Technical Updates and Naomi model development

Report and recommendations from a meeting of the UNAIDS Reference Group on Estimates, Modelling, and Projections
Montreux, Switzerland - 8-10th October 2019

REPORT & RECOMMENDATIONS

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

IN MEMORY OF JACOB DEE

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

Oli Stevens, October 2019

Abbreviations

ADR	AIDS Data Repository
ANC	Antenatal clinic
ART	Antiretroviral therapy
CDC	US Centers for Disease Control and Prevention
CLHIV	Children living with HIV
EPP	Estimation and Projection Package
IeDEA	International Epidemiology Databases to Evaluate AIDS
PEPFAR	US President's Emergency Plan for AIDS Relief
PHIA	Population-based HIV Impact Assessment
PLHIV	People living with HIV
(P)MTCT	(Prevention of) Mother to Child Transmission
UNAIDS	Joint United Nations Programme on HIV/AIDS
WHO	World Health Organization

Background

UNAIDS Reference Group on Estimates, Modelling, and Projections

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

- 'Technical update' work streams: These work streams are oriented to conducting research and providing technical feedback and guidance on specific updates for the suite of tools used for annual UNAIDS estimates, i.e. Spectrum, which includes the AIDS Impact Module (AIM), the Estimation and Projection Package (EPP), and the Case Surveillance and Vital Registration tool (CSAVR).
- 'Thematic' meetings: These meetings are focused on convening new research to catalyse innovation on specific aspects of HIV estimates that require substantial conceptual or methodological development

Meeting Objectives

The purpose of this meeting was to provide technical recommendations for updates for Spectrum and accompanying estimation tools, used by countries to furnish annual HIV estimates.

Objectives of this meeting were to:

- Review Naomi model development since the May 2019 UNAIDS Reference Group meeting and make recommendations for its implementation in the 2019/20 estimates round.
- Plan model development and implementation for Spectrum, EPP, CSAVR and AEM for usage in the 2019/20 and 2020/21 estimates rounds
- Identify and coordinate promising research directions for further longer-term development.

Outline

The UNAIDS Reference Group held its thematic meeting on *Technical Updates and Naomi model development* in Montreux, Switzerland from 8-10th October 2019. The meeting featured presentations and group discussion to generate consensus recommendations. The programme was divided into the following sessions:

1. Spectrum updates
2. EPP development and implementation
3. Treatment cascade estimation
4. Naomi overview and data inputs

5. Naomi development, data, and model testing
6. Implementation and scaleup
7. Interaction of Spectrum and Naomi
8. Asia Pacific and AIDS Epidemic Model review
9. CSAVR model development
10. Projecting estimates to 2020, 2025, 2030

This report presents a summary of the meeting presentations and discussions. The presentations are available to meeting participants at www.epidem.org (others, please contact the Secretariat). The final recommendations can be found at the end of this report. The recommendations drafted at these meetings provide UNAIDS with guidance on generating HIV estimates, provide an opportunity to review current approaches, and help to identify the data needed to further improve the estimates. Previous meeting reports are available at www.epidem.org. This transparent process aims to allow the statistics and reports published by UNAIDS and partners to be informed by impartial, scientific peer-review.

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

Session 1: Spectrum updates

1.1 Spectrum code and model interfaces

Code for tracking several key adult dynamics—aging, new infections, progression through CD4 categories, mortality on and off ART—is presently replicated in several places within the Spectrum software. Changes to AIM need to be manually replicated in the EPP, shiny90, and CSAVR models. Simplifying these codebases into a single library to be called by all models would obviate the need to rewrite code and could be released into the public domain, stimulating wider use and faster innovation. Further consideration needs to be given to the age disaggregation requirements of each model and the role of the paediatric model within the library.

Consolidating codebases could extend beyond harmonising those used for computation, to include input data parsing and the broader agenda of harmonising the many interfaces for existing HIV modelling tools. Aligning the unification of model interfaces and codebases with the development of a new unified model which reflects aspects of both generalised and concentrated epidemics is pragmatic. The Reference Group agreed to revisit harmonisation of code bases following the 2019/20 estimates round.

1.2 Changes to Shiny90 workflow

Shiny90, a model for estimating knowledge of status, was used in the 2018/19 estimates round. Users uploaded a partially completed Spectrum file to a standalone web interface, followed the model workflow, and downloaded model outputs for re-upload into Spectrum. It was noted, however, that users often did not redo Shiny90 following modification of their Spectrum file, despite invalidation of the results. Further, unlike EPP, Spectrum is not disabled during the use of the standalone web interface, permitting forking of file versions. Shiny90 is now launched directly from Spectrum, and any edits to programme data invalidates the uncertainty analysis, indicating to the user that the Spectrum file is no longer finalised. Additionally, in the 2020 version of Spectrum, the Shiny90 output file uploaded to Spectrum will be saved inside the Spectrum PJNZ file, such that the HIV testing survey and programme data and Shiny90 results are documented, reproducible, and preserved for future use similar to other data inputs and outputs from Spectrum.

1.3 Spectrum ART inputs and outputs

A substantial number of PLHIV not on ART may have previously been on ART, and may have differential mortality and initiation rates with respect to truly treatment naïve individuals. A new display in Spectrum visualises the proportion of previously treated populations, with user-controllable dropout, mortality and re-initiation rates. The previously treated population is not used in model fitting and is calculated post hoc.

Number newly initiating ART (intended to be treatment naïve individuals) can now be entered in the Spectrum programme data editor, but countries have limited capacity to

account for treatment experienced individuals within initiators. Input fields should be revised to allow countries to enter “First time” and “Total” initiators – if disaggregation is not possible,

“Total” alone can be filled. At present, treatment naïve initiator data are not used in model fitting, but can be used to calculate dropout rates and/or validate the percentage dropout rate supplied by countries. Countries should be encouraged to enter all sources of dropout data at their disposal - discrepancies between dropout rates calculated from programme numbers and directly entered proportions are likely to arise, and guidance should be appropriately provided.

Bar charts visualising birth estimates alongside programmatic data in a single year were included in Spectrum for the 2018/19 estimates round and received positive feedback during workshops. In addition to the existing single year displays, trends over time and a larger number of PMTCT related indicators will now be displayed. To reduce figure complexity, users will be able to select and view individual indicators over time. A wider discussion followed on the underutilisation of the *Validation* menu, currently positioned after *Results* in the Spectrum workflow, and the utility of displaying validation charts earlier in the programme data entry screens.

A small number of countries have ART data by 5-year age groups. Where data are available, Spectrum estimates of ART age distribution match programme data closely. Countries should be encouraged to enter age-disaggregated treatment data for validation, though it is not yet recommended to be included in model calibration.

1.4 Changes to definition for reporting current on ART

A topic raised during discussion of Spectrum ART inputs was changes in the definition used to report the number currently receiving ART. PEPFAR programmatic data is beginning to record lost from, and returning to, treatment with a new threshold of 28 days since last contact, down from the existing definition of 90 days. Numbers on treatment will be impacted by this change in threshold definition, and countries with large changes will require support to adjust current year numbers on treatment, and to consider retrospective smoothing. There is expected to be little impact on incidence or mortality in response to the numbers on ART. As a key epidemic indicator, countries will also require communications support regarding flatlining treatment numbers in the approach to the 2020 targets.

1.5 UN World Population Prospects 2019 demographic inputs

The demographic projection module within Spectrum, DemProj, sources inputs from the UN Population Division’s (UNPD) World Population Prospects (WPP), updated biennially. Key changes in the use of WPP 2019 over WPP 2017 in DemProj include:

- The availability of migration data disaggregated by sex and 5 year age group, permitting calculation of net migration by single year of age;
- Calculation of death rates from single age-disaggregated survivor numbers, rather than from central death rates, to eliminate discrepancies between life tables calculated in DemProj and WPP; and
- An expanded list of countries for which WPP explicitly accounts for HIV

- For countries with HIV prevalence >4%, HIV effects included in age-specific mortality rate (21 countries)
- For countries with HIV prevalence >1%, life expectancy projections account for HIV prevalence and ART (58 countries, including the 21 countries above)

The national populations of some countries have seen large changes in the 2019 WPP update (some in excess of 10%), and webinars between UNAIDS, UNPD, HIV estimates teams and National Statistics Offices will be held to review these changes.

1.6 HIV natural history assumptions

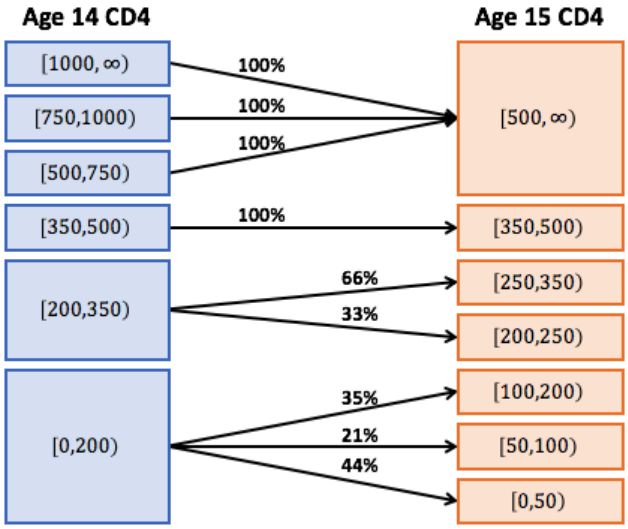
Rob Glaubius presented potential updates to the natural history models within Spectrum:

- smoothing discontinuities in the paediatric and adult models
- a full revision of the adult natural history model.

