Technical updates for UNAIDS HIV estimation tools

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

11 – 13 October 2022

Virtual

REPORT & RECOMMENDATIONS

Index

Abbreviations

AIM	AIDS Impact Model
ANC-RT	Antenatal Clinic Routine Testing
ART	Antiretroviral Therapy
ART-CC	ART Cohort Collaboration
ASM	Age Structured Model
CDC	US Centres for Disease Control and Prevention
CSAVR	Case Surveillance and Vital Registration
DHIS	Division of Health Informatics and Surveillance
DQA	Data Quality Assessment
DTG	Dolutegravir
EPP	Estimation and Projection Package
FSW	Female Sex Worker
GAM	Global AIDS Monitoring
GBD	Global Burden of Disease
HIC	High Income Country
IBBS	Integrated Biological and Behavioural Surveillance Survey
IeDEA	International Epidemiology Databases to Evaluate AIDS
IRR	Incidence Rate Ratio
KOS	Knowledge of Status
KP	Key Population
LMIC	Low- and Middle-Income Countries
LTFU	Loss to Follow Up
MENA	Middle East/North Africa
MSM	Men who have Sex with Men
PEPFAR	President's Emergency Plan for AIDS Relief
PHIA	Population-based HIV Impact Assessment
PLHIV	People Living with HIV
PMTCT	Prevention of Mother to Child Transmission
PWID	People Who Inject Drugs
SSA	Sub-Saharan Africa
TGW	Transgender Women
UNAIDS	Joint United Nations Programme on HIV/AIDS
VLS	VL Suppression
WHO	World Health Organization
WPP	World Population Prospects

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 SACEMA, Imperial College London and the University of Cape Town. Participants of the meeting are listed at the end of this document (**Appendix B**).
Cari van Schalkwyk, June 2022

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 SACEMA, Imperial College London and the University of Cape Town.

Meeting Overview

The UNAIDS Reference Group held the meeting virtually on the Microsoft Teams platform, from 11 – 13 October 2022. The meeting featured presentations and group discussions to generate consensus recommendations, divided into the following 8 sessions:

This report presents a summary of the meeting presentations and discussions that underpin recommendations by the Reference Group. The presentations are available to meeting participants (**Appendix B**) at www.epidem.org (others, please contact the Secretariat via epidem@sun.ac.za). The final recommendations can be found at the end of this report. The recommendations (**Appendix A**) drafted at these meetings provide UNAIDS with guidance on generating HIV estimates, review current approaches, and identify required data to further improve HIV estimates. The meeting agenda and objectives are in Appendix C. 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.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.

Introduction

Mary Mahy opened the meeting by thanking participants for their commitment and emphasizing the importance of the Reference Group in guiding UNAIDS with best science about the assumptions and designs of the model used for the global HIV estimates.

Mahy raised five points:

1. Countries use the data generated by the models to assess their progress in their HIV response. What is especially important this year is that countries will be using it for their Global Fund applications.
2. These estimates are also important as a component of our global advocacy efforts and helps to ensure that HIV stays a high priority for countries.
3. One of the objectives of UNAIDS is to increase the availability of data to inform modelled estimates, recognizing that more country specific data will generate more accurate estimates for the setting. Mahy urged participants to keep this objective in mind and consider where stronger linkages between program data and the models can be considered.
4. UNAIDS plans to have in person estimates development country workshops for the upcoming 2023 estimates round, recognizing that these workshops and the teamwork that they inspire build the quality of the estimates that countries produce. Regional workshops will be conducted between December 2022 and early March 2023.
5. Mahy's final point recognises, similar to the country workshops, the value of in-person meetings for the Reference Group and she notes that the next meeting in May 2023 will be in-person, but that the option to join remotely will remain available.

Cari van Schalkwyk presented a brief overview of meeting objectives. In most sessions, the group will be presented with minor model changes or Spectrum interface updates that will be reviewed and recommended for implementation in the next estimates round.

Session1 - Incorporating WPP 2022 into Spectrum, EPP, and Naomi models: At the May 2022 meeting the Reference Group was presented with an overview of the major changes to the methods in the WPP 2022 estimates. These included switching to single year and single age estimates instead of 5x5 as in previous versions and switching to calendar years instead of mid-year intervals as in previous versions. The objective of this session is to review the impact of these changes to the estimates provided by the modelling tools (Spectrum, EPP, Naomi, CSAVR, Shiny90).

Session 2 - On-ART mortality: At the May 2022 meeting the Reference Group recommended to investigate the default on-ART mortality rates in high-income countries, which resulted in over-estimating AIDS deaths in several countries, and the on-ART mortality rates in Asia Pacific countries that resulted in underestimating AIDS deaths in several countries. Session 2 presentations will report back on those recommendations. The objectives of this session are:

1. Recommendations for updating Spectrum on-ART mortality in high income countries.
2. Recommendations for updating Spectrum on-ART mortality in SSA.

Another recommendation from the May 2022 meeting was to convene a working group to prepare suggestions for required data and model structure of on-ART mortality in the future. Leigh Johnson had a proposal for this, but the data required to test the proposal using the

Thembisa model were not available in time to complete analysis before the meeting. This proposal will be presented at the May 2023 Reference Group meeting.

Session 3 – Improving prevalence measures from ANC-DQAs: In May 2022 the Reference Group recommended to support countries to conduct data quality assessments (DQAs) of their routine ANC and PMTCT data before 2023 HIV estimates round and a working group was established to start this process. This session will provide updates on the progress of this working group.

Another recommendation was to develop an algorithm on weighting of ANC routine testing data in EPP estimation, but the working group concluded that there is not sufficient data available from existing DQAs to inform this algorithm. Regardless, the objective of this session is to provide guidance on the use of ANC-RT data in the next estimates round.

Session 4 - Test and treatment churn: In this session, presentations will cover the recommendations from May 2022, and objectives for the session are, for countries where number on ART is greater than PLHIV to:

1. Decide whether to update EPP model fitting to reject EPP fits where ART coverage is over 100%,
2. Make recommendations about proposals to incorporate survey ART coverage data (by age/sex) into the Spectrum incidence rate ratio fitting tool.

Session 5 - HIV+ Migration in AIM & Enhanced User guidance for concentrated epidemics: In this session the Reference Group will:

1. Review and make recommendations for proposed guidelines to account for migration of people with HIV.
2. Review user guidance for concentrated epidemics.

Session 6- Key populations in concentrated epidemics: The topic of this session will be on creating estimates for key populations in concentrated epidemics. After the May 2022 meeting, a working group was established to decide on countries to focus on, facilitate data collection and to review and revise initial model results. However, data collection in countries took up all the time between May and a couple of weeks prior to this meeting, and the working group did not meet to review. This session has the objective to review the data and initial results from CSAVR-KP.

Session 7 - CSAVR: In this session the group will review and make recommendations on a new approach to improve fits to CD4 distribution at diagnosis in CSAVR.

Session 8 - Key populations in sub-Saharan Africa: A recommendation from the May 2022 meeting was to review outputs of workbook data submitted in 2022 round against Goals and Optima estimates and the group will be presented with these results and make recommendations on the KP workbook process in 2023.

A recommendation from May 2022 that will not be addressed in this session was to consider adding data on clients of sex workers to the workbook. A poster at IAS of a systematic review of this data was identified, but the authors were not available to attend this meeting. The Reference Group will revisit this recommendation in a working group meeting or at the latest at the May 2023 Reference Group meeting.

Other objectives of this session are to:

1. Review UNAIDS 'donut' estimates for distribution of infections by key population
2. Review evidence and strategies for incidence trends among key populations over time and develop short- and medium-term workplan

Session 1 – Incorporating WPP 2022 into Spectrum, EPP, and Naomi models

In June 2022, revisions to the World Population Prospects were released by the UN Population Division and the UNAIDS estimates tools were updated to account for these changes. This session had the objective to review the impact of changes to the WPP 2022 methodology on HIV estimates calculated by the modelling tools (Spectrum, EPP, Naomi, CSAVR, Shiny90).

Rob Glaubius reviewed the main changes to the WPP 2022 methodology, and how each impacts **Spectrum** estimates.

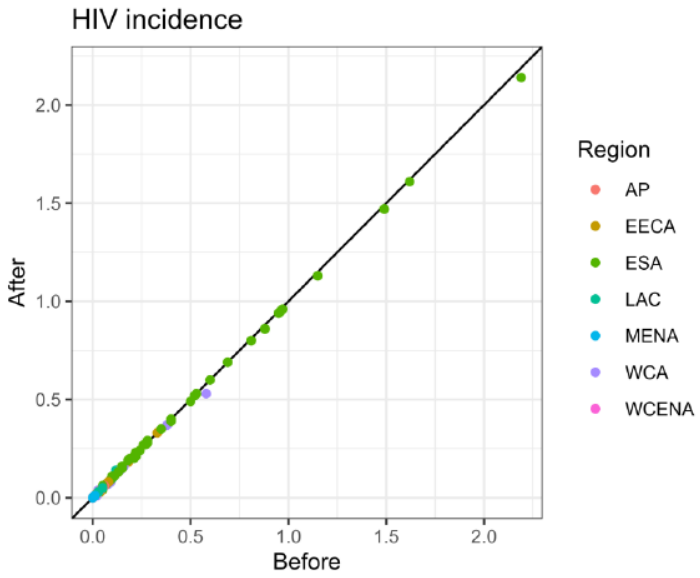
1. WPP 2022 estimates are by single age & year instead of 5x5 year: This simplifies preparation of demographic input files and have no direct implications for Spectrum calculations.
2. Fertility modelled at ages 10-54, not 15-49 like in Spectrum: Due to very small fractions of births in the age groups 10-14 and 50-54, the working group's recommendation was to include these births into the adjacent age groups in Spectrum.
3. Estimates pertain to calendar years, not mid-year intervals: Spectrum stock outputs (PLHIV, number on ART, HIV prevalence, ART coverage) will now represent December 31 values, aligning with ART programme data inputs. Event counts, such as new infections or deaths in year 202x will sum events happening between 1 January and 31 December. Rate outputs (HIV incidence and mortality rates) will use mid-year population denominators by calculating the average of consecutive end-year populations.
4. Net migrants added at end of year instead of throughout the year: Migrants in Spectrum are currently added during the calculation of demographic dynamics throughout the year, before any of the HIV dynamics are calculated. This means that migrants in the projection are exposed to half a year's fertility and HIV unrelated mortality, and a full year of HIV related dynamics. Moving to end-year migration calculations after the HIV dynamics have been processed for a year helps resolve this inconsistency.

Glaubius summarised the changes in key indicators due to the mid-year to calendar year change as follows:

Indicator	Before	After
People living with HIV	$\text{PLHIV}(\text{Jun } 30, 2021)$	$\text{PLHIV}(\text{Dec } 31, 2021)$
ART coverage	$\frac{\text{ART}(\text{Dec } 31, 2021)}{\text{PLHIV}(\text{Jun } 30, 2021)}$	$\frac{\text{ART}(\text{Dec } 31, 2021)}{\text{PLHIV}(\text{Dec } 31, 2021)}$
15-49 HIV incidence	$\frac{\text{NewHIV}(\text{Jul } 1, 2020 \text{ to Jun } 30, 2021)}{\text{Pop}(\text{Jul } 1, 2020) - \text{PLHIV}(\text{Jul } 1, 2020)}$	$\frac{\text{NewHIV}(\text{Jan } 1, 2021 \text{ to Dec } 31, 2021)}{\text{Pop}(\text{Jul } 1, 2021) - \text{PLHIV}(\text{Jul } 1, 2021)}$

These changes will have to be carefully communicated to users of the tools.

Globally, the change resulted in a median <1% decline in ART coverage (the PLHIV denominator is median 1% larger, reflecting growth between June 30 and December 31). Changes in HIV incidence are illustrated below:



Glaubius noted the implications for **CSAVR** and **Shiny90**. Both tools use EPP-ASM, which will need to be updated to match changes to DemProj and AIM. Programme data entered into these tools likely present calendar years rather than mid-year intervals, and the change should improve consistency between the model and the data. No changes are required in the model fitting process.

Next, **Jeff Eaton** described implications for **Naomi**. This tool produces quarterly results and would require simple changes to adapt to the new WPP estimates. Naomi end-year ART coverage will now match the estimate from Spectrum, which was always mid-year.

Tim Brown followed with a presentation on implications for **EPP**. Spectrum will pass new population files to EPP and pressing the existing 'Adjust for changed pop' button in EPP will make necessary changes to the population estimates. For countries that use the urban-rural split, consecutive mid-year estimates from the World Urbanisation Prospects 2018 will be averaged. Since survey and surveillance prevalence data are assumed to be mid-year values, the likelihood in EPP will be adjusted by half a year (average of two December 31 values to get July 1 value). Brown raised for discussion how EPP should best handle migration: continue to do calculations at tenth of year timesteps, or match WPP 2022 and Spectrum assumptions to add all migration at the end of the year (potentially causing arbitrary jumps).

Key points from the discussion:

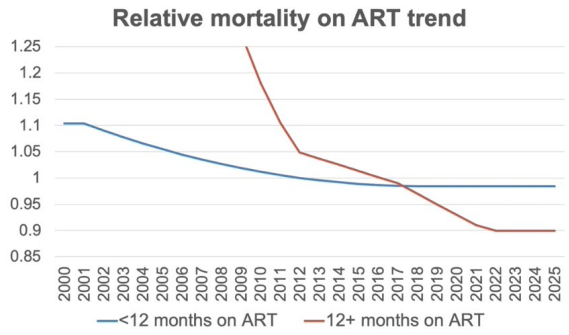
- Patrick Gerland noted that WPP2022 mortality rates include spikes due to crises (conflict, natural disasters etc) and recommended that the HIV estimates tools also consider these short-term increases.
- Gerland asked whether Spectrum uses one-year or five-year fertility rates, and Glaubius and Stover confirmed that Spectrum uses five-year rates and that these should be updated to one-year rates for the 2023 estimates round, while all the other recommended demographic changes are implemented. The impact of this change to subnational files should be carefully considered.

- Mahy noted that the presentations did not show results for the impact of the WPP 2022 estimates on the number of births and how that could affect the paediatric estimates.
- Mahy also noted that in countries where HIV is increasing (e.g., Philippines), the shift to calendar year may make a significant difference.
- Eaton raised the point that the South African Thembisa model uses mid-year to mid-year in estimates and that this will need to be adjusted for. Mahy raised the question of the approach in the ECDC model, and Ard van Sighem confirmed that this model uses calendar year in estimates, so no adjustments are required.
- Stover noted that there are quite a few countries that use an external population adjustment file so that Spectrum exactly matches their official population estimates. These countries will need to be warned about the changes.
- Stover recommended that EPP continues to calculate migration in tenth of year timesteps to maintain a smooth trend for curve fitting.

