

The Spectrum projection package: improvements in estimating incidence by age and sex, mother-to-child transmission, HIV progression in children and double orphans

J Stover,¹ P Johnson,² T Hallett,³ M Marston,⁴ R Becquet,^{5,6} I M Timaeus⁷

¹Futures Institute, Glastonbury, Connecticut, USA

²US Census Bureau, Washington, DC, USA

³Department of Disease Epidemiology, Imperial College London, London, UK

⁴Department of Population Studies, London School of Hygiene and Tropical Medicine, London, UK

⁵INSERM, unité 897, Centre de recherche 'Epidémiologie et Biostatistique,' Bordeaux, France

⁶Institut de Santé Publique Epidémiologie Développement (ISPED), Université Victor Segalen Bordeaux 2, Bordeaux, France

⁷Department of Population Studies, London School of Hygiene and Tropical Medicine, London, UK

Correspondence to

Mr John Stover, Futures Institute, 41A New London Turnpike, Glastonbury, CT 06033, USA; jstover@futuresinstitute.org

Accepted 1 September 2010

ABSTRACT

Background The Spectrum program is used to estimate key HIV indicators from the trends in incidence and prevalence estimated by the Estimation and Projection Package or the Workbook. These indicators include the number of people living with HIV, new infections, AIDS deaths, AIDS orphans, the number of adults and children needing treatment, the need for prevention of mother-to-child transmission and the impact of antiretroviral treatment on survival. The UNAIDS Reference Group on Estimates, Models and Projections regularly reviews new data and information needs, and recommends updates to the methodology and assumptions used in Spectrum.

Methods The latest update to Spectrum was used in the 2009 round of global estimates. This update contains new procedures for estimating: the age and sex distribution of adult incidence, new child infections occurring around delivery or through breastfeeding, the survival of children by timing of infection and the number of double orphans.

INTRODUCTION

UNAIDS projections of adult HIV prevalence and incidence are prepared using the Estimation and Projection Package (EPP)¹ for generalised epidemics and the Workbook for low-level and concentrated epidemics.² The Spectrum projection package is used to determine the consequences of these projections, including the number of people living with HIV by age and sex, new infections, AIDS deaths, AIDS orphans, the need for treatment and prophylaxis, and the impact of treatment. Earlier updates to Spectrum have been described previously.^{3, 4} A full description of the methodology is provided in the manual for the AIDS module of Spectrum (AIM).⁵ The purpose of this paper is to describe the updates to Spectrum made in the last 2 years and used in the 2009 global estimates of HIV/AIDS.

Spectrum is a modular program used to examine the consequences of current trends and future program interventions in reproductive health. The program and manuals are updated regularly and available in multiple languages free of charge at <http://www.FuturesInstitute.org/>. Model development has been funded primarily by USAID, UNAIDS and the Bill and Melinda Gates Foundation with technical collaboration from UNAIDS, WHO, Unicef, UN Population Division, US Census Bureau, UNFPA and other organisations.

The major inputs and outputs of the AIDS module of Spectrum are shown in figure 1. Demographic

projections are based on user inputs or projections prepared by the United National Population Division.⁶ The projections start with an estimate and projection of adult incidence, which is combined with information on the age and sex distribution of incidence and progression to death to estimate the number of new adult infections by age and sex. New infant infections are estimated from prevalence among pregnant women and the rate of mother-to-child transmission, which is dependent on infant feeding practices and the coverage of prophylaxis with antiretrovirals (ARVs). New infections progress over time to a symptomatic stage where antiretroviral therapy (ART) is needed. Those who receive first- and or second-line ART experience extended survival. People at any stage are subject to non-AIDS mortality at the same rates as those who are not infected. Adult deaths result in orphans.

Several new features have been added to Spectrum for the 2009 round of estimates in response to new data, new treatment guidelines and new needs by national programmes for planning. There are new procedures for estimating the age and sex distribution of adult incidence, new child infections occurring around delivery or through breastfeeding, the survival of children by timing of infection, the effects of HIV infection in fertility and the number of double orphans.

