Generating national estimates of HIV prevalence in concentrated epidemics

Report and recommendations from a meeting of the UNAIDS Reference Group on Estimates, Modelling and Projections held in Barcelona, Spain, 7 October 2013

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, October 2013

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, the World Health Organization (WHO) 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, Imperial College London.

Aim of the meeting

To identify methods and the use of additional sources of data which may be utilised for alternative estimation approaches in concentrated epidemics.

The specific objectives of this meeting were:

1) Discuss the challenges, key issues and data available in concentrated epidemics and the potential applicability of the use of novel methods to utilise the available data sources.

2) To review and discuss the methods used by the European Centre for Disease Prevention and Control (ECDC) for generating national estimates of HIV prevalence in this region, results from application of these approaches and the potential for application in other regions.

3) To review and discuss the Euro-Coord-SSOPHIE modelling approach and results from applying this method to data from the UK.

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 and WHO 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.

Generating national estimates of HIV prevalence in concentrated epidemics

In 2015, countries will report on their progress towards the Millennium Development Goals. These targets include: *Reduce sexual transmission of HIV by 50%* and *Halve the transmission of HIV among people who inject drugs.* There are only two years remaining before countries will need to report on progress towards these goals. Currently, the majority of countries in the world use EPP/Spectrum to generate their national estimates of prevalence (and other indicators) as part of the UNAIDS Global Report. Some countries with concentrated epidemics are not entirely satisfied with the estimation tools currently available and have requested methods that utilise the data they have available, for example, case report data. The objectives of this meeting were to discuss the key issues, challenges and data available in concentrated epidemics and to review and discuss alternative approaches for estimation with the overall aim to improve the current process and methods used.

Key issues, challenges and data availability in countries with concentrated epidemics

Inconsistent data collection and the use of different survey methods can result in widely varying results which can influence the national (or even regional) trend. While data collection is increasing, many countries will have only two data points from the same site and do not actually have a trend. Most data collection efforts have occurred recently and there is often limited information for the epidemic structure in the early years. While countries are getting appropriate information for how the data should be utilised, more guidance is still needed for how to handle imperfect data situations, data-light settings, and for how to incorporate additional data available that may not be for model fitting, but may be useful to inform the epidemic trends and level.

Discussion: The use of EPP instead of Workbook has been an improvement. Workbook uses ranges of values and is easy to manipulate (higher ranges result in higher levels of prevalence) and difficult to assess externally because it is difficult to identify where the data are coming from.

Regions handle their key data issues in different ways. The Asia Pacific region adds two extra days to the workshops to focus on data issues while Latin America uses an interactive phone/WebEx consultation process before the workshops to address data issues and prepare for estimation.

In the EPP context, countries can utilise case report data for model validation (timing of the epidemic) and to assist with calibration.

Estimation issues in the MENA region - data available to inform estimates and alternative approaches

HIV in the Middle East and North Africa region (MENA) is estimated to be low, less than 0.2%, but the epidemics have only recently emerged and are currently growing. HIV is focussed in key populations (sometimes only a single key population), may be geographically clustered, and is often only discovered after prevalence has reached considerable levels. In particular, HIV appears to be emerging among injecting drug users (IDU) and men who have sex with men (MSM). Since 2003, epidemics among MSM have emerged in at least half of MENA region countries.

Data challenges: Some countries have bio-behavioural surveys, but these are scattered in time and geography, usually focussed on key urban settings, and mostly funded by global donors which may

not be sustainable. There are no nationally representative data (e.g. DHS) for HIV infection or sexual behaviours. However, there are data available that have not previously been used which can be utilised to better inform estimation. HIV testing rates are extremely high in this region with mandatory testing before marriage, for all migrants, for all pregnant women and for drivers licences (Syria). These testing databases are generally not utilised to inform estimation and the data are not routinely reported. There are also case report data and a new standardised form has been developed for case notification (but is not currently being used).

Discussion: Testing policies vary by country as does the quality of the data and databases but the data are usually electronic and contain detailed information. There was consensus that available databases are not being used appropriately to inform estimates (as validation, to inform timing of the epidemic in key populations), but it is unclear how to make this use systematic. It was discussed that migrants could be used as a proxy for incidence – all migrants are tested and must be HIV- to remain in the county and are re-tested annually. The pre-marital testing and driver's license testing databases may also be particularly useful.

The ECDC modelling project - Results from use of these methods in the EU & EE/EFTA region, *Ard van Sighem*

Countries with concentrated HIV epidemics are often in the situation of not having enough data for estimating HIV prevalence using methods based on prevalence surveys (DHS, AIS) but do have other information – case report and CD4 data – that can be utilised to inform alternative estimation models. The ECDC modelling project seeks to utilise the data available to generate national estimates of HIV prevalence using systematic approaches.