The recommendations following discussions are captured in **Appendix A**.

Session 2 – On-ART mortality

Jeff Eaton gave an overview of the objectives of the on-ART mortality session.

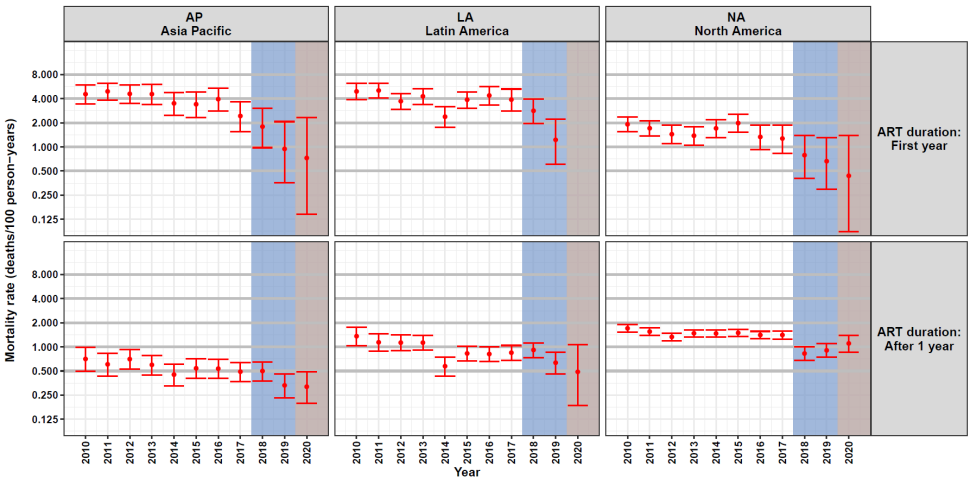


Firstly, the 2022 Spectrum assumption was that mortality for those on ART for more than 12 months (most people on ART) declines by ~10% between 2017 and 2021 in SSA, and then remains constant from 2021. Based on a preliminary analysis of pooled clinical cohort data from the IeDEA network, the group should reach consensus and make a recommendation about this assumption for the 2023 estimates.

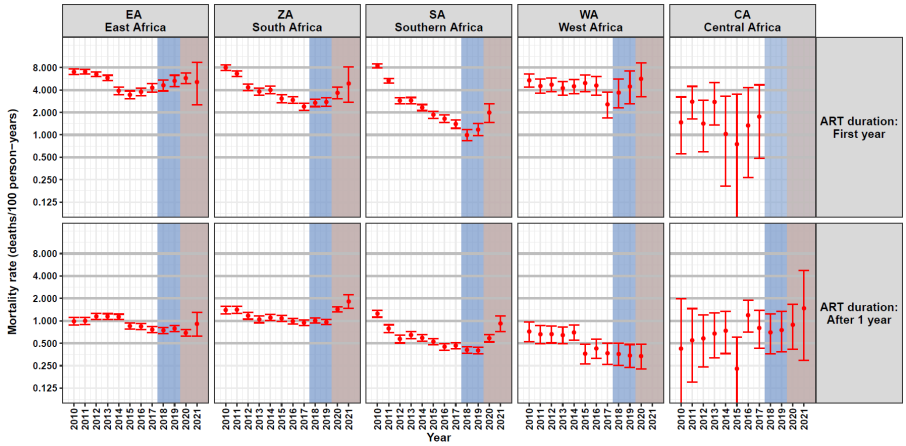
The second objective of the session was to review the comparisons with overall mortality, to decide if there should be changes to the default mortality rates in Spectrum (with presentations relating to Asia-Pacific, WCENA and Sub-Saharan Africa), and a discussion and recommendation for the way forward with how Spectrum accounts for excess non-HIV mortality among PLHIV.

Reshma Kassanjee presented an **update on the IeDEA analysis of adult on-ART mortality**, focusing on trends in all-cause mortality rates amongst adults on-ART who started ART at age 15 or older. Kassanjee presented the 'unadjusted analysis' of the data—a multivariable mixed effects Poisson regression model with random effects accounting for heterogeneity in mortality by programme, but not correcting for biases that arise from the unascertained deaths among people lost to follow-up (in countries in SSA). Data for people with a recorded CD4 count at ART initiation and who met the eligibility criteria at the time were included in the analysis (n=631,644). In the figures below, the blue shading indicates data not included in previous IeDEA mortality analyses and the brown shading indicates the COVID19 pandemic period.

There is a downward trend in estimated mortality in recent years in Asia-Pacific and Latin America. It was noted that linkage to viral registration systems in North America was not complete for 2018-20, which may result in under ascertainment of deaths. In contrast, estimated mortality among those on ART increased in some African regions in recent years.



	Hazard ratio (95% CI)	
	First year of ART	After 1 year
2015-2017 (Reference)		
2018-2019	0.65 (0.5,0.85)	1.01 (0.89,1.14)
2020-2021	0.32 (0.12,0.87)	1.04 (0.83,1.31)



	Hazard ratio (95% CI)	
	First year of ART	After 1 year
2015-2017 (Reference)		
2018-2019	1.03 (0.96,1.11)	0.97 (0.93,1.02)
2020-2021	1.46 (1.3,1.63)	1.18 (1.11,1.26)

At the May 2022 Reference Group meeting, results from the 2022 round of UNAIDS estimates showed that default Spectrum parameter values (based on IeDEA data) lead to too low on-ART mortality in the Asia-Pacific region. The next two presentations relate to this issue.

Renee de Waal described a survey of participating cohort studies **about mortality ascertainment at the Asia Pacific IeDEA sites** that was conducted between May and October 2022. In previous analyses of the IeDEA on-ART mortality data, no correction was made for unascertained death in AP because ascertainment in the region was considered to be very complete with most cohorts conducting death registry linkage or active follow-up of

patients with unknown outcomes. The survey asked questions to site managers about linkage to vital registration and processes following loss-to-follow-up. Sixteen of the 21 sites and 11 of the 13 countries in AP completed the survey. The survey found that mortality ascertainment practices varied, even within the same country; vital status data is probably not as complete as previously thought; and in addition, a recently published study performed in six of these sites found that nearly a third of people reported as lost-to-follow-up had died, implying that routine data underestimated mortality.

De Waal suggested potential next steps in the IeDEA mortality analysis:

1. Update the current analysis by including all the sites as was done previously, without any adjustment for an ascertained mortality;
2. Exclude sites with poor ascertainment and only include sites that have the capacity for registry linkage;
3. Include all sites but consider an adjustment similar to what is done for some of the African regions; or alternatively
4. Produce separate mortality estimates according to whether countries were considered higher income or lower-middle income countries.

Tim Brown gave an update on **on-ART mortality from programme data in Asia-Pacific countries**. Brown explored the age-sex structure of the ART mortality in AP as it is seen in countries' programme data and then compare it with the current and new on-ART mortality parameter estimates from IeDEA for AP. Four countries, Cambodia, the Philippines, Thailand and Vietnam were included in this analysis. Some of the findings were:

- Age distributions on ART over time compared well with final Spectrum files for Cambodia, Philippines and Thailand, but poorly for Vietnam where they will need to work on fitting incidence rate ratios in the 2023 estimates round.
- On-ART mortality increased with age as expected in Cambodia, while the Philippines have a very young HIV age distribution and mortality is noisy at older ages. In Thailand and Vietnam there is some elevation in mortality among 20–29-year-old women.
- In all four countries overall on-ART mortality declined over time
- Using the on-ART mortality multipliers as discussed at the May 2022 UNAIDS Reference Group meeting, total on-ART deaths from 2022 Spectrum files compare well with observed deaths in Cambodia and the Philippines if additional deaths from LTFU are added (15% and 10% respectively). In Thailand, the numbers match without any additions (where total LTFU is also much lower), but in Vietnam Spectrum underestimates observed deaths over time.
- Without using the multipliers (using the default parameters derived from IeDEA data) on-ART deaths are underestimated, falling significantly below the observed deaths.
- Using the new default parameters based on Kassanjee's new analysis of the IeDEA data, Spectrum compares well to observed on-ART mortality in Cambodia and underestimates by a factor of two in the Philippines and Thailand.
- Differences in age patterns of on-ART deaths and LTFU may mean that adding a percentage of LTFU to the deaths may cause some bias.

Brown concluded that the on-ART age data and on-ART death inputs and their validation are valuable additions to Spectrum and that countries that have this data should be fitting their incidence rate ratios using this data.

Adam Trickey presented data from **ART-CC** relating to **AIDS and non-AIDS mortality by sex/age** from European ART cohort studies. From the ART-CC cohort, data of 127,175

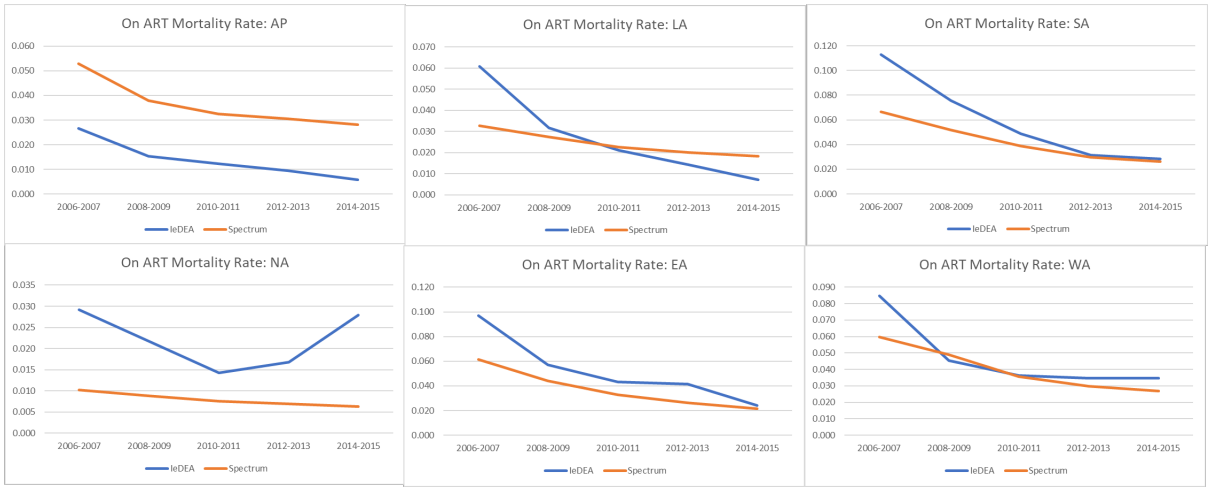
people with HIV from Austria, France, Greece, Germany Italy, Netherlands, Spain, and Switzerland were included in this analysis. Weighted mortality rates for the general population based on the same age distribution of the ART-CC cohort in each country were calculated from www.mortality.org. Trickey's analysis showed that over 2/3 of the excess mortality for men with HIV vs the general population, and 3/4 for women with HIV was due to non-AIDS mortality.

The rates of all-cause mortality were higher at older ages, and AIDS mortality were surprisingly low, particularly in the highest age groups where mortality causes related to age were more prominent than AIDS. The impact of ART on the study population, over time, showed a steep decline in AIDS mortality, and a less steep decline in non-AIDS mortality, whereas the age-matched general population mortality rates increased (when matching the ART-CC age distribution which has become older over time).

An analysis of the impact of CD4 and viral load measurements showed that mortality rates for those with CD4 above 500, in ART-CC, were higher than those with CD4 between 350 and 500, but with wide and overlapping confidence intervals. This was regardless of the duration on treatment.

Rob Glaubius compared ART-CC and Spectrum on-ART mortality, considering the same 8 countries as in Trickey's presentation. When standardising to the same age distribution as people in ART-CC, all-cause mortality rates in the total population in Spectrum is similar to that in ART-CC (~4.1/1000py among men and ~1.9/1000py among women). It was noted that people on ART in Spectrum in these countries are older than in ART-CC, resulting in higher on-ART mortality. When standardising to age distribution again, ART-CC and Spectrum estimates of all-cause mortality for those on ART are similar (~7.2-8/1000py among men and ~5-6/1000py among women).

John Stover showed a **comparison of on-ART mortality rates pre-2016 between Spectrum** (2022 files) **and IeDEA.**

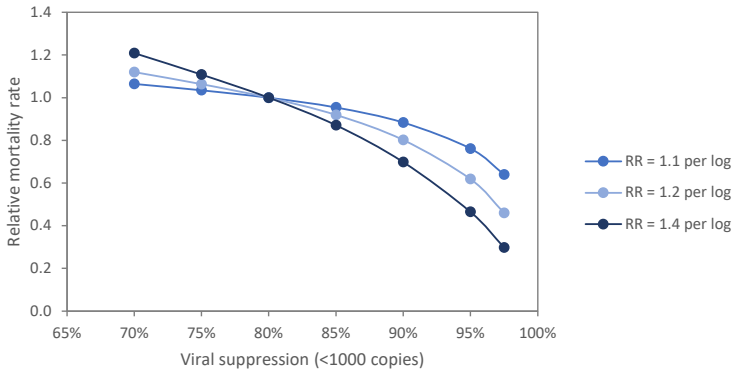


- As discussed by de Waal and Brown, countries in Asia-Pacific have been increasing the default Spectrum on-ART mortality rates (derived from IeDEA data) and in addition China (which accounts for 63% of PLHIV in AP) have been using their own (much higher) on-ART mortality rates.

- There are large differences in North America because USA uses custom patterns for progression, off-ART mortality, and the distribution of new infections by CD4 count.
- In other regions, mortality on ART in Spectrum is close to IeDEA rates in recent years and differences in earlier years are due to different patterns of CD4 count at treatment initiation (the IeDEA analysis excludes people without a CD4 count at initiation, while the Spectrum estimates above reflect all people on-ART).

Stover showed a new **validation** figure added to Spectrum. Countries can input numbers of all-cause deaths to people on ART and the validation figure will display these numbers against the Spectrum estimated number (sum of AIDS and non-AIDS deaths among people on ART).

Leigh Johnson presented on a simulation exercise investigating **VLS as a driver of on-ART mortality**. At the May 2022 UNAIDS Reference Group meeting, Johnson presented a literature review which showed that after controlling for differences in baseline CD4 count and ART duration, each log10 increase in viral load (after ART initiation) was associated with a 10-40% increase in mortality. The question was posed: what would these results imply for mortality on ART if (say) viral suppression increases from 80% to 90%?
Johnson created a simulation model to approximate the relationship between the effect of viral load on mortality per unit change in log viral load and calibrated that to the observed hazard ratios in each of the four studies included in the review (separate simulation for each study). Ten thousand ART patients were simulated with VLs drawn from a reverse Weibull distribution with a shape parameter derived from IeDEA data, and scale parameter varied to produce different levels of VLS in the simulated population.



