# Technical updates for UNAIDS HIV estimation tools

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

9-12th October 2023

# Table of Contents

# Abbreviations

| | |
|---|---|
| AIM | AIDS Impact Model |
| ANC (-RT) | Antenatal Clinic (Routine Testing) |
| ART | Antiretroviral Therapy |
| ART-CC | ART Cohort Collaboration |
| ARM | Age and Risk Structured Model |
| ASM | Age Structured Model |
| CDC | US Centres for Disease Control and Prevention |
| CSAVR | Case Surveillance and Vital Registration |
| DQA | Data Quality Assessment |
| EPP | Estimation and Projection Package |
| FSW | Female Sex Worker |
| GAM | Global AIDS Monitoring |
| IBBS | Integrated Biological and Behavioural Surveillance Survey |
| IeDEA | International Epidemiology Databases to Evaluate AIDS |
| KP | Key Population |
| (K)PSE | (Key) Population Size Estimate |
| 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 |
| PrEP | Pre-Exposure Prophylaxis |
| PWID | People Who Inject Drugs |
| SSA | Sub-Saharan Africa |
| TGW | Transgender Women |
| UNAIDS | Joint United Nations Programme on HIV/AIDS |
| VLS | Viral Load Suppression |
| WHO | World Health Organization |

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**).

# 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 an online meeting on the Microsoft Teams platform, from 9 – 12 October 2023. The meeting featured presentations and plenary discussion to generate consensus recommendations. The meeting agenda was organised into the following 7 sessions:

1. Key Populations data synthesis in sub-Saharan Africa
2. Dynamical modelling of HIV trends in key populations in sub-Saharan Africa
3. ART coverage data discrepancies
4. Definitions and reporting of HIV mortality among PLHIV
5. Mortality among people on ART
6. ART interruption rates
7. Concentrated epidemics

This report contains a comprehensive summary of presentations and discussions from the meeting, forming the foundation for the Reference Group's recommendations. Meeting participants (see **Appendix B**) can access these presentations on www.epidem.org. For those not in attendance, please reach out to the Secretariat at epidem@sun.ac.za for access requests. **Appendix A** presents the final recommendations. These recommendations guide UNAIDS in the reviewing HIV estimates, assessing current methodologies, and identifying data for refining HIV estimates. The meeting agenda and objectives are in **Appendix C**. For records of previous meetings, visit www.epidem.org.

# Meeting Introduction

**Mary Mahy** opened the meeting by thanking participants for joining and providing an update on the progress made by UNAIDS. In 2023 approximately 130 countries completed estimates and approved results to be published by UNAIDS, but participation remained lower among Western European and North American countries. These countries either did not approve publication of results or did not have updated data since 2021, highlighting a persistent challenge with timely data reporting.

Existing models in Spectrum use different data to estimate HIV epidemic trends, depending on availability and context. The vision of the Reference Group since 2020 has been to develop one model for all settings that can incorporate multiple data sources. Looking ahead to 2024, Mahy noted the objectives of the Symphony model:

1. Model that incorporates multiple data sources: The model will aim to capture diverse data sources, encompassing case reporting, mortality data, empirical or measured VLS (viral load suppression), and HIV incidence data that is measured directly and representative of the population. The CSAVR model and Key Population Workbook are steps towards this objective but have not yet fully realized the Symphony model's vision.
2. Produce estimates of HIV indicators stratified by critical populations: With the emphasis of the global AIDS strategy on inequalities, estimates for critical populations by age, sex, and geographical area are essential to identify, reduce, and monitor inequalities.
3. Incorporate coverage and impact of major interventions: The model should incorporate major interventions, including VLS, pre-exposure prophylaxis (PrEP), condoms, and voluntary medical male circumcision.

There are ongoing challenges in the quality and consistency of HIV data, which have significant implications for the models and estimates. The key issues identified included:

- Discrepancies in antiretroviral treatment numbers: There are often discrepancies where ART coverage from treatment program service provision data is higher than the estimated number of people receiving ART from household surveys. This highlights the need for continued scrutiny and validation of data within countries.
- The first 95 estimate: The proportion of PLHIV aware of their status will be overestimated if ART data—and hence ART coverage—are over-reported, impacting the reliability of outputs from Shiny90 and CSAVR. The Reference Group needs to explore more accurate methods to determine the proportion of people aware or different approaches for guiding HIV testing program gaps and strategies.
- Antenatal clinic data: Longstanding challenges about accuracy of routine ANC data reporting has not been resolved. It is hoped that the continued review of ANC data through the Naomi tool will improve the accuracy of data input to models for the sub-Saharan African countries.
- Challenges determining representation of key population integrated bio-behavioural survey (IBBS) results: Key populations are often hidden or marginalized, resulting in challenges in data interpretation. The results from IBBS surveys frequently raise questions about whether they accurately reflect prevalence or ART coverage in specific urban or rural areas. Extrapolating this data to a broader context continues to be a complex task. Additionally, discerning true population size estimates compared

to those engaged by programs adds to the complexity. Another important phenomenon for modelling key population HIV epidemiology is turnover, which refers to the dynamics of people joining or leaving certain activities or roles, such as women entering or exiting sex work, or individuals starting or stopping drug use. This turnover affects the duration of engagement in these activities, which needs to be accounted for in model accuracy. Mahy highlighted the importance of innovative tools that assist country teams in reviewing and visualizing data. Such tools not only facilitate improved data input but also lead to more accurate estimates. The use of visual aids like the Key Population Workbook has proven invaluable in initiating discussions about regional prevalence differences and other critical data points.

Regarding plans for UNAIDS HIV estimates in 2024, Mahy confirmed that UNAIDS will not organize regional workshops and country teams will receive support via virtual trainings and videos. To support this, UNAIDS aims to enhance guidance documents, emphasizing the data review aspect of the process. She expressed gratitude for the continued efforts, noting the complexity of generating annual estimates and discussing model updates.

**Cari van Schalkwyk** introduced the meeting's structure and primary objectives to all participants, outlined the objectives for each meeting session.

**Session 1 - Key Populations data synthesis in sub-Saharan Africa**:
Objective: Discuss the implementation of the Triangulator tool for sub-national key population size estimates and the Aggregator tool for national level estimates.
Key focus: Updates on the Key Population Workbook, analysis of key population size trends, and urban-rural ratios in key population size estimates (KPSE).

**Session 2 - Dynamical modelling of HIV trends in key populations in sub-Saharan Africa**:
Objective: Review the development of the Goals-Age and Risk-stratified Model (Goals-ARM) as per the Reference Group's recommendation (May 2023 meeting; Stellenbosch).
Key focus: Updates on the model's development, population turnover, and its significance in the model's accuracy.

**Session 3 - ART coverage data discrepancies**:
Objective: Address disparities between ART program data and survey evidence, using specific country examples.
Key focus: Recommendations for data source selection and handling discrepancies.

**Session 4 - Definitions and reporting of HIV mortality among PLHIV**:
Objective: Examine review results covering the definitions of AIDS deaths in the context of increased ART coverage and estimates of the fraction of deaths among PLHIV that are due to AIDS.
Key focus: Discuss implications and potential recommendations following the review.

**Session 5 - Mortality among people on ART**:
Objective: Update on mortality rates observed among individuals on ART, from the IeDEA adult mortality study.
Key focus: Analysis of specific mortality trends in ART re-initiators and the broader ART cohort.

**Session 6 - ART interruption rates**:
Objective: Discuss ART interruption data to recommend non-zero regional default rates for upcoming 2024 estimates.
Key focus: Exploration of data on ART interruption and its implications.

**Session 7- Concentrated epidemics**:
Objective: Updates on EPP-concentrated and CSAVR models, and the ECDC model, to refine data collection and analysis methodologies in the context of concentrated epidemics.
Key focus: Challenges in key population data and estimates, especially PSE, prevalence, and ART coverage, specifically focusing on the challenges and nuances associated with non-EPP countries.

# Session 1: Key Populations data synthesis in sub-Saharan Africa

**Chair Jeff Imai-Eaton** highlighted the central objectives for Session 1:

- Recommendations on use of the Triangulator tool in the key population estimates process in sub-Saharan Africa (SSA)
- Review of data on population size estimates trends and urban/rural population size estimate proportion ratios

Imai-Eaton underlined the significance of incorporating a thorough review and assimilation of key population data into the HIV estimates process, a top priority for the Reference Group over the past two years. The Key Population Workbook's development, extensively discussed in previous Reference Group meetings, has been instrumental for the estimates teams in reviewing the available key population data within their countries. It enables them to compare key population data against other sets of data. The comparison process involves evaluating data on PSE, HIV prevalence, and ART coverage. These are then compared against other modelled estimates derived from models such as the Goals model, the Optima model, or other country-specific models. The objective of this exercise is to reach a consensus on estimates for population size, prevalence, treatment coverage, and new infections, specifically segmented by population group. This approach ensures a more comprehensive and accurate understanding of the HIV epidemic across different key populations.

One of the main gaps identified in the workflow of the Key Population Workbook pertains to systematic methods for extrapolating key population surveillance data, such as that obtained from IBBS surveys to national estimates. Imai-Eaton highlighted that the approach to this extrapolation has been somewhat ad hoc, relying on methods like taking medians or using other existing national assessments. The first objective of this session was to build upon discussions from the May 2023 meeting, particularly around the development of the Triangulator and Aggregator tools. These tools are being evaluated to determine if they can be incorporated as more systematic approaches for data synthesis.

The session also focused on priority questions related to key population size estimate (KPSE) data from the May 2023 meeting, including evidence about changes in KPSE over time and the urban versus rural ratios of these population sizes. This information is vital for accurately extrapolating data, which is predominantly urban, to national estimates.

**Key population data quality assessment, guidance on inclusion/exclusion, and data synthesis**

**Ian Fellows**, in collaboration with **Carl Corcoran,** described the Triangulator and Aggregator tools and workflow.

The Aggregator tool is an expansion of the Triangulator tool and aims to aggregate KPSE at the national level, while addressing the varying levels of data quality and richness within key population data.

Motivation:

Critical metrics for informed decision-making include population size, prevalence, incidence, and VLS. While national estimates are needed for decision making, typically only local

information is available, with urban information being more well populated. Available estimates among key populations are at the subnational unit (SNU) or at the city level and may be rich or sparse.

***Goal***

*Create a tool for stakeholders to use that:*

1. *Generates national estimates from local information.*
2. *Can be used in countries with limited data (e.g. only one estimate in one city) and in data rich environments.*
3. *Is statistically valid and produces uncertainty estimates.*
4. *Is easy and intuitive to use.*

The primary input data are estimates of population size proportions. The main outcome of interest is the proportion of the general population that constitutes a member of a key population, along with the corresponding statistical uncertainty ranges. These estimates, including both lower and upper limits, are currently aggregated at the SNU level. The process involves using the Triangulator to create local estimates, which are then aggregated using the Aggregator tool. The proportion of this population residing in urban versus non-urban environments is a required input assumption.

The tool is designed to be flexible, operating with data from every SNU, a single SNU or a subset. The data requirements are minimized to reduce the data burden. Thus, only essential data points like population size and the proportion of the population in urban areas are necessary, in addition to whatever key population estimates are available. This approach ensures the tool is user-friendly and does not demand excessive external information.

The Aggregator uses a Bayesian model that operates on the logit scale, integrating a hierarchical structure. The model assumes that SNUs come from a distribution with specific means and variances and can also incorporate user-supplied correlation matrices for related SNUs. The model considers the difference between urban and rural data, allowing users to define the relationship between these datasets.

A web-based Shiny application has been developed to implement this model. A demonstration of the tool showed how users can input data into a table or upload their data, adjust priors, and run the model to obtain aggregated estimates, as seen in Figure 1.1. The application is still under development, with plans for further refinements based on user feedback.

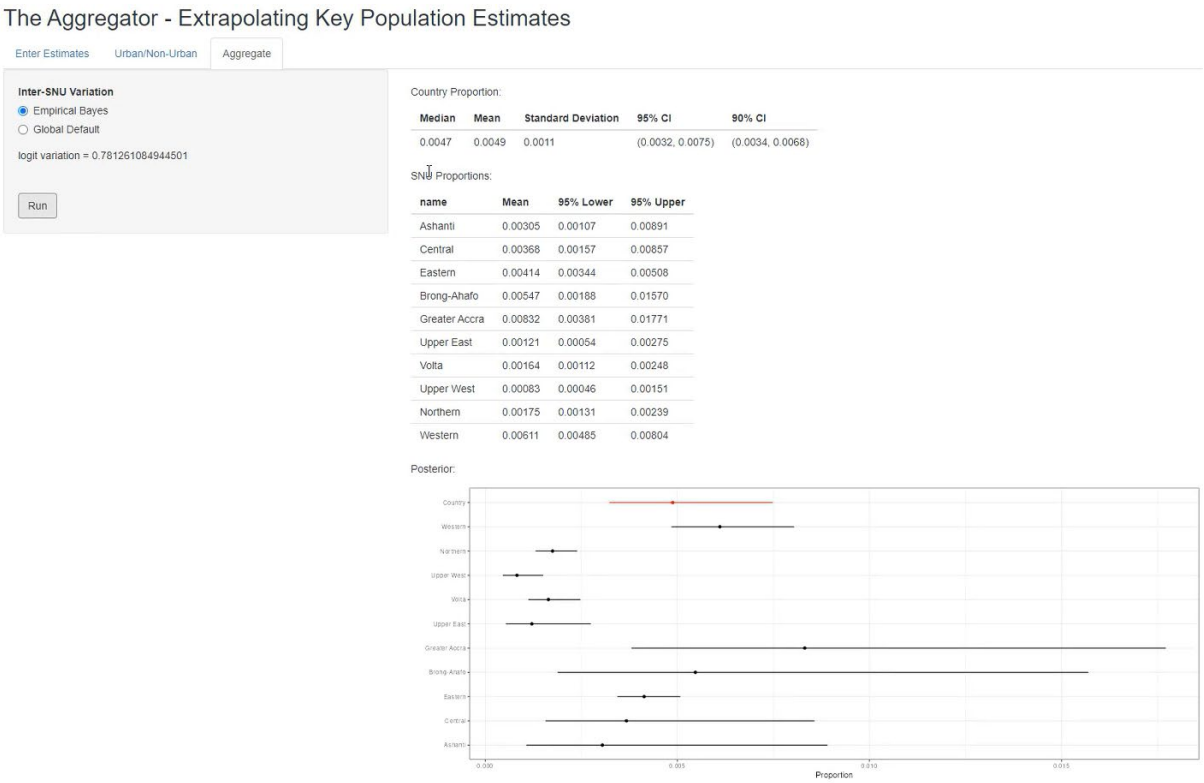



**Figure 1.1: The Aggregator interface and key population estimate outputs.**

**Carl Corcoran** presented the step-by-step workflow of the Triangulator, which statistically combines data from multiple studies to derive consensus subnational estimates, which form the inputs to the Aggregation tool.

Triangulator inputs:

- Name or method of the estimate
- Median, lower, and upper bounds of the estimate
- Confidence score (used as a weight when pooling data)

The Triangulator combines these inputs and prior information using a Bayesian model to produce a consensus estimate for a particular population quantity, such as population proportions, at a single SNU. This Bayesian approach generates a distribution that represents the estimated key population proportion.

Country teams work on a SNU-by-SNU basis, inputting data into the Triangulator to generate consensus estimates for each location, where data is available. The outputs for each SNU are then tracked to create a comprehensive dataset.

Confidence scores are on an absolute scale out of 100, with higher scores indicating greater confidence in the results. Confidence scores should be determined collaboratively, country teams should scrutinize individual study designs to determine the reliability of the data. For large datasets, confidence scores are set uniformly based on the study method and are generally kept low to facilitate data processing across multiple countries. Although this approach is not perfect, it yields reliable results for most countries using these default settings.

Standard default priors have been assumed for certain key populations, such as about 1% for men who have sex with men and 2% for female sex workers.

Corcoran demonstrated how to combine the Triangulator and the Aggregator into potential workflows, using country-specific case studies to show adaptability across different countries with varied data availability.

Corcoran presented results for several countries with varied data availability (see presentation).

*Case study: Female sex workers in Zimbabwe*

Several key population studies were considered, and the Triangulator processed these studies to produce district-level estimates of key population proportions. Corcoran showed the plot below (Figure 1.2) that displayed different estimate types on the vertical axis.

The Triangulator adjusts these estimates based on confidence scores, which scale the variation in the estimates and reflect the confidence in their representativeness. The dotted lines in the plot represent extended confidence intervals, influenced by the assigned confidence scores.

From this process, the Triangulator generated pooled estimates for the female sex worker proportion in Zimbabwe, marked on the plot with circular dots and confidence intervals, which then become inputs for the Aggregator tool.

In Zimbabwe, where substantial data was available at the SNU level, the Triangulator produced results that were comparable to other studies. He mentioned specific studies for comparison, including Stevens *et al*. (1) on key population size, HIV prevalence, and ART coverage, and the Laga et al. (2) estimate, which involved mapping the number of female sex workers across SSA.

The results from the Triangulator in Zimbabwe showed a narrow confidence interval around the national level estimate.

(1) Stevens et al. (2022) Key population size, HIV prevalence, and ART coverage in sub2 Saharan Africa: systematic collation and synthesis of survey data. medRxiv https://doi.org/10.1101/2022.07.27.22278071
(2) Laga et al. (2023) Mapping the number of female sex workers in countries across sub-Saharan Africa. PNAS https://doi.org/10.1073/pnas.2200633120

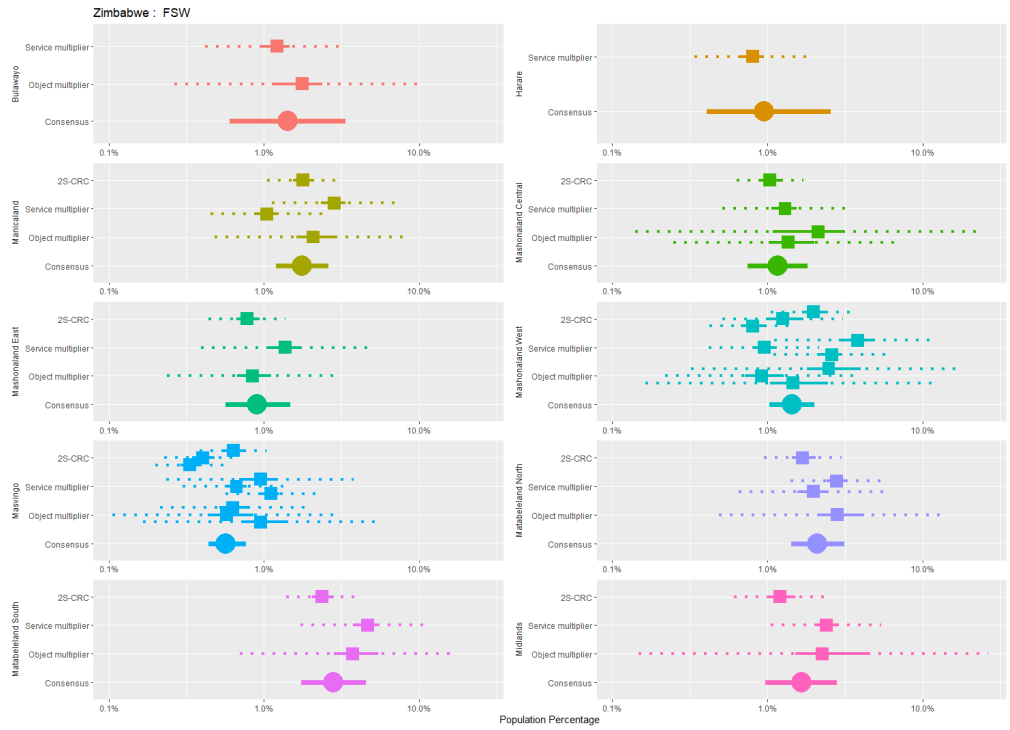



**Figure 1.2: Consensus estimates for female sex worker proportions in Zimbabwe.**

Following the presentation, several key points were discussed:

1. Integration of data lacking standard errors: The challenge of including data missing standard errors in data synthesis was highlighted. This functionality is not yet accommodated by current Triangulator methods.
2. Time component in models: A query was raised about whether the model considers changing key population proportions over time. It was acknowledged that while time is considered at the Triangulator level, extracting trends is challenging due to data sparsity. The relevance of study timeliness and age as study quality indicators was also discussed.
3. Use of strong prior distribution: The discussion broached the subject of strong prior distributions on national level parameters. An excessively strong prior might limit the model's responsiveness to new data. It was clarified that while there are priors for proportions at the Triangulator level, the national-level prior in their hierarchical model is uniformly flat, allowing new information to be integrated without being overly influenced by the prior.
4. Use of informative vs. non-informative priors: The implications of employing non-informative priors were examined, with a consideration of how they might influence proportion estimates. The model uses a flat logit scale prior to address this issue. The potential benefits of using informative priors when they are well-supported and documented were also discussed.
5. Comparison of multi-country analysis: The challenges of comparing pooled multi-country analyses were discussed, given the variance in datasets. A suggestion was made to reevaluate the model using only datasets that account for uncertainty, facilitating a more precise comparison.

While the Aggregator is a methodological advancement, it is still under development and it is recommended to encourage teams to use the tool in a pilot phase, but not yet as a

standardized recommended methodology to produce estimates. Specific feature additions and methodological refinements are available in **Annex A - recommendations**.

**PSE over time: multi-country analysis of trends, numbers, and proportions**

**Oli Stevens** presented an analysis assessing changes in PSEs over time based on population size proportion data collated for the Key Population Workbook. The study aimed to evaluate the assumption that key population size estimates (KPSEs) remain constant over time, a common assumption used in several mathematical models used by UNAIDS to produce estimates, such as EPP, CSAVR, and AEM. These models often assume a balance between new entries and exits within a key population group due to cessation or death. This study assessed linear trends in KPSE proportions and investigated whether socioeconomic or structural changes within a country were associated with changing in KPSEs.

The regression analysis adjusted for changing PSE study methods over time classified into three groups: empirical methods, mapping-based methods (such as programmatic mapping and place), and a combination of various methods. A random slope by method over time was included to capture changes in implementation. A study-level random effect was used to account for variation in individual study implementations.

Key findings:

- No evidence of change in PSEs over time, either by country or method. The log odds ratios were near zero, and the 95% confidence intervals included zero in all cases.
- No significant differences were found across any of the country-specific random slopes or method time slopes.
- PSE proportions could not be estimated for transgender women due to insufficient data.
- Data for female sex workers, men who have sex with men, and people who inject drugs across SSA was sparse and temporally scattered, with large heterogeneity, reflected by wide uncertainty ranges in modelled estimates.

Based on the Key Population Workbook data, the study concluded that there is no empirical evidence for a time trend in PSEs, segmented either by country or method (Figure 1.3).

Stevens noted that the dataset used in the analysis did not undergo any data quality assessment or exclusion, which could have filtered out low-quality studies. Including such studies without screening might influence the overall analysis and interpretation of trends.

The linear time trend used in the analysis may be insufficient to capture potential spikes or significant fluctuations in population sizes due to changing economic conditions, conflicts, or other structural changes within a country. Higher temporal and spatial observations of KPSEs, ideally from routine data collection, are needed. However, most settings lack such a data collection framework. Cross-sectional survey data currently available cannot detect changes in KPSEs due to limitations in data quality and frequency.

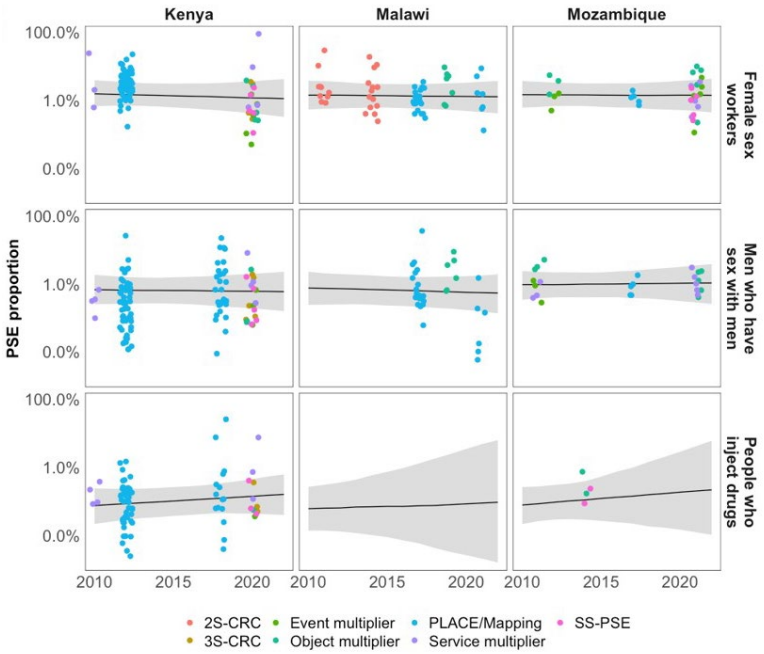



**Figure 1.3: PSE estimates for female sex workers, men who have sex with men, and people who inject drugs over time across Kenya, Malawi, and Mozambique.**

Figure 1.3 displays the data and modelled estimates for PSE proportions of female sex workers, men who have sex with men, and people who inject drugs in Kenya, Malawi, and Mozambique. There was wide heterogeneity in the observed PSE proportions (note log scale in Figure 1.3), particularly for mapping studies (blue points) which use diverse specifications in the denominators and changes in methodologies over time. The observation-level heterogeneity impedes the ability quantify a time trend.
Overall, the presentation concluded that there was no statistical evidence for a time trend in KPSEs across countries, but the heterogeneity in observations and infrequency of surveys limit confidence in this conclusion.

**Analysis of national PSE proportions reported via Global AIDS Monitoring (GAM)**
**Arias-Garcia** presented an overview of KPSEs reported by countries through the annual Global AIDS Monitoring (GAM) process coordinated by UNAIDS, WHO, and partners to measure progress against HIV targets.

The collected data included five key populations: sex workers, gay men and other men who have sex with men, people who inject drugs, transgender people, and those in prisons or closed settings. Countries submit national data as **population size counts** annually to the GAM through an online tool. Each value is approved by national governments as official reporting. Metadata accompanying the data submission includes definition of the key population, region the estimate covers, derivation method, the year of estimation and any subnational disaggregation. After submission, UNAIDS initiates a validation process that examines queries linked to incomplete or contradictory information, inconsistencies across indicators, and any substantial deviations from previously reported data. Discrepancies exceeding 25% from previous reports are scrutinized, and countries may be asked about reasons for changes or the possibility of national extrapolation from subnational estimates. For estimates falling outside regional bounds published in Spectrum guidelines or below 1%

of the male adult population for men who have sex with men, countries are requested to adjust their figures in accordance with WHO and UNAIDS technical briefs.

A mixed-effects logistic regression analysis was used to estimate time trends in logit-PSE proportions as reported to GAM.

