Integrating key population HIV estimates into the national HIV estimates process for sub-Saharan Africa

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

16-19th May 2023

Stellenbosch, South Africa

REPORT & RECOMMENDATIONS

Index

Abbreviations

AIM	AIDS Impact Model
AIDS	Acquired Immunodeficiency Virus
ANC	Antenatal Clinic
AP	Asia and the Pacific
ART	Antiretroviral Therapy
ASM	Age Structured Model
CDC	US Centres for Disease Control and Prevention
CSAVR	Case Surveillance and Vital Registration
EECA	Eastern Europe and Central Asia
ESA	Eastern and Southern Africa
EPP	Estimation and Projection Package
FSW	Female Sex Worker
HIV	Human Immunodeficiency Virus
IRR	Incidence Rate Ratio
KP	Key Population
LTFU	Loss to Follow Up
MENA	Middle East/North Africa
MSM	Men who have Sex with Men
PHIA	Population-based HIV Impact Assessment
PLHIV	People Living with HIV
PrEP	Pre-exposure Prophylaxis
PSE	Population Size Estimates
PWID	People Who Inject Drugs
SSA	Sub-Saharan Africa
TGW	Transgender Women
t-PAF	Transmission Population attributable Fraction
UNAIDS	Joint United Nations Programme on HIV/AIDS
VLS	Viral load suppression
WCA	Western and Central Africa
WHO	World Health Organization
WPP	World Population Prospects

The meeting of the UNAIDS Reference Group on Estimates, Modelling, and Projections was organised for UNAIDS by the Secretariat of the Reference Group (www.epidem.org), managed at SACEMA, Imperial College London and the University of Cape Town. Participants of the meeting are listed at the end of this document (**Appendix E**).
Cari van Schalkwyk and Akim Lukwa, May 2023

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 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. They guide UNAIDS and partner organisations in the development and use of the tools that countries employ for annual HIV estimates, which form the basis of UNAIDS Global HIV epidemic statistics. 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 convened a hybrid meeting held in-person in Stellenbosch, South Africa with remote participation through Microsoft Teams, from 16-19 May 2023. The meeting featured presentations and group discussions to generate consensus recommendations, divided into the following five sessions:

This report summarises the meeting presentations and discussions that underpin recommendations by the Reference Group. Meeting participants can access all presentations at www.epidem.org. Others can direct inquiries to the Secretariat at epidem@sun.ac.za). The final recommendations can be found at the end of this report. The recommendations (**Appendix B**) provide UNAIDS with guidance on generating HIV estimates, reviewing current approaches, and identifying required data to further improve HIV estimates. The meeting agenda and objectives are in Appendix F. Previous meeting reports are available at www.epidem.org. This transparent process ensures that the statistics and reports published by UNAIDS, and partners are informed by impartial, scientific peer review.

Introduction

Mary Mahy welcomed meeting participants and introduced the meeting participants. **Jeff Eaton** gave an overview of the meeting objectives centred on key populations. The Reference Group prioritizes gathering accurate data on key and vulnerable populations, especially in Sub-Saharan Africa, to quantify population groups experiencing infections, and determining their significance in HIV transmission, epidemic control and enhancing a national HIV response. The 2021-2026 Global AIDS strategy, centred on inequalities and key populations, has renewed efforts to equip countries with the support, tools, and guidance they need to monitor and address their key population data effectively.

Session 1: Estimation of new infections among Key Population and their partners (donuts): refined time-dynamic methods

Since 2016, UNAIDS has been using specialized in-house analyses to release estimates on new infections among Key Populations (KPs). Recently feedback emphasized the inclusion of time-trends in KP estimates, a structured collation of data from diverse sources, and addressing critical gaps in data assumptions. In response to this, UNAIDS have proposed revised methods for producing estimates in 2023, focusing on the distribution of new infections among KPs and their partners.

The objective of this session was to critically assess the proposed methods for UNAIDS Global AIDS Update 2023. Through productive discussions, attendees weighed in on technical approaches, inherent assumptions, methodological strengths and weaknesses, and effective communication of the derived estimates and changes.

Session 2: Population Size Estimation, HIV prevalence and ART coverage synthesis methods

Rooted in the discussions from the April 2021 Reference Group meeting, this session and the subsequent third session covered the intricacies of Key Populations (KPs) data. In the referenced meeting, primary topics included KP indicators' definitions, review of the Incidence Patterns Model, the availability and quality of KP data in SSA, indicators quantifying KP transmission and epidemic drivers, and models for KP stratified estimates in dense epidemic contexts. A collective recommendation emerged: the development of a comprehensive guide for collating, reviewing and synthesizing data, adoption of a systematic process, importance of permanent documentation of data sources and the rationale behind estimates. The aggregation of KP data was initiated in 2021 and will soon transition to the next phase - the development of an innovative model-based estimation tool, designed to seamlessly merge KP data to yield national-level estimates. The target users of this model are the national HIV estimates teams using the Spectrum Model. Therefore, the objective of this session was to delineate a recommended methodology and framework for these teams to ensure the effective synthesis and extrapolation of KP data into tangible metrics: population size, HIV prevalence and ART coverage estimates.

Session 3: Dynamic modelling of HIV epidemic trends in Key Populations in sub-Saharan Africa

This session's primary goal was to recommend an approach and development process to integrate a KP model into the UNAIDS estimates process, steered by individual countries. Based on the approaches and methodologies for synthesis and extrapolation of KP estimates, this session focused on the model process, framework and developing a set of assumptions for producing tools that can be used for inference.

Session 4: Review of 2023 Estimates
This session gave an overview of the 2023 HIV estimates focusing on global epidemic trends, Spectrum models used, challenges encountered with national incidence trends and knowledge of disease status.

Session 5: Transmission-dynamic developments to Estimation and Projection Package (EPP)
The objectives for this session were fourfold:

- Formulate recommendations regarding the proposed methodology for time-varying reductions in HIV transmission when under ART, as a function of (time-varying) VLS.
- Evaluate the possible non-zero default assumptions associated with treatment interruption in adult ART.
- Analyse the Spectrum's adult ART coverage estimates in settings with extensive coverage.
- Conclude with recommendations to set development priorities for EPP.

Session 1: Estimation of new infections among Key Population and their partners

Chaired by **Jeff Eaton**, the primary goal of this session was to review and offer insights on the proposed methods for the UNAIDS Global AIDS Update 2023 Report. Since 2016, UNAIDS has published the distribution of new infections. However, for 2023, a shift in approach is anticipated by:

1. Implementing a more systematic approach to collate data from multiple sources.
2. Responding to the growing demand for historical trends in the distribution of new infections.
3. Leveraging additional evidence to address key information gaps concerning data and assumptions.

Prior to the meeting session, the Secretariat of the UNAIDS Reference Group confidentially shared a draft working paper with participants of the Reference Group, including several KP experts who newly joined for this topic. Participants provided written comments and participated in a two-hour virtual meeting to discuss the working paper methods and the group's objectives. Written comments and discussion from the pre-meeting were received by UNAIDS as feedback and the Secretariat used this feedback to set the session's agenda. Minutes of the pre-meeting can be found in Appendix A.

Presentations and discussions in Session 1 revolved around:

1. Description of the proposed methods.
2. Changes to estimates from previous years.
3. Description of the Goals model (proposed for sub-Saharan Africa).
4. Data and modelling evidence.
5. Feedback from discussants.
6. Working group discussions and recommendations.

Opening the session, **Keith Sabin** emphasized the day's overarching goals:

1. Creating a tool to determine evidence-based distribution of new infections among key populations, both in the present context (year with latest data) and historically tracing back to 2010.
2. Grasping the assumptions required to produce these distributions and how to obtain the needed data.

Sabin pointed out that this initiative is designed with a brief lifecycle in mind (targeting publication in the July 2023 Global AIDS update report), with the expectation that it will be superseded by a more systematic and country-centric continuous process. This would be part of the annual national HIV estimates, using methods that will be elaborated upon in subsequent sessions. The specific objective of this process is to get defensible distributions of new infections among key populations and their partners, using data up to 2022 (including Spectrum, Goals, and other model outputs) and the group would review the most appropriate inputs and acceptable assumption levels.

1. Description of the proposed methods

Eline Korenromp began by outlining the methods proposed, preliminary results and rationale for deviations from previous results. She provided an overview of the model and data sources used for each KP and country. For every country, the total number of new infections among 15–49-year-olds, as determined by Spectrum (used as the upper limit at country level) is categorized into new infections among KPs, their sexual partners, and the remaining population. The proportions among KPs (FSW, MSM, TG and PWID) of the relevant genders' new infections are obtained from:

- Within Spectrum, EPP or AEM files for 54 concentrated epidemics.
- The Goals model for 47 SSA epidemics (FSW, MSM, PWID).
- New diagnoses from Modes of Transmission for European countries and Russia, USA, Canada, New Zealand, and Singapore.
- For countries lacking KP model or data, the method involves applying either regional median or average percentages of adult male/female infections. This results in approximately 40% of new infections among KPs and their partners relying on extrapolations. In some cases, the initial estimates of new infections among KPs and their partners exceeded the total new adult infections as estimated by Spectrum.

New infections among non-KP sexual partners of KPs (pooling clients of FSW and other partners) are derived from regional, time-constant multipliers (applied to the infection estimates for KP). This is based on a non-systematic literature review performed in 2018. For instance, it is assumed that for every 100 new infections among FSW in the Asia-Pacific there would be 25 new infections among their clients.

Overall limitations include:

- National EPP estimates give strong initial incidence peaks aligning with high prevalence observed in early surveillance data, which tends to underestimate post-2010 incidence and prevalence rates.
- Regional imputations for missing data weaken the analysis. Some regions and groups had no single country estimate, leading to imputations across regions. For example, it is assumed that the proportion of all new infections that were among TG women in

EECA, MENA, ESA and WCA is the same as the proportion of all new infections that were among TG women in Canada – 1.2% of all new infections among women.

- The proposed multiplier approach is basic. It lacks sufficient data for various regions to validate specific regional differences and lacks a time-specific aspect.
- Probable underestimation of group turnover rates in EPP, Goals and AEM models may result an underestimation of new infections among KPs given a specific prevalence rate.
- Case diagnoses by Modes of Transmission tend to underestimate the KP proportions, particularly for TG individuals.
- Goals was not yet updated to reflect 2022 estimates and is still calibrated to 2021 Spectrum files.

Korenromp revisited the recommendations from the 17 April 2023 Working Group meeting, (minutes in Appendix A) showcasing outcomes from those already implemented (including the switch for South Africa from using Goals to using the Thembisa model and the review of assumed client/partner multipliers against model outputs).

2. Changes to estimates from previous years

Keith Sabin compared results from the 'donuts' in the 2022 UNAIDS global report to the current revised results. The most notable shifts were seen in the ESA and WCA regions. In these regions, there was a marked reduction in the proportion of new infections attributed to KPs and their partners. The primary drivers behind these shifts include:

1. Using time-dynamic Goals models instead of time-ignorant outdated modes-of-transmission studies,
2. Methodological refinements:
   - to assumptions for interpolation.
   - partner multipliers.
   - constraints applied to ensure consistency between KP estimates and total national estimates at the country (not only regional) level.

