
Method Development for the UNAIDS Estimates: May 2016

Report and recommendations from a meeting of the
UNAIDS Reference Group on Estimates, Modelling and Projections
Geneva, Switzerland, 16-19 May 2016

REPORT & RECOMMENDATIONS

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) based at Imperial College London. Participants of the meeting are listed at the end of this document.

Kelsey Case & Soraya Rusmaully, June 2016

Introduction

The Joint United Nations Programme on HIV/AIDS (UNAIDS) *Reference Group on Estimates, Modelling and Projections* exists to provide impartial scientific advice to UNAIDS and other partner organisations on global estimates and projections of the prevalence, incidence and impact of HIV/AIDS. The Reference Group acts as an ‘open cohort’ of epidemiologists, demographers, statisticians, and public health experts. It is able to provide timely advice and also address ongoing concerns through both *ad hoc* and regular meetings. The group is co-ordinated by a secretariat based in the Department of Infectious Disease Epidemiology at Imperial College London.

Aim of the meeting

This meeting had two key aims:

1. To support the further development and refinement of the current methods used to generate UNAIDS Estimates.
2. To review and discuss method development surrounding the main theme areas including:
 - Spatially-specific estimates of HIV
 - Age-structured models
 - Use of programme service data

Approach

The meeting featured presentations combined with group discussion to generate consensus recommendations. The list of participants is included in Appendix I and the meeting agenda is included in Appendix II.

The recommendations drafted at Reference Group meetings give UNAIDS guidance on how best to produce estimates of HIV/AIDS, provides an opportunity to review current approaches and also helps to identify information needs (earlier reports are published on the Reference Group website www.epidem.org). This transparent process aims to allow the statistics and reports published by UNAIDS and WHO to be informed by impartial, scientific peer review.

I. Pre-Meeting, Key Issues arising from the UNAIDS Estimates

For the 2016 UNAIDS Estimates, estimates were conducted in over 160 countries with 145 countries submitting Spectrum files. Shadow files (estimate files produced by UNAIDS) were prepared for 18 countries which importantly include Russia, China, Ethiopia and Congo. At the global level, the 2016 and 2015 estimates are very similar, but there are important differences in the child estimates due to changes made to the model parameters¹. The changes include:

- *Implementation of a new default distribution of children starting ART by age* based on leDEA data. This resulted in higher mortality historically and fewer children living with HIV (CLHIV).
- *Updated probabilities of mother-to-child-transmission (MTCT) of HIV*, including a reduction in the probability of MTCT with incident HIV infection (18% compared to 30%), which resulted in a reduction in new child infections.
- *Post-natal transmission beginning after six weeks* which also resulted in a reduction in new child infections.

The overall effect of these changes is a reduction in new child infections and AIDS deaths, and a substantial reduction in CLHIV. The new child prevalence estimates generally fall within the confidence intervals of available survey data, but may underestimate prevalence as a result of the lack of reliable data to represent patterns of breastfeeding among HIV-positive women and dropout rates of postnatal prophylaxis. Discrepancies were observed between Spectrum estimates and survey data in Botswana (prevalence amongst 5-9 year-olds in the 2004, 2008 and 2012 surveys), and in Mozambique and South Africa for 10-14 year olds.

Recommendation: Develop manuscript which describes the updates to the child model in detail and explains the changes in child estimates for the UNAIDS Supplement in AIDS. *Follow-up: Mary Mahy and John Stover to lead, submission end June 2016*

Key issues arising in Spectrum

The fit incidence to programme data tool

The *fit incidence to programme data* tool in AIM fits to case-report data (adjusted for reporting lag), AIDS deaths and estimates of PLHIV (adjusted for undiagnosed). The advantage of this approach is that countries with robust HIV case surveillance and vital registration systems are able to directly incorporate this information in the fitting. An increasing number of countries (N=60) used this tool to produce their 2016 UNAIDS Estimates. Several key issues were raised in this process:

- *Is it appropriate to fit to estimates of PLHIV?* It was discussed this is somewhat of a circular process (fitting to estimates). It also requires knowing the proportion undiagnosed, and requires strong linkage between case-report and vital registration systems.
- *Reconciliation between reported new infections and AIDS deaths:* Many countries with strong surveillance systems have reported AIDS deaths that are not in agreement with case-reports of new HIV infections.

¹ *Modelling paediatric HIV and the need for ART*. Report and recommendations from a joint consultation with WHO and the UNAIDS Reference Group on Estimates, Modelling and Projections held in London, UK, 28-29 October 2015. Available: <http://www.epidem.org/resources>

- *Epidemic patterns in EPP vs the fit to programme data tool*: Use of different approaches can result in very different fits. In most cases, the fits obtained using the *fit to programme data tool* were deemed more reasonable.
- *Non-convergence*: Challenges arose with non-converging files, particularly in countries where incidence has not yet peaked (Czech Republic, Malta, Chile, Austria and Bulgaria). It was discussed that the double logistic curve can be over parameterized and recommended that a simple logistic curve is used in these situations. Changing the starting conditions can also help with convergence.
- *Infections acquired abroad and out-migration*. Recent studies have shown changing patterns in Europe of sexual mixing and probable place of HIV acquisition (e.g. estimates acquired abroad), which have important implications for HIV estimates. Other countries, Saudi Arabia for example, are looking for guidance with regards to how to handle out-migration of PLHIV.
- *Outbreaks*. This approach fits a smooth curve to the data and thus is unable to appropriately capture outbreak situations, for example in Greece and Romania.

Issues to consider for future estimates include:

- *Define regional defaults from the literature (time from infection to diagnosis, undiagnosed)?*
- *How to better capture HIV outbreaks (currently smoothing over)?*
- *Develop recommendations for capturing infections acquired abroad and out-migration of PLHIV?*
- *Provide option to fit to either AIDS-related deaths or all deaths among PLHIV?*

Recommendations:

- ✓ **Avenir Health to extract starting conditions from *fit to programme data tool* files and assemble these in a database so that this information can inform future estimates.**
- ✓ **Develop manuscript which describes the *fit to programme data tool* (defining the likelihood functions) and describes changes to the estimates. *Follow-up: Guy Mahiane, Kim Marsh, Kelsey Grantham, submission to UNAIDS Supplement end June 2016***

Estimates of PMTCT coverage

In the 2016 UNAIDS Estimates, many countries with high quality PMTCT programme data, had estimates of PMTCT coverage of 100% or more. There are many potential factors which can contribute to overestimation of PMTCT coverage including:

- Incorrect population or prevalence
- ART coverage by age
- Fertility adjustment factors
- Programme data
- Migration for services

Anna Radin at Avenir Health is conducting a comprehensive review of these factors in countries with >100% coverage of PMTCT. A validation screen will be added to Spectrum to facilitate user comparison of the estimates with population data, survey prevalence and age-specific fertility.

Recommendation: Review results of this review, *end August 2016*

Mortality Comparisons

The World Health Organisation (WHO) produces abridged life tables for Member States as part of its mandate to monitor and report on global progress in improving health. Colin Mathers, Dan Hogan and colleagues compared Spectrum outputs produced from country-specific life tables prepared by the United Nations Population Division. The results appear consistent across most countries, but in

countries where ‘mortality shocks’ have been observed (Rwanda, for example), these shocks that occur in a single year are not captured in the 5-year smoothed projections produced by the Population Division. It is possible to use the WHO projections for annual mortality schedule, but it is not clear if this will have an effect on current estimates.

Malawi is a high-burden country where Spectrum estimates of all-cause mortality appear over-estimated compared to WHO estimates. Further investigation revealed that this may be due to incorrect EPP fits. The EPP fits for Malawi are now produced by region and impose a downward trend that is not observed in the survey data. Re-fitting to more closely match the survey data, resulted in a different trend in HIV incidence, but did not solve the discrepancies in mortality. This issue requires further investigation.

The Population Division is longer producing non-AIDS life tables for medium-HIV burden countries. WHO has estimated non-AIDS mortality (using Spectrum AIDS deaths) which can be used for these countries. The alternative is to employ the interpolation methods previously used (3 year ago).

Key issues arising in EPP

Implementation of hierarchical approach in EPP

Tim Brown and Robert Pucket have been working to implement Le Bao’s approach for analysing data with a hierarchical structure or spatial dependence in EPP. This approach uses generalised linear mixed models (GLMM) which call on R which is licenced under GNU public licence (GPL). As a result, if EPP invokes R directly, it is believed EPP may require release under GPL-licencing which would have important implications for both EPP and Spectrum. GPL specifies that all source code and libraries needed to build the programs must be made open source.