In contrast to results reported by Stevens based on study-level observations, population size proportions reported to GAM have significantly increased over time (Figure 1.4) for female sex workers and men who have sex with men (when excluding values below 1% reported after the 2018 guidance). The trend was not statistically significant for people who inject drugs or the total men who have sex with men dataset without exclusions. Data for female sex workers, men who have sex with men, and people who inject drugs across SSA was sparse and temporally scattered, with large heterogeneity reflected by wide uncertainty ranges in modelled estimates.

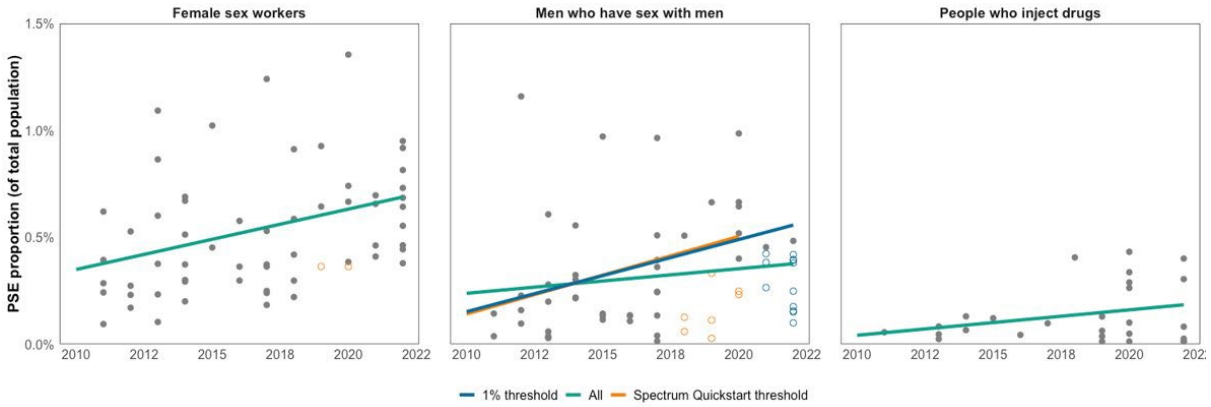



**Figure 1.4: Temporal trends in KPSE proportions for female sex workers, men who have sex with men, and people who inject drugs between 2010 to 2022. Graphs show data points for individual estimates, with lines indicating fitted logistic regressions. For men who have sex with men, lines represent separate regressions fitted to all reported data, and when exclude men who have sex with men proportions below the 1% threshold after 2021, and outside the Spectrum Quickstart manual ranges after 2019, respectively.**

There was considerable discussion about whether the significant time effect observed for sex workers and men who have sex with men datasets (Figure 1.4) might be due to reporting mechanisms - for example more consistent extrapolations from subnational to national size estimates in more recent reporting years, a true increase in the KPSE proportions, or increased societal openness leading to more honest reporting. The lack of a time effect among people who inject drugs (Figure 1.4, right), prompted questions about possible underreporting or whether the size of this population remained relatively stable. Stigma or fear of disclosure might still be influencing underreporting in these groups.

Hypotheses for the apparent increased KPSEs reported for female sex workers:

- Enhanced data collection methodologies, broader outreach in studies, and more consistent extrapolation to national estimates, rather than only reporting subnational size data measured directly in size estimation studies.
- Financial hardships in certain regions may have influenced more women to enter the sex work industry.
- Reduced stigma associated with sex work in some regions could lead to more accurate reporting.

Questions about GAM data and the assumptions underlying KPSEs:

1. Spatial correlation and regional analysis: Confirmed spatial correlation in the data; analysis for other regions was not done due to data not being collated similarly to SSA.
2. Data reporting and temporality: Clarified that data points reflect the time when the size estimation study was conducted, not the reporting year.
3. Use of program service delivery data: Program data could be incorporated into national KPSEs by countries, particularly for populations like people who inject drugs, but it is not common.
4. Confidence in estimates: Concerns about the reliability of estimates slightly above the 1% threshold for men who have sex with men and whether they could be subject to underestimation like lower estimates. The discussion highlighted ongoing data validation and the provisional nature of government-approved estimates, particularly for key populations, expecting yearly improvements.
5. Revisiting the 1% minimum threshold for men who have sex with men: Initially set based on limited data, but considering substantial additional new data, indicating lower population proportions, there was agreement on potential benefits of re-evaluating this threshold to improve estimate accuracy.
6. Conservatism in current estimate ranges: Any consideration for adjusting the range of 1% to 4% for population proportion estimates needs comprehensive global evidence.

The discussion highlighted the importance of continuous review and adjustment of methodologies and assumptions in GAM data to accurately reflect the changing dynamics of key populations.

**Urban:rural PSE ratios**

**Salome Kuchukhidze** presented analysis of data from Population HIV Impact Assessment (PHIA) household surveys addressing the rural-urban distribution and HIV outcomes among men who have sex with men. The study used survey questions regarding the gender of sexual partners to estimate population size proportions of men who have sex with men and understand their HIV risk, treatment coverage, and viral suppression rates.

Male survey respondents were classified as men who have sex with men if they self-reported one of their most three recent sexual partners were male. Logistic regression was used to estimate the urban:rural difference in each outcome among men who have sex with men, adjusted for age group and wealth. To create a balanced analytical sample, men who did not report same-gender partners were subsampled matched on rural-urban residence and wealth.

Among over 76,000 male survey respondents in 11 surveys who reported being sexually active in the past year, 744 reported having had a male sexual partner (slightly below 1%). About 20% also reported female sexual partners.

Key findings:

- Men who were currently married were more likely to report having sex with other men, signalling potential reporting errors in responses to sexual partner gender questions - a potential limitation for overall interpretation of further results.
- No strong associations between being reporting a male same-gender partner and living in households with children.

- Reporting a male same-gender partner was not associated with having two or more sex partners in the past year. There was significant heterogeneity across surveys in this finding.
- Men reporting a male same-gender partner had 50% higher HIV prevalence compared to men who do not report having sex with men in the past year, adjusted for rural-urban status, wealth, primary sampling units, and the number of sexual partners.
- Men with HIV reporting a male same-gender partner were 22% more likely to be on ART than men with HIV who do not report having sex with men in the past year.
- Men with HIV reporting a male same-gender partner were 17% more likely to be virally suppressed. This may be due to more targeted interventions and better engagement in HIV care in these populations.
- **Men who live in urban areas were not significantly more likely to report a same-sex partner than men who live in rural areas (aRR 1.05, 95% CI 0.90-1.22; Figure 1.5).**

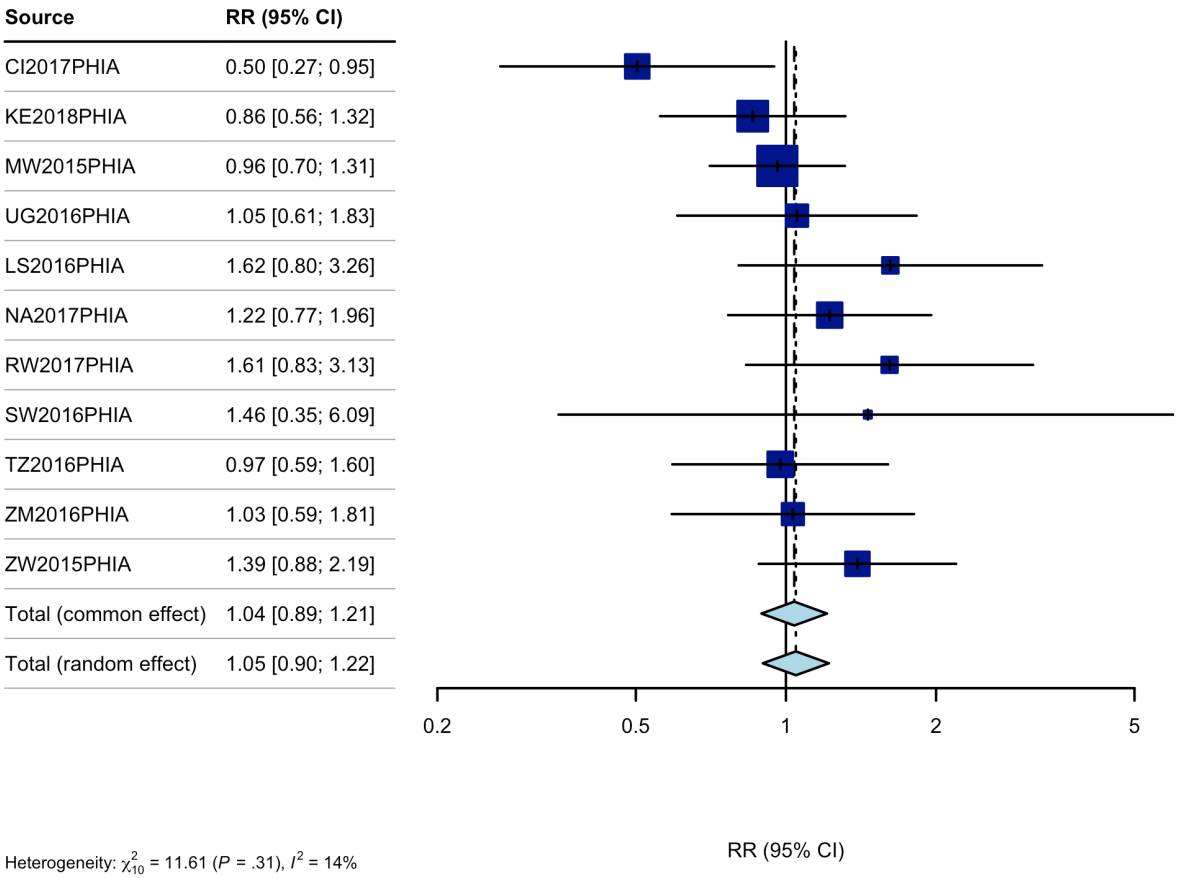



**Figure 1.5: Adjusted risk ratio of reporting a male same-sex partner in urban vs rural areas.**

Overall, no substantial differences were identified between rural and urban areas or by age groups in the proportion of men reporting a same-sex male partner.

Challenges and considerations:

- There were very few recent HIV infections recorded among men reporting a same-sex male partner, which poses a limitation to the conclusiveness of recent transmission trends.

- Concerns about misclassification and the sensitivity and specificity of survey questions were raised.

Next steps:

- Further analysis will explore the impact of focusing on past three sexual partners in the past year for estimates.
- There will be a further review of the accuracy of self-reports through triangulation with additional survey data, for example the gender of household members identified as spouses by men who reported being married and reported a same-sex partner.
- Additional surveys will be assessed for inclusion, such as recently released PHIA surveys from Botswana.
- The study aims to triangulate gender information from individual surveys with household roster data, despite potential imperfections in this method.

Below is a summary of questions and clarifications addressed in the discussion:

1. Proportion of missing data: A question was raised about the proportion of missing data on sexual partners' gender. Overall missing data was low (<5%). eSwatini had a larger percentage of missing data and was excluded from the analysis. With less than 5% missing data, a complete case analysis was used for the study.
2. Inclusion of additional household surveys: The potential inclusion of other household surveys, such as DHS surveys with serology or AIS surveys, was discussed. DHS surveys do not collect information on the gender of sexual partners, and therefore could not be included.

**Collated key population size estimation study data**

**Oli Stevens** presented an analysis of the association between population density and KPSE proportions from collated PSE studies to extrapolate evidence on urban-to-rural ratios of KPSEs proportions.

Existing KPSEs studies are conducted in predominantly urban settings. Extrapolating PSE proportions measured in size estimation studies to national estimates therefore requires an assumption about KPSE proportions in rural areas. The Key Population Workbook uses a standard urban/rural ratio (0.6% rural key population proportion for every 1% urban key population proportion) for all key populations based on expert consensus. This consensus suggests that rural PSE proportions will generally be smaller than urban areas. However, limited data is available for rural areas, making it difficult to empirically estimate this ratio.

This analysis examined the relationship between KPSE proportions and population density in locations with population size studies to assess evidence for whether PSE proportions were larger in more urbanized areas. KPSE study location area names from surveillance reports were matched to boundaries either from the Global Urban Rural Mapping Project (GRUMP) (~1000 KPSE observations) or Naomi area names (~700 observations). Matched boundaries provided total population sizes as denominators to convert the acquired KPSE into KPSE proportions. Population density for each KPSE observation was assigned as the population density in a 10 km radius around the matched area centroid, calculated from WorldPop 'constrained' population density rasters. A 100 km radius was used for sensitivity analysis.

Figure 1.6 shows the results. Higher population density was negatively associated with PSE proportion, the opposite of the hypothesised association. The negative correlation remained but was weaker with a 100 km radius to define population density.

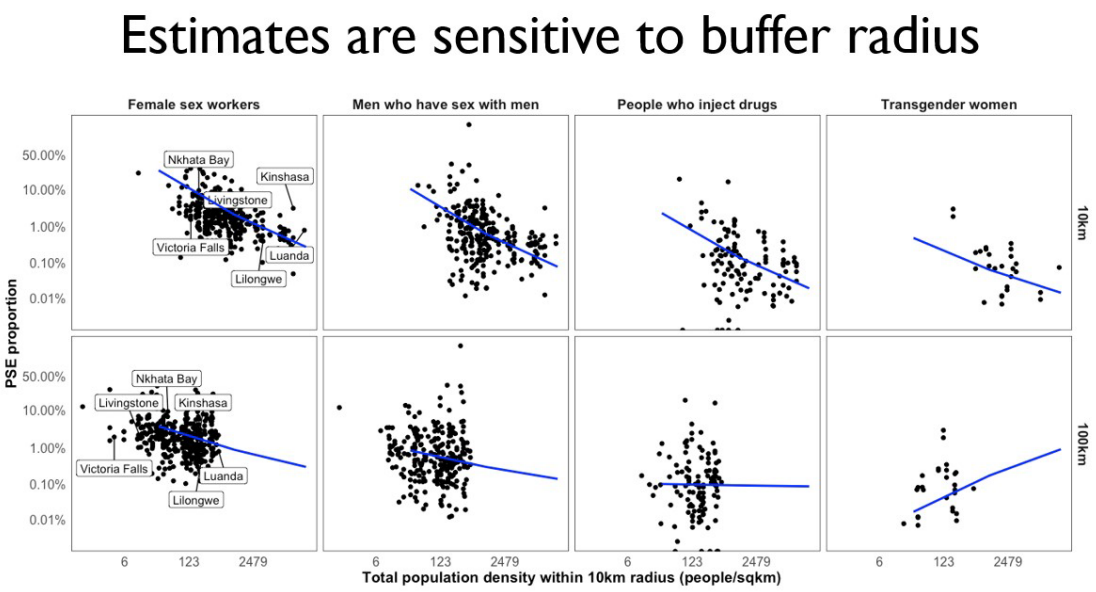



**Figure 1.6: Scatter plots of the relationship between key population size estimate (KPSE) proportions (female sex workers, men who have sex with men, people who inject drugs, and transgender women) and total population density within a 10 km or 100 km radius.**

Key findings:

- A negative correlation was found between total population density and KPSE proportion, with higher KPSE proportions in areas of lower population density.
- There was strong potential for bias in the analysis related to possible misspecification of area population totals, which were derived from the same source for both KPSE denominators and population density numerators. This could induce the observed negative correlation.

Conclusions and recommendations:

- The preliminary analysis indicating higher PSE proportions in lower-density areas challenges the current consensus and the 0.6 ratio assumption.
- Further investigation is needed, particularly around choosing an appropriate buffer radius for analysis. The approach should be reviewed and adjusted for potential biases, and variables like national border areas need better handling.

The discussion elicited the following key points:

1. Scepticism about rural-urban dynamics: A participant expressed scepticism on findings in countries like Uganda, Namibia, Zimbabwe, and Malawi.
2. Consideration of sampling methods: Another participant questioned sampling in rural areas and suggested that in very rural areas, there would be only one hotspot or bar, making it an obvious place of gathering. In contrast, urban areas have a lot of diversity in hotspots.
3. Under-enumeration in urban areas: Stevens also acknowledged the possibility of under-enumeration in large urban areas, where not all hotspots might have been fully accounted for, leading to potentially lower urban PSEs.
4. Challenges in comparing urban and rural dynamics: Stevens expressed scepticism about the steep negative correlation estimated, suspecting bias related to mis-specified population sizes.

**Network scale-up method PHIA in Nigeria/Cote d'Ivoire**

**Anne McIntyre** analysed urban-rural PSE ratios using data using the Network Scale Up Method (NSUM) applied in two recent nationally representative household surveys: Côte d'Ivoire PHIA (2017) and the Nigeria HIV/AIDS Indicator and Impact Survey (NAIIS) (2018). The two surveys estimated population sizes for female sex workers, men who have sex with men, people who inject drugs, transgender individuals, and clients of female sex workers.

The NSUM method works by quantifying the size of a respondent's total network through referencing how many people the respondent knows among populations with 'known sizes' derived from census data, national registries, or representative surveys (e.g., teachers, women giving birth). Then how many members of key population groups respondents know is asked for comparison.

Results regarding the urban:rural proportion ratio different substantially between the surveys. In CIPHIA, around 63% of the total population were urban and 37% rural (~2 times larger urban population), and key population sizes (in counts) ranged between 4.4 to 4.6 times larger in urban areas, **implying an urban:rural ratio around 0.3 for population proportion.** In NAIIS, 48% of the total population was urban and 52% was rural. Estimated key population size counts in urban areas ranged between 0.78 to 0.90 times those in rural areas, **implying approximate parity of urban:rural KPSE proportion ratios.**

In both surveys, the urban:rural ratios were remarkably similar across all key population groups, indicating potential limitations in the network size assessment.

Potential limitations:

- Disparity in representation of known populations in urban and rural areas, affecting the accuracy of urban-rural ratios. For instance, university students might be more prevalent in urban areas, while the incidence of childbirth could be more evenly distributed.
- Outdated known population data, particularly in Nigeria where they were derived from the 2006 census. This discrepancy raised concerns about the changing distribution of rural and urban populations.

Future considerations for NSUM:

- To improve urban-rural ratio calculations, McIntyre suggested investigating whether known population data are available by urban and rural status.
- She recommended selecting known populations that would equally reflect urban and rural responses.
- Including an urban and rural component in the module questions might be worth considering, to refine the ratios further.
- Current NSUM instructions prompt broad geographic responses, but more localized questions might yield more accurate local data.

Primary inquiries centred around:

1. Urban-rural ratios: Clarification was sought on whether the urban-rural ratios were based on absolute numbers or proportions of the population that were PSE. McIntyre indicated that the ratios were in absolute numbers (reflected in interpretations above). As proportional comparisons are more appropriate, McIntyre agreed to do further analyses but added that it was an approximation still because of the use of urban and rural information in EAs.

2. Urban-rural determination: Clarification was sought on whether the urban-rural determination referred to the percentage of population within EAs or the national population in those urban and rural EAs. McIntyre clarified that it pertained to the national population distribution.

Overall, McIntyre's presentation was inconclusive about evidence on urban:rural ratios given the large difference in results between the surveys and similarity of result across all populations in the same survey. Some of this could be related to inherent challenges of the NSUM methodology.

**Inclusion of key population programme data in the SSA Key Population Workbook**
**Oli Stevens** discussed options for incorporating a new worksheet for key population programme data into the Key Population Workbook, which was recommended in the May 2023 Reference Group meeting.

Three primary objectives were discussed for the proposed new programme data worksheet:

1. Collating data for future use: Stevens proposed including a template sheet in the Workbook, stratified by Naomi District area ID and calendar quarter. This would allow for simple validation checks, such as ensuring the number of key population members reached is greater than other recorded indicators.
2. Review of data quality and trends: Data can be aggregated to the national and annual level for easy visualization. While this option is useful for viewing large-scale trends, it poses challenges in identifying specific data quality issues. Stevens noted the potential for revisiting the issue of extensive data review at the national level, as previously experienced with Naomi.
3. Comparison with survey data: This would involve comparing consensus national estimates of PSE, HIV prevalence, and possibly ART coverage with program data at the end of the workbook process. However, Stevens expressed concern about the potential loss of precision and understanding due to comparing aggregated district-level program data (programs do not have full coverage of all districts) with district-level survey data that has been extrapolated to national level.

Recommendation and suggestions for future data collection:

- At this stage, the focus should be on collating key population program data for future use only, collecting data on numbers reached by programs, numbers on treatment, and a prevention indicator. Stevens invited suggestions on which prevention indicator might be most reliably and consistently recorded across countries.
- He also suggested recording these data at the district level and calendar quarter, possibly using separate tools outside of Excel for future review of trends.
- Additionally, the suggestion from May 2023 was brought up again to change the timing of Key Population Workbook completion to ensure the right stakeholders are involved in the data review process.

Incorporation of the Triangulator:
Stevens proposed that if the Triangulator is integrated into the consolidation process, confidence scores should be entered in the workbook to ensure study quality decisions used in the data synthesis are recorded. He suggested copying and pasting data directly into a Shiny app rather than relying on file uploads, due to potential internet connectivity issues during workshops.

Key points of discussion:

- Country-specific work and nuances: Reflecting on engagements with country estimates teams, the importance of understanding the nuances in different countries was emphasized, such as the portion of key population members attending general versus key population specific services and comparing age distributions in cross-sectional surveillance versus program data.
- Data collection and future use: Stevens highlighted the need to start collecting program data now, even though the systems may not be fully ready to use them. He pointed out the increasing implementation of unique identifiers and the importance of being prepared to use data as systems strengthen.
- Concerns about "Reached by Programs" indicator: A concern was raised about the differentiation between number of program contacts and unique individuals, and it was emphasised that it must be made clearly in the workbook that the number of unique individuals should be entered. This is difficult in many programmes where clients do not have unique static health IDs, and in some cases where services are accessed anonymously by design.
- Importance of qualitative information: It was suggested that by having a comments field for collecting qualitative information on populations coming for the services or programs, one would be able to appreciate how these populations are different from the entire key population in the area and hence providing crucial inputs for future systematic data collection.
- It was noted that even if unique identifiers are used, they may not be shared between different implementers of key population programmes, which may be especially challenging in the case of mobile populations, such as sex workers.

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

# Session 2: Dynamical modelling of HIV trends in key populations in sub-Saharan Africa

**Chair Cari van Schalkwyk** highlighted the central objective for Session 2:

- Review assumptions and model structure decisions for the Goals-RSM and Goals-ARM models

This session followed a session about design of future transmission dynamic estimates model during the May 2023 meeting, where the Reference Group recommended the development of the Goals Age and Risk stratified model (Goals-ARM) for the estimation of HIV indicators among key populations in SSA. The May meeting recommended to establish a dedicated working group to guide the model's development. The working group has been active, with their first session convened in September 2023 to review and provide feedback on progress developing Goals-ARM.

The first three presentations in this session reviewed assumptions in Goals-*RSM* and data on key population turnover.

**Duration at risk: Assumptions in Goals-RSM 2023**

Following the May 2023 Reference Group meeting, UNAIDS adopted a new approach to estimate distribution of new infections by population group (the 'donuts') using outputs from dynamic transmission models. The Goals-RSM model was used for most countries in sub-Saharan Africa. To improve the evidence base on model assumptions, the May 2023 meeting recommended reviewing data on risk duration and 'turnover' of key populations.

**John Stover** described the population size and turnover (or duration in risk group) assumptions in the Goals Risk Structured Model (Goals-RSM) used for the UNAIDS 2023 estimates.

The structure of the RSM model represents distinct risk groups, including:

- Individuals with only one partner in the previous year
- Individuals with more than one partner in the preceding year
- Female sex workers and their clients
- People who inject drugs
- Men who have sex with men

Updates post-Reference Group recommendations:
Following recommendations from the previous Reference Group meeting (May 2023, Stellenbosch), UNAIDS specified new key population size proportions for each country, which were fixed regional medians proportions (Figure 2.1).

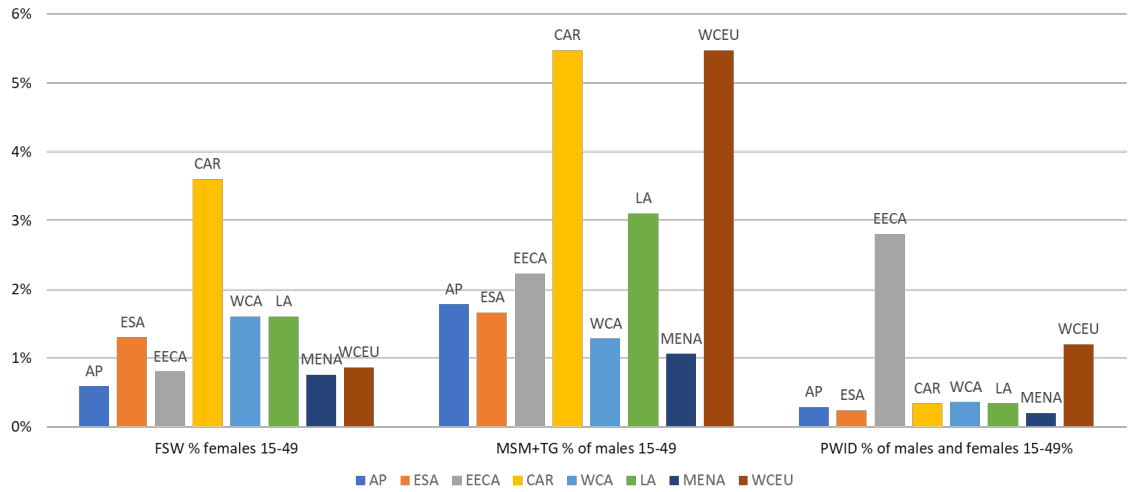



**Figure 2.1: Key population sizes as percent of 15-49 population; UNAIDS estimates based on country-specific studies.**

Duration estimates for risk groups (Figure 2.2):
For SSA, the same durations were applied across all countries, with 8.2 years for sex workers and 13.6 years for people who inject drugs. The durations of sex work are derived from a 2012 review by Erica Fazito and colleagues (3), while the durations for people who inject drugs are based on a 2020 review by Hines and colleagues (4). These estimates vary considerably across regions and are country-specific for the Hines estimates for people who inject drugs.