3. Description of the Goals model (proposed for sub-Saharan Africa)

John Stover provided insights into the **Goals Risk-Structured Model,** a model that was developed 20 years ago to guide national programs in developing their National Strategic Plans. The model calculates HIV indicators in the adults between the ages of 15 and 49 years. It segregates the population into:

- **Men:**
  - Not yet sexually active
  - In a stable partnership
  - Having multiple partners in the last year
  - Clients of Female Sex Workers (FSW)
  - Men who have Sex with Men (MSM)
  - People Who Inject Drugs (PWID)

- **Women:**
  - Not yet sexually active
  - In a stable partnership
  - Having multiple partners in the last year
  - Female Sex Workers (FSW)
  - People Who Inject Drugs (PWID)

Each risk group is characterized by various behavioural traits such as the number of partners annually, frequency of acts with each partner, condom use, and instances of needle sharing. Transition between groups is based on average duration within each group. Recruitment into each group happens at sexual debut after age 15, assuming all 15-year-olds are HIV negative. Partners are chosen from within the same risk group, with an exception for those in stable partnerships whose partners are from any risk group, influenced by marriage rates. HIV transmission is determined by the number of partners, the number of contacts per partner, the probability of interacting with an HIV-infected partner, and the transmission risk per sexual act, adjusted for partner's stage of infection, type of sex, existence of another sexually transmitted infection in either partner, effective ART usage by the infected partner and condom use, voluntary medical male circumcision, use of clean needles, and pre-exposure prophylaxis (PrEP) in the uninfected partner.

4. Data and modelling evidence

James Stannah then briefly presented results from a systematic review on **HIV testing, treatment cascade and HIV incidence among MSM in sub-Saharan Africa.** The review employed Bayesian hierarchical regression models to synthesise 39 incidence estimates from 31 studies. Random effects for region, country, and study were included as well as a random slope for time. The HIV incidence observed among MSM over 2011 to 2020 appears to be higher in comparison to the incidence rates among all males aged 15-49 (from Spectrum, for the same period). This surge in incidence occurred despite observable increases in testing and treatment**.** Evidence suggests a higher HIV incidence among MSM in Western and Central Africa (7.8py/100; 95% CI: 2.8-36.4) in 2020. In contrast, Eastern and Southern Africa reported a slightly lesser incidence of 4.7py/100 (95% CI: 2.3-11.9) in the same year. There is a non-significant downward trend in incidence over time (IRR_{time}=0.96 (0.63-1.50)). When compared to all men aged 15-49 from Spectrum national estimates in 2020, the incidence rate ratio for MSM was 199 (73-932) in WCA, while in ESA, the ratio was 27 (13-67).

Stannah also presented Oli Stevens' re-analysis of this data, which included adjustment for age of MSM in the study populations, as well as comparisons to district level Naomi HIV incidence estimates for all men aged 15-49. This lowered the estimates of HIV incidence to 4.3py/100 in WCA and 3.1py/100 in ESA and incidence rate ratios to 136 (44-418) in WCA and 21 (13-33) in ESA (relative to Spectrum-estimated overall national male estimates in 2020).

The few studies (performed between 2006 and 2020) that provided a treatment cascade for MSM suggested that knowledge of status, ART use and viral suppression have increased among MSM, however knowledge of status remains lower among MSM than among all men living with HIV aged 15-49 as of 2020. There was insufficient evidence regarding differences in ART use and viral suppression rates between MSM and all HIV-positive men aged 15-49.

Oli Stevens presented a comparison of the empirical cohort study-based and model-based estimates related to HIV incidence and the distribution of new infections in SSA settings. His presentation aimed to answer three questions:

1. Do UNAIDS regional estimates of HIV incidence and the percentage of new infections align with empirical and model-based estimates?

Transmission dynamic models included in this analysis comprised Goals (38 countries), Optima (15), Maheu-Giroux et al (1), Mishra et al. (3), Silhol et al. (4), and for South Africa specifically, EMOD, Thembisa and Stone et al. Additionally, empirical incidence data for FSW was summarized in a meta-analysis by Harriet Jones and Rebecca Anderson, and for MSM by James Stannah. This data was crucial in determining regional incidence rate ratios (KP vs general population), which Stevens used to derive proportions of new infections relative to country-level Spectrum estimates.

The 2022 UNAIDS' estimates for FSW and MSM w*ere higher* compared to both the reviewed dynamic model estimates (which generally aligned with one another) and Stevens' estimates derived from the empirical incidence data. Similarly, for PWID, UNAIDS 2022 estimates were higher than dynamic model estimates. New infections among clients of SW were not estimated by UNAIDS in 2022, and dynamic model estimates for this subgroup vary widely due to paucity of data and different model definitions of being a client.

2. Are model estimates of HIV prevalence and population size aligned within settings and what is their impact on the percentage of new infections?

There is poor agreement between dynamic models regarding HIV prevalence and PSE within similar contexts. Model-estimated proportions of new infections among KPs strongly correlate with inputted PSEs. This raises the concern that modelled KP estimates are largely driven by an input with high uncertainty, due to data quality and scarcity.

3. What is the relationship between KP HIV incidence and cross-sectional KP HIV prevalence?

The incidence-to-prevalence ratio of the empirical estimates and the dynamical models are generally far below the ratios implied in UNAIDS 2022 estimates. This implies that UNAIDS' prior estimates might have stipulated a higher incidence of HIV for a given prevalence among KPs.

Stevens noted that these findings support UNAIDS' ongoing/pending move away from the current cross-sectional calculations of new infection proportions to a dynamic model framework, that reconciles prevalence and incidence data. The correlation between PSEs and estimated proportion of new infections across dynamic models emphasizes the importance of utilizing that country-approved consensus model inputs.

Keith Sabin's presentation on the work led by **Jerry Jacobson** provides insights into estimating HIV incidence among female partners of MSM newly infected during the year. Exploratory analyses considered transmission to a female partner as a function of stage of infection (acute/chronic), probability of transmission per sex act by stage of infection, presence of sexually transmitted infections, condom use and their effectiveness and total number of sex acts per partner. Most of these data are not widely available in most settings. However, an example calculation with inputs from the USA suggested a multiplier above that used by

UNAIDS in 2022 for WCENA, EECA and MENA regions but below values assumed for LA, CAR, ESA and WCA.

Romain Silhol presented an overview of dynamic model estimates of onward transmission from KPs to clients and other partners, as well as acquisition of infections by those groups from non-KP partners, for SSA settings. For FSW, the dynamic model results suggest that 2022 (and earlier) UNAIDS assumptions underestimated new infections occurring in the sex work periphery, in both Eastern and Southern, and West and Central Africa. However, for MSM, dynamic model estimates of transmission to female partners were below UNAIDS-assumed multipliers in ESA. Model-estimated multipliers of infections decreased over time in almost all settings for clients and non-client partners of FSW (a feature not yet reflected in UNAIDS estimates up to 2022). On the other hand, transmissions from MSM to their cisgender female partners remained consistent over 2010 to 2020.

The analysis presented is limited by the absence of estimates outside SSA for partners of PWID and TG, however similar comparisons for the Asia-Pacific region, using the AEM model are anticipated to be included soon.

Sharmistha Mishra and **Mathieu Maheu-Giroux** presented a meta-analysis of HIV prevalence among clients of FSW in SSA. Two systematic reviews were performed to improve understanding of the complex HIV transmission dynamic arising from sex work.

A review presented by Maheu-Giroux synthesized national population-based surveys conducted in SSA from 2000 to 2020, with information on paid sex. This covered 87 surveys from 35 countries. The pooled HIV prevalence among men who had ever engaged in paid sex was 5%. This group were 50% more likely to be living with HIV compared to those who did not pay for sex. Few surveys had measured ARV biomarkers (n=8) or viral load (n=9) and these revealed no significant differences in ARV or viral load suppression between men who paid for sex and those who did not. This analysis was limited by recall/social desirability bias and heterogeneity between surveys. Its strengths include large sample size, exhaustive analysis of all population-based surveys, and controlling for effects of age and residence type.

The review presented by Mishra focused on client-specific surveys and national household surveys in SSA (2004-2019). A slightly higher pooled HIV prevalence of 6.5% was found among men who had paid for sex in the past year or had sex with FSW in past year. Client-specific surveys (conducted only in urban settings,) reported a higher HIV prevalence among clients of FSW compared to national-level household surveys. This study is yet to assess effects of (confounding by) clients' age.

The presenters acknowledged that although there are imperfect data on clients of FSW, the data presented can be used to parameterize and calibrate models and to check face-validity.

5. Feedback from discussants

The Secretariat enlisted four discussants to offer diverse insights on the proposed methodology for calculating new HIV infections by population group. Their feedback guided key considerations and recommendations that the Reference Group make to UNAIDS.

Joshua Kimani spoke from the viewpoint of a researcher and key population programme implementer overseeing clinics in Nairobi (with 30,000 FSW, 10,000 MSM and 2,000 TG people enrolled in services). With biometric enrolment in Kenya’s programmes, there is accurate counting without duplication. Kimani’s emphasised the necessity of enhancing quality

of data inputs for models. He highlighted challenges in obtaining accurate population size estimates in many SSA due to criminalisation of sex work and same-sex relationships. However, Kenya's involvement of KP communities in the programmes greatly improves local data and estimates. Kimani also pointed to evolving self-identifications, such as transgender women increasingly being empowered to identify as transgender instead of MSM. This implies that PSE for MSM in older data may have been the sum of transgender women and cisgender MSM. He also raised concerns about the definition of FSW across settings and studies (ranging from sex work as a full-time job to occasional transactional sex) and potential biases this may cause.

Stef Baral, an epidemiologist specializing in implementation research on key populations in SSA expressed concerns about exclusion of programmatic data and more recent epidemiological data in the proposed methods. As definitions and methods of sampling, data analysis and interpretation evolve, Baral underscored the importance of cautious trend interpretation. He noted the contribution of the Goals model to the HIV response as undeniable. However, pragmatic motivations to design models simply as Goals – distinguishing only selected sub-populations with sufficient data – should be balanced against the risk of downweighing the role of certain communities, which likely results in decreased attention to their needs. Baral also advised careful consideration when publishing significantly different results from past UNAIDS estimates, to maintain credibility if results were to substantially change in a second-next round.

Mathieu Maheu-Giroux represented a group of modellers with expertise in modelling HIV among key populations from McGill University, Imperial College London, University of Toronto and Bristol University. The group acknowledged the challenges of constructing estimates due to sparse data, lack of standardisation across surveys, differences in assumptions and definitions and the challenging timeline. Maheu-Giroux emphasized the importance of transparency and clarity in reporting. He stressed the need to clearly state objectives, justification of the chosen approach, and ensure it aligns with those objectives. Additionally, any adjustments to initial KP infection estimates, especially if they're reduced to fit within total adult infection envelopes from Spectrum, should be explicitly communicated, with all assumptions and their supporting data clearly cited. The group recommends using the GATHER (guideline for accurate and transparent health estimates reporting) checklist in reporting these estimates. Maheu-Giroux also highlighted several assumptions warranting further discussion, such as definitions of clients and partners, group turnover rates in models used, the use of case diagnosis data from high-income countries, the extrapolation of TG estimates across regions. Finally, he recommended that estimates should be validated, and uncertainty accounted for.

