Methods for estimation of ART's impact on deaths averted/delayed; and Developments in EPP 2007

Report of a meeting of the UNAIDS Reference Group on Estimates, Modelling and Projections held in Baltimore, USA, July 13th 2007

TECHNICAL REPORT AND RECOMMENDATIONS

Joint United Nations Programme on HIV/AIDS

UNAIDS

UNICEF • UNDP • UNFPA • UNDCP
ILO • UNESCO • WHO • WORLD BANK

UNAIDS

The meeting of the UNAIDS Reference Group on Estimates, Modelling and Projections (the 'Epidemiology Reference Group') was organised for UNAIDS by the UK 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. The recommendations in this document were arrived at through discussion and review by meeting participants and drafted at the meeting.

Dr Peter White, London, October 2007.

Introduction

The Reference Group on Estimates, Modelling and Projections
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 the World Health Organization WHO) 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, Imperial College London www.epidem.org).

Aim of the meeting
The aim of this meeting was to bring together experts to produce recommendations on methods for estimation of ART's impact on deaths averted/delayed; and on developments in EPP that may be required to enable it to reproduce more-complex dynamics of HIV transmission in various countries.

Approach
The meeting featured both presentations of recent data and group discussions, which focused on specific technical issues. Presentations and discussion topics are listed in Appendix I.

The meeting was attended by 17 experts see Appendix II for a list of participants). Each contributed, not only data, insights and analysis, but also worked hard to produce a set of recommendations, drafted at the meeting. We would like to thank them for their hard work and attendance at the meeting.

The recommendations drafted at Reference Group meetings give UNAIDS and WHO guidance on how best to produce estimates of HIV/AIDS, an opportunity to review current approaches and also help 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.

Methods for estimation of ART's impact on deaths averted/delayed

1. Estimating lives saved in by ART programs

There is now considerable interest in estimation of the overall health benefits from different interventions supported through the Global Fund. However, it is still too early to measure them directly. As an interim approach, an attempt was made to consolidate service delivery results for the key interventions into a single measure.

The Global Fund uses 'lives saved' (rather than life-years gained) as its standard measure, as it is easily communicated and is more-easily compared across its three diseases of interest (HIV/AIDS, TB, and malaria), since much malaria mortality occurs in the young – unlike for HIV and TB – so estimating life-years gained leads to large uncertainty. However, 'added years of life due to ART' is also used. An estimated 230,000 lives had been saved by ART, with an estimated 290,000 added years of life due to ART.

For their end-2006 report, WHO and UNAIDS have estimated "life-years gained" on the basis of the number of people on ART in each year, combined with the assumption that people on ART have an annual survival probability of 0.9. The recent changes in assumptions recommended at the December 2006 Reference Group meeting differential survival probability by years on ART; time to ART eligibility of 8 years out of 11 years overall survival) would make a similar approach much more complex.

It was noted that there is potential for confusion if different measures are used by different agencies, and that it is important that robust methods are used. For AIDS, it was suggested that the metric "life-years gained" would be preferable to 'deaths averted' since deaths are only averted as long as patients continue to be on ART, and because current ART service statistics do not keep track of individuals staying on ART they only provide the total number on ART).

2. Estimating the time from seroconversion to ART eligibility and death in untreated patients from lower income countries: The eART-LINC Collaboration

There is uncertainty about the length of time from seroconversion to ART-eligibility and death in untreated patients in resource-limited countries.

Individual participant data (IPD) were obtained from four cohort studies in resource limited countries (one from Côte d'Ivoire, one from Uganda, and two from Thailand). Weibull survival models in a Bayesian framework were used to estimate time from seroconversion to ART eligibility (defined in this work as CD4 counts dropping < 200 cells/μl; current WHO criteria are broader) and from ART eligibility to death or, from seroconversion to death, depending on the available information. Consistency constraints were set by assuming that the mean survival time from seroconversion to death should equal the sum of the mean survival time from seroconversion to ART eligibility and from eligibility to death. Hierarchical models were used to accommodate cohort heterogeneity and implemented models in WinBUGS.

A total of 1512 patients from resource-limited settings were analysed. Mean survival times were 5.67 years (95% credibility interval 2.85-9.96) from seroconversion to ART eligibility, 2.49 years (1.09-5.06) from ART eligibility to death, and 8.29 years (5.65-12.48) from seroconversion to death, with no substantial heterogeneity between cohorts.

The approach chosen for combining natural history information from several cohorts using IPD could successfully be implemented. However, patient numbers are small at present, and estimates are uncertain. It is hoped that more cohort studies will share their data with the eART-LINC collaboration in the future, enabling more-precise estimates to be obtained and the effects of potentially-important covariates examined.