As expected, the average mortality rate (shown relative to 80% VLS) declines with increasing population VLS. For the scenario where HR = 1.2 per $\log_{10}$ increase in viral load, average mortality will be 20% lower at 90% VLS than at 80%.

However, IeDEA data shows that viral load distributions are very different in different regions of the world and the shape parameters in the reverse Weibull distribution tends to be higher when VLS is higher. Using the shape parameters and VLS found in each region, the relationship between VLS and average mortality is not so clear. Johnson concluded that the fraction virally suppressed in a population by itself will not give a lot of explanatory power, and we may need to know the proportions of treated patients in different viral load categories to get more accurate estimates of mortality on ART.

During the discussion, the most urgent recommendations required for implementation in the 2023 estimates round were:

- In the figure showed by Kassanjee on page 11 there are apparent increases in the IeDEA estimates of mortality in 2017-21 in SSA, while the current Spectrum assumption is decreasing mortality in the same period (down to 90% of 2017. Mahy raised the question whether the apparent increases could be due to increased loss-to-follow-up during COVID; Kassanjee said it is possible that

healthier people were LTFU and the people remaining on treatment during this period were sicker, resulting in higher apparent mortality. Johnson said that the IeDEA executive committee warned to be cautious about interpreting the data in the most recent years in all regions not only because of COVID, but also because some cohorts were censored early or stopped contributing in the most recent years. Zaidi and Phillips raised concerns that it is difficult to reconcile increasing viral suppression from PHIA surveys with increasing mortality. The recommendation was made to decrease mortality in SSA between 2017 and 2021 to 95% of the 2017 value and then remain constant thereafter. The current Spectrum assumption for regions outside of SSA will stay the same.

- Referring to the four possible ways forward in the Asia-Pacific region suggested by De Waal (page 11), Stover suggested to only use data from countries with linkage to vital registration. Johnson agreed but flagged that these countries may be higher income countries, which may cause biases and that an analysis of all options, with consideration of low/high income should be considered. These analyses will be performed in 2023. For the 2023 estimates round, the recommendation is a default mortality adjustment of 4x based on the adjustments in the 2022 estimates and to encourage AP countries to use the new validation screen (Spectrum vs programme all-cause on-ART mortality) introduced by Stover.
- Accounting for excess non-AIDS mortality among PLHIV. Eaton asked – should the UNAIDS estimates explicitly distinguish between AIDS and non-AIDS mortality, and if so, what are the necessary assumptions? Johnson referred to the declining non-AIDS mortality over time (controlling for CD4 at initiation) presented by Trickey and asked what could be driving these declines. Trickey suggested reductions in cardiovascular diseases, infections, and cancers due to increased reductions in viral loads in this population could explain the decline. Trickey reiterated that it will be important to carefully define the difference between what is classified as AIDS mortality and what is HIV related mortality. Stover made the point that if these non-AIDS related mortality is declining due to reductions in viral load, the deaths should still be called HIV-related mortality. After a lengthy discussion about ways forward with this problem, it was recommended that a thorough review to clarify definitions and assumptions for estimating AIDS deaths, including a review of global data on excess non-AIDS death among people on ART should be performed before making recommendations on a modelling approach.

Further recommendations from the discussion are captured in **Appendix A**.

Session 3 – Improving prevalence measures from ANC-DQAs

There is increasing concern about the quality of the routine antenatal testing data used to inform antenatal HIV prevalence and MTCT estimates. At the similarly themed May 2022 Reference Group meeting session, some of these data quality concerns were discussed and the group recommended that countries should be encouraged to do data quality assessments (DQAs) or supportive supervision exercises to improve the quality of the routine data, recognising that results of these exercises may not be available in 2023. In the short term, anticipating that these data quality problems will persist for the 2023 estimates round, this session has the objective to make recommendations on how to address the most significant problems.

Mary Mahy gave an overview of progress in **Protocol development and plans to improve the quality of ANC data.** Several groups are looking at ANC-RT data quality, but inputs from the Reference Group on the design of supportive supervision exercises are crucial since the ANC-RT data are an important component of the modelling. After briefly listing some of the issues that have been elaborated on in previous Reference Group meetings (e.g., more first HIV tests than births, inability to identify first test at ANC vs follow-up tests, inability to deduplicate across facilities, etc), Mahy stressed that good quality ANC data are essential for tracking the HIV epidemic and that the most cost-efficient way to obtain these data is to use the data that are produced from routine visits.

The proposal for the supportive supervision exercises has multiple purposes:

1. To improve the current routine data and systems to provide better services and have more reliable data;
2. To immediately start collecting high quality trend data (ancillary data) so that we don't wait until 2028 to know where the epidemic is heading;
3. To start moving toward individual level electronic systems.

The steps in the proposed exercise are:

1. Select sites to visit
2. Ask regional or national managers to visit sites to collect and review data to identify and address issues
3. Set up individual level electronic systems (where feasible) to collect **prospective data** for additional data required in ancillary exercise
4. Compare facility-level, aggregate, routine data from patient charts over same period as prospective exercise to validate
5. Make recommendations and develop data improvement plan
6. Follow-up activity: Extract data for facility from national system to identify systematic errors in reporting from facility to central level

Proposed ancillary data to collect in step 3 are:

- Sociodemographic (age, parity, marital status, timing and results of most recent HIV test, results of first test at ANC).
- For women who self-report being HIV positive, the source of evidence of previous HIV diagnosis (ART patient card, etc), time of first HIV diagnosis and time of first ART initiation.
- For all HIV+ women, the results of most recent viral load test.

Remaining issues that the groups are still considering include:

- The method of sampling sites that are representative of the country as a whole (i.e., not focus on only high burden sites).
- At the same time, these sites should have the infrastructure to roll out a simple electronic system.
- These exercises should not have the purpose to set up new electronic systems, but should expand on existing plans to implement electronic systems in countries.
- Data included prospectively should be useful not only for surveillance, but also provide actionable information for facilities.

Comments following Mahy's presentation included:

- Stover: When compiling the protocol for the supportive supervision exercises, it will be useful to add some examples of how the data will be used.
- Eaton: The main benefit of having the electronic data capture is that is provides the opportunity for more real time oversight of the quality (vs the regional director needing to drive to the facilities on a monthly basis).

Ian Wanyeki presented on **Guidance on the correct indicators that should be reported on**. Every year during the estimates workshops, guidance is shared on the definitions of indicators and where they are used in the modelling tools. However, many of these indicators are used only once a year and if issues are discovered it is difficult to find solutions for older data (if going back to clinics to go through registers). This is one challenge that may be overcome when moving to electronic registers. Five ANC indicators are requested: Number of ANC clients, Number of ANC clients with known HIV, Number of clients with known HIV already on ART, Number of ANC clients tested for HIV and Number of those tested who test positive. Issues that arise are that sometimes the Number of ANC clients are fewer than those tested for HIV. In some cases, 'Number of known HIV positive' is not captured and 'Number already on ART' is used as a proxy. If a country uses paper registers, capturing new indicators means that they must print and distribute new registers. Botswana started capturing 'Number of known HIV negative' which had to be incorporated into the Naomi data capture tools.

Oli Stevens gave an **update from the Mozambique ANC data quality exercise** on behalf of **Makini Boothe**. In Mozambique, ANC and PMTCT coverage has been above 100% in nearly all provinces over time and persistently above 100% for the last 3 years and the country has seen very sharp declines in ANC-RT HIV prevalence since 2015. For the data quality assessment, data for one quarter (Q4) in 2019-2022 in 3 sites per province were collected. Headline results of the quality assessment that Stevens presented included HIV prevalence and ART coverage trends matching those routinely reported. An interesting finding of the more granular data obtained during the exercise, is that HIV prevalence appears to increase as the total number of ANC visits per woman increases, indicating better care seeking behaviour among women with HIV.

In the final presentation of the session, **John Stover** presented on the **Spectrum updates** related to ANC data as listed in the May 2022 meeting recommendations.

- A row was added to the ANC testing editor for the programme reported total number of in-facility births. This number is visually compared to the total number of estimated births from Spectrum, the number of first ANC visits and the number receiving at least one HIV test. Stover noted that he anticipates large discrepancies between these indicators in many countries and that a few countries should be identified to assist with developing guidance for such cases. Isaac Taramusi (Zimbabwe) and Wilford Kirungi (Uganda) confirmed that they will assist.

- Stover reported that ‘known HIV negative’ was not added, given that very few countries are collecting this information, but after discussion it was decided to add this indicator to the ANC testing editor.
- The HIV-related fertility adjuster data editor was removed, to instead read-in ANC prevalence and denominator from the (national-level) ANC testing editor. This point was contested by Eaton, Mahy and Korenromp and it was decided to have a smaller group discussion on this issue.

Key points from the discussion:

- Monita Patel raised the question of whether we should also be capturing ‘unknown HIV status’ at first ANC if we are capturing ‘known positive’ and ‘known negative’. Mahy responded that ‘first ANC’ in this context means ‘result at first HIV test’ and that all women should get this first test at some point during the pregnancy.
- Leigh Johnson, the chair of the session, raised the question to the group of what options there are to adjust ANC-RT data in the short term, while we wait for the data quality exercises to inform decisions.
    - Mathieu Maheu-Giroux who was involved in a previous assessment of MER data quality for PEPFAR, noted that if there was less than 10% missing or implausible data in a year, there was not substantial impact on HIV prevalence estimates if this year was excluded during calibration.
    - Mahy suggested that large changes in prevalence from one year to the next, given large number of tests, should flag that the ANC data should not be used.
    - Eaton suggested to move away from using the PMTCT inputs directly and rely more on survey data about ANC testing and PMTCT to make adjustments to the input data.
    - Stover suggested that large discrepancies between the program and Spectrum estimated numbers of births should be queried and those data probably dropped.
- Regarding the representativeness of sites selected for the supportive supervisory exercises, Johnson asked whether a fixed selection of sites will be visited every year, or if sites will be changed every year. Mahy noted that using the same sites every year may result in just improving data quality at those sites, but Eaton believes that having a time series of data from the same sites would be highly valuable.

The recommendations that followed this session are captured in **Appendix A.**

Session 4 – Testing and treatment churn

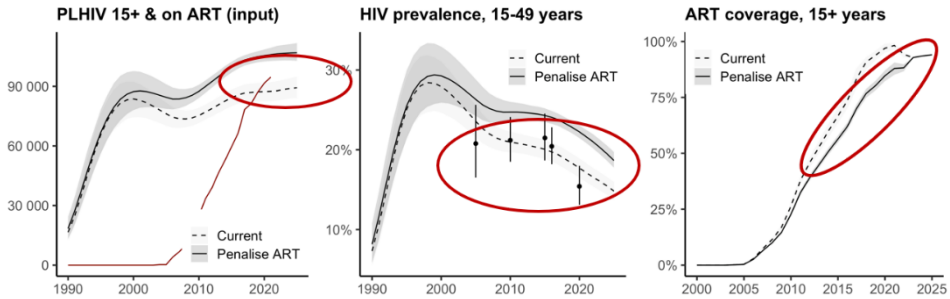
This session follows from discussions and recommendations from the May 2022 Reference Group meeting. At this meeting data was shown of incompatibility between programme data of people on ART and estimates of PLHIV, specifically that estimated ART coverage was above 100% in some countries, and a way to account for this will be discussed in this session.

In the first presentation, **John Stover** showed **Spectrum updates** that were suggested during the May 2022 meeting:

- Visualisation of 'waterfall' results estimated by Spectrum: the components of change in total number of people on ART in a year (newly on ART, re-engaged, died, disengaged, transferred out and unattributed) can be entered in an editor and will be displayed.
- ART DQA results: countries will still enter their DHIS programme data into the main editor but have the option to open an adjustment editor where they can enter a scale factor to adjust the programme data (by time and sex).

Rob Glaubius presented on an investigation to incorporate **survey ART coverage data (by age and sex) into the calibration of incidence rate ratios (IRRs),** which are used in Spectrum to disaggregate input HIV incidence trends by age and sex. Currently, these ratios are estimated from HIV prevalence in household surveys (by age and sex). These ratios can be estimated as fixed values over time, or as time-varying ratios if more than one survey estimate is available. Glaubius showed that adding ART coverage data as calibration data had minimal impact on the model's fit to HIV prevalence or ART coverage and did not impact PMTCT coverage estimates. However, allowing time-varying IRRs did have an influence on fits which suggests that the use of the time-varying IRR model should be reviewed.

Jeff Eaton presented an analysis of the May meeting proposal to **reject EPP fits with numbers of people on ART greater than the number of PLHIV (coverage over 100%)** to enforce consistency with ART programme data. The group anticipated that this might inflate the number of people with HIV to accommodate the input number on ART, but by imposing this, the user would be forced to wrestle with the inconsistency of the ART programme data, the prevalence data, and the population data, and seek some consistency between those. Eaton approached the problem by penalizing parameter combinations where the number on ART was greater than the number of PLHIV by taking the sum of the absolute difference between the model number on ART minus the input number on ART and penalize that with a normal distribution with a small standard deviation. This difference will be 0 and thus the best fit if the number of PLHIV is greater than the number on ART. Eaton showed the impact on estimates for Matebeleland South, noting that the approach will allow ART coverage to increase steadily and remain below 100%, but that the model has a poor fit to HIV prevalence data and results in much higher HIV incidence rates.