CD4 counts in the paediatric model are disaggregated into different categories to those in the adult model. As children age into the adult model at age 15, they are proportionally allocated to the adult categories, which can result in a substantial increase in HIV mortality, particularly at low CD4 counts (Fig 1).

Age 14 HIV mortality rates

CD4 category	Off ART	On ART ≥12m
[1000,∞)	0.026	0.002
[750,1000)	0.032	0.004
[500,750)	0.039	0.004
[350,500)	0.048	0.005
[200,350)	0.058	0.005
[0,200)	0.071	0.010



Age 15 HIV mortality rates

CD4 category	Off ART	On ART ≥12m
[500,∞)	0.005	0.005
[350,500)	0.011	0.011
[250,350)	0.026	0.016
[200,250)	0.061	0.016
[100,200)	0.139	0.021
[50,100)	0.321	0.035
[0,50)	0.737	0.055

Figure 1: CD4 category transition between Spectrum paediatric and adult models, with exemplar change in HIV mortality rates from Botswana

Adults similarly experience HIV mortality discontinuities at 10-year age group boundaries, seen in kinks in AIDS deaths by age. Mortality rates can be smoothed in the adult model without changing overall mortality by interpolating cumulative mortality risk. This somewhat smooths AIDS deaths, but discontinuity remains as progression parameters remain estimated by age group rather than single year of age. Neither changes to paediatric nor adult mortality rates are recommended for this year's updates, but are to be revisited in 2020. It is recommended that on-ART mortality rates are to be re-estimated directly from IeDEA data by single year of age, rather than smoothing existing rates for the 2020/21 estimates round. A further discussion of smoothing paediatric mortality rates can be found in the UNAIDS Reference Group 2019 Paediatric Meeting report.

Comparison of national Spectrum results about the CD4 distribution of untreated adults with PHIA survey data indicate Spectrum underestimates the proportion untreated PLHIV with CD4 <200 cells/µL and overestimates the proportion with CD4 >500 cells/µL compared to PHIA survey data. Natural history parameters—the distribution of CD4 at seroconversion, the rate of HIV progression between CD4 category, mortality rates by CD4 category in the absence of ART—were re-estimated better to fit the CD4 distribution in PHIA surveys while also retaining overall consistency with data about HIV survival by age at infection.

Continuous models, stratified by ages 15-24, 25-34, 35-44, 45+, were fit to data about:

- CD4 category at seroconversion (data from CASCADE Collaboration)
- All-cause mortality rates by CD4 category (data from CASCADE Collaboration)
- Survival after seroconversion with untreated HIV (data from ALPHA network)
- CD4 categories among untreated PLHIV (data from PHIA surveys)

Parameter estimates fitted to the above data, when compared to Spectrum defaults, have:

- More PLHIV with lower CD4 counts at seroconversion;
- Faster CD4 progression, except for at ages 25+ with CD4<200;
- Lower HIV mortality rates;
- Mortality concentrated in lower CD4 categories; and
- ART allocation weight that prioritises treatment eligibility over expected mortality.

When comparing fitted parameter estimates to Spectrum defaults in 13 high burden countries (Fig. 2):

- Large increase in adult AIDS deaths throughout the epidemic (+30% in 2018).
- Decreases in adults living with HIV and new adult infections (-5.9% and -5.8% in 2018).
- Historical increases in child AIDS deaths, CLHIV, and new paediatric infections, minimal differences in 2018.

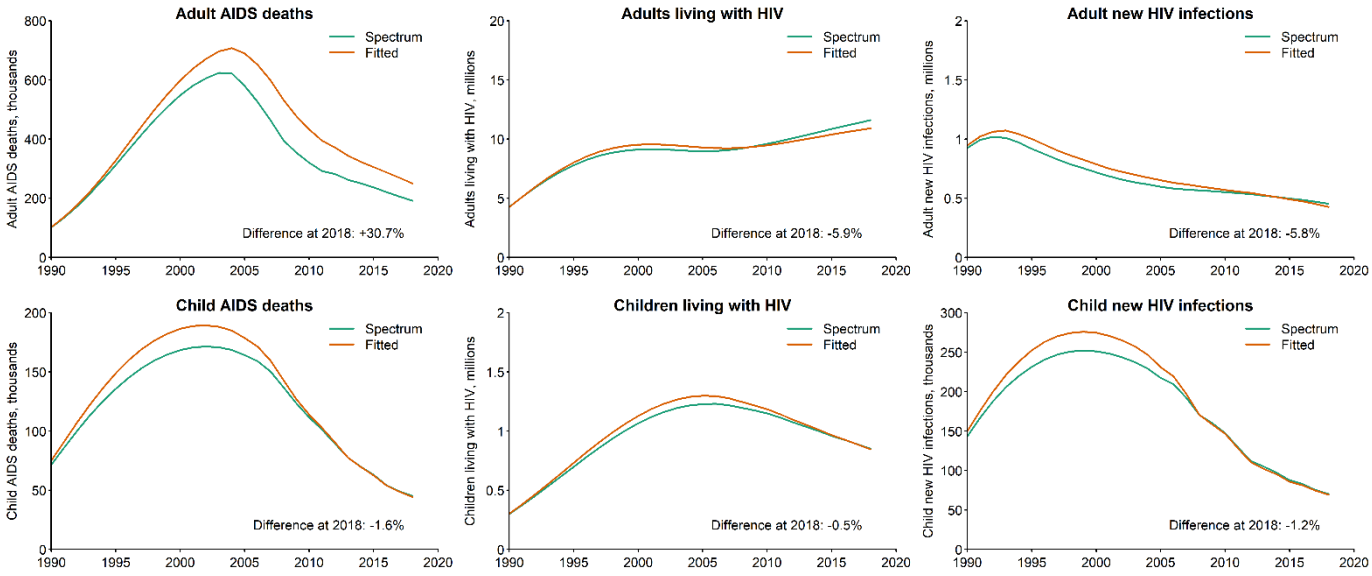


Figure 2: Comparison of adult and paediatric AIDS deaths, individuals living with HIV and new infections between existing Spectrum defaults and revised natural history parameters

Though good fits are obtained to calibration data, fit to validating data (CD4 count at initiation from the IeDEA network) is poor - fitted estimates over estimate the share of ART initiators at low CD4 counts and under estimate at high CD4 counts. This appears irreconcilable within the current modelling framework with PHIA data: PHIA data suggest more untreated PLHIV at low CD4 counts, whilst IeDEA data suggest the opposite. When the IeDEA data are included as calibration data, the model fails to converge. Due to the substantial change in AIDS deaths, the fitted parameter set is not recommended for adoption until a parameter set that reflects all data sources can be found. It is noted that viral load is not included in the model: differentiation between high and low viral load within a single CD4 category may permit better model fits.

Session 2: EPP development and implementation

2. EPP development for use in generalised epidemic settings

2.1 Fertility rate ratio adjustment

At present, Spectrum passes EPP the FRR local adjustment factor as fit to ANC-RT data, which is used in EPP-ASM to calculate prevalence in pregnant women. A probit offset to the routine data had been used so that the prevalence curve for pregnant women passed through the ANC-RT data. A further improvement has since been recommended. EPP is passed the FRR local adjustment factor by Spectrum, and EPP fits a FRR scaling factor to adjust for EPP subpopulation deviation from Spectrum's national or provincial factor.

2.2 Fitting to age/sex structure prevalence data

EPP fits to 15-49 prevalence data and passes a single 15-49 incidence and prevalence for each year to Spectrum, whereupon incidence rate ratios, fit to age/sex prevalence data, allocate the new infections. When fitting to age/sex structured prevalence data directly, EPP will pass age/sex specific incidence and prevalence, and fitted IRRs to Spectrum.

Due to the additional data, the associated SPU files are large (100-300MB), and it is proposed that:

- EPP only writes the SPT file during normal operation
- Spectrum calls EPP during the uncertainty analysis for it to write the SPU file
- The SPU file contains 300 samples, rather than the current 3000 samples
- The SPU file is destroyed after the uncertainty analysis is complete and not included in the final PJNZ.

The required development for EPP to fit to age/sex structured data, including updating the surveys database, and for Spectrum to accept age/sex structured outputs, is substantial, and will be implemented for the 2020/21 estimates round.

2.3 Joint modelling of prevalence and ART coverage in EPP

ART coverage is presently estimated by Spectrum using treatment numbers as a numerator with PLHIV estimate as a denominator. Recent household surveys, however, furnish direct survey-based estimates of ART coverage, which can be incorporated into EPP fitting. Fitting to both prevalence and ART coverage from survey provides a small precision increase in both estimates, but can produce dramatic changes to HIV prevalence and incidence if survey estimates of ART coverage are disparate from those calculated from programme data. HIV prevalence and incidence are affected as, in order to furnish different estimates of ART coverage, the number of PLHIV must change as numbers on treatment remain unchanged. This effect is pronounced in subnational regions with individual Spectrum files between which there is significant cross-region ART attendance, as shown in Fig 3, where PLHIV in Maputo Province seek care in Maputo City. The programme numbers reflect all those seeking treatment in the region, rather than those only residing in the region, whilst the survey measures resident ART coverage, and the model must markedly increase or decrease HIV prevalence and incidence to reconcile the data sources.