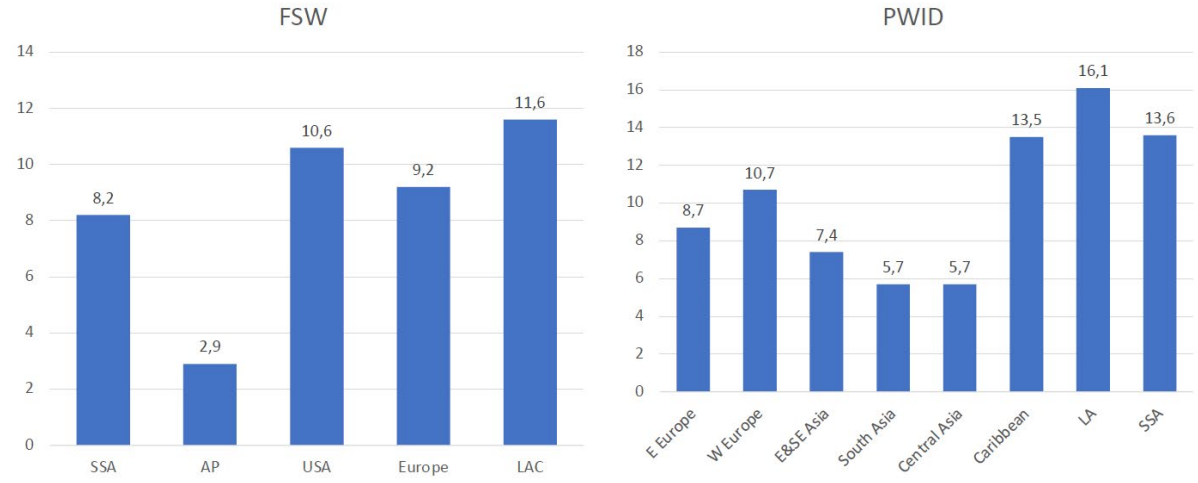



**Figure 2.2: Durations in risk group assumed by Goals-RSM. men who have sex with men are assumed to remain lifelong, at constant level of risk. Female sex worker estimates outside SSA are from Fazito *et al* (3); people who inject drugs estimates from Hines *et al* (4).**

Model updates and outputs:

- All 108 models were updated with the new KPSEs and duration estimates.
- The models estimated the number of new infections occurring in 2022 among key populations and their partners (Figure 2.3)



(3) Fazito et al. (2012) Analysis of duration of risk behaviour for key populations: a literature review. Sexually Transmitted Infections. https://doi.org/10.1136%2Fsextrans-2012-050647
(4) Hines et al. (2020) Associations between national development indicators and the age profile of people who inject drugs: results from a global systematic review and meta-analysis. The Lancet Global Health https://doi.org/10.1016/S2214-109X(19)30462-0

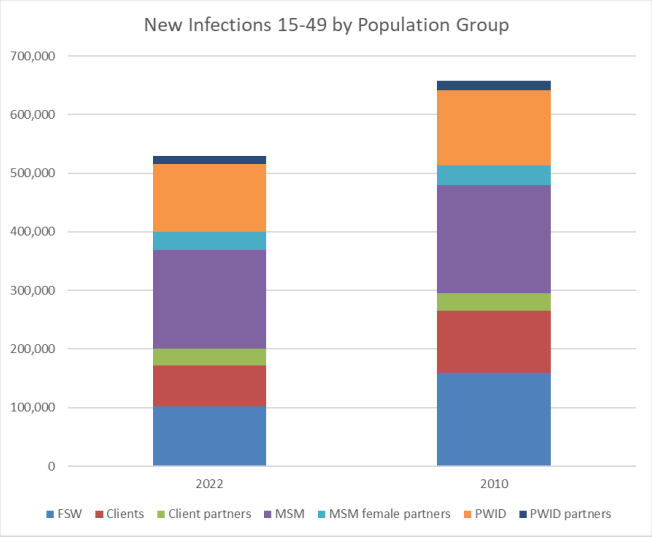



**Figure 2.3: Distribution of new HIV infections among adults aged 15-49 by key population groups (female sex workers and their clients, men who have sex with men and their female partners, people who inject drugs, and their partners) for 2010 and 2022.**

The discussion highlighted the complexities involved in accurately modelling partnerships and key population behaviours. More detailed data is needed to fully use the model's capabilities and accurately represent the diverse dynamics of partnerships and risk behaviours.

The following questions were discussed during the discussion:

1. Partner modelling in Goals-RSM: A question was raised regarding how the mixing assumptions were specified for rates of partners formation across different risk groups. It was explained that, for sex workers, clients are considered a distinct group, with this distinction based on national household survey data related to individuals who pay for sex. For other partner categories, the model uses marriage rates and data on the proportion of key population members engaging with partners outside their key population group.
2. Key population partnership outside Group: In the Goals model, individuals classified as "low risk" have the potential to form partnerships across any group; men who have sex with men can also partner with women. Despite the model accommodating four risk strata for men who have sex with men, this is not currently used due to limited data.
3. Turnover and HIV status of new female sex workers: It was queried whether women entering sex work are drawn from the entirely susceptible population, equally from all populations, or from already higher risk populations, which has been reported in some data. The current model assumes entry into sex work coincides with the onset of sexual activity, thus excluding the possibility of HIV-positive individuals entering as new sex workers. Updates to this assumption may be considered for future iterations of the Goals-ARM model but remains unchanged in the current RSM version.

**Duration at risk: Variation across countries (SSA; female sex workers, men who have sex with men, people who inject drugs)**

**Rebecca Anderson** presented estimates of the distribution of duration at risk among key populations using pooled individual-level data from key population surveys across.

Objective:

The research aimed to determine the average survey-reported duration at risk among key populations and assess if these estimates varied across regions or over time.

Data was sourced from individual-level key population surveys:

- 53 surveys for female sex workers across 25 countries.
- 27 surveys for men who have sex with men from 20 countries.
- 19 surveys for people who inject drugs from 15 countries.

The earliest data for female sex workers was from 1995, but most data were from after 2010, especially for men who have sex with men and people who inject drugs. Studies used different survey questions to capture "duration at risk." For instance, female sex workers were asked about their age when they started sex work or how long they have been selling sex, men who have sex with men about the age they first had sex with another man, and people who inject drugs about the age they first injected illegal drugs.

Methodology:
The team extracted duration at risk data from surveys, prioritizing certain questions over others to maintain consistency. They accounted for different sampling methodologies, applying Respondent-Driven Sampling (RDS) weighting when appropriate. Kernel density estimation was used to smooth study-specific duration distributions, and a random effects meta-analysis was used to derive pooled medians of duration at risk.

Key findings:

- The median duration of risk for female sex workers was approximately 4.5 years, with no substantial regional variation observed. For men who have sex with men, a median of 5.5 years was reported, with slightly longer durations in Western and Central Africa compared to Eastern and Southern Africa. People who inject drugs reported a longer median duration at risk of about seven years, with no significant regional variation.
- A preliminary temporal analysis suggested an increase in duration over time, especially for female sex workers and men who have sex with men, though the data were sparse, and further analysis is needed to confirm these trends.

Key points from the discussion:

1. Methods of duration calculations: **Jesse Knight** highlighted his recent work (5) on methodological considerations and biases in calculation of duration at risk, such as the cyclic pattern of entering and exiting sex work. He noted that they found opposing impacts of biases: on the one hand, reported duration reflects only duration before date of survey and on the other hand, those with shorter durations or those with cyclical behaviours may not be captured in the cross-sectional survey sample. Under certain assumptions about the distribution of duration at risk, the reported mean duration at risk is equal to the mean true distribution. The doubling of the median duration (like suggested in Fazito *et al*.) may be an inappropriate assumption. Based on the findings from his study, Knight suggested that there might be a need to re-evaluate and potentially lower the estimates of the duration of sex work, considering the quick entry-exit cycle observed in many cases.
2. Impact of cross-sectional studies: It was noted that the over-sampling of younger men who have sex with men and the underrepresentation of older individuals might skew the duration at risk data.



(5) Knight J, Wang S, Mishra S. Adjusting for hidden biases in sexual behaviour data: a mechanistic approach. medRxiv https://doi.org/10.1101/2023.08.16.23294164

**Duration at risk: Variation over time (South Africa; female sex workers)**

**Nanina Anderegg** presented analysis of age and duration at risk among female sex worker studies in South Africa.

Data source:
The analysis included aggregate data from 19 studies conducted between 1996 and 2019, providing 23 estimates on the mean duration of sex work and 33 on mean age of female sex workers.

Methodology:

- Fitting a Bayesian hierarchical model with a log-linear relationship between study year and duration of sex work/age of female sex workers.
- Sensitivity analysis focusing on age at entry into sex work, addressing the gap in data regarding entry ages.
- Simulation exercise to assess the impact of observed trends on HIV incidence rates, assuming both constant and variable durations of sex work and age.

Key findings:

- The median duration of sex work increased over time (approximately 46% over a decade).
- Significant increase in the mean age of female sex workers (approximately 16% over a decade).
- The mean duration of sex work was around three years in 1996, rising to about seven years in recent times. Similarly, the mean age increased from approximately 26 years to 33 years over the same period.

Sensitivity analysis:

- Data indicated that while the age at entry into sex work is lower than age during sex work, the increase over time was similar.

Simulation exercise:

- Showed that assuming constant age and duration results in a less pronounced decrease in HIV incidence compared to assuming an increase in age and duration over time.

Implications:

- The findings suggest that mathematical models should account for changes in age and duration of sex work to accurately capture incidence trends among female sex workers.
- These trends may have various explanations, including improved community mobilization and rights advocacy or limited alternative employment options.

Key points from the discussion:

1. Characteristics or patterns of sex work over time: A question was raised on whether the increasing duration in sex work over time might be associated with different patterns of sex work. The importance of examining how cohorts entering the field at different times may vary was highlighted, to understand these trends better and address biases.
2. Heterogeneity in data: Anderegg acknowledged the heterogeneity of the included studies, suggesting that the trends observed could be due to different study populations and characteristics within populations. A suggestion was made to mitigate this heterogeneity by comparing with results from studies that have continuously

surveyed sex workers longitudinally, for example the CeSHHAR programme in Zimbabwe and longstanding sex worker cohorts in Kenya.

**Goals-ARM development**

**Rob Glaubius** provided an update on the development of the age and risk-stratified Goals model (Goals-ARM).

The Goals-ARM has a demographic population projection component integrated with a compartmental HIV epidemic model. The model, stratified by age and sex, incorporates nine behavioural risk groups, each classified by varying criteria, such as marital status and behavioural tendencies.

Key features of Goals-ARM and updates since May 2023 recommendations:

- Heterosexual risk groups: Despite ongoing deliberations on inclusion of a marital status stratification, the model has retained marital status-based risk groups.
- Transgender people: The model stratification has been updated to include a risk category for transgender women.
- HIV status stratification: The model stratifies HIV stages by CD4 cell count categories, and men are further stratified by circumcision status.

Key observations of model fits for Malawi and Zimbabwe, presented as two pilot countries:

1. Delayed epidemic onset: In both countries, the initial rate of epidemic growth in the fitted Goals-ARM model was slower than estimated by other models.
2. Flexibility in fitting key population data: The model successfully fits sparse key population data on HIV prevalence and supports arbitrary age ranges in which data is reported. However, the model has not yet made any specific accounting for the limited national representativeness of the key population study data.

Model improvements and recommendations shared from the Reference Group September 2023 meeting (still to be implemented):

- The need for dynamic variation in key population sizes over time.
- Differentiating male-to-male transmission risk by insertive or receptive role to reflect the varied risks among men who have sex with men and TGW.
- Expanding visualizations to enhance understanding of population sizes and the impact of sexual balancing on partnership rates.
- Using collated key population survey data to inform prevalence priors, especially for populations like people who inject drugs among whom data is scarce.
- Considering additional data on mixing patterns and sexually transmitted infections as a co-factor for HIV acquisition.
- Addressing the slow epidemic take-off with various proposed mechanisms, including examining marital and cohabiting structures, and potentially using an exogenous time-varying multiplier on the force of infection as a last resort.

Next steps:

1. Testing methodologies to address the slow start of the epidemic in the model.
2. A focus on South Africa as the next pilot country to measure the influence of vital registration data vs antenatal survey data on the model's early epidemic representation.
3. A detailed technical document is in the pipeline to provide an in-depth understanding of the model and gather feedback.

Discussion revolved around ideas to address the challenge in achieving the rapid exponential growth at the start of the epidemic, with key points:

1. The role of STIs in the early epidemic: The role of high rates of sexually transmitted infections (STIs) among sex workers played in the rapid transmission of HIV during the initial phase of the HIV epidemic in Thailand and likely in many African countries. The potential impact of considering STIs as a cofactor for HIV acquisition on the dynamics of the epidemic, particularly in its early stages, was discussed.
2. Interpretation of Key Population HIV prevalence data: Questions were raised regarding the approach to interpreting data on HIV prevalence among key populations from IBBS as nationally representative and the inclusion of random effects in the analysis to account for possible variability. It was noted that the data was treated as unbiased survey observations, with suggestions that the likelihood could be modified to account for biases in the data.
3. Weight of data in likelihood: The query was raised about whether all data, including key population HIV prevalence and household survey estimates, carried the same weight in the likelihood. The weighting is primarily based on effective sample sizes, without introducing additional complexity in comparing data from household surveys and key population-specific surveys.
4. Female sex worker population age distributions: It was clarified that a log-normal distribution fitted to reported age distributions from a single IBBS survey was used to derive the female sex worker age distribution in the model. This could be enhanced with the analysis reported by Anderson.

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

# Session 3: ART coverage data discrepancies

**Chair Leigh Johnson** introduced Session 3's objective:

- Review options and progress in assuring quality of ART data input to models and ART coverage estimates

The session revisited previously discussed discrepancies between survey estimates and program data related to people on HIV treatment. There are noticeable variations not just in absolute figures but also in age distributions. The aim is to equip countries with tools to discern the origins of these discrepancies.

**Brief update on Leapfrog – common code for simulation models (Spectrum, EPP-ASM, CSAVR, Shiny90)**

**Maggie Walters** provided an update on LeapFrog, a tool serving as a unified code base for the simulation models integrated within the Spectrum suite. LeapFrog consolidates various models within the suite that use the same epidemic simulation code, specifically focusing on population projections and disease progression models.

The core purpose behind LeapFrog's introduction was to address the recurring challenges that arose from having multiple separate code bases. Whenever any modifications were made to one, they had to be laboriously replicated across the others. This repetitive process was not only tedious but also heightened the risk of inconsistencies.

Much of Spectrum's 'DemProj' and 'Impact' module was originally developed in Delphi during the early 2000s. However, between 2017 and 2021, EPP, Shiny 90, and CSAVR shifted to the more versatile C++ language. The present endeavour, LeapFrog, seeks to reimplement this simulation code in templated C++, enhancing efficiency and accuracy. While a simulation typically took approximately two seconds, the C++ version drastically reduces this to 0.002 seconds, which significantly expedites the modelling process.

The advanced LeapFrog framework offers multiple benefits:

- Flexibility: It enables users to seamlessly enter and exit simulations.
- Customization: Users can activate or deactivate specific modelling components, thus tailoring the simulation. This includes components like EPP, CSAVR, Shiny 90, and the paediatric model.
- Improved documentation: Recognizing the complexities of managing such an extensive tool, efforts are underway to strengthen the codebase documentation.
- Integration: The templated C++ format supports the integration with various automatic differentiation libraries, which will be pivotal for model fitting.

Plan for integration:

- For the 2024 estimation round, a collaboration between Imperial and Avenir Health is underway to incorporate the Leapfrog code.
- Leapfrog is also expected to be used as a standalone external tool for estimating uncertainty in the 2024 estimation round.

**Rob Glaubius** presented advantages of LeapFrog and its strategic integration plan for Avenir's software suite.

LeapFrog can serve as a single computation engine across various tools, ensuring uniformity in demographic and HIV epidemic calculations. This consolidation presents two primary challenges: seamless integration into both Spectrum on the web and Spectrum desktop.

Spectrum on the Web:
There is a preference to begin with the web version of Spectrum as it offers a relatively straightforward integration process. Over the past year, efforts have been dedicated to transitioning Spectrum on the web to Python. As part of this shift, both the AIM and DemProj calculations within Spectrum were segregated into an individual Delphi module. This modular structure allows Python to call upon these calculations as needed. Given the distinct isolation of AIM calculations, there should be a straightforward transition where LeapFrog can be effortlessly inserted in place of the existing module, with some minor adjustments for effective communication with the broader Spectrum on the web framework.

Spectrum Desktop:
The integration for Spectrum desktop presents more complexities. Having been under development for an extended duration, Spectrum desktop's modules, especially AIM, are intrinsically linked to every other module due to the software's organic evolution over time. Before LeapFrog's integration, there is a need to meticulously isolate AIM calculations from the other interconnected modules. Once achieved, LeapFrog can replace the existing structure, thereby streamlining the computation process.

Key points from the discussion:

1. There was a proposal to use LeapFrog as a tool to run uncertainty analysis using an R or Shiny app. The intent is to produce results consistent with Spectrum but in a significantly reduced time frame. The discussion highlighted variability in the time observed by different users for uncertainty analysis in Spectrum, indicating the potential value of such a tool.
2. While the integration of LeapFrog into Spectrum is ongoing, there was no clear indication of making immediate changes to the current uncertainty calculations in Spectrum for this year.
3. Questions arose about how Spectrum users will be guided to explore adjustment options—whether this guidance would be incorporated within the software, provided in a user manual, or communicated informally. It was emphasized that it should be clearly communicated that there would not be significant changes that could drastically alter estimates.

**Visualising program and survey data by age (ART coverage, prevalence)**

**John Stover** elaborated on visualizing program and survey data, with a specific focus on ART coverage by age. Stover showed the two validation screens in Spectrum related to ART by age. In Figure 3.1 below, the first graph displays program data on the number of individuals on ART by age against Spectrum's estimates, while the second graph compares ART coverage from survey data against Spectrum's estimates. This juxtaposition allows for a clear comparison between data and the estimates generated by Spectrum for any given year, thus inconsistencies can be quickly identified, helping inform refinements to improve future estimates.

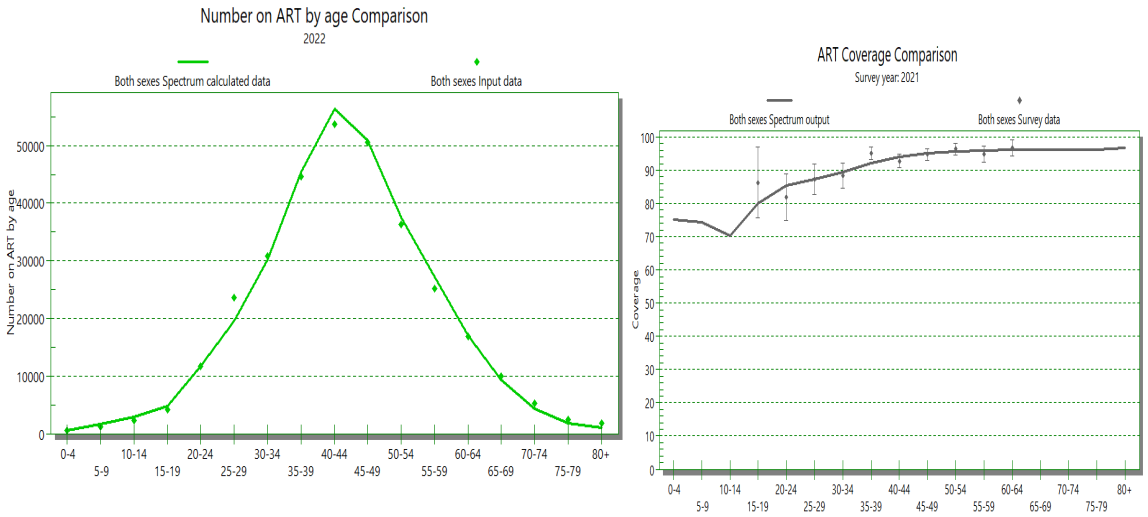



**Figure 3.1: Botswana ART coverage, showing the successful alignment of empirical data with model estimates.**

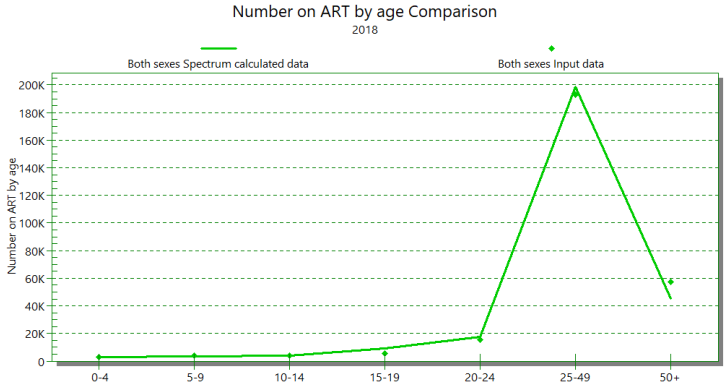



**Figure 3.2: Age-specific ART coverage in Cameroon, 2018.**

Botswana was presented as a model example with both program and survey data available, demonstrating a strong fit and suggesting high ART coverage (Figure 3.1). Challenges were noted with countries like Cameroon that provide data in broader age groups, which often appear adequate but offer limited detailed insight due to the extensive range of the 25-49 age group (Figure 3.2). While 56 countries have entered program data on ART by age group, only four countries have entered survey data on ART coverage by age into their Spectrum files, indicating a significant gap in making use of available data for analysis and validation.

Stover's presentation concluded with a call for more countries to use detailed age group data to improve their HIV treatment program assessments. The lack of such data in 80 countries underscores the need for continued efforts in data collection and validation to provide a more accurate picture of the HIV epidemic and its treatment coverage.

The key inquiries from the discussion reflected a concern with the accuracy and completeness of the data being used to inform HIV epidemic models and the need for enhanced tools and methods to ensure the data's utility in program planning and assessment. These are summarised below:

1. Visualizing time series and age groups: The discussion focused on the usefulness of visualizing time series data in larger age groups and how discrepancies in data could impact interpretations around incidence, especially in young adults. This led to a conversation about the best ways to visualize and analyse data over time to capture trends and discrepancies.
2. Goal for more countries to enter survey data by age group: The goal to motivate more countries to submit their survey data categorized by age group was brought up, along with strategies for facilitating this process. The availability of such data within the Spectrum database was mentioned, along with the simplicity of importing this data, suggesting a need for enhanced guidance to encourage the use of validation screens.
3. Comparing program and survey data: The possibility of directly comparing program data and survey data was queried, with a suggestion to use bar plots to compare model estimates with both data sources, to identify discrepancies in the data sources more clearly.
4. Improving fit to ART by age data: The discussion acknowledged the complexity in adjusting models when discrepancies in ART by age data are identified. The need for a clear, step-by-step approach to address discrepancies was discussed, recognizing that it may not be a straightforward process.

**Implications and recommendations for country data entry (survey and program) in 2024 estimates round**

**Jeff Imai-Eaton** provided an analysis of ART coverage data discrepancies, focusing on the difference between program data and PHIA survey ART coverage estimates. He discussed the impact of these discrepancies on HIV estimates and proposed methods to address them within the Spectrum model.

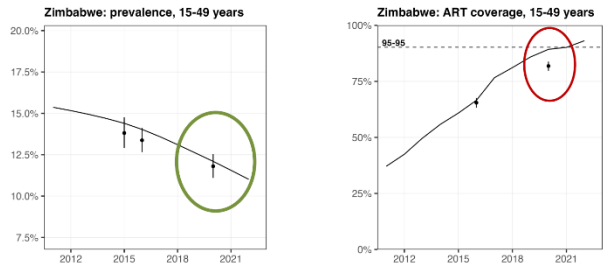
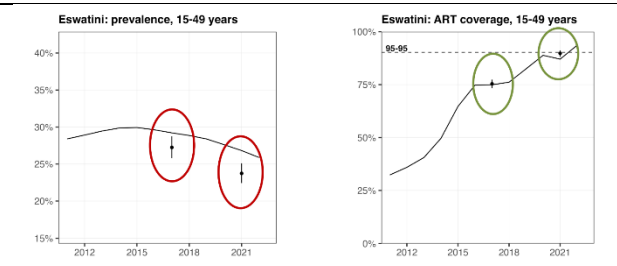
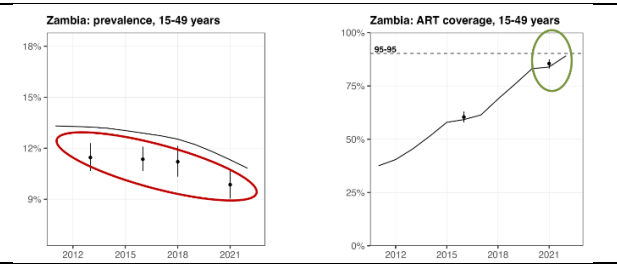
***Spectrum uses programme data to estimate ART coverage***

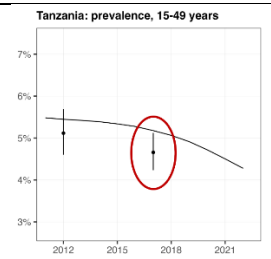
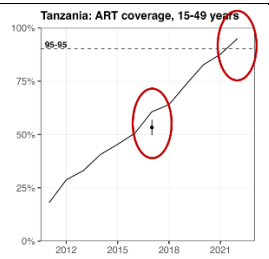
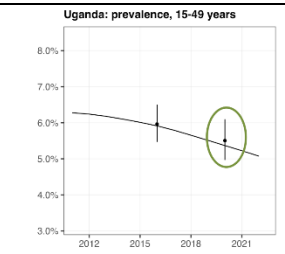
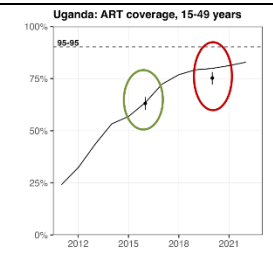
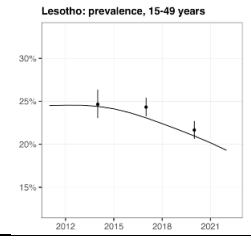
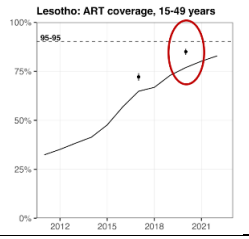
$$\text{ART Coverage} = \frac{\text{ART programme data}}{\text{Prevalence} \times \text{Population size [Modelled PLHIV]}}$$

Imai-Eaton underlined the significance of ART coverage as a measure for setting and evaluating program targets. He emphasized that as ART coverage reaches high levels, even small discrepancies become crucial.

Various scenarios observed in different countries were presented, reflecting on the biases that affect ART coverage estimates. Imai-Eaton presented specific case studies of countries with ART coverage discrepancies and the adjustments made to reconcile these data, and a summary is presented in Table 3.1.