While the group agreed it is ideal to ultimately have open-source software, the full ramifications of this issue cannot be ascertained by this group. It was noted that because UNAIDS supports East-West Centre for the development of EPP, they are technically the owner of the source code. It should be noted that GPL is not the only option. There are other open-source licenses (MIT or Apache) which allow software to be made open-source, but do not have the viral nature of the GPL.

Recommendations:

- ✓ **The Reference Group encourages open-source software, but a further technical and legal discussion is needed.**
- ✓ **UNAIDS to work with Avenir Health and East-West Center to draft a brief summary of the issue to present to WHO lawyers. *Follow-up: UNAIDS, John Stover, Tim Brown, May 2016***
- ✓ **UNAIDS to seek legal advice regarding GPL, the impact for the commercial libraries that Spectrum uses and the potential issues in different countries with regards to copyright. *Follow-up: UNAIDS, June 2016***

Implementation of AEM in Spectrum

The AIDS Epidemic Model (AEM) can now be used within Spectrum and has been updated to improve synchronization. However, many countries want to observe continuity in their estimates of HIV prevalence, incidence and AIDS deaths from previous AEM files. Because AEM is not age-structured, the same incidence pattern in AEM and Spectrum will develop variations over time. The age distributions of PLHIV and the general population are very different and thus there are differences in the background mortality (age-effect of background mortality) which contributes to this discrepancy between AEM and Spectrum. For the 2016 Estimates, East-West Center worked

directly with countries to adjust transition and mortality parameters in order to reduce this variation. It was agreed this approach should not be considered a long-term resolution.

The next steps are to investigate generating a database of mortality ratios over time from Spectrum for use in AEM to produce a pattern of background mortality pattern that is more representative of the population of PLHIV and better aligned with Spectrum.

Recommendation to investigate methods to better match background mortality between AEM and Spectrum. *Follow-up: Tim Brown and John Stover, review by Aug 2016*

Newly implemented features in EPP and Spectrum

Implementation of variance inflation

Le Bao previously identified EPP may not capture all sources of variability in ANC prevalence data, which was observed with ANC values falling outside the confidence intervals (CIs). As a result, EPP is overly sensitive to spurious trends in ANC prevalence, underestimates epidemic uncertainty (particularly before surveys), and may not give enough importance to higher quality data sources (e.g. national surveys). Jeff Eaton previously proposed² to add additional variance to account for non-sampling error in ANC prevalence. The variance inflation parameter has now been implemented in EPP. Preliminary results from Botswana and Zimbabwe illustrate these changes result in greater uncertainty in both prevalence and incidence, particularly during earlier stages of the epidemic, as expected. Further testing will continue to observe the effects across many countries. The user will be able to turn this feature on/off and will be able to set the prior.

Recommendation: Review results from implementation across many countries, *Follow-up: Review results in next Reference Group teleconference, July 2016*

Spectrum for web

Spectrum is already able to run on the cloud, and Avenir Health is currently developing Spectrum for the web. This would allow users to run the software directly from their browser. It was agreed that this additional availability would be beneficial for many users. UNAIDS, Avenir Health and East-West Center will further liaise to discuss in detail the implications of web-based Spectrum and strategies for improving efficiencies, for example file sharing.

STI estimation model in Spectrum

A new tool has been developed in Spectrum to generate estimates of gonorrhoea and syphilis at the national level in order to inform programme planning and target setting. These national estimates would also feed into WHO's global and regional estimates of sexually transmitted infections (STIs). This tool, *Spectrum-STI*, is currently being piloted and is planned for incorporation at the 2017 regional HIV estimation workshops.

² Method development for the UNAIDS estimates: October 2015. Report and recommendations from a meeting of the UNAIDS Reference Group on Estimates, Modelling and Projections held in London, UK, 26-27 October 2015. <http://www.epidem.org/resources>

II. Updates from collaborative partners and organisations

The work of the Reference Group occurs in coordination with other groups including the European Centre for Disease for Disease Prevention and Control (ECDC), the Measurement and Surveillance of HIV Epidemics (MESH) Consortium, ICF International, the United States Center for Disease Control and Prevention (US CDC) and the Institute for Health Metrics and Evaluation (IHME), among others.

Update from ECDC

The ECDC HIV modelling tool is available online and provides two approaches to estimate HIV incidence. A workshop was held in February 2016 to discuss and apply both the Spectrum software and the ECDC modelling tool in six countries. While both approaches generate results for key indicators, the results often differ, and the quality of the estimates will depend on the quality of the input data. The effects of migration, infections acquired abroad and missing data remain key issues for generating HIV estimates in this region. It was highlighted that the ECDC model estimates the proportion undiagnosed which could be used to inform the input data for the *fit to programme data* tool in Spectrum.

The ECDC modelling tool is applied to sub-populations (risk groups) while Spectrum's *fit to programme data tool* is used to fit estimates at the national level. Many countries in this region also have their own country-specific modelling processes used to generate estimates. While ECDC will continue to support the method development and provides some support for use of its tool, it does not oversee the development and production of estimates. It is unlikely that most EU countries will generate routine annual HIV estimates in a coordinated fashion due to human resource constraints.

Recommendation: ECDC and UNAIDS to continue to hold workshops together and work together for the production of robust HIV estimates in the European region.

The MESH Consortium

The MESH Consortium is comprised of working groups which focussed on the following areas:

- 1. Measuring HIV-related mortality**
 - Mortality on ART: Key focus is lost to follow-up, systematic review and meta-analysis
 - HIV attributable mortality: Methods to estimate mortality among PLHIV starting ART
- 2. Routine case-based surveillance**
 - Protocol and data collection guide
 - Situational analyses in-country to inform case-based surveillance guidelines
 - Systematic review of papers related to care cascade in SSA
- 3. Key populations**
 - Size and HIV epidemic dynamics
 - Stigma

These working groups feed into guidelines and dissemination activities. The UNAIDS Reference Group draws on the information gleaned across these working groups.

Public Health Impact Assessments (PHIAs)

Led by the US CDC, PHIAs are household-based surveys which focus on HIV in 13 PEPFAR countries.

The primary objectives are to:

- 1. Measure adult HIV prevalence and incidence (LAg assay)**
- 2. Estimate viral suppression**

The secondary objectives are to describe HIV risk behaviours and uptake of HIV services (care cascade), and optionally, to estimate paediatric prevalence at the national level. Data collection has now begun (Zambia) and countries will soon need to incorporate results from PHIA in Spectrum files. Some countries (Zimbabwe, for example) have simultaneous or adjacent household surveys occurring which could yield different results.

Home-based testing is used in PHIA which may increase participation bias compared to past DHS, and could result in different trends. PHIA can identify testing uptake in general but will not have detailed information about those who do not consent in order to measure this bias. The PHIA testing algorithm is also different from DHS; it relies upon the national testing strategy but then also includes a highly specific confirmatory test done remotely. There is the potential for this algorithm to also affect prevalence trends as it differs from the traditional testing strategy.

Recommendation to OGAC for ICAP to coordinate with Till Barninghausen/Mark McGovern to investigate selection biases and the use of Heckman methods for adjustment due to selective non-response. *Follow-up: Jacob Dee to bring this to the attention of OGAC, Joy Fishel to raise with DHS*

Performance of EIAs for HIV serology in Demographic and Health Surveys

HIV testing algorithms which render specimens positive after two positive enzyme immunoassays (EIAs) have the potential to result in overestimation of HIV prevalence. The potential for overestimation appears largely due to operator error during testing. It is now recommended that a third, more specific, confirmatory assay is used to confirm dual-EIA positive results. The DHS programme reviewed data available from 20 surveys since 2010 to assess the potential quality of testing in recent surveys. The full report is available at:

<http://dhsprogram.com/publications/publication-cr39-comparative-reports.cfm>.

The indicators of quality investigated were agreement between the 1st and 2nd EIA, and the percent of HIV-positive specimens with a low signal-to cut-off value (<5.0). The general finding was the initial assay tended to have many false positives, which should be excluded with the 2nd assay; however, the results of this process varied across countries. The data available do not allow for firm conclusions regarding overestimation or the magnitude of overestimation, but it is probable that some systematic overestimation is occurring. Surveys from Zambia, Senegal, Sierra Leone and Niger are the most suspect (10-20% false positive, higher in Niger). Surveys from Malawi, Uganda and Burundi could have a false positive rate around 10%. Re-testing of samples from recently completed surveys from Zambia, Ghana and Lesotho is currently underway.

In the future, the EIA testing algorithm will change which raises the question of how to handle past surveys and the potential need for adjustments to survey prevalence. Further, in future DHS surveys only those who are willing to receive their test result will be eligible for participation.

