# Estimating uncertainty in national HIV burdens using Bayesian Melding

Report of a meeting of the UNAIDS Reference Group for Estimates, Modelling and Projections, held in London, UK, 30th October 2006

## TECHNICAL REPORT AND RECOMMENDATIONS

Joint United Nations Programme on HIV/AIDS

**UNAIDS**

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

**UNAIDS**

# Estimating uncertainty in national HIV burdens using Bayesian Melding

*Report written by Dr Peter White, Coordinator, UNAIDS Reference Group on Estimates, Modelling and Projections (the 'Epidemiology Reference Group'), Imperial College London. In case of any queries please e-mail p.white@imperial.ac.uk.*

Present: Leontine Alkema, Tim Brown, Simon Cousens, Geoff Garnett, Peter Ghys, Nicholas Grassly, Adrian Raftery, Joshua Salomon, Peter White.

Participating by phone: Daniel Barth-Jones, Rob Dorrington, Leigh Johnson, Meade Morgan, John Stover.

Presentations were received from Tim Brown, Leontine Alkema, and John Stover.

Introduction

Estimates of HIV burden should include information on the certainty associated with the estimate. The Reference Group has been working to improve the techniques used to estimate the degree of uncertainty, whilst ensuring that its software remains appropriate for widespread use, in terms of the statistical expertise and computing power required. A Bayesian Melding (BM) approach is being implemented in EPP (Estimation and Projection Package), and Monte-Carlo methods are being implemented in Spectrum.

Progress in implementation of Bayesian Melding in EPP

Initial implementation will be completed by the November Reference Group meeting, at which recommendations on details of user interface will be sought.

The algorithm draws 200,000 candidate parameter sets from the specified ranges of values (the 'priors'), generates a curve for each and calculates its likelihood. Resampling then selects from this set of 200,000, weighted according to likelihood.

BM is computationally demanding and it is important to ensure that the algorithm is suitable for computers that are available to in-country teams. Computational efficiency has been increased by creating a look-up table of the integrals of Weilbull functions rather than performing calculations in real time. Storing results for all 200,000 requires a large amount of memory or hard-drive storage space; discarding those with negligible likelihood – which would not be resampled in any case – would reduce this requirement. Currently EPP is programmed in Java; using C++ would be expected to increase its speed but this does not appear to be necessary at present, since BM is expected to take only a few minutes on most computers that are likely to be used.

Users need to be able to output the 'best fit' curve, which may not necessarily be the highest-likelihood curve. The 'best fit' should be a generated curve, rather than the median of the resampled curves, to ensure consistency with Spectrum. It was noted that the most-likely parameter set obtained by BM will not be identical to the maximum-likelihood parameter set obtained by the curve-fitting algorithm because the latter is obtained by adjusting parameters to fine-tune the fit, whilst the former is

not fine-tuned. The difference between the two should be investigated. It would be possible to obtain a best-fit curve using the curve-fitter whilst using BM for estimation of plausibility bounds.

There is a need to explore the effects of different input prior distributions for each parameter, to produce default sets. Also, a default set of output priors (constraints on plausible curve shapes: e.g. date-ranges for peak prevalence) needs to be developed. Users will be allowed to alter default settings via an "advanced" page in EPP.

Adjusting for bias in ANC prevalence in EPP

ANC clinics typically provide a biased estimate of population prevalence. To adjust for this, the probit transformation of the ratio of the prevalences estimated by ANC and local population-based surveys is the most consistent, outperforming the arcsin transformation in terms of consistency of variance.

It was agreed that DHS prevalence estimates should be included in the likelihood function, rather than forcing the EPP curve to pass through it. This would account for uncertainty in the DHS estimate because the EPP curve would be 'pulled towards' it with a strength dependent upon its degree of uncertainty.

Implementation of Monte Carlo methods in Spectrum

In Spectrum, scalar quantities are specified by their mean and standard deviation. For temporally-varying values, a library of curves from which the model will select randomly is used to reduce computation requirements. Performing 1000 runs is expected to take around 15 minutes. Outputs are the median and 95% range (i.e. 2.5%, 97.5% values).

# Appendix: List of Participants

**Leontine Alkema**
Department of Statistics, University of Washington,
Seattle, USA

**Daniel Barth-Jones**
Assistant Professor and Epidemiologist,
Wayne University, Detroit, USA

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

**Simon Cousens**
London School of Hygiene & Tropical Medicine, UK

**Rob Dorrington**
Director, Centre for Actuarial Research
University of Cape Town, South Africa

**Geoff Garnett**
UNAIDS Epidemiology Reference 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

**Nicholas Grassly**
Department of Infectious Disease Epidemiology
Imperial College London, UK

**Leigh Johnson**
Centre for Actuarial Research
University of Cape Town, South Africa

**Meade Morgan**
Health Scientist, Global AIDS Program,
Centers for Disease Control and Prevention
Atlanta, GA 30333, USA

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

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

**John Stover**
Futures Institute
Glastonbury CT 06033, USA

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