DISTRIBUTING ADULT INCIDENCE BY AGE AND SEX

EPP is used to fit a smooth prevalence trend to surveillance and survey data and to estimate the associated incidence among adults aged 15–49. Then, the incidence trend is passed to Spectrum. Since Spectrum calculations are disaggregated by age and sex, the adult incidence from EPP also needs to be converted to age- and sex-specific incidence. To do this, we used a mathematical method for estimating incidence by age from two cross-sectional measurements of prevalence.⁷ The model is based on the assumption that individuals of age a in the earlier survey will be represented by individuals aged $a+\tau$ in the later survey, where τ is the interval between surveys, even though the surveys do not include the same individuals. The change in HIV prevalence among individuals aged a in the first survey and $a+\tau$ in the second survey can then be attributed to incident infections and AIDS deaths. By finding an approximate value for the rate of AIDS deaths, based on the observed distribution of survival after infection,⁸ the HIV incidence for

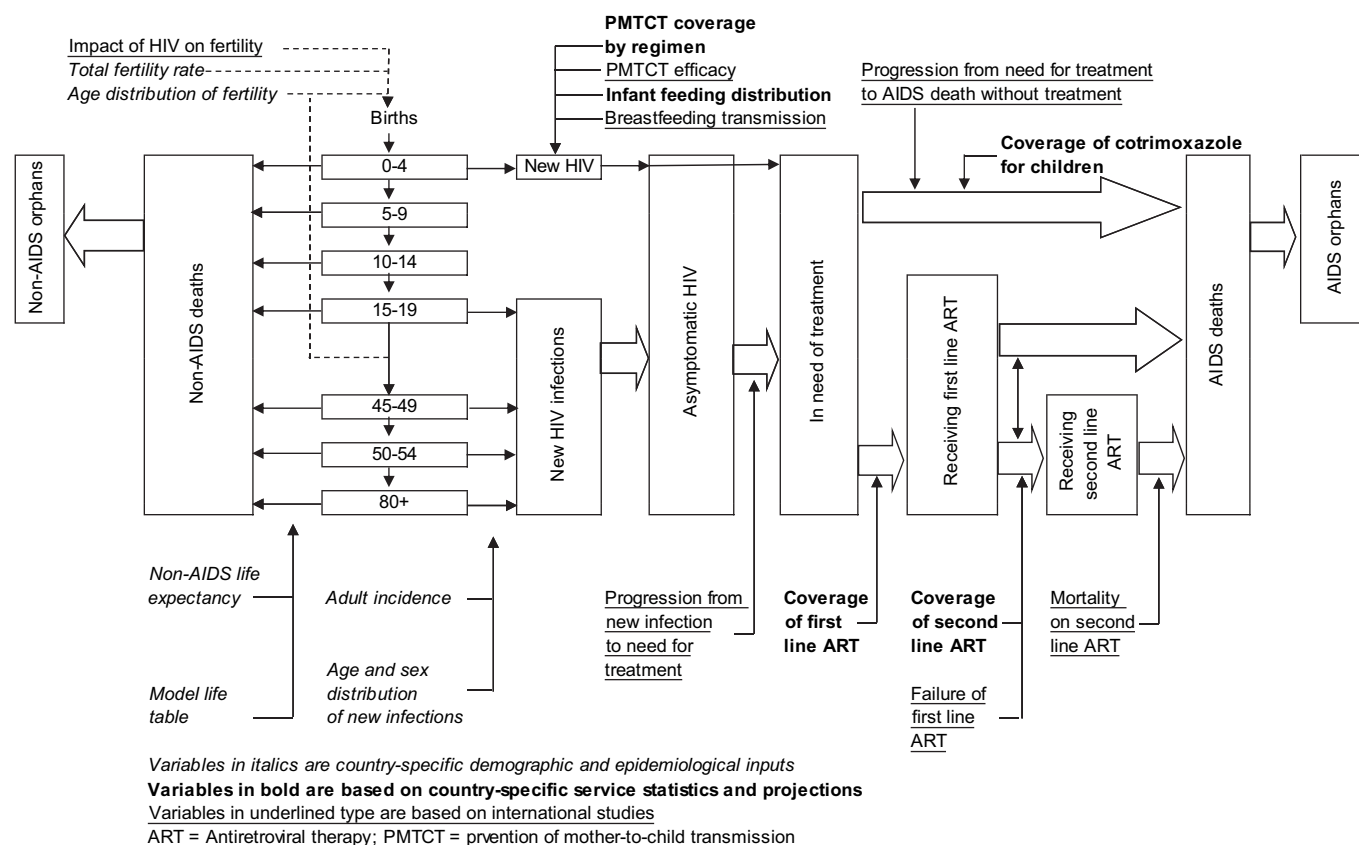


Figure 1 Inputs, outputs and process of the demographic and AIDS modules in Spectrum.

that age group can be estimated. This method has been thoroughly validated by comparing derived incidence estimates with actual incidence measurements in cohort studies in Tanzania, Uganda and Zimbabwe.⁷ If there is only one survey, the same procedure can be applied, making the additional assumption that HIV prevalence was constant in the 5 years preceding that survey. (If prevalence is not constant, the pattern for some age groups could be over- or underestimated, but prevalence is roughly constant in most of the countries with Demographic and Health Surveys (DHS).)