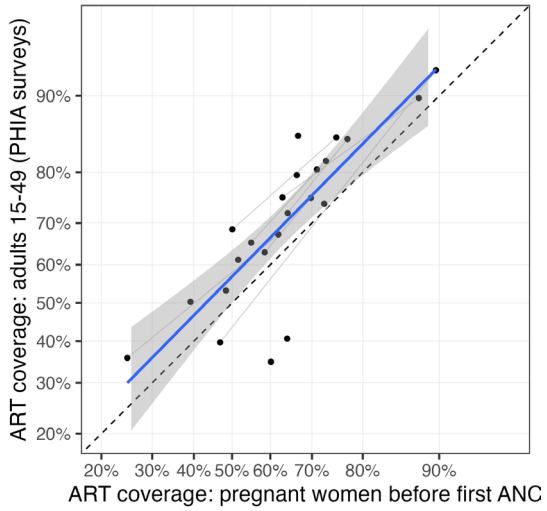
**Table 3.1: ART coverage reconciliation strategies in various countries.**

| Approach | Expected impact | Country | Observation | |
|---|---|---|---|---|
| 1. Calibrate to only to survey prevalence | The model should fit well to the survey prevalence data and typically slightly overestimate the survey ART coverage data in cases where the ART program data are higher. | Côte d'Ivoire | | Overestimates ART coverage by 8-10% in 2017. |
| | | Zimbabwe |  | Substantially overestimates ART coverage compared to the 80% ART coverage via survey in 2020. |
| 2. Fit to both survey HIV prevalence and ART coverage | Improved fit to ART coverage in survey year. Model prevalence higher than survey. | Eswatini |  | Increased number of PLHIV in the model than survey to reconcile ART coverage. Spectrum estimates 3% higher HIV prevalence than survey. |
| | | Zambia |  | Spectrum estimates 2% higher HIV prevalence. |

| | | | | |
|---|---|---|---|---|
| | | Tanzania |   | Spectrum estimates slightly overestimated for both ART coverage and HIV prevalence. ART data has continued increasing very steadily towards 95% without any saturation as seen in other countries. |
| 4. Adjusting ART input data | Improved model fit to survey prevalence and ART coverage. Deviation from reported program data. | Botswana |   | Fits well to ART coverage and HIV prevalence. Reduced number on ART by 5% in recent years. |
| | | Uganda |   | Fits well to HIV prevalence, slightly overestimates ART coverage. Reduced number on ART by 12% from 2017 onwards. |
| 5. Adjusting total population | Larger population → reconcile ART coverage | Lesotho |   | Applied *de facto* population, but ART coverage still underestimated. |

**ANC routine testing data as a predictor of population ART coverage**

Imai-Eaton discussed the hypothesis that ART coverage among pregnant women could act as a proxy for overall population ART coverage trends. The logic being that if pregnant women are accessing ART before their first ANC visit, it might reflect the broader ART coverage in the population. The Naomi model currently employs ART coverage data among pregnant women before their first ANC visit to model spatial patterns of ART coverage.




Methodology:

Population (15-49) ART coverage (from PHIA surveys) is highly correlated with year-matched ANC ART coverage, although on average population coverage is slightly higher (pregnant women are younger and more recently infected), as shown in Figure 3.3. A logit-scale linear regression was fit to this data and this model was used to predict population ART coverage using ANC coverage as input data in a few example countries, as shown below.

**Figure 3.3: Correlation between ART Coverage in Pregnant Women and Adults Aged 15-49.**

Country-specific observations:

In Figure 3.4A, Côte d'Ivoire presented a picture of ART coverage consistency, with Spectrum's estimate of 72% ART coverage among adults aged 15-49 being closely aligned with the predicted range based on ANC data. On the other hand, Zimbabwe's ART coverage by Spectrum (Figure 3.4B) appeared inflated at 93%, with ANC data indicating a stagnation in ART coverage, suggesting a probable overestimation. Tanzania's case presented in Figure 3.4C was particularly intriguing as Spectrum's 95% ART coverage estimation was significantly higher than the ANC data suggested, casting doubt on the country's attainment of the 95-95-95 targets.

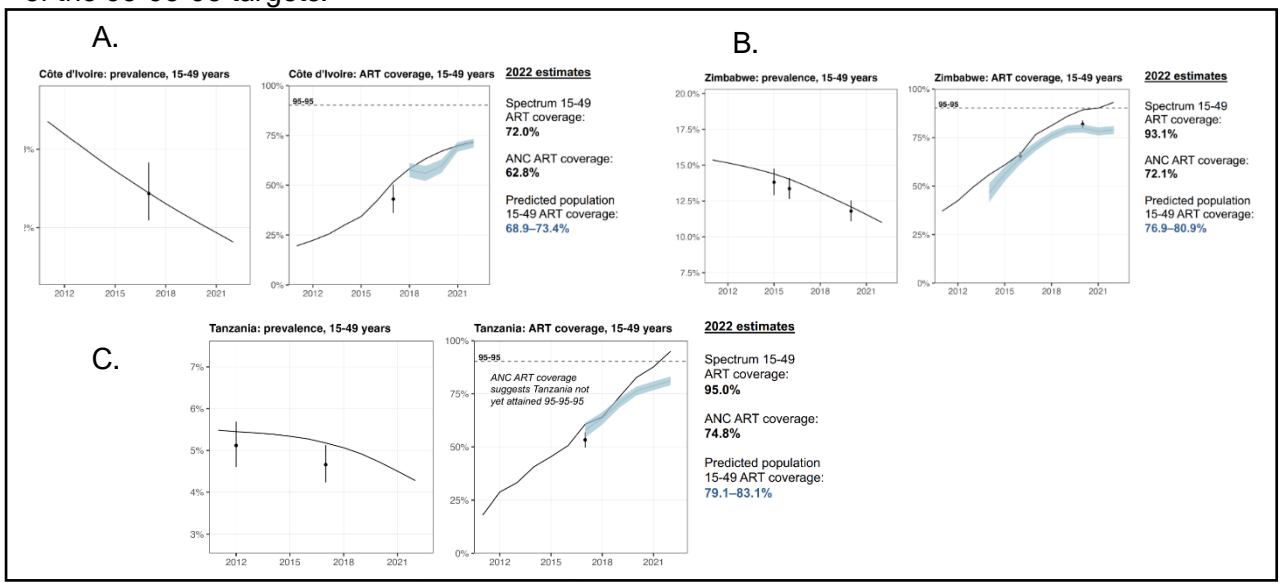



**Figure 3.4: Comparison of HIV prevalence and ART coverage estimates among adults aged 15-49 for A) Côte d'Ivoire, B) Zimbabwe C) Tanzania.**

Overall, it was underscored that ART program data might be systematically higher than survey estimates in the recent past. Using predicted population ART coverage based on ANC data in Spectrum can act as a visual validation tool. However, caution is recommended in using such data, especially when reaching high ART coverage. Imai-Eaton emphasized the importance of supporting country-level deep dives into ART data counting processes to better understand cross-sectional numbers on ART.

The presentation concluded with a call to action to address systematic disparities between program and survey data on ART coverage. The potential systemic trend of program data exceeding survey estimates in recent years was highlighted, alongside the uncertainty surrounding program data and population interpretations.

The discussion highlighted the complexity of reconciling ART coverage estimates with survey and program data. The need for further analysis and careful consideration of data quality was a recurring theme.

Key points from discussion:

1. Time variation in ART coverage estimation: There was an inquiry about the possibility of time variation in the relationship between ANC routine testing data and population ART coverage. However, it was noted that there's currently insufficient data to establish a definitive time trend.
2. Proposal for validation, not calibration: There was a clarification that the proposal to use ANC data is intended for validation purposes, not as a primary source for calibration or to confirm that 95-95-95 were met.
3. Examining sex differences in ART trends: In response to a suggestion, there was agreement on the need to analyse data for sex-based differences in ART trends, especially considering efforts to increase coverage among men.
4. Reconciling ART coverage estimates with ongoing incidence: Some countries' high ART coverage estimates were difficult to reconcile with ongoing HIV incidence. This mismatch suggested that the model's estimates of ART coverage might be overly optimistic, aligning more plausibly with the observed data trends when using the new approach.

Proposed recommendations:

- Support country deep-dives into ART counting processes.
- Plot predicted population ART coverage range based on ANC ART coverage trends.
- Make greater use of ART program data adjustments, especially when increases are unexpectedly steep or nearing high coverage without saturation.
- Exercise caution in certifying country attainments of 95-95-95 targets if ART coverage estimates are not corroborated by routine ANC ART coverage.

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

# Session 4: Definitions and reporting of HIV mortality among PLHIV

**Chair Cari van Schalkwyk** outlined Session 4's objective:

- Review definitions used in death data collections, and the need to realign or adjust for valid Spectrum fits to death data

Setting the scene for the session, van Schalkwyk underlined that in countries with high quality case-reporting and vital registration data, UNAIDS models are calibrated so that modelled deaths among PLHIV (minus age-sex matched background mortality) fit to AIDS-attributed deaths reported through vital registration. However, if the definition of model outputs representing 'excess mortality' among PLHIV are different from the definition of AIDS-related deaths recorded in vital registration, models may underestimate HIV-related deaths and as a result underestimate total epidemic size.

**Review of definitions and reporting of HIV mortality among PLHIV**

The primary goal of the review presented by **Adam Trickey** was to summarize, on a global and regional basis over time, the extent of differences between AIDS deaths and excess mortality among PLHIV. The review aimed to assist the Reference Group to make recommendations to UNAIDS regarding the definitions of AIDS deaths and how these should be communicated in mortality estimates. A full report is available at www.epidem.org, in the October 2023 meeting.

The method used for determining modelled mortality rates among PLHIV on ART, based on clinical cohort data minus sex/age-matched population mortality, has led to challenges. Notably, in Western and Central Europe and North America, a significant portion of excess PLHIV mortality is now from non-AIDS causes.

This presents two critical issues:

1. The CSAVR model within Spectrum, used by countries with high quality case-reporting and vital registration, calibrates Spectrum-modelled excess deaths to AIDS-attributed deaths reported through vital registration.

   However, if the definition of model outputs representing 'excess mortality' among PLHIV are different from the definition of AIDS-related deaths recorded in vital registration, CSAVR-based Spectrum estimates will underestimate total epidemic sizes (except in countries where VR-recorded AIDS-related deaths include excess non-AIDS deaths).

2. UNAIDS communicates the model-estimated deaths as 'AIDS-related deaths' and not as 'excess mortality among PLHIV', even though the latter is calculated by the models.

   This made little difference in early years of the epidemic, but over the past two decades, as ART has reached high coverage, these two indicators have increasingly diverged, in all world regions.

The review focused on three distinct aspects concerning the mortality among PLHIV:

**Review 1: Organizational definition of AIDS-related mortality**

The first review involved reaching out to various organizations that report causes of death among PLHIV to gather their definitions of AIDS mortality in relation to the WHO ICD-10 codes and guidelines for how HIV mortality should be measured.

A categorization framework by **Leigh Johnson**, for understanding deaths among PLHIV, includes four categories:

a) Expected deaths comparable to HIV negative population: This category includes deaths that occur at rates like those in an age and sex-matched control group without HIV. These deaths are not influenced by HIV status and represent the baseline mortality that would be expected in any population.
b) Excess deaths not attributable to HIV: These are deaths that occur at a higher rate in the PLHIV population compared to the general population but are not directly caused by HIV. For example, deaths due to substance use disorders might be more prevalent among PLHIV, not because of HIV itself, but due to related social and health factors.
c) Excess deaths attributable to HIV but not coded as AIDS: This category captures the indirect impact of HIV on mortality. It includes deaths resulting from conditions that are not classified as AIDS-defining but may occur more frequently in PLHIV due to the virus's effect on the body. An example is certain viral cancers that, while not directly caused by HIV, may have an increased incidence due to the compromised immune system of PLHIV.
d) Excess deaths due to AIDS: This final category specifically addresses deaths directly attributable to AIDS-related complications. It represents the most direct impact of HIV on mortality, encompassing conditions and illnesses that arise because of the advanced stage of HIV infection.

Review 1 findings:

1. WHO definitions and ICD-10 codes: The WHO uses ICD-10 codes (B20 to B24) to categorize AIDS-related deaths. The underlying cause of death, as defined by WHO, is the disease or injury initiating the events leading to death or the circumstances of the accident causing the fatal injury. Specific conditions like Kaposi sarcoma and Burkett lymphoma are considered direct consequences of HIV.
2. Responses from various organizations:
    - Institute for Health Metrics and Evaluation (IHME): Uses reclassification of "garbage codes" to address the varying quality of cause-of-death data across countries. Their rules after reclassification align with WHO's definitions.
    - South African MRC: Codes all pneumocystis and Kaposi's sarcoma deaths as AIDS-related due to their low incidence in the non-HIV population.
    - CDC: Uses a range of definitions, with *CDC Wonder* platform roughly following WHO's ICD-10 codes (B20-24) and other research using codes (B20-B24, O98.7, and R75). However, CDC's Atlas Plus platform defines AIDS deaths as any death among people who have ever received an AIDS diagnosis, differing significantly from WHO's approach.
    - UK Health Security Agency: Reports HIV-related mortality, including both AIDS and non-AIDS related mortality, but plans to start using a definition similar to the WHO definition with an additional category for possibly HIV-related deaths, particularly concerning non-AIDS defining viral-driven cancers.

3. Cohort-based reporting:
    - ART-CC Collaboration (Europe and North America): Uses the CoDe (Coding Causes of Death in HIV Pro Scholar), which deviates from WHO's definitions by incorporating additional patient information, such as CD4 counts at death.
    - IeDEA cohorts: Varies in approach, sometimes using HIV or AIDS-defining events on death certificates to indicate AIDS mortality. Other regions have not analysed cause-specific deaths or lack data.
4. Overestimation of AIDS-related deaths: Comparing the WHO definition with the CoDe protocol, a previous study indicated that WHO's ICD-10 based approach tends to overestimate the percentage of AIDS-related deaths.

In response to a query about the term 'overestimates' in the study findings, Trickey elucidated that the CoDe protocol is perceived as a more precise standard for identifying AIDS deaths compared to WHO's ICD-10 codes. The ICD-10 method, reliant solely on death certificate data, is considered less comprehensive than the CoDe protocol. CoDe involves a thorough review by clinicians of patient histories, including CD4 counts, to verify the cause of death and avoid inaccuracies, such as incorrectly attributing an accidental death to HIV.

**Review 2: Percentage of all mortality due to AIDS**

The second part was a rapid review and meta-analysis of 39 studies post-2016, predominantly from Eastern and Southern Africa (ESA) and Western and Central Europe and North America (WCENA), with a median participant age between 30 and 40 years, employing varied methodologies for classifying AIDS mortality. This review aimed to establish a clearer understanding of how significant AIDS is as a cause of death for PLHIV, including variations based on different factors such as geographical region, time, and ART status. There was no consistent definition of the minimum length of time individuals were categorized as being "on ART".

Review 2 findings:

1. Use of different protocols for classifying AIDS mortality:
    - 14 of the 39 studies used the CoDe protocol, mainly in high-income settings, with a few in China.
    - Six studies solely used ICD codes.
    - Two studies used verbal autopsy.
    - Seven studies based their classification on assessments by HIV physicians.
    - One study followed the WHO case definition.
    - One study classified AIDS mortality based on prior WHO stage 3 or 4 diagnosis.
    - Eight studies did not provide clear definitions for classifying AIDS mortality.

2. Trends in AIDS mortality over time:
    - The overall percentage of **on-ART mortality due to AIDS was found to be 52%**, showing regional variations.
    - The proportion of **deaths due to AIDS decreased over time** and with increasing national ART coverage at the study midpoint.
    - The percentages of on-ART mortality due to AIDS had a good agreement with Spectrum estimates, although they were slightly higher in the 2015 Spectrum estimates.
    - Off-ART mortality percentage was 51%, similar to on-ART mortality, but with high uncertainty. There was concern whether these figures accurately captured off-ART mortality or included individuals who were ART naïve at baseline but started ART during the study.

Limitations of the review:

- Variability in definitions of AIDS-related mortality across studies.
- Some studies included malignancies in their definitions, potentially overestimating AIDS mortality.
- A mix of prospective and retrospective studies, with varying eligibility criteria.
- Lack of data from certain regions, particularly Eastern Europe and Central Asia.
- Unreliability of data for off-ART mortality, with some studies showing implausible similarities in mortality percentages for on and off ART cases.

Key points from discussion:

1. Definition of minimum ART duration in ART papers: There was no consistent definition for the minimum ART duration in the papers, potentially resulting in individuals being classified as "on ART" even if they had been receiving treatment for only a brief period before their death.
2. Spectrum results aligning with review results: An inquiry raised the possibility of adjustments made in Spectrum artificially influenced the correspondence of review results with Spectrum results. The need for further examination was emphasized.
3. Consistency in relative levels across years: Lastly, the discussion touched upon the consistency of relative levels across regions in the years 2010, 2015, and 2020. It was observed that while relative levels remained consistent across these years, proportions in the 2015 period, where most studies were centred, appeared systematically lower, prompting questions about the reasons behind this discrepancy.

**Review 3: Percentage of excess mortality due to AIDS**

The third part of the review differed from the second review in that it specifically aimed to quantify what percentage of excess mortality among PLHIV could be attributed to AIDS.

Scope and sample of the review:

- Only 10 studies were initially eligible, with 7 finally included after removing duplicates.
- All studies were from high-income countries, with five from the Western and Central Europe and North America (WCENA) region, one from Japan, and one from Korea.
- The studies predominantly involved men who have sex with men.
- No studies reported data for people not on ART.

Methodology and classification:

- Two studies used the CoDe protocol, one used a physician panel, and four relied on ICD-10 codes for classifying AIDS-related mortality.
- The studies did not consistently report the percentage of participants on ART.

Review 3 findings:

- Over 1.3 million person-years and 16,700 deaths were analysed.
- The overall mortality rate was 0.0129, with AIDS-related mortality at 0.0045.
- The pooled percentage of all mortality due to AIDS was 40%, slightly higher than the 32% found for the WCENA region in the previous review.
- **The pooled percentage of excess mortality due to AIDS in these studies was 51%.**
- There was a variation in the percentage of AIDS-related excess mortality, ranging from 29% to 73% across studies.

Limitations and considerations:

- No data were available from SSA or LMIC, limiting the generalizability of the findings.

- Most of the studies were set in men who have sex with men-majority epidemic settings, and the applicability of findings to other epidemic contexts (like IDU-driven epidemics) is unclear.
- Variations in the definitions of AIDS-related mortality across studies, with a systematic higher percentage of AIDS mortality likely in studies using only ICD-10 codes.
- Inability to exclude deaths where cause of death data was missing.
- An additional aspect of Spectrum modelling excess mortality related to overdoses among people who inject drugs was noted, though not directly relevant to the reviewed studies.

Key points from discussion:

1. Inclusion of studies post-2016: Clarification was sought on whether the studies included in both reviews were published after 2016. Trickey confirmed this and suggested the possibility of expanding future reviews to include earlier data. He noted, however, that incorporating older data might not significantly contribute to the current review's relevance. Earlier ART regimens were less effective in preventing mortality compared to current ones, and the mortality due to AIDS in older studies would likely be higher, making them less useful for comparison with recent estimates.
2. Analysis of broader time periods: It was proposed that there's no reason not to analyse a broader range of time periods, provided the comparisons are made within the same period.
3. Understanding time trends in ART-CC analysis: There was a query about the ART-CC analysis, specifically whether there were trends observed in rates of excess non-AIDS mortality over time. Trickey indicated that AIDS-related mortality drastically reduced while non-AIDS mortality either increased or remained steady, particularly in high-income countries.
4. Impact of stable ART on mortality trends: Trickey pointed out that as people on ART age, they are more likely to die from various causes. However, being on stable ART generally decreases the likelihood of death, balancing the increased risk due to aging.

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

# Session 5: Mortality among people on ART

Session 5, **chaired by Jeff Imai-Eaton**, focused on two primary objectives:

1. Review updated IeDEA mortality analyses and impact on Spectrum estimates

This topic reviewed the latest analysis of mortality among people on ART data from the IeDEA cohort collaboration globally. The goal was to understand how this updated analysis influences the estimates in the Spectrum model.

2. Review the impact of assumptions about ART interruption on AIDS-related mortality estimates in different settings

The session reviewed evidence on the mortality outcomes among individuals who have previously interrupted ART, compared to those who have never received ART treatment and initiated ART for the first time. Currently, the Spectrum model assumes those interrupting treatment and reinitiating ART experience the same mortality as those treatment naïve in the same sex, age, and CD4 category.

**IeDEA adult mortality analysis**

**Reshma Kassanjee** presented an updated analysis of all-cause mortality rates among adults (aged ≥ 15 years) on ART who initiated ART at age ≥ 15 years. Mortality was segmented by region, sex, current age, CD4 count at ART initiation, current ART duration, and calendar time.

Data and methods:
Using data from IeDEA (https://www.iedea.org/regions/), the study encompassed PLHIV who accessed care across seven global regions. The time span of person-level longitudinal data was from 2001 to 2021, capturing person characteristics, clinic visits, ART starts, and CD4 and viral load tests.

The analysis involved multivariable mixed effects Poisson regression, considering time-varying covariates and a random effect for treatment programs. The adjusted analysis included both IeDEA routine care data and tracing study data, with the latter aimed at determining outcomes for samples of people classified as loss to follow-up (LTFU).

**Base case analysis:**

- The base case analysis changed the ART eligibility exclusion criteria from previous IeDEA analyses to include individuals who started ART with a CD4 count above the threshold used to determine ART eligibility at the time. This change aims to include individuals with potentially high mortality who were previously excluded.
- Individuals with low viral load (<1000copies/ml) measure close to ART start date (indicating they are already on ART) were excluded from the analysis.
- Individuals who interrupted treatment are no longer excluded from the analysis.
- Previous analyses censored individuals at the time of last ART visit, but the current analysis assumes people are on ART for 1 year post last visit.

**Two scenarios were compared to the base case:**

1. **First time ART initiators**: This scenario excluded individuals who previously interrupted treatment.

2. **Alternative ART Eligibility Rule:** This scenario excluded individuals who started ART with a CD4 count above the threshold used to determine ART eligibility at the time.

Study results (in regions outside Africa):

- In Asia-Pacific, Latin America and North America, mortality rates (for those on ART <1 year or >= 1 year) continue to decline. Impact of using updated mortality estimates for these regions in Spectrum was evaluated in the next presentation by **John Stover**.

Study results (in regions within Africa):

- Mortality rates were analysed and reported by time and region, and on ART duration.
- Tracing studies aim to find ART clients who were LTFU and estimate the proportion who died. A parametric survival model was fitted to data from several such tracing studies conducted in Africa, and IeDEA mortality estimates were adjusted using this information in a simulation model.
- Recent mortality trends in African regions are highly sensitive to these tracing study adjustments and the definition of LTFU.
- Before and after adjustment using tracing study data, mortality rates in all regions within Africa, for ART durations of <1 year and >=1 year, have slightly increased in recent years.

Kassanjee emphasized how both the direction and magnitude of mortality trends can be swayed by adjustments based on tracing study data and definitions of LTFU.

Challenges and considerations:

- The study noted a decrease in the availability of CD4 count data at ART initiation, particularly after the implementation of 'treat all' guidelines in African regions (excluding South Africa), leading to an increasingly unrepresentative sample size.
- There is an ongoing need to consider whether to remove individuals without CD4 data from the analysis or to impute this information, acknowledging that imputation requires assumptions that may introduce bias. A funded study is set to measure CD4 in individuals initiating or re-initiating ART to provide a clearer picture of the actual distribution of CD4 counts and to identify re-initiators.
- The reliance on older tracing study data for recent years might not be appropriate, particularly considering COVID-19 disruptions, changes in care models, and recent database closures.
- The presentation emphasized the need for more post-2020 data and engagement with IeDEA regions for optimal use of IeDEA data in future UNAIDS estimates.
- Kassanjee provided preliminary data indicating that the proportion of time during which people have not been seen for a year was estimated to be between 5% and 10% across different regions.
- Removing this proportion from the denominator would increase mortality estimates by about 5% to 10%.
- Despite this adjustment, the relative increase in estimates was not particularly large compared to the overall uncertainty and other methodological decisions being made.

Kassanjee concluded that the estimates for regions outside Africa might be acceptable for use in Spectrum's next update, pending validation. However, for regions within Africa, the current estimates were deemed too preliminary due to the challenges presented. More engagement with the IeDEA regions and the Reference Group is needed to determine how to best use the data to inform UNAIDS estimates moving forward. Given that the mortality

correction applied in the study is based on tracing studies conducted mostly before 2010, applying these to data up to 2021 may not be appropriate without considering recent changes in care delivery and the impact of the COVID-19 pandemic.

**Impact of updated IeDEA analysis on Spectrum estimates**

**John Stover** then provided insights into the potential impacts of the updated IeDEA estimates on Spectrum estimates in Asia Pacific and Latin America.

Four different scenarios comparing time trends were highlighted:

1) The current Spectrum mortality estimates from the 2019 IeDEA analysis including data up to 2017,
2) The new base case from the IeDEA analysis presented by Kassanjee,
3) Relative to 'base case', applies previously used rules to exclude people who may be returning to care rather than starting ART for the first time (rules since revised for the 2023 analysis).
4) Relative to 'base case', applies previously used exclusion rules to exclude people who started ART for reasons other than eligibility based only on CD4 (exclusion rule removed in 2023 analysis).

Stover noted that for both Asia and Latin America, **time trends** in overall mortality estimates were very similar across all four scenarios. A comparison of annual mortality rates by CD4 count at ART initiation showed that the current values and the new base case values were fairly similar. However, the alternative scenarios suggested lower mortality rates. This pattern was consistent across different age groups and for both males and females.

The updated rates were applied to 22 countries in the Asia-Pacific and 12 in Latin America (Table 5.1) to examine their effects on the estimates of deaths. The analysis involved:

- Maintaining custom mortality patterns for countries that have them, such as China, DPR Korea, Sri Lanka, PNG, Thailand, Vietnam, and Brazil.
- Using mortality multipliers where countries adjusted rates to match program data better, with 13 countries in Asia-Pacific using a multiplier of 4 and others like Laos, Myanmar, and Singapore applying custom multipliers.

**Table 5.1: Countries included to calculate effects of new rates on estimates of deaths.**

| | **Asia-Pacific** | **Latin America** |
|---|---|---|
| Number of countries included | 22 | 12 |
| Number using custom mortality pattern | 7<br>China, PDR Korea, Sri Lanka, PNG, Thailand, Vietnam, Brazil | 1 |
| Number using mortality multiplier | 13 countries: 4<br>Laos: 5.66<br>Myanmar: 11<br>Singapore: 2 | 0 |

The presentation detailed how the new base case rates and alternative scenarios (such as the alternate starting definition and confirmed CD4 eligibility) would affect the number of deaths among those on ART, shown in Figure 5.1.