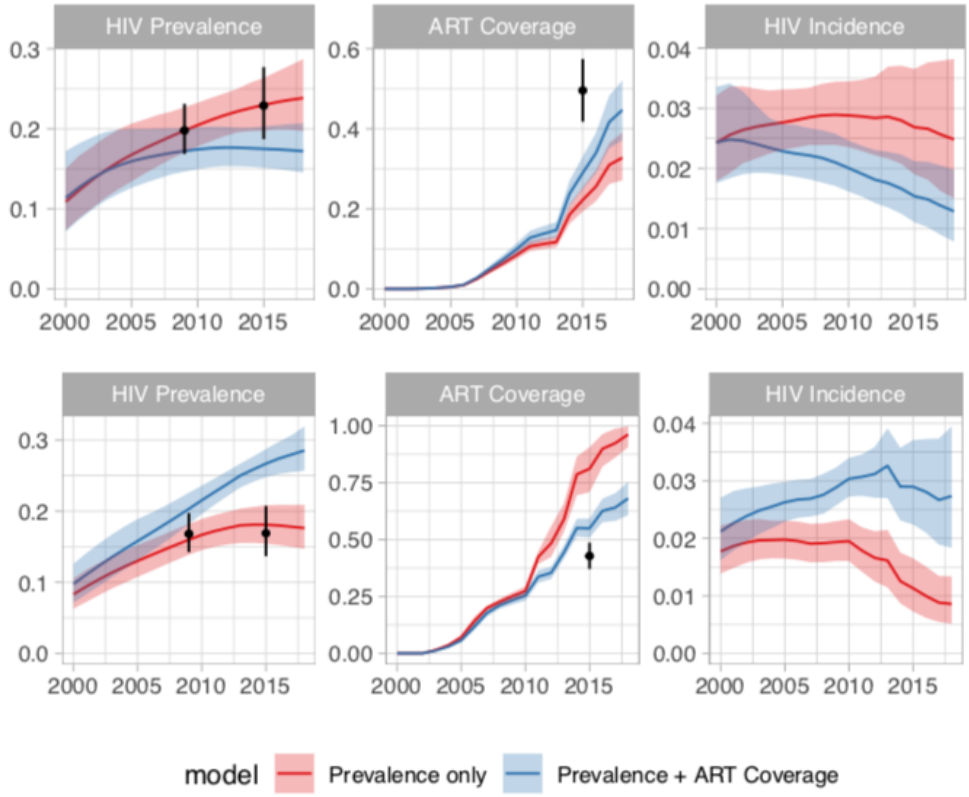


Figure 3: Disparity between ART coverage estimates from survey and from programme data due to cross-province care seeking. Upper panels: Maputo Province. Lower panels: Maputo City

Joint modelling in national Spectrum files with subnational EPP regions, where national programme numbers are allocated proportionally by prevalence and population (i.e. should better reflect survey estimates of ART coverage), is recommended for further testing. The use of joint modelling in subnational Spectrum files is not recommended.

2.4 Adjustment for ANC-RT bias

As the use of routine antenatal HIV testing (ANC-RT) data within HIV estimation and inference increases, there is a need to analyse data completeness, quality, and testing coverage. The dominant presentation of invalid data within PEPFAR ANC-RT data is HIV testing coverages in excess of 100% (number of known statuses exceeding number of clients). The magnitude of retesting is difficult to correct for – systematic differences in retesting behaviour between HIV- and HIV+ clients – and difficult to detect if testing coverage remains under 100%. A model has been developed to adjust for overestimation of prevalence at low testing coverages, and is offered within an R package, *ANCRTAdjust*, created to visualise ANC-RT data and offer adjustment options for invalid or missing data.

2.5 EPP in concentrated epidemics

Countries with concentrated HIV epidemics that use EPP without survey data rely heavily on data from antenatal care HIV testing. ANC-RT data are currently underutilised in the estimation of epidemic level. Presently, ANC-SS data are used as the estimator of epidemic level in EPP, with the probit offset adjusting the ANC-RT data to fit sentinel surveillance

data. However, ANC-SS are typically from a non-representative selection of sentinel sites. In settings with high ANC testing coverage, the HIV prevalence observed among routine ANC testing may be interpreted as representative of prevalence among all pregnant women. Consequently, it is recommended that ANC-RT data is used to inform the overall level of the epidemic, with an offset for prevalence in ANC-SS sites. HIV estimates teams should be mindful of the quality of ANC-RT data, including re-testing and representativeness. The review of HIV prevalence amongst ANC attendees relative to population prevalence in concentrated epidemic settings should be updated.

Tobi Saidel addressed key population inputs to EPP, outlining existing challenges with the workflow. Key population size estimates can be entered by estimates teams as proportions or absolute numbers. EPP does not currently record which format countries provide, and the introduction of a flag for size estimate format should be considered. Some countries have key population size estimates considerably below “regional norms” and it was suggested that these countries are entering local, non-representative absolute numbers, which, when converted to proportion of total population, are very small. Further guidance is needed for countries on the age and sex definitions of denominators used by EPP for key population size estimates, and countries should be encouraged to document the source or methodology used to derive all estimates. The Reference Group will review how Spectrum allocates incident infections by age in key populations, and consider whether key population estimates should be outputted using user entered 15-49 size estimates as denominators.

In concentrated epidemics, sex ratios of infection in Spectrum are derived from EPP. When modifying key population turnover and calibration, sex ratios in Spectrum change, and are then discordant with those in EPP, and this is should be made clearer to users. Spectrum could block subpopulation prevalence results when EPP and Spectrum sex ratios are discordant. The sex ratio of infection can also be used as a data quality indicator—concentrated epidemics are expected to have more male infections than female—and countries with the inverse should be encouraged to interrogate their input data.

Session 3: Treatment cascade estimation

Kim Marsh presented an overview of the use of the shiny90 model in the 2018/19 estimates. The Shiny90 model uses survey and testing programme data to estimate knowledge of status in adults. All countries in Eastern and Southern Africa (ESA) and 17 of 23 countries in Western and Central Africa (WCA) produced shiny90 estimates. However only five countries supplemented survey data with additional HIV testing programmatic data. Compared to imputed first 90 estimates in 2018, estimates in ECA were similar and estimates in WCA were higher. For the 2019/20 estimates round, countries should be encouraged to bring HIV testing programme data to improve the shiny90 estimates.

Results from, and recent updates to, the shiny90 model were presented by Katia Giguère. Knowledge of status is lowest in WCA, men, and younger age groups. In addition to the indicators estimated during the 2018/19 estimates round, shiny90 now produces regional estimates of:

- Time to from HIV infection to diagnosis
- Probability of being diagnosed before HIV-related death

- Probability of being diagnosed within a certain time frame
- Probability of being diagnosed before reaching a specific CD4 threshold

Significant progress across all indicators and regions has been made since 2000, though regional heterogeneity remains, with worse outcomes in WCA compared to ESA. To maximise the utility of model results for programme planners, it is recommended that results are reported with reference to an individual infected or diagnosed today:

- Probability of being diagnosed within 6 or 12 months and CD4 count >350 and >200 cells/ml;
- A stacked bar showing proportions diagnosed within discrete time thresholds
- Simple statistics in current year e.g. 1 in *x* tests are positive.

The shiny90 web interface is unlikely to be updated before the upcoming estimates round to display the new outputs, and results will instead be made available in a PDF and CSV downloads.

WHO recommend using a viral load suppression (VLS) reporting threshold of 1000 copies/ml. Countries, however, routinely report viral load suppression data at different, lower, thresholds (see Estimating population viral load suppression), and no adjustment exists to transform all reported rates of VLS relative to a 1000 copies/ml threshold. Leigh Johnson presented a Weibull adjustment model calibrated to South African viral load data, which performed well with a phi parameter of 1.5 in adults (Fig 4). Outside the South African setting, the model performed acceptably in countries reporting VLS at a threshold of 400 copies/ml, but may underestimate VLS when applied to national thresholds of 50 copies/ml. The threshold adjustment in adults is recommended for this estimates round, and additional phi parameters will be estimated from IeDEA data which may better suit lower reporting thresholds.

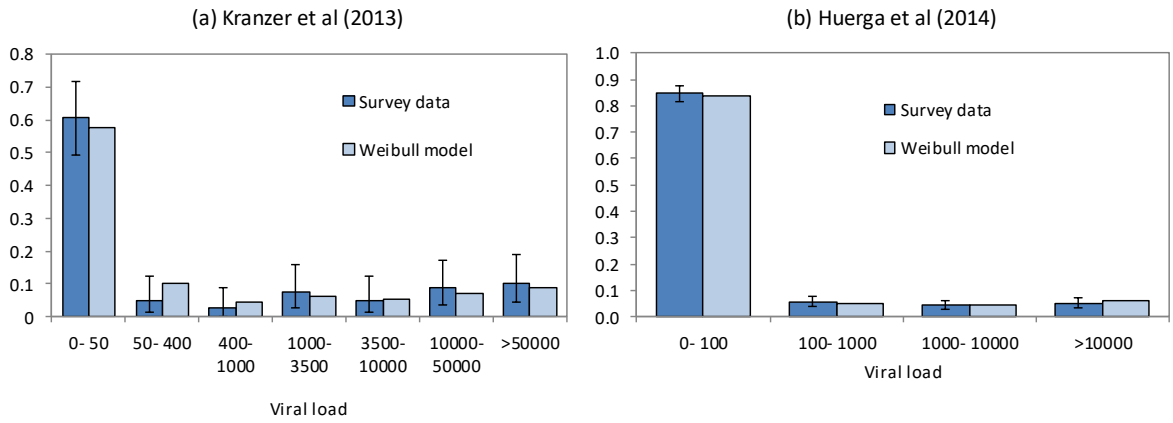


Figure 4: Weibull adjustment model calibrated to South African viral load data

A model for adjusting rates of VLS in South Africa for missing data, previously presented by Johnson, was applied to PEPFAR data 2017-2019 from 18 African countries. VLS was found to be positively correlated with testing coverage. However, due to concerns of unmeasured confounding within the model (e.g. quality of testing programmes), Johnson recommends against adjusting routine viral load data for incomplete data. Future approaches may look to

Bayesian approaches to appropriately represent uncertainty around relative rates of suppression in tested vs missing, untested, individuals.