The effects of ARV treatment are accounted for by removing the fraction of infected individuals that are alive due to treatment from the prevalence data in each survey.⁹ To do this, data on the scale-up of treatment in each country are used, and it is assumed that treatment is typically initiated 1 year before expected AIDS death¹⁰ and that the age/sex distribution of those on treatment is approximated by the distribution of AIDS deaths projected using Spectrum and published HIV prevalence trend data.⁴

These estimates of incidence are associated with uncertainty arising from the sampling error in the prevalence estimates. Further, in the countries with smaller epidemics, household surveys could underestimate the true HIV prevalence, given the concentration of infection in populations that are not fully captured (eg, hostels, brothels, military/police barracks).¹¹ In all settings, non-response could also lead to underestimation of prevalence, although we expect this effect to be small,^{11–13} and if it is constant over age and time, the estimated age distribution will be reliable. The method also does not account for the possibility of different rates of international migration by HIV status and age, which could introduce bias. Such information on international migration by HIV status is not available, but

experimentation with the method suggests that the magnitude of any errors introduced by migration is likely to be small.

We applied this method to data on HIV prevalence by age and sex from 28 DHS and AIDS Indicator Surveys. Four countries had more than one survey (Dominican Republic, Kenya, Mali and Zambia). For countries with adult prevalence above 4% the sample sizes are large enough for the incidence patterns to be relatively smooth. Figure 2 shows the resulting patterns for males and females. For countries with lower prevalence the patterns are more scattered. We calculated average patterns for these two sets of countries and also calculated a pattern for the injecting-drug user (IDU)-driven epidemics of Eastern Europe based on new HIV case reports.¹⁴ The resulting patterns are shown in figure 3.

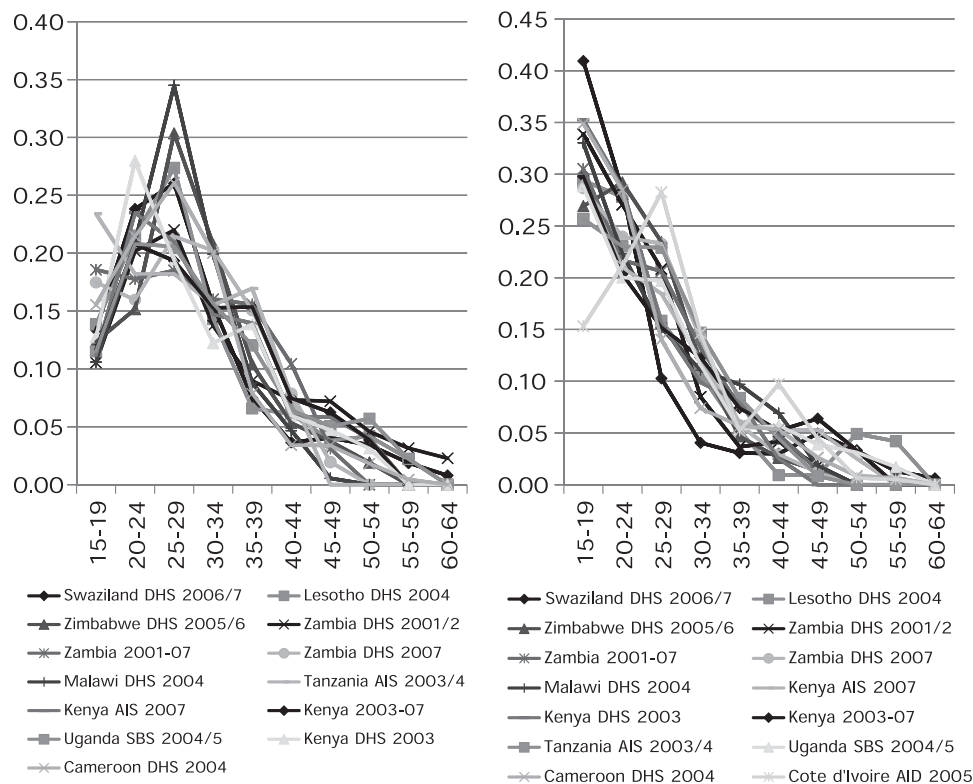
The same analysis indicates that the ratio of female to male incidence among adults 15–49 varies from 0.5 to 2.4 with a median of 1.38 in generalised epidemics, 0.84 in low level and concentrated epidemics, and 0.42 in IDU-driven epidemics. The ratios vary considerably across countries. This variation is reflected when uncertainty analyses are run.