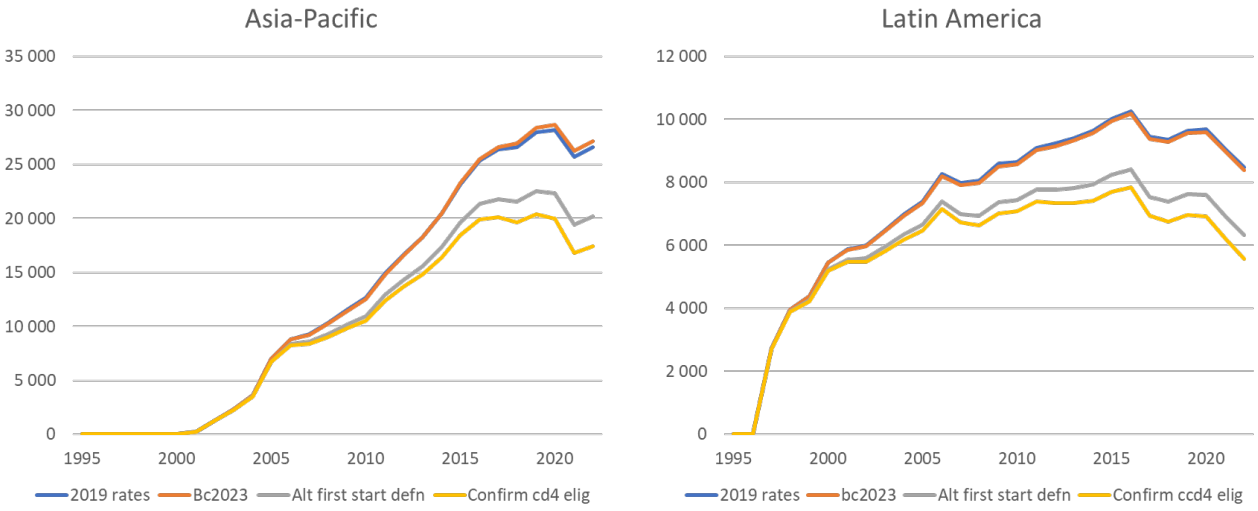



**Figure 5.1: Deaths to those on ART.**

The 2022 differences illustrated in Figure 5.1 were:

| | **Asia-Pacific** | **Latin America** |
|---|---|---|
| Base case 2023: | +2% increase | -1% decrease |
| Alternate first start definition: | -24% decrease | -25% decrease |
| Confirm CD4 eligibility: | -35% decrease | -34% decrease |

Similar patterns were observed when looking at all excess mortality among PLHIV as seen in Figure 5.2, although the differences were smaller due to a larger proportion of deaths occurring among those not on ART.

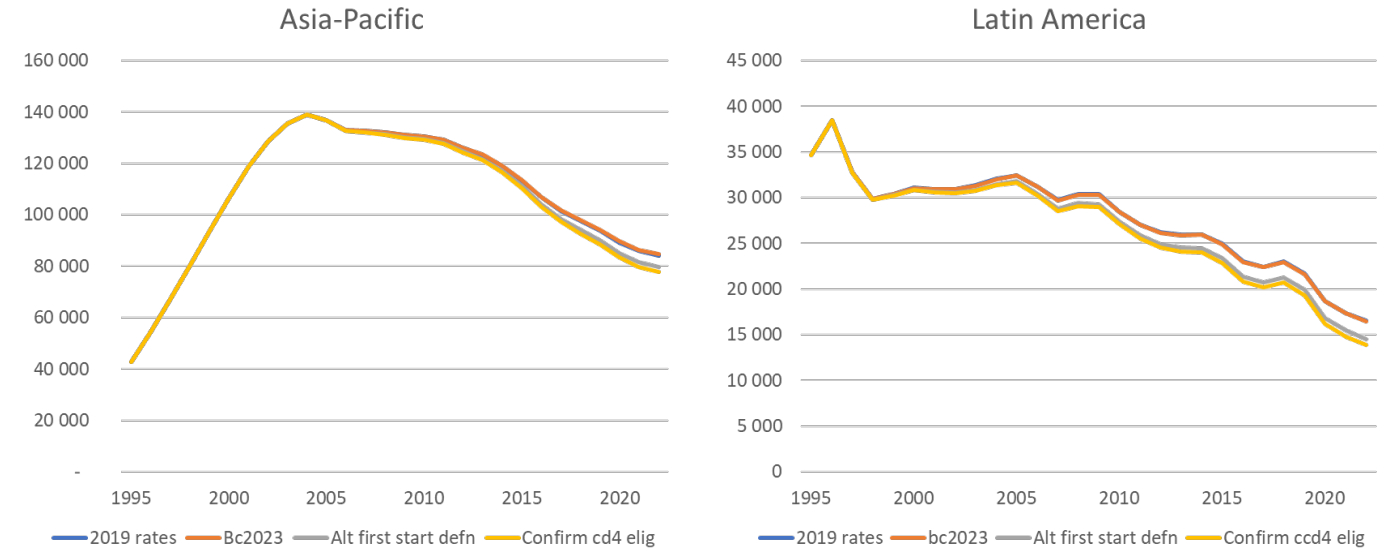



**Figure 5.2: All excess mortality among PLHIV.**

While the base case updates compared to current rates showed less than 2% difference for most countries, the alternative scenarios presented potential for more significant changes and considerable differences at the country level in some countries.

Conclusions drawn from the analysis:

- A standard update of the IeDEA analysis would have minimal impact on estimates of AIDS deaths in the Asia-Pacific and Latin America regions.
- Alternative definitions could potentially reduce the estimated number of deaths to people on ART by about 30%, though the actual differences may be less if countries adjust their regional patterns.
- The analysis underscored the need for custom adjustments to cater to the specific data and trends of individual countries.

After both presentations, participants engaged in discussion, which is summarised below:

1. The discussion began with an inquiry into the accuracy of past projections, specifically the 2018 estimates, compared to current time trend changes. Stover confirmed the projections closely matched the new data, demonstrating the trends were well anticipated.
2. Commentary was provided on Kassanjee's presentation regarding the decline in people getting CD4 counts at ART initiation. It proposed that if the rate ratios by CD4 count are stable, the analysis could be rerun without excluding those without CD4 count data, leveraging the entire dataset to discern trends over recent years. Kassanjee agreed that re-running the models without CD4 data would be possible if beneficial.
3. Model assumptions about the distribution of CD4 among individuals initiating treatment was questioned, as data suggests that this might have changed over time. It was clarified that the model itself estimates the changing distribution, implying that such an analysis is built into the modelling process.
4. The approach was supported by another participant, who suggested that time trends are crucial for the outcomes, hinting at potential relevance to subsequent discussions on current mortality trends in Spectrum.
5. Potential implications of spacing ART visits to a six-month interval were raised, and how IeDEA might account for this when considering a shorter LTFU definition. It was addressed by indicating the possibility of defining LTFU as a one-year period but also acknowledging the complexities introduced by COVID-19 and recent data effects, which necessitate further exploration with a more recent dataset unaffected by COVID disruptions.
6. A new NIH-funded study within IeDEA African regions was highlighted. This study will measure CD4 counts at ART initiation, providing valuable insights into the current distribution of CD4 counts and to identify re-initiators, with results expected in a year.

The table below shows a side-by-side comparison.

- **Kassanjee's results (regions within Africa)**: Highlighted the sensitivity of mortality estimates to methodological changes and data completeness, especially in African regions.
- **Stover's results (Asia Pacific and Latin America)**: Focused on the practical application of these mortality estimates in Spectrum models, confirming that the new base case and alternative scenarios could significantly impact national estimates, particularly in Asia-Pacific and Latin America.

Kassanjee and Stover compared mortality rates and their respective Spectrum model outputs, noting discrepancies that required further investigation. Kassanjee suggested that decisions made might need revision upon closer examination of mortality rates and their impact on Spectrum output results.

Stover highlighted that in the Asian files, many countries adjusted their default mortality rates significantly. Some countries used a multiplier to fine-tune deaths on ART, often multiplying by four, which led to substantial changes. He questioned the IeDEA team about why country data seemed to differ so significantly from the default rates provided by Spectrum.

Kassanjee responded by referencing a survey conducted earlier in the year to assess the completeness of mortality data in Asia Pacific programs. The survey results were unclear, but literature suggested some mortality might not be captured. Revisiting this issue would require a separate analysis and identifying individuals to lead this new project. Despite the challenges, the potential need for adjustments was acknowledged.

It was also noted that the reported deaths among those on ART were significantly higher than what the default rates produced, indicating a need for further discussion and analysis. There was agreement on the necessity for some adjustments but also recognition that it was outside the current scope due to existing challenges.

**Recommendations for non-sub-Saharan African regions:**

1. Adoption of the new base case analysis for non-sub-Saharan African settings, subject to further review and validation, was recommended. This adoption would depend on a thorough examination of the raw data and the impact of methodological decisions on the final mortality estimates.
2. There was a proposal for a smaller group discussion for methodological points of interest:
    a. The decision on who to exclude from the analysis.
    b. Within Africa, how to perform adjustments.
    c. Whether to remove individuals without CD4 data from the analysis or to impute this information, acknowledging that imputation requires assumptions that may introduce bias.
3. There was an emphasis on the need to revisit the Asia Pacific multiplier in light of the new mortality rates. Establishing a default scalar based on the collective experience of countries was discussed, with a suggestion to retain the multiplier of four but revisit and recalculate this assumption if the new rates differed significantly from the old ones.

**Quantifying mortality in sub-Saharan Africa:**

The discussion highlighted challenges in interpreting sub-Saharan African data. Key concerns included:

- The study noted a decrease in the availability of CD4 count data at ART initiation, particularly after the implementation of 'treat all' guidelines in African regions (excluding South Africa), leading to an increasingly unrepresentative sample size.
- The reliance on older tracing study data for recent years might not be appropriate, particularly considering COVID-19 disruptions, changes in care models, and recent database closures.

**Recommendations for sub-Saharan Africa**

- Spectrum should continue to use the current mortality rates for sub-Saharan Africa.
- Proposal to conduct a comprehensive review of the mortality rate assumptions in Spectrum, particularly for South Africa to see if further information can provide more accurate estimates.
- Conduct a focused analysis on South Africa to identify unique trends and validate regional estimates.
- Potential implications of spacing ART visits to a six-month interval were raised, and how IeDEA might account for this when considering a shorter LTFU definition. It was indicated that defining LTFU as a one-year period could be considered, but this would need to acknowledge the complexities introduced by COVID-19 and recent data effects.
- Engage with IeDEA regions to gather additional data and insights, focusing on post-2020 data to account for COVID-19 impact.

Presentations then focused on the mortality outcomes for adults who resumed ART after experiencing interruptions in treatment to evaluate whether Spectrum's current assumptions need adjustments, presented by Haroon Moolla (IeDEA South Africa) and Adam Trickey (ART-CC).

**Mortality in ART re-initiators: IeDEA South Africa**

**Haroon Moolla's** presentation focused on the mortality outcomes in adults who resumed ART in South Africa after experiencing interruptions in treatment. The focus was on mortality in isolation, unlike previous studies that combined mortality with other outcomes like progression to AIDS or LTFU. IeDEA-SA data was used, focusing on adults who initiated ART from 2004 to 2019. The study included 4,700 participants who resumed ART and were virally suppressed upon return. A notable observation was that a significant portion of these individuals were 'club patients' who are stable enough for six-monthly follow-ups, many of whom returned to care just within two weeks of the 180-day threshold, implying they might be part of this stable group.

Moolla highlighted adjusted hazard ratios indicating that both early (within 6 months of 1st ART initiation) and late interrupters (after 6 months) had a reduced risk of mortality compared to those who never interrupted. Sensitivity analyses revealed that assuming higher CD4 counts upon return to care did not significantly alter hazard ratios; longer interruption durations increased hazard ratios.

The presentation also addressed the limitations of the study, such as the challenge in distinguishing gaps in care from actual ART interruption and potential confounding factors like socioeconomic status. The exclusion of a substantial number of participants due to missing CD4 count data posed a risk of bias if the missingness was systematic.

**Mortality in ART re-initiators: ART-CC**

The aim of **Adam Trickey's** presentation mirrored that of **Moolla's** study using the Antiretroviral Therapy Cohort Collaboration (ART-CC) data from the WCENA region.

Key definitions and analytical approach:

- Interruptions in HIV specialized health care was defined as breaks of at least 180 days without lab CD4, RNA, ART, or visit records, followed by a return to care without VLS.
- LTFU was distinct from interruptions and was based on the last contact date plus 365 days if the date was before the administrative censoring date.
- Sensitivity analyses were conducted using interruptions of varying lengths (270 and 535 days) to understand the impact of definition duration on results.
- The analysis method used was Cox regression to compare mortality rates across different interruption groups, adjusting for various factors such as sex, mode of HIV acquisition, year of ART initiation, age, and CD4 count.

Trickey approached the study with an initial definition of care interruptions as breaks of 180 days, yet he expressed reservations about this threshold. In the ART-CC context, he suggested that such a duration might not sufficiently represent a substantial gap due to the longer intervals between follow-up visits commonly observed in these cohorts. To address this, he conducted sensitivity analyses using alternative definitions of interruptions — 365 days and varying intervals up to 535 days — to provide a broader understanding of the impact of care gaps.

Findings:

- The results indicated that mortality rates were higher after interruptions in care, with early interruptions (reinitiating within less than six months of first ART initiation) and late interruptions (reinitiating after six months or more) both showing increased mortality compared to no interruption periods.
- Sensitivity analyses yielded very similar results to the main analysis, demonstrating the robustness of the findings.
- For individuals with CD4 values available post-interruption, the mean CD4 counts were higher after the interruption than at baseline.

Comparison between studies
In South Africa, re-initiators had lower mortality than those initiating for the first time, whereas the opposite was observed in WCENA.
One critical difference in Trickey's study, compared to Moolla's, was the linkage to death registries. While South African cohorts had consistent linkage, ART-CC cohorts varied, with most (but not all) having this connection. Despite this inconsistency, Trickey was confident in the ascertainment of mortality data within the ART-CC cohorts.

The presentation highlighted stark differences between the results obtained from ART-CC data and those from the IeDEA-SA analysis. For instance, dropping post-gap follow-up of individuals returning from a gap with VLS led to differing interpretations of mortality rates between the datasets.

Trickey’s efforts to compare with Moolla's findings underscored a common goal in HIV research: to understand and accurately portray the implications of treatment interruptions on mortality. The discussion brought to light the intricacies in understanding ART interruptions and their consequences on patient outcomes, and key points are summarised below:

1. There was an inquiry about whether the 180-day interruption definition included the number of days of medication given to the client. It was confirmed that the analysis did account for this, and concerns were expressed about higher mortality rates upon return to care with higher CD4 counts in the ART-CC data, which seemed

counterintuitive compared to South African data where suppressed viral loads upon return indicated higher survival rates. This discrepancy pointed to the need for careful consideration of the implications of treatment interruptions and their impact on mortality.

2. In some contexts, stable clients receive 180 days of treatment and visit the facility biannually. Clarification was requested on whether these patients would count as interruptions based on an interruption definition of 180 days**.** It was explained that in the analysis, this issue was addressed by including the first 180 days of all gaps in treatment. Additionally, to mitigate potential overestimation of treatment interruptions, they excluded individuals who were VLS from their analysis. It was acknowledged that this might be an overcorrection, but in the South African context, it was deemed not too excessive.
3. A participant noted the apparent changes in the adjusted hazard ratios for risk of mortality among those who interrupted ART, which differed from what Moolla presented at a previous conference. Moolla clarified that the current analysis differed from the previous one as it reset the duration to zero at each ART restart, rather than measuring from the first ART initiation.
4. The discussion acknowledged the need to distinguish between treatment interruptions and LTFU, emphasizing that treatment interruption implies a period without ART, which impacts mortality.
5. Kassanjee provided preliminary data, indicating that the proportion of time during which people have not been seen for a year in their dataset was between 5% and 10% across different regions. Removing this proportion from the denominator would increase mortality estimates by about 5% to 10%. This suggests that treatment interruptions are relatively common and should be carefully considered in mortality models.
6. Current Spectrum assumptions: Spectrum currently does not distinguish between first-time initiators and re-initiators. All individuals are treated as part of a single pool with equal mortality rates. **Concern:** By resetting the duration to zero every time a patient restarts treatment, there is a risk of overstating mortality. However, Trickey's findings showed the opposite.

**Recommendations**

1. More detailed work is required to reconcile the differences between the results from IeDEA-SA and ART-CC analyses. The group agreed that it is premature to make immediate changes to Spectrum’s assumptions without a deeper understanding of the factors driving the observed differences in mortality outcomes between Moola and Trickey’s results. The recommendation was to keep the current model unchanged.
2. Future work:
    - Conduct a detailed analysis of the distribution of death timings on ART and consider whether longer durations on treatment should be stratified in the model.
    - Explore the possibility of adding an assumption about the proportion of people reinitiating ART to refine the model further.

Perform analysis around the impact of reinitiating ART on mortality to ensure accurate representation in Spectrum.

**On-ART all-cause deaths validation: 2023 data entries and Spectrum fits**

The presentation by **Eline Korenromp** focused on the validation of all-cause deaths on ART using 2023 data entries and Spectrum fits.

Overview of the study:
A new AIM validation option in Spectrum was introduced. This feature was used to analyse data from national HIV programs from 27 countries across various regions including Asia-Pacific, Western Central and Eastern North Africa, the Middle East, and the Caribbean. Notably, no data from SSA or Latin America was included in this study. These data included records of all-cause deaths among people on ART, ART initiation, and interruption rates. Program data were compared with Spectrum estimates to identify discrepancies and validate the model's accuracy. The validation process focused on analyzing data entries from 2023, considering various factors such as age, sex, and regional differences.

Korenromp outlined the process for integrating program data into Spectrum:

1. Data formatting:
    - Standardize data, ensuring consistency in units, age categories, and demographic variables to match Spectrum's format.
    - Use provided templates for consistent and easy data entry compatible with Spectrum.
2. Validation checks:
    - Verify all required fields are filled, incomplete data leads to inaccuracies.
    - Ensure data aligns with expected patterns and historical data for age distributions and sex ratios.
    - Identify and correct outliers, duplicates, and other anomalies.
3. Choose appropriate algorithms matching data structure and analysis objectives.
    - Map program data to corresponding fields in Spectrum using the chosen algorithms.
4. Data entry:
    - Enter formatted and validated data into Spectrum as per the user guide instructions.
    - Perform validation within Spectrum, to check data integrity and identify issues.
    - Calibrate the model with new data entries and adjust model parameters as needed to align with validated data.

Data analysis and findings:
Each country's data was assessed for the fit between the incidence model used (such as direct incidence input, ECDC, CSAVR, AEM, EPP-Concentrated) and the reported on-ART deaths. A detailed breakdown of the analysis was given, highlighting whether the Spectrum model's estimates were above, below, or in line with the program-reported on-ART deaths.

Country-specific insights:

- Good fits: Spectrum model showed several good fits in countries with high-quality and rich data. Notable examples include the UK, the Netherlands, and Iraq. These countries demonstrated a strong alignment between the Spectrum model estimates and the reported on-ART deaths.
- Spectrum above reported on-ART deaths: In most cases, Spectrum's estimates were above the reported on-ART deaths, particularly in recent years where Spectrum's estimates were rising but the reported on-ART deaths were either stable or falling. This trend was observed across different incidence models including CSAVR (New Zealand, Egypt), AEM (Myanmar, Cambodia), and EPP (Tunisia).

- Exceptions: There were exceptions where Spectrum's estimated on-ART deaths were lower than what was reported. This was specifically noted in France (using a direct incidence model) and Cuba (using the CSAVR model). Both these countries had a long series of high-quality data reported on on-ART deaths, which contrasted with Spectrum's lower estimates.

Korenromp discussed the challenges in ensuring consistent definitions and data quality across countries. The presentation noted trends such as the potential incompleteness of on-ART death data, possible underestimation of early mortality on ART by Spectrum, and the absence of certain data entries in the majority of the 2023 files. Additionally, the allocation of ART by mortality risk in the model was questioned, considering its time-constant parameter.

Korenromp suggested streamlining Spectrum results options for deaths among PLHIV. The entry of all-cause mortality data, particularly for countries using the CSAVR, was emphasized as crucial for distinguishing between AIDS-related deaths and deaths among those on ART.

The discussion highlighted the need for robust validation methods to ensure the accuracy of Spectrum estimates, and the importance of consistent application of these methods across all regions to maintain data integrity was highlighted.

- It was recommended that countries strengthen their data collection and reporting mechanisms, ensuring that all relevant data on ART initiation, interruptions, and mortality are accurately recorded and reported to UNAIDS.
- The need for clear guidelines and support for countries to improve their data management practices was highlighted.
- Capacity-building initiatives were proposed to train national program staff on data management and the use of Spectrum, highlighting the need for ongoing training and support to maintain high standards of data management.

Regular review and feedback mechanisms should be established to continually assess the accuracy of Spectrum estimates. This includes periodic validation exercises and the incorporation of feedback from national programs and other stakeholders. The importance of continuous improvement and adaptation based on new data and insights was emphasized.The full recommendations arising from this session are shown in **Appendix A**.

# Session 6: ART interruption rates

The session's primary objective was outlined by **chair Leigh Johnson**: To recommend realistic, non-zero regional default rates for interruption from ART for use in the 2024 estimates.

**ART interruption: ART-CC**

**Adam Trickey** presented an analysis focusing on the rates of ART interruptions, particularly examining the rate of the first interruption for simplicity. This analysis, building upon his earlier presentation, sought to provide a clearer understanding of ART interruption patterns.

Trickey used Cox regression models for his analysis, opting for a 365-day threshold to define interruptions instead of the 180-day threshold used in his previous analysis (which was chosen to correspond to the IeDEA-SA analysis). The decision to use a longer duration aimed to capture more meaningful interruption events, and he also employed sensitivity analyses to strengthen his findings.

***Results***

*- 89,187 people living with HIV*
*- 17% female, 44% men who have sex with men*
*- 8709 (9.8%) had a ≥365-day treatment interruption*
*- 536,437 person-years included in total*
*- Crude rate of interruption of 16.2 (95% CI: 15.9-16.6) per 1000 person-years.*

Notably, about 10% of the participants experienced a treatment interruption. However, Trickey conveyed uncertainty with these results, indicating the complexity and challenges in accurately measuring and interpreting ART interruption rates.

Furthermore, the analysis delved into hazard ratios to identify predictors of ART interruptions. Consistent with existing literature, Trickey found that older age was associated with a lower likelihood of experiencing an interruption. Conversely, injecting drug use emerged as a significant predictor for interruptions, aligning with previous research findings in this area.

The session concluded with Trickey highlighting the overall rate of 16.2 interruptions per thousand person-years.

**ART interruption: IAS**

**Anna Grimsrud** presented on ART Interruption Rates, emphasizing the critical need for a nuanced understanding of these interruptions and their impact on patient outcomes, particularly mortality.

Grimsrud emphasized that ART interruption rates cannot be zero. She highlighted the movement of patients in terms of whether they're early, late or established on ART, in an interruption or back in care, and the crucial need to understand how this movement correlates with subsequent morbidity and mortality. The lack of a universal definition for treatment interruptions complicates the understanding of patient experiences and the impact of interruptions on mortality.

Terms like "Loss to Follow-up," "Interrupted," "Late," and "Disengaged" are used variably across studies and contexts, ranging from a few weeks late for a scheduled appointment to more extended periods without contact with healthcare services. Thus, there is variability in

mortality risk based on the duration of interruption and the point in the treatment cascade at which individuals find themselves.

Grimsrud advocated for a paradigm shift in how we view the HIV care cascade, proposing a cyclical rather than a linear model of treatment that works for the patient. This perspective acknowledges the reality of patients moving in and out of care, which is crucial for understanding and addressing the challenges in HIV treatment retention and outcomes.

The influence of various factors like “Treat All” policy, same-day start, improved ART regimens, differentiated service delivery, decentralization of services, as well as the client experience on return would also affect the rate and nature of interruptions. The negative experiences of patients returning to care, such as being reprimanded for lateness or missing documents, were noted. These experiences can influence the likelihood of future interruptions and impact overall treatment outcomes.

PEPFAR data which does not include South African data, shows an increasing trend in the proportion of clients on longer multi-month dispensing intervals, suggesting a shift towards less frequent ART pickups over time, as seen in Figure 6.1.

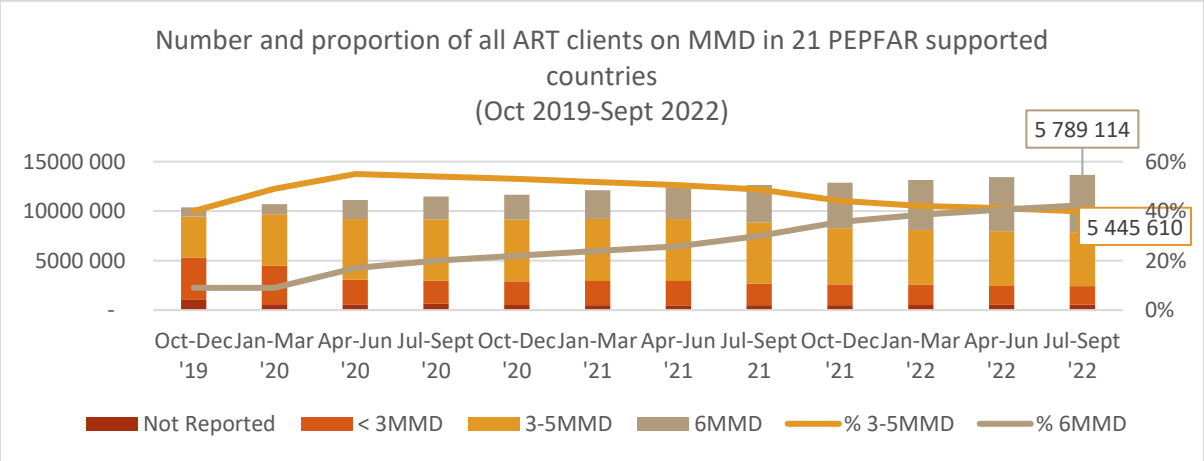



**Figure 6.1: The number and proportion of all ART clients on Multi-Month Dispensing (MMD) across 21 PEPFAR-supported countries from October 2019 to September 2022.**

In South Africa, a significant number of PLHIV (1.6 million) are collecting ART from community pick-up points and may not be recorded in facility records, potentially affecting interruption data.

In South Africa, data reveals that people aware of their HIV status often retest and many starting ART have prior ART experience. The gap between diagnosed and treated HIV individuals remains a challenge, highlighting the cyclic nature of engagement in care. PEPFAR data shows comparable numbers of returns to care and "interruptions," suggesting a need for improved retention strategies. Additionally, interruptions are often short, with many returning within two weeks, but the duration varies. Understanding these patterns is crucial for enhancing ART programs and ensuring continuity of care for PLHIV. Grimsrud illustrated the cyclical HIV cascade (Figure 6.2) and highlighted the proportions of individuals at different stages of treatment, including those with and without previous treatment interruptions, providing insights into where interruptions are most common and where the highest mortality risks lie.

