Method Development for the UNAIDS Estimates: October 2015

Report and recommendations from a meeting of the
UNAIDS Reference Group on Estimates, Modelling and Projections
London, UK, 26-27 October 2015

REPORT & RECOMMENDATIONS

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

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, November 2015

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:
    - Age-structured models
    - Use of case report and mortality data
    - Use of programme service data
    - Spatially-specific estimates of HIV

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.

Report and Recommendations

The UNAIDS Reference Group on Estimates, Modelling and Projections provides guidance and recommendations to UNAIDS and other partners on estimates and projections of the prevalence, incidence and impact of HIV/AIDS. This includes ongoing updates and improvements to the tools currently used – the Estimation and Projection Package (EPP) and the AIDS Impact Module (AIM) in Spectrum – and new method development which can be categorised into four main theme areas:

- Age-structured models
- Use of programme service data
- Use of case report and mortality data
- Spatially-specific estimates of HIV

The work of the Reference Group will occur in coordination with other groups including the Measurement and Surveillance of HIV Epidemics (MESH) Consortium. The MESH Consortium has four working groups which will focus on the following areas:

1. Measuring HIV-related mortality
    - Mortality on ART: Lost to follow-up (LTFU) is the key focus with a systematic review (to be presented at CROI, February 2016) and analyses of IeDEA data
    - HIV attributable mortality: Cause of death data from ALPHA; HIV-negative mortality to provide a counterfactual in absence of ART from ALPHA
2. Routine case-based surveillance
    - Initial assessment of patient information and surveillance systems
    - Recommendations for pilot activities
3. Key populations
    - Size and HIV epidemic dynamics: Network-based vs venue-based vs programme data to produce national estimates
    - Stigma metrics
4. Support for guideline development and dissemination:
    - Mapping of existing and planned guidelines for HIV measurement and surveillance
    - Support for further guideline development

The work of the MESH Consortium focuses on identifying the data that are available and how to best utilise these data while the UNAIDS Reference Group focuses on improving estimates using the data available. The two groups will continue to coordinate to support these aims.

I. Updates and improvement for EPP and Spectrum

EPP

In the 2014 Estimates, there were inconsistencies between EPP and AIM in Spectrum which resulted in differences in prevalence and subsequent discontinuities in incidence. These inconsistencies included different implementation of HIV progression rates, ART eligibility, and ageing out at age 50. To address these issues, the progression rates have been updated in EPP and are now in sync with AIM, and a database of age-outs has been implemented which will pass the number of 50 year olds living with HIV (PLHIV) by ART status from AIM to EPP (in the same manner as new entrants at age 15). Eligibility for ART is also now in sync (it was off by one year compared to AIM). The effect of these changes is there is now a much closer agreement in prevalence and incidence. The prevalence adjustment is required as small adjustments are still needed. EPP will be further updated to capture special populations eligible for ART regardless of CD4, in sync with eligibility in AIM, and will be updated to accommodate data inputted by countries for those on ART by CD4 compartment. These modifications will result in invalidation of curve fits when old files are opened in the new software.

A further consideration is to move from passing numbers to rates among PLHIV for HIV-positive 15 year-olds entering and HIV-positive 50 year-olds exiting. This will better support apportionment in concentrated epidemics and subnational projections where it is more complex to distribute entry and exits. It was also discussed that countries may want to input specific ART scale-up targets for different key populations; it might be useful to consider if this feature can be added in the future.

Recommendations for EPP:

- ✓ **Implement ART eligibility for special populations and ART distribution by CD4 in EPP.** *Follow-up: East-West Center, Dec 2015*
- ✓ **Investigate use of rates for entry/exit at age 15/50 among subpopulations.** *Follow-up: East-West Center, Avenir Health, Jan 2016*
- ✓ **Add guidance regarding inclusion of subpopulations with low prevalence (i.e. 0%), or small samples sizes and low prevalence in EPP (see below).** *Follow-up: East-West Center to incorporate in EPP help file, UNAIDS to incorporate in country guidance documents, Jan 2016*

Spectrum

Incidence menu: There are now many options for generating incidence in Spectrum – using EPP, AEM, the *fit to programme data* tool and a direct incidence input. Users must now define where incidence is being read from in the incidence menu. This menu also now includes two options for EPP populations – fitting to ages 15-49 (as usual) or 15+ (used in China). The EPP prevalence adjustment is found here and it is recommended to keep the adjustment on for UNAIDS Estimates.

WPP 2015: The UN Population Division has published the World Population Prospects (WPP) 2015 which will be updated in Spectrum. Globally, the change in the world population is a 25 million increase with notable changes occurring in India, China, Egypt, Eritrea and Moldova among other countries. The UN Pop Division is now providing the no-AIDS mortality projections for 21 high HIV-burden countries, which will remove the need for Spectrum to create these projections (estimated by matching to UN Pop Division all-cause mortality). Due to the different estimation cycles, the WPP2015 projections will continue to be one year out of sync with Spectrum which is unavoidable at present. This will result in small differences in all-cause mortality in Spectrum compared to

WPP2015. For the remaining 22 countries with lower HIV-burdens, Avenir Health will test whether it is necessary to estimate no-AIDS mortality projection files.

It was discussed that WPP2015 came out at the same time as the 2014 UNAIDS estimates, thus differences can be expected between the two particularly for countries which had substantial differences in HIV prevalence in the 2014 estimates. There may also be differences between UN Pop Division and national census data, which may be further exacerbated at the subnational level. Countries should continue to adjust the population data (total population) to match census data where appropriate.

