) and AIDS deaths (B) fit to surveillance data from Bulgaria.

Calibrating to CD4 count data

In versions of CSAVR used for previous UNAIDS estimates, CD4 count at diagnosis data were used as calibration data. Simulation studies showed that these data strongly influenced estimated HIV incidence trends and awareness of HV status.

The current version of CSAVR removed CD4 count at diagnosis because the model often struggled to reproduce these data and there were concerns that poor data quality, completeness, and selective use could lead to biased estimates. CD4 count at diagnosis is missing for large proportions of cases in the ECDC TESSy database (around 30% in recent years, and upwards of 75% 2000-2005), and it is not clear how biased these data may be or how best to impute them. A tool developed by the ECDC assumes the data are missing at random, and additional consideration with country teams should be given to why data may instead be systematically missing.

It was noted that another factor contributing to poor fit to observed data about CD4 count at seroconversion could be the parameters for the natural history model in Spectrum. CSAVR uses the same natural history progression parameters as those in Spectrum/AIM. The Spectrum/AIM natural history model is currently under revision (UNAIDS Reference Group on Estimates Modelling and Projections, 2020). The updated parameters reflect a lower mean CD4 count at seroconversion than the

current defaults, which may improve the consistency of CSAVR estimates of CD4 count at diagnosis compared to TESSy case report data.

Session 3: Data inputs and user interface

Kim Marsh presented a summary of the completeness and quality of case surveillance data, and key challenges for data interpretation.

- The HIV case definition has evolved over time in many countries. Though the majority of countries introduced case surveillance of HIV diagnoses and AIDS cases contemporaneously or with minimal delay, some countries have large intervals (e.g. Italy introduced HIV reporting 20 years after AIDS case reporting, whilst the Netherlands introduced AIDS case reporting 17 years after HIV diagnosis reporting).
- Underreporting remains an issue including in some high-income countries with strong health systems. For example, France is estimated to have underreported HIV diagnoses by a third in 2017.
- National capacity to deduplicate case reports and exclude previously diagnosed PLHIV is key to high quality CSAVR estimates – capacity exists in EECA and LA (15 of 26 countries), but is limited in MENA.
- There are often reporting delays such that official case numbers change retrospectively for the two to three most recent years. Adjusting for reporting delays in Portugal increased new HIV infections by ~30% in recent years. Excluding the most recent year from estimation may help to minimise bias and is common practice in some CSAVR applications. More formal guidance is needed.
- Of countries that have HIV as a notifiable condition, a third do not have surveillance systems in place for subsequent sentinel events including CD4 count, linkage to care, or HIV-related mortality.

Proposed solutions to these issues could include:

- The development of metrics to flag data quality issues. for example when the ratio of HIV diagnoses to deaths exceeds 10:1.
- A systematic approach to adjusting for reporting delays or excluding recent data.
- Improved documentation and metadata to accompany case notification data,

For AIDS deaths data inputs to CSAVR, it is recommended to use estimates of the total AIDS deaths, adjusted for reclassification of AIDS deaths miscoded as other causes or as unknown causes (‘garbage codes’) (UNAIDS Reference Group on

Estimates Modelling and Projections, 2018). Hmwe Kyu described the methods for the annual AIDS death estimates produced by the IHME Global Burden of Disease Study based on death registration data reported by countries to the WHO Mortality Database. Deaths assigned 'garbage codes', those deaths which are assigned to intermediate or immediate causes of death rather than the underlying cause of death, are recoded to HIV/AIDS. The age pattern of mortality of a given cause of death is compared to the age pattern in the pre-HIV era (1980-1984), and any excess mortality above 5% is recoded to HIV/AIDS. Following garbage recoding, the data are smoothed, taking a Bayesian average between the prior estimate of deaths by age and year and the observed deaths. An online data visualisation of the IHME cause of death data is available here.

It is challenging for national estimates teams to understand the process by which the recoded mortality estimates by age and sex are produced. Expanding the visualisation tools produced by IHME to include the Sankey diagrams showing the causes from which AIDS deaths have been recoded disaggregated by age and sex would help users to understand the model inputs.

Jonathan Berry presented the AIDS Data Repository (ADR), currently used by national estimates teams in sub-Saharan Africa to manage, store, and validate HIV programme data. ADR does not currently support data management needs for CSAVR. Berry outlined the potential use-cases of a repository system to manage and extract case surveillance and mortality input data. While health information systems in sub-Saharan Africa (e.g. DHIS2) exist to manage programmatic data, data sources for case surveillance data in CSAVR countries differ, and close collaboration with country teams is required to understand what data management solutions would be most useful.

Proposed updates to the CSAVR interface in Spectrum were presented by Rob Glaubius:

- Divide the model fitting display into three tabs
    - Model fitting, where indicators included in fitting are displayed,
    - Model comparison, where fits with all incidence model options are overlaid with data by indicator, and
    - Model validation, where additional model outputs, including mean CD4 at diagnosis and knowledge of status are displayed.
- Simplify the interface by removing unused elements.
- Visualise comparison to previous year fits similar to those offered in EPP.