3. Implications of changed ART eligibility: if and how to include CD4 levels

Modelling the impact of ART on mortality, taking account of the stage of disease at treatment initiation shows that ART-eligibility criteria, frequency of testing individuals, accessibility of care, and care-seeking behaviour of individuals are all important. As ART provision increases, treatment tends to be initiated at an earlier stage of infection. Spectrum is not able to model the effects of stage-at-initiation at present.

It was noted that model estimates of ART impact may be over-optimistic due to parameters having been estimated from studies in trial settings, where the infrastructure is superior to what will be available in standard ART provision. When estimating mortality in those on ART, loss to follow-up (LTFU) will have a major impact. As ART provision is rolled-out in harder-to-reach locations and population groups, rates of LTFU are likely to increase, requiring active tracing of individuals to detect outcomes.

4. Estimating the effects of ART using Spectrum

A simple method of estimating the effects on ART on AIDS deaths is to compare two Spectrum projections that are identical except that one has ART coverage set to zero for all years and the other has actual ART coverage. The difference between the two projections is the effect of ART on AIDS deaths, AIDS orphans, and other indicators. This comparison depends on proper handling of the effects of ART on prevalence. One approach is to input prevalence from EPP or the Workbook into the no-ART Spectrum projection and calculate the resulting incidence, then use this incidence as the input to both the ART and non-ART projections. Alternate approaches include having Spectrum automatically perform this transform, or removing the effect of ART from the prevalence curve prepared by EPP or the Workbook.

For most countries (except e.g. Brazil), ART provision to date will have had little impact on the HIV epidemic, so it will not have affected model estimates. Soon this will no longer be the case. It was noted that increasingly there will be women who are receiving ART and whose improved health has enabled them to become pregnant, causing ANC-sampled HIV prevalence to increase. Currently there are few systems for recording whether a sampled woman is receiving ART, and studies should be done to examine whether increases of ART in the population are reflected in sentinel surveillance (perhaps by testing blood samples for the presence of ARVs or their metabolites).

5. From life years gained to disability-adjusted life-years (DALYs)

To date, achieving or maintaining a low HIV prevalence has been seen as a measure of success in HIV control (although incidence is a better indicator, it is much harder to measure). However, ART provision increases HIV prevalence by prolonging lives of those with HIV, confounding the interpretation of prevalence, and so more-sophisticated measures will be required in future. (An increase in prevalence due to

HIV transmission is undesirable, whilst an increase due to longer survival of infected persons is desirable.)

DALYs are typically used for comparing magnitude of different health problems and burden of disease, and combine mortality and morbidity. They capture total health losses associated with disease, injury or risk factors in units of time. Benefits of interventions may be computed as differences in DALYs between intervention scenario and some counterfactual. DALYs are the sum of years of life lost (YLL) due to premature death and years lived with disability (YLD) adjusted for severity. One DALY represents one lost year of healthy life.

DALY calculations typically rely on a simple classification of disease stages / sequelae, with constant disability weight for duration of each stage. The *Global Burden of Disease* study has included only two states for HIV infection: pre-AIDS and AIDS. For capturing benefits of ART, having at least three states (pre-AIDS, untreated AIDS and treated AIDS) would be useful.

There are different schemes for calculation of DALYs, depending on social value choices: age weights and discounting. It was proposed to calculate DALYs according to different schemes, to enable comparison and to avoid disagreements over which is the 'correct' one to use: (i) no age weights, no discounting; (ii) no age weights, with discounting; (iii) both age weights and discounting.

The recently initiated project to update the Global Burden of Disease for 2005 is expected to update and standardize the methodology for cross-disease comparisons.

6. Recommendations

- There is a need to use modelling to examine how the impact of ART on life-years gained / deaths averted can change as a function of time and programme characteristics.
- It is desirable that major international organisations use common metrics and methods for assessing the impact of ART. For AIDS, the preferred metric is life-years gained.
- Spectrum should calculate life-years gained through ART by comparing two projections, with and without ART, using incidence curves as input.
- In Spectrum, where ART coverage has been high for a long time, the prevalence is assumed to reflect the effect of ART. Where coverage has not been high for a long time, calculate the incidence curve without ART and then calculate prevalence on basis of that incidence and the effect of ART on survival of those on ART.

- In the medium term, Spectrum should be modified to account for stage of HIV infection and its effects on prognosis on ART.
- Survival by stage of HIV infection should be based on separate survival patterns for people starting on ART with low and high CD4 counts. These survival patterns should be derived from relevant datasets (e.g. ART-LINC).
- The individual participant data analysis of time to ART eligibility (as per 2006 WHO criteria) should be finalised.