These patterns are used in Spectrum for all countries in each epidemic type and are assumed to be constant from the beginning of the epidemic. These assumptions produce a good match to the survey measures of prevalence by age and sex 20–30 years after the start of epidemic, as shown in figure 4.

HIV INFECTION AMONG CHILDREN

Previous versions of Spectrum estimated new HIV infections among children as occurring around delivery, through the combined probability of infection in utero, peripartum and intrapartum, and through breastfeeding. The survival pattern

Figure 2 Distribution of new HIV infections by age for males (left) and females (right) in high-prevalence countries. AIS, AIDS Indicators Survey; DHS, Demographic and Health Surveys; SBS, Sexual Behaviour Survey.



was modelled as the proportion of children infected at birth who survive to each future age and was based on fitting a double Weibull curve to longitudinal data.¹⁵ The double Weibull curve was used to account for the finding that some HIV-infected children are fast progressors (mostly those infected in utero, peripartum and intrapartum), and some are slow progressors (mostly those infected after birth through breastfeeding).

In the latest version of Spectrum, infections around delivery are tracked separately from those occurring through breastfeeding in order to follow fast and slow progressors separately. The probability of HIV infection at birth for a child born to an HIV-positive mother is assumed to be 20% in the absence of prophylaxis, 11% with single-dose Nevirapine, 4% with dual

ARV prophylaxis or 2% with triple ARV prophylaxis or treatment.⁴ The probability of infection through breastfeeding is assumed to be 1.5% per month for mixed feeding during the first 6 months, 0.75% per month for exclusive feeding during the first 6 months, 0.75% per month for months 7 and later, and 0.3% per month when the mother is on triple therapy.¹⁶⁻²⁰

New survival curves for children infected at birth or through breastfeeding were constructed using data from 12 sub-Saharan African clinical trials and cohort studies.²¹

HIV-infected children were divided into categories of early or late infection, and non-AIDS mortality was removed to focus on net AIDS mortality. Double Weibull curves were fitted to each set of data to determine the pattern of survival. As there was little follow-up beyond 2.5 years, there is limited information to inform the survival pattern beyond age 2.5. Therefore, the survival pattern was extended by assuming that children that have survived at least 2.5 years would have survival similar to young adults aged 15–24, with a median survival time of about 20 years.²² The resulting survival curves are shown in figure 5. The curves indicate a median survival of 1.1 years for early infection and 9.4 years for late infection, and 9% survival at 20 years for early infection versus 16% for late infection.

ESTIMATING THE PREVALENCE OF DOUBLE ORPHANHOOD

Spectrum estimates the number of maternal, paternal and double orphans by parent’s cause of death (AIDS and non-AIDS) based on methods developed for Africa by Grassly and Timaeus in 2001.²³ The number of new maternal orphans is estimated from the number of women dying each year, their fertility history and the age-specific mortality of their children. Paternal orphans are calculated in a similar manner using estimates of men’s fertility. If the deaths of the parents were completely independent, the probability that a child is a double orphan would be simply the product of the probability of being a maternal orphan and the probability of being a paternal