Two methods are used for this work:

Method 1: Reconstruct the incidence curve for the entire epidemic
Method 2: Utilise the relationship between CD4 counts and AIDS

Method 1: HIV cases can be calculated from observed AIDS cases until the introduction of ART (1996). HIV diagnosis can be used as an alternative marker with the caveats that the distribution of time between HIV infection and diagnosis is unknown and may change over time and that changes in annual number of infections cannot be distinguished from changes in diagnosis rates. Adjustments can be made using markers for how recent a newly diagnosed infection is (CD4 count or simultaneous HIV/AIDS diagnosis). Results obtained using Method 1 can be used to generate estimates of prevalence by 1) inputting incidence curve into Spectrum, 2) Using the incidence curve and diagnosis rate as an input into the SOPHIE model, 3) Cumulating incidence and subtracting deaths amongst those infected with HIV. The data required for Method 1 include:

- Annual total number of AIDS cases up to 1996
- Annual number of HIV diagnoses
- Annual number of HIV/AIDS cases
- CD4 count at time of diagnosis (if available)
- Demographic data: Mode of transmission, sex, country of birth (migrant vs non-migrant)

Limitations of this approach include that CD4 counts are only reported from 2005; before 2005 it is not possible to estimate separate diagnosis rates and CD4 intervals, thus assumptions are needed. Underreporting, delayed reporting, incomplete information, double counting, mortality before HIV

diagnosis and large proportions of migrant populations can also substantially affect the estimates. An in depth knowledge of the data system utilise (how information is collected and recorded and inputted into data systems) is necessary in order to best inform the use of these methods.

Method 2 uses the relationship between CD4 count and AIDS to capture those who present for testing due to AIDS symptoms. This method is straightforward with the main caveat that people may be testing earlier and the relationship between pre-AIDS diagnosis and time to AIDS is less straightforward. A key benefit of this method estimates can be generated with only one year of surveillance data. The data required include:

- Annual number of HIV/AIDS cases
- CD4 count at time of diagnosis
- Number with pre-AIDS symptoms
- Demographic data: Mode of transmission, sex, country of birth (migrant vs non-migrant)

Four countries have currently been tested using these approaches. The next steps include: Using Method 1 with data from Estonia; testing the methods on simulated data generated with different underlying true HIV incidence and diagnosis rates; conducting sensitivity analyses on input parameters; and investigating smoothing the HIV incidence curve.

Discussion: It was discussed that it probably is possible to combine methods 1 & 2 and that the SSOPHIE model is doing this to a degree, but not explicitly. To date, the methods have not been tested in in sub-optimal conditions. The ECDC timeline for software is October 2014. In advance of this time frame, it would be useful to apply this method in other regions where countries have expressed their desire for alternative estimation methods (Latin America and MENA) and compare the results to those from EPP/Spectrum.

EuroCoord-SSOPHIE project: Stochastic simulation of outcomes of people with HIV in Europe, *Fumiyo Nakagawa & Andrew Phillips*

The aim of the SSOPHIE project is to build an individual-based, stochastic model of HIV infection and the effect of ART that recreates the historical trend and projects the status of HIV-infected individuals in countries throughout Europe. The SSOPHIE model includes estimated outcomes such as treatment usage, drug resistance, pregnancy, rates of AIDS and death.

This modelling approach is more data and computationally intense (compared to the ECDC work) and the intention is for the SSOPHIE team to work directly with the countries for application. Fitting is based on Approximate Bayesian Computation whereby multiple simulations sample unknown parameter values from suitable distributions (incidence and diagnosis rates, probability of starting ART when eligible, CD4 threshold at ART initiation, rate of loss to follow-up, etc), then the sets of parameter values for which the simulated outcomes best fit the observed data are selected (using a fit score) to generate point estimates and plausibility ranges. The simulations can also be done by risk group if the data are available.

Results: Fitting the SSOPHIE model to UK data resulted in a fairly good fit given the wide range of flexibility allowed. More input into the prior distribution, more simulations (and a more stringent fit score) will further improve this fit. The model is also able to be utilised without having “perfect” data conditions (having all data required) but the plausibility ranges quickly expand and some “imperfect” data situations perform better than others – for example having all data required only for a single

year replicates the trend observed but with wide plausibility ranges. One of the key caveats of this approach is that the fitting procedure is very time intensive. An alternative option to improve efficiency is to use the incidence and diagnosis rate from Method 1 in the ECDC approach.

Discussion: The IMIS approach (developed by Le Bao and Adrian Raftery) may be able to improve computational time by improving sampling efficiency. The model will also be less computationally intense after the initial investments - once the model is set up for a country, only new data will need to be added in future years, thus this could be a potential future option for generating estimates in some countries. If prevalence were added to the SSOPHIE model, this method could be used as comparison to estimates from EPP/Spectrum.