The final discussant was **Leigh Johnson,** speaking from the perspective of a South African epidemiologist and modeller. Johnson highlighted the challenges of using multipliers for add-on infections among partners within a region. He cited a survey in the ESA region that showed varied acceptance of LGBTI individuals as neighbours, ranging from 5% to 80%. This disparity suggests significant variations in societal pressures for MSM to have female partners in less accepting settings. Johnson advocated for the use of dynamic models for UNAIDS KP estimates but expressed concerns regarding turnover assumptions. Specifically, the zero turnover for MSM and PWID, along with the long durations of sex work assumed by the Goals model, might lead to underestimates of new infections among KPs. Conversely, the assumption that all women who enter sex work are HIV-negative could result in overestimating new infection estimates for FSW. Johnson referred to Silhol's presentation which highlighted a decline in infection multipliers for partners of FSW in SSA over time. This decline aligns with

the observation that male HIV incidence has typically declined more rapidly than female HIV incidence in many settings. Johnson pointed out that estimates of these multiplies by Thembisa (a model Johnson developed and applies in South Africa) were outliers. This was because Thembisa broadly defines clients of sex workers (all men who *ever* paid for sex) but has a more restricted definition for sex workers (not including transactional sex). This leads to a higher numerator and lower denominator compared to other models. Finally, Johnson cautioned against using models-of-models which may lose sight of the data and assumptions driving individual model estimates.

6. Working group discussions and recommendations

During the meeting, participants were allocated to seven working groups, with six convening in-person and one online. Each group was tasked with evaluating the proposed methodology in comparison to previous years, by:

1. Highlighting a few of the main strengths of the proposed methods.

2. Identifying 2-3 priority limitations or areas of concern for the proposed methods.

3. For each concern, describe the potential implication for interpretation of results, suggest additional data or analysis that would address this concern, and (optionally) an alternative outcome/information that could be reported that would mitigate the concern. A summary of the working group responses and resulting recommendations follow, with full responses at the end of this document (Appendix C).

The group did not review final draft estimates produced by UNAIDS, and therefore does not provide any consensus recommendation about endorsement or interpretation of results.

Key strengths of the proposed methodology

The UNAIDS Reference Group reviewed proposed methods and noted several strengths of proposed revised methodology:

- Systematic and documented selection of sources for estimates in each country.
- Use of country-specific estimates underpinning regional aggregates which adds rigour to the internal consistency of the estimates (across KPs) and paves the way for country-led processes.
- Increased sourcing of estimates from dynamic models including transmission models, relying on epidemiologic data and across KP and non-KP populations, including their interaction with changes over time.
- Catalysing discussion of topics around KP data, model assumptions and structure, and consistency between models and methods used for burden estimates, UNAIDS target setting and resource mobilization.

Recommendations to address the key limitations of the proposed methodology

The group noted several limitations in proposed methods, clarity of reporting, assumptions, and data inputs. These limitations, their potential implications for results and their interpretation, and recommended strategies to address are described below.

Transparency and clarity of reporting

To enable understanding and transparency of the results and facilitate communication and understanding of future changes, the Reference Group recommends reporting of the below in an appendix. Those in **bold** are considered minimum essential information to report at *country-level*.

- **Source used for each country, population group and indicator (model or extrapolation specifying the location 'borrowed'), for the indicators:**
    - Population size, HIV prevalence, new infections, ART coverage/VLS [if applicable], in 2010 and 2022
    - **For population size inputs, indicate locations where population sizes have been adjusted from values reported by countries to Global AIDS Monitoring, but deemed implausible as inputs to estimates.**
- Country-specific estimates for population size, HIV prevalence, new infections, and ART coverage/VLS in 2010 and 2020 (the country-specific values that are aggregated to form regional estimates)
- Values for key model assumptions used in each country: for example, average duration at risk (1/turnover) for each modelled risk group, rates of sexual partnership formation & sex acts per partnership
- **Note locations where estimates of number of infections among key populations have been capped to not exceed total population estimates of new infections** and report pre/post capped values.
- **Where values are not able to be shared; specify why**
- Publish the source code of all models (without data)

The Reference Group recognizes the considerations for UNAIDS, as an international non-governmental organization, to balance full publishing of country-specific details underpinning calculations for transparency and clarity with concern about reporting country-specific numbers that have not been reviewed or endorsed by member states.

Additional recommendations surrounding reporting were made:

- Communicate the (1) timeliness of data inputs and (2) ability of data to indicate a trend in each country versus trends primarily derived from transmission dynamic model assumptions. To support this, report the years of data inputs used over time in each country for each input (population size, prevalence, ART coverage).
- Strengthen articulation of rationale for using dynamic models as preferred sources to make it clear that trends represent consistent reconstruction of trend from a single source (a dynamic model) and not a combination of estimates produced from multiple sources at multiple times.
- Clearly state if/where infections among clients and partners reflect transmission *from only* members of key populations and not from other non-KP partners.
- Clarify that infections among partners represent infections arising from all KP with HIV (not only from KP infected in past year).
- Provide a clear intended definition for each population group with notes about challenges implementing this in reporting where applicable.
- Report clear and transparent description of why estimates have changed, and strengths and limitations of new methodology compared to previous.
- Clearly describe changes by region and distinguish where changes are primarily related to:

    - Changes to model source
    - Changes to extrapolation assumptions
    - Changes to input data or assumptions
- Consider using the GATHER checklist when reporting new infections by key populations.
- Consider visualising model choices and data inputs by country with flow diagrams.

Turnover assumptions: FSW, PWID and other populations

In Goals model results used for preliminary analyses presented by UNAIDS, there was assumed to be no population turnover for MSM and PWID, and turnover for FSW is assumed to be around 10 years. The Reference Group was concerned that these durations in the risk population were longer than reflected in empirical data on duration at risk. This may result in underestimating HIV incidence when calibrating to observed prevalence. Conversely, recruitment from an all HIV-negative population into a key population may lead to overestimated HIV incidence (which is the case for FSW in Goals).

Short term **recommendation**:

- Perform rapid review of turnover data and assumptions for PWID and FSW and refit the Goals model with updated turnover duration assumptions. Solicit advice and input from surveillance experts, modellers, and community organisations.
- For countries that use population risk group structured EPP files which specify turnover durations, use the same turnover rates for Goals model calibration.
- The Goals model assumes no risk heterogeneity for MSM. Note this as a limitation for interpreting estimates and that this assumption will likely change in future.
- Avoid using the term 'turnover' (modelling terminology) in MSM when communicating to wider audience.

Medium term **recommendation**:

- Priority research area to collate, review, and meta-analyse turnover data.
- Model development priority for MSM: capture risk and age stratification reflected in surveillance data.

Infections among partners & clients / multiplier assumptions

In the proposed methods, transmission to clients/partners of KPs were approximated as a regionally fixed ratio of new infections in key populations, based on a non-systematic literature review performed in 2018. With this approach, it was assumed that no transmission occurs from KPs with existing infections, and that clients/partners are only infected from the KPs.

Recommendations from the group are to:

1) Update to using time-varying multipliers from mechanistic models where available (Goals, AEM, Optima, etc).
2) For remaining countries, inform by other dynamic models where available.
3) For remaining countries where no dynamic model is available, continue with infection multiplier approach.

Clearly communicate the method chosen for each country, including the implications about transmission from prevalent cases and non-KP partners.

Transgender extrapolation approach

Current assumption that 1.2% of all new infections in EECA, MENA, ESA and WCA are among TG populations based on data from Canada is not tenable and the Reference Group recommends using a more epidemiologically principled extrapolation approach.

Recommendation: Convene consultation meeting with (1) GATE, (2) TG epi and surveillance experts, (3) modellers to discuss:

- Preference and alternatives for reporting estimated distribution of new infections:
    - Report combined cis-MSM and TG proportion.
    - Report separate cis-MSM and TG proportion in some countries (with country-level epidemic models that estimated both groups) but combined for countries that only estimated an overall MSM group (e.g., Goals) and regions composed entirely by such models without TG (i.e., sub-Saharan Africa) ['shadow' estimates of TG new infections will be required for all countries to inform global aggregation]
    - Report distinct TG proportion in all regions, noting limited evidence from current modelling and data.
- Discuss data and assumptions to estimate:
    - TG population size
    - TG prevalence [e.g., meta-analysis prevalence ratio] / new infections

Population size estimates: rapid assessment / review

Distributions of new infections by key population group are very sensitive to population size inputs. Comprehensive presentations and discussions about population size estimation were covered in another session of the May 2023 Reference Group meeting, and **recommendations** below refer only to limitations that can be addressed in the short term.

- Review with expert group all PSEs from Stevens et al. data synthesis. Anticipate lower PWID proportion than current 0.75% in ESA and 0.1% (of men and women 15-49 years) in WCA.
- Recommend against implementing the 'sunsetting' method for deriving PSEs in which PSE estimates older than 5 years are removed from informing country estimates.
- Rerun Goals with PSEs updated to latest GAM-reported country data and aligned with benchmark percentages where needed.

Reflecting uncertainty

The Reference Group recommends reporting on key areas of uncertainty and how these may affect results. Particular attention should be given to noting which results may be sensitive to changes in future data and methods revisions.

Methods for extrapolating estimates to countries with no country-specific modelling or data for regional aggregates

The current extrapolation method uses the median of estimates for all countries in a region with estimates. If the median is 0, the mean of estimates for countries in the region is used. The Reference Group recommends using extrapolations based on epidemiologic similarity rather than UNAIDS regions (Avenir global strategy analysis example: correlated prevalence trends). This extrapolation should ideally be rule-based and there should be clear documentation of what was used as a proxy for what.

Review and validation of results

The Reference Group recommends validation of results against some publications/preprints:

- Stevens et al. - Key population size, HIV prevalence, and ART coverage in sub-Saharan Africa: systematic collation and synthesis of survey data.
- Stannah et al. - Trends in HIV testing, the treatment cascade, and HIV incidence among men who have sex with men in Africa: A systematic review and meta-regression analysis.
- Jones et al.[1] - HIV incidence in female sex workers in Sub-Saharan Africa: a literature review and meta-analysis.
- Silhol et al. - Presentation during the May 2023 Reference Group meeting.
- Degenhardt et al.[2] - Epidemiology of injecting drug use, prevalence of injecting related harm, and exposure to behavioural and environmental risks among people who inject drugs: a systematic review.

Clients of FSW

The Reference Group recommends reporting clients as distinct population groups based on a distinct epidemiologic role from other partners of KPs. This should create an opportunity for creating awareness and advocacy for programming needs amongst this population.

References

1. Jones HS, Anderson R, Stevens O, McClelland RS, Richardson BA, Thirumurthy H, et al. HIV incidence in female sex workers in Sub-Saharan Africa: a literature review and meta-analysis. *In preparation*. 2023-24.
2. Degenhardt L, Webb P, Colledge-Frisby S, Ireland J, Wheeler A, Ottaviano S, et al. Epidemiology of injecting drug use, prevalence of injecting-related harm, and exposure to behavioural and environmental risks among people who inject drugs: a systematic review. The Lancet Global Health. 2023 Mar 27.

Session 2: Population Size Estimation, HIV prevalence and ART coverage synthesis methods

Leigh Johnson chaired this session. Its objective was to recommend approaches and processes for national HIV estimates team users to synthesize and extrapolate key population survey data for population size, HIV prevalence and ART coverage.

Discussions centred on: 1) Key population data collation and 2) Methods and tools for key population data synthesis extrapolation.

Keith Sabin summarized the need and use of key population size estimates. He reiterated the need to have a process that guides harmonization of estimates, because these are critical for programming.