The treatment cascade display in Spectrum shows the time trends in the three 90s. Additionally, Spectrum will now show each 90 with a detailed disaggregation. Knowledge of status will be disaggregated by incident and long-standing infection, both of which further disaggregated by on treatment, previously treated, never treated. On treatment will be disaggregated by newly initiated, previously treated, on treatment for over a year, and never treated. Viral suppression will be disaggregated by suppressed based on routine VL test, assumed viral suppressed, and not virally suppressed. A CSV download will be provided for rapid extract of the detailed cascade for programme managers.

Naomi model development

Naomi is a district-level HIV estimation model implementing the recommended approaches for further development at the Spring 2019 UNAIDS Reference Group meeting (see "The District Model"). The objectives of these sessions were to review model development and implementation since May 2019, and to make recommendations on the interaction between Spectrum at the national or subnational level with Naomi.

Session 4: Naomi overview and data inputs

Jeff Eaton provided an overview of the relational data structure created for Naomi. Naomi requires input data on population structures over time, and HIV prevalence and ART coverage from household surveys, treatment programmes, and antenatal care. The full data structure can be found here.

Several population products exist that offer pixel level populations that can be flexibly aggregated up to any given set of boundaries. Differences between population products when used as denominators can give rise to divergent HIV estimates, and remains an underappreciated source of uncertainty in the estimation process. Megan O'Driscoll presented a comparison of population products whose sources and methods are summarised in Table 1.

Source	GPW	Facebook HRSL	WorldPop	Landscan	U.S. Census Bureau
Regions	Global	Selection of countries	Global	Global	Selection of countries
Years	2000, 2005, 2010, 2015, 2020	2015	2000-2020	2000-2017	2000-2015
Spatial dis-aggregation	Pixel level	Pixel level	Pixel level	Pixel level	Subnational administrative units
Key data inputs	Harmonised subnational census and boundary data from CIESIN			Unknown	N/A
Spatial dis-aggregation methods	Minimally modelled. Uniform distribution population within administrative unit.	High resolution satellite imagery & machine learning techniques to identify settlements	Machine learning with census, survey and satellite datato estimate spatial distribution of populations.	Remote sensing imagery analysis and spatial modelling to estimate "ambient" spatial distribution of populations.	N/A

Table 1: Summary of subnational population products in sub-Saharan Africa

HRSL populations are similar, with small absolute difference, whilst LandScan differs (consistent pattern across several countries, see Rwanda as an example in Fig 5). However,

despite concordance in total population, substantial differences between WorldPop and GWP are seen in sex ratio and under 30 proportions (Fig 6), which is concerning. Difficulties arise when boundaries used in population products differ from those used in Naomi. Populations using raw or minimally modelled counts (USCB and GPW respectively) are problematic with area misalignment, whilst more heavily modelled sources (WorldPop and Facebook HRSL) are less likely to be influenced by boundary changes.

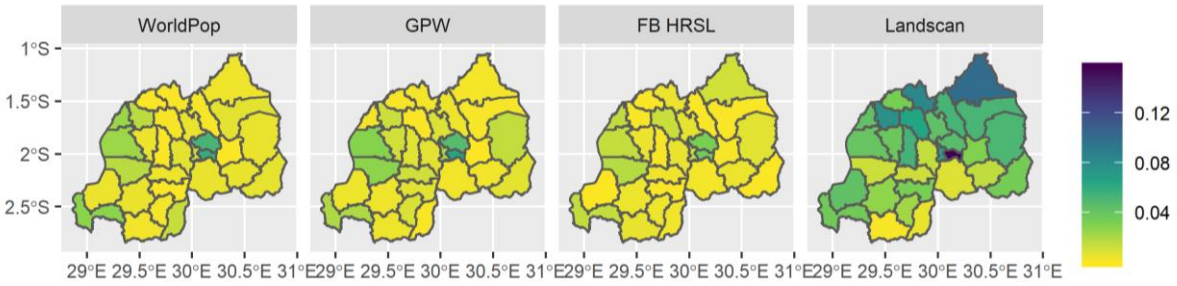
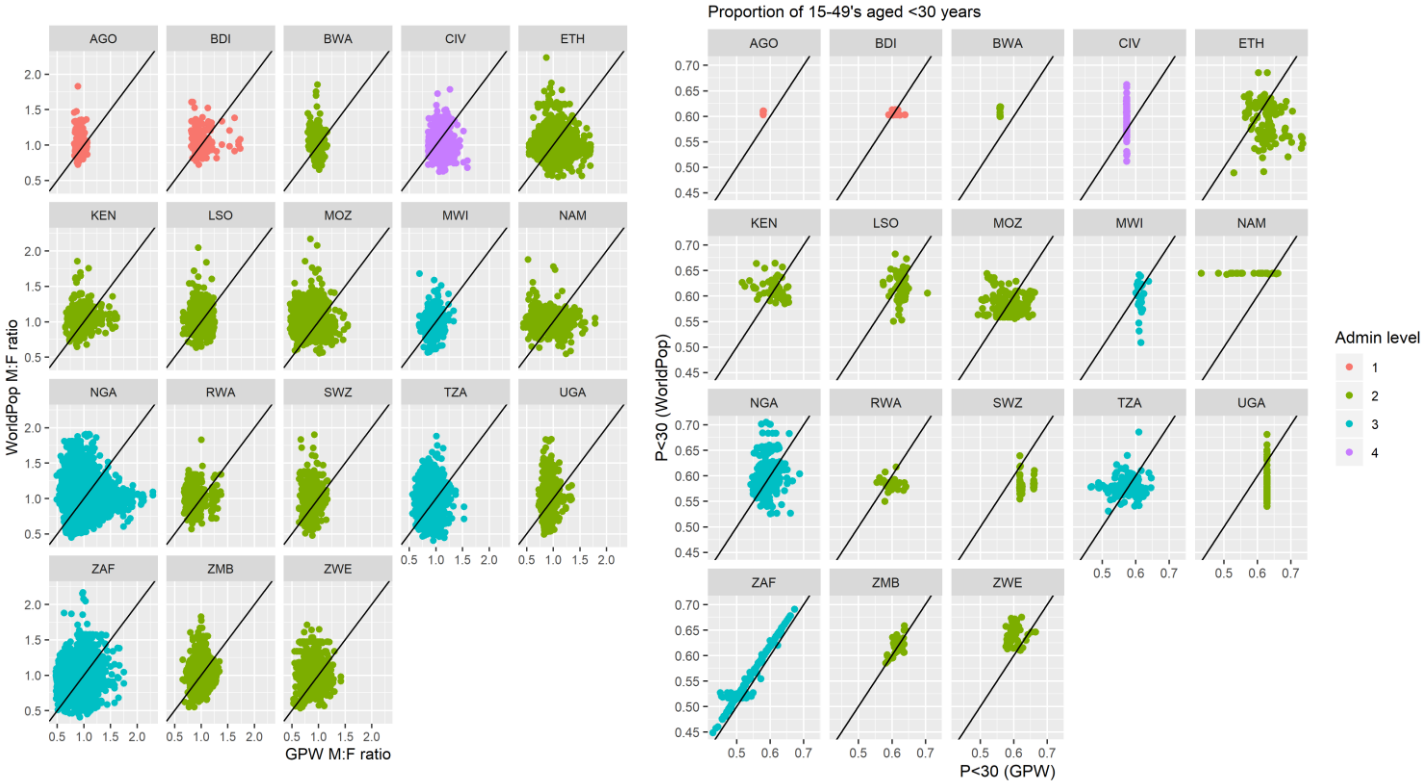


Figure 5 (above): Percent difference in total population in Rwanda, using US Census Bureau populations as reference.

Figure 6 (below): Comparison between WorldPop and Gridded Population of the World of 15-49 sex ratio (left), and proportion of 15-49 that is under 30 (right)



Evaluation and recommendations of a single population product remains difficult. Countries may look to use their most recent unmodelled census counts as denominators, and an enumeration of input sources will assist in showing that national censuses form the key data inputs to these products, and differ only in disaggregation methodology.

Age-specific fertility is required as in input to Naomi for district-level estimates of paediatric prevalence and births to HIV$^+$ women. A spatiotemporal model of district-level fertility from household survey data was presented by Oli Stevens. The model identifies substantial heterogeneity between districts, and identifies districts with an age pattern of fertility that differs from the national average. It is recommended that births to HIV$^+$ women is not included as a user-accessible output.