Recommendations

Recommendations for the ECDC modelling approach:

- **Test without perfect data (missing large chunks of CD4 data) and test the assumption that data are missing at random – evaluate performance.**
- **Consider incorporating incidence assays and treatment into this approach**

Follow-Up: Ard van Sighem, *further discussions in Nov, review results April 2014*

- **Explore the possibility of using ECDC approach in other countries, in MENA test in Morocco and amore data scarce country, also consider testing in Brazil (only has AIDS reports), Mexico, and an Eastern European country.**

Follow-Up:

- UNAIDS to contact ECDC to confirm the use of this model in additional countries.
- Laith Abu-Raddad to test in Morocco and an additional MENA country (Oman, Tunisia?).
- Andrew Phillips and Ard van Sighem to advise regarding additional individuals to conduct country investigations.
- Tim Hallett to contact Juan Vesga re availability for consultanct to conduct country investigations using Ard's method and also incorporating prevalence and treatment in Brazil and Mexico.
- UNAIDS to identify additional person(s) to conduct country investigation in Ukraine.
- UNAIDS to contact new colleague in Moscow regarding Russia participating in this analysis.
- ECDC group to contact MoH in Russia regarding participating in this analysis.
- Compare investigative results from ECDC model to results from EPP/Spectrum.

Review results in April 2014

Recommendations for additional investigation in the MENA region:

- **Investigate the use of data from migrants as a proxy for incidence.**
- **Investigate the use of other data bases to inform estimates (databases from marriage, drivers licence).**

Follow-Up: Laith Abu-Raddad, *review in April 2014*

Hierarchical model for concentrated epidemics – information sharing to improve model fits

Follow-Up: UNAIDS to provide list of concentrated epidemic countries with reasonable data for each of the key populations.

Follow-Up: Le Bao *to* present proposed methods *November 2013, review results December 2013 at Reference Group meeting*

Appendix I: List of Participants

Le Bao
Penn State
State College, Pennsylvania, USA

Laith Abu-Raddad
Weill Cornell Medical College
Ar-Rayyān, Qatar

Tim Brown
East-West Center,
Honolulu, Hawaii, USA

Kelsey Case
Department of Infectious Disease Epidemiology
Imperial College London, UK

Geoff Garnett
Bill & Melinda Gates Foundation
Seattle, WA, USA

Peter Ghys
UNAIDS
Geneva, Switzerland

Simon Gregson
Department of Infectious Disease Epidemiology
Imperial College London, UK

Tim Hallett
Department of Infectious Disease Epidemiology
Imperial College London, UK

Mike Isbell
Consultant
New York, NY USA

Mary Mahy
UNAIDS
Geneva, Switzerland

Bradley Mathers
University of New South Wales
Sydney, Australia

Mark McGovern
Harvard School of Public Health
Boston, Massachusetts, USA

Fumiyo Nakagawa
University College London
London, UK

Ryosuke Omori
Weill Cornell Medical College
Ar-Rayyān, Qatar

Katrina Ortblad
Institute for Health Metrics & Evaluation
Seattle, WA, USA

Andrew Phillips
University College London
London, UK

Robert Puckett
East-West Center
Honolulu, Hawaii, USA

Keith Sabin
UNAIDS
Geneva, Switzerland

Karen Stanecki
Consultant
Arlington, VA, USA

John Stover
Futures Institute
Glastonbury, CT, USA

Ard van Sighem
Stichting HIV Monitoring
Amsterdam, Netherlands

Appendix II: Meeting Agenda

Generating national estimates of HIV prevalence in concentrated epidemics

Monday, October 7th, 2013

Start	Duration	Subject	Speaker
Estimation in concentrated epidemics (Chair: Tim Hallett)			
1300	15	Meeting Opening - Background and context for this meeting, aims and objectives	Peter Ghys, *UNAIDS* Tim Hallett, *Imperial College London*
1315	10	Introductions	ALL
1325	15	Data availability and key issues and challenges in countries with concentrated epidemics	Keith Sabin, *UNAIDS*
1340	10	*Questions*	ALL
1350	20	Estimation issues in the MENA region - data available to inform estimates and alternative approaches utilised	Laith Abu-Raddad, *Weill Cornell, Qatar*
1410	10	*Questions*	ALL
1420	25	Results from application of ECDC modelling approach, precise data required, characteristics of results in the countries utilised, key challenges	Ard van Sighem, *Stichting HIV Monitoring*
1445	10	*Questions*	ALL
1455	*20*	*Coffee break*	-
1515	25	The SSOPHIE modelling approach – fitting using the the SSOPHIE approach, use of the SSOPHIE approach with externally estimated incidence and diagnosis rate curve and potential for future application	Andrew Phillips & Fumiyo Nakagawa, *University College London*
1540	10	*Questions*	ALL
1550	40	**Group discussion - recommendations, follow-up and next steps**	ALL
1630		*Close*	-