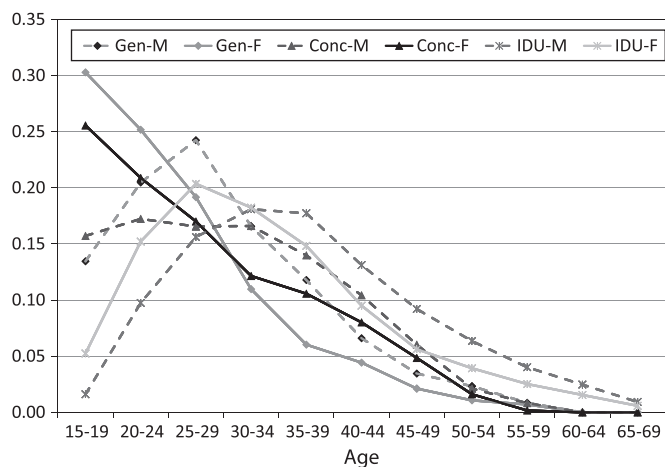
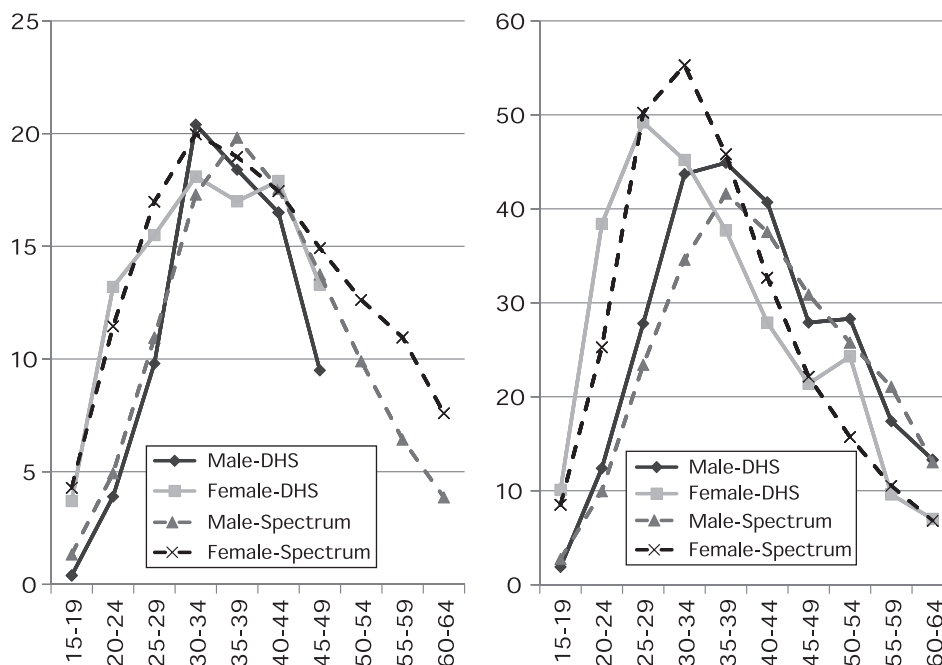


Figure 3 Patterns of the distribution of new HIV infections by age and type of epidemic. Conc, low level and concentrated epidemics; F, female; Gen, generalised epidemics; IDU, injecting-drug-user-driven epidemics; M, male.

Figure 4 Comparison of the age pattern of prevalence from Spectrum projection from 1970 (dashed lines) to survey results (solid lines) in Malawi 2004 (left) and Swaziland 2006/2007 (right). DHS, Demographic and Health Surveys.



orphan. However, because the parents usually share socio-economic and other exposures, their risks of death are correlated for non-AIDS causes. Moreover, due to the sexual transmission of HIV, AIDS deaths of fathers and mothers are highly correlated. Therefore, Grassly and Timaeus fitted a regression model to DHS data from countries in sub-Saharan Africa to estimate the excess risk of being a double orphan as a function of adult HIV prevalence and aggregate indicators of marriage patterns. Since that work was carried out, many more DHS surveys have become available for countries with severe HIV epidemics and high adult mortality, necessitating a new analysis.

The new analysis uses orphanhood data from 55 African DHS conducted in 29 countries, each of which collected data on 8000 to 33 000 children. This database does not represent a random sample of African countries or of points in time. Nevertheless, the surveys cover countries from all parts of the region. The only populous African country for which no data are available is the Democratic Republic of Congo.