Re-estimation of child survival with HIV: Milly Marston reviewed the available data for survival of children infected at older ages and used these data to update child survival patterns. Data from haemophiliacs (from Europe, US and Australia) infected at ages 5-14 years illustrated that survival was higher compared to young adults infected with HIV at ages 15-19 (currently used to inform child survival past ages 2-3 years). Re-estimating the child survival curves to fit to these new data results in longer survival; implementation of the new patterns will result in an overall increase of over 150,000 children living with HIV (2014).

It was discussed that from a programme perspective, it is problematic to have changes in child estimates due to model parameter changes, particularly with all the investments to reduce MTCT. In addition, countries still cannot find the children they are estimated to have already, let alone additional children. It was further discussed that the data from haemophiliacs are not necessarily representative of sub-Saharan Africa where the majority of children infected with HIV are.

Adolescent indicators: Adolescent indicators can be displayed in Spectrum but have been hidden due to uncertainty surrounding their validity. These outputs are still questionable, but there is often strong demand for these outputs (UNICEF, WHO). The ALPHA Network may have opportunities to investigate long-term non-progression vs recent infection in adolescents.

Recommendations for Spectrum:

- ✓ **Use the UN Pop Division trends in no-AIDS survival, recognising there will be differences in all-cause mortality between Spectrum and WPP 2015.** *Follow-up: Avenir Health*
- ✓ **Encourage countries to compare the population data with national census data and to consider if adjustments are needed.** *Follow-up: UNAIDS and Avenir Health to include in guidance documents*
- ✓ **Do not implement the new child survival curves; wait until more empirical data are available for validation.**
- ✓ **Bring the adolescent indicators issue to the Paediatric Estimates meeting.**
- ✓ **Investigate long term non-progression vs recent infection in ALPHA Network.** *Follow-up: Avenir Health, Basia Zaba*

Joint estimation of CD4 progression and survival implemented in Spectrum

Estimates from Tara Mangal's analysis which jointly estimates CD4 progression and survival has been implemented in Spectrum. The CD4 progression rates are similar for both approaches; however, the jointly estimated patterns vary by region, but do not vary by age while the Spectrum patterns vary by age but not region. Tara found that once differences in initial state probabilities were controlled

for, there were not any differences by age (initial states and mortality still vary by age, but actual progression does not). For mortality in the absence of ART, the jointly estimated rates have lower mortality compared to Spectrum, and across regions the mortality rates for Europe are lower than Africa and Asia. The example of Zambia was used to illustrate that implementation of this approach resulted in similar results for AIDS deaths but greatly improved uncertainty bounds (previously informed by IHME methods and resulted in implausibly wide and often bizarre bounds).

The next steps include modifying Spectrum to accommodate the regional patterns, applying the Africa pattern to East, South and West Africa and assigning a pattern for missing regions. It was discussed that mortality on ART vs not on ART should be investigated to ensure the latter is not lower than the former. In addition, to compare mortality with ART scale-up using the new parameters to observe differences as a result of possibly different CD4 distribution and ART allocation. Further comparisons are also needed for countries in Asia.

Recommendation: Conditionally agree to implement the new patterns in Spectrum and EPP following further comparisons of country results in Asia and the impact of treatment on mortality (ensuring mortality on ART is not greater than ART-naïve mortality). *Follow-up: Tim Brown, John Stover, Nov 2015*

Use of Spectrum to generate estimates of HIV in China

China used a modified version of EPP/Spectrum this year to generate provincial-level estimates. Modifications were made to the model parameters using a patient database of over 500,000 PLHIV. The HIV data systems in China include HIV/AIDS case-reporting and follow-up (individual level data, CD4 testing every 6 months), treatment and care system (ART and regimens, linked to case-reporting and death certification) and annual sentinel surveillance (key population prevalence survey).

There are an estimated 300M internal migrants in China which is particularly important when generating subnational files. Thirty-two Spectrum files were produced and very detailed subpopulations were included. Modifications were made to the parameters for CD4 progression and mortality on/off ART. For CD4 progression, the default pattern in AIM was maintained for higher CD4 counts, but modified (by age and sex) with local data for lower CD4 counts. Mortality in the absence of treatment and on treatment from the Chinese data is generally higher than AIM. The incidence rate ratios in AIM were used. All files were aggregated together to produce national estimates. The results are in the range of estimates produced using other methods with approximately 800,000 PLHIV, 60,000 new infections (which is 10,000 higher than estimated with other methods), and around 28,000 deaths.

The key challenges encountered were estimating the CD4 progression parameters, survival patterns and age and sex patterns of incidence. In China, there is a lack of information for those who have died shortly after diagnosis, and longer survivors provide more information. There is also a lack of information for key populations and there may be a need for different ART uptake and adherence among different subpopulations.

It was highlighted that it may not be appropriate to fit to populations that have 0% prevalence, or separately enter sites with low prevalence subpopulations that have small sample sizes relative to prevalence (e.g., sample size of 200 with prevalence of 0.5%). It is recommended to either aggregate the sites for these subpopulations, reconsider their inclusion in the fitting if they have no prevalence, or consider grouping prefectures/counties into epidemiologically similar regions (e.g. could cluster similar prefectures, for example those with high, medium, and low IDU epidemics, as in Yunnan). It

was also recommended to triangulate different data sources available and consider whether additional data may be able to inform the fitting – for example data from the provincial surveillance systems and study data. A further suggestion was to compare cumulative incidence vs reported cases to ensure there are not more reported cases than estimated. It was discussed that mortality on ART has a very different pattern compared to Spectrum and further investigation is warranted. Further, these data provide an opportunity for comparison with IeDEA Network data.