It was highlighted there are surveys going forward which allow for the comparison of prevalence with and without the confirmatory test. For example, Ghana and Malawi have added a confirmatory test. It is anticipated the difference will depend on the quality of lab testing and will vary across countries. It is also not clear how/if countries will report the different outputs.

Testing strategies have also changed over time. Previous confirmatory testing strategies were generally to re-test or to confirm with Western blot. Prior to 2010, older generation tests were used which detected antibody only while the new tests detect antibody and antigen. It is believed that

these newer tests detects false positives more often, thus this potential problem would have increased over time.

It was queried whether it would be possible to obtain the full suite of lab testing results. This information could inform a Bayesian framework and a more powerful analysis (continuous variables and cut-offs). The DHS programme has have talked about this but have not had data agreements with countries/labs in order to publicly release these data. It might be more realistic to use the data used for this systematic analysis to do this type of analysis (which does have the continuous variables), however the structure of the data are quite difficult for epi analyses.

Recommendations:

- ✓ **Bayesian analysis of results of tests from DHS**
- ✓ **Seek optical densities of the tests**
- ✓ **Comparison of the DHS methods with the PHIA's**
- ✓ **DHS to provide data for this group with regards to testing algorithms (past vs future)**
- ✓ **Revision to handling of uncertainty in DHS (and overall projection)**

Follow-up: DHS and CDC to lead on this endeavour with additional support from Secretariat

GBD2015 HIV estimation process

The Global Burden of Disease study (GBD) is informed by estimates of HIV from EPP/Spectrum. However, modifications are made in order to produce internally consistent estimates of incidence, prevalence and mortality. In GBD2013, modifications made to Spectrum included:

- UNAIDS Spectrum model recoded in Python for parallel computing
- Modified assumptions for mortality on/off ART from systematic review of published studies
- Propagated uncertainty (vary all key input parameters over plausible ranges)
- Concentrated epidemics with near complete vital registration data: Calibrate UNAIDS models based on cause of death data adjusted for HIV misclassification
- Generalised epidemics in southern Africa: Adjust so that prevalence/incidence in agreement with reported mortality

Key challenges which remained in GBD2013 included a mismatch between all-cause mortality and HIV estimates even after attempts for adjustment (selecting most consistent estimates from uncertainty intervals). For example, in Botswana HIV mortality had to be reduced 20% to be lower than all-cause mortality. This suggests all-cause mortality is too low, prevalence is too high, or survival is too low. In developed countries, implausible ratios of incidence to mortality remain.

For GBD2015, the key goals were to improve the consistency between EPP and Spectrum processes, eliminate the “matching” process used for generalised epidemics, and in high-income countries, update the mortality assumptions and incidence adjustment process for countries with vital registration data. In GBD 2015, there were methodological improvements to resolve inconsistency between EPP/Spectrum mortality estimates and all-cause mortality estimates from the GBD demographic estimation process. An ensemble model was used to combine HIV estimates from EPP/Spectrum and all-cause mortality and a Cohort Incidence Bias Adjustment (CIBA) was used for developed countries. Overall, in GBD2015 there is a stronger correlation with the UNAIDS estimates, results are more similar in agreement and uncertainty has increased.

Changes in the methods used for GBD2015 include:

- Linked EPP and Spectrum runs with identical input parameters
- In EPP, use of weighted parameters instead of a simple mean
- In Spectrum, use draw-level HIV-free mortality input consistent with all-cause mortality process which effectively links HIV estimates from Spectrum and demographic processes
- Ranked draw-level inputs including mortality on/off ART to increase uncertainty interval
- Mortality on ART updated for developed countries using ART-CC data
- Use of ensemble model for generalised epidemics
- For all-cause mortality estimation, use of single year sibling survival estimates and draw level HIV crude death rate to expand uncertainty for countries with limited mortality information

Ensemble model for generalised epidemics: GBD2015 gives less weight to the all-cause mortality envelope in deriving final HIV mortality and all-cause mortality estimates. Equal weight is given to HIV mortality from Spectrum and all-cause mortality with final HIV mortality estimates obtained by averaging estimates from Spectrum and the mortality envelope process.

Concentrated epidemics: Mortality on ART updated for developed countries using data available from 10 countries from the Antiretroviral Cohort Collaboration (ART-CC), which resulted in much lower estimates compared to GBD2013. In countries without national survey data, UNAIDS estimates are combined with GBD-Spectrum runs with incidence adjustments where necessary. Data from the National Health and Nutrition Examination Surveys (NHANES) in the United States are used to adjust CD4 progression.

Cohort Incidence Bias Adjustments (CIBA): Scales the sizes of each incidence cohort based on raw estimates of HIV mortality from Spectrum using incidence curves without any adjustment and those observed in VR with adjustments for incompleteness and cause misclassification. Spectrum is re-structured to follow cohorts of PLHIV by year and age-group.

Recommendations:

- ✓ **Working group to form, *Follow-up: UNAIDS & Secretariat to coordinate***
- ✓ **Further comparison of methods used and differences in estimates. Manuscript describing differences in approaches, flagging key difference in estimates for the UNAIDS Supplement. *Follow-up: Kim Marsh & Haidong Wang to lead, submission end June 2016***
- ✓ **Comparison of GBD model assumptions (including induced parameters for natural history, effect of ART) with available country data (east Africa, ALPHA) *Follow-up: Haidong Wang***

III. Ongoing development of EPP and Spectrum

Incorporation of incidence assays in EPP

Data from incidence assays in national population-based surveys can currently be included in EPP. Countries need to input the number of infections disaggregated by recent vs non-recent, number of negative samples, the mean duration of recent infection (MDRI), false recency rate (FRR) and defined length of recent infection for the assay. While these data will soon be available, at present, countries like Rwanda and Swaziland have incidence data available from a cohort sample. In Swaziland, the cohort incidence is measured at the national level but Spectrum estimates are produced at the provincial level. This raises the questions:

1. *How to incorporate incidence data from a cohort-type measure?*
2. *How to incorporate national incidence measures in subnational projections?*

Le Bao highlighted that a cohort study sample, representing a subset of the general population, cannot be addressed by the current implementation in EPP which assumes a random, independent draw from the general population. The likelihood would need modification with a joint probability to incorporate this type of data.

It was discussed that in the future there may be other types of data from incidence assays, for example a rapid incidence assay (rapid test kit with additional band for recent infection) is soon headed to the field and may later occur in routine surveillance. Case-based surveillance may also have incidence measures in the near future, and there may be incidence assays amongst key populations. It was further discussed that the PHIAS will be conducted at the national level and thus an increasing number of countries will require guidance on how incorporate national-level incidence for a subnational projection (and likely huge uncertainty at subnational level). The incidence measure from PHIAS may also introduce bias if there is a high proportion of non-consent.

Recommendations:

- ✓ **Investigate reformulation of likelihood to account for sampling frame and survey design to incorporate incidence from cohort studies, PHIAS, and incidence measures in routine surveillance, with the aim to better understand what incorporation of these different types of data sources will entail.** *Follow-up: Le Bao, Jeff Eaton, Tim Brown*
- ✓ **Working group to conduct these further investigations and generate recommendations for how to use incidence data.** *Working Group: Tim Hallett, Tim Brown, Jeff Eaton, Le Bao, Mary Mahy, Secretariat to coordinate, review progress at Fall meeting, Sept 2016*

Multi-state Markov models for estimation of mortality and disengagement from care

The implementation of adult mortality on ART in Spectrum is stratified by factors measured at the start of ART including age, sex and CD4. At present, there is no further CD4 progression when on ART. This work aims to incorporate discontinuation of ART (with CD4 decline) and re-engagement in care in Spectrum. The deliverables are to provide estimates of mortality (by age and sex) while in care and disengagement from care. Estimates will be produced for all leDEA regions with double-sampling. For other regions, estimates of the classification probabilities from external studies or programmes with double-sampling designs can be used; however, it is unlikely that estimates of mortality are generalisable from one programme or study to another.

Double-sampling (east Africa) is used to ascertain the status of patients disengaged from care. This information is used to estimate transition intensities, transition probabilities, and the effect of covariates. Giorgos Bakoyannis, Ying Zhang and Constantin Yiannoutsos have developed multi-state Markov models to estimate mortality and disengagement from care in the presence of outcome misclassification. This includes the development of:

1. Non-parametric pseudolikelihood estimator of the transition intensities and transition probabilities of a Markov process under outcome misclassification
2. Semiparametric pseudolikelihood estimator for the proportional hazards model for the transition intensities under outcome misclassification
3. Parametric version of both of the above for piece-wise exponential models (similar to Spectrum software assumptions)
4. Goodness of fit test regarding the multinomial logit model for the predictions

The next steps are to:

- Improve estimation of arrival to each CD4 state based on a joint model with B-splines for the trajectory of CD4 on ART
- Develop easy-to-use functions in R and STATA for implementation of methods outside leDEA
- Relax the Markov assumption
- Incorporate re-engagement in HIV care
- Incorporate transfer to another facility as an additional state?