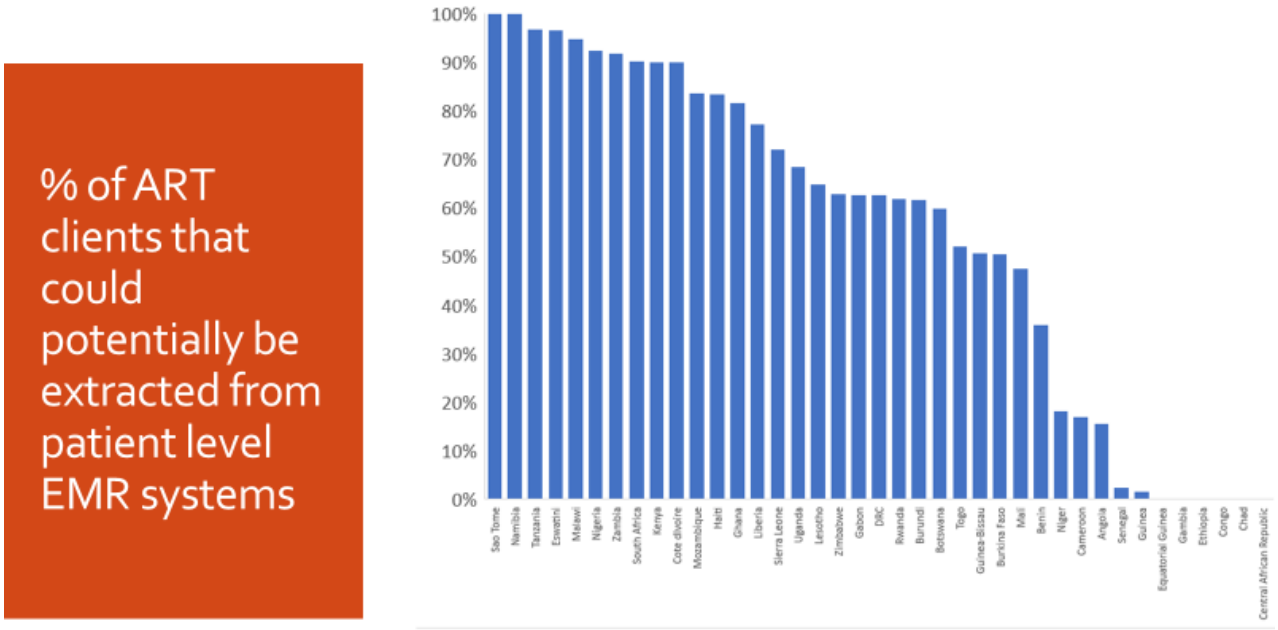


Eaton's conclusions are that although this approach enforces consistency with the ART programme data, values of just under 100% may still be implausibly high and may not be more useful for planning than a value of 110%. This approach also does not allow a robust estimate of the gap in treatment coverage and although it highlights a problem to the user, there are no convenient tools to understand or fix the problem.

Eaton's proposals for future directions are listed below:

- Move away from the urban/rural or sub-region stratified explicit EPP structure to align EPP and Spectrum populations.
  - EPP stratification is not consistent with current programme data structure and attendance patterns.
  - The stratification is a relic of historical surveillance data when oversampling of urban locations with different trends occurred. Fitting well to the historical data can be addressed instead by weighting the ANC likelihood.
- Move towards likelihood-based solution: modelling ART initiation rates and fit to ART data.
  - Examples: Thembisa, Wolock model
- Implement full EPP-ASM model: simultaneously estimate incidence trends and rate ratios.
- Continue focus on programme data review and quality, clear & noisy validation in Spectrum, judicious estimates file review.
- Deeper integration of programme data into epidemiologic inference
  - Programmes still diagnosing a lot, with positivity >>1%, initiating a lot of PLHIV are not near >95% ART coverage.

Ian Wanyeki presented on the **Current levels of electronic medical records data use in SSA countries**. Wanyeki sent a short questionnaire to in-country estimates teams to gather information on their use of electronic medical records. Systems in use varied across countries, and also within countries over time. Some countries said they used EMR systems, but data capture happened at ad hoc intervals instead of continuously.



Source: Country HIV estimates focal points, Oct 2022

A notable challenge recognised is the use of non-networked systems within a country. For example, districts have their own systems that are not linked to each other, which means that if someone moves, the system in their new area won't recognise them as an existing ART client.

Key points from discussion:

- Kirungi noted that 'transferred in' is still an important indicator in settings with sub-national files and should be included in the waterfall editor and visualisation.
- Korenromp noted that it would be more informative to show both the Spectrum and the programme 'waterfall' visualisations. In support of this, Eaton noted that this way the country can explicitly see their programme data and how Spectrum has reconciled it, and this will motivate conversations about large differences.
- Maheu-Giroux asked whether including ART coverage data to the IRR fit reduced uncertainty in the estimates, which could be an argument in favour of using this data. Glaubius responded that the methods implemented here do not allow estimation of uncertainty, and this question will need to be investigated externally. Only point estimates from the IRR fitting process are used in Spectrum, and the uncertainty bounds that are generated when uncertainty analysis is done do take into account some uncertainty in the IRRs, but that isn't propagated from the fitting process.
- Tim Brown agreed with Eaton that the misallocation between urban and rural could contribute to the problem of ART coverage exceeding 100% in countries with subnational files, but these countries may be reluctant to do one national projection. Brown suggested that some checks can be built in that can flag if some sub-region has coverage exceeding 100% and this will encourage the user to go back and look at allocations in Spectrum and change using programme data or Naomi estimates.

A summary of the discussions is captured in the recommendations (**Appendix A**).

Session 5 – HIV+ Migration in AIM & Enhanced User guidance for concentrated epidemics

Following the May 2022 meeting's recommendations, a working group met in August to further discuss approaches to handle migration of people with HIV in the modelling tools. This session provided feedback on this group's discussions, as well as present results on age and sex distributions of migrants with HIV in a few countries with such data available. In this session, the group also reviewed the guidance that UNAIDS will provide to countries for dealing with data on migrants with HIV, as well as other issues such as model choice, knowledge of status methods, demographic scope of national estimates, and CSAVR inputs.

Eline Korenromp presented an **update from the working group on migration of people with HIV.** The issue mainly concerns countries with concentrated epidemics (17 countries), and fewer than 3% of PLHIV in these countries are new migrants. Korenromp gave an overview of the recommendations from the May meeting, and the points below combine the input from the working group as well as the discussion after the presentation:

1. Despite data challenges, country immigrant data more likely reflect actual immigration, and so should substitute AIM's calculation of multiplying the country's HIV prevalence by age and sex to net migration numbers. If users input data on migrants with HIV into AIM, AIM should not apply this background calculation. If countries have missing data for intermediate years, users should impute values.

2. CSAVR will use the AIM-entered data on migrants with HIV, assuming all are diagnosed, to calculate its estimate of PLHIV, KOS and (fitted) deaths. Guidance to users will be to enter in CSAVR all new (first-time in-country) diagnoses, regardless of citizenship.

3. Expanded user guidance of which countries should use migration data include:
    - Recommended scope for all national estimates is to use the *de facto* population, including non-citizens.
    - Immigrant candidates sent back home by some ‘destination’ countries upon testing HIV positive should not be entered into the ‘destination’ country’s estimation. Instead, they get captured in the ‘origin’ country estimation, through surveillance, surveys and/or routine data.
    - Countries that use CSAVR or ECDC, fitted to new diagnoses including among immigrants, can add known positives/previously diagnosed into AIM. If directly entering KOS numbers (as opposed to adopting CSAVR’s KOS estimate), then also include these in the KOS calculation.
    - Countries using EPP with Shiny90 should never enter data on immigrants with HIV to AIM (generalised epidemic countries usually have net **e**migration).
    - Countries using another incidence model, or EPP without Shiny90, can enter all data on migrants with HIV, regardless of place of first diagnosis, into AIM’s immigrant tab and should then include the same in KOS entries (cumulative diagnoses minus cumulative deaths and emigrations).

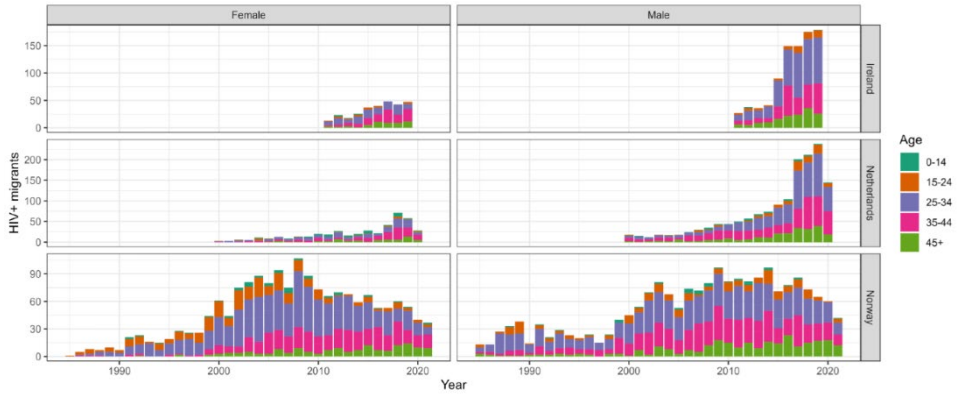
4. AIM should allow entry of data on migrants with HIV, with or without age/sex disaggregation.

- Based on this recommendation from May 2022, **Rob Glaubius** presented an **exploratory analysis of data of migrants with HIV**, to examine if any countries have good data on age/sex patterns of migrants, to develop default age patterns.

As of 2022, data on migrants with HIV must be entered by year, sex, and five-year age-group. However, some countries know only total numbers, without sex/age disaggregation. Glaubius explored two questions:

- Can default age/sex patterns be derived from countries with disaggregated data that countries with aggregated data can apply?

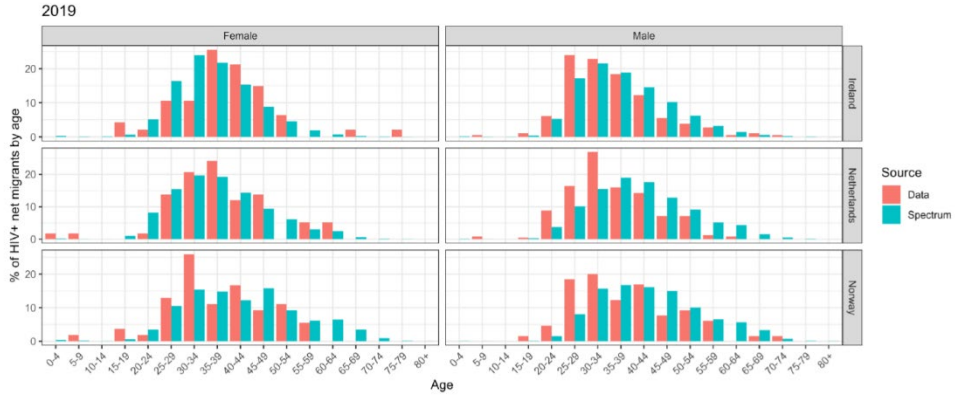
Data were available for 5 countries, and the data for 3 of the 5 countries with public files are shown below.



In general, there were more male migrants with HIV than females, and the majority were aged 25 to 44. The median age of migrants has been increasing over time, consistent with an aging global HIV epidemic. However, there is heterogeneity in the median age by sex and country among these three countries in the WCENA region.

- How do these data compare to net migration of people with HIV calculated by DemProj?

For these countries, Spectrum's net migration calculation by age and sex underestimates the numbers of immigrants with HIV that were entered. Spectrum's age distribution may be an acceptable approximation of the distribution in the data, as shown for 2019 below. However, in some years Spectrum calculates net **e**migration in some age groups, for example in Ireland in 2015.



Glaubius concluded that, based on the heterogeneity in data from three countries in WCENA, it will be hard to justify deriving regional default patterns. An additional observation should be considered in the first recommendation above (users should impute missing data): the

observed age distributions gradually shift over time, implying that while imputing for nearby years may be robust, recent years' data should not be used to impute early missing data. Spectrum's calculated age/sex distribution of migrants with HIV can be used to disaggregate country immigrant data if there is a net influx of migrants, but further investigation is required for when there is a net outflow.

Key points from discussion:

- Mahy asked whether the sending location of the migrants (if known) could be used to inform sex/age patterns. Sabin and Brown commented that even if sending locations were known, the age/sex of people who tend to migrate may be different to the overall age/sex pattern of PLHIV in the country.
- Glaubius commented that it is generally assumed that migration to Europe is largely from Africa, but the data show that 75% of immigrants with HIV are male, which does not reflect the epidemic pattern in Africa.
- Brown noted that although countries track immigrants with HIV, they don't track emigrants with HIV, which can be problematic. Glaubius highlighted again that Ireland had net emigration in 2015, but the user inputted data on migrants with HIV is of course positive. This issue warrants further investigation.

Eline Korenromp presented an overview of updated **user guidance for concentrated epidemics** (updates to the *Quick start guide for Spectrum*) on the following issues:

1. Model choice and triangulation

    - If enough data are available over time, by sex and by key population, and capacity allows, it is advantageous to run two suitable models (e.g., EPP and CSAVR), compare results and try to explain differences and uncertainties.
    - Models can be compared for coherence of the cascade of care: reported numbers of PLHIV knowing their status should not exceed estimated PLHIV, and ART/PMTCT coverage should not exceed 100%, for any year and group.
    - The final model choice should be based on plausibility of results, e.g., coherent cascades over time and across groups, not on precedents of past estimates. Nevertheless, large changes from last year(s) estimates should inspire scrutiny and warrant a clear consensus explanation for the direction and magnitude of change.

2. Knowledge of Status: preferred estimation method

    - Countries with more than one national population-based serosurvey should estimate KOS using the Shiny 90 model.
    - Countries that do not have serosurvey data (mostly concentrated epidemics) can enter cumulative case surveillance data covering 2015-2021 or longer and subtract cumulative (all-cause) deaths and emigrations among diagnosed PLHIV. All new, first-time diagnoses (citizens and non-citizens) should be included. When entering data for immigrants with HIV into AIM, these should also be included in calculated KOS for those years.
    - If case and/or death surveillance have limitations (e.g., KOS drawn from case and death reports exceed the country's modelled total PLHIV for any year), model estimates should be used. Model estimates, based on diagnoses, deaths, and emigration, cover all years and are comparable between countries.

- If countries do not have a household survey, are unable to subtract deaths and (if applicable) emigrations from case reports, and do not use CSAVR or another model to estimate KOS, they should leave KOS blank.

3. CSAVR inputs, given data availability and quality

New case diagnoses	**Required**. Enter all deduplicated, first-time HIV diagnoses among all residents, returning working migrants, and immigrants. Previously positive immigrants diagnosed abroad should not be entered, but should instead be entered into AIM
AIDS deaths	**Recommended.** If complete, good-quality causes of death (IHME groups 2A and 2B). Up to 3 data sources can be entered. Latest GBD inputs adjusted for misclassification should be used, for years available. These can be supplemented with original VR data for recent years, if there were no major historical difference between the two time-series.
CD4 at diagnosis	**Optional**. *Enter* data for all available years but *use in fitting* for only those years where >80% and preferably >95% of new diagnoses had a CD4 count recorded to be representative.
Annual HIV test volume	**Optional.** This provides context to trends in diagnoses but does not influence the fit. With disruptions in testing services, during COVID in 2020-21 or humanitarian crises, fewer diagnoses may be due to less testing – evident as a stable (CSAVR-calculated) test positivity rate. CSAVR's estimate is robust against temporary falls in diagnoses and using all diagnoses data including years with dips is important to capture (not overstate) historic progress in KOS. Drop years of diagnosis data if reduced tests and diagnoses reflect recent reporting delays from service disruptions to prevent biasing CSAVR's incidence estimate downward.

4. Demographic scope of national estimates (*de facto* vs. *de jure*)

- The WPP 2022 country demographic estimates are for the national *de facto* population, including citizens and non-citizens (born abroad), with the latter including long-term residents.
- UNAIDS recommends that national HIV estimates be for *de facto* populations, which matches the demographic estimates coming from WPP 2022.