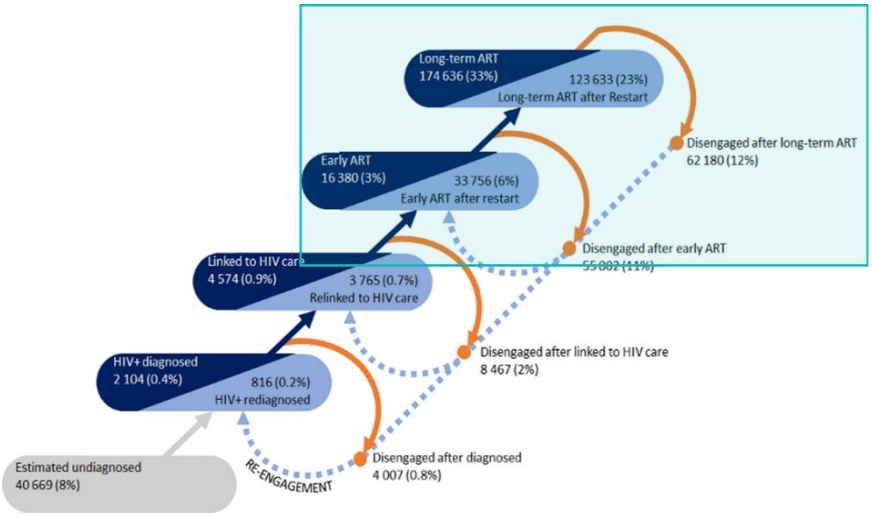



**Figure 6.2: Cyclical nature of HIV care engagement among patients in the Western Cape, contrasting those who have never experienced treatment interruption (dark blue) with those who have re-entered care (light blue) or are currently disengaged (orange).**

Grimsrud stressed the need to understand the profile of people returning to care, their experiences, and the duration of their interruptions, which requires a differentiated approach. Understanding the nuances of interruptions and the patient experience upon return is vital for improving long-term outcomes, including morbidity and mortality.

**AIM entries and indicator names: Program stats vs. Validation>Waterfall**

**John Stover** presented modifications to the Spectrum and Waterfall Analysis in light of feedback and observations from the 2023 estimates round. The waterfall plots were designed to compare and validate HIV/AIDS program data with Spectrum's estimates. However, the intended functionality and user-friendliness of the tools appeared to be lacking in practice.

The waterfall plots (one for program data and one for Spectrum estimates) show the number of people on ART at the beginning and end of the year, the number of individuals who had recently begun ART treatment, the number who passed away while on ART and those who were no longer actively engaged in care or had disengaged.

Stover noted the lack of evidence showing that users were entering program data into the Spectrum editor for the waterfall analysis. He proposed adding the waterfall plots to the ART editor and to simplify the process by merging the program data and Spectrum estimates charts, ensuring the tool's utility aligns better with user needs.

The discussion centred on the following key points:

1. Clarification on how Spectrum would handle data on LTFU and re-engagement was sought. Stover explained that while Spectrum can calculate those previously on treatment, it cannot differentiate between newly treated and previously treated individuals without making an assumption.
2. The suggested showing of both programme-recorded and model-estimated data for treatment initiation and interruption could aid discussions by highlighting discrepancies between recorded program activities and the modelled population dynamics.
3. The current process of entering program data into Spectrum was discussed, highlighting issues with data entry and the use of estimates. It was noted that the current display served more as a confirmation of data entry rather than as a detailed

analysis tool. A proposal was made to simplify the data entry process in Spectrum by consolidating program data inputs into a single chart. This chart would then be used to validate ART data and identify any inconsistencies between Spectrum estimates and program data.

4. The challenge of making assumptions for re-engagement from previously treated versus treatment-naïve populations was addressed, indicating the need for an assumption about their distribution.

**Recommendations:**

- Simplify the data entry process in Spectrum, chart to display key metrics such as the number of people on ART, newly initiating, re-engaging, disengaging, and deaths.
- Encourage entering program data directly into the ART editor to ensure accurate Spectrum estimates and validation.
- Add bars to the chart to display both program-recorded and Spectrum-estimated values for initiations, deaths, and treatment interruptions.
- Enhance Spectrum to incorporate real-time feedback on data entry and visualization errors, allowing for immediate corrections.

The discussion around treatment interruptions included:

<u>Treatment interruption definitions</u>

- Further discussions highlighted the importance of differentiating between "loss to follow-up" and "treatment interruption," with some participants suggesting that these terms are not interchangeable and should not be conflated. The distinction is critical because "loss to follow-up" could mean the patient has not shown up for appointments but might still be receiving ART elsewhere, whereas "treatment interruption" implies a cessation of treatment.
- The current WHO definition of loss to follow-up (28 days after the last missed visit) might not fully capture the clinical implications of treatment interruptions, particularly in the context of multi-month dispensing of ART.
- Shorter interruptions might not have a substantial impact on mortality, but longer interruptions (e.g., greater than 90 days) significantly increase mortality risk.The importance of considering viral rebound was emphasized, which can occur as soon as 14 days after stopping medication. This aspect is crucial for ongoing transmission dynamics and individual patient health.

<u>Default treatment interruption rates</u>

The discussion emphasised the importance of having a non-zero default rate for treatment interruptions. A tentative proposal was made to use the 1.6% rate for high-income countries as a starting point, based on estimates from ART-CC. For other countries, a 5% default treatment interruption rate was suggested, based on the median rate entered by countries who entered non-zero values in Spectrum. The discussion also touched on the variability of disengagement rates among countries, noting that in some cases, rates were significantly higher, suggesting a need for careful consideration in applying a default rate across different regions.

The main recommendations following these presentations are captured in **Appendix A.**

# Session 7: Concentrated epidemics

Chaired by **Cari van Schalkwyk**, the objectives of this session were to:

- Discuss planned and possible refinements to EPP-Concentrated (including EPP-ASM), CSAVR, and ECDC models.
- Advise the 2024 plan for key population data collection and estimates for (non-SSA) countries not using EPP-Concentrated or AEM

The distinction between "generalized" and "concentrated" epidemics remains a focal topic in the context of HIV/AIDS research and interventions, prompting Cari van Schalkwyk to present the various incidence models in use for 2023, as seen in Table 7.1.

**Table 7.1: Incidence models used in 2023.**

| | EPP Gene-ralized | EPP Concentr. | AEM | CSAVR | ECDC | Non-Spectrum 'Direct Incidence' | Total |
|---|---|---|---|---|---|---|---|
| Countries | 39 | 36 | 13 | 67 | 5 | 12 | **172** |
| PLHIV 2021 (M-2022) | 18.1M | 4.8M | 2.1M | 3.0M | 0.08M | 10.3M | **38.4M** |
| Regions | **ESA, WCA**, CAR, AP | **MENA, CAR, LA**, WCA | **AP** | **MENA**, CAR, LA, **WCENA, EECA** | WCENA, EECA | WCENA, **Brazil, South Africa** | |
| Shadow files | 1 | 1 | 0 | 13 | | 4 | |

- India & China use EPP-Concentrated by state/province; UNAIDS summarizes national totals as Direct Incidence.
- Shadow file = not approved by country; produced by UNAIDS with no or limited input from country; not published on https://aidsinfo.unaids.org/
- Besides these 172 with population >250,000 in 2022, CSAVR models completed for: Grenada, Saint Kitts & Nevis, Palestine.

The primary rationale behind retaining the terminologies "generalized" and "concentrated" in the Reference Group discourse is rooted in data available and the specific tools harnessed for epidemic estimation. The tools, a vital part of the Spectrum suite, are designed to cater to the specific characteristics and requirements of either a generalized or concentrated epidemic, depending on the prevalent conditions in the regions they serve. As an illustrative point from the slide, the EPP Generalized model is aptly employed in 39 countries, given the availability of HIV prevalence surveys in these regions.

**EPP-Concentrated: early trends and recommended curve types**

**Tim Brown's** presentation focused on the importance of model selection in estimating HIV epidemic trends, especially in concentrated epidemic settings with limited data. He emphasized the necessity of scrutinizing early trends in epidemics, proposing that the term "focused" be used instead of "concentrated" to describe such epidemics.

Brown's key objectives were to:

1. Evaluate the impact of different model choices within EPP-concentrated and the repercussions of sparse data.
2. Highlight issues raised by using the *classic* incidence model through examples from Bolivia and Jamaica.

3. Suggest revised guidance for incidence model choices in sparse data settings and the evaluation of early epidemic behaviour.

Brown stressed that all EPP incidence models primarily calculate the force of infection.

- *Classic* keeps the transmission rate fixed throughout, limiting its flexibility.
- *R-hybrid* employs a logistic function for incidence calculation, offering adaptability.

Current guidance for low data situations:

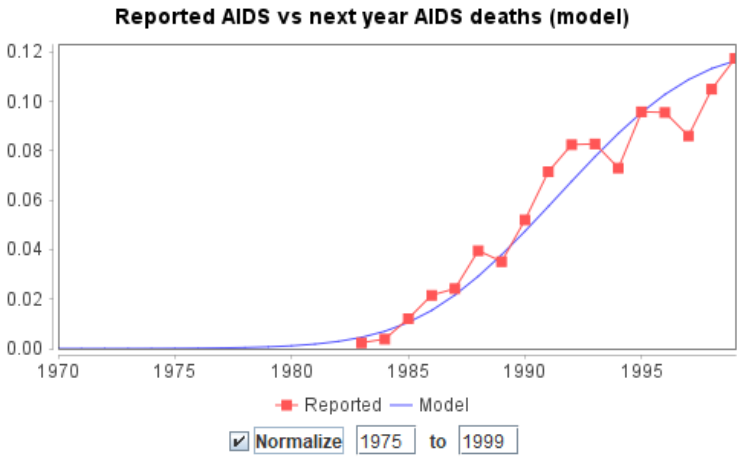
- For data spanning less than three years, the *classic* model is recommended.
- With at least three years of data, the *r-spline* model is recommended.
- In generalized epidemics, *r-hybrid* model is preferred.
- For countries with abundant data, mostly in SSA, *r-hybrid* is the popular choice.
- Brown highlighted that among the 36 countries discussed, many used *classic* despite the sparse data available, presenting challenges in accurate modelling.

Challenges with the *classic* incidence model:
**Bolivia**: When finalizing 2023 estimates, concerns arose about an early high incidence peak for men who have sex with men. While changing parameters could move it around, it remained strong, even at national level. Review of the data showed that the spike came from the rapid rise and plateauing of prevalence. Though some had concerns, these are common incidence trends in populations like men who have sex with men and people who inject drugs.

The *classic* model tends to produce flat curves if the data does not define a clear trend.

**Peru**: If the data more clearly defines a trend, *classic* can follow it. But having four fixed parameters limits its ability to adapt to subsequent changes in the epidemic's trajectory. Since many countries use ANC census data to estimate male HIV prevalence, any changes in the epidemic's trajectory indicated by later data could pose a challenge for *classic*.

**Jamaica**: Another rapid incidence spike among men who have sex with men was noted in the May 2023 meeting. Results from the *classic* model showed a quick rise and plateau, but a more gradual increasing trend seemed more consistent with the data. This issue was compounded by a temporal ordering problem, where the model inaccurately depicted the epidemic's initiation among remaining males and females prior to the men who have sex with men group, which Brown suggested was the actual initiator. Furthermore, the late rise and subsequent decline in the epidemic among female sex workers, as indicated by *classic*, did not align with the expected temporal dynamics, raising concerns about the model's capacity to accurately capture the progression and interrelations of the epidemic across different population groups.




Brown used the fitting page's Data Check feature to analyze early AIDS case trends based on WHO-reported data up to the year 2000, shedding light on the initial dynamics of the epidemic that other methods might miss. Data from Jamaica shows reported AIDS cases with deaths, which, in the pre-ART era, closely followed diagnoses due to the high mortality rate, as seen in Figure 7.1.

**Figure 7.1: EPP's Data Check feature, to validate HIV epidemic models with historical AIDS case data.**

Brown's experimentation with model conversion to *r-hybrid*, from *classic*, revealed significant improvements in the representation of the epidemic's dynamics, particularly for the men who have sex with men population in Jamaica. As *r-hybrid* model employs a logistic function for incidence calculation, it presented a more gradual and distributed rise in HIV prevalence that aligned well with early stages of AIDS deaths. This adjustment also corrected temporal ordering issues, more accurately depicting the epidemic's initiation and progression among different population groups.

The case of Trinidad and Tobago further illustrated the practical implications of model choice. Switching to *r-hybrid* from *classic* model demonstrated a better fit with reported AIDS case data, raising questions about the accuracy of early epidemic modelling and the reliability of data reporting practices.

Brown concluded his presentation with recommendations for potential adjustments to current modelling guidance. If any of these occur, revision should be considered:

- Data shows a trend to the eye, but *classic* fits flat through it.
    - This may mean that *classic* is too rigid for this data set.
- Strong, very sharply peaked incidence spikes
    - Explore which populations give rise to them and examine them more closely.
- An epidemic rise that occurs too early and/or trends substantially faster or slower than AIDS cases
    - Use the "Check data" button on EPP's "Fitting Results" page to compare against AIDS cases. For many "concentrated" epidemics, this is a pre-data period.
- Epidemiological implausible comparative trends across sub-populations
    - Clients coming up faster than female sex workers.
    - Remaining populations (male/female) rising before key populations.

Revision of models could entail:

- Test out other choices of EPP model (*r-hybrid, r-spline, r-trend*).
- In the model you're using consider if the start year is set too early or late.
- Consider using prevalence conditions, according to guidance in the Spectrum guide for updating HIV projections.
- They are to be used with caution.

- Review your input carefully.
    - If data don't define a trend, then fitting any model may give spurious results.
    - If so, it is critical to explore your epidemic against other sources of info: AIDS cases in early years, ART trends over time, case reporting trends, etc.
- Use the Spectrum validations to the max.

The *classic* incidence model is normally used for geographic areas with relatively few data points. However, IF

1) *Classic* gives an initial prevalence rise that is too sharp and implies an implausibly early and sharp incidence peak,

OR

2) the data themselves show a consistent time trend that could be cleanly fit with a single curve, but *classic* does not follow the data trend well,

THEN countries might consider another model such as *r-hybrid*, *r-spline* or *r-trend* and select that model that best fits their data. This should happen with guidance from UNAIDS, as it is possible the model may overfit the data if there are too few data points.

Brown suggested changes to the selection figure (Figure 7.2). He concluded by emphasizing the need for individualized model choices based on each specific scenario. He cautioned against blanket model changes and stressed the importance of country consultations.

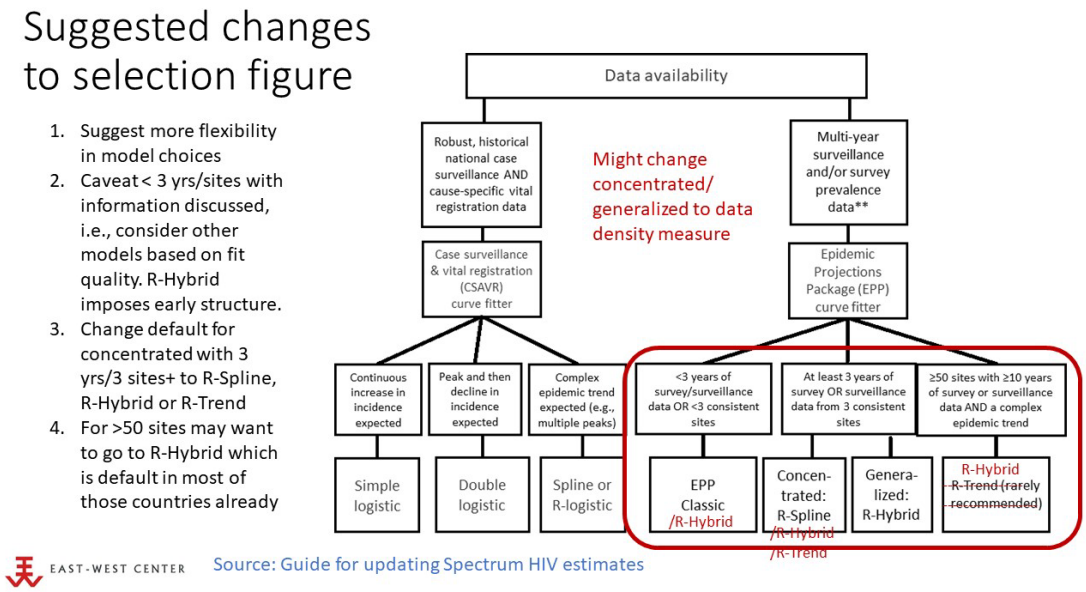



**Figure 7.2: Decision tree to choose the best model for deriving incidence in Spectrum, and suggested changes in red text. Source: Guide for updating Spectrum HIV estimates (hivtools.unaids.org).**

The discussion highlighted the relevance of the presented models not only for understanding early epidemic trends but also for addressing unexplained fluctuations in regional level trends. In cases of potential data or model representation confusion (e.g., the Trinidad case), it's crucial to review and verify the details in the provided resources, such as the Excel file with all the model fits, to ensure accurate understanding and interpretation. Overall, there was a clear agreement on the value of the proposed adjustments to model selection and guidance for improving epidemic estimations.

**Updated Age-Sex stratified EPP-Concentrated model**

**Deepa Jahagirdar's** presentation on the updated Age-Sex stratified EPP-Concentrated model (EPP-ASM) continued the dialogue on improved modelling for concentrated epidemics.

Key features of the EPP-ASM model:

- Developed mainly for concentrated epidemics.
- It integrates demographic projections from Spectrum and the transmission dynamics from EPP.
- Offers age-sex specific HIV and evolution over time.
- No interactions between key populations in the model.
- Different models for transmission rate can be implemented, as seen in Table 7.2. below.

**Table 7.2: Differences between EPP-concentrated and EPP-ASM.**

| | **EPP-concentrated** | **EPP-ASM** |
|---|---|---|
| **Data** | Prevalence from surveillance and household surveys | Prevalence from surveillance and household surveys (can be age-sex specific) |
| **Parameters** | Incidence model | Incidence model and incidence rate ratios for age and sex |
| **Incidence model r (t)** | *R-spline, classic, r-trend, r-hybrid* - function of prevalence, ART, and transmission rate | Same options |
| **Key populations** | Concentrated: Model is **run separately for each key population** (no dynamic transmission) | Concentrated: Model is **run separately for each key population** (no dynamic transmission) |

The data fitting process is shared between EPP and EPP-ASM, which both use the IMIS procedure to derive the posterior distribution. This fitting process incorporates multiple data sources such as household surveys and HIV Sentinel surveillance or ANC sites or routine testing data.

Since May 2022, where results for Senegal were presented, Jahagirdar has continued developing the model, particularly for countries and subpopulations lacking survey data, where model fitting requires more discretion and comparative analysis to EPP. The recent country AIM files generally used the *r-spline* model, while testing often employed the *r-hybrid* model. Efforts to reconcile these approaches have been made where possible.

***Key findings***

*- EPP-ASM demonstrates promising fits where data is ample, occasionally outperforming EPP.*

*- In limited data scenarios, model choice impacts projections significantly, and there is no consensus on treating EPP as the gold standard.*

Country case studies and observations:

- **Dominican Republic**: Showed lower HIV prevalence among men who have sex with men in EPP-ASM compared to EPP. The absence of survey data led to significant challenges, with EPP-ASM predictions sometimes collapsing towards zero, especially in populations like transgender people.
- **Niger**: In the 'remaining populations', with ample data to fit to, good fits were obtained against both surveillance and survey data with EPP. The average age of PLHIV has increased in more recent years. The differences in model fits between male and female populations might be explained by turnover rates or variations in prevalence/incidence coming from different populations, given the identical data sets used for both genders.
- **Ecuador**: Initial model predictions showing a flatline trend among men who have sex with men, indicating no significant changes in prevalence over time. Adjusted priors in EPP-ASM led to better agreement with EPP despite limited data.
- **Honduras**: Difficulties in accurately estimating prevalence for female sex workers, failing to reconcile with higher ANC data points. Despite applying a tighter prior and using refined spline techniques, persistent discrepancies were observed between EPP and EPP-ASM, indicating the need for potential model revisions.
- **Peru**: EPP-ASM estimated a near-zero prevalence for trans populations and men who have sex with men due to the absence of survey data.

Potential fixes and recommendations:

1. Tighten the prior on surveillance data when survey data is not available to better align EPP-ASM with EPP results.
2. Consider revising the model for specific countries, especially if the fit does not align with available data.
3. Explore the possibility of integrating or transitioning towards the Goals-ARM model, particularly in the context of generalised epidemics.
4. Continue to refine the EPP-ASM model, ensuring its suitability and effectiveness for concentrated epidemics as an alternative to EPP.

The conversation after the presentation focused on the differing data requirements between models (EPP-ASM vs. Goals-ARM) and the implications for countries' readiness to adopt new models were highlighted as critical factors. Although Goals-ARM might be ready for use by the end of 2024, its data requirements are significant, there could be value in having a backup, such as an improved EPP model, that does not have as extensive data requirements.

## CSAVR

The CSAVR session addressed several challenges and areas for potential adjustment, including:

- HIV incidence functions in CSAVR: 1) splines may overfit fluctuations in recent diagnoses and 2) double logistic produces strong swings in incidence early in the epidemic (Guy Mahiane)
- Instances of ART coverage >100% of PLHIV (Guy Mahiane)
- Aggregating uncertainty across CSAVR models (Rob Glaubius)
- Fitting AIDS death data using AIM's excess mortality rates: possible adjustments (Rob Glaubius)

HIV incidence functions in CSAVR

**Mahiane** narrated the developmental journey of incidence options available in the CSAVR tool however, constraining models could, in some contexts, degrade their performance rather than enhance it. This led to the exploration of alternative incidence options, including splines with

three or four knots, to avoid impeding the models' effectiveness across different settings. However, the challenges of time and data availability persist, complicating the precision and reliability of these models.

Mahiane's analysis, covering a sample of countries, demonstrated that while the various modelling options could closely fit the number of new HIV diagnoses within the data period, projections for years outside the available data could significantly diverge, as seen in Kuwait and Tunisia, indicating that fewer knots do not necessarily prevent overshooting in projections. The incidence trends in Argentina and Saudi Arabia suggested that splines with fewer knots might align with the best model, as indicated by the Akaike Information Criterion (AIC), yet they could also deviate and project unreasonable trends, as observed in Spain, Bosnia and Herzegovina. Notably, none of the spline options managed to eliminate the pronounced bump around 2020 observed in Spain, highlighting the limitations of the current modelling options in capturing sudden changes or anomalies in incidence trends.

Issues highlighted:

- **Double logistic**

The double logistic option typically generates moderately rising or falling incidence rates without sharp fluctuations. It is less sensitive to fluctuating ART coverage compared to the spline or R-logistic options. Despite its general stability, this option sometimes predicts sharp historical peaks in incidence, pre-2000 and pre-ART, as observed in Venezuela, Saudi Arabia, and Chile. Such anomalies are challenging to mitigate without imposing additional constraints on the model parameters, which could affect the model's performance across different settings.

- **Spline**

The spline option produces good fits to the number of new HIV diagnoses, but incidence shows fluctuations that do not align with societal or programmatic changes.
In Spain, the spline option suggested an unrealistic decrease in incidence by more than 90% between 2014 and 2022, leading UNAIDS to withhold publication of these estimates. Similar steep incidence declines were noted in Cuba, Croatia, Japan, Belgium, and France. Conversely, unexpected rises in incidence following increased diagnoses were observed in Saudi Arabia and Lebanon. This prompted the question whether constrain the latest spline should be further constrained to avoid spurious claims after 2010.

<u>Instances of ART coverage exceeding PLHIV estimates</u>
Certain cases, notably in Sweden, displayed ART numbers surpassing the estimated number of PLHIV from 2010 to 2020. This anomaly was observed across all four incidence options, but mainly with the single logistic curve. Implementation of a constraint to the likelihood showed that regardless of the incidence option used, it was possible to achieve estimates where the number of PLHIV was at least equal to or greater than ART numbers for all countries, including Sweden. The relative difference in the number of PLHIV under the constrained model versus the unconstrained model could exceed 30%, particularly for the R-logistic option in Sweden.

Recommendations
Mahiane suggested making it optional for CSAVR users to fix splines with fewer than five knots and to constrain the number of PLHIV by ART numbers. This recommendation encourages countries to review their ART data critically and if confident in their data, countries can opt to apply the constraint to ensure the number of PLHIV is estimated to be greater than ART coverage.

The discussion centred around the constraints on the number of PLHIV relative to ART numbers, the configuration of spline knots, and preferences for the double logistic model.

Key inquiries & clarifications:

1. Spline knots configuration: An alternative approach to handling spline knots were suggested: keeping five knots but forcing them to be positioned before 2019. This suggestion aimed to address the anomalies in case diagnosis trends observed during the COVID-19 pandemic starting in 2020. Mahiane noted that while this approach could be considered, the focus has been on estimating the position of knots without setting a knot after five years from the present to avoid overly constraining the model.
2. Adjustments to modelling approaches: Mahiane clarified that the addition of three and four knot spline options was evaluated alongside the best incidence curve chosen using the AIC for each country. This was illustrated with examples like Croatia, where comparisons among different spline options were made against the best model.

Aggregating uncertainty across CSAVR models (**Rob Glaubius)**

Different incidence curves, although fit to the same data, may yield varied estimates. Currently, countries select one of the four models based on AIC and expert review. The uncertainty depicted in estimates reflects only the selected model, omitting structural uncertainties present across different models. This omission can lead to an underestimation of the true uncertainty in HIV prevalence and incidence estimates.

Approximate method for model averaging:

To explore the utility of model averaging, an approximate method was employed. This involved sampling 1000 posterior incidence curves from each CSAVR model for 20 randomly selected countries. These were then resampled proportionally to their likelihood, comparing uncertainty bounds from these resamples against those derived from the best-fitting model alone. This process aimed to capture a broader spectrum of uncertainty by incorporating insights from multiple models.

1. Case study: In Barbados, the double logistic model dominated the resampling process due to its superior performance, suggesting that uncertainty bounds would closely mirror those of the double logistic model. Conversely,
2. Case study: In Iceland, a more equitable distribution of resampled curves across three different models indicated a diverse representation of uncertainty.
3. In Netherlands, most resampled curves leaned towards the spline model, with a limited representation from the double logistic after resampling. The combined uncertainties indicated a shift from a stable epidemic to one that's decreasing rapidly.

Recommendations and reflections

Glaubius acknowledged the limitations of the exploratory method used in his analysis, and although provisional, this analysis could serve as a valuable pilot. Glaubius questioned whether prioritizing the development of a formal model averaging approach to better aggregate uncertainty across CSAVR models should be a focus area. He acknowledged the potential for wider uncertainty bounds to offer a more realistic representation of future uncertainties. However, he also noted the predominant emphasis on point estimates in decision-making processes, questioning the practical benefits of expanding uncertainty bounds for policy and planning purposes.

The discussion underscored that while model averaging presents a promising avenue for reflecting structural uncertainties more comprehensively, practical challenges and limitations necessitate further exploration and development. Additionally, the case of Spain illustrates the

need for CSAVR models to adapt to changing data landscapes, such as those influenced by global events like the COVID-19 pandemic.