All proposed updates to the interface were recommended for implementation. UNAIDS will organize a review session with country estimates teams to solicit user feedback on proposed CSAVR user interface developments.

Recommendations

Technical review

- **Model selection and goodness-of-fit**
  - The use of AIC for selection of CSAVR model variants should be reviewed.
  - Metrics should be considered for overall model goodness-of-fit.
  - The short-term incidence projections of different model options should be systematically reviewed, particularly with respect to observed sharp projected incidence declines following the end of data availability.
  - More formal justification should be developed for the inverse Gaussian distribution for the likelihood function and choice of variance specification.
  - Out-of-sample coverage estimates should be reviewed separately for new diagnoses and AIDS deaths data. This may guide revisions to model specification to improve the low reported out-of-sample coverage of prediction intervals.
  - More explicit modelling the observation process and incorporating expert knowledge about data collection may improve estimates (for example changes in case definition, reporting delays, levels of AIDS death reclassification).

- **Inclusion of CD4 count data in CSAVR estimation**
  - Model fit to observed data about CD4 at diagnosis should be reviewed with revised natural history parameters in Spectrum.
  - The potential effects of missing data about CD4 at diagnosis and approaches to adjusting for missing observations should be reviewed

- **Knowledge of status**
  - Work with country teams to find testing numbers and positivity for the early epidemic
  - Consider estimating the fraction of individuals diagnosed when asymptomatic

- **Age/sex stratified estimates**
  - Existing model fit to sex and age stratified data should be more systematically reviewed, particularly noted poorer fits in women.
  - The usage and results of incidence rate ratio fitting to case surveillance and vital registration data should be further reviewed.

- **Key population stratified estimates**
    - The key population stratified model should be tested with data from additional countries.
    - Literature about misreporting of transmission risk factors in case report data, potential effect on key population stratified CSAVR estimates, and possible adjustments should be reviewed.
    - The Reference Group should further consider assumptions of the key population stratified model about population size, incidence patterns, and diagnosis rate.

Data quality and user interfaces

- Review estimation of HIV-related deaths in countries with IHME mortality classifications 2B and 2C
- A tool should be developed to assist countries in visualising and understanding adjusted AIDS death estimates, including Sankey diagrams showing redistribution of other causes of death to AIDS by age and sex.
- Metrics should be created to identify potentially inconsistent data inputs to model users.
- Systematic guidance should be developed for how to adjust or exclude recent case data to account for reporting delays.
- Improved metadata to accompany case surveillance data should be systematically captured, including when HIV case notification started, completeness of CD4 count reporting, estimates of reporting delays if available.
- UNAIDS should organise sessions to review proposed model user interface developments with country estimates teams.

References

Maheu-Giroux, M., Marsh, K., Doyle, C. M., Godin, A., Lanièce Delaunay, C., Johnson, L. F., … Eaton, J. W. (2019). National HIV testing and diagnosis coverage in sub-Saharan Africa. *AIDS*, *33*, S255–S269. https://doi.org/10.1097/QAD.0000000000002386

UNAIDS Reference Group on Estimates Modelling and Projections. (2018). *Technical updates and method development for the UNAIDS Estimates*. Retrieved from http://www.epidem.org/sites/default/files/reports/Report_Technical Updates_Bern_FINAL.pdf

UNAIDS Reference Group on Estimates Modelling and Projections. (2020). *Consolidating data and estimation tools for next-generation models*. Retrieved from http://www.epidem.org/sites/default/files/reports/UNAIDSRG_Spring2020_report_0.pdf

UNAIDS Reference Group on Estimates, Modelling and Projections

CSAVR technical review and model development

Friday 5th June 2020

All times are GMT+1 (London)

Time	Duration (mins)	Topic	Presenter(s)/ Lead Discussant
15.30	15	Welcome and introductions Meeting objectives CSAVR in 2020 UNAIDS estimates	Mary Mahy Jeff Eaton Kim Marsh
Session 1: Technical review (chaired by Leigh Johnson)			
15.45	30	Technical description of CSAVR 2020 model: • Model parameterisation and implementation • Inference methodology • Age/sex stratification	Guy Mahiane
16.15	25	Model validation • Out-of-sample validation for incidence models • Review "edge cases"	Guy Mahiane
16.40	25	Proposed methods developments • Key population stratification • Reporting uncertainty in diagnosis and AIDS death data • Calibration to CD4 count data	Guy Mahiane
17.05	45	Discussion	
17.50	10	Break	
Session 2: Data inputs and user interface (chaired by Josh Salomon)			
18.00	15	Quality and completeness of case surveillance data	Kim Marsh
18.15	15	IHME mortality estimates	Hmwe Kyu
18.30	10	AIDS Data Repository	Jonathan Berry
18.40	10	Interface development	Rob Glaubius
18.50	25	Discussion	
Session 3: Future model development (chaired by Jeff Eaton)			
19.15	30	Discussion: • Alignment with diagnosis indicators in Shiny90 • Integration of CSAVR and EPP for concentrated epidemic settings	
19.45	15	Discussion and recommendations	Jeff Eaton
20.00	CLOSE		