Session 5: Naomi development, data, and model testing

Jeff Eaton detailed Naomi model developments following the recommendations of the Spring 2019 UNAIDS Reference Group. Model changes are as follows:

- Spatial models on HIV prevalence and sex ratio of HIV prevalence.
- Autoregressive models on age patterns of prevalence and differences in age pattern of prevalence by sex.
- Re-implementation of model fitting in Stan with MCMC to posterior mode optimisation via Template Model Builder (TMB).
- Using survey VLS data for surveys without ART coverage.

Three versions of Naomi of increasing complexity are available, depending on data availability and robustness, and are summarised in Table 2.

	Model A - Enhanced Small Area Estimation	**Model B - Joint Model**	**Model C - Joint Model with ART attendance reallocation**
Data	- Survey HIV prevalence - Survey ART coverage - ANC testing prevalence - ANC ART coverage	- Survey HIV prevalence - Survey ART coverage - ANC testing prevalence - ANC ART coverage - ART programme numbers	- Survey HIV prevalence - Survey ART coverage - ANC testing prevalence - ANC ART coverage - ART programme numbers
Advantages	- Converges fast and reliably	- Converges fast and reliably - Reconciles survey and ART programme data	- Reconciles survey and ART programme data - Accounts for cross-district treatment seeking
Disadvantages	- Does not reconcile survey and ART programme data	- Assumes all ART clients are resident in district where they receive ART	- Fails to converge in the majority of settings

Table 2: Comparison of increasingly complex Naomi model variants

ART attendance reallocation within Model C currently fails to converge in the majority of settings, and it is recommended that more structure be imposed on the prior, potentially including urban/rural care seeking, or a gravity distance model. If Model C can be used successfully, it should return:

- Number on ART: unmodelled programme numbers; and
- Numbers receiving ART: reallocated numbers receiving ART

If Model C cannot be used, Model A is preferred over Model B as the issues summarised in Fig 3 of joint modelling in EPP are similarly applicable to Naomi. Model A should use the survey-based estimate of ART coverage and return:

- Number on ART: estimated ART coverage x prevalence x population
- Numbers receiving ART: unmodelled programme numbers

The deadline for a working testing version of the model is 1st November 2019, and if Model C is later made to work, it can be added without any impact on the backend API. Comparisons of Naomi results to 2018/19 HIVE results, age-specific prevalence estimates from Spectrum, and CDC SAE results will be prepared and circulated to members of the Reference Group.

Session 6: Naomi implementation and usage

The AIDS Data Repository (ADR) is a data repository to centralise, harmonise, and store country data used by HIV estimation tools, and will be debuted in the 2019/20 estimates round. To support the data needs of Naomi, the ADR will store and preview geographic and programme data. Recent development includes improving data validation with the addition of foreign key validation and issue tracking, and a 'harvester' to facilitate the extraction and processing of data from user-specified pivot tables in DHIS2. It is not required, though beneficial, for countries to use the ADR to use Naomi, therefore similar validation checks will be required within Naomi and ADR. The ADR may, in future, implement rolling validation, where users are permitted to upload unvalidated data, which is then flagged by the next validation cycle.

The user interface for Naomi was previewed by Rob Ashton. Countries will:

- Upload inputs including shapefiles, and population, programme, survey, and ANC data.
- Visualise, review, and validate input data.
- Select data to be used in model fitting.
- Fit the model.
- Visualise, review, and validate model outputs.

Naomi's backend architecture permits the model to run without locking the user interface, and flexibly scales with demand. This represents improvement over the Rshiny architecture on which the shiny90 model was built. Future development will include closer alignment with the ADR, including data validation and API integration, and the creation of a version history, so that users may run the model several times with differing model options and compare results.

Session 7: Interaction of Spectrum and Naomi

A discussion on the interaction of Spectrum with Naomi, both regarding the interface and raking of district-level results to national Spectrum values was led by John Stover. It was agreed that:

- Naomi should be launched from the Tools menu within Spectrum.
- Spectrum should display simple representations of Naomi model results, complex visualisations and validation of model results to be held within the Naomi web interface.
- Spectrum should not estimate deaths or PMTCT coverage at the district-level by disaggregating national level results.
- The PEPFAR Data Pack should be produced from Spectrum, not the Naomi web interface.
- Countries unable to run Naomi will use the Spectrum District Estimates tool, disaggregating national results based on district-level ANC prevalence.
- Spectrum should retain visualisation capacity for these non-Naomi countries.

Addressing the future of subnational Spectrum files and calibration of Naomi estimates to Spectrum, Jeff Eaton presented a comparison of Naomi results aggregated to provincial level with subnational Spectrum files. As expected, the results from Model A were closer to Spectrum results than Model B, as both Model A and Spectrum fit to prevalence only. Whilst raking Naomi results provides consistent estimates and time trends, Naomi fits to more data sources than Spectrum, and raking may be counterproductive. It was agreed that:

- Moving away from subnational Spectrum files is a substantial change and will require further data preparation and model estimates review work with countries.
- Subnational files should be retained for this estimates round, and UNAIDS engage countries on whether they are amenable to moving away from them in 2020/21.
- ART numbers should not be calibrated to Spectrum, but total numbers on ART should correspond to those in Spectrum.
- Mean total PLHIV should be calibrated to match Spectrum.
- Population denominators should be calibrated to Spectrum before model fitting.

Session 8: Asia Pacific and AIDS Epidemic Model review

The 23 countries in the Asia/Pacific region that produce HIV estimates with support from UNAIDS use either EPP with subpopulation or key populations, or the AIDS Epidemic Model (AEM). AEM is a behavioural, risk differentiated model, for use in key population driven epidemics, and primarily used by South East Asian nations. As a behavioural model, AEM requires more input data than EPP to inform its highly paramaterised structure, and some countries are unable to provide national data, instead using defaults from Thai data. Sabin raises five points for discussion:

- The HIV epidemic in the general population in AEM derives only from key population turnover.
- Paediatric estimates do not use ANC testing data, and PMTCT data are often limited.
- HIV prevalence in each subpopulation is aggregated into a single prevalence point per year.
- To produce a national Spectrum file, Spectrum mortality is adjusted to fit AIDS deaths from AEM in an opaque manner.
- The use of AEM typically relies on technical assistance from East West Center.

Tim Brown provided an overview of AEM's model structure, parameters, data inputs and workflow. AEM simulates HIV transmission in ages 15+ through a contact-based model, using a mechanistic compartmental structure (Fig 7). It requires data on:

- Biological data trends (e.g. HIV and STI prevalence);
- Sexual and injecting behavioural trends (e.g. condom use, duration of sex work, average client numbers for FSW); and
- Size estimates for key populations

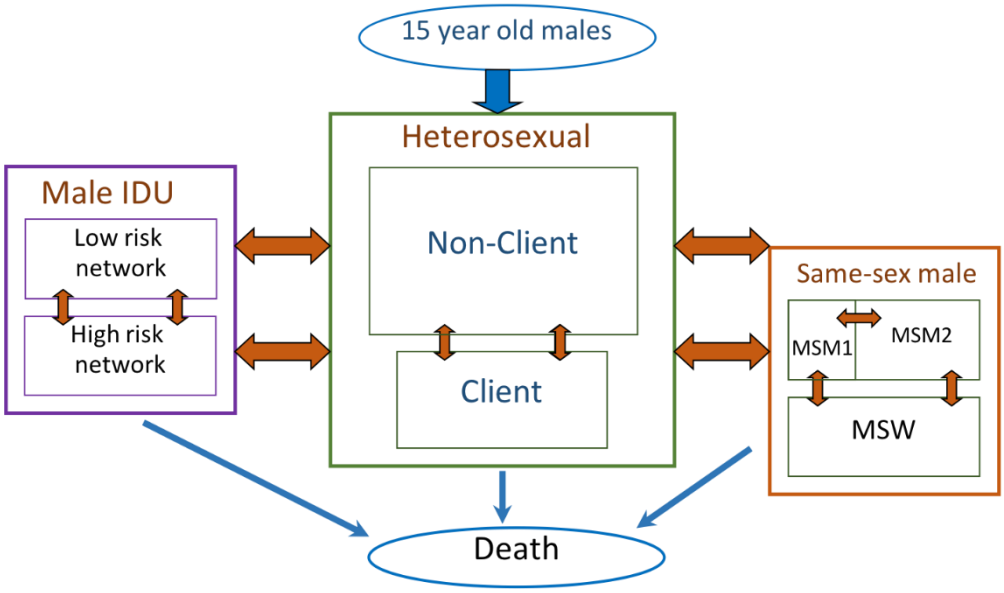


Figure 7: Risk compartments for males in the AIDS Epidemic Model

Data on these trends are developed through in-country consultation, consolidating all available surveillance data and published studies into a single prevalence point for each key population per year, noting that each individual data source is often non-representative of the wider subpopulation. Beyond the calculation of HIV estimates themselves, AEM also offers policy and planning analysis use to inform National Strategic Plans and investment cases for the Global Fund. Echoing the issues raised by Keith Sabin, Tim Brown further added that in-country consultations to ensure adequate data preparation, interrogation, and analysis are time intensive and exceed current capacity. AEM demands large quantities of data—insufficient documentation of inputs is common—and is being applied in countries with limited data about key model inputs.