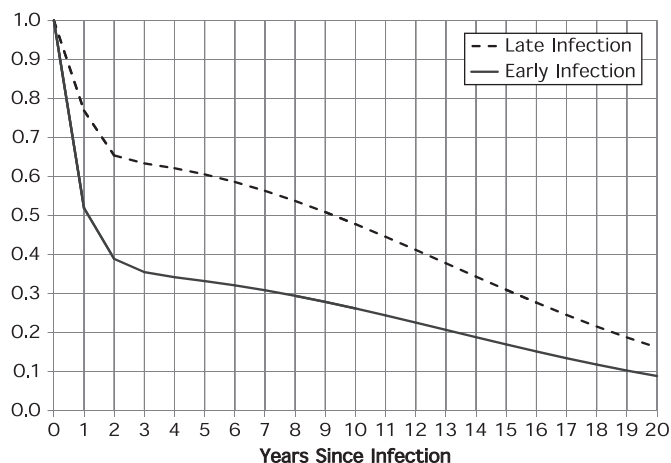


Figure 5 Proportion of HIV-infected children surviving by time since infection for children infected around the time of birth (early infection, solid line) or later through breastfeeding (late infection, dashed line).

The new model includes a somewhat simpler set of explanatory variables than the earlier regression model, requiring only children's age (a) and adult HIV prevalence (P) at two points in time to predict the risk of double orphanhood (λ).

$$\ln \lambda_{i,a} = \theta_0 + \theta_1 \ln \tilde{\lambda}_{i,a} + \theta_2 a_i + \theta_3 \ln P_{t-1} + \theta_4 a_i \ln P_{t-1} + \theta_5 \ln P_{t-12} + v_k + u_t + \gamma_i$$

where $\tilde{\lambda}$ is the risk of double orphanhood assuming independent risks of maternal and paternal orphanhood; the subscripts t , a , i and k refer to dates, ages, individual children and countries respectively; the θ values are coefficients to be estimated; v_k and u_t are country-level and survey-level distributions of a random intercept term (assumed to be normally distributed with variances σ_k^2 and σ_t^2), indicating the degree to which the excess risk of dual orphanhood varies between countries and over time within countries; and γ_i is an individual-level residual that is permitted to have extra-Poisson variation, ω .

The coefficients of the revised model are shown in table 1. The results of applying this equation are shown in table 2 for a low prevalence and a high prevalence setting. Teenaged orphans are about three times more likely to be double orphans than they

Table 1 Parameter estimates for random-intercept Poisson regression models of the percentage of double orphans among children aged 0–17 using Demographic and Health Survey data from countries in mainland sub-Saharan Africa with generalised HIV epidemics.

Variable	Coefficient	SE
Intercept	1.3478	0.1563
ln (percentage of children who are expected double orphans)	0.7539	0.0283
Children's age	-0.02108	0.0131
ln (HIV prevalence in year $t-1$)	-0.1506	-0.0540
Children's age × ln (HIV prevalence in year $t-1$)	0.007225	0.0050
ln (HIV prevalence in year $t-12$)	0.09887	0.0113
Inter-country variance of the intercept	0.0196	0.0059
Intersurvey variance of the intercept	0.0000	0.0000
Extra-Poisson variation	1.7701	0.2071

Table 2 Estimates of the excess risk of double orphanhood (with 95% CIs) from the random-intercept Poisson regression model at two levels of HIV prevalence for mainland African countries with generalised HIV epidemics

Children's age group (years)	1% HIV prevalence	25% HIV prevalence
0–2	12.8 (12.3 to 13.3)	10.9 (10.5 to 11.3)
3–5	6.2 (6.0 to 6.4)	5.7 (5.4 to 5.9)
6–9	4.3 (4.1 to 4.5)	4.2 (4.1 to 4.4)
10–14	3.1 (3.0 to 3.2)	3.4 (3.3 to 3.5)
15–17	2.5 (2.4 to 2.6)	3.0 (2.9 to 3.1)

would be if the risks of their two parents dying were independent. For younger children, the excess risk is even higher.

Compared with Spectrum's previous estimates, the revised model produces higher estimates of numbers of double orphans. The relative differences are greatest for groups in which numbers of double orphans are small (young children and populations with low HIV prevalence) and shrink to about 10% in populations with severe epidemics. The estimates are now more precise. Conditional on the accuracy of the estimates of maternal and paternal orphanhood and HIV prevalence generated by Spectrum, the regression model will usually predict the number of double orphans to within $\pm 4\%$.

The validity of the predictions made by this regression model is contingent on the accuracy of the DHS data on orphans. However, the effects of under-reporting of double and single orphans tend to cancel out in the estimation of the coefficients, and various attempts to assess these data have only found evidence of small to moderate biases in them.²² Moreover, it is the estimates of AIDS deaths, and maternal and paternal orphanhood made from surveillance and survey data on HIV prevalence that largely determine how many children are estimated to be double orphans, not this regression model.