A few countries (2022 round: Bahrain, Kuwait, Oman, Qatar, Saudi Arabia, UAE) have historically limited their estimate to the *de jure* subpopulation (citizens only). Such estimates do not cover the full national need of HIV prevention and care services, neither at country nor at regional and global levels. However, some were used in conjunction with surveillance and program datasets also covering non-citizen residents, thus creating internally incoherent estimates or inconsistency between Spectrum and GAM.

5. Surveillance requirements for a valid, publishable incidence trend

- Incidence trends from EPP or AEM models are published if four or more survey prevalence estimates were entered and fitted, including one or more datapoint within the past 4 years (2023 round: ≥ 2018)
- For low-level and concentrated epidemics estimated based on case and death reports, incidence trends are published if based on at least 8 datapoints on AIDS-related deaths (CSAVR) or AIDS diagnoses (ECDC model) within 1990-2021.
  - After discussion following this presentation, it was recommended to add 8 years' case diagnosis data, in addition to 8 years of HIV/AIDS death data, as a requirement to publish a CSAVR-based incidence trend.

6. Fitting fertility to ANC data.

   - In settings without universal routine testing of all women in all ANC clinics, fitting to ANC data requires data that are representative of all pregnant women.
   - If HIV testing is limited geographically (e.g., only urban areas) or if women's ANC attendance or likelihood of being tested is selective (e.g., if higher-risk women do not attend early in pregnancy, or testing is targeted to higher-risk women) then ANC-based prevalence may not reflect pregnant women overall, and the Local Adjustment Factor should not be fitted.
   - Recommendations for fitting to ANC data from Session 3 will also be listed in the user guidance.

7. Validate low-risk women estimate on ANC data, in AEM.

   - This was put forward by Korenromp as a point for discussion. In Asian countries ANC data are not included for comparison with the low-risk female population. An existing spreadsheet tool developed in 2009 did a conversion from ANC data by age into the expected low-risk women prevalence considering age patterns in population level marriage and fertility rates. Could this spreadsheet be used again in those countries where low-risk women are an important part of the national epidemic?

Key points from discussion:

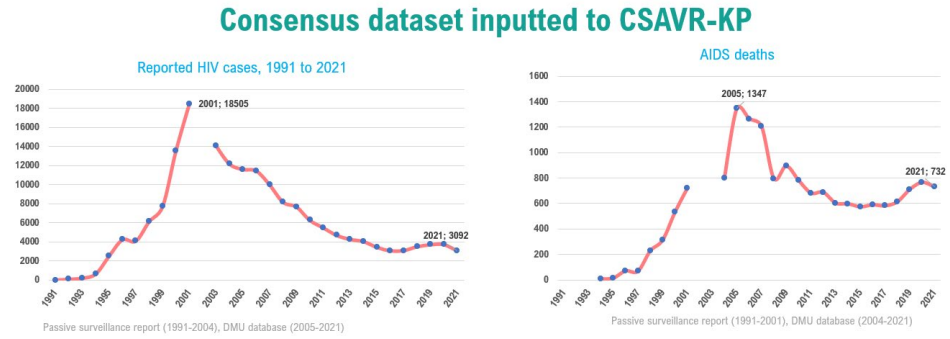
- Brown responded to the question under point 7 above that a small working group may need to investigate the availability of ANC data by age, over time. Age patterns in both women with HIV and the low-risk female population evolve over time and getting this data together may be difficult for countries. Brown suggests that 2 or 3 countries should be identified to pilot.

The recommendations arising from this session are shown in **Appendix A**.

Session 6 – Key populations in concentrated epidemics

At the May 2022 reference group meeting, results from a review on quality and biases in HIV case and death reports by mode of transmission in countries with concentrated epidemics were presented. A working group was set up to decide on countries whose data may be of good quality (Armenia, Cambodia, Morocco, and Peru), to facilitate data collection and to review and revise initial model results. However, data was not collected in time to reconvene the working group, and this session had the objective to review the data and initial results from CSAVR-KP for three countries that were able to provide data in time for analysis – Cambodia, Morocco, and Peru.

Khin Cho Win Htin presented on the **data collection process and challenges faced in Cambodia**. Cambodia is particularly interested to join this pilot because the national programme needs an improved understanding of mode of transmission of HIV, particularly to estimate progress towards the 1st of the 95% and to understand who the unreached populations are. From 1991 to 2004, data on HIV cases, AIDS cases and AIDS deaths were collected through passive surveillance, with possibilities of duplications and incomplete data reported. From 2005, a data management unit (DMU) was established in the national programme which largely manages data from voluntary counselling and confirmatory testing sites, data on ART clients, and data from implementers of prevention interventions since 2018. In the last 5 years, data were mainly collected electronically. Thirteen parallel and fragmented databases are on track to become integrated and streamlined into one.



The main challenges with data collection and management in Cambodia highlighted by Win Htin were:

- Validity and accuracy of data before establishment of data management unit (pre-2005)
  - Duplication vs under-reporting
  - Non-standardized definition on AIDS case
  - Limited institutional memory/documentations to better understand data quality and caveats.
- Early years of DMU establishment
  - Data entry was mainly focused on the data received from ART sites.
  - Had to deal with backlog and year of reporting vs actual event (new diagnoses/death) may not necessarily match.
  - Delayed reporting from some sites
- AIDS deaths data
  - Not all AIDS deaths are recorded as AIDS-related deaths even in the vital registration system.

  - Aging cohort with co-morbidities and NCDs has implications on cause of death recorded at the hospitals.
  - Exact proportion of AIDS deaths among LTFU is unknown and may vary over time.
- CD4 at the time of diagnosis
  - Early years – data was weak and incomplete.
  - Past 5 years – Skewed towards "target testing" due to issues with reagent stockout and run-down machines.
- Key populations data (new diagnoses)
  - Electronic, deduplicated data using UIC is available starting from 2018 only (National Prevention Database)
  - For CSAVR pilot, new diagnoses among KP were extracted from National Prevention Database though some KPs may walk-in to VCCT sites/ or referred through other services (OPD, surgical wards, ANC etc) and may not capture in NPD database – this is likely a small proportion
  - Though VCCT database collect the KP status, it is again underestimated due to
    - reluctance to disclose their KP status.
    - VCT counsellors may not have enough time to build rapport and trust with clients.

In the 2022 estimates round, **Cambodia** used the AEM model, and **Morocco and Peru** used EPP-concentrated. In the next presentation, **Eline Korenromp** showed a table of the indicators from which **consistency and gaps** can be inferred**; and impressions from fitting the original CSAVR model** in these countries.

Indicator	Cambodia	Morocco	Peru
HIV/AIDS deaths	1994-2021, M/F & age, Deaths only for PLHIV on ART, and do not adjust for possible deaths among those LTFU (>3 times deaths)	2015-2021, M/F & age	2000-21
Misclassification-adjusted deaths; IHME quality Cause of Death (CoD) group*	No GBD data; CoD quality *1A*	GBD2019 *adjusted data* 2000-2014 3-fold higher than raw; CoD quality *2C*	GBD2020 1999-2017; CoD quality *2B*
Case diagnoses	1991-2021, 2000+ by M/F & age (suppress unverified high numbers 2005-07)	1987-2021: M/F & age	1983-2021: M/F & age
Case diagnoses by KP	2018-2021: MSM, FSW, PWID	1987-2021: MSM, BSM, FSW (defined as all non-PWID F with multiple partners), PWID	1989-2021: MSM, FSW, PWID, TG
KP size & prevalence	From AEM, no change	From EPP, no change	From EPP, no change

* 1A=highest quality; 2C=lowest quality

Fitting the original CSAVR model to reported AIDS deaths (or Cambodia: on-ART deaths) gave later and (mostly) lower epidemics for Cambodia, Peru, and Morocco. This observation remained in Morocco and Peru, when fitted with the GBD adjusted AIDS deaths.

Houssine El-Rhilani (from the UNAIDS Morocco country office) made the comment that case diagnoses significantly increased since 2015, when visas were no longer required to enter the country. This may have attracted migrants with HIV, since ART is provided free in Morocco.

Next, **Guy Mahiane** presented the **initial fit of CSAVR-KP to data from Cambodia and Peru and discussed some model refinements**.

New features of the model, based on recommendations from the May meeting, are that 1) key population sizes can be entered by the user and vary over time and 2) prevalence data (entered or pulled in from EPP) are included in the fitting. The latter includes prevalence among IDU (aggregated for both sexes) and ANC prevalence.

For the two countries, CSAVR-KP was fitted to data on new diagnoses by key populations, as well as population sizes and HIV prevalence from EPP or AEM files. The model fit well to HIV diagnosis and prevalence data of key populations, but the CSAVR-KP incidence and prevalence trends overall were higher than for AEM (Cambodia) and EPP (Peru). HIV-related deaths (not included in the fit) produced by both CSAVR-KP and AEM/EPP were much higher than country reported HIV/AIDS deaths.

Next step in model development is to include transgender women as a key population. The group agreed that data collection and fitting for Morocco and Armenia should continue, followed by a more detailed interrogation and discussion of the model fits.

The main points arising from the discussions following these presentations are captured in the recommendations in **Appendix A.**

Session 7 – CSAVR

In this session, the group reviewed CSAVR interface updates, and a new model proposal to improve fits to CD4 at diagnosis data. This sought to address a longstanding challenge of CSAVR typically overestimating the distribution of CD4 count at HIV diagnosis compared to observed CD4 distributions from case reports.

First, **Guy Mahiane** presented on CSAVR updates recommended at the May 2022 Reference group meeting.

Some recommendations were not implemented and are postponed:

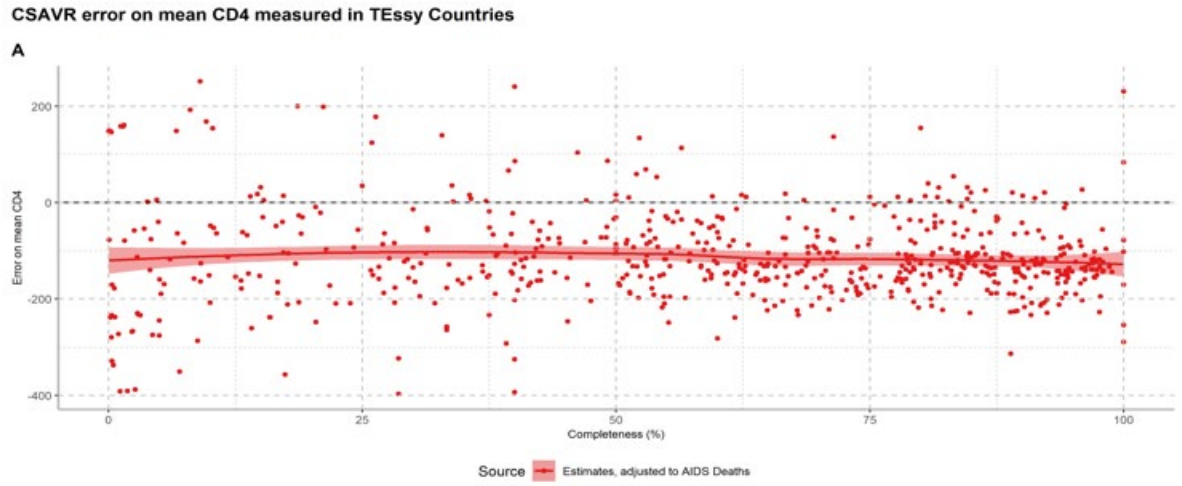
- Visualisation of HIV testing data (annual numbers tested and positivity) in CSAVR.
- Allow for CSAVR entry and saving of mortality data from multiple sources (original VR, misclassification adjusted GBD, etc.) in parallel, and choose one – or a user-defined hybrid – for fitting.
- Enable CSAVR data editors to retain data on diagnoses, deaths and CD4 for all years, but selectively exclude some data points from model fitting.

The following recommendation was implemented:

- To avoid people on ART outnumbering PLHIV with knowledge of status, CSAVR sends to AIM estimates for proportions of PLHIV not on ART knowing their status, rather than overall proportions knowing their status. Knowledge of Status then gets calculated as
    - KoS = ART + (PLHIV – ART) * (KoS among PLHIV not on ART)

Mahiane illustrated the implementation for Qatar and Kuwait.

Mahiane went on to present a **CSAVR development to improve fits to CD4 distribution at diagnosis**. A recommendation from the May 2022 Reference Group meeting was to determine a completeness threshold for CSAVR to be fitted to observed distribution of CD4 at diagnosis. Mahiane illustrated (below) that the average overestimation of CD4 count at diagnosis by CSAVR remains consistent regardless of completeness of data. In the graph below, each point represents a country and a year. This suggests that setting a completeness threshold won't solve the overestimation by CSAVR.



Mahiane suggested the addition of an extra parameter to be fitted by the model that will scale the CD4 distribution in the model to match the data and illustrated examples from Croatia, Denmark, Germany, and Portugal.

Discussion following this presentation included the observation that it is clear from the plot above that overestimation by CSAVR is independent of the completeness of the CD4 at diagnosis, which may reflect structural issues that will not be overcome by introducing a scaling factor (Eaton).

The main points arising from the discussions following these presentations are captured in the recommendations in **Appendix A.**

Session 8 – Key populations in sub-Saharan Africa

Recommendations from the May 2022 Reference Group meeting regarding estimates for key populations (KPs) in sub-Saharan Africa (SSA) included that the workbook-based data collation should be continued in the annual estimates process; that workbook outputs should be compared to Goals and Optima estimates, and the latter 2 models should be updated to use the workbook data. This session includes a presentation relating to these recommendations.

In addition, the session will review the UNAIDS 'donuts' analysis and suggested methods to show time trends in the donuts.

This session had 3 main objectives, which Mahy briefly introduced.

1) To review the UNAIDS 'donuts' analysis (global and by region)
2) To comment on strategies to produce incidence trends among KPs
3) To make recommendations about the KP workbook process in SSA

Keith Sabin presented **on the methods used to estimate the proportion of new infections that are among key populations (the 'donuts')** since they were first published in the UNAIDS global report since 2016. These donuts show the proportion of new HIV infections in sex workers, people who inject drugs, gay men and other men who have sex with men, transgender women, and the remaining population, *globally and by region*. The proportion of new infections attributed to key populations have increased over time, but due to fluctuations in estimates over rounds, the question was raised whether these increases are real or due to the analytic approach. Until the July 2022 report, data sources for the donuts included:

- Spectrum: EPP-Concentrated or AEM
- Modes of Transmission (MoT) or Incidence Pattern model (IPM)
- In 2022 included the KP workbook which collated GOALS and/or Optima estimates from 31 SSA countries
- Data on new diagnoses (WCENA, Russian Federation) from ECDC/national reports
- Total adult (15-49) male & female new Infections from Spectrum
- Countries without a relevant model or data: apply regional median or average of adult new F/M infections.