1. Key population data collation

Oli Stevens presented KP data collated in the 2023 UNAIDS-supported estimates round from Excel workbooks deployed by sub-Saharan Africa estimation teams. First, he summarised the data extraction and analysis of 2022 round KP data collection, the first round that the Excel workbook approach was used, earlier presented at the May 2022 Reference Group meeting and submitted as a preprint. The goal of the KP workbooks was to create national-level estimates of new HIV infections for key populations, by consolidating data and facilitating and standardizing a country-led review of data availability and quality. Workbooks used in 2022 and 2023 estimates had been prepopulated with earlier published data extracted by ICL, and countries could add data and scrutinise the quality of the prepopulated data. In each country's workbook, the country's estimates can be compared to estimates from other countries in the region, to help flag potential outliers. The hardest part of the KP workbook is creating consensus national estimates – limiting to data that pass quality review and extrapolating any subnational estimates to nation-wide. Where local data are not available, the workbook offers dynamic model estimates (from Goals and Optima in latest available calibration by Avenir Health or the Burnet institute), and where local data were available, users could calculate weighted average of subnational data including if from different time points. Firm guidance on what countries should do is necessary. Countries could compare their consensus estimates with the estimates from Goals and the Incidence Pattern Model, and if the numbers were close, use the modelled estimates of new infections per key population. If consensus estimates were far off the (Goals/Optima/IPM) models estimates, countries could request refitting of those models using their consensus data.

In the 2023 estimates round, only 7 countries in WCA and 5 in ESA submitted workbooks with completed consensus estimates to UNAIDS.

An important default assumption in the workbook, based on expert consensus, is that there are fewer KPs in rural areas than in urban areas, with a universal rural-to-urban ratio of 0.6 (rural PSE proportion=0.6*urban PSE proportion).

Despite few countries completing the workbook up to the consensus estimation step, the process had numerous successes including highlighting the utility of data quality review and triangulation across SSA countries, demonstrating the need for countries to start reporting PSE proportions alongside counts (thus enabling standardization and extrapolation over time),

and providing a tool to replace earlier move regional estimates by country-owned estimates. However, there is still reluctance in integrating nationally owned surveillance with academic studies, and extrapolating PSE estimates remains a challenge. Stevens recommended reconsidering the timing of workbook completion viz-à-viz HIV Spectrum estimates, exploring ways to better integrate with Spectrum, and for UNAIDS to consider reviewing and publishing of consensus population size and HIV prevalence estimates.

2. Methods and tools for key population data synthesis extrapolation

Carl Corcoran presented a web-based tool called The Triangulator: a flexible statistical tool for reaching consensus on KP population size and prevalence, based on a Bayesian hierarchical model. As illustrated in Stevens' presentation about the KP workbook, synthesizing expert knowledge with empirical estimates to create a consensus estimate can be challenging and this tool aims to combine these in a statistical rigorous framework. Expert knowledge is quantified through the choice of a prior as well as indications of confidence in each empirical estimate. The current version of The Triangulator combines data from a single location into a consensus estimate for that location, but Corcoran and team are extending the tool to combine estimates from several locations into a national estimate and will consider a time dimension.

Kate Rucinski presented on KP data synthesis and extrapolation to estimate population sizes, using Namibia as an example. This research and country team thoroughly considered all data that could inform estimates, including KP programme statistics. The group used The Triangulator to reach consensus PSEs, using the KP programme data to inform the prior. The inclusion of programme data generally increased PSEs for FSW but decreased PSEs for MSM. Rucinski pinpointed that extrapolation approaches are sensitive to key decisions including: inclusion of empirical estimates, chosen denominators, the choice of auxiliary data to inform relevant strata, and covariate selection for the Triangulator.

Le Bao presented on statistical methods for KP indicators in SSA, focusing on two areas: capturing spatial heterogeneity at district level and doing estimation over time. Bao described the mixed effects model, that captures spatial heterogeneity and allows for uncertainty by study methods. In their a recently published paper, estimated FSW population proportions ranged from 0.3% (of adult women 15-49 years) in Malawi to 2.8% in Burundi, and from 1.1% to 2.3% in different regions within Nigeria. This challenges the validity of assuming fixed proportions by region. A time component may be added to their model next, but longitudinal data on KP size is limited even for data-rich settings and real time trends may be confounded by changes in study design. Longitudinal estimates of HIV prevalence and ART coverage for KP may be better achieved by applying dependence structures between key populations and the general population.

Oli Stevens presented on data synthesis methods to estimate HIV prevalence and ART coverage in SSA, using data collated through the 2023 round KP workbooks. Age-location-year matched total population HIV prevalence or ART coverage can predict KP prevalence or coverage, with fixed effects for region and study method, spatially correlated province-level effects and random study-level effects.

Finally, Stevens presented on adjusting MSM surveillance and survey prevalence data for age group, before inputting them in AIM or Goals calibration. The average age of MSM in surveys

are young, with a large majority under 30. He showed that changing the population denominator to men aged 15-29 years could increase the MSM population proportion from around 0.5% to 1% in SSA, although with limited impact on HIV prevalence and ART coverage. To address the large impact of age on PSE, he recommended age-matching.

Topics presented on during this session were further explored in working group discussions in Session 3; these are summarised here and in Appendix D.

Session 3: Dynamic modelling of HIV epidemic trends in Key Populations in sub-Saharan Africa

This session, chaired by **Cari van Schalkwyk,** aimed to recommend a proposed approach and development process to integrate a key population model in the UNAIDS estimates process.

Discussions centred on:

1. Modelling approach and process for key population estimates
2. Review of Goals Age-Risk Model and other model approaches
3. Indicators for quantifying role of key populations in epidemic control
4. Proposal and work plan for the next year.

Jeff Eaton introduced the session and reiterated the need to develop a new model-based estimates tool within 2023-2025. This tool should synthesize key population data and produce national KP estimates critical for national planning and global reporting; its target users are national HIV estimates teams using the Spectrum model. Eaton summarized the country-led HIV estimates process. At the April 2021 and May 2022 meetings, the Reference Group recommended development of transmission dynamic models to estimate the distribution of new infections among KPs, since these combine data/outcomes over time and can produce counterfactual-based analyses.

As technical requirements of proposed transmission dynamic models, Eaton proposed:

- Estimates and process be consistent with national Spectrum estimates of PLHIV, ART coverage and new infections.
- Produce estimates minimally from 2010 to present for KP size, HIV prevalence and incidence, and ART coverage.
- Estimate the distribution of new infections across groups.
- Estimate the distribution of likely source of transmission across groups.
- Produce estimates for KP even if local data is limited/not available.
- Account for population dynamics within each KP (entry/exit from risk or 'turn over').
- Be calibrated to locally collected KP surveillance data or consensus estimates.
- Visualize comparison to locally available KP surveillance data.

Additional desired features for the proposed model are:

- Estimates should be age stratified.
- Include estimates for former KP members and their partners.
- Represent intersectionality of risk.

- Produce output indicators such as the transmission population attributable fraction (t-PAF).
- Reflect sexual risk heterogeneity among the remaining population and key population partners.
- Represent prevention interventions (PreP, condom use etc.), with usage/coverage and impact amongst key populations.
- Involve some formal statistical data synthesis (such as HIV prevalence, PSE).
- Relate modelled intervention coverage estimates to KP programme data.

The target timeline is to:

- develop model and demonstrate its application on a subset of countries (as a desk-review exercise), at the UNAIDS Reference Group technical meeting in October 2023,
- pilot the model with several SSA country teams alongside the 2024 UNAIDS HIV estimates process (December 2023-March 2024).
- implement the model in a software tool for use by national teams as part of the 2025 UNAIDS HIV estimates process (by December 2024).

A request for proposals was widely distributed in March 2023, and three proposals were received. Proposals were presented to a subset of the Reference Group during an online meeting on 21 April. The three submitted proposals had different, but complimentary aims: modelling methods (Avenir), modelling process (Mishra et al.) and data extrapolation approaches (Bao et al.). Bao presented during Session 2, and Avenir (Glaubius and Jahagirdar) and Mishra presented their proposals in this session.

Sharmistha Mishra presented a proposal for the KP estimates process and approach.

Mishra proposed as **guiding principles** for model development:

- Model development/modifications: To minimise the potential to underestimate t-PAF of KP (e.g., in decisions about force of infections, mixing assumptions, turnover and KP members recruitment)
- Data preparation: To capture uncertainty in calibration targets and behavioural inputs by formalising consistency checks, triangulations and adjustments. (e.g., variability in sampling design across sources).

Mishra proposed **steps for model development** that can be separated into these two arms, also noting as important ‘step 0’ identifying a few reference KP models for a range of settings. She recommends engaging KP community advisory boards, KP programmes and country teams in the development process.

- Model development/modifications:

  Step 1. Design new model or modify existing model:

  - Map KP and wider population strata that should be modelled
  - Map overlapping exposure routes and the force of infection
  - Implement model for efficiency and sensitivity analyses

Step 2. Model calibration and standardized outputs

Step 3. Model checks:

- Internal validity - checks of outputs
- External validity - compare fits with reference KP models (e.g., ratio of 1-year new HIV infections among FSW vs clients)

Step 4. Estimation of indicators with uncertainty (i.e., tPAFs):

- Define counterfactuals
- Compare tPAFs with reference KP models
- Data visualization for tPAF

- Data preparation:

  Step 1. Calibration targets from KP Workbooks

  - Develop formal checks + adjustments (e.g., FSW to client ratio)

  Step 2. Input parameter synthesis

  - Per-country & regional (for missing country data)
  - Develop formal triangulation checks + adjustments (e.g., condom-use during sex work, from different sources)

Mishra highlighted some key **process steps** to be followed by country teams when implementing the tool, including:

1. Generating calibration targets by applying KP workbook using formal guidance, checks, and adjustments.
2. Options for revising parameters (editable priors for country- and region-specific parameters)
3. Calibration and guidance-set of checks on outputs
4. Estimation with uncertainty
5. Sensitivity analyses

She then laid out some considerations raised by her working group on this process:

1. Developing a new model versus adapting an existing model
2. Minimizing model complexity to maximize ability to use while minimizing biases in generating KP indicators.
3. Resource availability and time constraints
4. Identifying and engaging with reference KP models with country-level partnerships
5. Leveraging KP epidemiological data and evidence synthesis expertise across teams.

Deepa Jahagirdar presented the Estimation Projection Package-Age-Sex Structured Model (EPP-ASM) for key populations, first introduced to the Reference Group at the May 2022 meeting. This model is an extension of the EPP-concentrated epidemics model, including age structure, by integrating the demographic projection model of Spectrum. Using Senegal as example, Jahagirdar showed that the model fits well to available data and results are sensitive to assumptions about age-distribution and turnover rates. To implement this model, it will need

to be developed for generalised epidemics, implemented in C++ and go through a vetting of data and assumptions. She highlighted the pros and cons of pursuing this. The pros are that EPP-ASM slots into the current estimates framework and tools that countries know, captures age/sex dynamics and supports moving towards a comprehensive model. This tool may be ready to launch in a short timeframe. The cons are that EPP-ASM requires assumptions that do not have much data (e.g., age distribution of KPs) and the model still lacks dynamic transmission between subgroups.

Rob Glaubius presented the Goals Age-Risk Model, developed to answer questions at the intersection of age and risk. Avenir Health' first presentation on Goals-ARM was at the May 2022 Reference Group meeting. Based on feedback at this meeting, clients of FSW were added as a modelled population. Recruitment to and exit from KPs are like what Stover described in Session 1 for the Goals Risk-Stratified model, except that data on the age distribution of KPs are also used to inform transition rates. Glaubius gave an overview of HIV transmission dynamics within and across different types of partnerships, sexual mixing by age, and by behavioural risk group. He mentioned that several of the model's current assumptions have been challenged over the preceding two meeting days; these will be refined with continued input from the Reference Group. Avenir aims to have a working prototype by the October 2023 meeting.