OTHER ENHANCEMENTS

There are a number of other updates to Spectrum for the 2009 estimates. All the demographic inputs have been updated to use the 2008 round of estimates and projections from the United National Population Division.⁵ Procedures for estimating the number of adults and children eligible for treatment have been updated based on new analyses of cohort data and can now be tailored to national eligibility guidelines.²⁴

Survival on ART has also been updated based on the latest studies for adults and children.²⁵ Estimates of the fertility-inhibiting effects of HIV infection on fertility have been updated with analysis of DHS data.²⁶ A new procedure for estimating ranges for regional and global estimates has been implemented which uses the same random number streams for global patterns (eg, progression, treatment effectiveness, prevention of mother-to-child transmission effectiveness) for each set of Monte Carlo projections at the country level, allowing the resulting runs to be summed across regions and the resulting uncertainty ranges calculated for the combined results. This approach assumes no correlation across countries in the incidence estimates produced by EPP, so it may underestimate the true uncertainty.

STRENGTHS

Spectrum is a tool that can be used by national programmes to estimate key planning indicators using their own data and the results from international studies. It incorporates the latest research findings and is regularly reviewed by the UNAIDS

Key messages

- ▶ The Spectrum software is used to calculate HIV indicators of interest on the basis of previously prepared prevalence estimates.
- ▶ New calculations and assumptions have been added in order to incorporate the latest data and respond to the international commitment for expanded treatment.
- ▶ New methods have been developed for estimating the age distribution of incidence, child survival and double orphans.

Reference Group on Estimates, Modelling and Projections. Modifications to Spectrum respond to new data availability and the needs of country programmes.

LIMITATIONS

Many of the assumptions in Spectrum are derived from a small number of studies that may not be representative of all populations or all regions. Some estimates of the need for treatment rely on guidelines as they existed in 2008 and 2009, and ought to be updated to accommodate the new guidelines for treatment and PMTCT programmes. Spectrum calculations rely on incidence generated by another model, EPP, so it is important to ensure that the same assumptions are used in both models. It is impossible to validate Spectrum estimates for mortality, need for treatment or orphans for most countries because of a lack of survey or vital registration data with which to compare.

CONCLUSIONS

Spectrum is one of a package of tools that has been developed to assist national programmes and international organisations to estimate the impact of the AIDS epidemic and the needs for treatment and prophylaxis. It continues to be updated to incorporate the latest research findings and provide indicators needed for programme planning. Recent updates have improved the estimates, particularly for child infection, treatment need and double orphans.

Funding The AIM module in the Spectrum software and manuals and training materials were developed with funding from USAID (Geneva) and UNAIDS (Washington, DC).

Competing interests None.

Contributors All authors reviewed the final manuscript. JS and PJ were primarily responsible for the introduction and the sections on uncertainty, strengths and weaknesses. Each of the other authors contributed a section of the paper: TH on methods for estimating incidence by age, MM and RB for estimates of child survival by mode of transmission, and IMT for updated methods for estimating double orphans.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. Brown T, Bao L, Raftery AE, *et al*. The UNAIDS estimation and projection package 2009. *Sex Transm Infect* 2010;**86**(Suppl 2):ii3–10.
2. Lyerla R, Gouws E, Garcia-Calleja JM, *et al*. The 2005 Workbook: an improved tool for estimating HIV prevalence in countries with low level and concentrated epidemics. *Sex Transm Infect* 2006;**82**(Suppl 3):iii41–4.
3. Stover J, Walker N, Grassly NC, *et al*. Projecting the demographic impact of AIDS and the number of people in need of treatment: updates to the Spectrum projection package. *Sex Transm Infect* 2006;**82**(Suppl 3):iii45–50.
4. Stover J, Johnson P, Zaba B, *et al*. The Spectrum projection package: improvements in estimating mortality, ART needs, PMTCT impact and uncertainty bounds. *Sex Transm Infect* 2008;**84**:i24–30.
5. Stover J. AIM: A Computer Program for Making HIV/AIDS Projections And Examining The Demographic and Social Impacts Of AIDS. Glastonbury CT: Futures Institute, 2009. <http://www.FuturesInstitute.org> and <http://www.unaids.org>.