Transmission to clients/partners of KPs were assumed to happen as a regionally fixed ratio of new infections in key populations, based on a non-systematic literature review performed in 2018. For example, it is assumed that each new infection in FSW will result in a number of new infections in clients, a number that varies by region. It is assumed that no transmission occurs from KPs with existing infections.

Eline Korenromp then presented on proposed methods to show a **time trend in estimates from the donuts**. For this analysis, the focus shifted to country-level estimates (which can then be aggregated to regional and global level) which added the additional constraint that numbers of new infections per KP should add up to 100% of all KPs at country level, and not just regional. Data sources for this analysis included:

- For 54 concentrated epidemics: EPP-Concentrated or AEM
- For 47 SSA epidemics: Goals
- For ECDC countries and USA, Canada and New Zealand, new diagnoses were used as a proxy for new infections for MSM and PWID, but FSW data in those countries were problematic.
- In 24 countries that had no estimates, the regional median proportions were taken.

Preliminary results show that in most regions, between 2010 and 2021, proportions of new infections of the total for FSW remained stable or declined, while the proportion for MSM increased. The proportions of new infections in KPs of the total new infections are generally lower than in previous donuts included in the UNAIDS reports. This can be attributed to several reasons including the stricter limits (country level vs regional level should sum to less than 100%) and that Goals estimates for SSA are lower than the estimates from IPM/MoT.

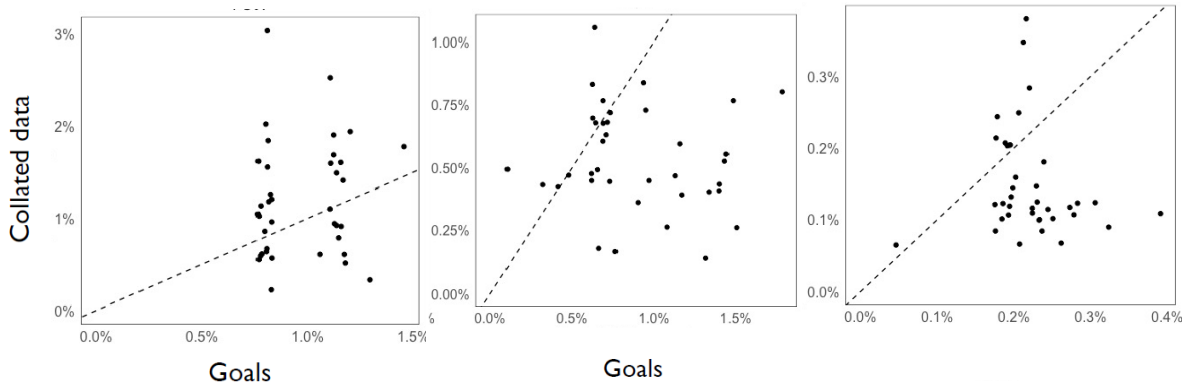
New infections among clients/partners of KP are calculated like Sabin described above. This assumption, as well as the extrapolation to regional medians for countries without estimates, resulted in some countries having more total new infections among KPs than estimated by Spectrum for the total population.

Sabin noted that the reason for bringing these results and assumptions to the Reference Group is to get feedback on whether these time-trend results are robust enough to be included in the 2022 World AIDS day report (with final results due a week after the Reference Group meeting).

Key points from discussion:

- John Stover noted that the new approach presented by Korenromp is a big improvement but recognises that the substantial change in results needs careful consideration before going to publication and that time should be spent to ensure that the best possible estimates are presented.
- Tim Brown noted that turn-over to general population from key populations may result in new infections in the general population, while these infections may actually result from people who were formerly in key populations.
- Going forward, a small working group will be set up to consider methodology and results of the time-dynamic donuts, with the aim to produce robust estimates for the July 2022 UNAIDS report.

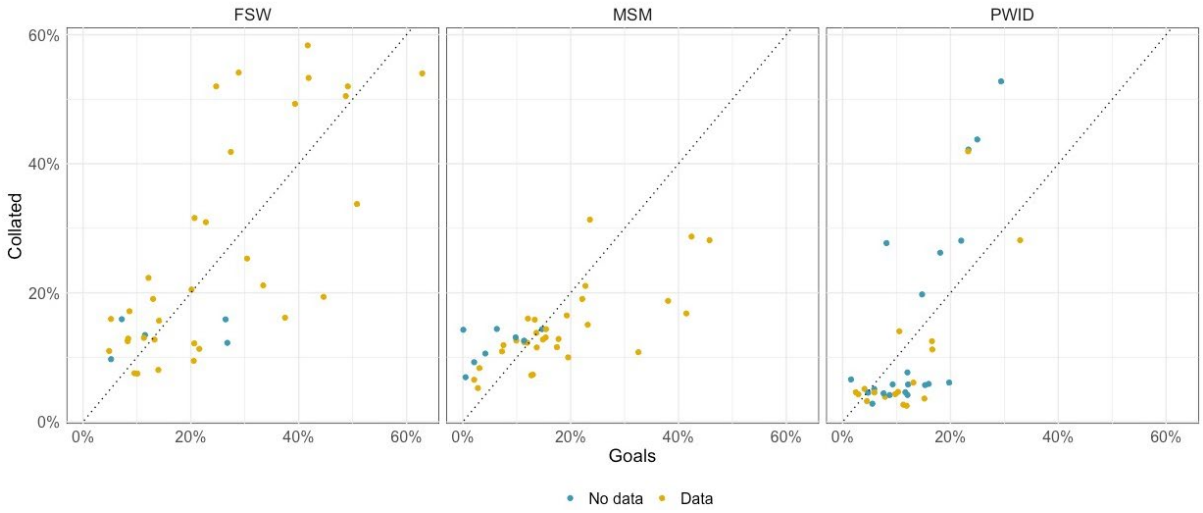
Oli Stevens next presented a **comparison of collated KP estimates with Goals estimates**.
Population Size Estimates
The data collation exercise estimated cross-sectional values for 2020, using data from 2010 onwards. Data was spatially smoothed to provide estimates for countries without surveillance data. Goals, which is a transmission dynamic model, initiates a population in 1970, with specified KP population proportions. Mortality, recruitment and turn-over determines the KP population sizes after 1970, and no calibration to country-level data is performed. The figures below for FSW, MSM and PWID show a wider range of population proportions of FSW from the collated data exercise than Goals and the proportions for MSM and PWID are generally larger for Goals than the collated data exercise.



FSW **MSM** **PWID**

HIV prevalence estimates

The data collation exercise regressed KP HIV prevalence against adult population prevalence and when a country had HIV prevalence data, that country could deviate from the regional trend, whereas if a country had no data, their estimate would lie exactly on the regional trend. By comparison, Goals calibrates to KP prevalence data over time obtained from the KP Atlas, or when working with countries on national strategic plans or investment cases. For FSW, nearly all countries have surveillance estimates of prevalence (yellow) and estimates between the two methods are similar. For MSM and PWID data are much more sparse (blue means missing data and regional values were taken) and in general this result in the estimates from Goals having wider ranges than the estimates from the collated data exercise.

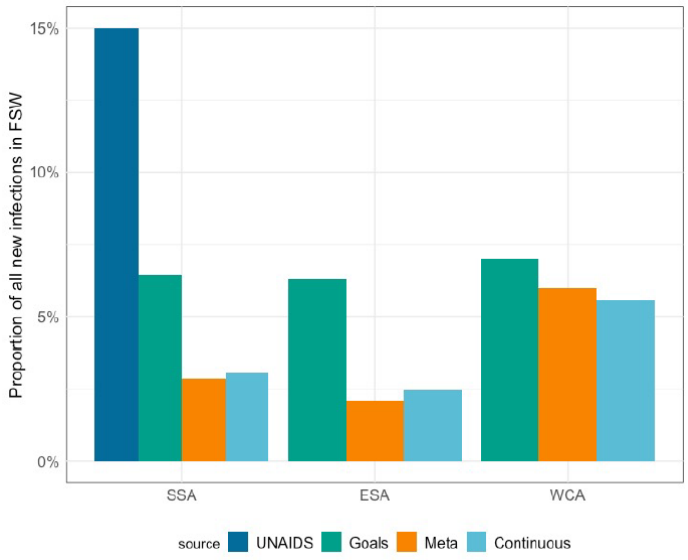


Stevens also presented on estimates of **the distribution of new infections among KPs in SSA**. The objective of the presentation was to investigate consistency between UNAIDS reported estimates of distribution of new infections (the donuts) and empirical KP data and modelling estimates. As Sabin highlighted earlier, the proportion of new infections attributed to new infections among KPs have increased over time, and the substantial increase in SSA (in ESA from 17% in 2017 to 51% in 2021) warrants further investigation.

Incidence data in SSA are quite sparse, and this work extrapolates existing data to make estimates of KP incidence for all countries in sub-Saharan Africa. Two methods (a regional fixed effect meta-analysis and a continuous log-linear model) for extrapolating data from systematic reviews of directly observed incidence (by James Stannah for MSM and by Harriet Jones for FSW) were used to produce KP incidence rate ratios (IRRs) and then these IRRs were applied to the whole region, using the HIV negative population size estimates for key populations.

Female Sex Workers

Combining data from Jones' systematic review (up to 2019) and more recent studies into a regional fixed effect meta-analysis, IRRs for FSW compared to total females aged 15-49 are estimated as ~5 in ESA and ~9 in WCA. Using regression methods (a continuous log-linear model), IRRs in countries with low total female incidence are estimated to be ~10 and in settings with high incidence about ~5. Since the estimates for the two methods are similar, and the continuous method does not dichotomise countries into ESA and WCA, the continuous method was used to compare to Goals estimates. The median estimate of incidence using the continuous method is 1% (IQR 0.5-2%) while the Goals median estimate is 2% (IQR 1-5%) which translates into proportions of new infections of median 4% (IQR 3-7%) and 8% (IQR 5-14).

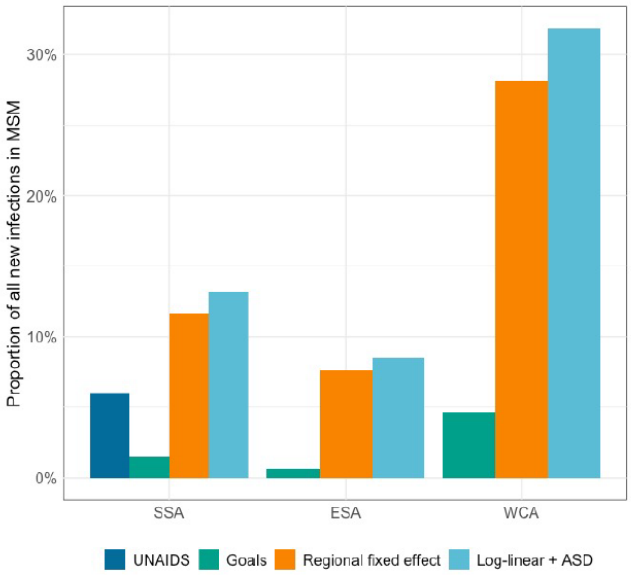


UNAIDS estimated in 2022 that **15%** of new infections in SSA occurred among FSW, while Goals estimated 6.5% and the two extrapolation methods estimated 2-3%. Back-calculating from UNAIDS' estimate of ~690,000 new infections in SSA and 2.1 million HIV-negative FSW as estimated by the data collation exercise (results shown at May meeting), the UNAIDS estimate translates to a SSA FSW HIV incidence of **5%** (vs 1% from data extrapolation and 2% from Goals).

Stevens compared results for only Malawi with another transmission dynamic model – Optima. This model's estimates of PSE of FSW in Malawi were similar to the data collation exercise (but double that of Goals) and prevalence was similar to Goals, but 10% lower than the data collation exercise at 40%. Estimates of proportion of new infections and incidence were not similar between the models or the extrapolation methods.

Men who have sex with men

Combining data from Stannah's systematic review into a regional fixed effect meta-analysis, IRRs for MSM compared to total males aged **15-29** are estimated as ~25 in ESA and ~150 in WCA. However, for this key population, regression methods (adjusting for age of sexual debut) do not produce similar estimates. In this case, IRRs are ~2000 fold in low total incidence settings and ~30 fold in high incidence settings and there are poor correlation of incidence estimates using the two methods. Both methods also poorly correlate with the Goals model's incidence estimates.



UNAIDS estimated in 2022 that 6% of new infections in SSA occurred among MSM, while Goals estimated 1.4%. The two extrapolation methods give implausibly high estimates that are inconsistent with prevalence among MSM, which may indicate bias in the risk of men enrolled in the cohort studies.

For Malawi, PSE (~40000) and HIV prevalence (~13%) were similar in the data collation exercise, Goals and Optima, but estimates of incidence and proportion of all new infections were implausibly higher with the two extrapolation methods than with Goals/Optima.

People who inject drugs
No empirical estimates on PWID incidence in SSA exist. Estimates of prevalence among PWID from the data collation exercise are lower than estimates used by Goals; and the estimate of proportion of all new infections in PWID is ~1% in Goals and ~3% by UNAIDS.

	Regional FE	Continuous	Goals	UNAIDS
FSW incidence	SSA: 6x 1% ESA: 5x WCA: 9x	5-10x 1.2%	2%	~5%
FSW prop. new. inf.	2.5%	3%	6.5%	15%
MSM incidence	SSA: 50x 5% ESA: 25x WCA: 150x	10-2000x 6%	0.5%	~4%
MSM prop. new. inf.	20%	17%	1.4%	6%

Stevens summarised his findings as:

- UNAIDS estimates of the proportion of new infections among FSW are implausibly high and not consistent with prevalence estimates.
- UNAIDS estimates of the proportion of new infections among MSM seems plausible but should be reconciled with MSM HIV prevalence and age dynamics of MSM surveillance.
- The lack of data for other key populations (PWID, transgender women, partners of KPs) raises serious concerns in the estimates of the proportion of new infections among these groups.

His proposed way forward with this work is that estimates from transmission dynamic models (preferably with age structure) should underpin the UNAIDS estimates.

Key points from discussion:

- Shufang Zhang, speaking from the viewpoint of a funder (GFATM), mentioned that many countries' national strategic plans rely on Goals, and whether these results should raise concern. To address this concern, a recommendation from this session is that Goals should be recalibrated using data from the collation exercise as inputs.
- Mahy notes the adjustment that Stevens made to compare MSM incidence from studies to total incidence among males aged 15-29 instead of 15-49, and that adjustments for age group in survey and surveillance data should also be considered in the dynamical models.

Jeff Eaton ended the meeting with a presentation on the recommendations, captured in **Appendix A**.