Guy Mahiane from Avenir Health has used a sample of data from China to estimate HIV progression and mortality rates to better inform the modified parameters in Spectrum. However, the data present many challenges – CD4 at time of seroconversion is unknown, time from infection to diagnosis is also unknown. There may also be biases that arise depending on how individuals fall into care. Recommendations were generated for further development of this work including the establishment of a working group.

Recommendations for China Spectrum estimates:

- ✓ **Compare estimated infections to case reports.**
- ✓ **Reconsider populations included in the fitting (e.g. 0% prevalence).**

  Follow-up: Wei Guo

Recommendations for China data:

- ✓ **Use of the China data to inform ART mortality rates (IeDEA).**
- ✓ **Further investigation into variation in CD4 and mortality by routes of care.**

  Follow-up: China Working Group, IeDEA, MESH

Recommendations for CD4 progression analysis:

- ✓ **Relax assumptions for misclassification of CD4 (allow true CD4 to be higher or lower than observed CD4); relax starting CD4 assumptions; further consideration of methods used to impute time of seroconversion.**
- ✓ **Consider collapsing older age categories to increase sample size which may improve estimates of progression rates and estimates of AIDS-related deaths.**
- ✓ **Work with Tara Mangal on MSM package; consider further insight from Sam Bhatt.**
- ✓ **Avenir Health to convene a working group to further investigate, Guy Mahiane to share summary of the work conducted.**

  Follow-up: Avenir to convene China Working Group, IeDEA, MESH

Analyses of mortality on ART from IeDEA data

The parameters for adult mortality on ART in AIM are informed by data from the IeDEA Consortium based on a 2012 analysis. The methods used in the 2012 analysis were to first model crude mortality then adjust for loss to follow-up (LTFU).

1. **Model crude mortality:** Using a piecewise exponential regression model with constant hazards for intervals (on ART <6 months, 6-12 months, 12-24 months, >24 months). The explanatory variables include baseline CD4, age at start of ART and gender and pre/post 6 months on ART were modelled separately (interaction).
2. **Adjust for LTFU:** Linkage data from South Africa were used to adjust estimates used in southern Africa. For eastern Africa, double sampling data from Kenya were used to inform

the adjustment. There was no further information to inform adjustments in other regions, thus the "inflation factor" from South Africa was used for all other regions.

Nina Anderegg conducted an updated analysis using the more recent IeDEA data available and including additional countries which now have data. The methods of the 2015 analysis were similar:

1. **Model crude mortality:** Same model and approach as previous but only three time intervals (on ART <6 months, 6-12 months, >12months).
2. **Adjust for LTFU:** Again, multiplying crude mortality estimates by inflation factors to obtain the adjusted mortality. The same approach as before was used for Kenya, but a new method was needed for South Africa because the original and LTFU data are now no longer fully distinguishable. A single model was fitted with an additional explanatory variable, "linkable/non-linkable". The inflation factors are the estimates of the effect of this "linkable/non-linkable" factor and a 2-way interaction. These adjusted estimates were calculated with:
    - *Correction 1:* 168 different inflation factors (all interactions)
    - *Correction 2:* 2 inflation factors (no interactions)

The results using *Correction 2* (approximately 1.98 inflation factor >6 months on ART) were more similar (than *Correction 1*) to the adjusted results from the 2012 analysis for South Africa. Applied to West Africa, the results are quite similar compared to the 2012 analysis, but in Asia Pacific, the adjusted mortality rates are quite different compared to the 2012 results, largely due to the additional mortality data now available. It was discussed that the inflation factors obtained from the non-interaction model (*Correction 2*) are similar to those identified in the previous analysis (1.5-2.5) thus in the range applicable for the sub-Saharan Africa region. However, this may not be appropriate for other regions (Latin America, Asia Pacific) where there may be different relationships requiring different inflation factors. It was further discussed that the IeDEA sites in-country may not be representative of what is happening in the rest of the country, let alone other regions. The expectation is that above average care is happening at these sites. Other data available from these regions (for example, China and Brazil) can be considered for comparison with the caveat these routine programme data will be incomplete.

A key question is whether to use this approach if there are significant differences between those linkable vs non-linkable in the South Africa data. Further, use of the South Africa adjustment for other regions does not feel entirely appropriate, but there are no other options at present. Further analyses are needed here which will be conducted as part of the MESH Consortium workplan.

It was recommended that the new analysis should cover the entire time period, as opposed to just the more recent period, and should investigate time trends, dividing the data at the point of rapid ART expansion (around 2009 but to be determined from the data). The implementation of time trends in AIM and EPP (if needed), is unlikely to occur by January 2016, but an update of the parameters currently used could be incorporated if time trends are not required.