6. **United Nations Department of Economics and Social Affairs, Population Division.** *World Population Prospects: The 2008 Revision, Highlights*, Working Paper No. ESA/P/WP.210. <http://esa.un.org/unpd/wpp2008/index.htm>.
7. **Hallett TB, Zaba B, Todd J, et al.** Estimating incidence from prevalence in generalised HIV epidemics: methods and validation. *PLoS Med* 2008;**5**:e80.
8. **Todd J, Glynn JR, Marston M, et al.** Time from HIV seroconversion to death: a collaborative analysis of eight studies in six low and middle-income countries before highly active antiretroviral therapy. *AIDS* 2007;**21**:S55–63.
9. **Hallett TB, Stover J, Mishra V, et al.** Estimates of HIV incidence from household-based prevalence surveys. *AIDS* 2010;**24**:147–52.
10. **Egger M.** Outcomes of antiretroviral treatment in resource limited and industrialized countries. *14th Conference on Retroviruses and Opportunistic Infections*. 2007, Los Angeles. http://www.retroconference.org/2007/data/files/webpage_for_CROI.htm.
11. **Mishra V, Barrere B, Hong R, et al.** Evaluation of bias in HIV seroprevalence estimates from national household surveys. *Sex Transm Infect* 2008;**84**(Suppl 1):i63–70.
12. **Marston M, Harriss K, Slaymaker E.** Non-response bias in estimates of HIV prevalence due to the mobility of absentees in national population-based surveys: a study of nine national surveys. *Sex Transm Infect* 2008;**84**(Suppl 1):i71–7.
13. **Reniers G, Eaton J.** Refusal bias in HIV prevalence estimates from nationally representative seroprevalence surveys. *AIDS* 2009;**23**:621–9.
14. **ECDC.** *HIV Surveillance in Europe 2007*. Stockholm: European Centre for Disease Prevention and Control, WHO Regional Office for Europe, 2008.
15. **Marston M, Zaba B, Salomon JA, et al.** Estimating the net effect of HIV on child mortality in African population affected by generalized HIV epidemics. *J Acquir Immune Defic Syndr* 2005;**38**:219–27.
16. **The Breastfeeding and HIV International Transmission Study Group (BHITS).** Late postnatal transmission of HIV-1 in breast-fed children: an individual patient data meta-analysis. *J Infect Dis* 2004;**189**:2154–66.
17. **Iloff PJ, Piwoz EG, Tavengwa NV, et al.** Early exclusive breastfeeding reduces the risk of postnatal HIV-1 transmission and increases HIV-free survival. *AIDS* 2005;**19**:699–708.
18. **Rollins N.** HIV transmission and mortality associated with exclusive breastfeeding: implications for counselling HIV-infected women. http://www.path.org/files/Nigel_Rollins.pdf (accessed 26 Dec 2006).
19. **UNAIDS.** *Consultative Meeting on Data Collection and Estimation Methods Related to HIV Infection in Infants and Children*. New York: UNAIDS, UNICEF, WHO, 2008.
20. **Mofenson LM.** Prevention in neglected subpopulations: prevention of mother-to-child transmission of HIV infection. *Clin Infect Dis* 2010;**50**(Suppl 3):S130–48.
21. **Becquet R.** UNAIDS Child Survival Working Group. Survival of children HIV-infected perinatally or through breastfeeding. A pooled analysis of individual data from Sub-Saharan Africa. *The 17th Conference on Retroviruses and Opportunistic Infections*, San Francisco, 2010. Paper#840 <http://www.retroconference.org/2010/PDFs/840.pdf>.
22. **Marston M, Todd J, Glynn JR, et al.** Estimating 'net' HIV-related mortality and the importance of background mortality rates *AIDS* 2007;**21**(suppl 6):S65–71.
23. **Grassly NC, Timaeus IM.** Methods to estimate the number of orphans as a result of AIDS and other causes in Sub-Saharan Africa *J Acquir Immune Defic Syndr* 2005;**39**:365–75.
24. **Grassly NC, Lewis JJC, Mahy M, et al.** Comparison of survey estimates with UNAIDS/WHO projections of mortality and orphan numbers in sub-Saharan Africa. *Popul Stud* 2004;**58**:207–17.
25. **May M, Lewden C, Brinkhof M, et al.** Derivation of parameters used in spectrum for eligibility for ART and survival on ART. *Sex Transm Infect* 2010;**86**(Suppl 2):ii28–34.
26. **Chen WJ, Walker N.** Fertility of HIV-infected women: insights from demographic and health surveys. *Sex Transm Infect* 2010;**86**(Suppl 2):ii22–27.