Sharmistha Mishra presented indicators for quantifying the role of key populations in HIV epidemic control for policy makers, answering the question: What is the disproportionate risk of acquisition, direct transmission and onward (indirect) transmission among subsets of the population? One metric is the transmission population attributable fraction or tPAF which translates a risk factor into dynamics of HIV transmission: the proportion of cumulative HIV infections over time in a population that stem directly and indirectly from the risk factor or prevention gap. This is calculated using a dynamic model and comparing new infections in the *base case* scenario with new infections in a *counterfactual* scenario – defined (for example) as zero transmission and acquisition risk among sex workers or zero transmission and acquisition risk among sex workers to and from their clients. Another metric is a prevention fraction, where the counterfactual scenario is not complete exclusion of transmission from the population of interest, but some assumed success of an intervention (such as 90-90-90 targets met).

Jesse Knight presented three inputs into models that have a substantial impact on tPAF estimates: turnover (higher tPAF with higher turnover), partnership durations (typical models overestimate tPAF of longer partnerships and underestimate tPAF for shorter partnerships) and population size (higher tPAF with higher PSE).

John Stover presented on other indicators that may be more familiar to policymakers and easier to understand, such as the proportion of new infections (donuts), numbers of infections averted if prevention targets are met, cost-effectiveness of interventions (should we invest in prevention in this population given constrained resources).

In discussion, Tim Brown noted as his personal experience that presenting graphs with different scenarios of what different interventions can achieve over a longer time frame, typically 10 or even 20 years is as or more effective at influencing policymakers than presenting a single number such as tPAF or cases averted.

Working group discussions

Meeting participants were allocated to seven working groups (6 in-person and one online) and asked to identify 2-3 priority areas to discuss and make recommendations from the list below:

Key population data synthesis:

1. *KP data assessment and inclusion/exclusion:*
    - Define criteria and process to guide quality assessment of key population survey data [e.g., in KP workbook] and determine inclusion/exclusion/weighting in data synthesis
2. *KP data synthesis guidance:*
    - Draw decision tree to define key population data synthesis method choice
    - Define tools required for implementation by national HIV teams
    - Review criteria to ensure appropriate application and objective results
3. *Urban/rural PSE assumption:*
    - Specify currently available data and analysis to improve 0.6 rural:urban population size proportion assumption
4. *Key population programme data in workbooks:*
    - Define programme data indicators that should be captured in workbooks
    - Define accompany metadata and contextual data that should be recorded
    - How should these be interpreted compared to estimates?

Key population model development:

5. *Counterfactual-based indicators:*
    - Define the counterfactual indicator(s) that the new model should produce
6. *Validating key population transmission model decisions:*
    - Identify one key high priority model structure decision based on morning presentation and discussion
    - Describe how to determine or validate the decision (model analysis, data synthesis, expert input)

A summary of the working group responses follows, with full responses (Appendix D) and summarised recommendations (Appendix B) at the end of this document.

Working group discussion: key population data synthesis

1. *KP data assessment and inclusion/exclusion:*
    - Inclusion/exclusion criteria by method of data collection alone should not be used, but a decision tree should be created, and national workshops run to move towards quality inputs to be synthesised.
    - Model definitions of the KPs should be considered when including/excluding data and therefore must be well communicated.
2. *KP data synthesis guidance:*
    - The use of a Bayesian synthesis tool such as The Triangulator was recommended.
    - Development of The Triangulator to integrate longitudinal estimates was recommended.
3. *Urban/rural PSE assumption:*

- Several groups recommended a systematic review of all available data
- One group made concrete suggestions of analyses:
    i. Urban/rural ratio for MSM in Hodgins 2021 review was 0.7
    ii. Some PHIAs asked about sex of last 3 partners. This data could be used under assumption of same willingness to report
    iii. Compare rural-to-urban migration of MSM in KP surveys to an appropriate comparator (non-KP population; e.g., DHS, year-matched) and obtain the ratio
    iv. Georeferenced KPSE could be overlaid with data on gridded population density and assess correlation
- Urban/rural differences in willingness to report should be considered when interpreting data

4. *Key population programme data in workbooks:*
    - If programme data is to be included, guidance on its use will have to be provided.
    - Concerns about data duplication, targeted programming (to areas with larger PSEs), and perverse incentives to meet targets may cause biases in data and therefore data should only be used contextually (not as calibration data or as priors in Bayesian tools).
    - Recommendation: add another sheet to the KP workbook with programme data and integrated visualisation where appropriate.

Working group discussion: key population model development

5. *Counterfactual-based indicators:*
    - Groups were divided on the use of tPAF as an indicator for routine reporting, but agreed that if it will be used, careful messaging should be prioritised to avoid the impression of blame.
    - Different indicators should be used for different purposes/questions: tPAF seems most useful for advocacy and broad program priorities while indicators such as prevention fraction, impact and cost-effectiveness will be more useful for deciding between interventions
6. *Validating key population transmission model decisions:*
    - Model differentiation of sexual transmission categories should be consistent with variation in risk (model outputs compared to survey data)
    - Trends in PSE over time should be considered (rural to urban migration may play a role; changes in acceptance of MSM over time etc).
    - Important to include age structure among KPs to account for varying levels of risk.
    - Identify existing, complex models and determine how simplifications within those models affect the outputs.
    - Create simulated data sets and see if complex and simple models can capture the ground truths of the simulated data.

Session 4: Review of 2023 Estimates

Josh Salomon chaired this session with the objective to review challenges arising during the 2023 estimates process.

Eline Korenromp gave an overview of the 2023 HIV estimates, focusing on global epidemic trends, Spectrum models used and challenges with some national incidence trends and knowledge of status.

Korenromp noted as main changes in 2023 Spectrum software:

- most countries updated demography from WPP 2017 to WPP 2022.
- adult mortality on ART was reduced for high-income countries and increased for Asia-Pacific (based on refined IDeEA analysis).
- the transmission impact of ART (omega) can now be made country-specific, based on a country's viral load suppression data over 3 recent years (detailed in Session 5).

Korenromp noted as improvements in 2023 concentrated epidemic Spectrum estimates:

1. Increased entry and use of routine ANC testing prevalence;
2. More countries adopted EPP's or CSAVR's IRR by sex, instead of the concentrated epidemic global default or an unexplained custom pattern;
3. To report knowledge of status, more countries switched to CSAVR thus substituting absent or questionable program data;
4. MENA Gulf countries switched from *de jure* (nationals only) to *de facto* (all residents) as population basis. This resulted in up to 5-fold lower incidence, prevalence, mortality rates for similar numbers of infections.

Using Jamaica as an example, Korenromp illustrated typical differences in epidemic trends estimated by EPP versus CSAVR. EPP is heavily influenced by high prevalence in early years (probably biased by oversampling high risk areas) which implies lower incidence in recent years to maintain high levels of prevalence. CSAVR estimates flatter and later epidemics, possibly biased by under-reporting in early years.

Session 5: Transmission-dynamic developments to Estimation and Projection Package

This session, chaired by **Mathieu Maheu-Giroux** aimed to:

1. reach recommendation on a proposed method for time-varying reduction in HIV transmission when on ART.
2. review possible non-0 default assumptions about interruptions in adult ART.
3. review Spectrum's adult ART coverage estimates in high-coverage settings.
4. recommend priorities for development of EPP.

In Spectrum, the effect of ART on transmission has up to 2022 been an average reduction in transmission per 1% increase in population ART coverage:

incidence = transmission rate * prevalence * (1 – 0.8*ART coverage). (Equation 1)

This globally fixed, time-constant assumption ignores that both age and risk distribution of PLHIV on ART, and viral load suppression on ART vary over time. The 0.8 parameter, or omega, was derived using a meta-analysis of transmission dynamic models results over the period 201x to 201y (October 2020 Reference Group meeting). Now several countries have reached very high levels of VLS, it is apparent that the omega value of 80% must be reassessed and possibly be made flexible. For 2023 estimates, based on an interim modelling analysis where the marginal effect of ART increased roughly by 0.01 per 1% increase in VLS, a country-specific calculation of omega was introduced, calculated as the percentage of VLS among people on ART (averaged between the 3 most recent years of VLS data entered to Spectrum) minus 0.7, to a minimum of 70%.

Eline Korenromp presented on omega values used in 2023 country estimates. In 65/151 files extracted, the default of 80% was still used, and new averages per region across countries varied from 85% in ESA to 78% in the Caribbean.

John Stover then presented a possible method for time-varying HIV transmission when on ART. Varying the calculated omega over time as a function of increasing VLS for SSA countries had a negligibly small effect on estimated new infections. The analysis will be extended to countries outside SSA.

Jeff Eaton presented on ART coverage based on programme-reported numbers versus household surveys and ART in ANC. Spectrum uses programme data to estimate ART coverage and as ART coverage becomes high, small margins become very important. Small errors, such as overcounting in ART programme data or in the PLHIV denominator applied to surveys' prevalence, become critical. Eaton compared ART coverage based on programme data with PHIA surveys, which revealed relatively large discrepancies in some (of the SSA) countries.

Strategies used for addressing such discrepancies in Spectrum include:

1. Adjusting ART programme data downwards by a fixed % in recent years to align with a recent survey (e.g., in Uganda);
2. Calibrate to survey ART coverage in EPP fitting (e.g., Eswatini, Botswana);
3. Increase Spectrum total population size for more consistency with HIV prevalence and ART coverage measured in a PHIA (e.g., Nigeria, Zambia).

The second strategy in Eswatini increased estimated PLHIV to reconcile programme and survey ART data, but increased Spectrum prevalence to be above the survey prevalence estimates (especially among women of reproductive age in recent years).

Eaton showed that ART coverage among adults aged 15-49 from surveys is highly correlated with year-matched percentage of ANC clients with HIV already on ART prior to first ANC at the national level, although pregnant women is slightly higher, as expected given they are younger and so more recently infected. A plot of national ART coverage from programme data vs percentage of ANC clients with HIV already on ART prior to first ANC highlighted several countries of concern, where the former was 80% or higher, but the latter below 40%. Eaton concluded that the pattern of programme data suggesting ART coverage above PHIA surveys is concerning. Apart from too high estimates for the treatment (second 90 and third 86) cascade estimates, if left unadjusted, these program data lead Spectrum to underestimate current and future incidence -- probably more so than an incorrect omega value. Solutions to this problem will be context-specific, with countries having to decide which data sources they believe are most reliable (programme, survey, or population estimates).

Rob Glaubius presented on options and effects of calibrating program data on ART by age in Belize. Belize's baseline AIM fit gave a Spectrum-estimated age distribution skewed older than the programme data. Two incremental adjustments that improved fit were made: 1) fitting incidence rate ratios (IRRs) by age to programme ART by age and 2) changing the adult ART interruption rate from 0% to 20%. A third adjustment, changing the ART initiation allocation weights, produced a different CSAVR incidence trend than the other two adjustments and is not discussed below. IRR fitting alone may overcorrect for misspecification of treatment interruption and the ART uptake mechanism but allowing for treatment interruptions led to incidence patterns (by age) consistent with the baseline fit. Allowing for treatment interruption increased new infections and PLHIV (as expected), but unexpectedly decreased HIV deaths. This may be because Spectrum is constrained to match the programme data of number of people on ART. Without interruption, there is less capacity for younger people (with lower mortality) to initiate ART. Further investigations were suggested to understand the sensitivity of CSAVR incidence estimates to ART allocation weights. [In Spectrum, people initiate ART proportional to either the number of people in each CD4 category who are not on treatment, or the expected number of deaths absent ART in the CD4 category. The relative weights given to these respective options are called allocation weights.]