Recommendations for the short term:

- ✓ **Review results of Nina Anderegg estimates using correction method 2 based on data from the full time period and compared with differences pre/post ?2009 to better understand the time trend.** *Follow-up: Nina Anderegg Nov 2016*
- ✓ **If time-dependent mortality rates are not necessary, implement the updated parameters for mortality on ART using correction method 2.** *Follow-up: Avenir Health, East-West Centre, Nina Anderegg Nov 2016*

- ✓ **If time dependent mortality rates are required, implement and review after software delivery.** *Follow-up (if needed): Avenir Health, East-West Centre, Nina Anderegg Feb 2016*

Recommendations for long term:

- ✓ **Draw together estimates from large country programmes (Brazil, China) and studies outside IeDEA - ALPHA Network may be able to contribute more sites to these analyses. MESH has begun pooled patient-level database assembly; Reference Group to ask MESH steering committee to extend this project, conditional on evaluation of the quality of the programme data.** *Follow-up: Reference Group,MESH (and MESH Steering Committee), Basia Zaba to discuss with mortality group, Nov 2016*

Joint modelling approach with correction for incomplete death ascertainment

Spectrum has currently implemented adult mortality on ART (as above) which is stratified by factors measured at the start of ART including age, sex and CD4. This approach does not take into account changes in CD4 over time or changes in treatment regimen. As such, Constantin Yiannoutsos has proposed the following updates for adult mortality on treatment in Spectrum:

1. Incorporate longitudinally obtained CD4 count measurements
2. Incorporate information on change of ART regimen
3. Adjust for incomplete death ascertainment (non-random)

A joint modelling framework is proposed, jointly estimating the longitudinal trajectory of CD4 and ART regimen and survival. This approach will be tested with IeDEA data from East Africa. The adjustment for incomplete death ascertainment is informed by double-sampling of dropouts (sample of dropouts that was followed-up to ascertain vital status). The joint models are run on each data set (vital status known, imputed) and the results are synthesised into a single mortality estimate based on the covariate profile of each individual. The longitudinal population profiles take into account baseline data (age, sex, CD4 at start of ART) and most recent CD4 count, time from ART initiation to this CD4 and whether the individual has moved to 2nd line therapy. The longitudinal trajectory is then modelled with a linear mixed model with random effects for intercept and slope of time since ART initiation, while survival is modelled though a Cox proportional hazards model.

It was discussed that the many CD4 categories (used to match to Spectrum) make this work particularly challenging, but can be collapsed to better support this work. However, for implementation in Spectrum it was raised whether this approach could be simplified, perhaps with a rate of change of CD4 over time (i.e. increase). It was also queried whether mortality depends on how an individual got there, and how long it took to get there. In addition, if it is possible to show differences between this approach (disaggregation), and the aggregated averages in Spectrum. Additional suggestions included to directly estimate as a Markov model instead of a linear model and to incorporate viral load as a future predictor of mortality.

Recommendations:

- ✓ **This work to continue and further develop, review progress in early 2016.** *Follow-up: Constantin Yiannoutsos*

✓ **Consider direct estimation of Markov model representing Spectrum structure.**

II. Use of case report and mortality data to inform HIV estimates

Countries that have strong case report and mortality data would like to use this information inform estimates of HIV. There are different ongoing streams of work which aim to incorporate these data sources, including the development of new methods and further refinement of the tools available.

The *fit incidence to programme data* tool

The *fit incidence to programme data* tool in AIM fits an incidence curve to case-reports (adjusted for reporting lag), AIDS deaths and estimates of PLHIV (adjusted for undiagnosed). A key advantage of this approach is that countries with strong programme data (e.g. case-reports and vital registration) are able to directly incorporate this information in the fitting. Uncertainty is now formally captured.

This tool was applied in Brazil and Argentina, two countries with high quality programme data. In Brazil, trends in mortality remain stable in the recent time period, despite numbers on ART continuing to rise. It is possible to fit to the mortality data, PLHIV and estimates of new infections; however, the results for median CD4 count at ART initiation do not match the programme data. In order to match the mortality data and treatment scale-up, it is necessary to maintain high coverage of those at low CD4 counts. This same occurrence was found when applying the tool in Argentina.

It is not clear why this is happening. It may be due to different progression rates in Latin America, but regional variation has not been observed. Mortality may be underestimated but these countries have strong confidence in the mortality estimates. Prevalence may be over-estimated, but again there is confidence in the number of PLHIV.

Estimating incidence from routinely collected programme data in Brazil

Tara Mangal is leading this stream of work which aims to incorporate multiple sources of routinely collected data in order to better inform estimates of HIV incidence. Applied in Brazil, these data include: AIDS cases notifications; reported AIDS deaths; number receiving ART; CD4 counts and viral load. All data have been linked and are available at the individual-level.

Updates to the model include the addition of:

- Age-structure
- Natural mortality from every state (time/age dependent)
- Susceptible population (to fit to prevalence studies)
- Rates of moving to ART change over time (depending on ART availability and coverage)
- Effects of ART on mortality (time/age dependent).

Survival in the absence of ART, CD4 progression and initial state probabilities are informed by the joint estimation work previously conducted. Survival on ART and rates of ART initiation are estimated from the data. The age distribution of new infections and CD4 distributions at time of detection are also estimated from the data. The model is fitting to the number of detected cases, number of deaths and number starting ART.

Key challenges include misreporting of AIDS deaths (which changes over time) and death before diagnosis. Date of diagnosis is not available so detection is the first time an individual appears in a database. Many people first show up in surveillance when they pick up ART. From the data it appears there are two groups starting ART – a group that starts ART very late and a more “normal” group, which is not captured with use of an exponential distribution.