- Some countries report the cumulative number of treatment-initiation events, rather than current number of persons on treatment. These countries should be assisted in improving their systems to account for those ceasing treatment to enable them to estimate ART prevalence.

- A long-term goal is for software packages (Spectrum in particular) to allow calculation of DALYs, using the proposed different schemes and GBD standards.

Modifications to EPP

1. Rationale

In several countries EPP has problems fitting observed epidemics. For example, in some countries HIV prevalence has declined rapidly and then stabilized at a low but non-zero plateau, which EPP can't reproduce - if it matches the decline, EPP tends to predict zero prevalence in the future with high probability. Bayesian melding doesn't resolve this: almost all the selected trajectories converge to zero. Other problems include the sustained almost linear decline in Rwanda and the sustained rise in South Africa. The problem is that the model isn't flexible enough; it was agreed that allowing φ and/or r to change with time could be a solution.

In a first step, the following options for increasing EPP's flexibility were examined for four settings (urban Kenya, urban Zimbabwe, urban Rwanda and KwaZulu-Natal):

1. *Constrained* – This is the status quo model with constant values (over time) for both r and φ.
2. *r Jump* – This allows for a one-time step change in r and requires 2 additional parameters, one specifying the year in which the change happens and another specifying the percentage change in r.
3. *r Drift* – This allows for a trend in r over some defined period. This variant requires 3 additional parameters specifying the start year and end year for the trend and a rate of change that is applied annually over this period.
4. *r Drift + φ Jump* – Includes the features of the "*r Drift*" scenario with the addition of a one-time change in φ, parameterized by the year of the change and the absolute change.
5. *Unconstrained* – Allows for trends in r and φ, each over some (independently-specified) time period.

Next, additional simulations were undertaken with wider ranges around most parameters, and with further flexibility to allow for a second change in r.

Overall, adding some flexibility to EPP will help improve fits for a range of different settings but there are some important limitations to note:

1. As has always been the case, the maximum likelihood estimate may not produce the pattern that is expected. In particular, where we may have some prior belief that an epidemic is not in continued decline, allowing for trending r and φ will by no means guarantee that our anticipated pattern will emerge as the one that best fits the empirical data.

2. More flexibility will obviously allow for curves with more diverse shapes, but there would need to be some reasonable constraints on any new parameters. One reason for constraints is to impose greater efficiency; for example, there is no real reason to allow for r to change too early in the epidemic because this is not really distinguishable from a higher starting value of r.

We need to proceed cautiously with any changes to EPP. At most we might introduce some relatively tightly controlled trend in r. This will add enough flexibility to better fit some of the more stubborn patterns. But it seems likely that even with a more flexible EPP there will be persistent cases where the MLE does not provide the best visual fit to the data, or does not conform to prior expectations about the direction of recent trends, particularly when extrapolating beyond the last data point.

Adding increased flexibility requires increasing the number of parameters, from 4 to at least 6. This should be possible for Bayesian Melding but will increase

computational requirements. It may be harder to find enough good curves in Bayesian melding with 6 parameters than with 4, sampling from the prior. However, Bayesian melding may actually work better if the modified model fits the data better. Increasing the dimensionality of parameter space may require a more-efficient sampling scheme, possibly using Adaptive Importance Sampling, e.g. Incremental mixture importance sampling. This requires further work, expected to take approximately 1 year.

It was noted that increasing sampling efficiency may solve problems with the current EPP for some countries where the number of unique curves is small, and could also allow fuller use of prior information on outputs, including the prior expectation of prevalence levelling-off rather than declining to zero. Therefore, it may actually be the case that the current EPP does not actually need to be changed.