The use of discrete states as opposed to continuous state was queried and it was agreed a joint model with continuous CD4 would be ideal, but the discrete states were used correspond with the structure of Spectrum (developed in response to the data previously available). Spectrum can be modified but it would be useful to first fully understand how important it is to make this change. One suggestion was to first consider fitting the most appropriate model to the data, then seek an appropriate way to incorporate into Spectrum.

It was discussed that in the double-sampling, the outcome status of those disengaged from care is either determined as either dead or alive and if the latter, reengagement is queried. However, there is no additional information with regards to the outcome of those re-engaging. In the model, those who re-engage in care are assumed to have the same future behaviour as all others in their state.

Recommendations:

- ✓ **Review potential for direct estimation of outcomes of those re-entering care, or propose best assumptions for use in absence of these data**
- ✓ **Produce estimates for East Africa and identify impact in Spectrum**
- ✓ **Identify if optimal to slightly alter the structure of the Spectrum estimation model for better alignment with Markov models**
- ✓ **Generate recommendations for regions without double-sampling**

Follow-up: Giorgos and Constantin (with input from John Stover). Present results at Fall meeting, Sept 2016

IV. Spatially-specific estimates

A key aim of UNAIDS is to support the production of estimates at lower levels. Consensus among stakeholders on the methods to use to produce these estimates is needed. As part of the UNAIDS Global Estimates process, approximately 20 countries have worked to develop estimates at the subnational level, 12 countries submitted subnational files for the 2016 UNAIDS Estimates, of which seven had subnational estimates of robust quality for publication. A key issue is the demography at the subnational level where there is often substantial uncertainty for both the historical period and the present projection. The US Census Bureau has supported the development of these subnational demographic projections. Migration presents a challenge, both in terms of demography and migration for services (>100% ART/PMTCT coverage). A lack of data at lower levels can result in strange curve fits, and survey data may strongly influence subnational estimates but are often not powered for use at these lower levels.

Countries are currently using Spectrum to develop estimates at the following levels:

- **1st level – National:** Prevalence estimates from ANC and national surveys are used in EPP/Spectrum with national programme data and WPP demographic data.
- **2nd level – Province/region:** Same process as national, but fewer ANC sites at this level, larger uncertainty around survey data. Programme data available (but more difficult to obtain, often do not sum to national). Subnational demographic projections required.
- **3rd level – District/county:** May/may not have ANC data, may have DHS cluster data, rely on prevR in most instances. Programme data should be available at this level. Do not have full demographic projections. Simple distribution of subnational estimates to lower level (e.g. multiply prevalence by population)
- **4th level – Sub-county** (e.g. Kenya): PMTCT data used and programme data from site-level. County indicators distributed to sub-county level. In Kenya, all data (DHS, KAIS, ANC, PMTCT) compiled into a chart and the outputs are compared with ART, PMTCT coverage.

Estimates of PLHIV at the 2nd subnational level

Health surveys are generally designed for estimation at SNU-1, the regional level, for example. However, these surveys are increasingly being taken down to lower levels – SNU-2/3. PEPFAR 2016 requires estimates of PLHIV at SNU-2 and different approaches have been used for these estimates including disaggregation of Spectrum estimates, small-area estimation and prevR.

1. **Spectrum disaggregation:** Uses national/subnational Spectrum file, demographic data and PMTCT site-level HIV prevalence.
 - Advantages: Can disaggregate any Spectrum projection to SNU, can be used in countries without surveys, all SNU will sum to national total
 - Limitations: Dependent on Spectrum modelling assumptions, no measure of uncertainty
2. **Small area estimation:** Fay-Herriot models which combine survey and auxiliary data
 - Compute appropriately weighted average of direct and model-based estimates
 - Advantages: Simple estimation from data, no Spectrum modelling assumptions, can improve accuracy and precision, incorporate spatial variation, and provide valid CIs.
 - Limitations: Estimates limited to timing of health surveys, requires benchmarking to sum to national total, and requires updated shape files (for spatial variation).

Changing geographic units/boundaries over time presents a key challenge to generating these estimates. There is a need for geographic alignment on boundaries as opposed to place names with common geographic hierarchies amongst organisations (i.e. UNAIDS, PEPFAR). It was discussed that the development of small area estimates will not be supported (by OGAC) for routine consensus use. In Mozambique, Spectrum disaggregation is the approach currently used; however, production of estimates at the district level remains a challenge as a result of data availability and quality. It was highlighted that all countries need empowerment with tools and methods that facilitate decision making across all funding streams.

PEPFAR perspective on subnational estimates

In PEPFAR countries, there is now a focus on geo-prioritisation with focussed scale-up of services (ART) at the subnational level. This targeted planning, and the figures used to develop targets and allocate resources, are almost entirely derived from modelled estimates (PLHIV and unmet need for ART). Subnational estimates are also used to measure performance and monitor impact.

2015 was the first year of geo-prioritisation whereby PEPFAR country offices brought their own subnational estimates derived from various methods including prevR, spreadsheet disaggregation (based on prevR), and PMTCT prevalence applied to population data with corrections. In 2016, select high-burden countries received (from headquarters) subnational estimates derived from Spectrum spreadsheet disaggregation or small area estimation methods. Further PEPFAR-supported processes which rely on subnational estimates include the DREAMS initiative, a US\$ 385 million partnership to reduce new HIV infections among young women (15-24 years) in ten PEPFAR countries. This initiative is focussed at the 2nd subnational unit (SNU-2) with the goal to reduce incidence by 25% in 2016 and 40% by the end of 2017.

Going forward, collaboration between UNAIDS and PEPFAR is needed to ensure alignment on global estimates, and the methods used and results generated, at the national and subnational level. It was discussed that uncertainty is a key component, for all estimates but particularly at the subnational level. However, in the PEPFAR work the ranges have been dropped and only point estimates are used. The question was also raised whether PEPFAR geo-prioritisation will result in a change in care-seeking behaviour with patients changing their patterns of service to attend well-funded clinics.

Development of methods to produce spatial estimates of HIV epidemics

An ongoing stream of method development is underway to provide more robust estimates at the subnational level. Samir Bhatt and Pete Gething are developing a Bayesian geospatial model which aims to incorporate a range of data sources (currently survey, PMTCT and ANC data) and predictor data (covariates) and combines this information to generate spatially-specific estimates of HIV prevalence, incidence, AIDS deaths and ART coverage at the most granular level, which can be aggregated to any geographic level of interest, and also provides rigorous representation of uncertainty. This approach is informed by existing UNAIDS parameters and approaches (where appropriate). The methods will initially be applied in 12 countries with generalised.

The methods developed take national (or subnational) level EPP/Spectrum prevalence by year, age, sex and incorporates survey prevalence (by year, age, sex), PMTCT prevalence (year, age), ANC prevalence (year, age), and spatial covariates and combines this information in a Bayesian space-time geostatistical model to generate pixel-level estimates of HIV prevalence through time, disaggregated by age and sex. The main model is a Gaussian process stacked generalisation

(publication forthcoming). A Gaussian process is used to fit functional relationships for age and sex (individual level), off-set by the Spectrum mean, boosted by the inclusion of covariates, and incorporating spatial and temporal dynamics (interaction with age and time).

ANC + PMTCT model: Combines ANC and PMTCT data each with a bias parameter, capitalising on the strengths of each. PMTCT (routine data) have a strong spatial resolution, while ANC data (sentinel data) have fewer sites but are temporally rich which is very useful to understand the historical period. Includes patio-temporal model, stacked covariates, Matérn spatial covariance, exchangeable constant temporal covariance.

Survey model: “Tethered” to Spectrum. An anomaly approach is used to de-trend the data. The mean from Spectrum is used, the data are de-trended based on this mean into an anomaly, in order to learn the spatial pattern, then sent back in. Includes bias adjustment terms, spatio-temporal mode, covariates, Matérn spatial covariance, Spectrum spatial trend (but no temporal trend).

Uncertainty and aggregation: Bayesian uncertainty, 95% credible uncertainty, ability to aggregate to any resolution while preserving uncertainty bounds

Role of covariates: The key predictors are PMTCT and ANC data (45%), night lights (15%), age (15%, lower predictive power than perhaps intuitive because model tethered to Spectrum mean), accessibility (10%), land aridity (5%), temperature fluctuations (5%), and population (5%).

The next steps are to then estimate ART coverage and generate estimates of HIV incidence.