The next steps are to:

- Include those progressing rapidly from detection to ART

- Include changing ART mortality rates over time and by age
- Reconsider the parameterisation of the distribution of CD4 counts at ART initiation to better describe late starters
- Include misreporting/underreporting of AIDS deaths
- Include prevalence studies
- Investigate whether proportion detected at death needs to change over time

Further considerations include time dependence and identifying how much is needed without overcomplicating the model. Additionally whether the age distribution can be kept constant or whether it needs to change over time. These methods will be tested at an upcoming workshop in Colombia which will also include Argentina and Mexico (Nov 2015).

Recommendations:

- ✓ **Review outcomes from Colombia workshop.** *Follow-up: Reference Group, John Stover, Juan Vesga, Tara Mangal, David Wilson, Dec 2015*
- ✓ **Compare Tara's results from Brazil to the results from the fit to programme data tool.** *Follow-up: John Stover, Tara Mangal, Jan 2015*

Update on the use of case-report data to estimate HIV incidence in the ECDC region

The European Centre for Disease for Disease Prevention and Control (ECDC) tool is now available online. Training workshops will occur in March 2016 and include WHO Europe countries. Validation of the estimates produced is occurring in advance of this workshop at a February workshop which will include Belgium, Luxembourg, Portugal, Greece and the Netherlands. Avenir Health and UNAIDS will attend this work shop to compare the Spectrum models with ECDC tool.

Modelling the clinical cascade

There are different global targets that focus on the clinical cascade, including the UNAIDS 90-90-90 targets and the WHO consolidated information guidelines which incorporate the HIV care cascade as part of 10 key global indicators. WHO is organising a workshop 17-20 November in Marrakesh which will bring together 28 countries and over 100 individuals with the aim to apply a simple framework to analyse clinical cascade data to help counties identify gaps and prioritise key interventions.

Jack Olney is developing a modelling framework which can be used to support this process. He has developed a simple deterministic model stratified by CD4 count. Baseline incidence is taken from Spectrum and transmission probabilities are derived based on CD4 count and treatment status (to allow for interventions to have an effect on incidence). Data from AMPATH in Kenya are used to inform diagnosis, linkage and dropout rates, which will be updated with country-specific information where available. Outputs include estimates of the care cascade (including 90-90-90 goals) with projections to 2020 and the identification of potential strategies to achieve future targets.

Six interventions are incorporated and a matrix of intervention combinations has been developed which allows simulation of the best combination of strategies for any given cost. John Stover recommended, from previous experience, to bring regional default cost data to the workshop instead of relying on countries to bring these data. David Wilson indicated that this approach is quite similar to the approach developed within his group and thus there is potential for collaboration and he may be able to further inform the costing estimates from his work.

It was highlighted that from a country perspective there are many challenges to reporting along the clinical cascade. In Zimbabwe, it is difficult to quantify the number of individuals diagnosed and alive

(have overall number of diagnoses), the number on ART, and those with suppressed viral load. It was discussed that the aim of the Marrakesh workshop should be to get the data required. The methods will further evolve depending on what happens at this meeting. There is future scope for linkage with Spectrum, for extension to the subnational level.

Recommendations:

- ✓ **Strong enthusiasm for this work, strongly endorsed by the Reference Group.**
- ✓ **Do work with the countries, identifying data available and further development of the methods.**
- ✓ **Collaboration with David Wilson and sharing of costing estimates.**
- ✓ **Longer term – will identify how to integrate with models used at national level for produce estimates of disease burden and impact.**

  Follow-up: Review outcomes of Marrakesh meeting, end Nov 2015

HIV data collection systems – Haiti and Zimbabwe

In Haiti, CDC has supported the HIV/AIDS reporting system (HASS) since 2008. HASS uses name-based reporting (no unique identifier) and data across different electronic systems are integrated providing information on sentinel events captured across the cascade (CD4, viral load, ART initiation, retention, viral suppression, death). A caveat is that the system does not capture negative tests (so cannot get a proxy for incidence) and there is a fair amount of repeat testing of known positives. There may be opportunities for further engagement with Haiti and exploring how to utilise the data available to better inform HIV estimates in this country.

In Zimbabwe, HIV programmes currently use a paper-based data collection system. With over 1400 facilities and over 800,000 patients on treatment, Zimbabwe is increasingly unable to cope with the burden of this reporting system. There is also a need for better monitoring and tracking. The long-term vision is to move to an electronic system for all health data. In the short term, only HIV and TB data collection will move to electronic systems. Key challenges include erratic power, need for IT support and training of staff and the workload required to convert paper records into electronic systems. As a result, the paper-based system will continue in parallel to scale-up of the electronic system which is currently underway.

It was discussed that a unique identifier is used for all facilities with electronic medical records, thus switching between electronic and paper-based clinics will result it duplicates. It will also not be possible to prevent a patient from getting two unique identifiers. It was suggested that it would be helpful if this system could also be used for clinical monitoring.

Recommendations:

- ✓ **These important developments in data collection and data systems are greatly welcomed.**
- ✓ **MESH and Reference Group should coordinate their engagement with countries in order to understand data availability, promote collection and develop suitable estimation methods.** *Follow-up: MESH, Reference Group Secretariat*

III. Use of ANC and PMTCT data in EPP fitting

Countries are currently transitioning from surveillance in ANC to PMTCT. Le Bao and Ben Sheng investigated how to accommodate this transition in EPP while Jeff Eaton and Le Bao has investigated effects of accounting for non-sampling error in ANC, work which will have further considerations when PMTCT data are incorporated.