2. Recommendations

- It was noted that EPP has been found to have good predictive ability - i.e. fitting to earlier data-points results in a good prediction of later data-points - for most countries, so we should beware over-elaborating the default model, except where absolutely necessary (e.g. for Ethiopia, Kenya, Uganda, Rwanda, other countries with steep recent declines in prevalence with or without evidence of prevalence levelling off and South Africa).
- There are two separate issues: (i) problems of optimising EPP's fit to data, and (ii) patterns that EPP simply cannot reproduce. Further work (e.g. on more-efficient sampling methods) is required to address these. It may be the case that more-efficient sampling means that increased flexibility is not generally required.
- In the short term, whilst sampling methods are assessed, EPP should continue to be used, but with separate analyses done for the few intractable countries.
- The Excel version of EPP (which allows for annual changes in *r*) should be applied to the Uganda data to test its performance.
- Allowing the *r*-jump option (i.e. a step-change in *r*) should be an 'advanced' option and should not occur by default. The parameters specifying *r*-jump (i.e. time of occurrence and 'new' *r* value) should not be varied by the automated fitting routine, but should only be manually changed by the user, post-fitting. It would be appropriate that use of additional parameters to increase flexibility is statistically justified or based on extrinsic evidence.
- Data on behaviour change are required to determine priors for parameter changes in the elaborated version of EPP; it was questioned how many countries have good data collected more frequently than the usual 5-year interval between DHS. However, some countries are adopting a schedule with more frequent behavioural data collection in national surveys (every 2-3 years).
- Investigation of the ability of using Bayesian Melding and potentially the use of Bayes factors to fit the elaborated version of EPP.
- Long-term, fitting of a more-flexible curve should be automated with robust statistical testing of whether it is justified or not.
- A long-term goal should be to consider the development of new models for fitting to prevalence as an alternative to continuously elaborating EPP beyond the limits for which it was designed.

Appendix I: Meeting Agenda

Friday July 13th: (1) Methods for estimation of ART's impact on deaths averted/delayed; (2) EPP 2007

Start	Duration	Subject	Speaker
900	25	Opening remarks	Peter Ghys
Session 1 - Estimation of effects of ART: Methods for estimates of deaths averted/delayed. *Chair: John Stover*			
925	15	UNAIDS/WHO method (as per 2006 Global Report and 2006 Epi Update).	Karen Stanecki
940	25	Estimating lives saved in Global Fund-supported programs: ARV focus	Ryuichi Komatsu
1005	35	Estimating the time from seroconversion to ART eligibility and death in untreated patients from lower income countries: The eART-LINC Collaboration	Simon Wandel
1040	*15*	*Coffee break*	-
1055	30	Implications of changed ART eligibility; if and how to include CD4 levels	Geoff Garnett
1125	30	Estimating the Effects of ART using Spectrum	John Stover
1155	20	From life years gained to DALYs	Josh Salomon
1215	30	Discussion	-
1245	*75*	*Lunch*	-
Discussions. *Chair: Geoff Garnett*			
1400	100	Discussion of how to interpret data	-
1540	*20*	*Coffee break*	-
Session 2 - EPP 2007: changing parameters over time to allow for bottoming out after declining prevalence *Chair: Peter Ghys*			
1600	10	Proposal for EPP modification	Josh Salomon
1610	20	Bottoming Out, Uncertainty and Model Choice in EPP	Adrian Raftery
1630	10	Kenya update	John Stover
1640	15	Discussion	-
1655	-	Close	-

Appendix II: List of Participants

Rick Berzon
HIV/AIDS Technical Advisor
USAID

Tim Brown*
Senior Fellow, East-West Center,
Honolulu, USA

Tim Fowler
Chief, Health Studies Branch
International Programs Center
US Census Bureau, USA

Geoff Garnett
UNAIDS Epidemiology Ref. Group Secretariat
Department of Infectious Disease Epidemiology
Imperial College London, UK

Peter Ghys
Manager, Epidemic and Impact Monitoring,
Policy, Evidence and Partnerships Department,
UNAIDS, Geneva, Switzerland

Peter Johnson
International Programs Center
US Census Bureau

Ryuichi Komatsu
Strategic Information Officer
The Global Fund to Fight AIDS, Tuberculosis & Malaria

Katherine Marconi
Director, Strategic Information
PEPFAR, Office of the Global AIDS Coordinator
US Department of State

Adrian Raftery
Departments of Statistics and Sociology,
University of Washington, Seattle, USA

Laura Robertson
UNAIDS Epidemiology Ref. Group Secretariat
Department of Infectious Disease Epidemiology
Imperial College London, UK

Joshua Salomon
Department of Population & International Health,
Harvard School of Public Health, USA

Michelle Sherlock
Strategic Information
Office of the U.S. Global AIDS Coordinator
US Department of State

Ray Shiraishi
Centers for Disease Control and Prevention
Global AIDS Program
Atlanta, Georgia, USA

Karen Stanecki
Epidemic and Impact Monitoring,
Policy, Evidence and Partnerships Department,
UNAIDS, Geneva, Switzerland

John Stover
Futures Institute
Glastonbury CT 06033, USA

Simon Wandel
Institute of Social and Preventive Medicine
Division of Clinical Epidemiology & Biostatistics
University of Bern, Switzerland

Peter White
UNAIDS Epidemiology Reference Group Secretariat
Department of Infectious Disease Epidemiology
Imperial College London, UK

*Participated in EPP discussion by telephone.