Estimating ART coverage

A two-step approach is proposed for estimating ART coverage. First, estimate the spatial coverage pattern from facility data. The denominators are key here and require defining a catchment area for each facility. The prevalence model provides the number infected per pixel, but then requires “allocation” to facilities, i.e. catchment modelling. Examples from malaria work (Haiti catchment modelling) illustrate the importance of how the catchment area is defined. By incorporating physical access to health facilities and replacing “crow’s flight” distance with realistic journeys across a network (incorporating transport network data, journey time modelling, care-seeking data), a more robust map of coverage per facility catchment can be derived. Note this approach will not address more complex definitions of coverage (e.g. different ART eligibility criteria). Next, the pattern of coverage across space and time is estimates. The estimates of HIV prevalence and ART coverage across space and time are then combined with migration to derive estimates of HIV incidence.

Estimating HIV incidence

Two conceptual approaches for estimating incidence are being considered:

1. Mini epidemic progression model within each pixel (i.e. mini EPP/Spectrum):

- Pixel level estimates of:
 - Number infected (age, sex, time)
 - ART coverage (age, sex, time)
- All other inputs/parameters are national-level/fixed
- Generate age-sex-time incidence (and death) per pixel
- Advantages: Closer representation of epidemiology, more conceptually aligned with Reference Group models, estimates of AIDS deaths.
- Disadvantages: High sensitivity to migration assumptions, computationally expensive, and no guarantees of consistency with national level estimates

2. Disperse national incidence according to relative transmission potential per pixel

- Incidence a function of (per pixel):
 - Number infected by age and ART status (from geospatial model)
 - Relative transmission rate for ART vs non-ART (*a priori* specification)
 - Relative transmission rate for each age group (*a priori* specification)
- Advantages: Parsimonious, constrained to national-level Spectrum/EPP incidence, more robust to migration biases
- Disadvantages: Less biologically principled (greater reliance on simplifying assumptions), no automatic pathway to AIDS deaths

Incorporating migration

Two potential approaches for consideration:

1. Cell phone and census data (Flowminder data, pixel to pixel migration)
2. Nuno Faria (Oxford), Bayesian Phylogeography: Regional classifications, build in transitions between geographies, tracing evolution back in time (to obtain spatial-temporal picture)

It was discussed that both approaches for estimating incidence should be investigated. The first approach may suffer from a lack of historical data (survey data, AIDS deaths), and an inability to appropriately capture the historical trend can affect current estimates. Test the methods in countries with rich data sets. Basia Zaba highlighted there is a strong difference between new locally-acquired infections and incidence as a result of moving into an area. In SSA, the latter often dominates, as high as a factor of 4-5.

Questions were raised whether this approach can incorporate new data available in the future, and also the availability of data at the facility level. The new surveys forthcoming are powered at the subnational level, but not at the facility level. Also, not all facilities provide ART (and ART availability may change over time at the facility level). It was discussed the framework under development aims to be able to incorporate new data sources when available, for example incidence assay data. By focusing at the facility level, this approach will provide optimal flexibility to aggregate to any SNU.

Modelling HIV epidemics at subnational and sub-population level

Le Bao is developing new methods to further strengthen and efficiently model HIV epidemics at the subnational and sub-population level in EPP. This approach applies hierarchical models to observed data (simplified approach which uses spline functions instead of EPP), and then incorporates the results into EPP via auxiliary data (“pseudo-site”). Generalised linear mixed models (GLMM) are commonly used for analysing data with a hierarchical structure or spatial dependence. The posterior distribution of prevalence from GLMM is used as the prior information. Results from use of this approach illustrate it can efficiently improve (but will not reduce) prediction accuracy. While results are generally very similar in the recent period, there can be very different historical trends.

New developments include a set of improved hierarchical models which allow for testing the heterogeneity of the epidemics. Three possible implementations of GLMM are compared in R – estimating the model parameters using MCMCglmm, STAN and INLA. INLA performed the fastest with stable performance, and provides easy extension to spatial-temporal models, but does not provide full posterior samples. MCMCglmm does provide full posterior samples, but is relatively slow with a large number of sub-epidemics and has limited flexibility (no spatial modelling). STAN is transparent and flexible but extremely slow computationally.

Next steps: finalise the set of models for concentrated epidemics (June), finalise the communication with EPP/Spectrum (June), test the models on multiple countries (August). Further refinements may be needed following this testing process – are all proposed models necessary (simplify?), define the recommended range of auxiliary data sample size (which determines the strength of the prior), and define the recommended form of the spline model. It was discussed that this approach could be done in R in advance of the country workshops, which would avoid the GNU licencing issue that arises with implementing GLMM in EPP.

Intersection with the geospatial model: This approach aims to improve estimates at finer geographic levels which will better inform the geospatial model developed by Pete and Sam (which will use the mean from this approach as the “new” Spectrum mean in the prevalence model). The method development from both streams of work will continue in parallel but also learn from each other to further inform. It was discussed it would be useful to observe the performance of both approaches in the same countries, as currently the methods are being applied to different sets of countries.

Patterns of service attendance

The aim of this work is to investigate if mobility for antenatal care (ANC) introduces bias in national and local HIV estimates which is relevant for not only HIV estimates but also resource allocation based upon geographic prioritisation. It was hypothesised that attendance at more urban ANC clinics by women from rural settings with lower HIV prevalence could result in an underestimation of HIV prevalence in urban areas. Data from Manicaland, Zimbabwe were first used to investigate this hypothesis, and then expanded to an analysis at the national level.

Previous findings from Manicaland illustrate an underestimation of HIV prevalence (in two towns) as a result of attendance of women from surrounding rural areas. Rural residents who attended non-local ANC services were younger, less educated and less likely to be married compared to those attending local services. The main reason stated for this mobility for services among rural residents was for better quality of care, while urban residents stated their mobility for services was as a result of staying with relatives. This same analysis was then conducted at the national level but did not find a bias in HIV prevalence as a result of non-local clinic attendance. Potential factors that may contribute to this finding include a narrowing of the difference in HIV prevalence between urban and rural areas, a decrease in levels of non-local attendance in towns over time, and higher HIV prevalence among rural women near towns attending for services in towns.

The next step was to compare data from DHS vs ANC. This requires defining an ANC catchment area and a corresponding selection of DHS clusters. For Zimbabwe, there were similar HIV prevalence estimates among ANC and pregnant women from DHS surveys. This analysis is an extension of the Montana et al (2008) analysis, updated to include the most recent DHS and ANC data available.

Further questions from the group included:

- *Do known positives have a characteristic pattern of local vs non-local attendance?*
- *Conduct the same analysis in a country where there is bigger difference in HIV prevalence in urban vs rural areas?*
- *Consider public vs private clinic attendance (may be differences here)?*

It was raised that there is also a need for patterns of service use in urban areas. Data from the ALPHA Network data are rich but are from rural areas.

Recommendations for subnational estimates:

- ✓ Longer-term aim is for a single set of robust subnational estimates. It is anticipated that the Bayesian geospatial approach will be the preferred method for generating these estimates and will be available in the future
- ✓ In the interim, make countries aware this approach is forthcoming and provide coherent guidelines advising the methods available, with emphasis on continuity and simplicity
- ✓ For countries where use of the geospatial model is not possible, a limited selection of alternative methods should be provided with careful guidance. *Follow-up: UNAIDS & CDC to produce simplify guidance*
- ✓ Subnational estimation working group to advise on the technical issues and to speak to this topic on a regular basis. *Follow-up: Le Bao, Samir Bhatt, Pete Gething, Jeff Eaton, UNAIDS, US Government, Secretariat*

Recommendations for geospatial modelling:

- ✓ Establish timeline for incorporating ART, migration, estimating incidence and schedule for country visits.
- ✓ Data for catchment patterns – rural (ALPHA), urban (TBD, leDEA have information on where people have come for services)
- ✓ Migration is a priority consideration
Follow-up: Samir Bhatt, Pete Gething, review progress with Subnational Working Group

Recommendations for Le Bao:

- ✓ Repeat analysis on multiple countries; confirm guidance on appropriate weight of pseudo-sites and general guidance for use. *Follow-up: Le Bao, review August 2016*
- ✓ Conduct analyses on the same countries as geospatial model development
- ✓ Continue development of this work in parallel to geospatial model but with increased integration over the longer term
- ✓ Application of the method will depend on the outcome of the GPL discussion – ideally integrated in EPP but potential for manual pre-processing. *Follow-up: Review outcome of GPL discussion, Tim Brown, UNAIDS, Avenir Health, Le Bao, Secretariat*

Recommendations for organisations:

- ✓ WHO/UNAIDS to focus on SNU-1, PEPFAR will focus on SNU-2
- ✓ PEPFAR to support spreadsheet disaggregation in the same spirit of continuity where appropriate
- ✓ UNAIDS to incorporate spreadsheet disaggregation at regional workshops *Follow-up: UNAIDS, guidance document end Dec 2016*

Recommendation for research: Identify how to best utilise the results from subnational estimates to inform programming – for countries, PEPFAR, Global Fund. *Follow-up: To be addressed at HIV Modelling Consortium meeting on HIV Models to Inform Programme Planning, July 2016*

HIV Data Repository

There is an increased demand for more granular information to inform HIV programme planning, and there is a need for these data in order to produce robust estimates at sub-national levels. A lot of data already exists, and more is on the way. These data are not presently available or accessible in one place which limits their utility. The aim is to create a publicly accessible data archive containing data about HIV accessible through a web interface.