Appendix A

Recommendations

Recommendation	Lead person(s)	Timeline
Session 1: Incorporating WPP 2022 into Spectrum, EPP and Naomi models (chaired by Mathieu Maheu-Giroux) **Objectives:** • **Review impact of updates to tools to match new WPP methods and estimates**		
Endorse and implement proposed changes to Spectrum demographic model to align to WPP 2022. • Transition to 1x1 year WPP 2022 demographic inputs. • Retain fertility among age 15-49 female population. Add fertility from 10-14y (0.4% births) and 50-54y (0.04% births) into 15-19y and 45-49y age group, respectively. • Switch to adding net-migrants at end of projection year (after HIV progression and mortality).	Avenir Health, Imperial, East-West Centre	November 2022
Change reporting dates and interpretation of Spectrum indicators: • 'Stock' outputs (population size, PLHIV, number on ART, ART coverage) correspond to **end year** (December 31) of the reported year. • 'Event' outputs (new infections, deaths, births, diagnoses, treatment initiations) correspond to **calendar year** (Jan 1 – Dec 31). • 'Rate' outputs (incidence rate, mortality rate) will be calculated by number of events during calendar year divided by **mid-year** population denominator as approximation for exposure. This is a change from 2022 Spectrum practice, which used population at start of exposure period as denominator.	Avenir Health, Imperial, East-West, McGill	November 2022
Transition to single-age fertility rates from WPP 2022, instead of constant fertility rates by 5-year age group.	Avenir Health	May 2023
ECDC incidence rates already correspond to calendar years. No changes required.	No action	
For countries using subnational Spectrum files (Ethiopia, Kenya, Mozambique, Nigeria, Zimbabwe), review and update subnational demographic inputs to align with WPP2022 and other recent subnational demographic data (e.g., subnational fertility from surveys, population distribution from recent censuses).	Avenir Health, Imperial, US Census Bureau, UNAIDS	December 2022

Recommendation	Lead person(s)	Timeline
Check if there are any large changes in numbers of births in WPP 2022 that will potentially affect MTCT or paediatric outputs for any countries.	Avenir Health	November 2022
In countries where HIV is increasing (e.g., Philippines), shift to calendar year may make a significant difference to estimates of ART coverage. Run all 2022 files using the new DemProj inputs to identify such countries for focused communication.	Avenir Health	November 2022
Clearly communicate changes in interpretation of Spectrum indicators related calendar-year projection to users and stakeholders.	UNAIDS	December 2022
Session 2: On-ART mortality (chaired by Jeff Eaton) **Objectives:** • **Recommendations for update Spectrum on-ART mortality in high income countries** • **Recommendations for update Spectrum on-ART mortality in SSA**		
Mortality on ART in **sub-Saharan Africa**: assume decline from 1.0 to 0.95 between 2017-2021 and no change after 2021. This is a smaller decline than in 2022 estimates, which extrapolated from 1.0 to 0.9 during 2017-2021 and flatlined only after 2021. The basis for the revised assumption is <u>preliminary analysis</u> of new IeDEA data including the 2018-2021 period, which found no evidence of ongoing mortality decline in SSA. Conditional on: • Further feedback and interpretation from IeDEA Executive Committee. • Review of mortality among persons in IeDEA cohorts without CD4 at initiation, to assess whether selection bias affected apparent mortality time trends. • Consideration of potential contribution of Covid-19 related mortality to higher mortality in 2020-2021. • Triangulation of trends from South Africa vital registration data and IeDEA South Africa analysis.	Avenir Health (update Spectrum) Reshma Kassanjee Reshma Kassanjee Reshma Kassanjee Leigh Johnson	November 2022 November 2022 May 2023 May 2023 May 2023
Mortality on ART in regions **outside sub-Saharan Africa**: assume continued mortality decline at same rate as assumed over 2017-2021, based on <u>preliminary analysis</u> of new IeDEA data. Conditional on: • Further feedback and interpretation from IeDEA Executive Committee.	Avenir Health Reshma Kassanjee	November 2022 November 2022

Recommendation	Lead person(s)	Timeline
Update assessment of mortality on-ART rates from **all IeDEA regions,** including further analysis of: • Mortality trends during 2020-21 period, including potential impact of Covid-19 related mortality or increased AIDS mortality. • Effects of incomplete data CD4 at diagnosis on mortality levels and trends.	Reshma Kassanjee, Renee De Waal	April 2023
Mortality on ART in **Asia Pacific region**: increase the 2022 default on ART mortality rates by multiplying by mortality adjustment factors that (select) countries used in 2022 to match national AIDS deaths data. This is anticipated to entail an around 4x increase in on ART mortality. Use this as new default for 2023 estimation round, but qualified as interim, pending new rate estimates from IeDEA Asia-Pacific Region (expected 2023).	Avenir Health, East-West Centre	November 2022
Communicate change to defaults and potential need for further country-specific adjustment to teams in **Asia-Pacific region**, notably the countries that had used a mortality adjustment factor (but should now no longer need it) and high-income countries where higher rates may not be appropriate.	UNAIDS	January 2023
Conduct further analysis to characterise ART mortality patterns in **Asia-Pacific Region**, including: • Mortality rates limited to cohorts with strong vital registration linkage, including adjustment for deaths hidden in Loss to Follow-up. • Stratification according to low- versus high-income country status.	Reshma Kassanjee, Renee De Waal	April 2023
Update default on ART mortality rates for **WCENA region** to reflect **excess mortality** rate among PLHIV, based on total mortality observed in ART-CC cohorts minus background total population non-HIV mortality.	Avenir Health	November 2022
Maintain Spectrum's operationalization to estimate **total excess mortality** among PLHIV, rather than only AIDS-related mortality. This is dictated by the method and source of the mortality rates, as total mortality in clinical ART cohorts minus background mortality.	(No action)	

Recommendation	Lead person(s)	Timeline
The approach that Spectrum ART mortality parameters reflect excess mortality among people with HIV and not AIDS cause specific mortality will increasingly become an issue in **all regions** of the world. UNAIDS need to work on messaging around this. Notably for WCENA region, where evidence from ART-CC indicates most excess mortality among PLHIV is now due to non-AIDS causes.	UNAIDS	Training materials & workshops; Global Report Methods annex; etc.
Develop TOR for proposals to clarify definitions and assumptions for estimating AIDS deaths including a review of global data on excess non-AIDS death among people on ART and making recommendations on a modelling approach.	Secretariat	TOR: January 2023 Results: May 2023
Encourage **Asia-Pacific and WCENA** country estimates teams to input into Spectrum **data on ART by age** and **fit incidence rate ratios** to those, to improve projected mortality among PLHIV.	UNAIDS	January 2023
Encourage **Asia Pacific** countries to input **AIDS deaths by age,** for specifying mortality adjustment factor and triangulating mortality age distribution.	UNAIDS	January 2023
Create an input editor in Spectrum for countries to specify numbers of AIDS deaths and non-AIDS deaths among PLHIV, if available from national case surveillance (anticipated: primarily European countries). Display these, and have Spectrum produce a corresponding split into AIDS versus non-AIDS causes in Spectrum-estimated mortality among PLHIV, which users can choose to communicate instead of the default overall excess mortality. (Spectrum already has output on AIDS and non-AIDS deaths among PLHIV. Isn't this recommendation supposed to be about AIDS and non-AIDS deaths among people on ART?)	Avenir Health	November 2022
For WCENA: following (1) improved alignment of sex/age distribution of ART population in Spectrum to programme data, and (2) collation of data on AIDS versus non-AIDS deaths among PLHIV, conduct further triangulation of Spectrum-estimated AIDS deaths compared to vital registration data. Based on 2023 data and estimates, consider updating default assumptions on AIDS and excess non-HIV mortality.	UNAIDS	May 2023

Recommendation	Lead person(s)	Timeline
PRELIMINARY RECOMMENDATION: Adjust default 'Developed country' ART mortality rate for CD4 >500 at initiation category to be equal to mortality rate of the CD4 350-500 category – instead of the 2022 default with mortality rate among CD4 >500 category above the CD4 350-500 category. Because regression analysis of ART-CC data did not confirm such a difference, which may have reflected unadjusted confounding of due to selective ART initiation at high CD4 counts of persons with symptoms. Conditional on: • Pre-testing if this indeed adequately redresses too high projected mortality for countries like Italy. • In parallel, further analyse ART-CC data to assess impact of ART eligibility, VL or symptoms at ART initiation on mortality in high CD4 categories.	Avenir Health Kelsey Case, UNAIDS Adam Trickey	November 2022 November 2022 April 2023 (if needed)
Continue development of alternative AIDS mortality modelling strategies, including: • A (time-and country-varying) effect of VL suppression on ART mortality. • Reducing reliance on data about CD4 at ART initiation. • Considering local mortality data or other correlates of ART programme effectiveness.	Working group	May 2023
Session 3: Improving prevalence measures from ANC-DQAs (chaired by Leigh Johnson) **Objectives:** • **Guidance about when to include/exclude ANC-RT data**		
Develop a protocol for countries to undertake focused data collection at a selection of ANC sites are to improve the quality of routine ANC data for inputs to HIV estimates (and audit routine data quality to strengthen services). Recommendations are: • To select a representative sample of antenatal clinics. • To set up individual-level electronic systems to collect prospectively key HIV variables for all pregnant women (integrated into existing EMR systems, as far as possible). • To compare individual-level data to data from routine monitoring and identify sources of discrepancies with routine reporting systems. The emphasis should be on supporting real-time monitoring of quality data capture, rather than periodic (typically annual) retrospective review of data quality.	Mary Mahy, Jeff Eaton	December 2022

Recommendation	Lead person(s)	Timeline
Prior to new data from ANC, we encourage cleaning of ANC data. Suggested checks include: • Dropping years/periods in which more than 10% of data are missing or implausible (e.g., more than 100% of HIV-positive women on ART). • Excluding data with implausibly large changes from one year to the next (e.g., in HIV prevalence among pregnant women). • Reviewing trends in new ANC clients and facility births to identify any evidence of changing trends in ANC attendance and reporting and coverage of facility birth. Pilot these data cleaning checks and their implications for estimates in a few countries. Suggested countries were Zimbabwe, Uganda and Mozambique.	UNAIDS	2023 estimates
HIV-related fertility reduction in concentrated epidemic settings: further seek and consider evidence about whether the default HIV fertility reduction in concentrated epidemics should be smaller than patterns derived from survey data in sub-Saharan African. (See further evidence on this in October 2022 paediatric HIV estimates meeting.)	Avenir Health, Imperial, UNAIDS	November 2022
Change the Spectrum interface to include "numbers of known negatives" in the ANC testing table (with a footnote explaining, this is not an essential input, i.e., recognizing that most countries don't collect this information). Do not add a row for "Women with unknown HIV status".	Avenir Health	November 2022
<u>Recommendations from May 2022 still to be addressed:</u> Develop a simple screening tool to identify health facilities with probable data quality problems, based on completeness of data, variability of results over time, and consistency across indicators (ratios <100%). • Validate screening tool using a classification of facility reporting quality (example: Rwanda ANC data quality review).	TBD	May 2023
Recommendation	**Lead person(s)**	**Timeline**
Session 4: Test and treatment churn (chaired by Josh Salomon) **Objectives:** • **Reach consensus about rejection of EPP fits with ART coverage over 100%** • **Recommendation about proposals for calibrating IRRs to survey ART coverage data**		
Waterfall: • Present both waterfall cascade (new initiations, treatment interruptions, deaths) from programme data and estimated by	Avenir Health	November 2022

Recommendation	Lead person(s)	Timeline
Spectrum to compare observed programmatic outcomes with model estimates of treatment programme changes. • Include transfer in and transfer out in in the waterfall categories. These are important for reconciling treatment programme changes in subnational files.		
Calibrating IRRs by age to survey ART coverage data: • Review recommendation to fit time-constant sex/age incidence rate ratios or time-varying IRRs (which changed the fitted IRR little, but increased fitting time), considering large number of additional household surveys since the previous analysis underpinning this recommendation in 2017.	Avenir Health	May 2023
Do not implement change in EPP fitting to enforce PLHIV greater than the input number on ART • Although this approach enforces consistency with the ART programme data, values of just under 100% may still be implausibly high and may not be more useful for planning than a value of 110%. This approach still does not allow a robust estimate of the gap in treatment coverage and although it highlights a problem to the user, there are no convenient tools to understand or fix the problem. Add a check in the EPP ART distribution page to notify the user if the number on ART in a region is greater than the PLHIV from the previously saved EPP fit.	(No action) East-West Centre	 November 2022
<u>Recommendations from May 2022 still to be addressed:</u> Develop example of using DQA results to adjust spatial ART by district inputs in Naomi.	Imperial	May 2023
ART data quality assessments: • Strongly encourage countries without recent DQA of ART data to conduct these; propose standardized methodology • Stratifying DQA results by sex/age/location is important.	UNAIDS	2023 estimates
Further disaggregate HIV testing data inputting to Shiny90: • self-testing • index testing • HCT by sex / age (GAM age groups) (We propose to do this by adding a new editor in the Program Statistics section. Then the data can be read by Shiny90 or any other component that needs them. If this is also a file in ADR, we can read from ADR into Spectrum editor.)	Avenir Health, McGill, Imperial, Fjelltop	2023 estimates

Recommendation	Lead person(s)	Timeline
Session 5: HIV-positive Migration in AIM & Enhanced User guidance for concentrated epidemics (chaired by Cari van Schalkwyk) **Objectives:** • **Review and make recommendations for proposed guidelines to account for migration of people with HIV** • **Review user guidance for concentrated epidemics**		
If users input HIV-positive migrant data to AIM, AIM should not apply its background calculation of net HIV-positive migration. If migrant data have intermediate years missing (or missing age/sex), users should impute values and record the imputation in Excel (building on data compilation and analysis XLS created for all high-income and all other countries with migration in 2022). Still to investigate: Should AIM also suppress calculated net *e*-migration for years applicable, given that country data record immigrants but do not subtract emigrants?	Avenir Health Avenir Health / UNAIDS	December 2022
Data on migrants living with HIV will only be entered in AIM, not CSAVR (which fits to all new diagnoses, including first-time diagnoses among immigrants). CSAVR pulls immigrant numbers from AIM, to add into KOS (assuming all were diagnosed) and into mortality.	Avenir Health	December 2022
Add as new result, HIV-positive migrants over time, so users can check and ensure a smooth sensible modelled pattern (from Spectrum combining user entries for some and background calculation for other years).	Avenir Health	November 2022
Investigate the use of Spectrum's calculated age/sex distribution of HIV-positive migrants to disaggregate country HIV-positive immigrant data.	Avenir Health	May 2023
Review the use of EPP's External HIV option. If used, then add recommendation that users should add the External HIV numbers (by subpopulation EPP-fitted) also into AIM (by age/sex).	UNAIDS	December 2022
Investigate utility of comparing/triangulating/validating AEM's low-risk female HIV prevalence estimate against ANC data. Obtain and use time series by age and sex in two or three countries to test.	New Working Group, led by East-West Centre	May 2023
Add 8 years' case diagnosis data, in addition to 8 years of HIV/AIDS death data, as a requirement to publish a CSAVR-based incidence trend.	UNAIDS	December 2022