Incorporating PMTCT data into EPP fitting

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. Some continuity across these data sources is needed to accurately estimate the trend; if there is no continuity, it can be difficult to distinguish the trend at the transition and whether it is due to real differences or due to the calibration parameters. The current work aims to investigate the level of continuity required and compare with different levels of data quality. A natural spline approximation was used instead of the EPP model. The model is fit to historical ANC data (Mozambique central region) and model parameters from a sample fit are assumed to describe the underlying prevalence trend. The year 2005 is arbitrarily chosen as the transition year with simulated data used to replace the ANC data pre-2005. Post-2005, PMTCT data are generated with some sites that overlap with ANC. Two data quality parameters are used, the number of overlapping sites (0, 8, 16, out of 100 sites), and the years of PMTCT data (1, 3, 5).

In general, higher data quality, especially the years of PMTCT data (3 and 5 years) improves the ability to recreate the trend and to estimate the calibration parameter while low quality data are problematic. Increasing the number of overlapping sites gave less improvement but this may be due to the limited number of overlapping sites (16/100 overlapping was the maximum tested), or it may be an artefact of the assumption that beta is constant across all sites. Some of the error in prevalence is due to estimation of the calibration parameter (beta).

Next steps include testing with different epidemic trends (increasing, decreasing, stable), further investigating the calibration parameter (use of different betas), deriving the informative prior distribution of beta from real data. A limitation of this work is that it has not been possible to obtain country datasets for testing (despite numerous ongoing attempts). It would also be helpful to view uncertainty around these simulations.

The question was raised whether allowing variation in beta across sites would result in a different conclusion. One suggestion was to use PEPFAR data (available at the site level) to further investigate the calibration parameter (beta) with the caveat that PEPFAR clinics may be different from other clinics. Country representatives expressed that it is still difficult to ascertain where the issues are coming from in the PMTCT data (labs, tests, etc). A further complication, increasing over time, is that HIV+ pregnant women attending PMTCT who know their status may not be captured. This is increasingly an issue and currently not appropriately captured in guidelines.

A PMTCT data entry page has been implemented in EPP but is currently turned “off”. While it is agreed that these are the right methods to move forward, there was concern surrounding implementation of new features in a year that countries do not have the support of regional workshops.

Recommendations:

- ✓ **Further analyses are needed; assemble a small team to continue these discussions. Datasets are needed to support this further development.** *Follow-up: Le Bao, Ben Sheng, Simon Gregson, Mary Mahy, Kim Marsh, Jeff Eaton, Wilford Kirungi, Mutsa Mhangara, Tim Brown , Nov-Dec 2015, Reference Group Secretariat to convene*
- ✓ **Aim for implementation in developmental Spectrum (conditional on data available to support further analyses).** *Follow-up: Le Bao, Ben Sheng, Tim Brown, Mar2016*

Accounting for non-sampling error in ANC prevalence in EPP

There is more variability observed in ANC prevalence than predicted by EPP. This is observed with ANC values falling outside the confidence intervals (CIs) and the incomplete coverage of the CIs which has been formally quantified[1]. Additional error may be due to sampling procedures, poor quality control in labs, contamination, changes in diagnostics. As a result, the model is overly sensitive to spurious trends in ANC prevalence, underestimates epidemic uncertainty (particularly before surveys), and may not give enough importance to higher quality national survey data. Jeff Eaton proposes to address this issue by adding additional variance to account for non-sampling error in ANC prevalence and to allow this additional variance to vary for each EPP fit.

Two methods were considered for estimating this additional variance and applied for 9 countries:

1. Calculate the residual sum of squares between model estimates and observed data and use this residual to identity the extra variance (MLE vs unbiased variance estimator)
2. Estimate the additional variance as a model parameter in EPP fitting

Including this additional variance (using all approaches) resulted in a good fit to survey data and coverage in the out-of-sample prediction increased from 75% to 91%. For the in-sample fit, the unbiased variance estimator and explicitly estimating the additional variance had similar performance while the latter performed slightly better for the out-of-sample prediction. This approach does not affect fitting times.

Explicitly estimating the additional variance can be incorporated into both the r-trend and r-spline models. For Botswana, the effect on prevalence is minimal but results in wider CIs. In contrast, for Kenya incorporating the additional variance (when the equilibrium prior is also included) results in better fits to the survey trend and much larger uncertainty in the historical period. Across all nine countries the uncertainty around estimates in the historical period increases, for both prevalence and incidence, but narrows in the more recent time period.

It was noted that this type of approach will be particularly important when PMTCT data are included in the fitting because the sample size of the data will be huge, with small sampling error but potentially large non-sampling error. Not appropriately capturing the additional error will force the fit to the PMTCT data. It was discussed that including this additional variance results in the timing of the Kenya rural epidemic to increase before the urban epidemic which may be problematic at the country level. There is a concern here from the group about the shift in timings of epidemic (rural Kenya), would want to see more of the rural fits – particularly regarding the timing of the rural and urban epidemics.

Recommendations:

[1] Bao, L., et al. Sex Transm Infect 2012;88:i58–i64. doi:10.1136/sextrans-2012-050689

- ✓ **Implement this approach in a developmental version of Spectrum.** *Follow-up: Jeff Eaton, Tim Brown, Feb-Mar 2016*
- ✓ **Further testing on more countries to review timing of rural/urban epidemics, and with large samples (PMTCT).** *Follow-up: Reference Group secretariat, Feb-Mar 2016*

IV. Spatially-specific estimates

The aim of this work is to estimate HIV epidemics at finer scales and capture epidemic patterns across different areas and among different subpopulations.