Scope: General population data in sub-Saharan Africa (in first instance, but infrastructure to include all countries and key populations). Raw data, at lowest level available, with geographic location

Ownership: UNAIDS to house, collaboration of MoH, WHO, PEPFAR, CDC, US Census Bureau. Initial set-up by Secretariat at Imperial and University of Oxford

Access: Cloud based, registration required, secure log-ins and access permissions at different levels

Resources: To be identified

Next steps include:

1. Consultation of stakeholders
2. Assembling of existing surveillance data (national surveys, ANC)
3. Create data archive infrastructure (interface, back end)
4. Pilot countries (2-3) for intensive data curation
5. Commission tool for routine extraction of DHIS2 indicators

It was discussed that the US Census Bureau has been digitising its HIV database which may also be incorporated. From the country perspective, the historical data (raw) may not be readily available. There are also ongoing confidentiality issues with regards to making data publicly available. The concern was also raised that the data repository had ballooned from an original aim to curate ANC data to a comprehensive effort to collate all HIV data which may be an overwhelming aim.

The Repository of Open Access Data for the Malaria Atlas Project (ROAD-MAP) presented a similar type of data warehouse that is used for malaria. It provides a flexible foundation to store many different type of data at the lowest level possible, including geographic coordinates. These data are used to inform the malaria mapping work conducted by Pete Gething's group at the University of Oxford. The data are protected with varying levels of user access.

Recommendations for the data repository:

- ✓ **Establish funding available** *Follow-up: UNAIDS, US Government*
- ✓ **Identify value proposition to countries**
- ✓ **Learn from malaria database design and aim to replicate in similar fashion for HIV**
- ✓ **Store data at lowest level possible; implement user access and permissions**
- ✓ **Coordinate with US Census Bureau with regards to archived electronic data**
Follow-up: Secretariat to steer this forward, summer 2016

V. Development of an age-structured model

EPP fits observed ANC/PMTCT and survey data (generally 15-49 years) to provide prevalence and incidence trends for Spectrum. The different age and sex structures in EPP and Spectrum produce divergence over time for the same incidence input. Jeff Eaton is developing methods to incorporate age-structure into EPP. This will provide age-specific estimates, allow for improved consistency between EPP and AIM, and will account for selection biases in ANC data.

Fertility trends among HIV-positive women are important for estimates because there is a need to, 1) Plan for PMTCT services needed and evaluate coverage and, 2) Understand how to generalise prevalence trends among pregnant women to general population prevalence. Jeff Eaton previously proposed a new model to represent effects of HIV on fertility³. The advantage of this approach is that it captures the effects of ART on the relationship between HIV and fertility which removes the need for ad-hoc parameter adjustments in the ART era. The model predicts a narrowing (but not removal) of fertility differences over time between HIV-negative and HIV-positive women, in line with data from the ALPHA Network. However, this approach was not implemented because it exacerbated the problem of having estimates of >100% PMTCT coverage. This work has now been further extended with two main aims:

1. *Directly estimate subfertility by stage of infection (Markov modulate Poisson process model)*
2. *Investigate reason for differences in subfertility estimates between DHS (Chen and Walker) and APHA network – do survivorship and misclassification biases in DHS analysis explain attenuated subfertility estimates?*

To address the first aim, the FRRs of women in each CD4 category is estimated relative to those with CD4 >500. This is done by calculating the probability of CD4 stage over time using individual HIV status, survival information and Spectrum CD4 progression parameters (Markov model). Observed fertility events are used to estimate the fertility rate in each stage based on probable state distribution (Poisson process). This approach directly estimates subfertility by age and duration of infection from cohort data (ALPHA Network) and provides robust uncertainty. One additional run will be made with the final dataset.

To address the second aim, a comparison of pre-ART ALPHA Network cohort data suggests an 18% greater fertility reduction compared to Chen and Walker's⁴ estimates from DHS. This resulted in the new model predicting lower estimates of need for PMTCT which exacerbated the issue of >100% PMTCT coverage. However, there are biases in sub-fertility estimates from DHS which represent cross-sectional surveys, observe all births over the past three years, and calculate the FRRs assuming the HIV status measured at the survey was the HIV status measured over the past three years. Women who died during this period will not be captured (survivorship bias), which systematically omits those with lowest fertility. In addition, for those who seroconvert during this time period, some HIV-negative person time will be assumed as HIV-positive person time (misclassification bias).

The ALPHA Network cohort data can be used to simulate a cross-sectional survey analysis. This allows estimation of age-specific FRRs with a 3-year retrospective period, replicating the Chen and

³ Method development for the UNAIDS estimates: October 2015. Report and recommendations from a meeting of the UNAIDS Reference Group on Estimates, Modelling and Projections held in London, UK, 26-27 October 2015. <http://www.epidem.org/resources>

⁴ Chen WJ, Walker N. Fertility of HIV-infected women: insights from Demographic and Health Surveys. Sex Transm Infect. 2010;86 Suppl 2:ii22-7.

Walker analysis. The results indicate that these biases explain half (9.5%) of the difference between the estimates. However, this bias is reduced in the more recent period (2006-2011) as a result of fewer deaths (less survivorship bias in more recent period).

Overall, this model predicts that with everyone on ART, fertility of HIV-positive women will remain approximately 30% lower than HIV-negative women. It is not clear whether there is any evidence in support or against this. Updated data are still urgently needed regarding the fertility of HIV-positive women the ART era. Data are also needed from more geographically representative areas including urban areas.

Development plan for the age-structured model

There is a two-phase plan for the development plan for the age-structured model:

Phase I: Incorporate demographic structure into EPP

- This will enable accounting for biases in ANC and PMTCT prevalence
- Internally consistent incidence, prevalence and mortality to pass to Spectrum

Phase II: Incorporate age-structured data into epidemic estimation

1. Incorporate age-specific **HIV prevalence data**, estimate age/sex incidence rate ratios (IRRs)
 - Aim: Estimation of IRRs by age/sex will provide better match to age/sex-specific prevalence
2. Incorporate age/sex-specific **adult mortality data** into likelihood-based inference
 - Aim: Internally consistent HIV estimates and all-cause mortality
3. Estimate **changes in age/sex pattern of HIV incidence** over time
 - Hypothesis: Greater concentration of HIV incidence in younger ages early in the epidemic
4. Estimate **HIV care cascade**; jointly synthesise data from PHIA and programme data

Producing demographic files for subnational estimates

Subnational HIV estimates require underlying demographic projections at the subnational level. The US Census Bureau has supported UNAIDS in the development of these files. Births, deaths and migration are the key components, and migration (international and internal) remains a key issue. For Spectrum, demographic projections are needed from 1970, but it is difficult to get subnational information by age and sex back this far. Geographical boundaries will also have changed over time. The overall strategy to produce base populations is to use the earliest census data by subnational area (age and sex), adjust the population to reconcile with the Spectrum 1970, population and further adjust to match to current census data. Consistency is maintained between the national and subnational total fertility rates (TFR) and mortality rates.

Migration is generally the most difficult aspect of population change to estimate and project. Internal migration is often poorly captured and is critical for appropriately capturing population growth and change over time by age and sex at the subnational level. Demographic changes in age-structure over time hinge on this issue which has great importance particularly for age-structured modelling. The age-structured model requires demographic inputs at the level EPP is being run, i.e. urban/rural or subnational.

It was discussed that Spectrum, and most cohort component projection models, assume entering migrant populations take on the same characteristics of the area where they migrate to - for

example, people who move from a low prevalence area to a high prevalence area are assumed to take on characteristics of high prevalence area, have the same fertility rates, HIV prevalence, etc. It was agreed that the best information to use is two censuses, and to then supplement with survey information. International migration is often used to “fix” inconsistencies between successive censuses. Putting all discrepancies into migration is a strong assumption; it assumes mortality and fertility are correct. Data from PHIA will be able to provide more information about populations at both national and subnational level.