Nash Montevirgin detailed the HIV surveillance systems in place in the Philippines, the use of case surveillance data in HIV estimates, and the validation of these data. Surveillance is broken down into two main sections—passive HIV/STI surveillance, and active case HIV surveillance—HARP (Fig 8). HARP is further distinguished by three sections: HIV diagnosis, with confirmatory centralised testing; monitoring the treatment cascade, and TB and pregnancy status; and mortality reporting. Of these, mortality is the weakest reporting system. In addition, IHBSS surveys are conducted every 2-3 years for MSM, PWID, FSW and incarcerated populations, through which size estimates are also generated.

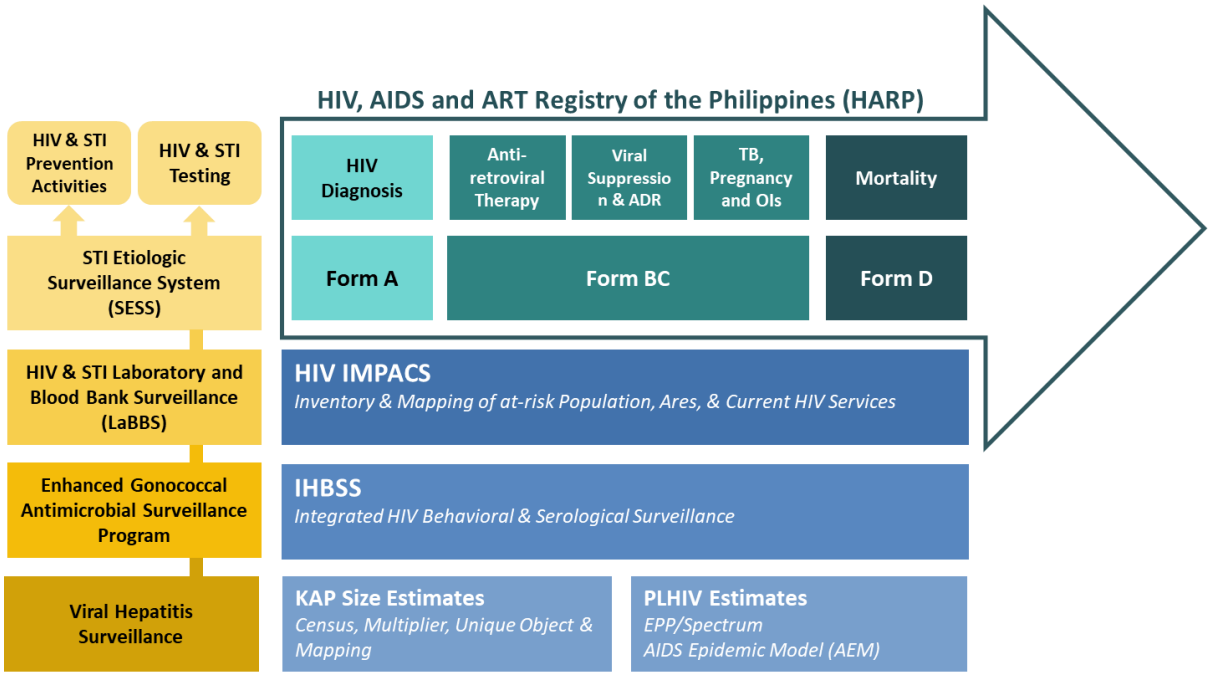


Figure 8: Summary of HIV/STI surveillance in the Philippines

Data from 150 treatment hubs are submitted through HARP are encoded, assessed for completeness, and deduplicated by fuzzing matching or by unique identifier code where available - a process to be simplified with the move to an online HARP system. Comparing surveillance data to estimates: number of PLHIV from HARP is around 30% lower than estimated PLHIV, number of reported AIDS deaths is around 60% lower than AEM estimates; and ART coverage can approach and exceed 100% when using HARP numerators and estimate denominators. Death registries are known to have large gaps and undercount AIDS deaths, and mortality reporting is an ongoing focus for Ministry health

system strengthening. At present, HARP data are used for validation only within AEM, rather than as calibration.

Subsequent discussion focused on two topics:

1) Spectrum-AEM mortality adjustment

At present, incidence calculated in AEM is imported into Spectrum, whereupon Spectrum prevalence and mortality do not match those calculated in AEM. On and off-ART mortality is iteratively adjusted until a match is found. Four recommendations were made in discussion:

- Review mortality adjustment in 2018/19 files.
- AEM estimates 15+, whereas Spectrum is age-structured. Asia/Pacific countries do not have household surveys with which to fit IRRs, and it is recommended that a new IRR fitting tool be developed. The tool will fit IRRs to any available age specific data – ART, case reports, or mortality data.
- AIDS death estimates should be validated against vital registration systems.
- Implement a prevalence/incidence adjustment as exists in EPP/Spectrum instead of adjusting mortality.

2) Paediatric estimation

There were concerns around the quality of paediatric estimates, which are derived from assumptions about HIV prevalence among pregnant women and rates of mother-to-child transmission. The primary concern was that little data exist about HIV+ non-FSW pregnant women—surveillance focuses on key populations. Recommendations included to:

- Validate paediatric estimates with paediatric ART treatment numbers where available.
- Encourage countries to enter ANC testing data in Spectrum so that raw, unadjusted ANC data can begin to be collected.
- Review breastfeeding assumptions in Asia/Pacific.

Jinkou Zhao presented the need for HIV estimation in small island countries for Global Fund allocation. Surveillance data are limited: partial case-surveillance and mortality reporting exists in some countries, and behavioural surveys and ART registers are available in Caribbean islands. Global Fund and UNAIDS will work with countries to collate data, and a simple statistical model will be developed to estimate PLHIV cross-sectionally from case-surveillance and vital registration.

Session 9: CSAVR model development

In the 2018/19 estimates round, CSAVR was used in 44 countries and utilised estimates of AIDS mortality adjusted for incompleteness and cause of death miscoding produced by the Global Burden of Disease study. The *r-logistic* and *double logistic* incidence models were used by the majority of countries. Model selection through AIC was well received. Nine countries produced estimates using both ECDC and CSAVR models and comparison shows that the ECDC model is more likely to estimate a decline in recent incidence more than CSAVR, but declines estimated by the ECDC model can be rapid, possibly driven by the

underlying spline fit rather than the data. Accounting for migration and foreign diagnoses requires more consideration and is becoming increasingly important in Latin America as well as Europe.
Recent CSAVR model development was presented by Guy Mahiane. Three substantial changes from the 2018/19 estimates round versions have been implemented:

1) Stratifying results by sex and age
2) Harmonising with shiny90

   Shiny90 tracks populations by HIV testing status, awareness, and opportunistic symptom status. CSAVR does not require tracking by HIV testing, but now reflects shiny90's structure tracking unsymptomatic and symptomatic populations. This in turn influences testing and diagnosis rate.

3) Flexibility of diagnosis and incidence trends

   Available incidence fitting options are logistic, double logistic, splines and rLogistic models. The *r-logistic* model models the transmission rate, while the others model incidence rate directly. To increase the flexibility of diagnosis and incidence trends:
   - Two shape parameters added to the double logistic model;
   - A fourth knot added to the spline model; and
   - Three parameters added to the diagnosis rate function.

Age/sex stratified infection data in migrants can now be entered, to be subtracted from total infections to produce estimates of resident infections. Parameter displays have been updated to show values on the natural and transformed scales, and a text description of each. Testing each of the four incidence models on a test set of countries finds that the r-logistic and spline models give good fits to new diagnosis and AIDS death data, though mean CD4 at diagnosis is continually overestimated. As CSAVR uncertainty analysis does not include uncertainty in Spectrum progression and mortality rates, it will not provide an exact match to Spectrum uncertainty bounds. It is recommended that CSAVR outputs only graphical, rather than tabular, results. Regarding future model development, fitting to age/sex disaggregated AIDS death and diagnosis data, and to the full CD4 distribution are recommended as priorities above including key populations in the model. It is noted that, given Rob Glaubius' earlier presentation on natural history model misspecification, fitting to the full CD4 distribution may not improve estimates.

Session 10: Projecting estimates to 2020, 2025, 2030

As we approach 2020, there is a need to conduct a 1 year projection from the 2019 programmatic data used in the 2019/20 estimates to report on the 90-90-90 targets. Looking beyond the 90-90-90 targets, to targets in 2025 and 2030, model development will look to integrate mechanistic and statistical approaches, consolidate key population data in all epidemics, and incorporate projection and intervention modelling into a single product.

10.1 Global Burden of Disease Study HIV forecasts

Projection methods within the Global Burden of Disease were presented by Deepa Jahagirdar. Incidence, ART coverage, PMTCT coverage, and paediatric ART coverage are forecast independently, and used in conjunction with demographic forecasts to estimate incidence, prevalence and mortality in a given projection year. Incidence is defined in terms of a projected counterfactual incidence (i.e. without ART) and projected ART coverage. ART coverage is a function of projected ART price, national treatment expenditure, and ART coverage caps. The ART coverage caps are produced by modelled frontiers based on projected national income. It is noted that the functional forms of income frontiers are largely model-driven with data only at low incomes (Fig 9). This constrains the maximum attainable ART coverage and may not represent a realistic upper limit.