Hierarchical prevalence prior for efficiently pooling data across subnational regions

Le Bao has developed a new approach to more efficiently pool information for model HIV epidemics at the subnational and subpopulation level. The approach now used is to apply hierarchical models to the observed data without using EPP, to then incorporate the results into EPP via auxiliary data (by implementing a "ghost site" with variable sample size to control the level of influence). This was done for a generalised epidemic (Nigeria with 37 areas) and for a concentrated epidemic (Thailand with 4 areas and multiple subpopulations). The initial phase uses a simpler model than the CD4 model in EPP in which the prevalence trends are modelled with splines that provide flexibility and relieve the potential of over-fitting. Generalised linear mixed models (GLMM) are useful for analysing data with a hierarchical structure or spatial dependence. This approach is efficient, but ignores the epidemic model thus it only provides inference for HIV prevalence trends. The predictive distribution of HIV prevalence for an area and subpopulation is derived from GLMM as the prevalence prior distribution. Prevalence priors are only defined during the data period. Auxiliary prevalence data (the additional "ghost" site added to EPP surveillance spreadsheet) is introduced to further inform the prior. The GLMM runs outside EPP and then provides the additional site to EPP to inform fitting. The strength of the prior can be adjusted with the sample size of the auxiliary data. The default sample size value will be determined based on prediction accuracy of test datasets in multiple countries.

Results from Nigeria and Thailand indicate that use of auxiliary data (sample size 200-500) can substantially improve the prediction accuracy and will not substantially reduce prediction accuracy. This approach is applicable to all models (classic, spline, trend) and all data sources (prevalence, incidence, deaths). Fitting GLMM takes 5-30 min (with no extra computing costs afterwards), and fewer MCMC iterations can be used at the regional workshops. Other forms of GLMM can also be considered (i.e. spatial dependence)**.** The main limitation is that the surveillance data from all areas and risk groups need to be assembled in advance to run the GLMM and create the auxiliary data.

It was discussed that this new approach has very positive improvement, good practical application and can now realistically be implemented in Spectrum/EPP. Further testing should commence in both concentrated and generalised epidemics (Uganda and Zimbabwe suggested) and countries should have the option to try this approach, particularly for subnational estimates and sparse data settings.

Recommendations:

- ✓ **UNAIDS to support identifying countries to begin testing this approach.** *Follow-up: UNAIDS, Nov/Dec 2015*

- ✓ **Implement in developmental Spectrum, TB & LB to identify how to integrate the auxiliary site, define the user interface and turn on/hide this feature.** *Follow-up: East-West Center, Le Bao, Feb-Mar 2016*
- ✓ **Test in collaboration with countries, mid-2016.**

Estimates at the 2nd subnational level

The US Centers for Disease Control and Prevention (CDC) has generated small-area estimates using survey and auxiliary data whereby an appropriately weighted average of direct and model-based estimates is computed. Fay-Herriot models, commonly used for small-area estimation, were used in this approach. The best predictive model included ANC prevalence and dependency ratio (spatial pattern alone was a poor predictor). Advantages of this approach are that it incorporates uncertainty in survey estimates, it can incorporate spatial variation, in can improve precision, and it produces valid confidence intervals. The limitations are that the estimates are limited to the timing of health surveys and the methods require updated shape files. It was discussed that there are currently many different methods of generating subnational estimates (Spectrum, prevR, Oxford work, and others). This approach is currently an internal USG activity and it is unclear what the future entails. The key benefit is that it provides a good measure of uncertainty.

Recommendation: UNAIDS to liaise with CDC regarding small-area estimation, consider further exploration of these methods, realising different methods exist, do want to have country-led results (prevR not being done at country level). *Follow-up: UNAIDS, CDC*

V. Software development

It was agreed that with many of the substantial changes to the methods used that are anticipated in the next year, a phased development approach will be the best way forward with use of a Developmental EPP/Spectrum for testing and evaluation of new methods before implementation in EPP/Spectrum. This will include:

1. Jeff Eaton's approach for capturing additional variance in ANC data. *Phase I, Feb/March 2016*
2. Le Bao and Ben Sheng approach for incorporating PMTCT data into EPP fitting. *Phase I, Feb/March 2016*
3. Le Bao's GLMM approach and ghost sites. *Phase I, Feb/March 2016*
4. Jeff Eaton's age-structured model. *Phase II, June 2016*