Key inquiries are listed below:

1. The low diversity in resamples raised concern, and it was suggested to increase the number of samples significantly to address this issue. Glaubius acknowledged this concern, noting that attempts to increase sample diversity through resampling both with and without replacement resulted in a reduced diversity of unique samples, indicating a potential limitation of the approach.
2. A suggestion was made to adopt model averaging of the confidence bounds of various incidence options, representing uncertainty across different models. While countries may prioritize point estimates, broader uncertainty bounds could offer organizations like UNAIDS more objective leeway in interpreting data. For example, in Spain, varying models led to different conclusions about the country's HIV epidemic decline, which emphasises the need for a methodological basis to support decisions rather than relying on subjective judgment.
3. It was suggested to incorporate testing data and test positivity rates as additional parameters to improve model accuracy, reiterating discussions from previous meetings about enhancing CSAVR's capability to address real-world data features

<u>Fitting AIDS death data using AIM's excess mortality rates: possible adjustments</u>
Glaubius then presented an analysis concerning mortality rates among PLHIV on ART, highlighting that Spectrum's "AIDS deaths" or "HIV-related deaths" represent excess deaths to PLHIV, and that CSAVR calibrates these *excess deaths* to reported/IHME estimated *AIDS deaths*.

Glaubius highlighted that the variability in mortality rates among PLHIV on ART necessitates accurate modeling to reflect real-world scenarios. The analysis incorporated insights from **Trickey's** analysis within the ART-CC, which detailed mortality rates broken down by AIDS and non-AIDS causes compared to age and sex-matched general population mortality. The analysis revealed that while AIDS mortality constituted only a fraction of all mortality among PLHIV on treatment, non-AIDS mortality among this group exceeded that of the general population, as shown in Figure 7.3.

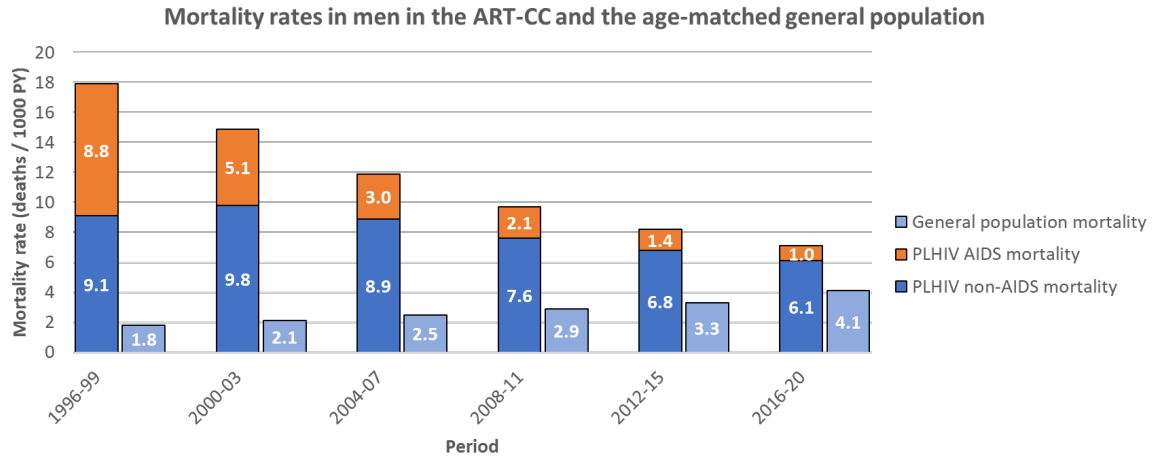



**Figure 7.3 From Trickey *et al* (6), Estimation of improvements in mortality in Spectrum among adults with HIV receiving antiretroviral therapy in high-income countries, JAIDS 2024.**

Defining mortality terms

- All-Cause Mortality: The sum of AIDS and non-AIDS mortality.
- AIDS Mortality: Represented as the orange segment of the all-cause mortality bar in Figure 7.3.
- Excess Mortality: The difference in all-cause mortality between PLHIV and the general population.

CSAVR's challenge with mortality data

A significant challenge for CSAVR is aligning excess death estimates with AIDS death data. The tool primarily adjusts its incidence estimates to match these figures, potentially leading to underestimations of the PLHIV population to align excess deaths with AIDS deaths. Glaubius underscored the variability of AIDS' contribution to all-cause mortality by age and sex, indicating the importance of considering these factors in mortality analyses.

Synthetic data study

To explore biases from calibrating excess death estimates to AIDS death data, Glaubius conducted a simulation study using synthetic data, given the limitations of current software and data challenges. This study aimed to assess the potential impact of different calibration approaches on CSAVR estimates, focusing on all-cause and excess deaths among PLHIV derived from Spectrum files for 2016-2022.

Findings from the simulation study

The study examined four countries (Algeria, Cuba, France, and Italy), generating synthetic vital registration data to reflect various scenarios of AIDS and excess deaths. The scenarios considered differences in AIDS mortality on and off treatment, using ART-CC estimates and assumptions about the proportion of AIDS deaths among off-treatment excess deaths.

Implications of the analysis

1. Impact on PLHIV population estimates:

- Calibrating CSAVR models to excess deaths, as opposed to solely AIDS deaths, could significantly affect estimates of PLHIV populations.
- For instance, Algeria and Cuba showed that fitting models to excess deaths could substantially increase the estimated PLHIV numbers compared to fitting solely to AIDS deaths.

2. France's estimates remained relatively insensitive to different death data inputs, which Glaubius attributed to three potential factors.
    a. Extensive case diagnosis time series: France's long-standing case surveillance data, which extends back to before the official recognition of AIDS, may have contributed to the model's insensitivity to changes in mortality data inputs.
    b. Inclusion of CD4 data: The use of CD4 data in France's modelling could have limited the sensitivity of CSAVR's fit to the vital registration data, thereby affecting the impact of different mortality data inputs on prevalence estimates.
    c. Bespoke model for incidence estimation: The synthetic data for France, generated by a model not based on CSAVR, could explain the different outcomes observed, highlighting the impact of model choice on fitting synthetic deaths data.

Glaubius set out four potential options for addressing the challenges identified in calibrating CSAVR models to mortality data:

1. Status quo: Continue fitting excess deaths estimates to AIDS death data, even though this may underestimate HIV prevalence.

2. **Modify CSAVR outputs: Make estimates of AIDS deaths that align with the current CSAVR input, potentially requiring mortality breakdown of AIDS vs. non-AIDS deaths in Spectrum.**
3. Input excess deaths: Though theoretically ideal, this option is unfeasible as excess deaths are not typically measurable but modelled.
4. Modify both CSAVR outputs and inputs: Fit all-cause deaths estimates to all-cause deaths data, sidestepping the need for breakdowns. The feasibility of this option remains uncertain.

Discussion key points:

1. The potential of analysing total deaths among PLHIV was underscored, especially if data could be disaggregated by age and sex. Several countries have such detailed data, which has previously been entered into CSAVR as AIDS deaths before being reclassified upon review.
2. Assumptions made in Spectrum regarding non-AIDS mortality among people living with HIV; non-AIDS deaths among people living with HIV are assumed to occur at the same rate as the age and sex-matched general population living without HIV. CSAVR does not currently have a breakdown to differentiate excess deaths that are AIDS-attributable versus non-AIDS-attributable.
3. The discussion highlighted the utility of all-cause mortality data among diagnosed PLHIV and those on ART for deriving parameters to distinguish between AIDS and non-AIDS deaths.

Recommendation: The group recommended pursuing Option 2, pending further insights from Trickey's review and discussions on quantifying the split between AIDS and non-AIDS excess mortality.

**ECDC model/platform: 2024 update**

**Ard van Sighem** presented on the ECDC modelling platform, which is an enhancement of the ECDC HIV modelling tool and is now accessible on the ECDC website. The platform combines functionalities from the ECDC modelling tool, and the HIV estimates accuracy tool.

1. **Modelling platform key features:**

- Available as an online/desktop application and R package.
- Uses case-based data, eliminating the need for aggregated datasets.
- Adjustment for missing data and reporting delay.
- Calculates the probability of HIV infection pre or post migration.
- Allows custom definitions for population modelling.
- Estimation of HIV incidence and undiagnosed population (e.g., MSM, transgender men).

2. **Data handling:**
    a. **Case-based data preference:** The platform now prefers case-based data over aggregated datasets, allowing for better adjustment for missing values and reporting delay.
    b. **Migration module:** Integrates parameters for pre- and post-migration infection probability, enhancing accuracy for mobile populations.
3. **Diagnosis matrix:**
    a. **Definition:** Specifies the probability of diagnosis over different time intervals and by CD4 count categories. Probabilities are estimated separately for various CD4 count categories.

    b. **Adjustment:** Allows for linear or constant probability adjustments, tailored to historical data trends.
4. **Reporting delay adjustment:**
    a. **Estimation:** Based on historical data, the platform adjusts recent data to account for reporting delays, ensuring more accurate and timely estimates.
    b. **Application**: The delay adjustment involves "topping up" the number of diagnoses in the most recent years to account for cases that are diagnosed but not yet reported. This method relies on the assumption that reporting delays are consistent over time. If a country undergoes a major reporting overhaul (e.g., a data cleaning event), this might complicate the delay adjustment process.

**Technical details:**

1. **Data imputation:**
    - **Multiple Imputation:** Predicts missing data values by analyzing distributions in complete cases and applying these predictions to incomplete cases. This method ensures consistency across multiple imputations.
2. **Population stratification:**
    - **Custom populations:** Users can define populations by gender, transmission route, region of origin, and place of residence, allowing for tailored modelling.
3. **Pre- and post-migration infection estimates:**
    - **Scenarios:** Different scenarios account for whether an individual was infected before or after migration, with probabilities adjusted based on CD4 count, age, and other factors.

Post-migration HIV acquisition: Initial findings indicate that around 33% of diagnosed individuals acquired HIV post-migration, with this proportion remaining relatively stable over time. Notably, individuals arriving post-1980 show a decreasing likelihood of post-migration HIV acquisition, reflecting the timeline of the HIV epidemic's emergence.

1. New enhancements permit excluding migrants with a high pre-migration infection likelihood and specifying multiple diagnosis matrices to accommodate data scarcity in certain populations.
2. Modelling results on migrant populations: Analyses excluding migrants with a significant pre-migration infection probability highlighted the distinction between undiagnosed individuals within the reporting country and undiagnosed migrants, offering insights into the epidemiology of HIV among migrant populations.

Open ends:

1. The tool currently lacks estimates for migrants with pre-migration HIV infections who remain undiagnosed within the reporting country. Addressing this requires assumptions about immigration rates and HIV prevalence among immigrants.
2. The model does not account for out-migration, which could be conceptually treated as "artificial deaths" to reflect its impact on the PLHIV count within the reporting country.
3. While the tool facilitates separate estimates for different populations, aggregating these to derive total undiagnosed or newly acquired HIV infection counts requires manual calculations. Discrepancies in aggregate estimates necessitate revisiting modelling assumptions.
4. The assumption of no diagnosis during early HIV infection may not be accurate anymore. Additionally, the model faces challenges in estimating for rapidly declining HIV epidemics, indicating areas for further development.

Key discussion points:

1. Concerns were raised about managing multiple key population outputs and their integration into Spectrum, specifically, how to handle different output files for each key population and ensure these are correctly designated and integrated. It was explained that when using the new ECDC HIV modelling platform, users need to upload one comprehensive data set from which various key populations can be defined. Each key population is modelled separately, resulting in multiple output files.
2. Concerns about integration were addressed by noting that the name of the output files, which currently are generic (e.g., "HIV model fix" followed by date and time), should be renamed by users to reflect the specific key population before importing them into Spectrum. This step ensures clarity and proper designation. It was suggested that a more intuitive file-naming system could be implemented in future versions of the platform to automatically include key population identifiers, facilitating easier integration.
3. Inquiries about the diagnosis matrix and reporting delay adjustments: Clarification was sought on how the diagnosis probabilities were defined and adjusted over time, and how the reporting delay was estimated and applied in the model.
4. Modelling probability and CD4 count categories: van Sighem clarified that the diagnosis matrix specifies the probability of diagnosis over different time intervals and by CD4 count categories. For example, a constant probability might be assumed between 1984 and 2000, followed by a piecewise linear change.
5. The probabilities are estimated separately for different CD4 count categories (e.g., <200, 200-349, 350-499, ≥500 cells/mm³), which allows the model to account for varying likelihoods of diagnosis based on disease progression. Reporting delay estimation: The methodology for estimating reporting delays was discussed, with van Sighem explaining that historical data on the lag between diagnosis and reporting inform adjustments to more recent data. This process assumes a relatively constant reporting delay, excluding anomalies like data cleansing events, which could skew delay estimates.
6. Differences in reporting practices: A discrepancy between CSAVR and ECDC regarding the treatment of diagnosis dates was highlighted. While CSAVR uses the reporting date as the diagnosis date, ECDC uses the original diagnosis date, offering a more precise timeline of infection and diagnosis.
7. Reporting delay variability: It was noted that reporting delay varies by country, affecting the model's applicability and accuracy. Some countries experience minimal delays, while others have significant lags, impacting the reliability of temporal analyses.
8. Potential for global application: It was suggested that this model should be presented to other global regions, particularly in Africa, to illustrate the benefits of using individual-level case reporting data for HIV surveillance and modelling.
9. Challenges with migration data: Concerns were raised about the treatment of migration data in the model, particularly for countries with significant migration like Greece. The discussion touched on the discrepancies between PLHIV estimates generated by different models due to varying treatment of migration data.

<u>Collating ANC testing and prevalence data in AEM</u>

**Tim Brown's** presentation was shared on DropBox for the Reference Group to peruse. No recommendations were made.

Key population data and estimates (PSE, prevalence, ART coverage) in non-EPP countries (**Sonia Arias-Garcia** and **Keith Sabin**)

Key population data is collected only in countries using the EPP model. Other countries using models like CSAVR, ECDC, or direct incidence models do not collect key population data. This presentation delved into using Excel files for collecting historical data on key populations during the 2023 workshops in the Caribbean (CAR), Latin America (LA), and Middle East and North Africa (MENA) regions. This method aimed to refine estimates and models by incorporating previously unused key population data, potentially enhancing country-specific EPP fits or upgrading to EPP-concentrated models.

The primary goal was to use untapped key population data to improve national estimates and models, thereby informing country goals and optimal calibration for infection distribution estimates and program targets.

Data collection: The file was a simplified version of the SSA Key Population Workbook and can be imported to Spectrum for graphical representation of the data. Among 43 countries participating in the workshops across the mentioned regions, 8 (19%) filled the Excel file, contributing 161 new key population data points, notably from Mexico and Djibouti, and there was a lack of data on ART coverage among key populations.

Arias Garcia highlighted a comparative analysis between workshop-collected data and GAM reported data. This comparison revealed discrepancies in 11% of the Excel file data with official submissions, categorizing data points based on their inclusion and matching across both sources.

While AIDSInfo criteria for data dissemination limit the inclusion of workshop-collected data, the Key Population Atlas offers a broader scope for integrating this data due to its encompassing historical and literature-sourced data.
There are challenges to consider, such as reluctance from countries to provide more data, the intricacy of the data, and potential modelling inconsistencies.

Future plans: A 2024 plan is set in place for non-GAM data collection with an aim to update platforms like the Key Population Atlas, GAM, or AIDS Info with data reported in the 2023-24 estimation context.

Arias Garcia concluded with open-ended questions about the continuation of Excel-based data collection, its integration into modelling efforts, and alternative strategies for promoting national key population databases in non-EPP modellable formats. She highlighted the need for further discussion and insights to address these questions effectively. Sabin emphasized the necessity to consider upcoming changes and requirements for the Symphony model. He defined the data elements currently used and desired for EPP models, suggesting a cautious approach towards requesting new data types to ensure relevance and practicality in modelling efforts. Sabin proposed exploring systematic reviews and modelling strategies to address data gaps and improve future models.

Recommendation was sought on the following questions: Should Excel-based data collection continue in next estimation rounds?

a) If yes, with the purpose of:
    - Adding the data to the Key Population Atlas
    - Basis for starting or expanding EPP-Concentrated models
    - Informing Goals and Optima country calibrations

b) If not,
    - Could the SSA Key Population Workbook be extended to CAR, LA and MENA?
    - Will UNAIDS refrain from promoting national key population databases in non-EPP countries?

The discussion centred around the following key points:

1. Integration of CSAVR KP and EPP ASM:
    a. Merge CSAVR KP and EPP ASM work streams to utilize strengths from both models, providing more comprehensive estimates.
    b. Focus on using case surveillance, vital registration, and Sentinel surveillance data in conjunction for better data integration and model utilization.
2. Data collection strategy**:**
    a. Provide countries with spreadsheets pre-populated with existing data, allowing them to review and update information more easily.
    b. Develop a comprehensive strategy for data collection that includes systematic reviews and well-documented sources.
3. Ensure that no data is published without proper documentation and verification, emphasizing transparency and accuracy.
4. Provide training for country teams on the new integrated system to ensure accurate data entry and analysis.
5. Model assumptions and data quality: It was proposed to extract and share model assumptions related to key population sizes, prevalence, and number of partners to allow countries to compare and potentially update these with better data. This approach aims to bridge the gap between the model's assumptions and the actual data countries might possess.

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

# Appendix A – Recommendations

| Recommendation | Lead person(s) | Timeline |
|---|---|---|
| **Session 1: Key Populations data synthesis in sub-Saharan Africa**<br>*Chair: Jeff Imai-Eaton*<br>**Objective:**<br>• **Recommendations on use of the Triangulator in key population estimates process in SSA**<br>• **Review data on PSE trends and urban/rural population size estimate ratios** | | |
| **Aggregator and Triangulator**<br>• Make Triangulator and Aggregator tools available to HIV estimates teams in 2024 UNAIDS estimates as an option for synthesising and pooling population size estimate (PSE) data for consensus Key Population Workbooks.<br>  ○ Deploy as a 'soft launch'. Encourage country team users to try the tools and provide feedback.<br><br>• Development requirements for tool implementation soft launch in 2024 UNAIDS:<br>  ○ [Statistical method] Develop approach to incorporate PSE observations that do not have standard error or sample size in Triangulator estimates. This requirement currently excludes many PSE observations.<br>  ○ [Statistical method] Specify global prior distributions or spatial correlation to enhance estimates for extremely data sparse populations or regions.<br>  ○ Enable smooth convenient data flow between tools: Key Population Workbook data tables → Triangulator input → Aggregator input → outputs to Key Population Workbook consensus sheets.<br>  ○ Ensure fields in Key Population Workbook to record all data elements required by Triangulator, including confidence score, standard error, sample size<br>  ○ Prepopulate Key Population Workbook with default quality scores.<br>  ○ Provide users guidance on how to determine and adjust confidence scores, with examples of impact of changing confidence scores. | Carl Corcoran, Ian Fellows, Oli Stevens, Ray Shiraishi, UNAIDS | 2024 estimates |
| For key population indicators other than PSEs, pre-populate the Key Population Workbook with results of existing data synthesis analyses, with an option for users to update:<br>• For HIV prevalence and ART coverage among key populations, prepopulate workbooks with results of Stevens *et al.* data synthesis.<br>• For new infections among key populations, pre-populate with results from other mathematical models [consistent with previous workbook iterations] | Oli Stevens | 2024 estimates |
| Several recommendations were made for consideration as priorities in future methodological development:<br>• Consider the merits of two-step modelling process for process data pooling (Triangulator) and extrapolation (Aggregator) versus one model that does both steps (similar to Laga *et al.* and Stevens *et al.* approaches).<br>• Understand sensitivity to confidence scores:<br>  ○ Consider a time decay on confidence scores such that more recent studies receive more weight. This may address the limitation of the current GAM approach of dropping studies before a certain window, resulting in abrupt changes to estimates based on different study inclusion.<br>  ○ Develop systematic approach to assigning confidence scores to be reproducible and consistent across countries. | Working group | May 2024 |
| **Time trends in PSE proportions (SSA)** | | |

| Recommendation | Lead person(s) | Timeline |
|---|---|---|
| • Analysis of collated key population size estimate (KPSE) study data for the Key Population Workbooks did not find evidence of a systematic trend in KPSE proportions over time.<br>  ○ Large heterogeneity in the study proportion estimates results in variations that are larger than the trend to be detected, which makes interpretation difficult.<br>  ○ The frequency of KPSE studies is not sufficient to observe changes that arise in response to specific events.<br><br>• Data on total PSE counts reported by GAM in SSA showed a rise in population size counts from 2010 to 2022.<br>  ○ This observed rise is likely due to the adoption of more systematic methodologies to obtain nationally representative counts in recent years and not necessarily an actual increase in the PSE percentage.<br><br>• Considered together, the **recommendation is to assume that key population proportions have not changed over time.**<br><br>• Regarding the men who have sex with men PSE minimum threshold of 1% of men 15-49 years, the recommendation from the UNAIDS/WHO policy brief in 2020 stipulated this threshold based on data determined 'nationally adequate', which included two countries in the Eastern and Southern Africa region. Collated key population study data in sub-Saharan Africa consistently find lower proportions of men who are men who have sex with men based on risk-based identification. The minimum PSE threshold remains important for policy and programming to ensure that men who have sex with men are not omitted from HIV programming in settings that lack data and where men who have sex with men are stigmatized and discriminated against.<br><br>Recognizing the importance of this threshold and the importance of not establishing programmatic targets that are unattainably high relative to risk-based definitions, it was recommended:<br>  ○ Retain the 1% benchmark as a minimum estimate for the proportion of men who are men who have sex with men in settings that do not have locally available data or epidemiologic synthesis.<br>  ○ In GAM reporting, do not exclude or reject reported men who have sex with men proportions below 1% where they are based on an empirical estimation process with appropriate extrapolation to national population.<br>  ○ UNAIDS/WHO repeat the analysis underpinning the minimum threshold analysis incorporating a large amount of key population size estimation data using high-quality methods available since the initial policy brief. | | |
| **Urban/rural PSE ratios**<br><br>Current inputs to the Key Population Workbook and Aggregator tools assume a default rural-to-urban ratio for PSEs proportion of 0.6 for all key populations, with wide uncertainty. This assumption is required because key population study data are almost exclusively conducted in urban areas. This assumption is based on expert consensus that female sex workers, men who have sex with men, and people who inject drugs are disproportionately concentrated in urban areas, which has also guided focus for key population surveys and program delivery. Most data reviewed implied modest or no systematic difference between in key population proportions between urban and rural populations:<br>• Analysis of men reporting male sex partners in PHIA household surveys implied a rural-to-urban ratio around 1.1; however, there was uncertainty about accuracy of reporting and interpretation of household survey data not developed to measure stigmatised behaviours.<br>• Analysis of key population survey PSE estimates identified a negative correlation between population density and key population proportions, which would imply smaller PSE proportions in more dense urban areas. However, this analysis could | | |

| Recommendation | Lead person(s) | Timeline |
|---|---|---|
| be very sensitive to misspecification of population denominators for PSE proportion estimates.<br>• Network scale-up population size estimates in the Cote d'Ivoire and Nigeria PHIA surveys implied higher PSE proportions in urban areas in Cote d'Ivoire for all key populations, but not in Nigeria. Interpretation was further limited by uncertainty in population projections of the urban/rural proportion.<br>• A meta-analysis of data from household surveys found that men in rural areas were 25% less likely to report paying for sex than men in urban areas (Hodgins *et al*. PLOS Med 2022).<br><br>While evidence was limited for a large systematic rural-to-urban difference in key population proportions, there was also perceived to be severe limitations to the interpretation of each data source. On this basis, it was recommended to:<br>• *Preliminary recommendation:* Retain the current default assumption of 0.6 rural-to-urban ratio, noting the extremely weak quality of evidence underpinning this assumption.<br>• Consult key population program experts and key population community organizations for guidance regarding assumptions for key population rural-to-urban ratios. | Sharmistha Mishra | November 2023 |
| **Key Population Workbook**<br><br>Add sheet to Workbook, for key population program data. Two indicators to include here:<br>• Number of key population members reached by HIV testing, treatment, and prevention programs.<br>• Number on treatment<br><br>Data to be recorded at district-level and by calendar quarter. | Imperial College London | 2024 estimates |

| Recommendation | Lead person(s) | Timeline |
|---|---|---|
| **Session 2: Dynamical modelling of HIV trends in key populations in sub-Saharan Africa**<br>*Chair: Cari van Schalkwyk*<br>**Objective:**<br>• Review assumptions about population group duration at risk and turnover, including variation by region within sub-Saharan Africa and changes over time<br>• Review assumptions and model structure decisions | | |
| Mean population duration assumptions for key population group turnover in Goals and EPP models have been derived as doubling the median duration at risk measured in key population surveys. Doubling the median reported duration is intended to account for that reported duration reflects only duration before date of survey, not the future duration at risk. However, this approach does not account for that shorter duration episodes are less likely to be captured in the sample. Under certain assumptions about the distribution of duration at risk, the reported mean duration at risk is equal to the mean true distribution (Knight et al. medRxiv).<br><br>Therefore, it was recommended that:<br>• Assumptions for mean duration at risk for female sex work and people who inject drug populations should not be derived as doubling the median reported duration in studies. Further analysis should be undertaken to determine whether the observed mean duration is appropriate or other adjustments are required. | Working group | May 2024 |
| Duration of risk for men who have sex with men:<br>• Rebecca Anderson showed a median estimate of 5.4 years at risk duration from surveys. However, men sampled in these surveys are typically younger than overall men who have sex with men populations. | Working group | May 2024 |

| Recommendation | Lead person(s) | Timeline |
|---|---|---|
| • **Consider potential for cyclical entry/exit patterns in men who have sex with men demographics**. | | |
| Changes in female sex worker duration over time:<br>• Evidence from South Africa indicated systematic increases over time in average duration and age of sex workers in South Africa:<br>  ○ Review whether reduced mortality among female sex workers following ART availability may have increased the average duration and age of sex workers in South Africa.<br>  ▪ *Update since meeting: Anderegg analysis of South African female sex workers data showed that AIDS mortality rates is too low to explain observed increases in duration over 1996 to 2019.*<br>  ○ Review whether changes in the underlying population age distribution due to demographic transition may explain changing sex worker age distribution.<br>• Review data on time trends in sex work duration from key population surveys from other countries | Nina Anderegg, Leigh Johnson<br><br>Oli Stevens, Rebecca Anderson | May 2024 |
| Goals-ARM development<br>• To address challenge creating realistic HIV epidemic growth rates during the early epidemic, consider extending model structure to allow for:<br>  ○ Heterogeneity in sexual risk behaviours in non-key population groups<br>  ○ Other STIs as cofactor for HIV transmission<br>• Convene a technical expert working group to provide feedback on model structure and assumptions, based on reviewing model technical appendix prepared by Avenir Health.<br>• Develop a briefing document about Goals-ARM, ensuring policymakers and decision-makers correctly interpret results or even use the model, thus avoiding potential misuse. | Avenir | Appendix circulated Feb 2024, Working group meeting March 2024 |
| **Session 3: ART coverage data discrepancies**<br>*Chair: Leigh Johnson*<br>**Objective:**<br>• **Review options and progress in assuring quality of ART data and accurate ART coverage estimates** | | |
| Support country-specific triangulation analyses to understand inconsistencies between survey estimates of ART coverage and program data, with an emphasis on identifying the mechanisms underpinning discrepancies between data sources.<br>• Eswatini is working on ART program data deduplication. Reach out to Julia Kohler from USAID. (Comment from William Miller)<br>• Ghana is working on DQA / deduplication exercise; expect a large ART reduction. (Comment from Ekow Wiah) | UNAIDS | 2024 estimates |
| Consider possible alternative displays of modelled versus program-reported age distributions of ART patients:<br>(a) ratios of modelled to recorded numbers in each age group, and<br>(b) time trends in modelled versus recorded numbers by 10-year age group.<br><br>This will address the concerns that:<br>(a) small absolute difference can be relatively large, and<br>(b) differential changes over time in modelled versus reported age distributions might indicate that dynamics are accurately modelled. | Avenir | 2024 estimates |
| | Avenir, | |