Rob Glaubius presented on ART interruptions in adults. ART dynamics in Spectrum are driven by the number of patients on ART at the end of each year and the annual percentage of patients who interrupt ART each year. ART interruption affects AIM's age distribution of PLHIV on ART: if ART interruption is understated, ART coverage and survival may be overestimated in older PLHIV and underestimated in younger PLHIV. In 2023 Spectrum files, almost 50% of countries retained the default of no ART interruption and another 30% entered values of 5% or less.

Glaubius showed how varying ART interruption impacts age-specific ART coverage and HIV-related mortality. At lower ART interruption rates, indicators are sensitive to small changes while at higher interruption rates, indicators seem relatively stable. Thus, accurate ART interruption estimates are preferred; barring that, overestimation seems safer than underestimation. Since countries don't often know what data to enter in Spectrum, use guidance is needed, aligning terminology in Spectrum with this guidance, and provide default

values based on literature review and/or primary data analysis for countries that lack national or nationally representative program data.

Jeff Eaton then reported on proposals for future development in EPP-ASM to account for transmission dynamics in population incidence trends in SSA. These proposals are motivated by the earlier presentations about the omega parameter: EPP's way of incorporating the impact of ART and VLS on new infections should be improved.
EPP infers incidence from trends in prevalence data and is not sensitive to small changes in recent prevalence data which is affected by ART through reductions in new infections and reductions in mortality. The time-constant omega parameter doesn't account for ageing populations living with HIV, changes in VLS with new ART regimens, or people initiating treatment earlier.

Some mechanistic features of Goals-ASM can be incorporated into the EPP-ASM model since it shares similar structure of sex, age, stage of infection and treatment status stratification. Instead of the simple transmission equation in EPP (Equation 1), the Goals transmission equation accounts for change in lifetime partners over time, age mixing, infectiousness by stage of HIV infection, viral suppression over time, circumcision coverage by age over time, condom use over time, STI co-infection, and PrEP use. It does not account for heterogeneity in sexual mixing and risk assortativity.

EPP development proposals to consider are:

1. **For generalized (SSA) epidemics, replace sub-national (urban/rural) by a national EPP structure.** This will address the following limitations of the sub-populations structure:

    a. EPP's Urban/rural stratification not consistent with how programme (ART) data is reported, in general and with as added difficult, patients attending ART services at location other than residence.
    b. Subpopulation demographic input files (produced by Avenir, applying urban/rural ratios to WPP 2022) are confusing to users and need updating.
    c. May inhibit moving towards more age-structured and detailed inputs.

However, there may be challenges in moving towards a national structure. ANC sentinel sites that informed historical epidemic patterns are not representative of the full national epidemic (but initially oversampled urban areas) and it will be important to capture that heterogeneity in historical prevalence data and weigh ANC likelihood appropriately to inform the national epidemic.

2. Incorporating transmission mechanisms into EPP-ASM, such as:

    a. *Time and age-varying viral suppression by ART.* Age-stratified data can be used but are not required of using default odds ratios by age from surveys (which are relatively consistent over time).
    b. *Infectiousness by stage of infection.* High infectiousness in early infection may become more important as ART coverage increases.
    c. *Variation in acquisition risk and transmission risk by age*. Not necessarily an explicit mixing matrix but considering shifting relative contribution to the aggregate force of infection as the HIV population ages.
    d. *Full, explicit age-mixing matrix.*
    e. *Circumcision coverage by age*. Important for recent trends and sex ratios in incidence.

3. Fitting to age-specific HIV prevalence and ART coverage and ART programme data

As shown in previous presentations, there are several countries where the incidence rate ratio fitting fails to fit to age patterns in prevalence and/or ART data. Previous analysis of fitting EPP-ASM to age-stratified prevalence data showed improved precision of estimates and improved out of sample prediction and was recommended for implementation by the Reference Group. This analysis will be updated with age-specific data from more recent household surveys and age-specific ART programme data.

The Reference Group recommended the listed EPP developments. This and other recommendations from this session are listed in Appendix B.

Appendix A: Minutes of working group meeting

Working Group of the UNAIDS Reference Group on Estimates, Modelling, and Projections: new infections by Key Populations and their partners (donuts)

17 April 2023
Virtual meeting

Objectives

- Solicit technical feedback on proposed approach and methods to estimate proportion of new infections by key population and their partners for the 2023 UNAIDS report [published July 2023]
- Identify key data or assumption gaps and identify opportunities to address these in the month following this meeting.

Agenda

1. Overview of objectives and proposed methodology
2. Discussion of some of the main comments made in the working paper:
    a. Terminology – Transgender, Men who have sex with men, prisoners.
    b. Proposed models and approach for deriving sources
    c. Data input quality: Size estimates and other data
    d. Assumptions about infections among non-KP partners of key populations

Action points

- Consensus to not explicitly represent transmen, male sex workers, and prisoners in the estimates due to data limitations.
- Trans/cis terminology:
    - Consult GATE to confirm preferred terminology for transwomen, transpeople or other, considering the limitations of the data and modelling process.
- Understanding changes in estimates from previous regional estimates:
- Review magnitude of changes in each region to identify where to prioritise further interrogation (UNAIDS)
- Document for each case what was the driver of change: (1) change in source for KP estimates (e.g., old MoT to Goals), or, (2) changes in inputs from previous estimates (population size, calibrated prevalence)
- Validate the spreadsheet multiplier approximation to KP partner transmissions with dynamic models.
- Request existing mathematical models to produce outputs for number of transmissions to partners of KPs where available (inquiry to be sent Goals, AEM, Thembisa, Optima, other models; some outputs available from existing KP model comparison).
- Circulate systematic review about number of non-KP partners of KPs to other modellers in the working group to inform model assumptions (expected mid-May).
- For WCENA/high-income countries, refine inputs from KP diagnoses focusing on first-time diagnoses only. Explore use of ECDC's new tool for imputation for missing transmission categories on country datasets.

Minutes

1. Overview of objectives and proposed methodology

The distribution of new infections among key populations and their partners has been published in annual UNAIDS global AIDS update reports since 2016. The proportion of new infections attributed to key populations have increased over successive rounds (which each estimated for the then most recent year only), raising the question whether these increases are real or due to the changes in methods or unput data.

A refined time-dynamic approach now being drafted will extend the estimation to cover 2010 and 2022 and build the estimation bottom-up from country-level distributions of new infections (instead of at regional level at once). The new approach also seeks to refine the estimation for clients of FSW, now as a group distinct from other non-KP sex partners of KP. All this, in the context of an ongoing move away among concentrated epidemic countries from key population stratified Spectrum (EPP) estimates, now often replaced by national-only CSAVR estimates are no longer available.

In an interim 2023 approach discussed, estimates for the Key Populations included:

- For 54 concentrated epidemics: EPP-Concentrated or AEM
- For 47 SSA epidemics: Goals
- For WCENA countries and selected other concentrated epidemics with case diagnoses by Mode of Transmission: KP-specific new case diagnoses by Mode if Transmission (including such data reported via ECDC)
- For 24 countries that had no estimates or reliable case diagnoses, regional median proportions were taken, or for some countries and some KPs, medians across regions.

The total number of new infections (male and female aged 15-49) estimated by Spectrum, 2022 round at the time, for each country constrains the totals.

In this interim 2023 approach, onward infections among clients and other partners of KPs is reflected as regional, time-constant ratios applied to new infections in each key population, based on a non-systematic literature review performed in 2018. For example, it was assumed that each new infection in a MSM in Asia Pacific will result in 0.15 new infections in female sex partners based on marriage rates of men with HIV in Asia Pacific (majority MSM). It is assumed that no transmission occurs from KPs with existing infections.

2. Discussion

a. Terminology

A common theme of comments concerned terminology for transgender people, including the contrast between cis and trans people, and stratification of estimates into trans men and trans women. The latter is constrained by lack of data, so estimates would mostly be based on assumptions. Since new infections among trans men are very few, an option would be to only provide estimates for trans women, and not trans people. Another concern raised in the comments was on the appropriate comparison population for trans women: adult men, or adult women. The recommendation on all these questions was to consult a community organization such as GATE to confirm preferred terminology and advise the preferred modelling approach considering limitations of data.

Comments questioned the lack of estimates for male sex workers and prisoners, but the consensus is that there are insufficient data on these groups for reliable estimates.

b. Proposed models and approach for deriving sources

The draft 2023 estimates discussed presented quite substantial changes from what was published in the 2022 UNAIDS Global Report for the year 2021. These changes are largely driven by the switch to using the Goals model for SSA, capping totals to Spectrum totals at the country instead of regional level, and improved KP population size estimates used as inputs to the models. The dimensions and reasons for the changes should be communicated clearly. The question was raised whether transmission from former key populations to their partners could be accounted for, but the consensus is that these estimates should be used to advocate for current key populations and dynamics among former key populations are of lower priority.

c. Data input quality

KP population size estimates assumed in EPP-Concentrated and AEM models are consensus estimates entered by national teams at the workshops. For countries where the Goals model is now used, PSEs are taken from GAM or from the KP workbooks. As mentioned above, changes in PSEs drive changes at the regional level and therefore the best possible estimates should be used.

For prevalence data that drive incidence estimates, Eline Korenromp raised the concern that the EPP model used for many concentrated epidemics is heavily influenced by high prevalence in early years datapoints (probably biased by oversampling high risk areas) which implies lower incidence in recent years to maintain high levels of prevalence. Other models fit for the same countries (where available, CSAVR or Goals) show trends of more gradual prevalence increase and so higher current incidence.

d. Assumptions about infections among non-KP partners of key populations

Several comments to the working paper and during the discussion were concerns about the assumed ratios to calculate new infections among partners of KPs. None of these ratios are informed by output from time-dynamic models. Modellers on the April call agreed that their model outputs of infections from KPs to their partners and the corresponding ratios could be made available to compare to the assumed ratios (e.g., Goals, AEM, Thembisa, Optima). UNAIDS planned a systematic review of data on number and gender of partners of KPs, condom use and proportions of MSM who are married to women which may help to inform models going forward.