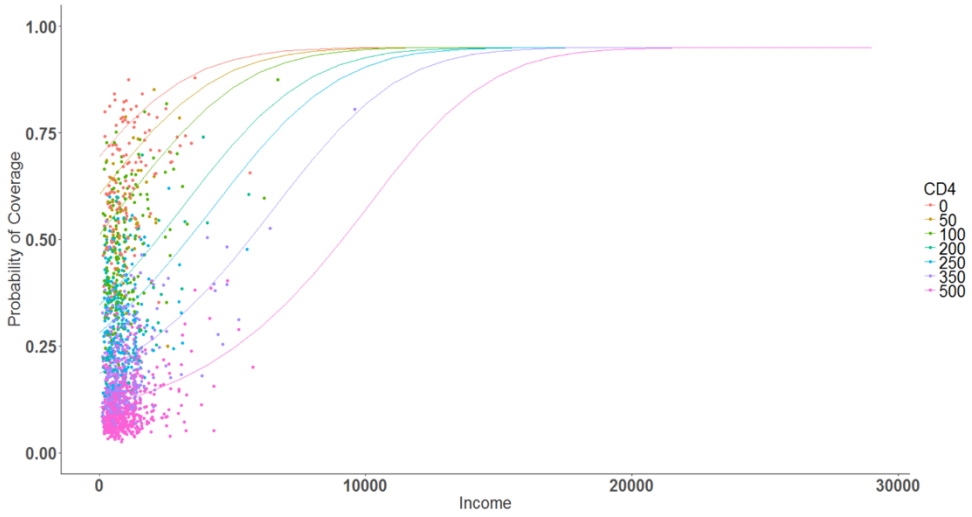


Figure 9: ART coverage caps determined by income frontiers

10.2 Short-term projections with UNAIDS tools

Josh Salomon presented methods and assumptions underpinning short term projections from existing models, summarised below.

	Model/input	Extrapolation method
Incidence		
	EPP	Extrapolate r(t), governed by random walk
	AEM	Extrapolation of risk behaviours and transmission probabilities
	CSAVR	Extrapolate incidence or transmission rate, depending on model selection
	ECDC	Recommend no extrapolation – extend last data point
	Direct incidence	Depends on incidence calculation method
Knowledge of HIV status		
	Shiny90 countries	Extrapolate random walk
	Non-shiny90 countries	User input, using expected attainment
ART coverage		
	Programme data	User input, using expected attainment

Flatline extensions or exogenously defined slope changes are easy to implement, but difficult to justify; while random walk formulations reflect increasing uncertainty as projection increases and have few structural assumptions. Projections can be obtained by extrapolating the model function, or by projecting data and fitting the model. The latter was preferred with spline-based models. Discussion about projecting ART coverage considered whether it is better to forecast ART coverage using probabilistic methods, or to extend the trend in programme numbers based on anticipated treatment scale-up that has informed 2020 budgeting exercises.

In discussion it was recommended that:

Incidence

- For non-spline based models: 1 year projection
- For spline-based models: 1 year data projection, fit model to projected data

First 90

- Use 1 year projections in Shiny90 and CSAVR
- Consider alternative approaches for remaining countries

Second 90

- Time series projection on ART programme numbers

- UNAIDS to publish both modelled projection and national programmatic targets for 2020

Third 90

- Project continuation of upward trend in viral load suppression

Appendix I: Recommendations

Recommendation/Action Item	Lead Person(s)	Proposed timeline
Session 1: Spectrum updates		
Previously treated populations		
• UNAIDS and PEPFAR to identify countries for whom new 28 day definition of "currently on ART" will incur large changes to on treatment numbers and support them to construct consistent time series	UNAIDS, PEPFAR	2019/20 estimates round
• Input fields for "First time" and "Total" ART initiations. If countries are unable to disaggregate, "Total" should be filled.	Avenir Health	1st November 2019
• Encourage countries to enter values for both percentage dropout rates and absolute number of reinitiators. Provide guidance if discrepancies arise	UNAIDS	2019/20 estimates round
ANC testing display		
• Maintain current single year display as default	Avenir Health	1st November 2019
• Explore options to display plots to display trends of single variable or relevant pairings over time as popout window	Avenir Health	1st November 2019
ART by age		
• Countries to enter ART by age to be encouraged to be used as validation data during workshops	UNAIDS	2019/20 estimates round
Natural history model		
• Smooth discontinuities in HIV mortality at 10 year age groups	Rob Glaubius	May 2020
• On ART mortality rates to be re-estimated from IeDEA data by single year of age	Leigh Johnson	2021 estimates round
• Revised adult natural history model not recommended for implementation in 2019/20 estimates round. Look to find parameter set which reconciles PHIA and IeDEA data	Rob Glaubius	2020
Session 2: EPP		
• FRR scaling factor: EPP to return both births to HIV+ women and fitted FRR scale factor to Spectrum	Tim Brown	1st November 2019
• Including survey ART coverage in fitting not currently recommended at a subnational level. Test further at national level and develop guidance for when it should be used by countries	Jeff Eaton	2020

EPP in concentrated epidemics		
• For regions without surveys and with high ANC-RT testing coverage, use ANC-RT as estimator of epidemic level, with an offset for ANC-SS data	Tim Brown	November 2019
• Update review on difference between population and ANC prevalences in concentrated epidemic settings	Kim Marsh	2020
• Ensure countries document the source/approach of all KP size estimates	UNAIDS	2019/20 estimates round
• Flag when sex ratios in concentrated epidemics differ from expected patterns	Avenir Health	
• Consider blocking subpopulation prevalence results when IRRs are adjusted in Spectrum and are subsequently inconsistent with EPP subpopulation estimates	Avenir Health	
Session 3: Treatment Cascade Estimation		
Shiny90 Expanded result set to include: • Probability of being diagnosed within 6/12 months and >350/>200 CD4/ml • Stacked bar of time to diagnosis over time • Simple statistics in current year relevant to programmes. E.g. 1 in X tests are positive. • Estimate of retesting rate • Proportion of tests that are retests Additional plots and data to be made available in PDF and CSV downloads	Mathieu Maheu-Giroux Katia Giguère	1st November 2019
Viral load suppression		
• Implement viral load threshold adjustment in Spectrum	Leigh Johnson/Avenir Health	1st November 2019
• Estimate phi parameters for additional IeDEA regions	Leigh Johnson	2019/20 estimates round
• Recommend against adjusting routine viral load for incomplete data		
• Remove testing coverage thresholds for entering viral suppression data in Spectrum. Percentage suppressed amongst those tested to be used as estimate of 3rd 90 in all cases	Avenir Health	1st November 2019
Cascade visualisation in Spectrum		
• Existing cascade visualisation to be retained. Single-year display for new, detailed, proposal of treatment cascade to be added	Avenir Health	1st November 2019
• Provide CSV extract of detailed cascade breakdown	Avenir Health	1st November 2019

Sessions 4-7: Naomi model inputs, development, and implementation

Population sources and birth estimation		
• Births to HIV+ women not recommended as Naomi output		
• Provide district populations across sources in tabular form for estimates workshops	Megan O'Driscoll	2019/20 estimates round
• Provide summary of districts with discordant populations across sources	Megan O'Driscoll	2019/20 estimates round
• Provide decision tree for estimates teams to use	Megan O'Driscoll	2019/20 estimates round
• Utilise population source comparison within UNPD country webinars with country estimate teams	Megan O'Driscoll/Mary Mahy	Nov 2019 -Feb 2020
Model outputs		

• Model A (fitting to prevalence data only)
  ○ Use ART coverage from survey
  ▪ Return number on ART = estimated coverage * prevalence * population
  ▪ Return number receiving ART = unmodelled programme numbers
  ○ Number receiving ART at district level should sum to national ART programme numbers
	Jeff Eaton	1st November 2019

• Model C (joint modelling with ART reallocation)
  ○ Return number on ART = unmodelled programme numbers
  ○ Return numbers receiving ART = modelled, reallocated numbers receiving ART
	Jeff Eaton	

Further development		
• Seek more programme data for countries for which last survey was HIV prevalence only, especially WCA	UNAIDS	
• Extract programme data from 2018/19 DataPack files	PEPFAR	1st November 2019
• Consider more structure on ART reallocation prior: urban/rural care seeking, gravity distance model	Jeff Eaton	
• Where comparable: compare Naomi results to HIVE, age specific prevalence estimates from Spectrum, CDC SAE results	Jeff Eaton	1st December 2019
Interaction of Spectrum and Naomi		
• Naomi to be launched from Tools menu	Avenir Health	
• Spectrum should have simple, tabular displays of Naomi results. District Estimates Tool must retain graphical capacity for countries unable to use Naomi	Avenir Health	
• Spectrum should not estimate deaths and PMTCT at the district level by disaggregating national results	Avenir Health	1st November 2019
• DataPack to be produced by Spectrum	Avenir Health	
• Population by age and sex to be adjusted to Spectrum before model estimation	Jeff Eaton	