| Recommendation | Lead person(s) | Timeline |
|---|---|---|
| Add a validation display in Spectrum comparing modelled female ART coverage over time with the trend predicted based on ART coverage in pregnant women entering ANC (already on ART before the current pregnancy), using a multi-country regression model (that fitted the correlation between overall female and ANC-based ART coverage based on national surveys that measured both indicators). Note, this validation may be piloted for ESA countries, where such surveys were conducted – but not elsewhere. | Jeff Imai-Eaton | 2024 estimates |
| UNAIDS to identify estimates of 95-95-95 target attainment as "uncertain" if Spectrum results based on ART program data and modelled PLHIV denominators indicate they appear close to reaching or surpassed the 95-95-95 targets, but Spectrum ART coverage estimates are inconsistent with survey estimates of ART coverage or with ART coverage in ANC. UNAIDS will recommend countries that need to prioritize DQAs. | UNAIDS | 2024 estimates |
| Provide guidance to countries on possible approaches to resolve discrepancies between survey estimates of ART coverage and program (ART and ANC) data, giving examples of how different countries have applied such corrections in previous Spectrum estimation rounds.<br><br>The guidance should indicate under what circumstances UNAIDS is likely to flag inconsistencies with survey data as problematic. | UNAIDS | 2024 estimates |
| Review evidence on different slopes of recent ART scale up between men and women to assess whether ART coverage trend among pregnant women (the regression model above) may differ from that in the overall adult PLHIV population.<br>• Systematically analyse odds ratios of ART coverage by sex and age in household surveys and evidence for change over time | Jeff Imai-Eaton | 2024 estimates |
| Use mathematical modelling to assess relationships between pregnant women ART coverage and population ART coverage:<br>• Changes over time<br>• Different epidemic contexts: male-to-female PLHIV ratio, more concentrated epidemics (West Africa)<br>• Impact of different incidence trajectories<br>• Impact of reporting biases (especially non-disclosure of ART at first ANC)<br>• Saturation among pregnant women and/or adults overall, at high ART coverage | Avenir, Jeff Imai-Eaton | 2024 estimates |
| **Session 4: Definitions, reporting and estimation of AIDS-related mortality among PLHIV**<br>*Chair: Cari van Schalkwyk*<br>**Objective:**<br>• **Review definitions of AIDS-related deaths used cause-of-death reporting to WHO, national, and in research studies**<br>• **Review definitions, interpretation, and calculation of AIDS-related mortality rates in Spectrum** | | |
| Meta-analysis of AIDS/non-AIDS mortality data:<br><br>• Systematic review and meta-analysis indicated that in high-income countries:<br>1. People with HIV have higher rates of non-AIDS mortality than people not infected with HIV (excess non-HIV mortality);<br>2. The proportion of deaths among people with HIV due to non-AIDS causes has increased steadily, consistent with Spectrum results reflecting the expected impacts of ART and ageing of PLHIV; and<br>3. The proportion of all-cause deaths among people with HIV that are due to non-AIDS causes was larger than UNAIDS 2023 Spectrum estimates. This is expected since Spectrum assumes that 100% of excess mortality is AIDS-related (apart from people who inject drugs in | Adam Trickey | 20 November |

| Recommendation | Lead person(s) | Timeline |
|---|---|---|
| selected EPP files), throughout all years and regardless of ART coverage.<br><br>• Further analysis of systematic review data was recommended to guide interpretation of results in modelled estimates: perform meta-regression to estimate proportions of AIDS-related deaths among PLHIV from studies, adjusting for:<br>1. Cause of death assignment methodology: whether study used 'CoDe' protocol[1] or other ICD-10 code assignment<br>2. Region<br>3. Year study was performed<br>4. ART coverage.<br><br>This analysis is recommended for mortality among both on-ART and off-ART, but data for off-ART were very limited and a larger share of mortality among those not on ART is anticipated to be AIDS-related. Therefore, the priority is to review the proportion of mortality due to AIDS-related and non-AIDS-related causes among people on ART.<br><br>• Expand review to include studies before 2016<br><br>• For high-income countries (i.e. WCENA + 3 countries in Asia-Pacific + 5 Gulf states in MENA), based on evidence from systematic review, consider approaches to adjusting Spectrum estimates of mortality among people living with HIV to account for excess non-AIDS mortality among people on-ART.<br>o Determine adjustment such that distribution of deaths among people with HIV from AIDS and non-AIDS causes is consistent with meta-regression results during the study period<br>o Review the potential impact of alternate approaches on estimates of AIDS deaths and total deaths among people living with HIV in Spectrum results.<br>o Among PLHIV not on ART, continue to assume that excess mortality reflects AIDS-related deaths.<br><br>• For high HIV burden countries in sub-Saharan Africa, do not make any adjustment for excess non-AIDS mortality among people with HIV (on and off ART).<br>o Continue to seek and review evidence on causes of death among PLHIV (on and off ART) in sub-Saharan Africa.<br><br>• For other regions, consider expert opinion on whether adjustments for excess non-AIDS mortality among PLHIV on ART in high-income should be applied.<br>o This has particular implications for countries in the Latin America, Caribbean, Eastern Europe and Central Asia, and Middle East & North Africa regions that use the CSAVR model to estimate HIV incidence trends from AIDS-related deaths (see Concentrated epidemic section below). | | |
| In UNAIDS reporting:<br><br>• UNAIDS reports estimates for 'AIDS-related deaths'. Clarify that UNAIDS-reported estimates of AIDS-related deaths aims to align to WHO cause of death definitions for AIDS-related deaths, based on ICD10 coding rules.<br>o Up to the 2023 round, the strategy for deriving excess mortality in Spectrum – but naming them, AIDS-related deaths -- has been to estimate excess all-cause mortality among people with HIV relative to national mortality by sex and age. This operational definition was imposed by the data underlying mortality rates coming from observed | UNAIDS | 2024 estimates |

[1] Protocol: Coding Causes of Death in HIV (CoDe) https://www.chip.dk/Portals/0/files/Code%20Protocol%202.3.pdf

| Recommendation | Lead person(s) | Timeline |
|---|---|---|
| all-cause mortality (by sex, age, CD4 category, and duration in ART) from sero-converter cohorts (for untreated mortality) and ART clinical cohorts (IeDEA and similar cohort studies), from which Spectrum derived excess mortality among people on ART as the difference between total mortality among people on ART minus mortality total population non-AIDS age-specific mortality (i.e., [AIDS-related mortality rates] = [all-cause mortality rates from cohort studies] - [background general population mortality rates]).<br>○ Continue to review data of non-AIDS excess mortality worldwide to refine approach to estimating AIDS-related mortality. | | |
| From countries that can report all-cause deaths among adult PLHIV alongside AIDS deaths (listed in Eline's presentation), see if the former is available by age and if we can learn something from this. | UNAIDS/Eline Korenromp | 2024 estimates |
| **Session 5: Mortality among people on ART**<br>*Chair: Jeff Imai-Eaton*<br>**Objectives:**<br>• **Review updated IeDEA mortality rates analysis and impact on Spectrum estimates**<br>• **Review estimated impact of ART interruption on mortality, age and CD4 of PLHIV in different settings** | | |
| Do not update ART mortality rates for regions within Africa with new IeDEA data.<br>• Data from newer tracing studies (distinguishing deaths from treatment disruption and transfers among those lost-to-follow-up) are not yet incorporated.<br>• Data from 2020/2021 sensitive to definitions of loss-to-follow-up (LTFU) and mortality assumptions for those LTFU.<br>• Increasingly limited CD4 count at enrolment, resulting in smaller fractions of cohorts included, possibly biasing estimated mortality towards sicker clients.<br><br>Review separately the patterns in South Africa's data, where unascertained mortality is lower than in most African IeDEA study sites due to linkage with vital registry data, to validate the assumption of extrapolating downward mortality trend from 2010–2017 estimated in 2018 analysis of IeDEA data up to 2022. | Reshma Kassanjee | |
| In regions outside of Africa, implement new default ART mortality rates in Spectrum based on new IeDEA estimates of mortality, conditional on:<br><br>• Simulating expected impact of updated mortality rates on Spectrum results and assuring any changes can be explained and justified.<br>• Further review and feedback from the IeDEA Executive Committee.<br><br>New mortality rates are based on IeDEA data up to 2021, with two changed data exclusions:<br><br>1. Excluding clients who had suppressed viral loads at enrolment (uncertain duration on ART)<br>2. Not excluding clients who were enrolled at a CD4 count higher than national eligibility criteria at the time. | Avenir | 2024 estimates |
| • The group reiterated the recommendation from May 2022 UNAIDS Reference Group meeting that a new ART mortality modelling strategy is required to analyse and interpret mortality trends from IeDEA cohorts that does not rely on baseline CD4 stratification now that baseline CD4 is no longer routinely collected. | | |

| Recommendation | Lead person(s) | Timeline |
|---|---|---|
| • Concern was raised about discrepancy between observed distributions of CD4 at ART initiation and modelled distributions in Spectrum results. This could result in inconsistent results for overall AIDS-related mortality. *IeDEA plans to implement future studies at selected IeDEA sites in SSA where CD4 measurement at ART initiation ore re-initiation will be offered to all patients for a 6–12-month period. The results of these studies will inform assumed CD4 distributions at initiation in recent and current years.* | | |
| Mortality among PLHIV who interrupt ART:<br>• Do not make any adjustment in Spectrum for differential mortality among people returning to care following ART interruption.<br>  ○ Contrasting findings between IeDEA South Africa and ART-CC (high-income countries) warrant further investigation into methods. | Adam Trickey, Haroon Moolla | May 2024 |
| **Session 6: ART interruption rates**<br>*Chair: Leigh Johnson*<br>**Objective:**<br>• **Recommend realistic, non-zero regional default rates for interruption from ART, for use in 2024 estimates** | | |
| Waterfall cascade visualisation:<br>• Simplify the visualisation to one chart and move it to the Program Statistics editor such that users view the chart while they are reviewing data. This should reduce duplicative data entry, and allow more intuitive use, as users will view the chart right after inputting the ART data.<br>• Encourage users to input their program data directly into the AIM Program Statistics editor. This data can be used for more purposes than validation display, such as for estimating numbers newly initiating ART.<br>• Present two bars: program data versus Spectrum, for indicators estimated such as deaths, new initiates, and unattributed/residual. | Avenir, UNAIDS | 2024 estimates |
| Default adult treatment interruption rates in Spectrum:<br>• In high-income countries, implement a default **annual treatment interruption rate of 1.6%,** based on estimates from ART-CC.<br>• *Provisional recommendation:* For non-high-income country regions implement a default **annual treatment interruption rate of 5%.** This rate was the median from countries that entered non-zero interruption rates in their 2023 Spectrum files (excluding a small number of countries and years with unique health system and sociodemographic issues, such as Venezuela). | Avenir | 2024 estimates |
| **Definitions of treatment interruption:**<br><br>There was substantial discussion about definitions of ART interruption and loss-to-follow-up from ART. There was consensus on several principles and requirements, but not on specific definitions or calculations for Spectrum inputs. Key recommendations from the discussion were:<br><br>• Spectrum users require clear guidance about the definition of treatment interruption for model estimates and how to calculate treatment interruption inputs for modelled estimates.<br>• For modelled estimates, rates of interruption seek to identify those individuals who have discontinued routine ART use, and therefore are at risk of viral non-suppression, adverse clinical outcomes, and onward HIV transmission. | | |

| Recommendation | Lead person(s) | Timeline |
|---|---|---|
| • WHO recommends defining loss-to-follow-up (LTFU) on ART as those who are more than 28 days late for a scheduled ART prescription refill (due to death, silent transferral, or treatment interruption). This concept of LTFU is primarily to guide active follow-up of patients at risk for interruption, and thereby avoid real treatment interruption:<br>  ○ LTFU in this definition overestimates treatment interruption, not only due to including deaths and transfers, but also because many individuals more than 28 days late for visit are continuing ART and have not truly interrupted treatment.<br>  ○ Conversely, it will not capture treatment interruption of individuals who have discontinued ART but are not yet late for a scheduled refill (the risk of which will increase with shift to longer multi-month dispensing).<br>• Using a threshold of more than 90 days after the last scheduled clinic or ARV pick-up visit may be a more accurate proxy indicator for real, clinically significant treatment interruption, by more specifically identifying individuals who have truly discontinued ART and are at risk of adverse outcomes.<br>• There is currently no consensus recommended definition for 'treatment interruption'. The group recommended WHO and UNAIDS consider definitions and indicators for calculating treatment interruption. | | |
| **Session 7: Concentrated epidemics**<br>*Chair: Cari van Schalkwyk*<br>**Objectives:**<br>• **Discuss planned and possible refinements to EPP-Concentrated (including EPP-ASM), CSAVR and ECDC models.**<br>• **Advise the 2024 plan for key population data collection and estimates for (non-SSA) countries not using EPP-concentrated or AEM** | | |
| **Guidance for use of EPP-Concentrated:**<br><br>Under Step 10 - EPP Incidence Curve Fitting of the UNAIDS estimates quick start guidance, nuance the earlier recommendations about use of the EPP classic model:<br><br>*EPP-Classic is normally used for geographic areas with relatively few data points. However, **if***<br>• *EPP-Classic gives an initial prevalence rise that is too sharp and implies an implausibly early and sharp incidence peak,*<br>*or*<br>• *the data themselves show a consistent time trend that could be cleanly fit with a single curve, but EPP Classic does not follow the data trend well, then another model such as R-Hybrid, R-Spline or R-Trend should be considered and the model that best fits the available data should be selected, while avoiding overfitting.*<br><br>*Use the Spectrum validations to the fullest. Notably, examine the match of EPP-estimated trends in infections and AIDS deaths, with reported HIV and AIDS diagnoses, respectively, during the pre-ART era, a validation within EPP. Utilize the "Data Check" button in EPP to compare early model estimations with actual early AIDS case trends.*<br><br>Regarding guidance for model choice:<br>• In cases where data are sparse, consider either EPP-Classic or R-Hybrid model.<br>• For countries with more consistent data over three or more years, any of the three main models (EPP-Classic, R-Spline or R-Hybrid) can be used.<br>• For data-rich countries, R Hybrid seems to be the preferred model – now including for concentrated epidemics. | UNAIDS | 2024 estimates |

| Recommendation | Lead person(s) | Timeline |
|---|---|---|
| **EPP-ASM-Concentrated**:<br><br>Merge Avenir's CSAVR-KP and EPP-ASM workstreams, which will enable use of case surveillance, vital registration and survey data. This model will be available for use sooner than Goals-ARM in concentrated epidemic settings and will provide incentive for countries to collect key population prevalence data. | Avenir | May 2024 |
| **CSAVR**:<br>• Add an option to fit CSAVR splines with less than 5 knots to reduce the flexibility on incidence patterns in settings with sparse data.<br><br>• Give users the option to constrain the number of PLHIV to be greater than or equal to the number on ART. This constraint is proposed as a new default, but with the option for users to dismiss it – in case their ART program data are believed less reliable (e.g. overstated) compared to case and death data to which CSAVR fits incidence and PLHIV.<br><br>• Continue exploring approaches to account for model uncertainty across different CSAVR model variants, because variation across models may more meaningfully indicate the true epidemic uncertainty.<br><br>• Prioritise CSAVR development of approaches capable of accounting for changes in HIV testing on numbers of diagnoses, for example reflecting that sharp declines in case diagnoses and reporting during COVID years 2020 and 2021 likely reflect reduced testing rather than true changes in HIV incidence a possible, optional covariate (reiteration of recommendation from May 2022).<br><br>• Fitting AIDS death data using AIM's excess mortality rates: Change CSAVR's outputs from excess mortality among PLHIV to AIDS deaths, matching the CSAVR-entered and fitted death data. With AIM's pending split of excess deaths into AIDS/non-AIDS, from Adam Trickey's review, this should address the current, structural under-estimation of PLHIV and incidence in pre-2023 CSAVR models. | Avenir | 2024 estimates |
| Excel-based key population data collection in 2024 estimates round:<br>• Pilot a new to design spreadsheet, prepopulated with prevalence and PSE data inputted in Avenir's Goals calibrations. Also show the resulting Goals fit for prevalence over time, as used for key population infection donuts in 2023. To entice and facilitate country feedback and updating, by illustrating the current use and relevance – with a view to 2024 key population infection donuts as well as future country-owned Goals-ARM calibrations.<br>• Time permitting, where the Avenir database and Excel extract did not specify the original/primary sources of the data (but only stated GAM or Key Population Atlas), try completing this, or invite the country to complete this. | Avenir<br><br>UNAIDS | 2024 estimates |

# Appendix B – Participants

| Name | Organisation |
|---|---|
| Adam Trickey | Bristol University |
| Alexander Viguerie | CDC |
| Alex Whitlock | University of Toronto |
| Anna Grimsrud | IAS |
| Annette Sohn | TREAT Asia |
| Ard van Sighem | Stichting HIV Monitoring |
| Austin Carter | IHME |
| Avi Hakim | CDC |
| Cari van Schalkwyk | SACEMA |
| Carl Corcoran | CDC |
| Constantin Yiannoutsos | Indiana University |
| Deepa Jahagirdar | Avenir Health |
| Ekow Wiah | NAC Ghana |
| Eline Korenromp | UNAIDS |
| George Siberry | PEPFAR |
| Guy Mahiane | Avenir Health |
| Haroon Moolla | University of Cape Town |
| Hmwe Kuy | IHME |
| Ian Fellows | CDC |
| Ian Wanyeki | UNAIDS |
| Irum Zaidi | PEPFAR |
| James Stannah | McGill University |
| Jeff Eaton | Harvard University |
| Jesse Knight | University of Toronto |
| Jinkou (Button) Zhao | Global Fund |
| John Stover | Avenir Health |
| Josh Salomon | Stanford University |
| Joshua Kimani | Partners for Health and Development in Africa |
| Juliana Daher | UNAIDS |
| Kate Rucinski | Johns Hopkins University |
| Keith Sabin | UNAIDS |
| Kelsey Case | UNAIDS consultant |
| Kyu Han Lee | Emory University |
| Le Bao | Penn State University |
| Leigh Johnson | University of Cape Town |
| Maggie Niu | Penn State University |
| Maggie Walters | Imperial College London |
| Marie-Claude Boily | Imperial College London |
| Mary Mahy | UNAIDS |
| Mary-Ann Davies | University of Cape Town |
| Mathieu Maheu-Giroux | McGill University |
| Michelle Morrison | BMGF |
| Nanina Anderreg | CIDER |
| Oli Stevens | Imperial College London |

| | |
|---|---|
| Otilia SCUTELNICIUC | UNAIDS |
| Rachel Esra | Imperial College London |
| Ray Shiraishi | CDC |
| Rebecca Anderson | Imperial College London |
| Reshma Bhattacharjee | USAID |
| Reshma Kassanjee | University of Cape Town |
| Rob Glaubius | Avenir Health |
| Romain Silhol | Imperial College London |
| Rowan Martin-Hughes | Burnet Institute |
| Salome Kuchukhidze | McGill University |
| Sharmistha Mishra | University of Toronto |
| Shona Dalal | WHO |
| Sidy Mokhtar Ndiaye | Enda |
| Sonia ARIAS GARCIA | UNAIDS |
| Tim Brown | East West Center |
| William Miller | USAID |
| William Probert | WHO |
| Wiwat Peerapatanapokin | East-WEST Center |
| Wolfgang Hladik | CDC |
| Ye Yu SHWE | UNAIDS |

# Appendix C – Agenda

## Monday 9 October:

| Time | Duration (mins) | Topic | Presenter(s)/ Lead Discussant |
|---|---|---|---|
| 14.00 | 15 | Welcome and introductions;<br>2023/24 estimates round: process and timelines | Mary Mahy |
| 14.15 | 5 | Meeting objectives and recommendation review | Cari van Schalkwyk |
| **Session 1: Key Populations data synthesis in sub-Saharan Africa (chair: Jeff Imai-Eaton)**<br>**Objective:**<br>• Recommendations on use of the Triangulator in key population estimates process in SSA<br>• Review data on PSE trends and urban/rural PSE ratios | | | |
| 14.20 | 60 | The Aggregator – extending the Triangulator to synthesize subnational key population size estimates | Carl Corcoran |
| 15.20 | 30 | PSE over time: multi-country analysis of trends, numbers, and proportions<br>• Key Population Workbook data<br>• GAM | Oli Stevens<br>Sonia Arias-Garcia |
| 15.50 | 10 | BREAK | |
| 16.00 | 50 | Urban/rural PSE ratios<br>• PHIA surveys<br>• Collated data<br>• Network scale-up method PHIA in Nigeria/Cote d'Ivoire | Salome Kuchukhidze<br>Oli Stevens<br>Anne McIntyre |
| 16.50 | 30 | Recommendations for the SSA key population data collection workbook<br>• New tab for PSE from program reach/targets/mapping<br>• Proposed use of the *Aggregator/Triangulator* and how this will fit into the workflow. | Oli Stevens |
| 17.20 | 40 | Recommendations | |
| 18.00 | | CLOSE | |

## Tuesday 10 October:

| Time | Duration (mins) | Topic | Presenter(s)/ Lead Discussant |
|---|---|---|---|
| **Session 2: Dynamical modelling of HIV trends in key populations in sub-Saharan Africa (chair: Cari van Schalkwyk)**<br>**Objective:**<br>• Review assumptions and model structure decisions | | | |
| 14.00 | 30 | Population group turnover<br>• Assumptions in Goals 2023<br>• Variation across countries (SSA; FSW, MSM, PWID)<br>• Variation over time (South Africa; FSW) | John Stover<br>Rebecca Anderson<br>Nanina Anderegg |
| 14.30 | 20 | Goals-ARM development | Rob Glaubius |
| 14.50 | 30 | Discussion and recommendations | |
| **Session 3: ART coverage data discrepancies (chair: Leigh Johnson)**<br>**Objective:**<br>• Review options and progress in assuring quality of ART data and coverage estimates | | | |
| 15.20 | 10 | Brief update on *Leapfrog* – common code for simulation models (Spectrum, EPP-ASM, CSAVR, Shiny90) | Jeff Imai-Eaton |
| 15.30 | 10 | Visualising program and survey data by age (ART coverage, prevalence) | John Stover |
| 15.40 | 10 | BREAK | |
| 15.50 | 110 | Implications and recommendations for country data entry (survey and program) in 2024 estimates round | Rob Glaubius, Mary Mahy, Jeff Imai-Eaton |
| 17.40 | 20 | Recommendations | |
| 18.00 | | CLOSE | |

## Wednesday 11 October:

| Time | Duration (mins) | Topic | Presenter(s)/ Lead Discussant |
|---|---|---|---|
| **Session 4: Definitions and reporting of HIV mortality among PLHIV (chair: Cari van Schalkwyk)**<br>**Objective:**<br>• Review definitions used in death data collections, and the need to realign or adjust for valid Spectrum fits to death data | | | |
| 14.00 | 110 | Review of definitions and reporting of HIV mortality among PLHIV | Adam Trickey |
| 15.50 | 10 | BREAK | |
| **Session 5: Mortality among people on ART (chair: Jeff Imai-Eaton)**<br>**Objectives:**<br>• Review updated IeDEA mortality analysis and impact on Spectrum estimates<br>• Review estimated impact of ART interruption on mortality in different settings | | | |
| 16.00 | 20 | IeDEA adult mortality analysis | Reshma Kassanjee |
| 16.20 | 20 | Impact of updated IeDEA analysis on Spectrum estimates | John Stover |
| 16.40 | 30 | Mortality in ART re-initiators<br>• IeDEA South Africa<br>• ART-CC | Haroon Moolla<br>Adam Trickey |
| 17.10 | 10 | On-ART all-cause deaths validation: 2023 data entries and Spectrum fits | Eline Korenromp |
| **Session 6: ART interruption rates (chair: Leigh Johnson)**<br>**Objective:**<br>• Recommend realistic, non-zero regional default rates for interruption from ART for use in 2024 estimates | | | |
| 17.20 | 30 | ART interruption rates<br>• ART-CC<br>• IAS | Adam Trickey<br>Anna Grimsrud |
| 17.50 | 10 | AIM entries and indicator names: Program stats vs. Validation>Waterfall | John Stover |
| 18.00 | | CLOSE | |

## Thursday 12 October:

| Time | Duration (mins) | Topic | Presenter(s)/ Lead Discussant |
|---|---|---|---|
| **Session 6 and 7 continued** | | | |
| 14.00 | 40 | Recommendations | |
| **Session 7: Concentrated epidemics (chair: Cari van Schalkwyk)**<br>**Objectives:**<br>• Discuss planned and possible refinements to EPP-Concentrated (including EPP-ASM), CSAVR and ECDC models.<br>• Advise the 2024 plan for key population data collection and estimates for (non-SSA) countries not using EPP-concentrated or AEM | | | |
| 14.40 | 20 | EPP-Concentrated: early trends and recommended curve types | Tim Brown |
| 15.00 | 20 | Updated Age-Sex stratified EPP-Concentrated model | Deepa Jahagirdar |
| 15.20 | 40 | CSAVR<br>• Fitting AIDS death data using AIM's excess mortality rates: possible adjustments<br>• Double Log curve: sharp artificial historic peak<br>• Splines: may overfit fluctuations in recent diagnoses<br>• Instances of ART coverage >100% of PLHIV<br>• Aggregating uncertainty across CSAVR models | Guy Mahiane,<br>Rob Glaubius |
| 16.00 | 10 | BREAK | |
| 16.10 | 15 | ECDC model/platform: 2024 update | Ard van Sighem |
| 16.25 | 20 | Collating ANC testing and prevalence data in AEM: user guidance to use for informal validation | Tim Brown |
| 16.45 | 25 | **Key population data and estimates (PSE, prevalence, ART coverage) in non-EPP countries**<br>• Data received through 2023 XLS (vs. GAM) in CAR, LAC and MENA workshops<br>• 2024 plan for (non-GAM) data collection; updating Key Population Atlas, GAM and/or *AIDS Info* with data reported in 2023-24 estimation context | Sonia Arias-Garcia<br>Keith Sabin |
| 17.10 | 30 | Recommendations | |
| 17.40 | | CLOSE | |