Appendix B: Recommendations Sessions 2-5

Recommendation	Lead person(s)	Timeline
Session 2: Population Size Estimation, HIV prevalence and ART coverage synthesis methods (chaired by Leigh Johnson) **Objective:** • **Recommend approach and process for national HIV estimates team users to synthesise and extrapolate key population survey data for population size, HIV prevalence and ART coverage estimates.**		
KP data quality assessment, guidance on inclusion/exclusion and data synthesis • Provide a 'decision tree' to guide users to appropriate approach depending on data availability and other factors. • Where applicable, recommend the use of the *Triangulator* by country teams, but allow for other approaches depending on data availability (e.g., unweighted/weighted average across data points). • The *Triangulator* requires users to specify uncertainty in inputted KP survey data: For PSE, a confidence interval or range and a score between 0 and 100 reflecting their confidence in the design of the study and therefore accuracy of the estimate. • Develop guidance for use of the tool, e.g., a study may be poorly designed to estimate population size (low confidence score), but give reliable HIV prevalence estimate (high confidence score)	Working Group	Feedback October 2023
Weighting older vs more recent KP data • For estimation purposes, move away from the 'sunsetting' approach of disregarding data older than 5 years, given evidence that proportional population sizes do not change quickly over time, and expected larger variation across countries (arguing against cross-country imputations to countries without recent data) than within countries. • Develop analytical approaches that improve comparability of old and new data sources (such as modelling population proportions instead of count sizes) and adjusts for any changes in measurement methods over time.	TBD	
Changes in PSE over time • Perform multi-country analysis of time trends in KP PSE proportions, considering whether changes over time represent true changes in PSE sizes versus changes in survey methods and coverage. • Determine whether to extend the Triangulator to include a time trend variable. • From multi-country analysis, determine typical or reasonable changes in PSE proportion over time, to inform a Bayesian approach to estimating a national PSE trend from available country data. Users can change Bayesian *priors* if they have good reason to, but default *priors* are needed to prevent heterogeneity being incorrectly interpreted as trend.	TBD	
Urban/rural PSE assumption Review urban versus rural PSE from the following sources: • Key population surveys that included rural location component: CRANE survey in Uganda (for FSW) and CeSHHAR data in Zimbabwe (for FSW), Malawi IBBS	Oli Stevens, Imperial College London and expert group	Feedback October 2023

Recommendation	Lead person(s)	Timeline
• Household surveys on men reporting male sex partners for relative MSM proportions • PHIA surveys that included network scale-up population size estimates (Nigeria, Cote d'Ivoire, Haiti) • Household surveys that included questions on selling and clients of FSW (also see Hodgins et al. PLOS Medicine) • Review how urban/rural is defined and consider moving away from binary urban/rural distinction, bearing in mind differences in population density and mobility between areas.		
KP programme data • Add a worksheet to the KP workbooks for users to record KP programme data (number enrolled, numbers tested HIV positive, numbers on ART, new diagnoses, VLS, etc.) and compare these with population/survey-based PSE • Develop guidance on interpreting comparisons of KP programme data with extrapolated KP estimates, considering that programme implementation is heterogenous across locations (e.g., less provision in rural areas), programme data are often not linked or anonymised and therefore the same individuals may be represented in multiple programmes, and KP members may access HIV services outside KP-focused services.	Working Group	Feedback October 2023
Refine conceptualisation of risk heterogeneity in MSM • Conceptualise model representation reflecting variation in risk among MSM (recent versus ever had sex with other men), as this affects average time spent in high-risk MSM population group. Avoid referring to the term "turnover" among MSM to avoid interpretation of changing MSM identity. • When interpreting or extrapolating MSM survey data, explicitly articulate assumptions about extrapolating from MSM survey data that disproportionately represent younger MSM. • Consider results from age-structured MSM models to guide decisions about modelling MSM.	Working Group	Feedback October 2023
Recommendations for the KP workbook: • Document assumptions and steps to define catchment areas and denominators for PSEs. • Add a field to enter study-specific or observation-specific information about data quality to preserve institutional memory about survey implementation and interpretation. This could include info on sampling/representativeness (e.g., brothel-based vs. street-based sex workers). • Decouple the timing of KP workbook completion from the HIV estimates cycle, so that national estimates team users are under less time pressure for producing estimates and targets when completing workbook process. • Convene dedicated meetings involving relevant national experts in KP surveillance and programmes in each country to complete the KP workbook exercise, rather than aiming to convene all national KP experts at the HIV estimates workshops.	Imperial College, Avenir Health, UNAIDS	Feedback October 2023
Session 3: Dynamical modelling of HIV trends in KPs in SSA (chaired by Cari van Schalkwyk) **Objective:**		

Recommendation	Lead person(s)	Timeline
• **Recommendation on a proposed approach and development process to integrate key population model in UNAIDS estimates process.**		
Recommend not further developing EPP-ASM-KP for the purpose of estimating key population indicators in SSA because the model does not represent transmission dynamics between population groups, and therefore is not capable to produce counterfactual indicators. Continue developing EPP-ASM-KP for application in concentrated epidemic settings that use EPP (which is not age structured) and consider pathway for integration of EPP-ASM-KP with CSAVR-KP.	Avenir Health, Working group	Meet in July to review

Develop Goals-ARM for estimating key population HIV indicator trends and transmission dynamics in SSA
 	
Convene a working group to provide guidance on development and review of Goals-ARM for estimating key population indicators, including:
• Structural decisions about the model, such as population group definitions, mixing, risk heterogeneity, and population recruitment and replacement
• Validation of model structure and assumptions, for example
  ○ Using simulated datasets to see if simple and more complex models can capture ground truths of the simulated data
  ○ Review existing models to see how adding/removing complexities affect model outputs.
	Avenir Health
	
	Working Group, TOR developed by UNAIDS	

Engage teams from 2-5 countries to develop and pilot a process to produce key population HIV estimates as part of the nationally led HIV estimates process. The process development and piloting should include: • Engaging the breadth of relevant experts and stakeholders in key population data and programmes, which extends beyond the groups currently represented in the UNAIDS national HIV estimates process and includes key population community members in model development and piloting. • Review and collation of relevant national data • Determine relevant counterfactual-based indicators (such as transmission population-attributable fraction) and ensure appropriate and non-stigmatising interpretation and communication of indicators.	Working Group, TOR developed by UNAIDS	
Session 5: Transmission-dynamic developments to EPP (chaired by Mathieu Maheu-Giroux) **Objectives:** • **Reach recommendation on proposed method for time-varying reduction in HIV transmission when on ART** • **Review default assumptions about LTFU from ART** • **Review Spectrum ART coverage estimates in high coverage settings** • **Reach recommendations on priorities for development to EPP**		
Reduction in HIV transmission when on ART (Omega) The global fixed value omega = 0.8 for the reduction in transmission per percentage increase in ART coverage did not account for increasing viral load suppression (VLS) among people on ART. During the 2023 estimates round, an update was implemented to calculate country-specific value for omega of VLS (among PLHIV on ART) minus 0.07. Avenir Health investigated the use of varying the calculated omega over time as a function of increasing VLS for SSA countries. The effect on new infections was negligibly small.		

Recommendation	Lead person(s)	Timeline
Recommend to not implement time-varying omega for SSA countries, but to explore its effect in other settings.	Avenir Health	October 2023
ART coverage (programme vs surveys and ANC-ART) In cases where ART programme data are discrepant with estimates of numbers on ART from household surveys, there are three options to address ART discrepancies in Spectrum: 1) adjust ART programme numbers down to align coverage with surveys; 2) calibrate HIV prevalence and PLHIV up (above survey-based prevalence estimates) to better match survey-based ART coverage (balancing fit against survey-based ART coverage and survey-based HIV prevalence); 3) increase total population size to increase PLHIV and decrease ART coverage for given ART numbers. Recommend the development of visualisations and guidance on choosing between these options, by including visualisation for: • ART coverage among pregnant women attending ANC vs. overall population ART coverage data. • Age-distribution in program-reported ART data vs in survey-based ART coverage • Plot incidence, prevalence and ART coverage, for the respective possible adjustment options, side-by-side Other recommendations following this presentation: • Investigate how alternative adjustment options affect estimates of incidence and prevalence across multiple countries. • Assess the potential impact of individuals interrupted treatment or registered at multiple locations on over-counting the number currently on treatment using individually linked electronic patient data implemented in Ghana.	Avenir Health Rob Glaubius, Jeff Eaton Ekow Wiah	October 2023
Calibrating to program data on ART by age in Belize Entering adult ART interruption, restoring defaults for initial ART allocation, and fitting IRRs by age to ART numbers by age improved Spectrum-estimated ART distribution by age and resulting PMTCT coverage. Belize's CSAVR incidence estimate trend was sensitive to how people are initiated on ART (by number in CD4 category, versus by mortality risk in CD4 category). Investigate if this is true for other CSAVR settings.	Avenir Health	October 2023

ART interruption in adults in Spectrum
 	
• Ask national HIV teams from each region who entered non-zero LTFU in their 2023 Spectrum estimate how they measured or calculated this fraction and examine factors that cause variations.
• Conduct global review of data to inform default treatment interruption rates for Spectrum, to replace the current default assumption of no ART interruption:
  o IeDEA, ART-CC cohorts
  o PEPFAR data
  o Research studies
  o If feasible, stratify by age, sex, time, time on ART.
• Refer to "treatment interruption" rate in place of "loss to follow-up" rate in Spectrum input editors and HIV estimates guidance materials.
	UNAIDS
	
	TOR by Secretariat
	
	Avenir Health
		October 2023
		
		October 2023
		
		October 2023

Recommendation	Lead person(s)	Timeline
• As shown in Belize example, ART interruption affects AIM's age distribution of PLHIV on ART. If ART interruption is underestimated, ART coverage and survival may be overestimated in older PLHIV and underestimated in younger PLHIV. ART interruption and ART allocation inputs should be reviewed, and revised if warranted, before artificially fixing misalignment with ART by age via IRR fitting.	UNAIDS	2024 estimates
Proposals towards accounting for transmission dynamics in EPP-ASM for population incidence trends in sub-Saharan Africa Proposed development investigations: • Review impacts of using national instead of subnational EPP structure (urban/rural stratification, regions) • Investigate potential transmission mechanisms into EPP-ASM. Five priorities in order of importance: 1. Time- and age- varying viral suppression 2. Variation in acquisition risk and transmission risk by age 3. Infectiousness by stage of infection, including primary infection 4. Circumcision coverage by age 5. Age mixing matrix. • Fit to age-specific HIV prevalence and ART coverage.	Kinh Nguyen, Jeff Eaton Avenir Health, Jeff Eaton	Meet to decide on priorities in July 2023

Appendix C: Working groups feedback for Session 1

WORKING GROUP	GROUP 1	GROUP 2	GROUP 3	GROUP 4	GROUP 5	GROUP 6
Key strengths of proposed methodology	• Move from regional to country-level estimates. • An explicit temporal component • Forcing discussions around unexplored topics					• Time-dynamic model • Easy to use; country ownership
Priority limitations of methodology	• Trends grounded in data points from a single point in time. This can be mitigated by doing empirical analyses to benchmark modelled output against external data. • Doesn't account for intersectionality. • While country level data are available, engagement with countries on the same is yet to take place.			• Clients of KP aren't separated from other partners. • GOALS is currently assuming low or zero turnover in most settings. This may lead to underestimation of HIV incidence in KPs. • No age-stratification in SW and MSMs in most models and surveys overrepresent younger KPs.		Assumption driven; may be difficult to model where data is lacking. Hence there may be a need to have a permissible range for some estimates.
PSE estimates recommendations		Endorsed the overall approach of PSE validation process from UNAIDS although better transparency is needed.				
Estimating infections in clients and			• Utilize dynamic models whenever possible.			

partners of key populations			• Transparency and clarity on methods used. • Have clients in separate groups from other partners of KP. • Turnover needs to be estimated for PWID as data is available. • Avoid sunsetting PSEs.			