Appendix I: List of Participants

Nina Anderegg
ISPM Bern
Bern, Switzerland

Le Bao*
Penn State
State College, Pennsylvania, USA

Tim Brown
East-West Center,
Honolulu, USA

Txema Calleja
WHO
Geneva, Switzerland

Kelsey Case
Imperial College London
London, UK

Anindya De
US Centers for Disease Control & Prevention
Atlanta, USA

Jacob Dee*
US Centers for Disease Control & Prevention
Atlanta, USA

Jeff Eaton
Imperial College London
London, UK

Liz Feron
London School of Hygiene & Tropical Medicine
London, UK

Peter Ghys
UNAIDS
Geneva, Switzerland

Simon Gregson
Imperial College London
London, UK

Wei Guo
Natl Center for AIDS/STD Control & Prevention
Beijing, China

Tim Hallett
Imperial College London
London, UK

James Hargreaves
London School of Hygiene & Tropical Medicine
London, UK

Peter Johnson
US Census Bureau
Washington DC, USA

Wilford Kirungi
MoH Uganda
Kampala, Uganda

Eline Korenromp
Futures Institute
Geneva, Switzerland

Guy Mahiane
Futures Institute
Glastonbury, USA

Mary Mahy
UNAIDS
Geneva, Switzerland

Tara Mangal
Imperial College London
London, UK

Kim Marsh
UNAIDS
Geneva, Switzerland

Milly Marston*
London School of Hygiene & Tropical Medicine
London, UK

Mutsa Mhangara
Ministry of Health and Child Welfare
Harare, Zimbabwe

Jack Olney
Imperial College London
London, UK

Chantal Quinten
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

Ben Sheng
Penn State
State College, Pennsylvania, USA

Ray Shiraishi
US Centers for Disease Control & Prevention
Atlanta, USA

Emma Slaymaker
London School of Hygiene & Tropical Medicine
London, UK

John Stover
Futures Institute
Glastonbury, CT, USA

Mahesh Swaminathan
US Centers for Disease Control & Prevention
Atlanta, USA

Jim Todd
London School of Hygiene & Tropical Medicine
London, UK

Juan Vesga
Imperial College London
London, UK

David Wilson
University of New South Wales
Sydney, Australia

Basia Zaba
London School of Hygiene & Tropical Medicine
London, UK

Appendix II: Meeting Agenda

Method Development for the UNAIDS Estimates: October 2015
26-27 October 2015

DAY 1: Monday, October 26th

Opening and introductions			
900	10	**Meeting opening - Meeting aims, key issues, introductions**	Peter Ghys, UNAIDS Tim Hallett, Imperial College London
910	10	**Introduction from the MESH Consortium, areas of focus**	James Hargreaves, LSHTM
Session 1: EPP and Spectrum (Chair: Peter Ghys)			
920	25	**EPP Updates and addressing compatibility between EPP and Spectrum**	Tim Brown
945	20	**Spectrum updates**	John Stover, Avenir Health
1005	25	Group discussion and recommendations	ALL
1030	*30*	*Coffee break*	-
Session 2: Mortality in the ART era (Chair: Tim Hallett)			
1100	20	**Updated Spectrum parameters for adult mortality on treatment (by CD4 at ART initiation and, age, sex, duration on treatment), by region, from the IeDEA Network data**	Nina Anderegg
1120	20	**Adult mortality on treatment by current CD4 from the IeDEA Network data , East Africa**	Constantin Yiannoutsos
1140	35	Group discussion and recommendations	ALL
1215	*75*	*Lunch*	-
Session 3: Issues surrounding data from ANC and PMTCT (Chair: Simon Gregson)			
1330	25	**Illustration of effects of non-sampling error in ANC prevalence**	Jeff Eaton, Imperial College London
1355	20	**Incorporating PMTCT data in EPP: Investigation of level of continuity required, data quality requirements**	Ben Sheng, Penn State
1415	20	Group discussion and recommendations	ALL
Session 4: Sub-national estimates (Chair: Tim Hallett)			
1435	20	**Further testing of the hierarchical model** **-** Testing on multiple countries with long time series and in countries where there are divergent trends in different areas. - Testing concentrated epidemic approach in Thailand - Update re investigations to provide guidance on appropriate weighting factor	Le Bao, Penn State
1455	15	**Estimates at the 2nd subnational level - prevR and Spectrum**	Ray Shiraishi, CDC
1510	20	Group discussion	ALL
1530	*30*	*Coffee break*	-

DAY 1 (cntd)

Session 5: Use of case-report and mortality data for estimating HIV incidence (Chair: David Wilson)			
1600	20	**AIM fit to programme data tool – incorporating mean CD4 at diagnosis and work with Brazil and Argentina**	John Stover, Avenir Health
1620	10	**AIM fit to programme data tool – generating uncertainty bounds**	Guy Mahiane, Avenir Health
1630	20	**Model development – use of multiple sources of routine surveillance to generate estimates of incidence in Brazil**	Tara Mangal, Imperial College London
1650	40	Group discussion	ALL
1730		**Close**	-

Day 2: Tuesday, October 27th

Start	Duration	Subject	Speaker
900	10	Opening and introduction	Tim Hallett, Imperial College London
Session 6: Modifications to progression parameters in Spectrum (Chair: Tim Hallett)			
910	15	**Comparison of results in Spectrum using jointly estimated progression parameters**	John Stover, Avenir Health
925	20	Group discussion	
945	25	**Methods used in China for generating estimates in Spectrum**	Guo Wei, NCAIDS
1010	15	**Estimating progression parameters using data from China - results and key challenges**	Guy Mahiane, Avenir Health
1025	20	Clarifying questions	ALL
1045	*30*	*Coffee*	-
1115	30	Group discussion and recommendations	ALL
Session 7: Case-based surveillance and modelling the service cascade (Chair: Tim Brown)			
1145	20	**WHO service cascade modelling**	Jack Olney, Imperial College London
1205	20	**Case-based surveillance in Haiti**	Jacob Dee/Mahesh Swaminathan, CDC
1225	20	**Introduction of electronic medical records in Zimbabwe**	Mutsa Mhangara, MOHCW
1245	30	Group discussion	ALL
1315	*60*	*Lunch*	
Session 8: Updates, recommendations, forward planning (Chair: Mary Mahy)			
1415	10	**Update from ECDC**	Chantal Quinten, ECDC
1425	35	Review of meeting recommendations	Tim Hallett, Imperial College London
1500	30	Next steps and follow-up items	ALL
1530		Close	