It was discussed that the age-structured model will be applicable for generalised epidemics in the first instance, but this alone will require approximately 40 demographic files produced at the level at which countries generate their estimates. Countries will increasingly aim to produce subnational estimates and thus fitting should focus at this level as opposed to urban/rural fits.

Recommendations for age-structured model development:

- ✓ **Pilot estimation of sub-fertility by age and stage of infection model in Spectrum, review results and compare with data in PMTCT.** *Follow-up: Jeff Eaton, Tim Brown, John Stover*
- ✓ **Jeff Eaton to produce code in appropriate format for Tim Brown**
- ✓ **Maintain all models in EPP in order to have flexibility to move between models**
- ✓ **Peter Johnson and Tim Fowler to continue to develop demographic models to support the development of an age-structured model – Lesotho and Tanzania in the first instance.** *Follow-up: Peter Johnson/Tim Fowler, Jeff Eaton, test age-structured model with use of these demographic projections for presentation at Fall Reference Group meeting*
- ✓ **Prototype coded, tested in countries where demographic data have been provided (Lesotho, Tanzania). Assess performance of model with the view that countries could try to use this at the workshops in early 2017.** *Follow-up: Jeff Eaton, Tim Brown, John Stover, review progress end July 2016 and end August 2016. Presentation at Fall Meeting, Sept 2016*

VI. Use of routine surveillance data in EPP fitting

Countries are currently transitioning from surveillance in ANC to PMTCT. Le Bao and Ben Sheng are investigating how to accommodate this transition in EPP. Incorporating PMTCT data results in an unbalanced longitudinal design, whereby ANC data are used for the historical period and PMTCT data are used for the current period. Analyses which use synthetic data to estimate the level of continuity required illustrate that higher data continuity improves the ability to estimate the calibration parameter (β) and to recreate the underlying prevalence trend. The next steps are to investigate different calibration parameters (use of different β s), define the minimum data continuity required and derive the informative prior distribution of β from real data.

Data from Zimbabwe, where there are overlapping years of prevalence data from ANC and PMTCT surveillance, were used to estimate the systematic difference of β . The data were fit independently by assuming a shared time trend between ANC and PMTCT and a fixed β across time. A key issue, encountered in this analysis but relevant for other countries, is that ANC data are collected at site level while PMTCT data are often collected at the census level (provincial level in Zimbabwe). This raises the question of how to handle PMTCT collected at census level in EPP fitting. The results from this approach using PMTCT data at census level indicate that while there appears some spatial variation across provinces, β is usually near 0 indicating it may be reasonable to assume there is no bias between ANC and PMTCT in Zimbabwe. The next steps are to estimate β using the EPP model, using data from Zimbabwe in the first instance and expanding to other countries as soon as datasets are obtained, to then add the variance inflation to the residual term, and to test informative prior distributions for β .

It was discussed that initially, counties will likely use only the routine surveillance data (PMTCT) from sentinel sites and will simply input these data into EPP as a continuation of the ANC trend, thus this analysis should also be investigated in more detail. Additional questions raised for consideration include:

- *If you allow β to vary across sites/areas, will this result in a different conclusion?*
- *How sensitive is β to prior – if you vary the dispersion of the prior, what is the distribution of β and is it sensitive to this?*
- *Are women who know their HIV status appropriately captured in routine surveillance data?*
- *How much of an overlap is needed if countries simply continue with the same sentinel sites in EPP but instead input PMTCT data?*
- *Should the sample sizes of data from PMTCT be modified? Should PMTCT be added as a different site?*
- *Should a variance inflation term be added for PMTCT data given that the sampling error will likely be small with the large sample sizes of these data?*

Jacob Dee indicated he is available to help with identifying countries with multiple years of routine surveillance data (PMTCT). However, the quality of these data is often variable. Kenya, for example, has multiple years of overlapping PMTCT data, but the data from early years are not trusted thus are not used (no overlap). It is also anticipated that many countries will switch immediately to PMTCT and will not have any overlap. Finally, while PMTCT data may be of poorer quality in past years, ANC sentinel surveillance will also be affected by the EIA issue discussed previously. Thus the biases in historical trends and the future trends will differ.

It was agreed that EPP will need to be able to accommodate continuation of data collection from sentinel sites, but also very large samples of aggregate data at different levels. It will be important to provide guidance to countries as soon as possible with regards to how to prepare their PMTCT data for inclusion in EPP. Further, in order to develop the interface in EPP, it is important to understand the data that countries currently have and how they will need to be entered.

Recommendations:

- ✓ **Generate recommendations for both scenarios – countries with site-level PMTCT and those with census-level PMTCT data**
- ✓ **For countries with site-level PMTCT: Generate recommendations regarding how the data should compare to ANC in order to be implemented, and how to implement in EPP (same site, different site, modification of sample size)**
- ✓ **Consider incorporating a pilot workshop to test the use of both approaches just before the Fall Reference Group Meeting, Sept 2016**

Follow-up: Ben Sheng, Le Bao with support from Working Group. Review progress in mid-July

Case-based surveillance

Both PEPFAR and WHO are supporting the development and expansion of case-based surveillance. WHO will release guidance for HIV case reporting and patient monitoring in December 2016. Haiti is one of the initial countries receiving support from PEPFAR and will soon produce their first national report (June 2016). Key issues identified in Haiti include matching (name-based surveillance and names not matching perfectly which can result in double-counting), mortality (absence of a strong vital registration system) and how to handle inconsistencies with other estimates. Discrepancies between case-based surveillance and UNAIDS estimates present a challenge for programme planning.

Pilot projects are underway in Kenya, Ethiopia and Senegal and protocols have been developed in Tanzania, Malawi and Uganda. Rwanda, Namibia and Mozambique are in pre-protocol discussions. Countries will want to take advantage of case-based surveillance to inform Spectrum estimates in the near future. It is not clear at present how these data will be incorporated.

Appendix I: List of Participants

Nina Anderegg

ISPM Bern
Bern, Switzerland

Giorgos Bakoyannis

Indiana University
Bloomington, IN, USA

Le Bao

Penn State
State College, Pennsylvania, USA

Samir Bhatt

Imperial College London
London, UK

Tim Brown

East-West Center,
Honolulu, USA

Txema Calleja

WHO
Geneva, Switzerland

Kelsey Case

Imperial College London
London, UK

Noela Chicuecue

Ministry of Health
Maputo, Mozambique

Anindya De

US Centers for Disease Control & Prevention
Atlanta, USA

Jacob Dee

US Centers for Disease Control & Prevention
Atlanta, USA

Peter Diggle

Lancaster University
Lancaster, UK

Jeff Eaton

Imperial College London
London, UK

Joy Fishel

ICF International
Washington DC, USA

Tim Fowler

US Census Bureau
Washington DC, USA

Geoff Garnett

Bill & Melinda Gates Foundation
Seattle, WA, USA

Peter Gething

University of Oxford
Oxford, UK

Peter Ghys

UNAIDS
Geneva, Switzerland

Simon Gregson

Imperial College London
London, UK

Steve Gutreuter

US Centers for Disease Control & Prevention
Atlanta, USA

Tim Hallett

Imperial College London
London, UK

Kelsey Grantham

Burnet
Sydney, Australia

James Hargreaves

London School of Hygiene & Tropical Medicine
London, UK

Dan Hogan

WHO
Geneva, Switzerland

Mehran Hooseini

The Global Fund
Geneva, Switzerland

Eline Korenromp

Futures Institute
Geneva, Switzerland

Guy Mahiane

Futures Institute
Glastonbury, USA

Mary Mahy

UNAIDS
Geneva, Switzerland

Kim Marsh

UNAIDS
Geneva, Switzerland

Colin Mathers

WHO
Geneva, Switzerland

Mutsa Mhangara

Ministry of Health and Child Welfare
Harare, Zimbabwe

Michelle Morrison

Bill & Melinda Gates Foundation
Seattle, WA, USA

Eby Pascal

UNAIDS
Geneva, Switzerland

Chantal Quinten

ECDC
Stockholm, Sweden

Anastasia Pharris

ECDC
Stockholm, Sweden

Brian Rice

London School of Hygiene & Tropical Medicine
London, UK

Soraya Rusmaully

Imperial College London
London, UK

Keith Sabin

UNAIDS
Geneva, Switzerland

Miriam Sabin

The Global Fund
Geneva, Switzerland

Ben Sheng

Penn State
State College, Pennsylvania, USA

Ray Shiraishi

US Centers for Disease Control & Prevention
Atlanta, USA

John Stover

Futures Institute
Glastonbury, CT, USA

Mike Thorn

University of Oxford
Oxford, UK

Haidong Wang

Institute for Health Metrics and Evaluation
Seattle, WA, USA

Kate Wilson

Imperial College London
London, UK

Constantin Yiannoutsos

Indiana University
Bloomington, IN, USA

Basia Zaba

London School of Hygiene & Tropical Medicine
London, UK

Appendix II: Meeting Agenda

Method Development for the UNAIDS Estimates: May 2016

16-19 May 2016

DAY 1: Monday, May 16th: Pre-Meeting on Key issues arising in the 2016 UNAIDS Estimates