Recommendation	Lead person(s)	Timeline
Session 6: Key populations in concentrated epidemics (chaired by Rob Glaubius) **Objectives:** • **Get feedback on interpretation of KP data and CSAVR-KP design**		
Pilot study of KP-CSAVR in four countries should continue, completing the ongoing model fitting and interrogating drivers of unexpected results in initial model fits presented. • This should include triangulation of model estimates against non-fitted sources of data (e.g., seroconversions in HIV prevention programs) if available and appropriate.	Avenir Health	May 2023
Prevalence and KP size data used in 2022-round country-approved EPP and AEM files can be used by CSAVR; ongoing review of sentinel surveillance and survey data in these concentrated epidemic settings is recommended, for use in any of these models.	Avenir Health	May 2023
<u>Recommendations from May 2022 still to be addressed:</u> Develop guidance for countries with concentrated epidemics on documenting test volumes and case diagnoses by mode of transmission, for 2024 Spectrum estimates and more broadly for meeting strategic priorities.	UNAIDS	Introduce at 2023 regional workshops
Session 7: CSAVR (chaired by John Stover) **Objectives:** • **Review CSAVR interface and methods updates**		
CSAVR should in some ways serve as database to document, quality-check, compare and select alternate data sets and/or points most valid for fitting		
• Priority for adding entry for annual Test volumes, and calculation of Test Positivity Rate, to help interpret fluctuations in case diagnoses. • Do not yet add functionality to enter and select alternative death data sets, or suppress selected unrepresentative data points (e.g., CD4 distributions among <x% of new diagnoses).	Avenir Health	November 2022
• **<u>Recommendation retained from May 2022:</u>** Explore using data on HIV testing volume and positivity in CSAVR fitting, to distinguish effects on diagnoses and knowledge of status from possible changes in new infections versus changes in testing effort.	Avenir Health, UNAIDS	May 2023
Recommend that CSAVR send percentage not on ART aware of status to AIM, as basis for AIM's calculation of KOS numbers consistent with ART numbers.	Avenir Health	November 2022

Recommendation	Lead person(s)	Timeline
AIM projects and outputs based on IDEA mortality definitions and estimates) *excess* mortality to PLHIV (AIDS and comorbidities). CSAVR uses these to fit AIM-based *excess* mortality to PLHIV, against VR (or IHME/GBD) *data of AIDS-attributable* deaths. Assess magnitude of the resulting underestimation of epidemics in CSAVR – including effects on CD4 at diagnoses and KOS.	Avenir Health, UNAIDS	May 2023
CD4 at diagnosis – further investigation of CSAVR's overestimation of CD4 (and of KOS), including of bias in CD4 data due to incomplete selective testing, the utility and mechanisms of adjusting CD4 fitting for such bias.	Avenir Health	May 2023 (secondary to DQA functionality, mortality definitions CSAVR-KP)
Session 8: Key populations in sub-Saharan Africa (chaired by Mary Mahy) **Objectives:** • **Recommendations on workbook process and tool changes for 2023**		
Update Goals and Optima calibrations with KP HIV prevalence and population size data from KP workbook data synthesis.	Imperial, Avenir, Burnett Institute	November 2022
Recommend caution about interpretation and next publications of refined but provisional KP shares in new infections ('donut') results. Establish sub-group to consolidate methods and estimates for distribution of new infections across population groups over time. • For MSM, adjust surveillance and survey data for age group, before using in AIM or Goals calibration. • Refine assumptions and values for onward transmission from KP to clients and partners. • Review 'turnover' rates about duration in key population groups and analyse how turnover assumptions affect incidence-prevalence ratios in KPs, clients and lower-risk groups. • Develop methods to estimate of proportion of transmissions that arise from individuals who acquired HIV as members of KPs and turned-over to lower-risk groups before prevalence measurement.	Working Group	
<u>Recommendations from May 2022 still to be addressed:</u>		

Recommendation	Lead person(s)	Timeline
Develop TOR for proposals for further development of the time-dynamic IPM concept as progress towards the Symphony model that consolidates all data sources, transmission dynamics, burden estimates and prevention and clinical impact, across multiple countries and epidemic types.	Reference Group Secretariat	TOR: January 2023 Proposals: May 2023

Appendix B

Participants

Name	Organisation
Adam Trickey	ART-CC
Adrien Allorant	McGill University
Andreas Jahn	MoH Malawi
Andrew Phillips	UCL
Ard van Sighem	Stichting HIV Monitoring
Avi Hakim	CDC
Brian Rice	LSHTM
Cari van Schalkwyk	SACEMA
Deepa Jahagirdar	Stanford University
Ehounoud Pascal Eby	UNAIDS
Eline Korenromp	UNAIDS
Faikah Bruce	SACEMA
Gatien Ekanmian	UNAIDS
Guy Mahiane	Avenir Health
Haroon Moolla	IeDEA UCT
Hind El Hajji	UNAIDS
Hmwe Kyu	IHME
Houssine El-Rhilani	UNAIDS (Marocco)
Ian Wanyeki	UNAIDS
Irum Zaidi	PEPFAR
Isaac Taramusi	MoH Zimbabwe
J Hoener	USAID
Jeff Eaton	Imperial College London
Jinkou Zhao	GFATM
John Stover	Avenir Health
Josh Salomon	Stanford University
Juliana Daher	UNAIDS
Katharine Kripke	Avenir Health
Keith Sabin	UNAIDS
Kelsey Case	Consultant
Khin Cho Win Htin	UNAIDS
Lahoucine Ouarsas	ALCS
Laura Porter	CDC
Leigh Johnson	University of Cape Town
Lev Zohrabyan	UNAIDS
Lucy Platt	LSHTM
Luisa Frescura	UNAIDS
Maggie Walters	Imperial College London
Makini Boothe	UNAIDS
Mark Wheldon	UN Pop Division
Mary Ann Seday	UNAIDS
Mary Felissa Reyes Vega	MoH/CDC Peru

Mary Mahy	UNAIDS
Mathieu Maheu-Giroux	McGill University
Mehran Hosseini	GFATM
Melaku Dessie	USAID
Michelle Morrison	Gates Foundation
Michelle Williams–Sherlock	CDC
Monita Patel	CDC
Nikos Pantazis	National and Kapodistrian University of Athens
Nina Anderegg	University of Bern / University of Cape Town
Noah Bartlett	CENSUS/POP FED
Nora Springstubb	CDC
Oli Stevens	Imperial College London
Parviez Hosseini	PEPFAR
Patricia Bracamonte	UNAIDS (Peru)
Patrick Gerland	UN Pop Division
Rachel Esra	Imperial College London
Ray Shiraishi	CDC
Renee De Waal	CIDER
Reshma Bhattacharjee	USAID
Reshma Kassanjee	University of Cape Town
Rob Glaubius	Avenir Health
Rowan Martin-Hughes	Burnet
Roza Babayan	UNAIDS (Armenia)
Sadhna Patel	CDC
Samuel I Dupre	CENSUS/POP FED
Sara Hertog	UN Pop Division
Shona Dalal	WHO
Shufang Zhang	GFATM
Thomas Spoorenberg	UNAIDS
Tim Brown	East West Center
Tim Wolock	Imperial College London
Tobi Saidel	Consultant
Wilford Kirungi	MoH Uganda
Will Probert	WHO
William Miller	USAID
Wolfgang Hladik	CDC
Yoko Shimada	Consultant

Appendix C

Agenda (all times are in GMT+2)

Day 1

Time	Duration (mins)	Topic	Presenter(s)/ Lead Discussant
14.00	15	Welcome and introductions 2022/23 estimates round: process and timelines	Mary Mahy
14.15	5	Meeting objectives and recommendation review	Cari van Schalkwyk
Session 1: Incorporating WPP 2022 into Spectrum, EPP, and Naomi models (chaired by Mathieu Maheu-Giroux) **Objectives:** • **Review impact of updates to tools to match new WPP methods and estimates**			
14.20	20	Spectrum, CSAVR, Shiny90	Rob Glaubius
14.40	10	EPP	Tim Brown
14.50	5	Naomi	Jeff Eaton
14.55	45	Discussion	
15.40	5	BREAK	
Session 2: On-ART mortality (chaired by Jeff Eaton) **Objectives:** • **Recommendations for update Spectrum on-ART mortality in high income countries** • **Recommendations for update Spectrum on-ART mortality in SSA**			
15.45	15	IeDEA updates • Revised parameters for all regions • Mortality among cohorts with and without baseline CD4 measurement • Mortality after interruption	Reshma Kassanjee
16.00	10	Results of IeDEA-AP survey about ascertainment of deaths	Renee de Waal
16.10	10	Comparison of IeDEA-AP data to ART programme data	Tim Brown
16.20	20	ART-CC: AIDS and non-AIDS mortality by sex/age. • Compare mortality from non-AIDS causes with population-wide mortality from these causes • Assess what share of excess on-ART mortality is attributable to AIDS vs. other causes • Assess whether this can explain overestimating AIDS deaths vs. VR data on AIDS cause of death Additional ART-CC analyses: • Review CD4>500 mortality rate • Impact of CD4 and VL measurement inclusion criteria on mortality	Adam Trickey
16.40	10	Comparison of ART-CC data to Spectrum output	Rob Glaubius
16.50	15	Comparison of crude on-ART mortality rates pre-2016 between Spectrum and IeDEA by age, sex, CD4 and ART duration	John Stover
17.05	5	Spectrum update: • Output for all-cause deaths among persons on ART (complementing the existing, HIV/AIDS-related deaths). With	John Stover

		corresponding validation plot, comparing this Spectrum estimate with country-inputted all-cause on-ART deaths.	
17.10	5	VLS as driver of on-ART mortality: simulation exercise	Leigh Johnson
17.15	45	Discussion	
18.00		CLOSE	

Day 2

Time	Duration (mins)	Topic	Presenter(s)/ Lead Discussant
Session 3: Improving prevalence measures from ANC-DQAs (chaired by Leigh Johnson) **Objectives:** • **Guidance about when to include/exclude ANC-RT data**			
14.00	20	Improving the quality of ANC data: Protocol development and plans	Mary Mahy
14.20	10	Guidance on the correct indicators that should be reported on	Ian Wanyeki
14.30	10	Update from Mozambique DQA	Makini Boothe
14.40	5	Spectrum updates • Add outcome 'total live births in health facilities' and 'known HIV negative' to ANC testing inputs in Spectrum/Naomi • Remove the HIV-related fertility adjuster data editor, to instead read-in ANC prevalence and denominator from the (national-level) ANC testing editor	John Stover
14.45	45	Discussion	
15.30	5	BREAK	
Session 4: Test and treatment churn (chaired by Josh Salomon) **Objectives:** • **Reach consensus about rejection of EPP fits with ART coverage over 100%** • **Recommendation about proposals for calibrating IRRs to survey ART coverage data**			
15.35	5	Spectrum updates: • Visualisation of 'waterfall' results estimated by Spectrum • ART DQA results - show and store both reported ART data and adjusted ART data time series after DQA	John Stover
15.40	10	Calibrating IRRs to survey ART coverage data	Rob Glaubius
15.50	10	Rejecting EPP fits with ART coverage over 100%	Jeff Eaton
16.00	10	Current levels of electronic medical records data in countries	Ian Wanyeki
16.10	40	Discussion	
Session 5: HIV+ Migration in AIM & Enhanced User guidance for concentrated epidemics (chaired by Cari van Schalkwyk) **Objectives:** • **Review and make recommendations for proposed guidelines to account for migration of people with HIV** • **Review user guidance for concentrated epidemics**			
16.50	20	Migration of people with HIV: • Update from Migration Working Group • Progress refining AIM and aligning CSAVR, including default age/sex patterns from available HIV+ immigrant data	Eline Korenromp Rob Glaubius
17.10	15	User guidance for concentrated epidemics: Model choice & triangulations, Knowledge of Status methods, Demographic scope of national estimates, migration, CSAVR inputs.	Eline Korenromp

17.25	35	Discussion	
18.00	CLOSE		

Day 3

Time	Duration (mins)	Topic	Presenter(s)/ Lead Discussant
Session 6: Key populations in concentrated epidemics (chaired by Rob Glaubius) **Objectives:** • **Get feedback on interpretation of KP data and CSAVR-KP design**			
14.00	5	Model Pilot: Selection of 4 countries and status	Keith Sabin
14.05	10	Data collection and challenges in Cambodia	Khin Cho Win Htin
14.15	20	Data quality, consistency, and gaps: impressions from fitting the original CSAVR model for Cambodia, Morocco and Peru	Eline Korenromp
14.35	10	CSAVR-KP: initial fit to pilot countries and model refinement	Guy Mahiane
14.45	30	Discussion	
15.15	5	BREAK	
Session 7: CSAVR (chaired by John Stover) **Objectives:** • **Review CSAVR interface and methods updates**			
15.20	10	CSAVR updates: • Visualisation of HIV testing data in CSAVR • Allow for CSAVR entry and saving of mortality data from multiple sources (original VR, misclassification adjusted GBD, etc.) in parallel, and choose one – or a user-defined hybrid – for fitting • Enable CSAVR data editors to retain data on diagnoses, deaths and CD4 for all years, but selectively exclude some data points from model fitting • To avoid people on ART outnumbering PLHIV with knowledge of status, CSAVR should send to AIM proportions of PLHIV *not on ART* knowing their status, rather than overall proportions knowing their status.	Guy Mahiane
15.30	15	CSAVR development: Improving fits to CD4 distribution at diagnosis	Guy Mahiane
15.45	30	Discussion	
Session 8: Key populations in sub-Saharan Africa (chaired by Mary Mahy) **Objectives:** • **Recommendations on KP workbook process for 2023, including process for updating Goals model estimates with workbook data** • **Review UNAIDS 'donut' estimates for distribution of infections by key population** • **Review evidence and strategies for incidence trends among key populations over time and develop short- and medium-term workplan**			
16.15	5	Key population workbook and process in 2023 estimations	Keith Sabin
16.20	20	Review outputs of workbook data submitted in 2022 round against Goals and Optima estimates; Goals and Optima to refit to workbook data	Oli Stevens
16.40	15	Global and regional trends in estimates of KP epidemic contributions	Keith Sabin/ Eline Korenromp
16.55	30	Discussion	

17.25	35	**Recommendations**	
18.00		CLOSE	