Appendix D: Working groups feedback for Sessions 2 and 3

WORKING GROUPS	GROUP 1	GROUP 2	GROUP 3	GROUP 4	GROUP 5	GROUP 6	GROUP 7
KP data assessment and inclusion/exclusion; Define criteria and process to guide quality assessment in KP survey data and determine inclusion/exclusion/ weighting					• Use a Bayesian synthesis tool. • Inclusion/ exclusion criteria by method alone should not be used instead a decision tree should be created. • Use of programme data in Triangulator • Apply Triangulator in each location then extrapolate to	Quality assessment important, not universal/automated decisions	Should be country-led A criterion to ascertain how data was collected' sample selection, methodology, time frame, location, context, reference population, self-reported or lab-based testing

					national level instead of using Triangulator plus at the national level.		
KP data synthesis guidance; Define tools required for implementation by national HIV team AND Review criteria to ensure appropriate application and objective results					• Method should be easy to use, transparent, reproducible and allow synthesis of existing estimates in a statistically sound method. • Use the currently available Triangulator		
Urban/rural PSE assumption; Specify currently available data and analysis to improve 0.6 rural/urban population size proportion assumption	• Proxy gradient for Strata • Systematic review available data to adjust along relevant strata	Analyze existing PSE data in large/medium sized cities for existing gradients based on population size/density; apply existing gradients to rural areas	• Ghana has a central individual-level tracker, implementing partners will report HIV testing data to tracker; data may be available in July • Value in pooling data from household surveys but may be challenging to interpret urban/rural differences if	• Hodgkins 2021 review used rural:urban ratio of 0.7 for men who ever paid for sex. • The 0.6 for MSM difficult to justify • The rural/urban stratification is convenient as we have the denominators		• Trying to extrapolate to rural areas is questionable. • Need to interrogate studies in rural areas	

			there are differences in willingness to report behaviours. • Mapping studies; Zimbabwe, Kenya, Malawi				
KP programme data in workbooks • Define programme data indicators that should be captured. • Define metadata and contextual data that should be recorded. • How should these be interpreted compared to estimates?	• Access relevant programme data for catchment areas • Biometrics in Kenya for de-duplication • Conduct incidence studies with repeat testing in Kenya with biometrics • Programmes are urban hence low yield in rural areas	• Agree with use of program data to evaluate sensibility of PSE. • Should only be used contextually			• Can use existing reported programme data in the workbook; PSE, ART coverage. • However, aggregate data quality checks are needed. • Identify countries with high quality programme data and develop best approaches to integrating data.	• Appropriate KP-specific MER indicators aggregated by the relevant geographical area. • Need to approach indicators with caution	
Define the counterfactual indicator(s) that the new model should produce		• Agree that a focus on 5-year projects is best • 5-year impact on a 2-year intervention that acknowledges	• Different indicators for different questions • tPAF – Advocacy & broad program priorities		• Past impact/prevention fraction • tPAF over time but we need a communication strategy		

		KP programming and funding cycles.	• PF – impact and cost-effectiveness • Need for careful messaging around indicators				
Validating key population transmission model decisions • Identify one key high priority model structure decision based on earlier sessions. • Describe how to determine or validate the decision	• How does GOALS differentiate sexual transmission categories? • Trends- getting PSE right over time • Are we missing important internal factors that affect PSE e.g., rural to urban migration?					• Validation depends on how the model is set up • Could potentially use KP HIV by age as an input to validate • Some age-specific IBBS data could be left out then used to validate model?	• MSM age structure • Review existing age structured MSM model and determine how simplifications within these models affect the output • Create simulated data sets and establish whether simple models can capture the ground truths of the simulated data.

Appendix E: Participants

Name	Organisation
In-person	
Akim Lukwa	SACEMA
Andreas Jahn	MoH Malawi
Avi Hakim	CDC
Cari van Schalkwyk	SACEMA
Carl Corcoran	CDC
Deepa Jahagirdar	Avenir Health
Ekow Wiah	NAC Ghana
Eline Korenromp	UNAIDS
Faikah Bruce	SACEMA
Guy Mahiane	Avenir Health
Ian Wanyeki	UNAIDS
James Stannah	McGill University
Jeff Eaton	Imperial College London
John Ojo	Africa CDC
John Stover	Avenir Health
Josh Salomon	Stanford University
Joshua Kimani	Partners for Health and Development in Africa
Kate Rucinski	Johns Hopkins University
Keith Sabin	UNAIDS
Kennedy Kipkoech Mutai	Bristol University
Leigh Johnson	University of Cape Town
Maria Au	USAID
Mary Mahy	UNAIDS
Mathieu Maheu-Giroux	McGill University
Michelle Morrison	Gates Foundation
Oli Stevens	Imperial College London
Peter Vickerman	Bristol University
Phelister Abdalla	NSWP
Ray Shiraishi	CDC
Reshma Bhattacharjee	USAID
Rob Glaubius	Avenir Health
Romain Silhol	Imperial College London
Sharmistha Mishra	University of Toronto
Sidy Mokhtar Ndiaye	Enda
Stef Baral	Johns Hopkins University
Wade Ivy	CDC
William Miller	USAID

Virtual	
Debra ten Brink	Burnet Institute
Hmwe H. Kyu	IHME
Irum Zaidi	PEPFAR
Jerome Milimu	USAID/SA
Jesse Knight	University of Toronto
Kelsey Case	Imperial College London
Kính Nguyen	
Lauren Parmley	USAID/SA
Le Bao	Penn State University
Marie-Claude Boily	Imperial College London
Rebecca Anderson	Imperial College London
Rowan Martin-Hughes	Burnet Institute
Shona Dalal	WHO
Sonia Arias Garcia	UNAIDS
Tim Brown	East West Centre

Appendix F: Agenda

UNAIDS Reference Group on Estimates, Modelling, and Projections

Integrating key population HIV estimates into the national HIV estimates process for sub-Saharan Africa

16-19 May 2023

All times are GMT+2 (Stellenbosch, South Africa)

Tuesday 16 May:

Time	Duration (mins)	Topic	Presenter(s)/ Lead Discussant
9.00	25	Welcome and introductions	Mary Mahy
9.25	20	Meeting objectives	Jeff Eaton
Session 1: Estimation of new infections by Key Population and their partners (donuts): refined time-dynamic methods (chaired by Jeff Eaton) **Objective:** • **Review and feedback of proposed methods for UNAIDS Global AIDS Update 2023 report**			
9.45	5	Objectives of the session	Keith Sabin
9.50	25	Description of methods and provisional results and reasons for changes from previous results, including: • A summary of model or data sources used for each KP and country • Changes from March to May draft (implemented and planned)	Eline Korenromp
10.15	20	Discussion	
10.35	15	• Summarize magnitude of change by KP and each region to identify where to prioritise further interrogation. • Document for each case whether changes are driven by: (1) change in source for KP estimates, (2) changes in inputs. • Advantages, caveats, and further possible refinements.	Keith Sabin Eline Korenromp
10.50	15	BREAK	
11.05	35	Discussion	
11.40	20	Description of Goals transmission and program impact model	John Stover
12.00	10	HIV testing, treatment cascade, and HIV incidence in MSM in SSA	James Stannah
12.10	20	Comparison of empirical and model-based estimates of HIV incidence and the distribution of new infections in SSA	Oli Stevens
12.30	30	Discussion	
13.00	60	LUNCH	
14.00	20	Systematic review about number of non-KP partners of KPs	Jerry Jacobson
14.20	20	Model estimates of onward transmission from KP to clients and other partners	Romain Silhol
14.40	20	Results of meta-analysis of HIV prevalence among clients of FSW in SSA	Sharmistha Mishra/ Mathieu Maheu-Giroux
15.00	40	Discussion	
15.40	15	BREAK	
15.55	60	Discussants	Stef Baral Leigh Johnson

			Mathieu Maheu-Giroux Joshua Kimani
16.55	65	Discussion	
18.00		CLOSE	

Wednesday 17 May:

Time	Duration (mins)	Topic	Presenter(s)/ Lead Discussant
Session 1 continued			
9.00	60	Working groups	
10.00	60	Working groups feedback and discussion	
11.00	15	BREAK	
Session 2: Population Size Estimation, HIV prevalence and ART coverage synthesis methods (chaired by Leigh Johnson) **Objective:** • **Recommend approach and process for national HIV estimates team users to synthesise and extrapolate key population survey data for population size, HIV prevalence and ART coverage estimates**			
11.15	5	Need and use for Key Population size estimates: program target setting, burden (Spectrum) and impact estimation etc.	Keith Sabin
11.20	30	KP data collated in 2023 Spectrum round from Excel workbooks – sub-Saharan Africa	Oli Stevens
11.50	30	Key population data synthesis and extrapolation	Kate Rucinski
12.20	40	Discussion	
13.00	60	LUNCH	
14.00	20	Key population size estimates using The Triangulator	Carl Corcoran
14.20	20	Statistical methods for key population indicators in sub-Saharan Africa	Le Bao
14.40	10	Data synthesis methods to estimate HIV prevalence and ART coverage	Oli Stevens
14.50	10	Adjust MSM surveillance and survey prevalence data for age group, before using in AIM or Goals calibration	Oli Stevens
15.00	60	Discussion	
16.00	15	BREAK	
16.15	60	Working groups	
17.15		CLOSE	
17.45		Pick-up for group dinner	

Thursday 18 May:

Time	Duration (mins)	Topic	Presenter(s)/ Lead Discussant
Session 2 continued			
9.00	60	Working groups feedback and discussion	
Session 3: Dynamical modelling of HIV trends in KPs in SSA (chaired by Cari van Schalkwyk) **Objective:** • **Recommendation on a proposed approach and development process to integrate key population model in UNAIDS estimates process**			
10.00	15	• Technical requirements • Summary of conclusions of 2021 UNAIDS Reference Group meeting • Overview of Working Group process	Jeff Eaton
10.15	10	Discussion	
10.25	20	Proposal for key population estimates process and approach	Sharmistha Mishra
10.45	20	Discussion	
11.05	15	BREAK	
11.20	30	Goals age-risk model and other model approaches	Rob Glaubius
11.50	20	Discussion	
12.10	30	Indicators for quantifying the role of key populations in HIV epidemic control	Sharmistha Mishra John Stover
12.40	20	Discussion	
13.00	60	LUNCH	
14.00	60	Working groups	
15.00	60	Working groups feedback and discussion	
16.00	15	BREAK	
16.15	30	**Recommendations**	
Session 4: Review of 2023 Estimates (chaired by Josh Salomon) **Objective:** • **Review challenges arising during the 2023 estimates process**			
16.45	30	2023 Estimates review and challenges arising	Mary Mahy/ Eline Korenromp
17.15	45	Discussion	
18.00		CLOSE	

Friday 19 May:

Time	Duration (mins)	Topic	Presenter(s)/ Lead Discussant
Session 4: Transmission-dynamic developments to EPP (chaired by Mathieu Maheu-Giroux) **Objectives:** • **Reach recommendation on proposed method for time-varying reduction in HIV transmission when on ART** • **Review default assumptions about LTFU from ART** • **Review Spectrum ART coverage estimates in high coverage settings** • **Reach recommendations on priorities for development to EPP**			
9.00	15	Time-constant Omega values used in 2023 round country estimates	Eline Korenromp
9.15	30	Proposed method for time-varying reduction in HIV transmission when on ART (Omega)	John Stover
9.45	30	Discussion	
10.15	20	ART coverage based on programme-reported numbers versus household surveys and ART in ANC	Jeff Eaton Rob Glaubius
10.35	10	Calibrating to program data on ART by age in Belize	Rob Glaubius
10.45	15	BREAK	
11.00	20	Review of LTFU from ART in the 2023 estimates round and implications	Rob Glaubius
11.20	60	Discussion	
12.20	70	LUNCH	
13.30	20	Accounting for transmission dynamics in EPP	Jeff Eaton Rob Glaubius John Stover
13.50	70	Discussion	
15.10	20	**Recommendations**	
15.30		CLOSE	