• Mean total PLHIV to be raked to Spectrum	Jeff Eaton	
• ART numbers should not be raked to Spectrum, but total numbers should correspond to those in Spectrum before use in Naomi	Jeff Eaton	
• Subnational Spectrum files to be retained for 2019/20 estimates round. Country to be engaged whether they are amenable to moving away from subnational files for 2020/21 estimates round	UNAIDS	2020
Session 8: Asia Pacific estimates and AEM review		
• Review Spectrum-AEM mortality adjustments in 2018/19 estimates files		AP workshop 2019/20
• Validate deaths in AEM and default Spectrum deaths against national vital registration data		2020
• Consider prevalence/incidence adjustment as currently applied in EPP/Spectrum instead of mortality adjustment	Tim Brown, Avenir Health	1st November 2019
• Spectrum to accept user uploaded IRRs	Avenir Health	1st November 2019
• Create IRR fitting tool for countries without survey data to consume: ○ ART programme data by age (available in 9 Asia/Pacific countries and PEPFAR data) ○ Case reports by age (available in Latin America) ○ AIDS deaths by age from vital registration data IRR fit to non-survey data to be fit outside of Spectrum and uploaded for 2019/20 estimates round	UNAIDS Reference Group	1st January 2020
• Compare paediatric estimates with paediatric programme data		2020
• Encourage countries to enter ANC-RT data in Spectrum programme data entry in 2019/20 estimates round	UNAIDS	2019/20 estimates round
• Review breastfeeding assumptions in Asia/Pacific	UNAIDS/UNICEF	2020
• Ensure countries document the source/approach of all key population size estimates	UNAIDS	2019/20 estimates round
Island nations		
• Global Fund and UNAIDS to work with countries to assemble data	Global Fund/UNAIDS	December 2019
• Develop simple statistical model for cross-sectional approach to estimate PLHIV from case-surveillance and vital registration	UNAIDS	2020
Session 9: CSAVR model development		
• Output graphical, not tabular results	Guy Mahiane	
• Fit to full CD4 distribution rather than mean CD4 ○ Consider and mitigate effects of natural history model misspecification	Guy Mahiane	
• Fit to age/sex disaggregate AIDS death and new diagnosis data	Guy Mahiane	
• Investigate differences in male/female and key population testing rates where data exist and suggest this is warranted	Guy Mahiane	
Session 10: Projecting estimates to 2020, 2025, 2030		

Recommendations for short term projections to 2020

Incidence

- For non-spline based models: 1 year projection
- For spline-based models: 1 year data projection, fit model to projected data

First 90

- Use 1 year projections in Shiny90 and CSAVR
- Consider alternative approaches for remaining countries

Second 90

- Time series projection on ART programme numbers
- UNAIDS to publish both modelled projection and national programmatic targets for 2020

Third 90

- Project continuation of upward trend in viral load suppression

UNAIDS Reference Group	2019/20 estimates round

Appendix II – Meeting agenda

Tuesday 8th October

Time	Duration (mins)	Topic	Presenter(s)/ Lead Discussant
9.00	30	Welcome and introductions Meeting objectives and overview UNAIDS 2020 estimates overview and key dates	Peter Ghys Jeff Eaton Mary Mahy
Session 1: Spectrum updates (chaired by Jeff Eaton)			
9.30	20	Model code and code workflow update	John Stover
9.50	60	• first90 interface • Previously treated compartment • ANC testing displays trends • Spectrum online	John Stover/Avenir Health
10.50	20	Coffee	
11.10	10	WPP 2019 • Updates to Spectrum default inputs	John Stover
11.20	10	Country webinars & guidance for HIV estimates teams	Mary Mahy
11.30	45	• Adult natural history model • Paediatric to adult natural history model transition • On ART mortality smoothing	Rob Glaubius
12.15	30	Discussion	
12.45	45	Lunch	
Session 2: EPP development and implementation (chaired by Tim Hallett)			
1.30	45	EPP • Fitting to age/sex stratified prevalence data • Display changes • EPP in the cloud • Inclusion of ART coverage from survey	Tim Brown Jeff Eaton
2.15	15	ANC-RT bias	Mathieu Maheu-Giroux
2.30	20	Discussion	

2.50	45	EPP in concentrated epidemics
		• ANC-RT interpretation at high testing coverages
		• Key population inputs
		  ○ Challenges and improvements
		  ○ Age group cross walk
		  ○ Sex ratios in EPP and Spectrum
			Tim Brown
			
			Tobi Saidel

3.35	30	Discussion	
4.05	20	Coffee	
Session 3: Cascade estimation (chaired by Mary Mahy)			
4.25	40	Shiny90/first90 in 2019 estimates first90 results Discussion	Kim Marsh Katia Giguère
5.05	30	Adjusting for VLS bias Controlling for variation in national VLS thresholds	Leigh Johnson
5.35	15	Programmatic cascade data entry in Spectrum	Avenir Health
5.50	CLOSE		

Wednesday 9th October

Time	Duration (mins)	Topic	Presenter(s)/ Lead Discussant
Session 4: Naomi overview and data inputs (chaired by John Stover)			
9.00	30	Model overview and data structure	Jeff Eaton
9.30	30	Admin-2 population sources	Megan O'Driscoll
10.00	20	Admin-2 fertility	Oli Stevens
10.20	20	Coffee	
10.40	30	Discussion Recommendations Selection of data sources	All
Session 5: Naomi development, data, and model testing (chaired by Leigh Johnson)			
11.10	30	Model specification • Overview of changes • Spatial structures	Jeff Eaton
11.40	45	Results from test countries	Jeff Eaton Steve Gutreuter
12.25	45	Lunch	
1.10	30	Discussion	All
Session 6: Implementation and scaleup (chaired by Mary Mahy)			
1.45	10	Timelines and process	Mary Mahy
1.55	20	AIDS Data Repository • Overview • Data availability, quality, and process • Current implementation	Fjelltop/Ian Wanyeki
2.15	30	User interface demonstration • Front end user interface • Back end infrastructure	Rob Ashton
2.45	15	Spectrum interface to Naomi and presentation of district estimates	John Stover
Session 7: Interaction of Spectrum and Naomi (chaired by Tim Hallett)			
3.00	20	Comparison of Naomi admin-1 aggregates with Spectrum subnational results • Indicators produced directly by Naomi • Indicators not produced by Naomi	Jeff Eaton
3.20	20	Coffee	
3.40	60	Discussion and recommendations on: • Calibration of district estimates to national. • Guidance on subnational Spectrum files • Interface and workflow between Spectrum and Naomi	All
4.40	20	Summary/Recommendations	Tim Hallett
5.00	CLOSE		
7.00		**Dinner**	

Thursday 10th October

Time	Duration (mins)	Topic	Presenter(s)/ Lead Discussant
Session 8: Asia Pacific & AEM review (chaired by Jeff Eaton)			
9.00	20	Asia/Pacific tools and results in 2019 estimates	Keith Sabin
9.20	40	AEM review • Model structure • Model parameters and sources • Calibration data • Methods	Tim Brown
10.00	30	Surveillance data in the Philippines	Nash Montevirgin
10.30	20	Coffee	
10.50	30	Discussion	
11.20	15	Estimation in small island countries	Jinkou Zhao
11.35	25	Discussion	
Session 9: CSAVR model development (chaired by Leigh Johnson)			
12.00	20	CSAVR in 2019 estimates	Kim Marsh
12.20	45	Lunch	
1.05	60	CSAVR development • Harmonisation of CSAVR/first90 diagnosis model • Diagnosis rate flexibility • Incidence model selection • Age/sex stratified outputs • Testing against simulated data • Migration Suggestions: Age/sex stratified inference	Guy Mahiane (remote) Avenir Health
2.05	40	Discussion & recommendations	All
Session 10: Projecting estimates to 2020, 2025, 2030 (chaired by Kim Marsh)			
3.00	20	Needs and requirements for projections	Mary Mahy
3.20	20	Coffee	
3.40	30	Projections in the Global Burden of Disease	Deepa Jahagirdar
4.10	30	Existing methods for projection and future approaches	Josh Solomon (remote)
4.40	45	Discussion	
5.25	20	Review of meeting recommendations	Leigh Johnson/Jeff Eaton
5.45	CLOSE		

Appendix III - Participant list

Guy Mahiane *	Avenir Health
John Stover	Avenir Health
Rob Glaubius	Avenir Health
Ray Shiraishi	Centers for Disease Control and Prevention
Steve Gutreuter	Centers for Disease Control and Prevention
Italia Rolle	Centers for Disease Control and Prevention
Katie Battey	Centers for Disease Control and Prevention
Tim Brown	East West Center
Jonathan Berry	Fjelltop
Jeff Eaton	Imperial College London
Kinh Nyugen	Imperial College London
Megan O'Driscoll	Imperial College London
Oli Stevens	Imperial College London
Rob Ashton	Imperial College London
Tim Hallett	Imperial College London
Laura Dwyer-Lindgren	Institute for Health Metrics and Evaluation, Seattle, USA
Deepa Jahagirdar	Institute for Health Metrics and Evaluation, Seattle, USA
Mathieu Maheu-Giroux	McGill University, Montreal, Canada
Katia Giguere	McGill University, Montreal, Canada
Newton Chagoma	Ministry of Health, Malawi
Irum Zaidi	PEPFAR
Nate Heard	PEPFAR
Parviez Hosseini	PEPFAR
Ian Wanyeki	UNAIDS, Geneva, Switzerland
Keith Sabin	UNAIDS, Geneva, Switzerland
Kim Marsh	UNAIDS, Geneva, Switzerland
Mary Mahy	UNAIDS, Geneva, Switzerland
Peter Ghys	UNAIDS, Geneva, Switzerland
Tobi Saidel	UNAIDS, Geneva, Switzerland
Leigh Johnson	University of Cape Town
Reshma Kassanjee	University of Cape Town
Josh Salomon	University of Stanford, USA
Brad Mathers	World Health Organization, Geneva, Switzerland
Morkor Newman	World Health Organization, Geneva, Switzerland
Alison Wringe	World Health Organization, Geneva, Switzerland
Cheryl Johnson	World Health Organization, Geneva, Switzerland

* Remote participant