Start	Duration		
1230	10	Opening	Tim Hallett & Simon Gregson, Imperial College London
Pre-meeting: Key issues arising in the current round of UNAIDS estimates (Chair: Tim Hallett)			
1240	35	Issues arising from current round of UNAIDS estimates - Brief overview of results - Paediatric estimates - Adjustment for pregnant women vs general population - Estimates obtained using the incidence fitting tool (fitting to match mortality)	UNAIDS
1315	20	Country perspective: Key issues in the estimates process in Mozambique	Noela Chicuecue, MoH Mozambique
1335	20	Group discussion	ALL
1355	30	Software Development: EPP - Incorporation of AEM into Spectrum - GLMM package and EPP (and GNU licence issue) - ANC variance inflation	Tim Brown, East-West Center
1425	35	Future considerations for EPP/Spectrum software - Discussion on consideration of open source code - Discussion on move to cloud-based software	ALL
1500	30	Coffee break	-
1530	45	Software Development: Spectrum - Development update from John Stover - STI estimation module, Eline Korenromp - Updates on the fit to programme data tool	John Stover, Avenir Health Eline Korenromp, Avenir Health Guy Mahiane, Avenir Health
1615	55	Estimates in high-income countries - Use of Spectrum to generate estimates in 45 high-income countries - ECDC estimates compared to estimates from the case-reporting tool - Incorporation of ECDC estimates into Spectrum - Group discussion	Kelsey Grantham, Burnet Institute Anastasia Pharris & Chantal Quinten, ECDC
1710	20	Final discussion and review of consensus recommendations	ALL
1730		Close	

Day 2: Tuesday, May 17th: Subnational estimates meeting

Start	Duration	Subject	Speaker
900	15	Meeting overview and aims	Mary Mahy & Peter Ghys, UNAIDS Tim Hallett & Simon Gregson, Imperial College London
915	5	Introductions	ALL
920	15	PEPFAR perspective on the global estimates process	Jacob Dee, CDC on behalf of Irum Zaidi
Session 1: Spatially-specific estimates of HIV (Chair: Tim Hallett)			
935	60	Oxford geostatistical model - Overall framework, road map - Description of data used and methodology - Results for South Africa and Mozambique - Description and explanation of covariates - Plan for inclusion of ART data - Plan for estimating incidence	Pete Gething, University of Oxford Samir Bhatt, Imperial College London
1035	25	Discussion	ALL
1100	30	<i>Coffee break</i>	-
1130	25	Hierarchical approach for generating spatially-specific estimates: Results from implementation across multiple countries	Le Bao, Penn State
1155	20	Use of urban vs rural ANC clinics in Zimbabwe - Changes in prevalence as a result of patterns of service use - Characteristics of pregnant women using local services vs those going elsewhere - Road map for further analyses: Size/shape of clinic draw; extension of Montana et al analysis; mapping of high/low prevalence areas	Kate Wilson, Imperial College London
1215	45	Discussion and recommendations	ALL
1300	60	<i>Lunch</i>	-
Session 2: Discussion on other sub-national estimation approaches (Chair: Tim Hallett)			
1400	15	Spectrum sub-national files: Brief overview of methods used to generate sub-national estimates using Spectrum	John Stover, Avenir Health
1415	25	Questions and discussion, implications for geostatistical model	ALL

DAY 2 (cntd)

1440	30	USG small-area estimates - <i>Methods</i> - <i>Results from countries</i> - <i>Discussion and implications for geostatistical model</i>	Steve Gutreuter/Ray Shiraishi, CDC
1510	30	<i>Coffee break</i>	-
1540	45	Group Discussion: - Recommendations for subnational estimates - Next steps for collaboration - Communication strategies for results from different approaches	ALL
Session 3: Data repository (Chair: Mary Mahy)			
1625	60	Discussion on data repository - <i>Scope: Review and discussion on data to include, geographic regions to include</i> - <i>Ownership: Discussion on where/how to house</i> - <i>Access: Mirroring country data systems, access</i> - <i>Resources: Discussion on resources available/required</i> - <i>Example of ROADMAP for malaria</i>	Discussion led by Mary Mahy, UNAIDS and Jeff Eaton, Imperial College London Presentation on ROADMAP, Mike Thorn, University of Oxford
1725		Close	

Day 3: Wednesday, May 18th: Ongoing development of EPP/Spectrum

Start	Duration	Subject	Speaker
900	10	Opening and further introductions	
Session 1: Ongoing development of Spectrum and EPP - Updates from MESH, UNAIDS & ICF Intl (Chair: Geoff Garnett)			
910	20	MESH Consortium update: Update on key streams of work related to UNAIDS Reference Group core themes, deliverables and timelines	Brian Rice, LSHTM
930	15	UNAIDS Update: Summary of new UNAIDS Estimates, implications from changes made to EPP/Spectrum	Mary Mahy, UNAIDS
945	20	Questions and discussion	-
1005	15	Update on Population-based HIV impact assessments (PHIAs)	Jacob Dee, CDC
1020	25	ICF International Update: Performance of EIAs for HIV serology in Demographic and Health Surveys	Joy Fishel, ICF International
1045	30	<i>Coffee break</i>	-
1115	25	Discussion and Recommendations	ALL

DAY 3 (cntd)

Session 1 cntd: Ongoing development of Spectrum and EPP (Chair: Josh Salomon)			
1140	50	EPP development: Incorporation of incidence assays in EPP - <i>Current implementation in EPP (Tim Brown)</i> - <i>Incorporation of cohort-based incidence measurement in Rwanda & Swaziland (Mary Mahy)</i> - <i>Accounting for survey design and sampling weights in the likelihood for incidence assays (Le Bao)</i> - <i>Future potential uses of incidence assays and future considerations</i> - Recommendations	Tim Brown, East-West Centre Mary Mahy, UNAIDS Le Bao, Penn State
1230	60	<i>Lunch</i>	-
1330	30	Spectrum development: Multi-state Markov models for estimation of mortality and disengagement from care	Giorgos Bakoyannis, Indiana University
1400	25	Discussion and Recommendations	
Session 1 cntd: Ongoing development of Spectrum and EPP - HIV estimates from GBD compared to UNAIDS (Chair: Tim Hallett)			
1425	30	IHME Update: Modifications to methods used, new results from Global Burden of Disease 2015	Haidong Wang, IHME
1455	15	Clarifying questions	-
1510	30	<i>Coffee break</i>	
1540	20	Global Burden of Disease estimates compared to UNAIDS : New estimates from GBD compared to new UNAIDS estimates	Kim Marsh, UNAIDS
1600	40	Discussion and Recommendations	ALL
1640	35	Review of consensus recommendations, follow-up items, timeline	ALL
1715		Close	

Day 4: Thursday, May 19th

Start	Duration	Subject	Speaker
900	10	Opening remarks	Tim Hallett, Imperial College London
Session 2: Development of an age-structured model (Chair: Tim Hallett)			
910	30	Age-structured model development: Update on the sub-fertility by age and stage of infection model, next steps, timeline	Jeff Eaton, Imperial College London
940	20	Implementation of age-structured model in EPP: Progress to date, key issues, next steps	Tim Brown, East-West Centre
1000	25	Questions & Discussion	ALL
1025	20	Age-structured model development: Producing sub-national demographic data files for an age-structured model	Tim Fowler, Census Bureau
1045	30	<i>Coffee break</i>	-
1115	30	Questions, group discussion, recommendations, timeline	ALL
Session 3: Use of programme service data in EPP fitting (Chair: Simon Gregson)			
1145	5	Opening remarks from MESH Consortium re use of programme service data	Brian Rice, LSHTM
1150	10	Update on case-based surveillance, WHO perspective	Txema Calleja, WHO
1200	10	Update on case-based surveillance, USG perspective	Jacob Dee, CDC
1210	20	Incorporating the transition from surveillance in ANC to surveillance in PMTCT in EPP fitting	Ben Sheng, Penn State
1230	30	Group discussion & recommendations	ALL
1300	75	<i>Lunch</i>	-
Closing session (Chair: Tim Hallett)			
1415	60	Review of consensus recommendations, follow-up items, future and ongoing work plans and timelines	ALL
1515		Close	